Studies of G Protein-Coupled Receptor Structure and Function: Receptor Modeling and Synthesis of Bivalent Ligands

Fiona Michelle McRobb

B. Med. Chem. (Hons)

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Medicinal Chemistry and Drug Action

Faculty of Pharmacy and Pharmaceutical Sciences

Monash University

2011

MONASH University

Institute of Pharmaceutical Sciences

Copyright

Addendum

p xiv, line 6: replace “GPCRs makes structure-based drug design” with “GPCRs make structure-based drug design”

p xiv, line 7: replace “cell membrane, indicate” with “cell membrane indicates”

p xv, line 14: replace “inhibitory potency observed” with “inhibitory potency was observed”

p xv, line 23: replace “as well as the discovery of a lead clozapine” with “as well as discovering a lead clozapine”

p 2, line 17: replace “invertebrate GPCRs, where the some” with “invertebrate GPCRs, where some”

p 5, line 12: replace “the second number represents the relative to” with “the second number represents the relative position to”

p 6, Figure 1.4 caption, line 4: add “the broken” and read “Arg 3.50 and Glu 6.30 of the broken ionic lock shown in magenta”

p 9, line 3: replace “ICL3 with more stable protein” with “ICL3 with a more stable protein”

p 9, line 6: replace “Thermostabillized” with “Thermostabilized”

p 9, line 15: replace “nanobody methods have been” with “nanobody methods has been”

p 9, line 22: replace “Not only was this useful from a drug design perspective, as adds” with “This was useful from a drug design perspective, as it adds”

p 9, line 24: replace “crystallization methodologies for GPCRs has allowed” with “crystallization methodologies for GPCRs have allowed”

p 9, insert new paragraph before Section 1.4.2:

“GPCRs are inherently flexible proteins and are likely to adopt a range of unique conformations, which may be influenced by the interaction of co-crystalized ligand or the G-protein binding.[43] Despite this, the crystal structures can be loosely classified into three main categories: the *apo* or ligand-free state, the “inactive” or antagonist/inverse agonist bound state and the “active” or agonist bound state. However, due to the induced-fit effect of ligand binding and the dynamic nature of GPCRs, it is likely that a number of conformers exist for each state.[43] Additionally, the use of thermostabilizing mutations and the crystallography conditions employed are not without risk, as they may alter the 3D structure of the receptor compared to the biologically relevant state.[49a] Ultimately, to gain a greater insight into the

structure and function of GPCRs, a number of crystal structures will be required of a variety of receptors, in complex with a variety of ligands, as well as data from other biophysical techniques."

p 15, lines 1-3: replace "The first crystal structure of the turkey β_1AR (Figure 1.9) in complex with an antagonist, cyanopindolol (**6**), as the turkey receptor was more stable" with "The first crystal structure of the β_1AR (Figure 1.9) in complex with an antagonist, cyanopindolol (**6**), was a turkey receptor as this was more stable"

p 18, line 4: replace "a number of GPCRs structures" with "a number of GPCR structures"

p 21, line 2: replace "antagonist bound structure" with "antagonist bound structures"

p 21, line 3: replace "contraction of orthosteric" with "contraction of the orthosteric"

p 21, line 12: replace "not as significantly as those" with "not as significantly as the changes"

p 32, line 22: replace "GPCR crystals structures" with "GPCR crystal structures"

p 36, line 20: replace "These techniques include chimeric studies" with "Chimeric studies"

p 36, line 23: replace "receptor was combined" with "receptor were combined"

p 42, line 10: replace "compounds with less" with "compounds with fewer"

p 43, Figure 1.20 caption, line 2: replace "by a linking group" with "by linking groups"

p 47, line 5: replace "therefore removing" with "therefore removing the"

p 50, line 2: replace "finely tuned" with "fine-tuned"

p 51, line 12: replace "but also the mechanism" with "but also on the mechanism"

p 54, line 8: replace "problematic as the form" with "problematic as they form"

pg 54, line 11: replace "increasing the influence on solubility" with "increasing the solubility"

p 55, line 8: replace "aromatic-containing" with "aromatic ring-containing"

p 55, line 22: replace "where as meta-substituted" with "whereas the meta-substituted"

p 56, line 2: replace "empirically bivalent ligand" with "empirically for each bivalent ligand"

p 56, Figure 1.28 caption, line 2 and p 57, line 6: replace "piperazines / piperidines" with "piperazine / piperidine

p 57, line 12: replace “and may present” with “and targeting these dimers may present”

p 57, line 14: replace “last 60 years and attracted” with “last 60 years and have attracted

p 57, line 15: replace “3D structural data of” with “3D structural data for”

p 57, line 19: replace “an even bigger role” with “an even greater role

p 59, line 8: replace “Clozapine considered” with “Clozapine is considered”

p 87, line 6: replace “it shares” with “rhodopsin”

p 88, line 8: replace “June 2011).” with “June 2011) structures.”

p 105, line 5: replace “site” with “sites

p 134, add after reference 49: “49a. Rosenbaum, D. M.; Rasmussen, S. G. F.; Kobilka, B. K. The structure and function of G-protein-coupled receptors. *Nature* **2009,** *459*, 356-363.”

p 191, line 11: replace “used identify two dimerization interfaces in the D_2R homodimer;” with “used to identify two dimerization interfaces in the D_2R homodimer:”

p 192, line 11: replace “Rosetta++[24]” with “Rosetta++,[24]”

p 192, line 12: replace “to atomic force microscopy model of rhodopsin dimer” with “to the atomic force microscopy model of the rhodopsin dimer”

p 192, line 24: replace “Filizola and co-workers,” with “Filizola and co-workers”

p 192, line 26: replace “and the guide” with “and to guide”

p 201, line 11: replace “contacts maintained” with “contacts are maintained”

p 207, line 4: replace “the increase affinity” with “the increased affinity”

p 203, Figure 4.5 caption, add to the end of caption: “Note: the initial pose of clozapine is closer to the extracellular side of the receptor.”

p 215, line 2: replace “a significant pharmaceutical target” with “significant pharmaceutical targets”

Table of Contents

Statement of originality

To the best of the author's knowledge and belief, this thesis contains no material which has been accepted for the award of any other degree or diploma in any university or other institution, and contains no material previously published or written, except where due reference is made.

Fiona M. McRobb

Part A: General declaration

Monash University, Monash Research Graduate School Declaration for thesis based or partially based on conjointly published or unpublished work.

General Declaration

In accordance with Monash University Doctorate Regulation 17/ Doctor of Philosophy and Master of Philosophy (MPhil) regulations the following declarations are made:

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes 1 original paper published in a peer reviewed journal and 2 unpublished publications. The core theme of the thesis is the development an evaluation of homology models of GPCRs and the synthesis, characterization and pharmacological testing of bivalent ligands of clozapine. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the candidate, working within the Department of Medicinal Chemistry and Drug Action, Monash Institute of Pharmaceutical Sciences, under the supervision of Dr Ben Capuano, Dr Elizabeth Yuriev and Dr Ian T. Crosby. The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research. The nature of assistance from co-authors has been duly acknowledged in Part B of this declaration, at the beginning of each chapter.

In the case of Chapter 2, my contribution to the work involved the following: the development and refinement of homology models, design and running of virtual screening experiments, analysis of results, and manuscript preparation. In Chapter 3, my contribution entailed the development and refinement of homology models, design and running of virtual

screening experiments, analysis of results, and manuscript preparation. In Chapter 4, my contribution to the work involved the design of compounds, as well as the synthesis, purification and analysis of all compounds and manuscript preparation.

Thesis chapter	**Publication title**	**Publication status**	**Nature and extent of candidate's contribution**
2	Homology modeling and docking evaluation of aminergic G protein-coupled receptors	published	Development and refinement of homology models, design and running of virtual screening experiments, analysis of results, and manuscript preparation. 80% contribution.
3	Predicting the structure of the dopamine D_3 receptor: An evaluation of virtual screening approaches to GPCR modeling	unpublished	Development and refinement of homology models, design and running of virtual screening experiments, analysis of results, and manuscript preparation. 75% contribution
4	Homobivalent ligands of the atypical antipsychotic clozapine: Synthesis and pharmacological evaluation	unpublished	Compound design, synthesis, purification and analysis of all compounds and manuscript preparation. 60% contribution.

I have not renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Candidate's signature: [redacted] Date: 05 / 09 / 2011

Acknowledgements

Firstly, I would especially like to thank my primary supervisor, Dr Ben Capuano for his continued patience and guidance over the past four years and for always encouraging me to achieve my best. I would also like to thank my associate supervisors, Dr Ian Crosby and Dr Elizabeth Yuriev for their continued assistance and guidance.

I would like to extend my gratitude to Dr David Chalmers, not only for his continued support with the computational aspects of the project, but also for his continued help with the preparation of manuscripts and guidance on this project. Likewise, thanks go to Dr David Manallack, Dr Mike Kuiper and Dr Ben Roberts for their assistance with the computational aspects of this project.

I am indebted to Dr Rob Lane and Prof. Arthur Christopoulos for performing the pharmacological testing and analysis for the compounds synthesized in this study and their assistance with manuscript preparation.

Thanks to Mark Agostino for his assistance with script writing and for throwing around ideas and his continued friendship, Yu Fang for his assistance with the molecular dynamics simulations and providing clozapine parameters and Kimberley McLean for sharing her thoughts on the virtual screening studies.

Thank you to Dr Jason Dang for performing all mass spectrometry analyses, as well as his assistance with the NMR, particularly when running high temperature experiments.

To all the members of Lab 103 / 106 past and present, especially Richard Conway, Vlad Moudretski, Jeremy Shonberg, Dragan Krsta, Tracey Huynh, Monika Szabo and Manuela Jorg, thanks so much for helping me along the way and for the fun times we've had in the lab. I have really enjoyed being part of such a fantastic group.

Thank you to my fellow PhD students for their support over the years, especially Susan Northfield, Diana Neale, Jo-Anne Pinson and Natalie Vinh for always being there for a chat and to share a coffee.

I would like to thank Amruta Bapat, Alison Chan, Michelle Ta, Godwin Lo and Des Hong for their friendship over the years and for the fun times we've had travelling together.

Finally, I wish to thank my family for their continued support, which made completing this thesis possible.

Abbreviations and Acronyms

α_{1A}AR	α_{1A}-adrenergic receptor
α_{1B}AR	α_{1B}-adrenergic receptor
α_{1D}AR	α_{1D}-adrenergic receptor
α_{2A}AR	α_{2A}-adrenergic receptor
α_{2B}AR	α_{2B}-adrenergic receptor
α_{2C}AR	α_{2C}-adrenergic receptor
β_1AR	β_1-adrenergic receptor
β_2AR	β_2-adrenergic receptor
β_3AR	β_3-adrenergic receptor
3D	three dimensional
5-HT	serotonin
5-HT_{1A}R	serotonin 1A receptor
5-HT_{1B}R	serotonin 1B receptor
5-HT_{1D}R	serotonin 1D receptor
5-ht_{1e}R	serotonin 1E receptor
5-HT_{1F}R	serotonin 1F receptor
5-HT_{2A}R	serotonin 2A receptor
5-HT_{2B}R	serotonin 2B receptor
5-HT_{2C}R	serotonin 2C receptor
5-HT_4R	serotonin 4 receptor
5-ht_{5a}R	serotonin 5a receptor
5-HT_6R	serotonin 6 receptor
5-HT_7R	serotonin 7 receptor
A_1AR	adenosine 1 receptor
A_{2A}AR	adenosine 2A receptor
BBB	blood-brain barrier
BRET	bioluminescence resonance energy transfer
cAMP	cyclic adenosine monophosphate
ClogP	calculated partition coefficient
CNS	central nervous system
D_1R	dopamine 1 receptor
D_2R	dopamine 2 receptor
D_3R	dopamine 3 receptor
D_4R	dopamine 4 receptor
D_5R	dopamine 5 receptor
DMEM	Dulbecco's modified eagle medium
DOR	δ opioid receptor

ECL	extracellular loop
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol tetraacetic acid
ERK	extracellular-signal-regulated kinases
Fab	fragment antigen binding
FDA	Food and Drug Administration
FRET	fluorescence resonance energy transfer
fs	femtosecond
GPCR	G protein-coupled receptor
GRAFS	glutamate, rhodopsin, adhesion, frizzled/taste2 and secretin
H_1R	histamine 1 receptor
H_2R	histamine 2 receptor
H_3R	histamine 3 receptor
H_4R	histamine 4 receptor
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	high-performance liquid chromatography
i.c.v.	intracerebroventricular
i.v.	intravenous
IC_{50}	half maximal inhibitory concentration
ICL	intracellular loop
IFD	Induced Fit Docking
K	kelvin
K_i	inhibition constant
KOR	κ opioid receptor
M_1 mAChR	M_1 muscarinic acetylcholine receptor
M_2 mAChR	M_2 muscarinic acetylcholine receptor
M_3 mAChR	M_3 muscarinic acetylcholine receptor
M_4 mAChR	M_4 muscarinic acetylcholine receptor
M_5 mAChR	M_5 muscarinic acetylcholine receptor
Mab	monoclonal antibody
MAS	MAS1 oncogene receptor
MCH	melanin-concentrating hormone
MD	molecular dynamics
MECA	melanocortin, endothelial, cannabinoid, adenosine
MHz	megahertz
MOR	μ opioid receptor
MSA	multiple sequence alignment
NAMD	Nanoscale Molecular Dynamics
NMR	nuclear magnetic resonance

OPLS	Optimized Potentials for Liquid Simulations
PDB	Protein Data Bank
PME	Particle Mesh Ewald
POPC	1-palmitoyl-2-oleoylphosphatidylcholine
rmsd	root mean square deviation
SBDD	structure based drug design
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEM	standard error of the mean
SOG	somatostatin, opioid, galanin
SP	standard precision
TFA	trifluoroacetic acid
TLC	thin layer chromatography
TM	transmembrane
vdW	van der Waals
VMD	Visual Molecular Dynamics
XP	extra precision

Abstract

G protein-coupled receptors (GPCRs) are therapeutically significant proteins and are targeted by over 25% of FDA approved drugs. GPCRs are highly druggable and are involved in a diverse range of disease states and, as such, are of immense interest to the pharmaceutical industry. GPCRs play an important role in cell signaling, mediating signals across the cell membrane. Recent advances in the high resolution X-ray crystallography of GPCRs makes structure-based drug design significantly more feasible. Additionally, increased understanding of the arrangement of GPCRs in the cell membrane, indicate that many GPCRs are likely to form dimers or higher order oligomers. In fact, dimerization is believed to be a common feature to GPCRs and may represent a novel therapeutic target for numerous disease states.

In this thesis, recent high resolution crystal structures of several class A GPCRs have been used in the development and evaluation of new homology models of pharmaceutically relevant GPCRs. Additionally, a series of homobivalent ligands have been developed as pharmacological tools to investigate GPCR dimerization.

Homology models for several therapeutically significant GPCRs were developed using the high resolution X-ray crystal structure of the β_2-adrenergic receptor as a template (Chapter 2; McRobb, F. M. et al. *J. Chem. Inf. Model.* **2010,** *50*, 626-637). Techniques to optimize the orthosteric binding site, such as flexible receptor docking and loop refinement, were investigated. Small scale virtual screening was undertaken to evaluate the developed homology models for use in a structure-based drug design campaign. Of the nine homology models developed, six showed moderate to good enrichment in virtual screening experiments (5-HT_{1B}R, 5-HT_{2A}R, 5-HT_{2C}R, D_2R, D_3R and M_1 mAChR). These newly developed aminergic GPCR homology models supplement the limited number of freely available GPCR homology models. It is hoped that these models will provide a better starting point for structure-based drug design.

As a continuation of our research, we have evaluated our GPCR modeling method using the GPCR Dock 2010 assessment (Chapter 3). GPCR Dock 2010 involved the prediction of the complex of the dopamine D_3 receptor with the small molecule eticlopride, prior to the release of the high resolution X-ray crystal structure. The five top ranked models from this prediction were submitted to the GPCR Dock 2010 analysis (Kufareva, I. et al. *Structure* **2011,** *19*, 1108-1126) and are also compared to eticlopride in the dopamine D_3 receptor crystal structure.

Three series of homobivalent ligands of the atypical antipsychotic clozapine, were designed, synthesized and pharmacologically evaluated (Chapter 4). Clozapine exerts its therapeutic effect by antagonism of dopaminergic and serotonergic GPCRs, however, clozapine only displays moderate (sub-micromolar) affinity for the dopamine D_2 receptor. Attachment of the spacer at the N4′ position of clozapine yielded a series of homobivalent ligands that displayed the most promising affinity and activity for the dopamine D_2 receptor. A spacer length-dependent relationship with affinity or inhibitory potency observed, with the 16 and 18 atom spacer bivalent ligands displaying low nanomolar affinity (1.41 nM and 1.35 nM) and a significant gain in affinity (75- and 79-fold, respectively) relative to the original pharmacophore, clozapine. Additionally, expanding on the modeling methods described in Chapters 2 and 3, four models of the dopamine D_2 homodimer were built and optimized using molecular dynamics simulations, to determine the approximate distance between the adjacent orthosteric sites of the dimer.

This project has successfully achieved the aims outlined, developing and evaluating homology models of aminergic GPCRs that are useful for structure-based drug design, as well as the discovery of a lead clozapine homobivalent ligand, with an appropriate attachment point and spacer length determined.

Chapter 5 provides a brief summary of this thesis, with an evaluation of outcomes, as well as directions for future work.

Chapter 1

Introduction

1.1 G protein-coupled receptors as drug targets

G protein-coupled receptors (GPCRs) are therapeutically significant membrane-bound proteins, with over 25% of FDA approved drugs targeting these receptors.[1] GPCRs play a key role in controlling cell signaling processes, conveying signals across the cell membrane in response to a diverse array of ligands, including hormones and neutrotransmitters.[2] GPCRs are implicated in a wide range of therapeutic areas including central nervous system (CNS) disorders, cancer, cardiovascular disease, inflammation, obesity and pain.[3] Following the full sequencing of the human genome in 2001,[4] over 800 human GPCRs have been identified.[5] As GPCRs are highly tractable drug targets,[6] they are one of the most thoroughly investigated targets in drug discovery.[3]

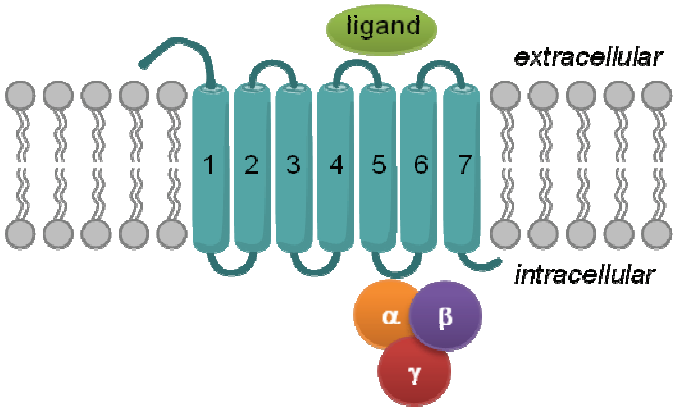


Figure 1.1 Schematic of a GPCR, displaying seven transmembrane helices, linked by intra- and extracellular loops and the G protein heterodimer binding intracellularly.

GPCRs are membrane bound receptors, with a characteristic structure of seven transmembrane (TM) α-helices (Figure 1.1), linked by intracellular loops (ICL) and extracellular loops (ECL). Whilst these receptors overall have a similar basic topology, they

share relatively low sequence homology. Typically, endogenous ligands bind to the orthosteric site on the extracellular side of the receptor and either induce or block a biological response. This response is regulated by the G protein, a heterotrimeric subunit (consisting of α, β and γ subunits), which binds to the GPCR intracellularly, and is involved in second messenger signaling.

GPCRs can exist in a number of states, ranging from fully active to fully inactive, and these states are reflected in a number of the available X-ray crystal structures (discussed in Section 1.4).[7-21] When an agonist binds to the receptor, it enables the binding of the G protein. Agonists mimic the response of the endogenous ligand and trigger a biological response. The binding of an antagonist to a GPCR has no intrinsic activity and it blocks other agonist or inverse agonist responses. Similarly, when an inverse agonist binds to the receptor, it inhibits agonist binding, as well as inhibiting the constitutive (agonist independent) activity of the receptor.[22]

1.2 Classes of GPCRs

Despite their similar topology, GPCRs have a range of functions and interact with a wide variety of ligands, and these receptors can be classified by their function.[23] The most well known classification system for GPCRs is the A-F,O classification system.[24,25] The A-F,O classification system aimed to describe all vertebrate and invertebrate GPCRs, where the some of the classes D, E and F (fungal pheromone receptors, cAMP binding receptors and archebacterial opsins, respectively) do not exist in humans, and the O class corresponds to "other" GPCRs that could not be classified.[5] More recently, following the sequencing of the human genome,[4] it has been demonstrated that human GPCRs can be classified into five families: glutamate, rhodopsin, adhesion, frizzled/taste2 and secretin; the GRAFS classification system (Figure 1.2).[5] Of these five families, three classes correspond to the original A-F,O classification system: rhodopsin (class A), secretin (class B) and glutamate

(class C).[5] The rhodopsin or class A family (Figure 1.2), is further broken down into four groups; α, β, γ and δ. The α-group contains five further classes of receptor; prostaglandin, amine, opsin, melatonin and the MECA (melanocortin / endothelial / cannabinoid / adenosine) receptor clusters. The β-group contains receptors that bind peptides. The γ-group contains three main clusters; SOG (somatostatin / opioid / galanin), melanin-concentrating hormone (MCH) and the chemokine receptor clusters. The δ-group contains four main clusters; MAS1 oncogene receptor (MAS), glycoprotein, the purine and the olfactory receptor clusters.[5]

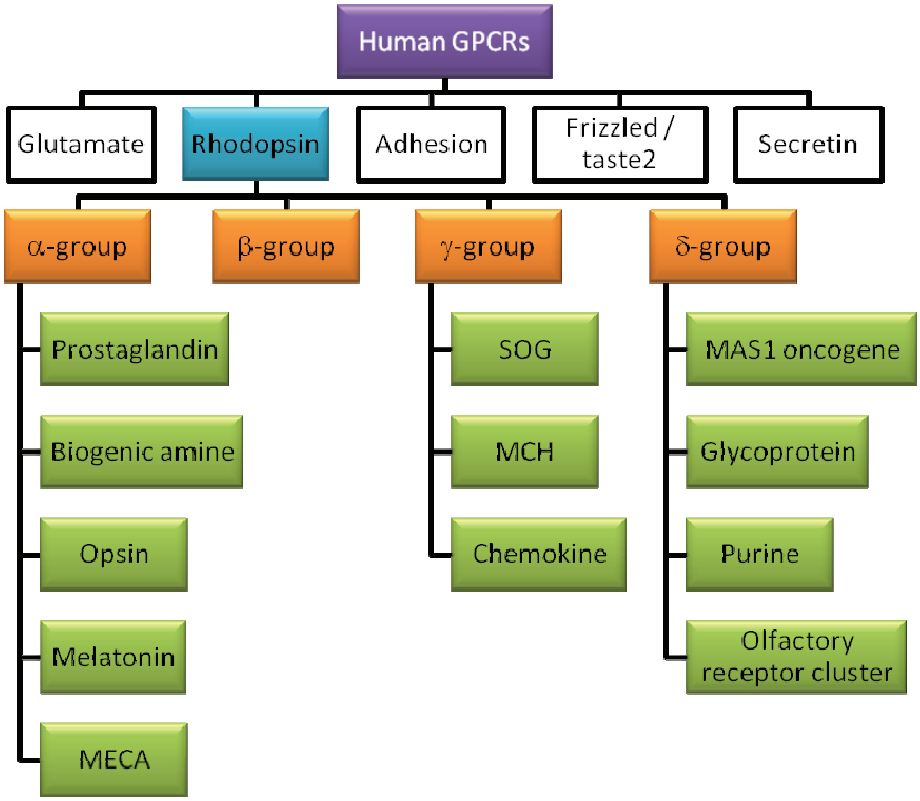


Figure 1.2 GRAFS classification of human GPCRs, highlighting the rhodopsin family subgroups.[5]

Over one quarter of approved drugs target rhodopsin-like GPCRs.[1] Of particular interest to our research group, and to this work, are the aminergic and purinergic receptors, which belong to the class A, α-group GPCRs according to the GRAFS classification system. Many of these receptors are implicated in CNS disorders,[26] such as schizophrenia[27] and Parkinson's disease,[28-30] which are a major focus in our research group.

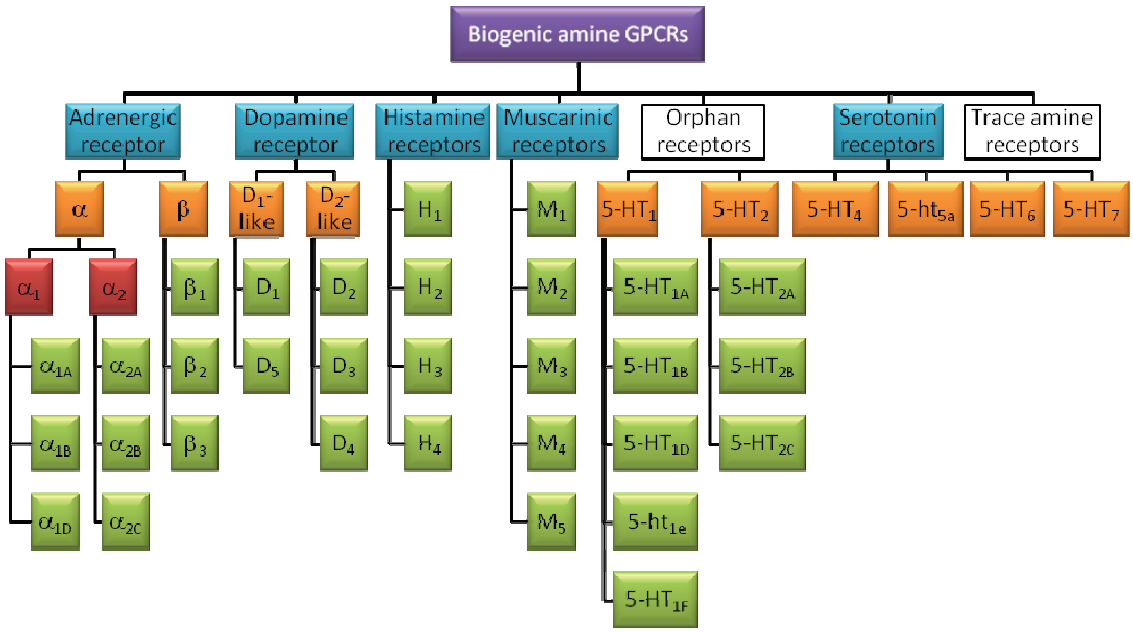


Figure 1.3 Receptors that constitute the biogenic amine cluster of human GPCRs, which includes the adrenergic, dopaminergic, histaminergic, muscarinic, serotonergic and trace amine receptors, as well as some orphan receptors (GPCRs for which an endogenous ligand has not been identified).[5]

The biogenic amine cluster of the α-branch of class A GPCRs consists of approximately 40 receptors and includes dopamine, serotonin, muscarinic and histamine receptors (Figure 1.3).[5,23] This GPCR cluster can be further broken down into receptor subtypes, based on similarity. The dopamine receptors are divided into two classes, the D_1-like and the D_2-like families. The D_1-like family consists of the D_1 and D_5 receptors. The D_2-like family contains the D_2, D_3 and D_4 receptors. The serotonin receptors have 12 receptor subtypes. The 5-HT_1 family consists of the 5-HT_{1A}, 5-HT_{1B}, 5-HT_{1D}, 5-ht_{1e} (5-ht in lowercase indicates that no clear functional role for this receptor has been determined in native tissue) and 5-HT_{1F} receptors and the 5-HT_2 family consists of the 5-HT_{2A}, 5-HT_{2B} and 5-HT_{2C} receptors. Also included in the serotonin family of receptors are the 5-HT_4, 5-ht_{5a}, 5-HT_6 and 5-HT_7 receptors. The adrenoceptors are divided into two classes α and β, which contain the α_{1A}, α_{1B}, α_{1D}, α_{2A}, α_{2B}, α_{2C} and the β_1, β_2 and β_3 adrenergic receptors. There are four histamine receptors, H_1, H_2, H_3 and H_4 and five muscarinic acetylcholine receptors, M_1, M_2, M_3, M_4 and M_5. Additional receptors in the biogenic amine cluster include trace amine receptors and

some orphan receptors (those receptors for which an endogenous ligand has not been identified). Over 18 of receptors from the α-branch of the rhodopsin-like receptors are significant drug targets,[23,31] and a number of cardiovascular, antihistamine, antipsychotic and antidepressant drugs act at these receptors.

1.3 GPCR sequence homology

Class A or rhodopsin-like GPCRs share a fairly low percentage sequence homology (< 30% for the non-visual class A receptors[32]), however, this is noticeably higher in the more conserved TM helices, because the majority of the conserved residues are located in this region.[31] Despite relatively low sequence homology, GPCRs share a similar topology, which is reinforced by the presence of conserved residues. In class A GPCRs, the most conserved residue in each helix has been identified and is used as a point of reference in the Ballesteros-Weinstein nomenclature.[33] In this numbering system, the first number corresponds to the helix number and the second number represents the relative to the most conserved residue, with the most conserved residue in each helix assigned the arbitrary number '50'.[33] However, this numbering system is not used for the variable loop regions, where receptor sequence numbering is used.

1.3.1 Conserved residues and motifs in class A GPCRs

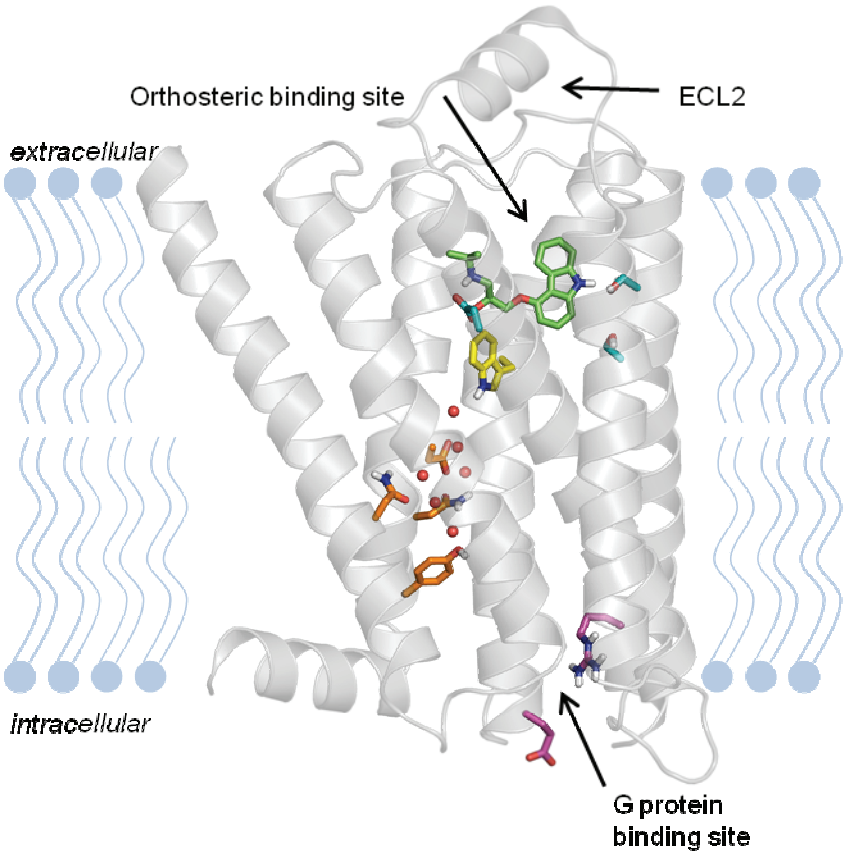


Figure 1.4 Key structural features of rhodopsin-like GPCRs, using the structure of the β_2AR (2RH1)[10] as an example, with co-crystallized carazolol displayed in green. Water hydrogen bonding network highlighted in orange (water molecules shown as red spheres), Arg 3.50 and Glu 6.30 of ionic lock shown in magenta, toggle switch (Trp 6.48) displayed in yellow and the aspartate residue (Asp 3.32) for aminergic GPCRs and Ser 5.42 and Ser 5.46 shown in cyan. Images of crystal structures were prepared using PyMOL.[34]

Conserved residues in the rhodopsin-like GPCRs form key motifs that are characteristic of these receptors and integral to their structure and function (Figure 1.4). The majority of conserved residues reside in the transmembrane region of class A GPCRs, with the exception of a conserved disulfide bridge between the conserved cysteine in ECL2 and Cys 3.25, which restricts the movement of ECL2, particularly above the orthosteric site.

The NPxxY$(x)_{5,6}$F motif is present at the cytoplasmic end of TM7 and contains the highly conserved residues Asn 7.49, Pro 7.50 and Tyr 7.53.[35] Asn 7.49 is involved in a hydrogen bond network, mediated by water molecules between Asn 1.50, Asp 2.50, which stabilizes

interactions between TM1, TM2 and TM7.[36] Additionally, Tyr 7.53 forms hydrophobic interactions with the conserved phenylalanine residue in helix 8.

The D(E)RY motif on TM3 (Asp 3.49, (Glu 3.49), Arg 3.50, Tyr 3.51) and Glu 6.30 form an important micro-switch in rhodopsin-like GPCRs, the "ionic lock".[37-39] Ionic interactions between Asp/Glu 3.49, Arg 3.50 and Glu 6.30 are thought to be critical for retaining the receptor in the ground or inactivated state, stabilizing the cytoplasmic portions of TM3 and TM6.[37-39] Rearrangement or breaking of the ionic lock has been proposed to be involved in receptor activation, allowing for the cytoplasmic movement of TM helices and binding of the G protein,[37] although the extent to which the ionic lock contributes to receptor activation is still debated.[38]

Another conserved motif in class A GPCRs, CWxP, is located on TM6 and includes the Trp 6.48 residue. Trp 6.48 resides at the base of the orthosteric binding site and is postulated to be a "toggle switch", changing the receptor between inactive and active states (*gauche*+ and *trans* respectively), although this change in rotamer has not been observed in any crystal structure.[40] This motif also forms part of the cluster of aromatic residues on TM6, present in biogenic amine GPCRs, and includes residues Phe 6.44, Trp 6.48, Pro 6.50, Phe 6.51 and Phe 6.52.[41]

For the biogenic amine receptors, the orthosteric site is located between TM helices 3, 5, 6 and 7, with contributions from ECL2, a portion of which (between the conserved disulfide bond and TM5) forms part of the orthosteric site. In the aminergic GPCRs, the key ligand-receptor interaction is between an ionized nitrogen and Asp 3.32.[42] Many ligands also interact with the less conserved serine residues on TM5, Ser 5.42 and Ser 5.46, particularly ligands containing a catechol moiety (such as the agonists dopamine and noradrenaline). Additionally, ligands interact with the cluster of aromatic residues on TM6.

Whilst there are sequences available for approximately 800 human GPCRs identified,[5] there are relatively few 3D high resolution crystal structures available of GPCRs, although this number has been steadily increasing over the past few years (refer to Section 1.4).

1.4 Available structural data for GPCRs

The majority of GPCR targeting drugs act at class A GPCRs, particularly the α-group,[23,31] thus stimulating a significant need to understand the structure and function of these proteins. However, because GPCRs are dynamic, membrane bound proteins, crystal structures are incredibly difficult to obtain, as they are notoriously difficult to crystallize due to low stability in crystallization conditions.[43] Significant technological advances have been made in the past 11 years, in particular since 2007, to determine the crystal structures of GPCRs. The noteworthy technological advances are discussed in Section 1.4.1. The list of crystal structures discussed in Sections 1.4.2 and 1.4.3 were current as of the 1st of August, 2011. However, these crystal structures only account for seven out of approximately 350 GPCR drug targets. Of particular interest to our research is the use of these crystal structures as templates for the development of homology models (Section 1.5). Thus, it is essential to have an appreciation of the available crystal structures, in order to select appropriate templates and methods for homology modeling (Section 1.5.3.1).

1.4.1 Methods of crystallization – protein engineering techniques

The first crystal structure of rhodopsin was obtained from bovine rod outer segment membranes, and crystallization was aided by the low activity of 11-*cis*-retinal-bound rhodopsin.[7,44] However, GPCRs that bind diffusible ligands proved more difficult to crystallize compared to the rhodopsin-based structures, due to their dynamic nature.[43] Additionally, the large ICL3 of these receptors (joining TM5 and TM6) further added to their instability. A number of approaches have been developed to crystallize GPCRs. In the non-rhodopsin crystal structures, ICL3 was often removed and used as a point to aid in the stability of the receptor. In one approach, an antibody fragment (Fab5) that was generated in

detergent from a monoclonal antibody (Mab5),[45] was used to recognize the intracellular surface of TM5 / TM6 and assist in the crystallization of β_2AR.[9,46] Another method was to replace ICL3 with more stable protein, the T4-lysozyme, to induce crystallization and stabilize TM5 and TM6.[8,10,43] This technique has been used for the successful determination of a number of GPCR crystal structures including β_2AR, A_{2A}AR, D_3R and H_1R (Table 1.1).[15,16,18-20] Thermostabillized receptors have also been used, in which a number of point mutations are introduced into the receptor to increase stability in detergent micelles.[47] These receptors were evaluated in binding and functional assays to assess the effect of the mutations on ligand binding.[47] This technique is also referred to as the **sta**bilized **r**eceptors (StaRs) method,[48] which has been used to determine structures of both agonist- and inverse agonist-bound structures. Because ICL3 is not replaced by the T4-lysozyme in this method, the structure of ICL3 can be determined. In determining the crystal structure of a GPCR in an active state, a camelid antibody fragment (nanobody), which selectively recognizes and stabilizes the G protein binding site of the active state of β_2AR (PDB ID: 3P0G) has been used.[49] A combination of the T4-lysozyme and nanobody methods have been used to determine the structure of the ternary complex of β_2AR, with the T4-lysozyme indirectly stabilizing the extracellular side of the receptor and the nanobody stabilizing the G protein.[21]

1.4.2 Key crystal structures

In 2000, the first breakthrough came when the crystal structure of bovine rhodopsin was solved,[7] which is a class A GPCR. This was the first high resolution crystal structure of a GPCR, which lead to significant advances for structure-based drug design. The next major advances in this field were the crystal structures of the first human GPCR, the β_2-adrenergic receptor, in 2007 (Table 1.1).[8-10] Not only was this useful from a drug design perspective, as adds to the number of available templates to use for homology modeling of GPCRs. Significant advancements in crystallization methodologies for GPCRs has allowed for the subsequent determination of a number of crystal structures.

Table 1.1 Details of the crystal structures of aminergic and purinergic GPCRs (chemical structures of ligands in Figures 1.5 and 1.6).

Receptor	PDB ID	Resolution (Å)	Stabilizing technique	Ligand	Ionic lock[a]	Ref.
β_2AR	2RH1	2.4	T4-lysozyme	carazolol	N	10
	2R4S	3.4/3.7	Monoclonal antibody	carazolol	N	9
	2R4R	3.4/3.7	Monoclonal antibody	carazolol	N	9
	3D4S	2.8	T4-lysozyme	timolol	N	12
	3NYA	3.1	T4-lysozyme	alprenolol	N	18
	3NY8	2.8	T4-lysozyme	ICI-118,551	N	18
	3NY9	2.8	T4-lysozyme	Novel inverse agonist[b]	N	18
	3KJ6	3.4	Monoclonal antibody	carazolol	N	46
	3P0G*[c]	3.5	Nanobody	BI-167,107	n/a	49
	3PDS^	3.5	T4-lysozyme	FAUC50	N	50
	3SN6*	3.2	Nanobody, T4-lysozyme	BI-167,107	n/a	21
β_1AR	2VT4	2.7	thermostabillized	cyanopindolol	N	47
	2Y00#	2.5	thermostabillized	dobutamine	n/a	51
	2Y01#	2.6	thermostabillized	dobutamine	n/a	51
	2Y02*	2.6	thermostabillized	carmoterol	n/a	51
	2Y03*	2.9	thermostabillized	isoprenaline	n/a	51
	2Y04#	3.1	thermostabillized	salbutamol	n/a	51
	2YCW	3.0	thermostabillized	carazolol	Y	52
	2YCX	3.3	thermostabillized	cyanopindolol	Y	52
	2YCY	3.2	thermostabillized	cyanopindolol	N	52
	2YCZ	3.7	thermostabillized	iodo-cyanopindolol	N	52
A_{2A}R	3EML	2.6	T4-lysozyme	ZM-241,385	N	15
	3QAK*	2.7	T4-lysozyme	UK-432,097	n/a	19
	2YDO*	3.0	thermostabillized	adenosine	n/a	53
	2YDV*	2.6	thermostabillized	NECA[d]	n/a	53
D_3R	3PBL	2.9	T4-lysozyme	eticlopride	Y	16
H_1R	3RZE	3.1	T4-lysozyme	doxepin	N	20

[a]Indicates if the ionic lock is intact (Y = yes, N = no, n/a = not applicable for agonist bound structures. [b]Novel inverse agonist ethyl 4-({(2*S*)-2-hydroxy-3-[(1-methylethyl)amino] propyl}oxy)-3-methyl-1-benzofuran-2-carboxylate (**5**).[54] [c]Symbols indicate agonist bound structures; *agonist, #partial agonist or ^irreversible agonist. [d]Adenosine-5′*N*-ethylcarboxamide.

Following the β_2AR structures in 2007, a steady stream of crystal structures of inactive rhodopsin-like GPCRs have been solved, including β_1AR,[47,52] A_{2A}AR,[15] D_3R,[16] CXCR4[17] and H_1R[20] bound to either antagonists or inverse agonists and these structures are discussed further in Section 1.4.3.1.

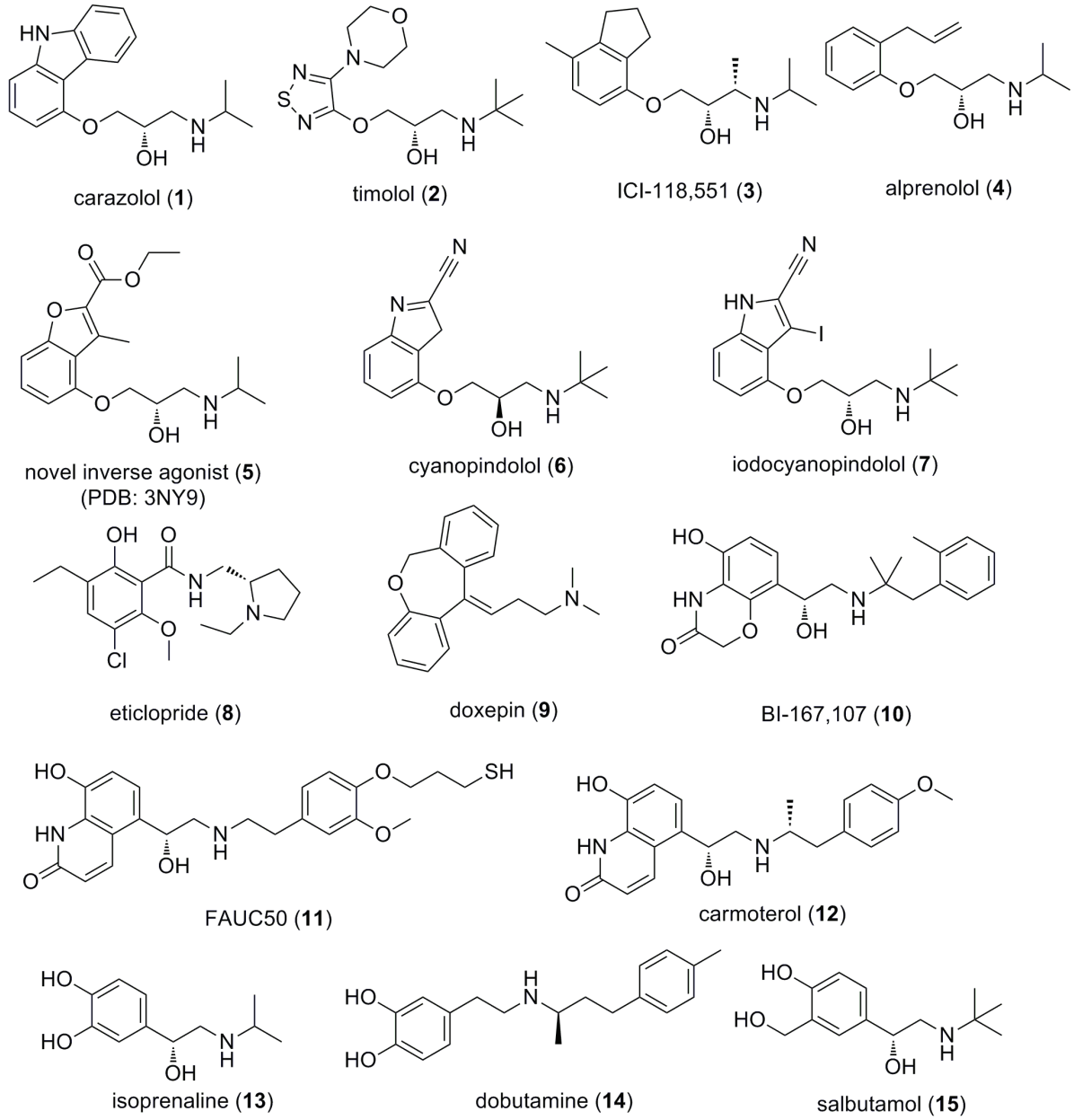


Figure 1.5 Structures of agonists, inverse agonists and antagonists co-crystallized in aminergic (adrenergic, dopamine and histamine) GPCR crystal structures.

The structure of β_2AR in complex with a camelid antibody fragment (nanobody) was the first structure of an active state of a human GPCR.[49] The nanobody binds to the intracellular G protein binding site, resulting in a 11 Å outward movement of TM6. Additional structures

in complex with agonists have been determined for the β_1AR, β_2AR, A_{2A}AR and opsin (active form of rhodopsin) receptors. An exciting new development in the determination of crystal structures in complex with agonists is the recently solved structure of β_2AR in complex with the Gs heterotrimer (G protein).[21] It is anticipated that these structures will further assist in the understanding of GPCR activation mechanisms and play a key role in structure-based drug design for agonists. These structures are discussed in further detail in Section 1.4.3.2.

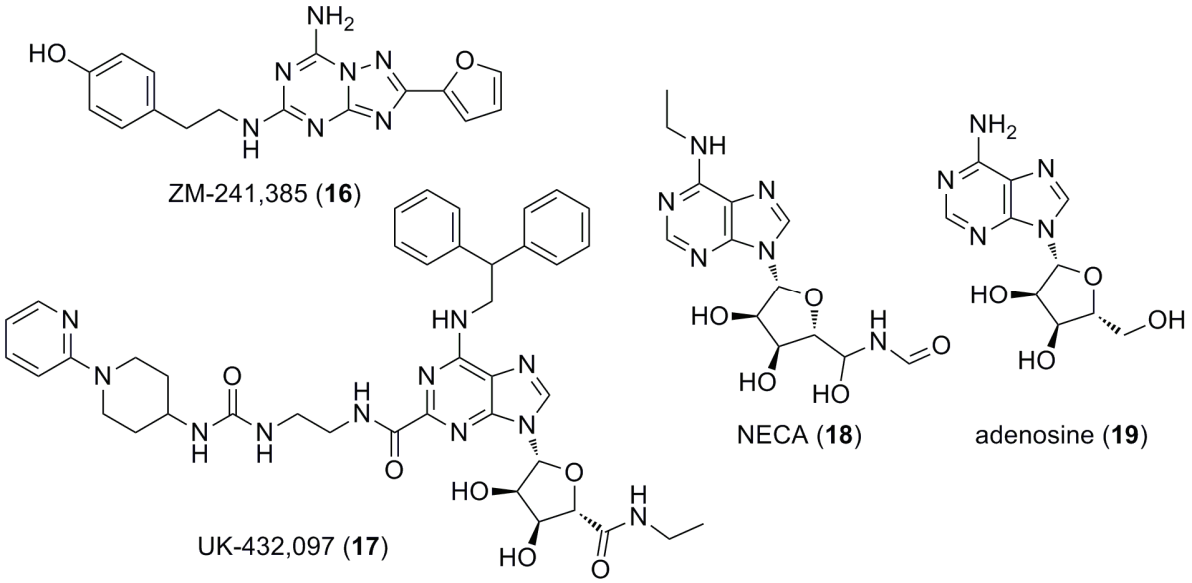


Figure 1.6 Structures of agonists and antagonists co-crystallized in adenosine GPCR crystal structures.

1.4.3 Description of GPCR crystal structures

1.4.3.1 Crystal structures of inactive GPCRs

Rhodopsin crystal structures. Palczewski et al.[7] solved the first high resolution crystal structure of a GPCR, bovine rhodopsin (Figure 1.7), which confirmed the previously predicted topology of a GPCR; that of seven transmembrane helices, linked by intra and extracellular loops. Several crystal structures of bovine (PDB ID: 1F88, 1L9H, 1GZM, 1U19, 2I37, 2J4Y, 3OAX, 3PXO, 3PQR)[7,55-59] and squid rhodopsin (PDB ID: 2ZIY, 2Z73)[11,60] have been solved. The ligand 11-*cis*-retinal binds covalently to Lys 7.43, forming a Schiff-base linkage (Figure 1.7b). In contrast to the bovine rhodopsin structures, TM5 and

TM6 are longer in the squid rhodopsin structures.[60] Due to the covalent nature of the ligand binding, retinal is bound tightly to rhodopsin with ECL2 forming a short β-sheet closing off the orthosteric site to prevent hydrolysis of the Schiff-base.[61] This can be problematic when using a rhodopsin crystal structure as a template for the development of homology models for receptors that bind diffusible ligands, as they often have larger, more solvent exposed binding sites (refer to Section 1.5.1).[62]

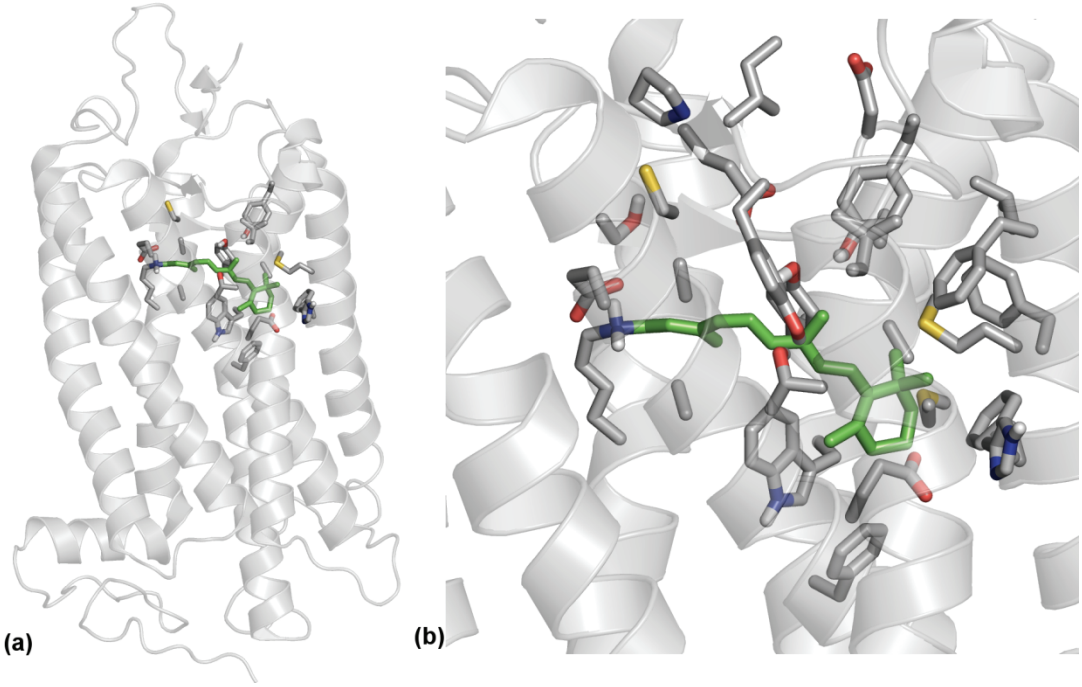


Figure 1.7 Crystal structure of rhodopsin (1U19),[56] displaying retinal in green and binding site residues within 4 Å of ligand (a) and close-up of binding site (b).

Inactive crystal structures of β_2AR. The first crystal structures of the human β_2AR were solved in 2007, in complex with the inverse agonist carazolol (**1**, Figure 1.5) and stabilized for crystallization using a monoclonal antibody (Figure 1.8a, PDB ID: 2R2S and 2R4S).[9] However, these structures lacked a fully resolved orthosteric site and extracellular loops, due to uninterpretable electron density as a result of significant anisotropy. Another structure of β_2AR in complex with carazolol (Figures 1.8b and 1.8c) was solved using the T4-lysozyme stabilization technique (PDB ID: 2RH1) at a higher resolution (2.4 Å), with a well-resolved orthosteric site and extracellular loops.[8,10] The 2RH1 structure revealed previously

unpredicted detail, such as a helix in ECL2 that stabilizes the loop and helps to keep the binding site open to the extracellular space. Additional structures of β_2AR have been determined using the T4-lysozyme technique in complex with other antagonists and inverse agonists; timolol (**2**), ICI-118,551 (**3**) alprenolol (**4**) and **5** (PDB ID: 3D4S, 3NY8, 3NY9, 3NYA, respectively). In the timolol-bound structure (3D4S), specific cholesterol binding sites were identified.[12]

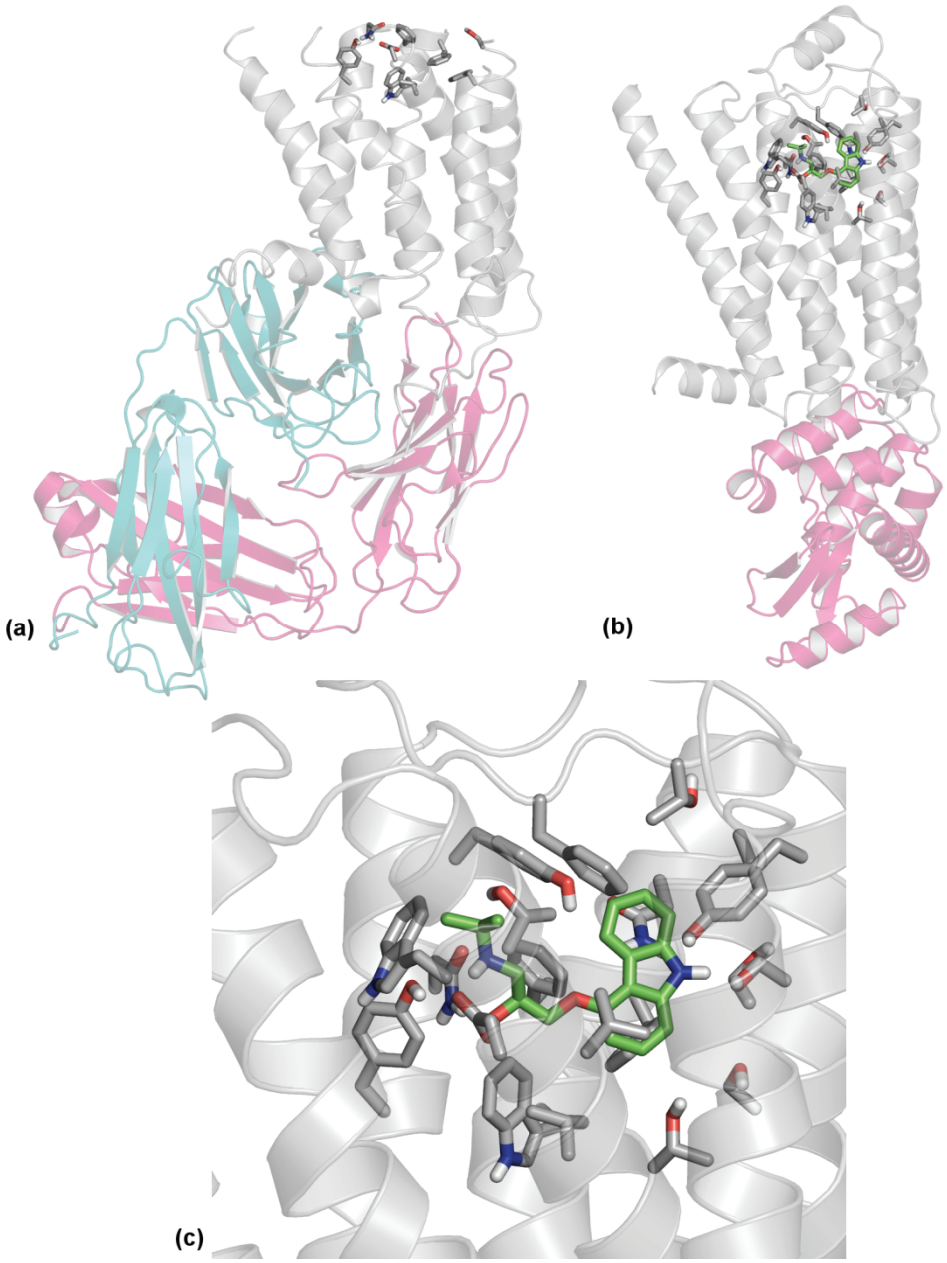


Figure 1.8 Crystal structures of β_2AR (a) 2R4S structure,[9] antibody colored in cyan and magenta (b) 2RH1 structure,[8,10] carazolol displayed in green, T4-lysozyme colored in magenta (c) close-up of orthosteric binding site (residues within 4 Å of ligand shown).

Inactive crystal structures of β_1AR. The first crystal structure of the turkey β_1AR (Figure 1.9), in complex with an antagonist, cyanopindolol (**6**), as the turkey receptor was more stable than the human receptor (PDB ID: 2VT4).[47] Additional structures have since been determined in complex with carazolol (**1**), cyanopindolol (**6**) and iodocyanopindolol (**7**) using the same thermostabilization technique (PDB ID: 2YCW, 2YCX, 2YCY, 2YCZ).[52] Similar to the β_2AR structure, β_1AR has a helix in ECL2 but in contrast to β_2AR, β_1AR has an additional short helix in ICL2, which interacts with the conserved D(E)RY motif. Additionally, in the more recent structures, ICL3 was resolved. Interestingly, in the more recent crystal structures TM6 extends further than in 2VT4, and has two distinct inactive conformations of β_1AR; straight and bent.[52] The β_1AR crystal structures with the bent conformation of TM6 (2CYW and 2CYX) are two of the few that display the ionic lock, albeit a weak ionic lock due to the distance between Glu 6.30 and Arg 3.50 (3.7-3.9 Å), compared to the ionic lock in rhodopsin (2.8-3.2 Å).

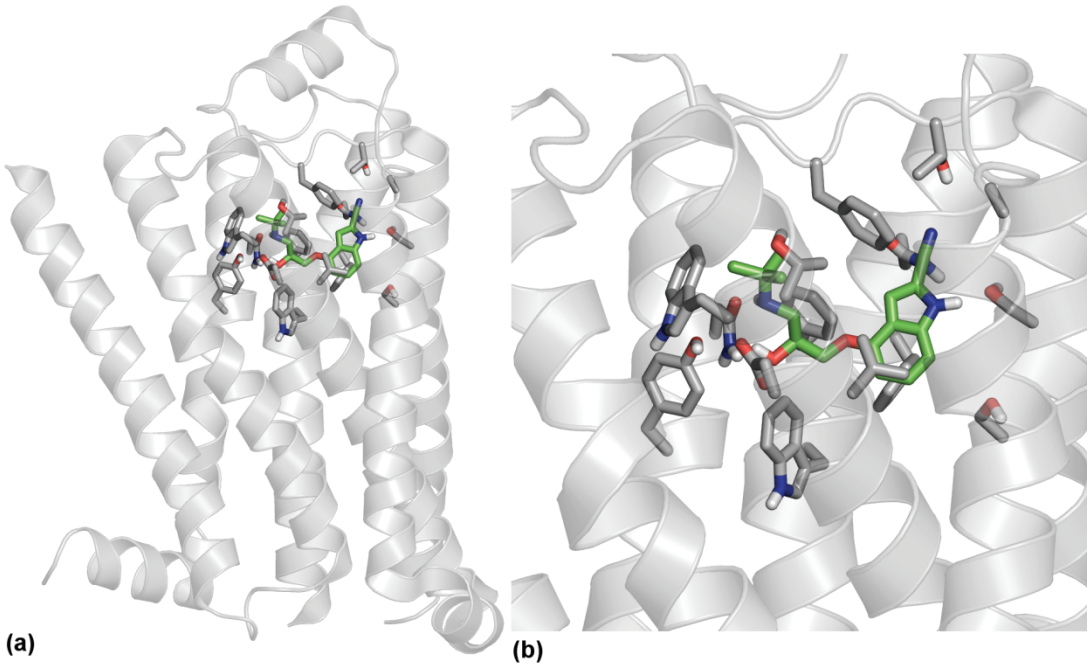


Figure 1.9 (a) Crystal structure of β_1AR (2VT4)[47], cyanopindolol shown in green (b) close-up of orthosteric binding site (residues within 4 Å of ligand shown).

Inactive crystal structure of the adenosine A_{2A} receptor ($A_{2A}AR$). A crystal structure of $A_{2A}AR$ (Figure 1.10) was determined in complex with the antagonist ZM-241,385 (**16**, Figure 1.6) and stabilized by a T4-lysozyme, revealed a different binding mode for the ligand in the orthosteric binding site, relative to rhodopsin and the β-adrenergic receptors.[15] In the $A_{2A}AR$ structure, ZM-241,385 binds perpendicular to the membrane and points into the extracellular space, making significant contacts to ECL2. Consistent with site-directed mutagenesis for this receptor, ZM-241,385 makes key hydrogen bonding interactions to Glu 169 (ECL2) and Asn 6.55. Unlike in the rhodopsin and the β-adrenergic receptor structures, a network of water molecules was observed in the $A_{2A}AR$ crystal structure, with some of these water molecules involved in mediating hydrogen bonds between the receptor and ligand.

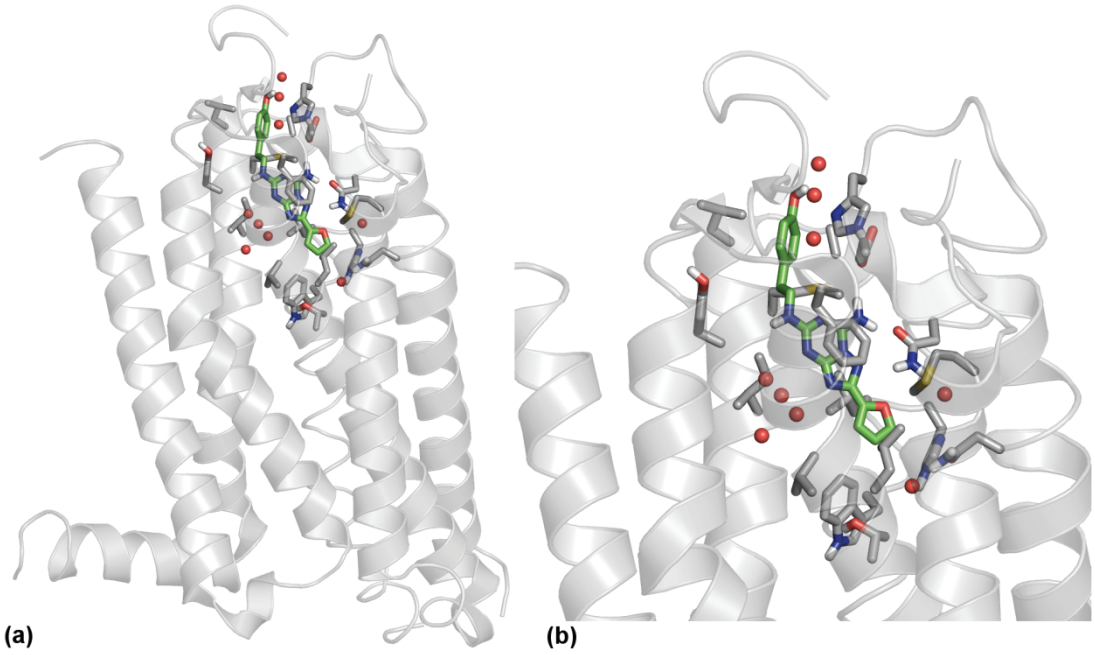


Figure 1.10 (a) Crystal structure of $A_{2A}AR$ (3EML),[15] ZM-241,385 shown in green (b) close-up of orthosteric binding site (residues within 4 Å of ligand shown). Water molecules in orthosteric binding site displayed as red spheres, T4-lysozyme omitted for clarity.

Inactive crystal structure of D_3R. A structure of the D_3R was solved in complex with the antagonist eticlopride (**8**) at 2.9 Å (Figure 1.11, PDB ID: 3PBL).[16] Overall, the structure is similar to the β-adrenergic receptor structures. However, unlike in the β-adrenergic

structures, ECL2 in the D_3R structure does not contain an ordered secondary structure because this loop is short in D_3R. Apart from the rhodopsin structures, the crystal structure of the D_3R was the first to display an intact ionic lock (Figure 1.11a). Additionally, ICL2 forms two distinct conformations in the different chains of the D_3R present in the crystal structure; in chain A ICL2 forms a short helix (Figure 1.11a), in chain B it is unstructured.

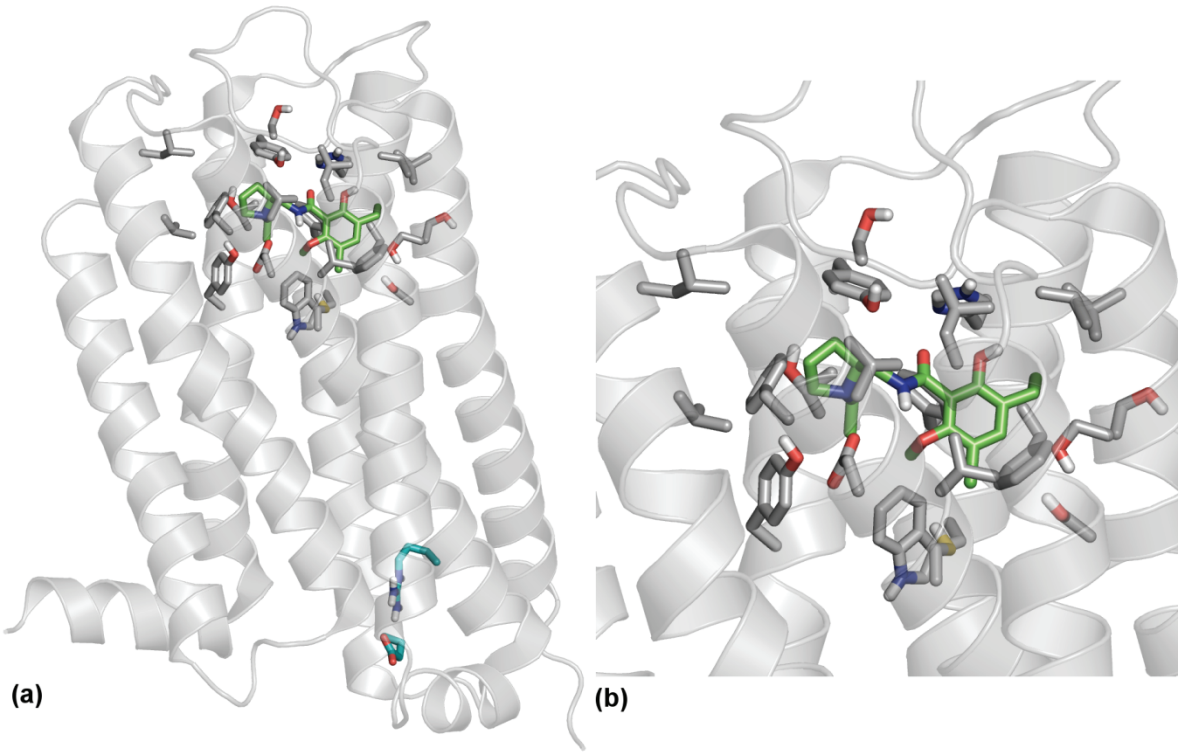


Figure 1.11 (a) Crystal structure of D_3R (3PBL),[16] with eticlopride shown in green and ionic lock shown in cyan and (b) close-up of binding site. Note: T4-lysozyme was omitted from image for clarity.

Inactive crystal structure of the histamine H_1 receptor (H_1R). The H_1R crystal structure has been determined (Figure 1.12) in complex with an antagonist, doxepin (**9**).[20] Unlike the ligands in the adrenergic and dopaminergic structures, the tricyclic ring of doxepin binds deeper in the binding site. Doxepin also directly interacts with Trp 6.48 and does not interact directly with ECL2. No ionic lock was observed in this structure. The crystal structure of the H_1R has revealed an anion binding region adjacent to the orthosteric site, which is not conserved in other aminergic GPCRs. In this crystal structure, a phosphate ion binds to the anion binding site, consisting of residues Lys 5.39 and Lys 179 (ECL2) and docking studies

have demonstrated that the second generation antihistamines interact with this site through a carboxyl group. Additionally, an increased distance between extracellular ends of TM3 and TM5 allows binding of larger ligands.

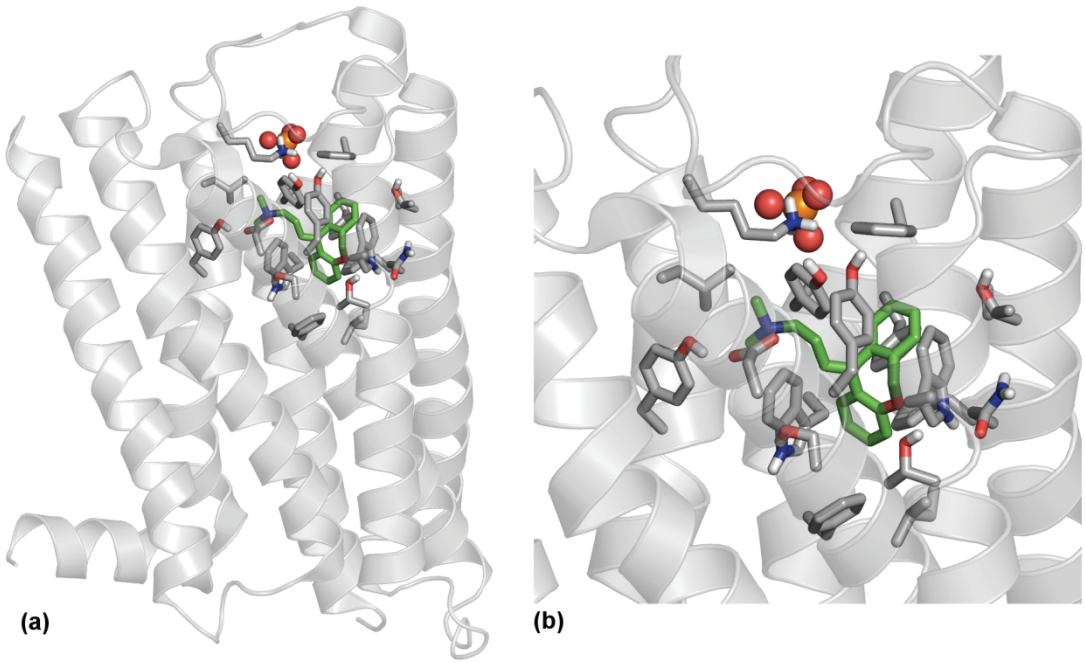


Figure 1.12 (a) Crystal structure of H_1R (3RZE),[20] with doxepin shown in green and the phosphate ion binding to the anion binding site shown as spheres; (b) shows a close-up of the orthosteric binding site. Note: T4-lysozyme was omitted from image for clarity.

1.4.3.2 Crystal structures of GPCRs in the active state

Interestingly, a number of GPCRs structures have been determined in their active state, using some of the crystallization strategies discussed in Section 1.4.1. Examination of these structures and comparison with the inactive structures has provided vital insight into the mechanism of GPCR activation and the ensuing structural rearrangements that occur. Although our research is primarily focused on inactive GPCR structures, an appreciation of the active states of these receptors can assist in the design of ligands that preferentially bind to a specific receptor state (e.g. active or inactive). For this reason, these structures are briefly discussed below.

Opsin crystal structures. The first crystal structure of an active GPCR was the structure of opsin in 2008 (PDB ID: 3CAP),[13] the active state of rhodopsin without the ligand retinal.

Another structure was solved of opsin attached to a peptide that mimics the G protein α-subunit (PDB ID: 3DQB),[14] and these structures give insight into the mechanism of activation of rhodopsin.

Active crystal structures of β_2AR. The first crystal structure of a human GPCR bound to an agonist was determined for the β_2AR (Figure 1.13a, PDB ID: 3P0G) in complex with a nanobody and the small molecule BI-167,107 (**10**).[49] The active state of β_2AR is inherently unstable, as agonists do not effectively stabilize the receptor in this state, particularly if there is no G protein attached. A nanobody that preferentially binds to active state of a GPCR at the G protein binding site was used to stabilize the active state of the receptor. Only minor changes were noted in the orthosteric binding site. Significant changes were observed for the cytoplasmic side of receptor, with TM6 moving by 11 Å, as well as rearrangements of TM5 and TM7.[49] A structure of β_2AR covalently bound to an agonist (FAUC50, **11**) by a disulfide bond (PDB ID: 3PDS) has also been determined.[50]

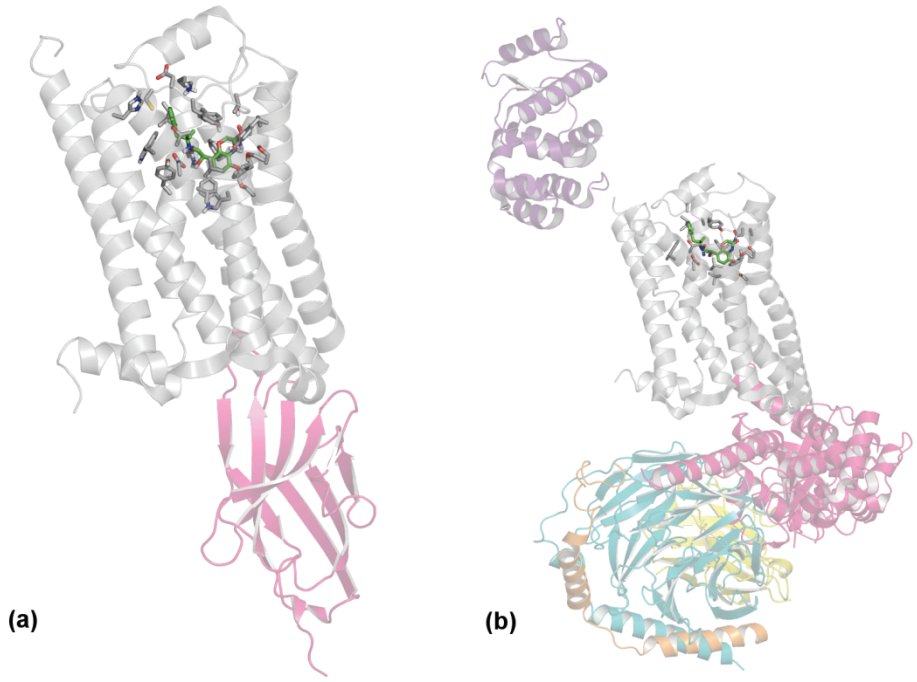


Figure 1.13 Crystal structures of β_2AR in the active state, bound to the small molecule BI-167,107, displayed in green (a) stabilized by a nanobody, shown in magenta (3P0G),[49] and (b) structure 3SN6 stabilized extracellularly by a T4-lysozyme (violet) and bound to the heterotrimeric G protein (highlighted in magenta (α), cyan (β) and orange (γ)), which is stabilized intracellularly by a nanobody (yellow).

In 2011, the first crystal structure of a human GPCR in complex with the Gs heterotrimer (G protein) was determined, using β_2AR in complex with BI-167,107 (**10**, PDB ID: 3SN6, Figure 1.13b).[21] Two different stabilization strategies were employed to obtain this crystal structure; a T4-lysozyme was attached to the amino terminus of the GPCR, which stabilized the extracellular side of the receptor and a nanobody stabilizes the intracellular side of the receptor through interactions with the G protein. Interestingly, the structure is similar to the 3P0G structure, with more significant variations on the intracellular side of the receptor; with TM6 moving outward by 14 Å and an extension of the TM5 α-helix. This crystal structure will significantly contribute to our better understanding of the interaction between a GPCR and its corresponding G protein.

Active crystal structures of β_1AR. Five crystal structures of the β_1AR have been solved (Figure 1.14, PDB ID: 2Y00, 2Y01, 2Y02, 2Y03, 2Y04)[51] in complex with full (carmoterol, **12**; isoprenaline, **13**) and partial agonists (dobutamine, **14**; salbutamol, **15**).

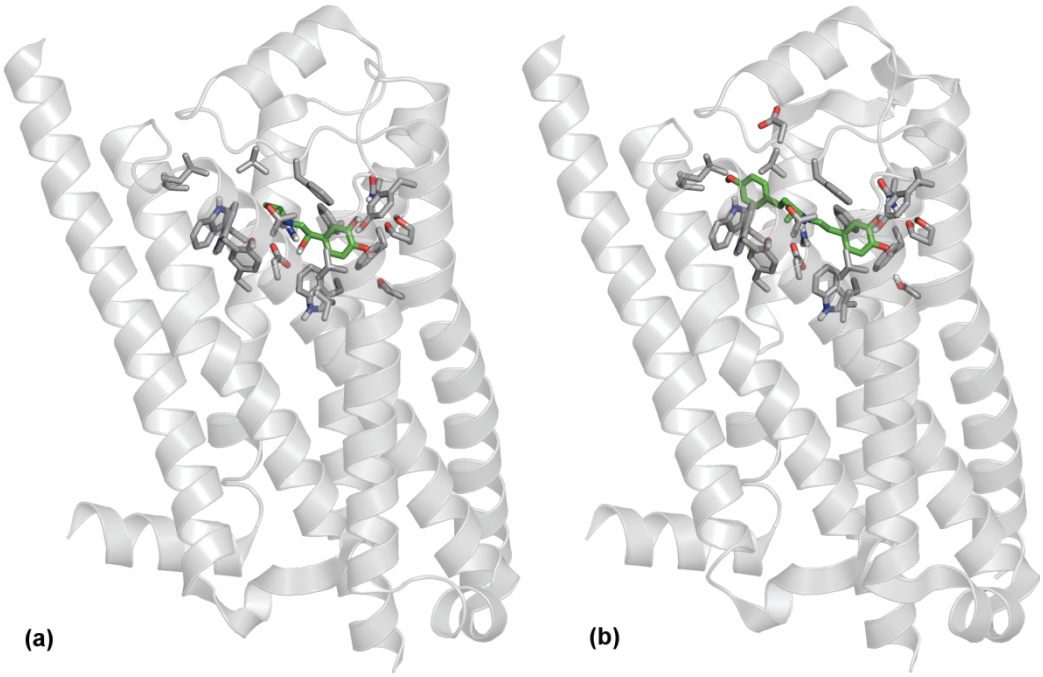


Figure 1.14 Crystal structure of β_1AR in the active state, bound to agonists (displayed in green); (a) isoprenaline (2Y03) and (b) dobuatamine (2Y00).[51]

In the β_1AR structures, three main changes were observed upon agonist binding compared to the antagonist bound structure; a conformational change in side chains Ser 5.43 and Ser 5.46 and a contraction of orthosteric binding site by ~1 Å.[51] Interestingly, differences were observed between full and partial agonist binding to TM5, where full agonists formed hydrogen bonds to residues Ser 5.42 and Ser 5.46, whilst partial agonists only formed a hydrogen bond to Ser 5.42. Also, of particular interest was the binding mode of the partial agonist dobutamine (Figure 1.14b), as the extended structure of the ligand formed interactions to a less conserved area of the receptor.

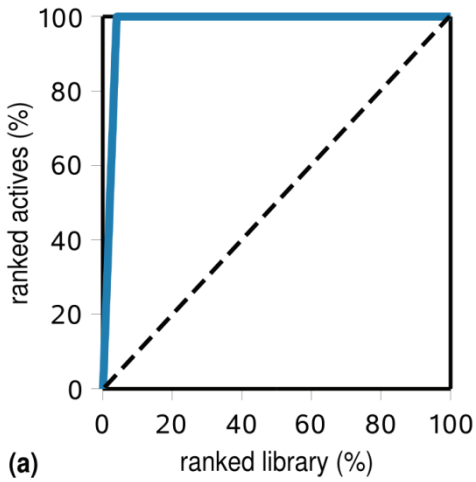
Active crystal structures of $A_{2A}AR$. The crystal structure of the A_{2A}AR has been determined in complex with three agonists. The first structure determined was bound to UK-432,097 (**17**, PDB ID: 3QAK),[19] with the TM helices moving slightly upon agonist binding, however not as significantly as those observed in β_2AR. Two additional structures, bound to NECA (**18**, PDB ID: 2YDV) and adenosine (**19**, PDB ID: 2YDO) have also been determined.[53]

1.4.4 Use of crystal structures in structure-based drug design and virtual screening

With the increasing number of crystal structures of class A GPCRs, it is not surprising that they are being used to identify novel compounds to target these receptors. These recent crystal structures are being used to aid the drug design process, rather than primarily for a retrospective analysis.[63] Although we have focused on structure-based drug design for GPCRs, it should be noted that these techniques are applicable to and have often been developed using other protein classes.

Initially, the crystal structures of GPCRs can and have been used for virtual screening studies, where a large virtual library of compounds is docked into the receptor and ranked by a scoring function. Ideally, compounds that are active at the receptor are ranked higher in the virtual screening study than the inactive compounds. There are a number of ways to assess the success of a virtual screening campaign. These include enrichment plots, which are a

graph of true positives vs. the virtual library (Figure 1.15a), from which enrichment factors can be calculated (Figure 1.15b, a high enrichment factor indicates more active compounds are identified).[64,65] This data can also be plotted as receiver operating characteristic (ROC) curves (graph of true positives vs. true negatives), and high values of the area under the curve (AUC) are indicative more active compounds are being identified in the virtual screen.[65] Ultimately, the goal of a virtual screening campaign is to identify novel active compounds for a given receptor. Some virtual screening campaigns have already been undertaken using the recent crystal structures of β_2AR and A_{2A}AR.



$$EF^{x\%} = \frac{(actives_{sampled})}{(library_{sampled})} \div \frac{(actives_{total\_database})}{(library_{total\_database})}$$

(b)

Figure 1.15 (a) Enrichment plot with an ideal graph shown in blue (actives identified before decoys) and a random enrichment shown as a dashed line; (b) the equation for calculating enrichment factors.

Two virtual screening studies have been reported using β_2AR. Topiol and Sabio docked two databases of approximately 400,000 and 4 million compounds, respectively, into β_2AR (PDB ID: 2RH1), and in the top 100 compounds, ligands were identified that occupied a similar region to the co-crystallized ligand.[66] Additionally, some of the identified compounds represented novel chemical classes, compared with known β-blockers. A selection of compounds from each virtual screening study were tested pharmacologically, identifying a number of active compounds, with activity ranging from low micromolar to low nanomolar.[67] In another virtual screening study with β_2AR (PDB ID: 2RH1), Kolb et al. docked a database of approximately 1 million lead-like compounds and 25 of the top

ranking compounds were tested pharmacologically, of which six compounds showed affinity ranging from a K_i of 9 nM to 3.2 μM.[54] One of the active compounds identified in this study has subsequently been co-crystallized with β_2AR.[18]

Katritch et al. carried out virtual screening studies using the A_{2A}AR (PDB ID: 3EML) and screening approximately 4.3 million compounds.[68] 56 high scoring compounds were tested in radioligand binding assays and 23 displayed K_i values under 10 μM, with two less than 60 nM, with a number of novel scaffolds identified. Carlsson et al. carried out a similar study, docking 1.4 million compounds and testing 20 compounds using radioligand displacement assays, where 7 ligands showed promising K_i values between 200 nM and 9 μM.[69] It is noteworthy that the most potent compound in this study was 50-fold more selective for A_{2A}AR than A_1AR and A_3AR.

From a structure-based drug design perspective, a drug-receptor complex can be used to optimize a lead compound.[63] An excellent example of lead optimization using the recent crystal structures of GPCRs, was the design of the irreversible agonist, FAUC50, using the 2RH1 crystal structure.[50]

Due to the ever increasing number of GPCR structures available, in both active and inactive states, structure-based drug design for GPCRs is becoming more practicable. Crystal structures in different states will assist in the elucidation of the mechanism of action of GPCRs. In addition to the virtual screening studies already discussed, crystal structures of GPCRs can be used to assist in the evaluation of novel compounds, such as elucidating potency and selectivity profiles.[70,71] This is particularly useful when investigating a family of receptors with closely related subtypes. However, compared to the number of druggable GPCR targets (~350 receptors),[5] the relative number of GPCRs for which crystal structures are available remains quite small (7 receptors). Because obtaining high resolution crystal structures can require years of work, homology models of target GPCRs can be utilized for structure-based drug design.[72] Additionally, there are now a number of templates that can be

used for the development of homology models that share higher homology with many GPCR drug targets, which ideally can assist in improving the quality of the models.

1.5 Homology models of GPCRs

1.5.1 Homology models based on the rhodopsin crystal structure

Following the determination of the structure of bovine rhodopsin in 2000,[7] and due to the similarity in topology of GPCRs, the crystal structure of bovine rhodopsin was widely used as a template to build homology models of other GPCRs. A large number of homology models based on the rhodopsin structure were built,[73-89] greatly assisting drug design campaigns for GPCRs.

Homology modeling is a technique used to generate a model of a target protein for which there is no 3D structure available and is reliant on structural homology within protein families; that is that related proteins have similar topology.[63] A 3D structure of a related protein is used as a template for modeling the target protein and high sequence homology is required: generally over 50% homology is considered good for the development of models for drug discovery,[90] although sequence identities over 30% are considered reasonable for GPCRs.[72]

Whilst the majority of GPCRs that are of interest from a drug discovery perspective share relatively low homology to rhodopsin (~20-30%),[72] this is somewhat compensated for by the overall similar topology of bovine rhodopsin to other GPCRs (the similar topology is confirmed by the recent crystal structures).[90] As a result of the low sequence identity, amongst other issues, the use of rhodopsin as a template for pharmaceutically relevant GPCRs was debated in the literature[91,92] and methods were developed to refine and improve upon the initial homology models based on the bovine rhodopsin structure.

A significant concern for homology models based on the bovine rhodopsin template was the positioning of ECL2, enclosing the orthosteric site, and the small size of the ligand cavity.[32,92-94] Additionally, the mechanism of endogenous ligand binding to rhodopsin is

different to other GPCRs, in that it binds covalently to 11-*cis*-retinal, whereas the majority of GPCRs bind diffusible ligands. This binding mechanism restricts the size of the orthosteric binding site in any homology model based on the rhodopsin crystal structure. To compensate for this, ECL2 has either been omitted from the homology model altogether,[72] the loop has been optimized using loop refinement protocols[95] or the binding site expanded by either ligand-steered homology modeling[96,97] or pressure-based steered molecular dynamics.[62] Despite a number of drawbacks, rhodopsin was the only available GPCR template for seven years and was used to successfully develop a number of homology models for structure-based drug design.[79,80,89,98,99]

1.5.2 Homology models of GPCRs based on non-rhodopsin structures

The determination of a number of crystal structures of rhodopsin-like GPCRs, starting with the crystal structure of β_2AR in 2007,[8-10] has significantly added to the understanding of this receptor class. Currently, there are crystal structures available for seven GPCRs, namely rhodopsin, β_1AR, β_2AR, A_{2A}AR, D_3R, H_1R and CXCR4 in an inactive state, as well as opsin β_1AR, β_2AR and A_{2A}AR in an active state (refer to Table 1.1, Section 1.4.2). From a structure-based drug design perspective, there is now a range of medium to high resolution structures that can be used for drug design purposes, including the development of new homology models using a template that shares higher sequence homology with a receptor of interest.

There are significant advantages in using templates other than rhodopsin, particularly for the development of homology models for class A GPCRs, as they often share higher homology with the non-rhodopsin receptors with solved crystal structures. For example, where a suitable template is defined as sharing > 30% homology with the target protein, β_1AR, β_2AR and A_{2A}AR crystal structures proved to be more suitable templates for developing homology models of a larger percentage of class A non-olfactory, non-orphan GPCRs (18, 16 and 12%, respectively), compared to rhodopsin (2%).[72]

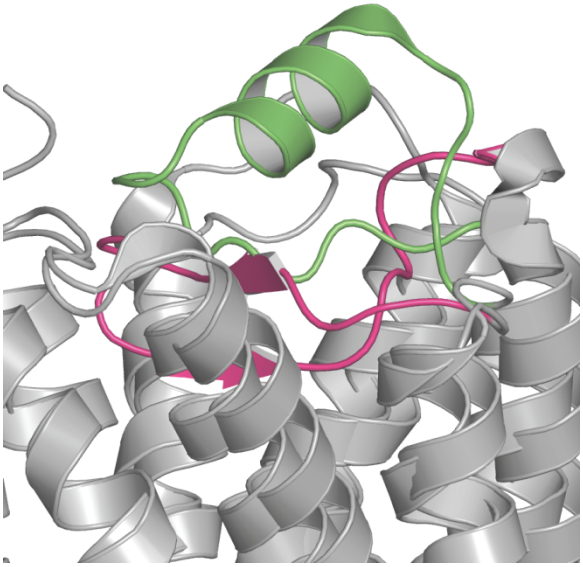


Figure 1.16 Comparison of the structure of ECL2 in the rhodopsin crystal structure (magenta, 1U19) vs. the β_2AR crystal structure (green, 2RH1).

Additionally, the β_1AR or the β_2AR crystal structures proved to be much better templates than rhodopsin for the development of homology models of biogenic amine receptors and not just because of the higher sequence homology, but because they bind ligands in a similar manner (i.e. diffusible ligands). As a result, the orthosteric site in these crystal structures larger and only partially enclosed by ECL2, leaving the orthosteric site exposed to the extracellular space (Figure 1.16). The placement of ECL2 above the orthosteric site (between the conserved disulfide bond and TM5) is also very similar in the D_3R and H_1R crystal structures. Examples of homology models of GPCRs based on non-rhodopsin templates are outlined in Table 1.2. Some of theses studies are discussed in Section 1.5.3.

Table 1.2 Examples of homology modeling studies using homology models of human GPCRs based on a non-rhodopsin template, identifying the crystal structure template used, as well as the modeling programs employed.

Receptor(s)	**Template**	**Modeling Programs**	**Docking Program**	**Ref.**
A_1AR and A_{2A}AR	2RH1	CAChe	CAChe	100
5-HT_{1A}R, 5-HT_{2A}R, 5-HT_{2B}R, 5-HT_{2C}R, 5-HT_6R, 5-HT_7R, α_1AR, α_2AR, D_2R, D_3R, D_4R, H_1R, M_1 mAChR, M_4 mAChR	2RH1	MODELLER	GOLD	101
5-HT_{2A}R	2RH1	MOE	AutoDock	102,103
5-HT_4R	2RH1	-	-	104
MC_4R	2RH1	QUANTA	QUANTA	105
H_4R	2RH1	MOE	-	106
H_1R	2RH1	SYBYL	Manual docking	107
NK_1R	2RH1	MODELLER	Glide	108
CCK_1R	2RH1 3EML	ICM	ICM	109
A_{2A}AR	2RH1	MOE	IFD, Glide MOE	110
CCR_5R	2RH1	InsightII	GOLD	111
Secretin	2RH1	ICM	ICM	112
D_1R	2R4R	Swiss-model/ Deep view	GRAMM	113
M_2R	2RH1	ICM	ICM	114
H_4R	2RH1	MODELLER	-	115
$P2Y_{14}$R	3EML	Prime	Glide	116
D_2R, D_3R, D_4R	2RH1	MODELLER	AUTODOCK	117
5-HT_{2A}R, H_1R	2RH1	MODELLER	GOLD	118
5-HT_{2A}R	2RH1	MODELLER	GOLD ICM Glide	119
D_3R	2RH1	MODELLER	GOLD	120

Table 1.2 continued

Receptor(s)	Template	Modeling Programs	Docking Program	Ref.
D_2R, D_3R	2RH1	MODELLER	GOLD	121
A_3AR	3EML	MOE	MOE-DOCK GOLD Glide Plants	71
$\alpha_{1A}AR$	2R4R	MODELLER	DOCK	122
$A_{2B}AR$	2RH1 3EML	Homer	FlexX	123
CB_1R	2RH1	MODELLER	-	124
H_4R	2RH1	SYBYL	-	125
Orexin receptor	2RH1	MODELLER	-	126
M_2R	2RH1	ICM	ICM	127
H_4R	2RH1	GPCRgen	GOLD	128
H_2R	2VT4	Prime	IFD	129
D_2R, D_3R	2RH1	MODELLER	-	130
$5\text{-}HT_{2A}R$	2RH1	SYBYL	Manual docking	131
D_2R	2RH1	SYBYL	GOLD	132
D_2R	2RH1	MODELLER	GOLD Glide	133
$5\text{-}HT_{2A}R$	2RH1	MODELLER	Glide	134
$\alpha_{2A}AR$, $\alpha_{2B}AR$, $\alpha_{2C}AR$	2RH1	MODELLER	IFD Glide	135

1.5.3 Development and refinement of homology models for GPCRs

Due to the increase in the number of high resolution crystal structures of GPCRs, the number of homology models of GPCRs published is rapidly growing each year.[63] The development of homology models is becoming a more routine process in drug discovery, as there are now more suitable templates to build structures. However, the sequence identity

between templates and targets still remains relatively low, which is mainly due to the highly variable loop regions. There are a few key considerations that need to be addressed when developing homology models of GPCRs. These are outlined below.

1.5.3.1 Selecting the appropriate template for GPCR homology models

Currently, one of the biggest questions when developing homology models of GPCRs is "which template should I use?". This problem has been addressed by an elegant study by Mobarec et al.,[72] where the crystal structures of rhodopsin, β_1AR, β_2AR and A_{2A}AR were assessed for their suitability as templates for the development of homology models of class A GPCRs. The β_1AR, β_2AR and A_{2A}AR crystal structures proved to be more suitable templates for developing homology models of a larger percentage of class A non-olfactory, non-orphan GPCRs (18, 16 and 12%, respectively), compared to rhodopsin (2%).[72] This choice is now expanded further to include the high resolution crystal structures of GPCRs in the active state (opsin, β_1AR, β_2AR and A_{2A}AR), as well as the additional inactive structures of D_3R, CXCR4 and H_1R.

The choice of template for GPCR homology models is now not only dependent on the target sequence but also the types of ligands under investigation (i.e. agonists or antagonists), with the most appropriate template available selected at the time of model development. Additionally, with the increasing number of crystal structures, multiple templates for homology model can also be considered.[72]

1.5.3.2 Developing models of extracellular loop 2

One of the major difficulties for the development of homology models of GPCRs is modeling the variable loop regions, which is particularly evident when developing models based on the rhodopsin crystal structure.[136] In fact, in some of the earlier rhodopsin-based models, the loops have been omitted altogether due to difficulties experienced in predicting their structure.[72,93,99,137]

In many of the crystal structures (rhodopsin, β_1AR, β_2AR, A_{2A}AR, D_3R, CXCR4) ligands interact with ECL2, which acts as a "gatekeeper" to the orthosteric binding site and is involved in ligand specificity.[138] As such, the modeling of these loop regions can be crucial to the development of GPCR homology models that are useful for structure-based drug design.

When developing protein models, the loop regions can be modeled either ab initio or based on a template. To develop a homology model of a loop, a reasonable template (usually a crystal structure) is required. For GPCRs with highly variable loop regions, this is challenging as there are no suitable templates.[139] Loop refinement can be used to improve the quality of a GPCR homology model.[93,139,140] However, loop modeling generally gives better results for shorter loops (< 12 residues),[95,139-141] because the difficulty of loop prediction (i.e. amount of conformational search space) increases rapidly with increasing loop length. As the average length of ECL2 is 20-35 residues,[139,142] the development of accurate models of this loop is incredibly difficult. A number of ab initio methods have been developed to generate models of loops that can lead to improved models of GPCRs.[93,139,140] Ultimately, more crystal structures are required so that additional templates are available for improved modeling of the loop regions.

1.5.3.3 The use of site directed mutagenesis data in the development of homology models of GPCRs

Site-directed mutagenesis data has provided a wealth of information regarding functionally significant residues and key motifs in GPCRs.[143] It involves selectively mutating residues and assessing the effects of mutation, usually the ligand binding properties, to assess the role of a specific residue. Commonly alanine scanning is used, where residues are sequentially mutated to an alanine residue, and if ligand binding is significantly reduced it is inferred that this residue is involved either in ligand binding or maintaining receptor stability. Additional techniques include the substituted cysteine

accessibility method (SCAM), which is used to determine whether a side chain is solvent accessible. This method is particularly useful in identifying residues lining binding site cavities.[143]

These techniques can be exemplified by a number of detailed studies for the dopamine D_2 receptor, where key binding site residues (such as Asp 3.32) were initially identified using alanine site-directed mutagenesis studies.[144-146] This work was followed up with extensive SCAM studies to identify residues lining the binding site cavity.[41,147-151]

As site directed mutagenesis data can be invaluable in determining key residues, it is commonly used in the refinement of homology models of GPCRs. However, the extent to which this data should be used in homology model refinement is debated, as it can be misleading (i.e. residues identified in site-directed mutagenesis may not be directly involved in ligand binding, refer to Section 1.5.4.2).[152]

1.5.3.4 Binding site optimization techniques

Due to the plasticity of the orthosteric binding site,[10] a number of refinement techniques have been developed to optimize the binding sites of GPCR homology models. The size of the binding site cavity was a significant problem for homology models based on rhodopsin, as the binding sites of many models were too small to dock active ligands identified in binding assays.

These methods include flexible receptor docking and molecular dynamics techniques. Cavasotto et al. used a ligand-steered homology modeling approach, where ligands that are known to be active at that receptor are used to shape the binding site through docking and energy minimization.[96] The Induced Fit Docking protocol, available as part of the Schrodinger software package,[153,154] similarly allows for flexibility in the side-chains of binding site residues and has been used for binding site optimization in GPCR homology models.[94,110,129,155] Molecular dynamics simulations can also be used to apply pressure to increase the size of the orthosteric binding site.[62,156] These techniques have allowed for the

generation of homology models of GPCRs that take into consideration the induced fit effect of ligand binding.[94,96,110,129,155] Some of these binding site optimization techniques have been explored in Chapters 2 and 3.

1.5.4 Evaluation of homology models and homology modeling methods of GPCRs

Despite the increasing number of GPCR crystal structures available, we are still heavily reliant on GPCR homology models for structure-based drug design. As such, it is important that we evaluate our modeling methods, using techniques such as virtual screening studies and ultimately comparison with high resolutions crystal structures.

1.5.4.1 Virtual screening evaluation

Many virtual screening studies have been carried out using GPCR homology models, either to evaluate structures or to identify novel drug candidates. It is becoming increasingly apparent that homology models of GPCRs are suitable for virtual screening experiments.

Small scale virtual screening can be used to evaluate homology models of GPCRs, by docking a small library (often a few thousand compounds) of inactive compounds that also contains a number of compounds that are active at the receptor studied.[78,96,129,133,157,158] This method is used to assess the viability of a model for large scale virtual screening purposes, which can involve docking millions of compounds from a virtual database such as ZINC.[159] Following homology model optimization and evaluation, large scale virtual screening can be undertaken to identify novel chemotypes, and has been used for a successfully for a number of GPCR targets.[79,84,97,98,160]

1.5.4.2 GPCR Dock modeling assessment

The determination of new GPCR crystal structures (discussed in Section 1.4), brings with it the opportunity for the GPCR modeling community to evaluate current structure prediction and ligand docking methods. Prior to the release of the coordinates and manuscripts for two GPCR crystals structures, $A_{2A}AR$ (PDB ID: 3EML)[15] and D_3R (PDB

ID: 3PBL),[16] two separate studies were undertaken to evaluate the current status of GPCR modeling methods.

In 2008, a community-wide critical assessment of GPCR modeling and docking methods (GPCR Dock 2008) was run, where participants were required to blindly predict the structure of $A_{2A}AR$ in complex with the ligand ZM-241,385 (**16**).[161] One of the key findings of this comparative study was that the accurate prediction of the loops, particularly extracellular loop 2 (ECL2), remained one of the more difficult aspects of GPCR homology modeling.[152,161]

A similar assessment was conducted in 2010 (GPCR Dock 2010), where participants were asked to generate models of D_3R in complex with eticlopride and of the CXCR4 chemokine receptor complex with a small molecule antagonist and with a cyclic peptide antagonist.[162] The key finding from GPCR Dock 2010 was that a combination of modeling techniques and biochemical and QSAR data were able to generate receptor-ligand complexes with almost similar accuracy to the experimental structures for proteins, where there was a template of reasonable homology to the target protein (i.e. D_3R). Our participation in GPCR Dock 2010 is discussed in Chapter 3.

1.6 Dimers and higher order oligomers of GPCRs

Until recently, GPCRs were commonly thought to act solely as monomeric proteins, however there is gathering evidence that GPCRs can form dimers and/or higher order oligomers.[163,164] Dimerization is now thought to be a common feature to the superfamily of GPCRs.[165,166] Using the currently available assays it is often not possible to distinguish between the dimers or higher order oligomers[167] and for the purposes of this research, dimers will be explored even though it may be possible that the receptor complexes contain higher order oligomers.

There are two forms of GPCR dimers (Figure 1.17). Homodimers form when two monomeric proteins of the same receptor associate. Heterodimers result from two different GPCR monomers interacting. However, since GPCRs are active as monomeric proteins,[167] it is currently uncertain whether GPCR dimerization is required for activity, particularly for class A GPCRs.[168] Although, in oligomerization experiments for rhodopsin, the rhodopsin dimer was found to be more active than the monomer.[169] The only unambiguous evidence that GPCR dimerization is required for activity is for class C GPCRs, such as the $GABA_B$ receptors.[164,170-172]

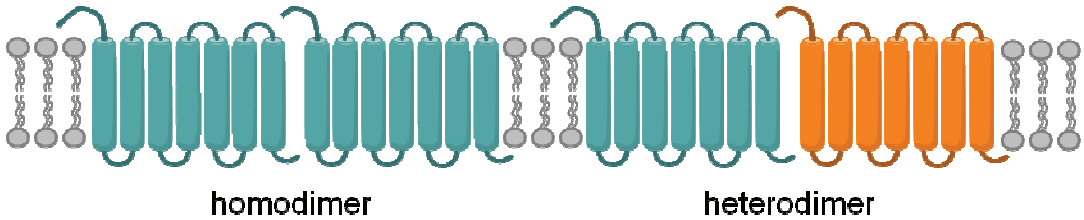


Figure 1.17 A schematic of a homodimer (two of the same monomers) and a heterodimer (two different monomers) embedded in a phospholipid bilayer.

1.6.1 Mechanism of dimer formation

GPCR dimerization is far from being understood, however two mechanisms of dimer formation have been proposed; domain-swapped dimers and contact dimers (Figure 1.18).[163] Domain swapped dimers occur when a monomer is divided into two domains and one domain is exchanged with the corresponding structure of an adjacent monomer, resulting in the interlocking of TM helices of two adjacent receptors.[165,167,173]

Contact dimers form through interactions on the surface of monomeric receptors without greatly changing the 3D structure of the monomers.[165,173] There are three different classes of contact dimers.[163] Type A contact dimers involve the formation of disulfide bonds at the N-terminus, commonly observed in the class C GPCR family (such as glutamate receptors).[163,174,175] Type B contact dimers are also observed in class C GPCRs (such as the $GABA_B$ receptor) and involve a coiled-coil interaction at the C-terminus.[163] The third

contact dimer, type C, is most likely to be the method of dimerization for rhodopsin-like (class A GPCRs). This contact dimer involves the formation of van der Waals interactions through the membrane exposed section of the receptor.

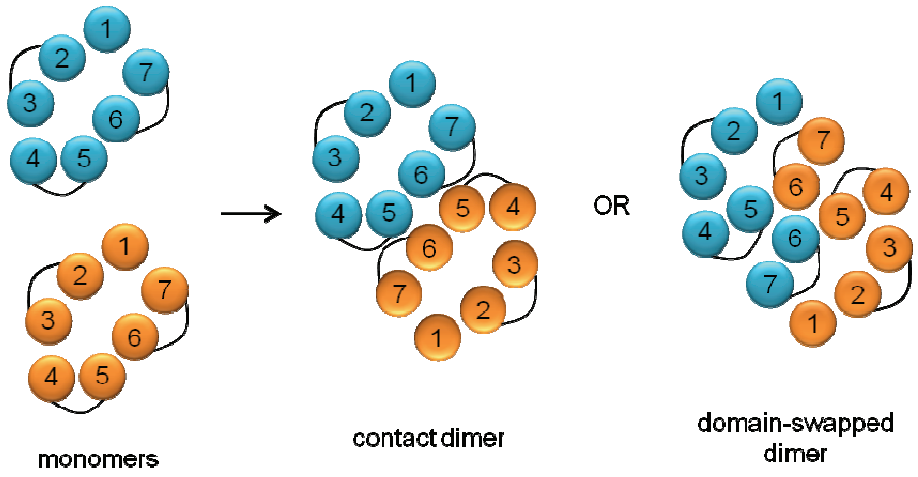


Figure 1.18 Types of GPCR dimers (image adapted from Szidonya et al.[167]).

1.6.2 3D structures of GPCR dimers

Currently, there is limited structural information for GPCR dimers. At this stage only contact dimers have been observed experimentally. The first structural data about GPCR dimers was obtained using rhodopsin and atomic force microscopy, with rhodopsin forming rows of dimers (Figure 1.19).[176] Transmission electron microscopy images demonstrated that rhodopsin also forms higher order oligomers.[169]

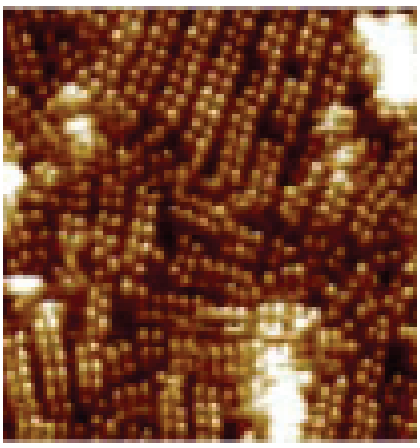


Figure 1.19 Atomic force microscopy image of rhodopsin dimers.[169,176] Image reprinted with permission from Jastrzebska et al.[169]

In many of the GPCR crystal structures (Section 1.4), the monomeric proteins are arranged in an anti-parallel orientation.[7] However, in some of the GPCR crystal structures,[10,57] dimers have been observed with the dimer interface formed by transmembrane (TM) helices 1 and 7. Interestingly, in the recent crystal structures of the chemokine receptor CXCR4 (PDB ID: 3ODU, 3OE0, 3OE6, 3OE8, 3OE9), a homodimer with a TM5 / TM6 interface was consistently observed in all structures.[17] However, different solvents and crystallization techniques can influence the formation of dimers in crystal structures,[177] and any dimerization observed may be dependent upon the monomeric proteins being investigated and the crystallization conditions used.

More robust structural evidence, such as a high resolution crystal structure of a class A GPCR dimers (to add to the CXCR4 dimeric structures), combined with other biochemical techniques is required to fully understand the mechanism of class A GPCR dimerization.[163] However, until high resolution structural data is available and based on the limited evidence available regarding dimer formation, dimerization of class A GPCRs currently focuses on the formation of type C contact dimers. For the purpose of this thesis, we have also chosen to focus on type C contact dimers for class A GPCRs.

1.6.3 Pharmacological evidence for GPCR dimerization

A number of homo- and heterodimers of GPCRs have been detected in transfected cells, with a limited number of studies in native tissue (Table 1.3). There are a number of experiments that are used to detect the formation of GPCR dimers.

These techniques include chimeric studies, such as those carried out by Maggio et al.,[163,178] have been used to elucidate dimer formation. For example, Maggio et al. developed chimeras of the α_2-adrenergic and muscarinic acetylcholine M_3 receptors, in which TM domains 1-5 from one receptor was combined with TM domains 6 and 7 from the other receptor and *vice versa*.[178] Only upon co-expression of the dimers was ligand binding and signaling reinstated.

Table 1.3 Homo- and heterodimers of aminergic GPCRs that have been identified pharmacologically. Table collated using data from the database GRIPDB.[179]

Dimer	Refs.	Dimer	Refs.
5-HT_{2C}R–5-HT_{2C}R	180	M_1 mAChR–M_1 mAChR	181
5-HT_4R–5-HT_4R	182,183	M_2 mAChR–M_3 mAChR	184
D_1R–A_1AR	185	M_3 mAChR–α_{2C}AR	178
D_1R–D_1R	186	α_{1A}AR–α_{1A}AR	187
D_1R–D_2R	188	α_{1A}AR–α_{1B}AR	187
D_2R–A_{2A}AR	189,190	α_{1B}AR–α_{1B}AR	187
D_2R–CB_1	191	α_{2A}AR–β_1AR	192
D_2R–D_2R	193-198	β_1AR–β_1AR	199,200
D_2R–D_3R	201	β_1AR–β_2AR	199,200
D_2R–SST2R	202	β_2AR–DOR	203
D_2R–SST5R	204	β_2AR–KOR	203
D_3R–D_3R	205	β_2AR–β_2AR	200,206,207
H_2R–H_2R	208		

Co-immunoprecipitation experiments are used to detect protein-protein interactions, however, due to the nature of these experiments and high levels of protein expression, artifacts can occur.[209,210] A limited number of experiments using native tissue have indicated that A_1AR,[211] D_2R[193] and $GABA_B$[212] may exist as dimers in brain tissue.[163]

Energy transfer-based techniques, such as fluorescence resonance energy transfer (FRET) or bioluminescence resonance energy transfer (BRET) have been used to identify and study a number of GPCR dimers in living cells.[209] These experiments are used to detect protein-protein interactions in live cells and relies on a transfer of energy between two labeled proteins that are in close proximity (100 Å).[167,213,214]

According to Palczewski,[177] some of the best evidence of physiologically relevant homodimers came from an in vivo study by Rivero-Müller et al.,[215] studying transgenic mice

that co-expressed binding deficient and signaling deficient mutant receptors of the luteinizing hormone receptor (a class A GPCR). In the absence of wild-type receptors, these binding and signaling deficient receptors re-established the normal luteinizing hormone function, indicating in vivo intermolecular interactions.

1.6.4 Functional consequences of GPCR dimerization – a new drug target?

Since GPCRs represent a significant drug target, their propensity to form dimers has major implications for drug discovery. GPCR dimers may possess unique properties that can be exploited in drug design, as dimers may be able to amplify signaling, as well as possessing different ligand binding and signaling properties.[165] Dimerization of GPCRs may change the pharmacology of the receptors, and could increase the functional diversity of the receptors, particularly for heterodimers.[216,217] Oligomerization can also change the type of G protein that associates to the receptor.[217] For example, the D_1R / D_2R heterodimer couples to a different G protein, compared to the monomeric proteins.[188] In addition to altered ligand binding and signaling properties, the GPCR protomers that constitute a dimer may interact, which is commonly referred to as cross-talk. Thus, the protomers may be allosterically coupled to each other and the binding of a ligand at one protomer can positively or negatively influence the binding of a ligand at the adjacent protein.[218]

Currently, selectively targeting GPCRs is challenging due to the high sequence homology of the orthosteric binding site, particularly in closely related receptor families. However, GPCR dimers could allow for the development of highly selective drugs that may have a reduced side effect profile.[172] Thus, GPCR dimers may offer a novel target with distinct pharmacological properties,[219] especially for complex disease states in which multiple receptors are implicated.[220]

1.6.5 Elucidation of class A GPCR dimerization interfaces

In addition to using the experimental techniques discussed in Section 1.6.3, such as BRET and FRET, other experiments including cysteine cross-linking have been used to elucidate dimerization interfaces for a number of GPCR dimers. Cysteine cross-linking experiments involve the mutation of specific residues to cysteine and if monomeric receptors are close to each other a disulfide bond can be formed, making a dimer. This dimeric form of the receptor can then be detected using techniques such as SDS-PAGE.[195,221]

Two dimerization interfaces have been identified for a number of receptors, namely TM1 and TM4 / TM5.[222] The pharmacological elucidation of GPCR dimerization interfaces can be exemplified by studies on the D_2R homodimer.[195] Truncated mutants[196] and cysteine cross-linking experiments[195,221] have been used to identify TM4 as a dimerization interface. However, in more recent studies, TM1 has also been implicated in D_2R oligomer formation, where the minimum size for an oligomer involved four monomeric proteins and there are two symmetric interfaces, TM1 and TM4, in the formation of higher order D_2R oligomers.[198] Additionally, D_2R dimers with a TM4 interface were found to be transient unless stabilized by cysteine cross-linking and, as a result, stable dimers may not be required for activity.[197,223]

Interestingly, in similar pharmacological studies, the TM4 interface has been implicated in dimerization interfaces for other class A GPCR homodimers including, 5-$HT_{2C}R$[180] 5-HT_4R[224] and $\alpha_{1B}AR$.[225,226] Higher order oligomers were also found for $\alpha_{1B}AR$, including a TM1 interface similar to D_2R.[225,226] Additionally, for the 5-HT_4R, cysteine residues on TM3 and TM4 have been implicated in forming the dimer interface.[224]

1.6.6 Molecular modeling of GPCR dimers

A number of models of GPCR dimers have been developed (Table 1.4), often using dimerization interfaces determined by the previously described experimental procedures.

These models are built using a number of different techniques, that vary based on the structural and experimental data available at the time of model development. Namely, dimeric GPCR models can be developed using protein-protein docking software, manually aligned using experimental data to align the two protomers or using a structural template (refer to Section 1.6.2), such as a crystal structure or atomic force microscopy images. Additionally, molecular dynamics (MD) simulations of dimeric structures can be used to study the dynamic nature of dimerization of GPCRs.

Table 1.4 Aminergic GPCR dimers for which molecular models have developed.

GPCR dimer	TM helices implicated in dimer/oligomer interface	Reference
D_2R–D_2R	TM1, TM4, TM5	195,198,221
5-$HT_{2C}R$–5-$HT_{2C}R$	TM1, TM4, TM5	180
5-HT_4R–5-HT_4R	TM4	168,227
5-$HT_{2A}R$–mGluR2	TM4, TM5	228
$\alpha_{1B}AR$–$\alpha_{1B}AR$	TM1, TM4	225,226
5-$HT_{2A}R$–5-$HT_{2A}R$	TM4, TM5	103

A number of models of the 5-HT_4R homodimer have been developed[168,224,227] using a homology model based on the rhodopsin crystal structure and the protein-protein docking software GRAMM (global range molecular matching).[229,230] Protein-protein docking studies predicted a TM2 / TM4 or a TM4 / TM6 dimerization interface[168] with the minimum distance between adjacent orthosteric binding sites determined to be 22 Å.[227]

A model of the 5-$HT_{2C}R$ homodimer was also constructed with a rhodopsin-based homology model.[180] Using a significant amount of experimental data, including extensive cysteine cross-linking experiments, two distinct interfaces were determined, TM1 and TM4 / TM5 and a model was developed for the TM4 / TM5 interface. Similarly, homology models of monomeric proteins of D_2R and δ opioid receptor (DOR), built using the crystal structure

of β_2AR (PDB ID: 2RH1), have been used to develop models of their respective homodimers, using cysteine cross-linking to identify the dimerization interface.[198,231]

Using a homology model of A_3AR, a homodimer was developed by superimposition on to the atomic force microscopy model of rhodopsin dimer (PDB ID; 1N3M[232]).[233] Based on the spacing and alignment of the TM4 / TM5 interface observed in the atomic force microscopy model of rhodopsin, a number of models of different symmetrical, contact dimer interfaces were built. Of these models, the TM4 / TM5 dimers were the most energetically favorable, followed by the TM1 / TM2 interface, which is consistent with the dimerization interfaces determined for other class A GPCRs.

Models of the 5-HT_{2A}R-mGluR2 heterodimer[102,228] and the 5-HT_{2A}R homodimer[103] were built using both protein-protein docking, as well as comparison to the theoretical model of the rhodopsin dimer and experimental data. Rosetta++[234] was used for protein-protein docking, and the complexes were visually assessed and compared to the atomic force microscopy model of rhodopsin dimer (PDB ID: 1N3M[232]). The heterodimers and monomers were all subjected to 40 ns of explicit molecular dynamics simulations in a solvated phospholipid bilayer and significant changes to the dimerization interface were observed.

Whilst a number of models of dimeric GPCRs have been developed, they are still quite limited due to the paucity of structural and biochemical data available regarding dimerization interfaces. However, increasing evidence is emerging indicating dimerization interfaces of TM 4 / TM 5 and TM1. Of particular interest are the molecular dynamics simulations carried out by Bruno et al. with the 5-HT_{2A}R-mGluR2 heterodimer[102] and the 5-HT_{2A}R homodimer,[103] where allosteric interactions were observed between the monomeric units of a dimer. Increased understanding of dimerization interfaces will assist in the development of dimeric models, however high resolution crystal structures of GPCR dimers, such as the recent CXCR4[17] will greatly enhance our ability to model these interactions.

1.7 Bivalent ligands

Portoghese et al. pioneered the development of bivalent ligands for opioid receptors as an alternative method of targeting GPCRs compared to more traditional methods.[166,235-240] Bivalent ligands were originally developed as an approach for medicinal chemists to increase the potency and selectivity of ligands.[241] For many class A GPCRs, there is a significant need for increased subtype selective compounds with reduced side effects, however this is quite challenging due to the very high sequence homology of the orthosteric site of related receptors.

With increasing evidence supporting GPCR dimerization, it has been proposed that bivalent ligands in fact target GPCR dimers, which leads to their increased potency and selectivity profiles.[241] Thus, targeting a specific dimer may result in compounds with less side effects,[242] particularly if GPCR dimers are localized in specific areas of the body.

However, the majority of bivalent ligands may be relegated to use as pharmacological tools, as they generally break most of Lipinski's guidelines for oral bioavailability,[243] and can have poor physicochemical properties. They are often high molecular weight compounds, commonly with high logP due to the introduction of long spacer units. And, depending on the types of spacer investigated, they can have a large number of rotatable bonds. However, there are strategies to improve upon these properties to some extent (see Section 1.7.4) and a number of promising bivalent ligands have been developed that display promising activity in vivo, as well as those that are useful as pharmacological probes.

Bivalent ligands are usually developed by linking two "pharmacophores" by a spacer, often incorporating a linking group between the pharmacophore and the spacer (Figure 1.20).[244] In this context, the term "pharmacophore" is used to describe an active compound or a privileged structure that is incorporated into a bivalent ligand (Figure 1.20); this terminology is commonly used in the field of bivalent ligands.[166,236,245] There are two types

of bivalent ligands that can be developed; homobivalent ligands (two identical pharmacophores) and heterobivalent ligands (two different pharmacophores).

Figure 1.20 A schematic depiction of a bivalent ligand, where a spacer links two pharmacophores, facilitated by a linking group.

More recently, bivalent ligands have been developed as pharmacological tools to investigate homodimerization and heterodimerization of GPCRs.[166,236] Bivalent ligands are designed to span the distance between two orthosteric sites on adjacent monomers, binding simultaneously at both sites. It has been postulated that two known pharmacophores linked together by a spacer of optimal length would show greater potency for the targeted receptors when compared to the activity of the individual pharmacophores (Figure 1.21).[166,246] Using these pharmacological probes, it may be possible to elucidate the approximate distance between adjacent orthosteric sites. Bhushan et al. suggested that bivalent ligands can be utilized as "molecular rulers" to assist in the elucidation of dimerization interfaces, as well as the approximate distance between orthosteric sites in a dimer.[237] In theory, the binding affinity of the bivalent ligand should be the product of the binding affinities of the two individual pharmacophores, if they bind simultaneously to adjacent orthosteric sites.[166,236] Additionally, bivalent ligands can have unique functional properties,[227] which may result from a change in selectivity for a specific G protein, and thus the regulation of a different signaling pathway.[242]

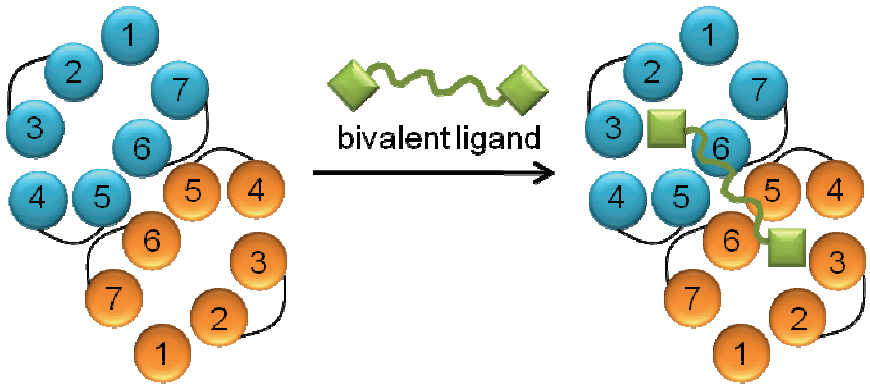


Figure 1.21 Schematic of a bivalent ligand binding to a GPCR dimer (adapted from Soulier et al.[168]).

1.7.1 Mechanism of action of bivalent ligands

The mechanism of action of bivalent ligands is not completely understood, however, the increase in potency and selectivity observed is likely to be a result of three possible binding events (Figure 1.22). The first binding hypothesis is that the local concentration of the pharmacophore is increased in the vicinity of the receptor binding site, because there are two pharmacophores covalently tethered, which increases the probability of a productive binding event (Figure 1.22b).[236] The second binding hypothesis is that following univalent binding of the first pharmacophore (Figure 1.22c), the second pharmacophore partially binds to the neighboring site or an allosteric site (neighboring accessory site), adding to the selectivity of the bivalent ligand (Figure 1.22d).[236] The third, and most commonly favored binding hypothesis, is that the bivalent ligand is simultaneously binding to a GPCR dimer at adjacent orthosteric sites (Figure 1.22e). This binding event is thought to be a two stage process where one pharmacophore of the bivalent ligand binds univalently to the receptor dimer (Figure 1.22c), thus lowering the entropy of the system and making binding of the second pharmacophore with the adjacent protein of the dimer favored over another molecule (Figure 1.22e).[236,247] Additionally, two bivalent ligands can bind at adjacent orthosteric sites (Figure 1.22f), however, this is thought to be less favorable entropically than one bivalent ligand binding to a GPCR dimer (Figure 1.22e). If two pharmacophores of a bivalent ligand

bind simultaneously to a GPCR dimer (Figure 1.22e), they could be used as pharmacological probes to elucidate the approximate distance between adjacent orthosteric sites in a dimer.[166,236]

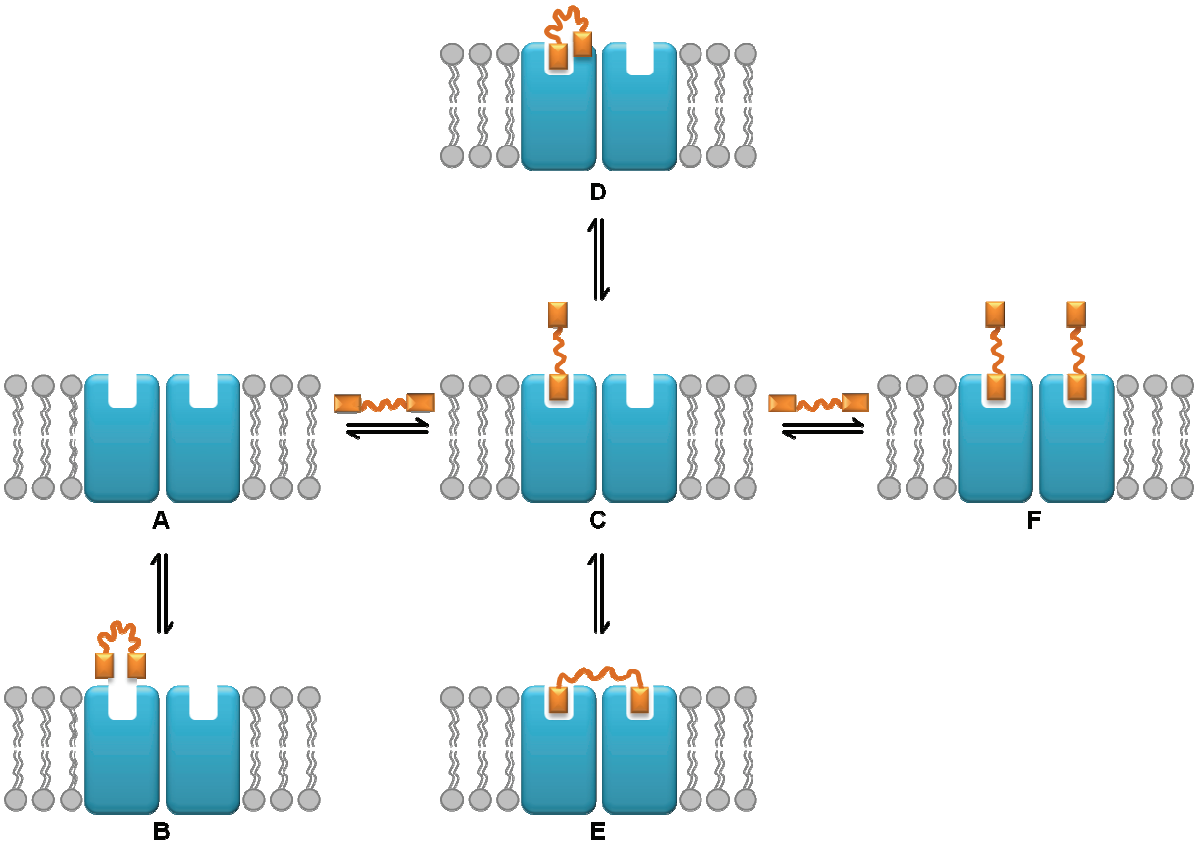


Figure 1.22 Binding of a bivalent ligand to a GPCR dimer (a), (b) local concentration of the pharmacophore is increased in the vicinity of the receptor binding site (monomer or dimer), (c) shows univalent binding of the compound, (d) the bivalent ligand can interact with an adjacent allosteric site (binding to a monomer or a dimer), (e) the second pharmacophore of the bivalent ligand binds to the orthosteric site of the adjacent monomer simultaneously, or (f) two bivalent ligands bind simultaneously to a GPCR dimer (adapted from Portoghese et al.[245]).

1.7.2 Pharmacological testing of bivalent ligands

When developing bivalent ligands, commonly a number of monovalent or capped spacers are synthesized to use as controls in pharmacological studies. This ensures that the impact of the linker of the monovalent ligand is evaluated to some extent. Alternatively, a bivalent ligand containing one active pharmacophore and one inactive or "dummy" pharmacophore

that is inactive at the dimer under investigation can be used as a control in pharmacological evaluation.[248]

Often the first type of pharmacological evaluation of bivalent ligands (and their monomeric counterparts) is a radioligand binding assay; that is, assessing the ability of the bivalent ligand to displace a radiolabeled ligand. As this type of experiment gives information about the affinity of the compounds, additional in vitro experiments, such as functional assays, are carried out. Functional assays give a measure of downstream effects of bivalent ligand binding. Ideally, a bivalent ligand should have activity that is the product of the binding affinities of the two individual pharmacophores, provided it binds simultaneously to two adjacent orthosteric sites (the third bivalent ligand binding hypothesis).[166,236] Other in vitro assays used to assess bivalent ligands include fluorescence resonance energy transfer (FRET) or bioluminescence resonance energy transfer (BRET). These experiments are used to detect protein-protein interactions in live cells and rely on a transfer of energy between two labelled proteins that are in close proximity (100 Å).[167,213]

1.7.2.1 Blood-brain barrier permeability of CNS bivalent ligands

For CNS targeting bivalent ligands, their propensity to cross the blood-brain barrier (BBB), and thus exert any desired activity, is often brought into question.[244] Consequently, the majority of bivalent ligands could be relegated to use as pharmacological tools. However, some CNS targeting bivalent ligands have been evaluated in vivo and have displayed BBB permeability.

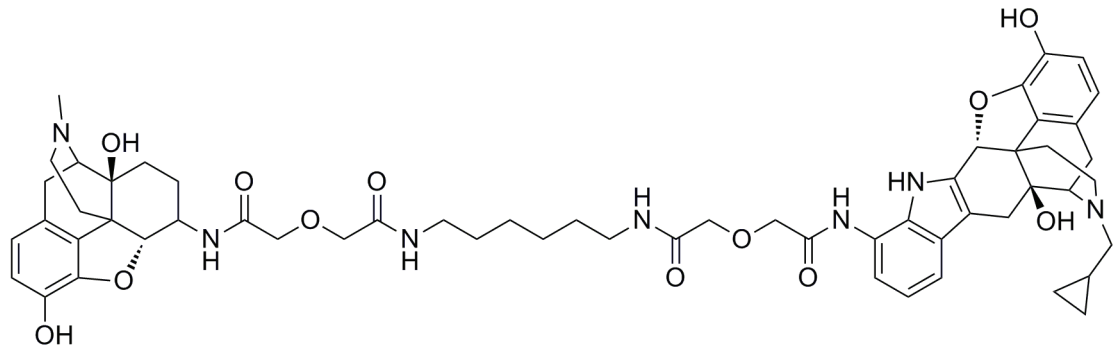


Figure 1.23 Chemical structure of MDAN-21 (**20**).[239]

In elegant pharmacological assays, MDAN-21 (**20**, Figure 1.23, a MOR-DOR heterobivalent ligand), was evaluated in animal models to evaluate BBB permeability.[239] MDAN-21 was administered intravenously (i.v.) to mice and found to be 50 times more potent than morphine.[239] MDAN-21 was also administered by the intracerebroventricular route (i.c.v., a method that bypasses the blood brain barrier), therefore removing influence of metabolic processes of the blood brain barrier. The potency ratio of i.v. compared to i.c.v. of morphine and MDAN-21 was the same, which indicates that morphine and MDAN-21 were able to cross the blood brain barrier to a similar degree.[239] This is a particularly promising study for the development of bivalent ligands, as it has been demonstrated that an optimally designed bivalent ligand may have in vivo activity and show BBB permeability.

1.7.3 Examples of bivalent ligands

A number of studies developing bivalent ligands have been reported in the literature, starting with the work of Portoghese et al. focusing on bivalent ligands acting at opioid receptor subtypes (δ, κ and μ opioid receptors, referred to as DOR, KOR and MOR, respectively).[166] A number of bivalent ligands have been synthesized and pharmacologically evaluated that target a number of different GPCR homodimers and heterodimers. Table 1.5 contains examples of bivalent ligands that have been developed to target GPCRs, including the receptor dimer under investigation, the disease state, as well as the optimal spacer length (if determined in the study).

Table 1.5 Bivalent ligand that have been developed, including the dimers at which they act, the ideal spacer length determined and the disease state they aim to treat.

Dimer	Pharmacophores	Spacer lengths (atoms)	Ideal spacer length (atoms)	Disease state	Ref.
5-HT_{1A}R-5-HT_7R	arylpiperazine[a]	2-12	N/C[b]	sleep disorders, depression	249
5-$HT_{1B/1D}$R	5-$HT_{1B/1D}$R agonist	6	6	migraine	250
5-$HT_{1B/D}$R	serotonin[c]	2-24	N/C	migraine	251
5-HT_4R	5-HT_4R agonist	6-29	N/C	gastrointestinal disorders	168, 227
A_{2A}AR-D_2R	A_{2A}AR antagonist D_2R agonist	26-118	N/C	Parkinson's disease	252
β_1AR	β_1AR antagonist	2-14	N/C	myocardial imaging agent	253
CB_1R	CB_1R antagonist	5-23	15	pain, inflammation	254
CCK_2R-MOR	CCK_2 antagonist MOR agonist	9-22	16-22	pain	255
CXCR4	CXCR4 antagonist	2-8 nm[d]	5.5-6.5 nm	cancer	256
D_2R	1,4-DAP[a,e]	18-25	22	schizophrenia	248
D_2R	D_2R antagonist	3-7	6	schizophrenia	257
GnRHR[f]	GnRHR antagonist	6-26	N/C	reproductive functions	258
KOR-DOR	KOR antagonist DOR antagonist	15-23	21	pain	237
KOR-DOR	KOR agonist DOR antagonist	12-20	18	pain	238
KOR-MOR	KOR antagonist MOR antagonist	20-23	21	pain	240
KOR-MOR	KOR agonist MOR antagonist	3-17	N/C	Pain	259, 260

Table 1.5 continued

Dimer	Pharmacophores	Spacer lengths (atoms)	Ideal spacer length (atoms)	Disease state	Ref.
KOR-MOR	KOR agonist MOR agonist/antagonist	6-12	N/C	Pain	261
KOR-MOR or MOR-MOR	KOR/MOR agonist or MOR antagonist, MOR antagonist	10	10	pain	262
LHR[g]	LHR agonist	11-28	N/C	pro-fertility	263
MOR-DOR	MOR agonist DOR antagonist	16-21	21	pain	239
Muscarinic receptors	Muscarinic agonist	11-12	11-12	Alzheimer's disease, schizophrenia	264
Muscarinic receptors	Aryl group[a]	21-30	26	cardiovascular disease	265
Opiate receptors	opioid antagonist	8 & 17	N/C	pain	266
Opiate receptors	opioid agonist	10-28	10	pain	245
Opiate receptors	opioid antagonist	10-28	16	pain	245
Y_1R[h]	Y_1R antagonist	8-36	N/C	-	267

[a]Privileged structure; [b]N/C = not conclusive; [c]endogenous agonist; [d]molecular rulers, spacer length measured in nm not atoms; [e],4-disubstituted aromatic piperidines/piperazines (1,4-DAPs); [f]gonadotropin-releasing hormone receptor (GnRHR); [g]luteinizing hormone receptor (LHR); [h]neuropeptide Y Y_1 receptor (Y_1R).

1.7.4 Design strategies for bivalent ligands

Bivalent ligands are initially designed by selecting pharmacophore(s) that are potent at the receptor(s) of interest.[241] Additionally, the pharmacophore(s) must have synthetically feasible attachment points for linking the spacer. Ideally, there should also be a significant amount of structure-activity relationship data for the pharmacophore(s), to assist in the selection of appropriate attachment points for the spacer.

Initial studies should be aimed at ascertaining an appropriate spacer length and attachment point. Following this, the spacer can be finely tuned including the exploration of different spacer types to include more hydrophilic and more conformationally restricted spacers, and ideally the spacer should not interfere with the binding of the pharmacophore to the receptor.[244] These considerations, as well as examples from the literature are discussed below.

1.7.4.1 Spacer attachment point

To attach the spacer to the pharmacophore often a linking group is introduced or the attachment point is chemically modified. These adjustments will ultimately be determined by the type of spacer selected and the pharmacophore-spacer linkage formed. Such linking options include amide, ether, ester bonds, as well as the formation of triazole linkers and direct alkylation of the pharmacophore.[241] Consideration must also be given when selecting the appropriate method of attachment for the bivalent ligand as certain linkages, such as esters, can be metabolized in the body. For example, in the butorphan bivalent ligands (e.g. **21**, Figure 1.24) developed by Decker et al. using ester linkages, methyl groups were introduced in the α-position to reduce the rate of ester hydrolysis.[259] However, whilst the hydrolytic stability increased, reduced activity of the bivalent ligand was observed.[259]

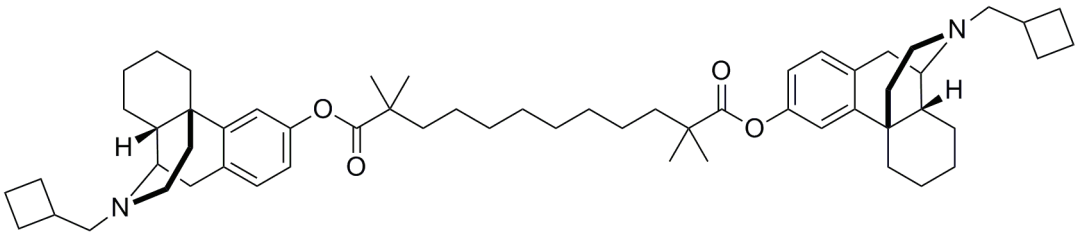


Figure 1.24 Example of butorphan bivalent ligand (**21**) containing α-methyl groups to reduce rate of ester hydrolysis.[259]

1.7.4.2 Optimizing spacer length

Next, the spacer length needs to be established. If the spacer length is too long it reduces the likelihood of both pharmacophores being attached to the dimer, and if it is too short the pharmacophores would not be able to interact with both monomers simultaneously.[166] The

initial approximation of spacer length has been demonstrated for the opioid receptors, where initial spacer lengths were determined using polyethylene glycol or polyglycine spacers.[235,245,266]

The fine tuning of spacer length can be exemplified by later studies on opioid bivalent ligands, such as the bivalent ligand KDAN-18 (**22**, Figure 1.25).[238] Bivalent ligands were developed, varying in length from 12 to 20 atoms, with spacers consisting of glycyl residues and methylenes. Incrementally changing the number of methylenes in the spacer, allowed for fine tuning of the spacer length.[238] From the extensive research of Portoghese et al., optimal spacer lengths for opioid receptors were determined to be in the range of 18 to 21 atoms.[237-240]

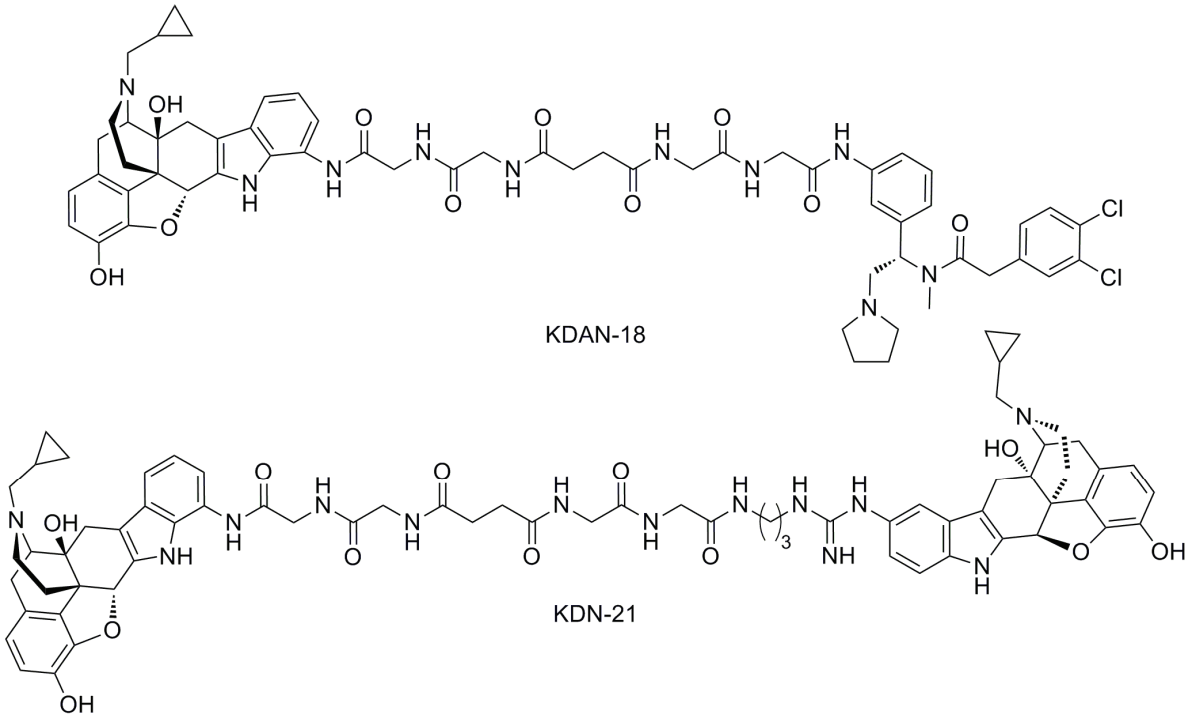


Figure 1.25 Structures of KDAN-18 (**22**)[238] and KDN-21 (**23**).[237]

However, the ideal spacer length determined was not only dependent on the dimer under investigation, but also the mechanism of action of the pharmacophore (i.e. antagonist vs. agonist). For example, the optimal spacer length determined for the heterobivalent ligands targeting the κ-δ heterodimer was determined to be 21 atoms (the KDN series, **23**), however,

when the κ-antagonist was changed to a κ-agonist (the KDAN series), the optimal spacer length was determined to be 18 atoms (**22**).[237,238]

1.7.4.3 Optimizing spacer type

Ideally, the spacer of a bivalent ligand can be used not only to optimize the distance between the two pharmacophores but also to tune the physicochemical properties. Initially, spacers should be easily incorporated into the bivalent ligand and relatively inert once introduced into the bivalent ligand. Preferably, it should be relatively easy to alter the spacer length so that a number of bivalent ligands of varying length can be synthesized to initially determine an approximate spacer length and attachment point. This is a critical step in bivalent ligand design, as the optimal spacer length varies and must be elucidated empirically for each bivalent ligand.

Following the determination of an approximate spacer length, different spacer types can be investigated. A number of properties need to be considered when selecting the appropriate spacer for the development of bivalent ligands, such as hydrophilicity and flexibility. As a result, many spacers have been explored in the literature, ranging from simple polymethylene chains and polyamide chains to more conformationally constrained aromatic-containing spacers as demonstrated in Table 1.6. Additionally, calculated logP (ClogP) values give an indication of the aqueous solubility of the spacer, and can be used to assist in the selection of an appropriate spacer.

Table 1.6 Examples of spacers used in the development of bivalent ligands (P1 = pharmacophore 1, P2 = pharmacophore 2) and ClogP values calculated with ChemDraw Ultra.[268]

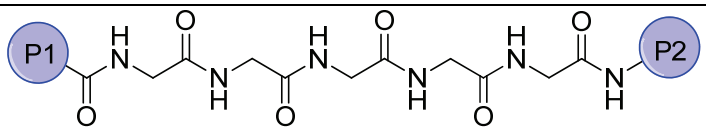
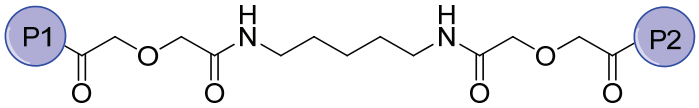
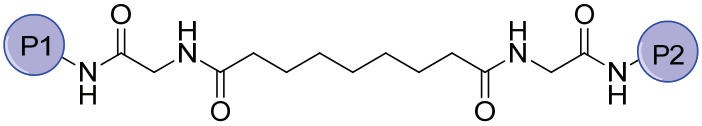
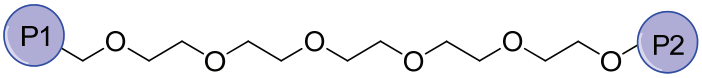
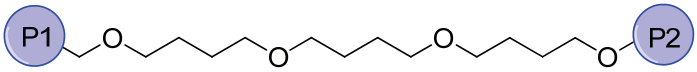
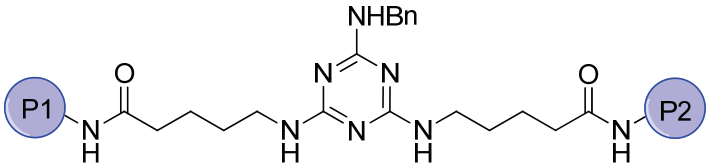
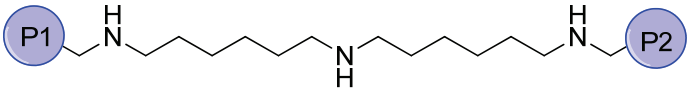
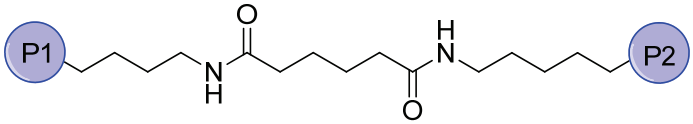
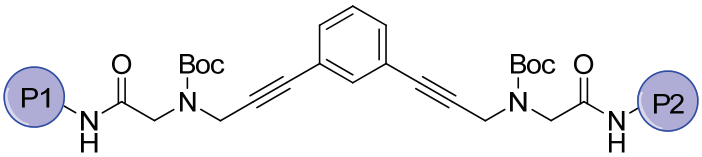
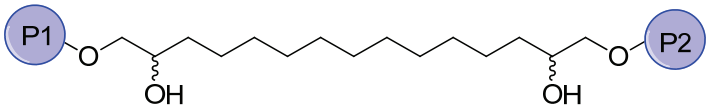
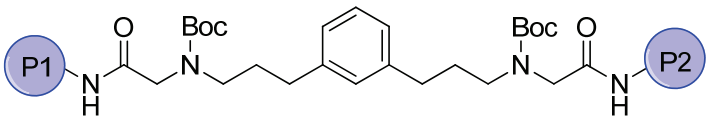
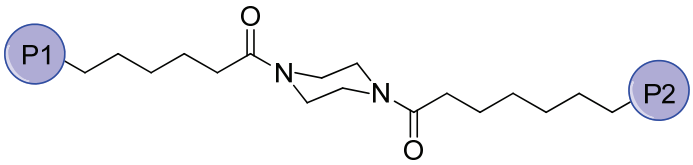
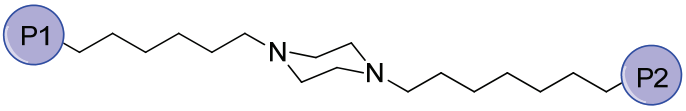
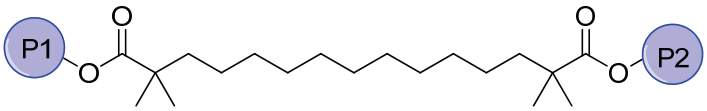
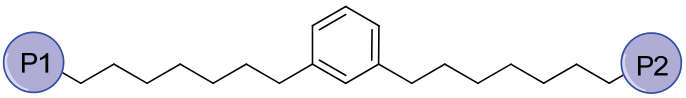
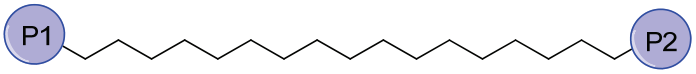
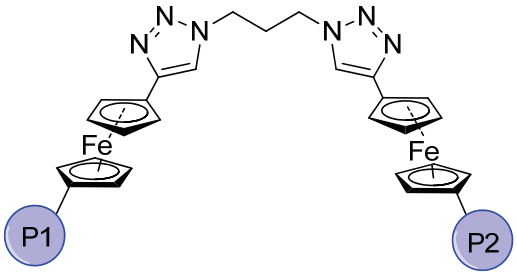
Spacer type	ClogP	Ref.
	-3.91	237,238, 240
	-1.71	239
	-1.34	236
	-0.20	258,263, 269-275
	1.01	275
	1.56	168
	2.35	254,265
	2.97	168
	3.02	227
	3.47	260
	4.54	227
	4.59	251

Table 1.6 continued

Spacer type	ClogP	Ref.
	6.12	251
	6.91	259
	10.55	168
	10.75	249,257, 269,270, 272,276, 277
	n/a	278

Increasing hydrophilicity of the bivalent ligand. The lipophilicity of the polymethylene spacers limits the aqueous solubility of these bivalent ligands, particularly with increasing spacer length. To combat this problem, Portoghese et al. balanced the hydrophobicity of the spacer by the incorporation of polyethylene glycol or polyglycine residues into the spacer (Table 1.6).[245,266] Additionally, in a series of xanomeline bivalent ligands targeting the muscarinic receptors, changing a polymethylene spacer to a polyethylene glycol spacer not only improved aqueous solubility, but also the affinity and potency of the compounds.[269]

However, polyethylene glycol spacers can be problematic as the form coils in aqueous solution.[275,279] As these coils may shorten the length of the spacer, LaFrate et al. investigated the use of polypropylene and polybutylene glycols, which are less likely to form these coils but still have the advantage of increasing the influence on solubility of the spacer, relative to

a polymethylene chain.[275] 1,3,5-Triazines, such as cyanuric chloride, as well as piperazines, can also be used to introduce a central hydrophilic core in the spacer to increase the hydrophilicity of the spacer (Table 1.6).[168]

Introducing conformationally restricted spacers. Conformationally restricted spacers can also be investigated as they may confer affinity and selectivity.[280] More conformationally restricted spacers can include oligoglycine spacers, such as those employed in the development of opioid bivalent ligands (Figure 1.25),[245] or less flexible spacers such as the aromatic-containing spacers used in the development of 5-HT_4 targeting bivalent ligands (Figure 1.26).[227] Particularly, using the more conformationally restricted spacers may result in an "all or nothing" binding event, as described by Bobrovnik[280] in an analysis of flexible and rigid spacers in bivalent ligands. Ideally, a more rigid spacer should increase the affinity of the bivalent ligand, but it could result in a lack of activity, as experienced by Berque-Bestel and co-workers when developing bivalent ligands for the 5-HT_4 receptor.[241,281] They found that if the spacer was too rigid, it could prevent the second pharmacophore from binding and that flexible spacers were "essential" for correct positioning of pharmacophores.[281]

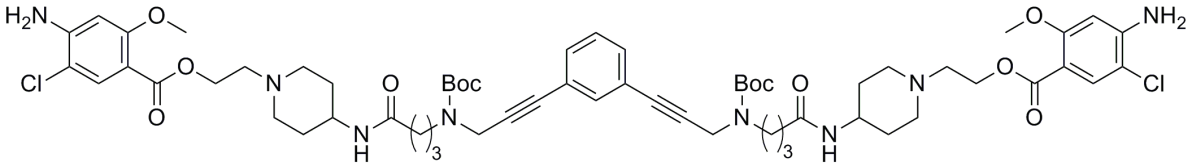


Figure 1.26 Structure of a bivalent ligand designed for the 5-HT_4R with a conformationally restricted spacer (**24**).

However, there are examples of bivalent ligands where conformationally rigid spacers resulted in enhanced activity. In the development of bivalent ligands for the muscarinic receptor, Christopoulos et al. employed a meta- or para-substituted phenylenedipropargyl spacer.[264] Interestingly, the para-substituted bivalent ligand (**25**, Figure 1.27) showed significantly increased affinity and selectivity for M_1 mAChR and M_2 mAChR subtypes, where as meta-substituted bivalent ligand (**26**) displayed functional selectivity towards the

M_1 mAChR and M_4 mAChR subtypes.[264] Similar to investigating spacer length, the conformational rigidity needs to be explored empirically bivalent ligand.

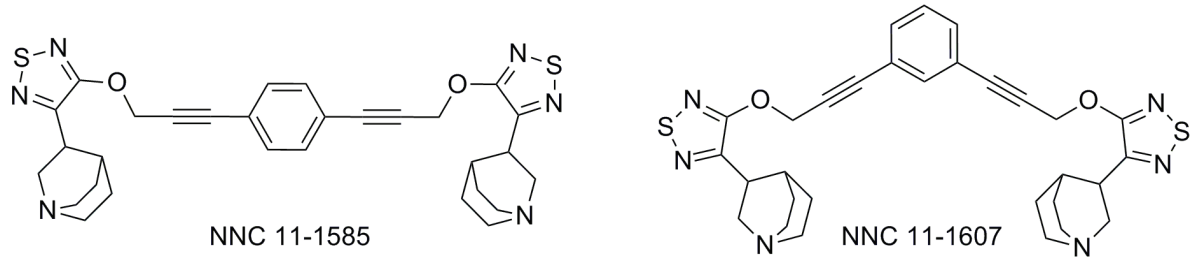


Figure 1.27 Structures of NNC 11-1585 (**25**) and NNC 11-1607 (**26**).

1.7.4.4 Bivalent ligands targeting dopamine receptor subtypes

Of significant interest to our research group has been the development of bivalent ligands that bind to D_2R for use as pharmacological tools. There have been three unique series of D_2R bivalent ligands have been developed in the past ten years (Figure 1.28).

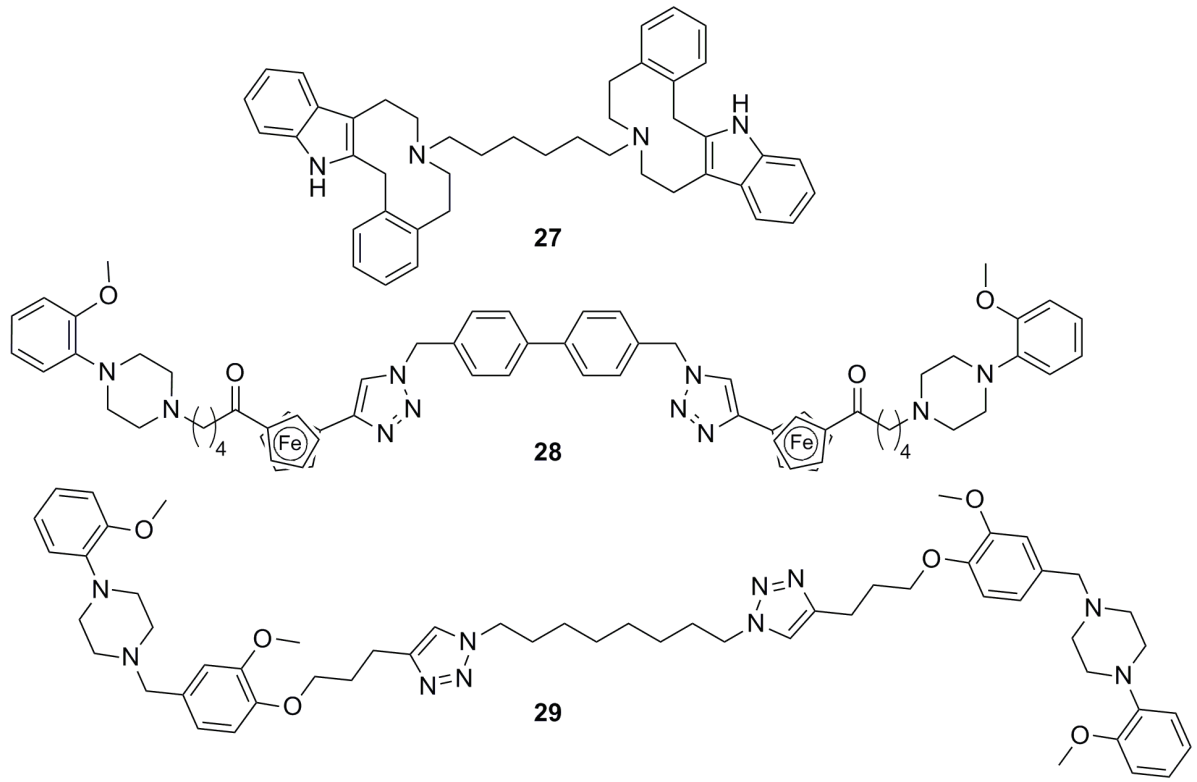


Figure 1.28 Example chemical structures of D_2R bivalent ligands; azecine bivalent ligand (**27**),[257] 1,1′-disubstituted ferrocenes (**28**)[278] and 1,4-disubstituted piperazines / piperidines bivalent ligand (**29**).[248]

Abadi et al. developed a series of bivalent azecine derivatives (**27**, Figure 1.28), with the six carbon polymethylene spacer displaying the best, if moderate, activity.[257] In more recent

studies, Gmeiner et al. developed two series of bivalent ligands to investigate the D_2R homodimers. A series of 1,1′-disubstituted ferrocenes (**28**) were developed,[278] where the ferrocene can act as a "molecular hinge" allowing for enhanced conformational flexibility of the spacer. The monovalent and bivalent ligands containing the butylene spacer with benzyltriazoyl spacer showed best affinity for the D_2-family of receptors. A series of 1,4-disubstituted piperazines / piperidines bivalent ligands (**29**) were also developed using click chemistry.[248] For this series, a spacer length of 22 atoms displayed the most promising activity and the more hydrophobic spacers exhibited binding properties indicative of a bivalent ligand targeting a dimer.

These bivalent ligands could be used as pharmacological probes to elucidate the formation and behavior of D_2R homodimers. Additionally, D_2R homodimers have recently been implicated in the pathophysiology of schizophrenia and may present an alternative method to treat this disease state.[282]

1.8 Summary

Class A GPCRs have been studied extensively over the last 60 years and attracted significant interest from pharmaceutical companies, yet high resolution 3D structural data of these proteins remained elusive until the last ten years. However, for the majority of this time the crystal structure of rhodopsin was the only high resolution template for GPCRs. Due to the ever increasing number of high resolution X-ray crystal structures of GPCRs since 2007, structure-based drug design will begin to play an even bigger role in the drug discovery process. It is currently an exciting time to work in the field of structure-based drug design for GPCRs, with approximately 30 non-rhodopsin crystal structures determined in the last four years. These additional structures also give molecular modelers the chance to evaluate and refine GPCR modeling techniques.

The role of GPCR dimers is becoming increasingly apparent, which gives us increasing opportunities to develop more selective drugs, with the potential to have fewer side effects.

Bivalent ligands may not be of significant interest to Big Pharma, due to their relatively poor physicochemical properties in relation to Lipinski's guidelines for oral bioavailability, however, as pharmacological tools, they can give a greater understanding of GPCR dimerization and oligomerization. Additionally, using the recent crystal structures, or models based on them, GPCR dimer models can be developed and used to investigate the behavior of these systems, using molecular dynamics simulations.

1.9 Thesis aims

This work explores two key hypotheses. The first being that homology models of GPCRs built using non-rhodopsin crystal structures would provide useful information regarding the function and structure of GPCRs, as well as providing insight into drug action for the development of novel pharmaceutical compounds. The second key hypothesis is that covalently tethering two pharmacophores by a spacer of a given composition and length may improve the potency and activity of the original pharmacophore and that any affinity gains observed could be a result of the bivalent ligand interacting with a GPCR dimer. To address these issues, this research had two main aims:

1. **To develop homology models of class A GPCR.** Specifically, to develop homology models of pharmaceutically relevant GPCRs, based on the β_2AR template and evaluate their potential for use in structure-based drug design using small scale virtual screening. Additionally, to use homology models of related GPCRs to increase our understanding of subtype selectivity. This also included developing a suitable binding site optimization protocol and procedure for method selection, as well as adding these new, optimized structures to the limited number of non-rhodopsin models that were freely available. These aims have been addressed in Chapters 2 and 3. Furthermore, these modeling techniques were applied to the development of models of the D_2R homodimer, which were used to evaluate the dynamics of D_2R homodimerization (Chapter 4).

2. **To develop a series of bivalent ligands for the dopamine D_2 receptor.** Expressly, to design, synthesize, structurally characterize and pharmacologically evaluate a series of clozapine homobivalent ligands (Chapter 4).

Clozapine (**30**, Figure 1.29) was selected as the pharmacophore for the development of bivalent ligands, as dopamine receptors are heavily implicated in the disease state of schizophrenia[283-285] and are principally targeted in the development of antipsychotics.[22,285] However, clozapine only displays moderate (sub-micromolar) affinity for the D_2R. Clozapine considered to be the leading atypical antipsychotic agent used for the treatment of refractory schizophrenia.[286] However, its clinical use is limited due to severe adverse side effects,[27,287-289] including a potentially fatal blood disorder, agranulocytosis.[290-292] This drug-induced dyscrasia is thought to result from the formation of a reactive nitrenium ion intermediate at the N5 position of clozapine.[291]

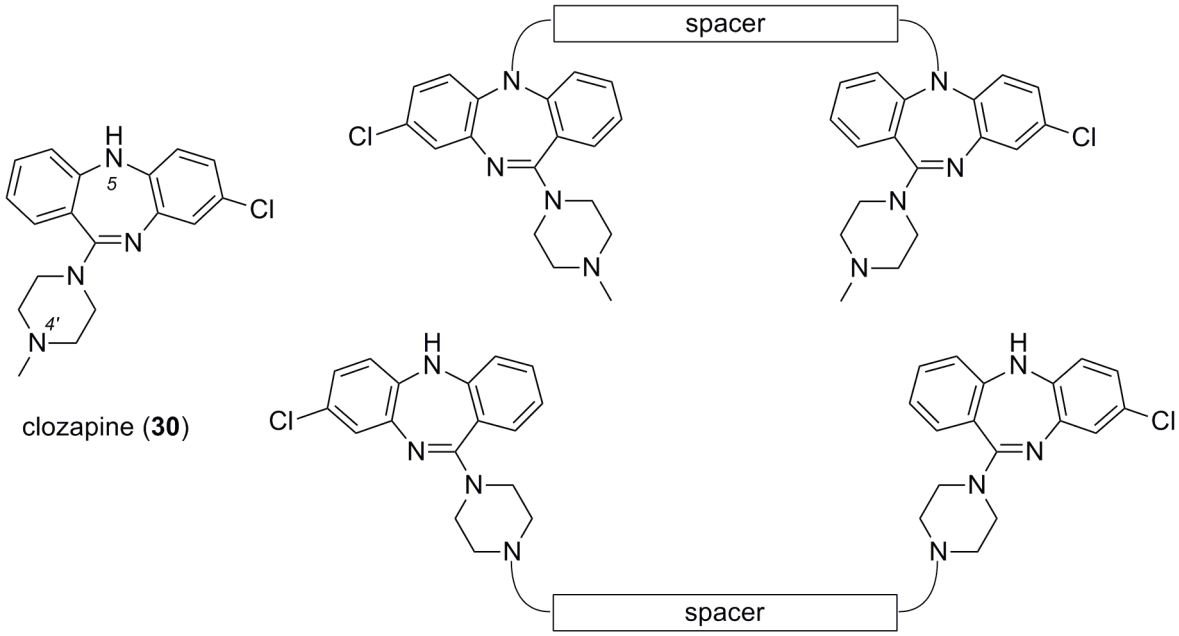


Figure 1.29 Structure of clozapine (**30**) and structures of the proposed clozapine bivalent ligands.

A wide range of clozapine analogues have been developed, including some from our own research group,[293-297] and have demonstrated that modifications at the N4′ position[293-295] and the N5 position[298,299] (amongst others) are tolerated. Both the N5 and

the distal piperazine nitrogen (N4′) positions of clozapine were synthetically attractive points for the attachment of spacers for the development of homobivalent ligands. Additionally, by developing homobivalent ligands with the spacer attached through the N5 position, it may be possible that the agranulocytosis side effect could be reduced or abolished.[291,300,301] The N4′ distal piperazine nitrogen (the ionizable nitrogen that interacts with the key aspartate residue on helix 3 (Asp $114^{3.32}$) at the entrance of the orthosteric binding site),[144] was also envisaged as an appropriate attachment point for the development of homobivalent ligands as the linker would be directed into the extracellular space to increase the probability of dual (cooperative) binding. Monovalent ligands were also synthesized for comparison in pharmacological assays.

In addition to the synthesis of homobivalent and monovalent ligands of clozapine, the aims of this project also included establishing an appropriate spacer attachment point, spacer length and linking strategy. Moreover, different spacer types were to be explored, with the aim of developing bivalent ligands with improved physiochemical properties. Ultimately, the homobivalent ligands were evaluated pharmacologically in functional and binding assays.

References

1. Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L. How many drug targets are there? *Nat. Rev. Drug Discovery* **2006,** *5*, 993-996.

2. Millar, R. P.; Newton, C. L. The year in G protein-coupled receptor research. *Mol. Endocrinol.* **2010,** *24*, 261-274.

3. Filmore, D. It's a GPCR world. *Mod. Drug Discovery* **2004,** *7*, 24-28.

4. Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J.; Sutton, G. G.; Smith, H. O.; Yandell, M.; Evans, C. A.; Holt, R. A.; Gocayne, J. D.; Amanatides, P.; Ballew, R. M.; Huson, D. H.; Wortman, J. R.; Zhang, Q.; Kodira, C. D.; Zheng, X. H.; Chen, L.; Skupski, M.; Subramanian, G.; Thomas, P. D.; Zhang, J.; Gabor Miklos, G. L.; Nelson, C.; Broder, S.; Clark, A. G.; Nadeau, J.; McKusick, V. A.; Zinder, N.; Levine, A. J.; Roberts, R. J.; Simon, M.; Slayman, C.; Hunkapiller, M.; Bolanos, R.; Delcher, A.; Dew, I.; Fasulo, D.; Flanigan, M.; Florea, L.; Halpern, A.; Hannenhalli, S.; Kravitz, S.; Levy, S.; Mobarry, C.; Reinert, K.; Remington, K.; Abu-Threideh, J.; Beasley, E.; Biddick, K.; Bonazzi, V.; Brandon, R.; Cargill, M.; Chandramouliswaran, I.; Charlab, R.; Chaturvedi, K.; Deng, Z.; Francesco, V. D.; Dunn, P.; Eilbeck, K.; Evangelista, C.; Gabrielian, A. E.; Gan, W.; Ge, W.; Gong, F.; Gu, Z.; Guan, P.; Heiman, T. J.; Higgins, M. E.; Ji, R.-R.; Ke, Z.; Ketchum, K. A.; Lai, Z.; Lei, Y.; Li, Z.; Li, J.; Liang, Y.; Lin, X.; Lu, F.; Merkulov, G. V.; Milshina, N.; Moore, H. M.; Naik, A. K.; Narayan, V. A.; Neelam, B.; Nusskern, D.; Rusch, D. B.; Salzberg, S.; Shao, W.; Shue, B.; Sun, J.; Wang, Z. Y.; Wang, A.; Wang, X.; Wang, J.; Wei, M.-H.; Wides, R.; Xiao, C.; Yan, C.; Yao, A.; Ye, J.; Zhan, M.; Zhang, W.; Zhang, H.; Zhao, Q.; Zheng, L.; Zhong, F.; Zhong, W.; Zhu, S. C.; Zhao, S.; Gilbert, D.; Baumhueter, S.; Spier, G.; Carter, C.; Cravchik, A.; Woodage, T.; Ali, F.; An, H.; Awe, A.; Baldwin, D.; Baden, H.; Barnstead, M.; Barrow, I.; Beeson, K.; Busam, D.; Carver, A.; Center, A.; Cheng, M. L.; Curry, L.; Danaher, S.; Davenport, L.; Desilets, R.; Dietz, S.; Dodson, K.; Doup, L.; Ferriera, S.; Garg, N.; Gluecksmann, A.; Hart, B.; Haynes, J.; Haynes, C.; Heiner, C.; Hladun, S.; Hostin, D.; Houck, J.; Howland, T.; Ibegwam, C.; Johnson, J.; Kalush, F.; Kline, L.; Koduru, S.; Love, A.; Mann, F.; May, D.; McCawley, S.; McIntosh, T.; McMullen, I.; Moy, M.; Moy, L.; Murphy, B.; Nelson, K.; Pfannkoch, C.; Pratts, E.; Puri, V.; Qureshi, H.; Reardon, M.; Rodriguez, R.; Rogers, Y.-H.; Romblad, D.; Ruhfel, B.; Scott, R.; Sitter, C.; Smallwood, M.; Stewart, E.; Strong, R.; Suh, E.; Thomas, R.; Tint, N. N.; Tse, S.; Vech, C.; Wang, G.; Wetter, J.; Williams, S.; Williams, M.; Windsor, S.; Winn-Deen, E.; Wolfe, K.; Zaveri, J.; Zaveri, K.; Abril, J. F.; Guigó, R.; Campbell, M. J.; Sjolander, K. V.; Karlak, B.; Kejariwal, A.; Mi, H.; Lazareva, B.; Hatton, T.; Narechania, A.; Diemer, K.; Muruganujan, A.; Guo, N.; Sato, S.; Bafna, V.; Istrail, S.; Lippert, R.; Schwartz, R.; Walenz, B.; Yooseph, S.; Allen, D.; Basu, A.; Baxendale, J.; Blick, L.; Caminha, M.; Carnes-Stine, J.; Caulk, P.; Chiang, Y.-H.; Coyne, M.; Dahlke, C.; Mays, A. D.; Dombroski, M.; Donnelly, M.; Ely, D.; Esparham, S.; Fosler, C.; Gire, H.; Glanowski, S.; Glasser, K.; Glodek, A.; Gorokhov, M.; Graham, K.; Gropman, B.; Harris, M.; Heil, J.; Henderson, S.; Hoover, J.; Jennings, D.; Jordan, C.; Jordan, J.; Kasha, J.; Kagan, L.; Kraft, C.;

Levitsky, A.; Lewis, M.; Liu, X.; Lopez, J.; Ma, D.; Majoros, W.; McDaniel, J.; Murphy, S.; Newman, M.; Nguyen, T.; Nguyen, N.; Nodell, M.; Pan, S.; Peck, J.; Peterson, M.; Rowe, W.; Sanders, R.; Scott, J.; Simpson, M.; Smith, T.; Sprague, A.; Stockwell, T.; Turner, R.; Venter, E.; Wang, M.; Wen, M.; Wu, D.; Wu, M.; Xia, A.; Zandieh, A.; Zhu, X. The sequence of the human genome. *Science* **2001,** *291*, 1304-1351.

5. Fredriksson, R.; Lagerström, M. C.; Lundin, L.-G.; Schiöth, H. B. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* **2003,** *63*, 1256-1272.

6. Hopkins, A. L.; Groom, C. R. The druggable genome. *Nat. Rev. Drug Discovery* **2002,** *1*, 727-730.

7. Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Trong, I. L.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **2000,** *289*, 739-745.

8. Rosenbaum, D. M.; Cherezov, V.; Hanson, M. A.; Rasmussen, S. G. F.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Yao, X.-J.; Weis, W. I.; Stevens, R. C.; Kobilka, B. K. GPCR engineering yields high-resolution structural insights into β_2-adrenergic receptor function. *Science* **2007,** *318*, 1266-1273.

9. Rasmussen, S. G. F.; Choi, H.-J.; Rosenbaum, D. M.; Kobilka, T. S.; Thian, F. S.; Edwards, P. C.; Burghammer, M.; Ratnala, V. R. P.; Sanishvili, R.; Fischetti, R. F.; Schertler, G. F. X.; Weis, W. I.; Kobilka, B. K. Crystal structure of the human β2 adrenergic G-protein-coupled receptor. *Nature* **2007,** *450*, 383-387.

10. Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G. F.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.; Stevens, R. C. High-resolution crystal structure of an engineered human β_2-adrenergic G protein coupled receptor. *Science* **2007,** *318*, 1258-1265.

11. Murakami, M.; Kouyama, T. Crystal structure of squid rhodopsin. *Nature* **2008,** *453*, 363-367.

12. Hanson, M. A.; Cherezov, V.; Griffith, M. T.; Roth, C. B.; Jaakola, V.-P.; Chien, E. Y. T.; Velasquez, J.; Kuhn, P.; Stevens, R. C. A specific cholesterol binding site Is established by the 2.8 Å structure of the human β_2-adrenergic receptor. *Structure* **2008,** *16*, 897-905.

13. Park, J. H.; Scheerer, P.; Hofmann, K. P.; Choe, H.-W.; Ernst, O. P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **2008,** *454*, 183-187.

14. Scheerer, P.; Park, J. H.; Hildebrand, P. W.; Kim, Y. J.; Krausz, N.; Choe, H.-W.; Hofmann, K. P.; Ernst, O. P. Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **2008,** *455*, 497-502.

15. Jaakola, V.-P.; Griffith, M. T.; Hanson, M. A.; Cherezov, V.; Chien, E. Y. T.; Lane, J. R.; Ijzerman, A. P.; Stevens, R. C. The 2.6 angstrom crystal structure of a human A_{2A} adenosine receptor bound to an antagonist. *Science* **2008,** *322*, 1211-1217.

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

17. Wu, B.; Chien, E. Y. T.; Mol, C. D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.; Wells, P.; Bi, F. C.; Hamel, D. J.; Kuhn, P.; Handel, T. M.; Cherezov, V.; Stevens, R. C. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **2010,** *330*, 1066-1071.

18. Wacker, D.; Fenalti, G.; Brown, M. A.; Katritch, V.; Abagyan, R.; Cherezov, V.; Stevens, R. C. Conserved binding mode of human β_2 adrenergic receptor inverse agonists and antagonist revealed by X-ray crystallography. *J. Am. Chem. Soc.* **2010,** *132*, 11443-11445.

19. Xu, F.; Wu, H.; Katritch, V.; Han, G. W.; Jacobson, K. A.; Gao, Z.-G.; Cherezov, V.; Stevens, R. C. Structure of an agonist-bound human A_{2A} adenosine receptor. *Science* **2011,** *332*, 322-327.

20. Shimamura, T.; Shiroishi, M.; Weyand, S.; Tsujimoto, H.; Winter, G.; Katritch, V.; Abagyan, R.; Cherezov, V.; Liu, W.; Han, G. W.; Kobayashi, T.; Stevens, R. C.; Iwata, S. Structure of the human histamine H_1 receptor complex with doxepin. *Nature* **2011,** *475*, 65-70.

21. Rasmussen, S. G. F.; DeVree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K. Y.; Kobilka, T. S.; Thian, F. S.; Chae, P. S.; Pardon, E.; Calinski, D.; Mathiesen, J. M.; Shah, S. T. A.; Lyons, J. A.; Caffrey, M.; Gellman, S. H.; Steyaert, J.; Skiniotis, G.; Weis, W. I.; Sunahara, R. K.; Kobilka, B. K. Crystal structure of the β_2 adrenergic receptor-Gs protein complex. *Nature* **2011,** *advance online publication*, DOI: 10.1038/nature10361.

22. Strange, P. G. Antipsychotic drug action: Antagonism, inverse agonism or partial agonism. *Trends Pharmacol. Sci.* **2008,** *29*, 314-321.

23. Lagerstrom, M. C.; Schioth, H. B. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat. Rev. Drug Discovery* **2008,** *7*, 339-357.

24. Kolakowski Jr., L. F. GPCRDb: A G-protein coupled receptor database. *Recept. Channels* **1994,** *2*, 1-7.

25. Attwood, T. K.; Findlay, J. B. C. Fingerprinting G-protein-coupled receptors. *Protein Eng.* **1994,** *7*, 195-203.

26. Fuxe, K.; Marcellino, D.; Borroto-Escuela, D. O.; Guescini, M.; Fernández-Dueñas, V.; Tanganelli, S.; Rivera, A.; Ciruela, F.; Agnati, L. F. Adenosine–dopamine interactions in the pathophysiology and treatment of CNS disorders. *CNS Neurosci. Ther.* **2010,** *16*, e18-e42.

27. Capuano, B.; Crosby, I. T.; Lloyd, E. J. Schizophrenia: genesis, receptorology and current therapeutics. *Curr. Med. Chem.* **2002,** *9*, 521-548.

28. Volpini, R.; Dal Ben, D.; Lambertucci, C.; Marucci, G.; Mishra, R. C.; Ramadori, A. T.; Klotz, K.-N.; Trincavelli, M. L.; Martini, C.; Cristalli, G. Adenosine A_{2A} receptor antagonists: New 8-substituted 9-ethyladenines as tools for *in vivo* rat models of Parkinson's disease. *ChemMedChem* **2009,** *4*, 1010-1019.

29. Joyce, J. N.; Millan, M. J. Dopamine D3 receptor agonists for protection and repair in Parkinson's disease. *Curr. Opin. Pharmacol.* **2007,** *7*, 100-105.

30. Joyce, J. N. Dopamine D3 receptor as a therapeutic target for antipsychotic and antiparkinsonian drugs. *Pharmacol. Ther.* **2001,** *90*, 231-259.

31. Jacoby, E.; Bouhelal, R.; Gerspacher, M.; Seuwen, K. The 7 TM G-protein-coupled receptor target family. *ChemMedChem* **2006,** *1*, 760-782.

32. Tebben, A. J.; Schnur, D. M. Beyond rhodopsin: G protein-coupled receptor structure and modeling incorporating the β_2-adrenergic and adenosine A_{2A} crystal structures In *Chemoinformatics and Computational Chemical Biology. Methods in Molecular Biology.*, 2011; Vol. 672, pp 359-386.

33. Ballesteros, J. A.; Weinstein, H.; Stuart, C. S. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. In *Methods Neurosci.*, Academic Press: 1995; Vol. 25, pp 366-428.

34. DeLano, W. L. The PyMOL molecular graphics system, DeLano Scientific: Palo Alto, CA, USA, 2002.

35. Fritze, O.; Filipek, S.; Kuksa, V.; Palczewski, K.; Hofmann, K. P.; Ernst, O. P. Role of the conserved $NPxxY(x)_{5,6}F$ motif in the rhodopsin ground state and during activation. *Proc. Natl. Acad. Sci. U. S. A.* **2003,** *100*, 2290-2295.

36. Angel, T. E.; Chance, M. R.; Palczewski, K. Conserved waters mediate structural and functional activation of family A (rhodopsin-like) G protein-coupled receptors. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106*, 8555-8560.

37. Ballesteros, J. A.; Jensen, A. D.; Liapakis, G.; Rasmussen, S. G. F.; Shi, L.; Gether, U.; Javitch, J. A. Activation of the β_2-adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6. *J. Biol. Chem.* **2001,** *276*, 29171-29177.

38. Vogel, R.; Mahalingam, M.; Lüdeke, S.; Huber, T.; Siebert, F.; Sakmar, T. P. Functional role of the "ionic lock"--an interhelical hydrogen-bond network in family A heptahelical receptors. *J. Mol. Biol.* **2008,** *380*, 648-655.

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

40. Holst, B.; Nygaard, R.; Valentin-Hansen, L.; Bach, A.; Engelstoft, M. S.; Petersen, P. S.; Frimurer, T. M.; Schwartz, T. W. A conserved aromatic lock for the tryptophan

rotameric switch in TM-VI of seven-transmembrane receptors. *J. Biol. Chem.* **2010,** *285*, 3973-3985.

41. Javitch, J. A.; Ballesteros, J. A.; Weinstein, H.; Chen, J. A cluster of aromatic residues in the sixth membrane-spanning segment of the dopamine D2 receptor is accessible in the binding-site crevice. *Biochemistry* **1998,** *37*, 998-1006.

42. Shi, L.; Javitch, J. A. The binding site of aminergic G-protein coupled receptors: the transmembrane segments and second extracellular loop. *Annu. Rev. Pharmacol. Toxicol.* **2002,** *42*, 437-467.

43. Kobilka, B.; Schertler, G. F. X. New G-protein-coupled receptor crystal structures: insights and limitations. *Trends Pharmacol. Sci.* **2008,** *29*, 79-83.

44. Okada, T.; Le Trong, I.; Fox, B. A.; Behnke, C. A.; Stenkamp, R. E.; Palczewski, K. X-ray diffraction analysis of three-dimensional crystals of bovine rhodopsin obtained from mixed micelles. *J. Struct. Biol.* **2000,** *130*, 73-80.

45. Day, P. W.; Rasmussen, S. G. F.; Parnot, C.; Fung, J. J.; Masood, A.; Kobilka, T. S.; Yao, X.-J.; Choi, H.-J.; Weis, W. I.; Rohrer, D. K.; Kobilka, B. K. A monoclonal antibody for G protein-coupled receptor crystallography. *Nat. Methods* **2007,** *4*, 927-929.

46. Bokoch, M. P.; Zou, Y.; Rasmussen, S. G. F.; Liu, C. W.; Nygaard, R.; Rosenbaum, D. M.; Fung, J. J.; Choi, H.-J.; Thian, F. S.; Kobilka, T. S.; Puglisi, J. D.; Weis, W. I.; Pardo, L.; Prosser, R. S.; Mueller, L.; Kobilka, B. K. Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. *Nature* **2010,** *463*, 108-112.

47. Warne, T.; Serrano-Vega, M. J.; Baker, J. G.; Moukhametzianov, R.; Edwards, P. C.; Henderson, R.; Leslie, A. G. W.; Tate, C. G.; Schertler, G. F. X. Structure of a β_1-adrenergic G-protein-coupled receptor. *Nature* **2008,** *454*, 486-491.

48. Robertson, N.; Jazayeri, A.; Errey, J.; Baig, A.; Hurrell, E.; Zhukov, A.; Langmead, C. J.; Weir, M.; Marshall, F. H. The properties of thermostabilised G protein-coupled receptors (StaRs) and their use in drug discovery. *Neuropharmacology* **2011,** *60*, 36-44.

49. Rasmussen, S. G. F.; Choi, H.-J.; Fung, J. J.; Pardon, E.; Casarosa, P.; Chae, P. S.; DeVree, B. T.; Rosenbaum, D. M.; Thian, F. S.; Kobilka, T. S.; Schnapp, A.; Konetzki, I.; Sunahara, R. K.; Gellman, S. H.; Pautsch, A.; Steyaert, J.; Weis, W. I.; Kobilka, B. K. Structure of a nanobody-stabilized active state of the β_2 adrenoceptor. *Nature* **2011,** *469*, 175-180.

50. Rosenbaum, D. M.; Zhang, C.; Lyons, J. A.; Holl, R.; Aragao, D.; Arlow, D. H.; Rasmussen, S. G. F.; Choi, H.-J.; DeVree, B. T.; Sunahara, R. K.; Chae, P. S.; Gellman, S. H.; Dror, R. O.; Shaw, D. E.; Weis, W. I.; Caffrey, M.; Gmeiner, P.; Kobilka, B. K. Structure and function of an irreversible agonist-β_2 adrenoceptor complex. *Nature* **2011,** *469*, 236-240.

51. Warne, T.; Moukhametzianov, R.; Baker, J. G.; Nehme, R.; Edwards, P. C.; Leslie, A. G. W.; Schertler, G. F. X.; Tate, C. G. The structural basis for agonist and partial agonist action on a β_1-adrenergic receptor. *Nature* **2011,** *469*, 241-244.

52. Moukhametzianov, R.; Warne, T.; Edwards, P. C.; Serrano-Vega, M. J.; Leslie, A. G. W.; Tate, C. G.; Schertler, G. F. X. Two distinct conformations of helix 6 observed in antagonist-bound structures of a β_1-adrenergic receptor. *Proc. Natl. Acad. Sci. U. S. A.* **2011,** *108*, 8228-8232.

53. Lebon, G.; Warne, T.; Edwards, P. C.; Bennett, K.; Langmead, C. J.; Leslie, A. G. W.; Tate, C. G. Agonist-bound adenosine A_{2A} receptor structures reveal common features of GPCR activation. *Nature* **2011,** *474*, 521-525.

54. Kolb, P.; Rosenbaum, D. M.; Irwin, J. J.; Fung, J. J.; Kobilka, B. K.; Shoichet, B. K. Structure-based discovery of β_2-adrenergic receptor ligands. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106*, 6843-6848.

55. Okada, T.; Fujiyoshi, Y.; Silow, M.; Navarro, J.; Landau, E. M.; Shichida, Y. Functional role of internal water molecules in rhodopsin revealed by x-ray crystallography. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99*, 5982-5987.

56. Okada, T.; Sugihara, M.; Bondar, A.-N.; Elstner, M.; Entel, P.; Buss, V. The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. *J. Mol. Biol.* **2004,** *342*, 571-583.

57. Salom, D.; Lodowski, D. T.; Stenkamp, R. E.; Trong, I. L.; Golczak, M.; Jastrzebska, B.; Harris, T.; Ballesteros, J. A.; Palczewski, K. Crystal structure of a photoactivated deprotonated intermediate of rhodopsin. *Proc. Natl. Acad. Sci. U. S. A.* **2006,** *103*, 16123-16128.

58. Standfuss, J.; Xie, G.; Edwards, P. C.; Burghammer, M.; Oprian, D. D.; Schertler, G. F. X. Crystal structure of a thermally stable rhodopsin mutant. *J. Mol. Biol.* **2007,** *372*, 1179-1188.

59. Li, J.; Edwards, P. C.; Burghammer, M.; Villa, C.; Schertler, G. F. X. Structure of bovine rhodopsin in a trigonal crystal form. *J. Mol. Biol.* **2004,** *343*, 1409-1438.

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

61. Choe, H.-W.; Park, J. H.; Kim, Y. J.; Ernst, O. P. Transmembrane signaling by GPCRs: Insight from rhodopsin and opsin structures. *Neuropharmacology* **2011,** *60*, 52-57.

62. Kimura, S. R.; Tebben, A. J.; Langley, D. R. Expanding GPCR homology model binding sites via a balloon potential: A molecular dynamics refinement approach. *Proteins* **2008,** *71*, 1919-1929.

63. Tautermann, C. S. The use of G-protein coupled receptor models in lead optimization. *Future Med. Chem.* **2011,** *3*, 709-721.

64. Pearlman, D. A.; Charifson, P. S. Improved scoring of ligand-protein interactions using OWFEG free energy grids. *J. Med. Chem.* **2001,** *44*, 502-511.

65. Kirchmair, J.; Markt, P.; Distinto, S.; Wolber, G.; Langer, T. Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection - what can we learn from earlier mistakes? *J. Comput.-Aided Mol. Des.* **2008,** *22*, 213-228.

66. Topiol, S.; Sabio, M. Use of the X-ray structure of the Beta2-adrenergic receptor for drug discovery. *Bioorg. Med. Chem. Lett.* **2008,** *18*, 1598-1602.

67. Sabio, M.; Jones, K.; Topiol, S. Use of the X-ray structure of the β_2-adrenergic receptor for drug discovery. Part 2: identification of active compounds. *Bioorg. Med. Chem. Lett.* **2008,** *18*, 5391-5395.

68. Katritch, V.; Jaakola, V.-P.; Lane, J. R.; Lin, J.; Ijzerman, A. P.; Yeager, M.; Kufareva, I.; Stevens, R. C.; Abagyan, R. Structure-based discovery of novel chemotypes for adenosine A_{2A} receptor antagonists. *J. Med. Chem.* **2010,** *53*, 1799-1809.

69. Carlsson, J.; Yoo, L.; Gao, Z.-G.; Irwin, J. J.; Shoichet, B. K.; Jacobson, K. A. Structure-Based Discovery of A2A Adenosine Receptor Ligands. *J. Med. Chem.* **2010,** *53*, 3748-3755.

70. Hattori, K.; Orita, M.; Toda, S.; Imanishi, M.; Itou, S.; Nakajima, Y.; Tanabe, D.; Washizuka, K.; Araki, T.; Sakurai, M.; Matsui, S.; Imamura, E.; Ueshima, K.; Yamamoto, T.; Yamamoto, N.; Ishikawa, H.; Nakano, K.; Unami, N.; Hamada, K.; Matsumura, Y.; Takamura, F. Discovery of highly potent and selective biphenylacylsulfonamide-based β_3-adrenergic receptor agonists and molecular modeling based on the solved X-ray structure of the β_2-adrenergic receptor: Part 6. *Bioorg. Med. Chem. Lett.* **2009,** *19*, 4679-4683.

71. Pastorin, G.; Federico, S.; Paoletta, S.; Corradino, M.; Cateni, F.; Cacciari, B.; Klotz, K.-N.; Gao, Z.-G.; Jacobson, K. A.; Spalluto, G.; Moro, S. Synthesis and pharmacological characterization of a new series of 5,7-disubstituted-[1,2,4]triazolo[1,5-*a*][1,3,5]triazine derivatives as adenosine receptor antagonists: A preliminary inspection of ligand-receptor recognition process. *Bioorg. Med. Chem.* **2010,** *18*, 2524-2536.

72. Mobarec, J. C.; Sanchez, R.; Filizola, M. Modern homology modeling of G-protein coupled receptors: Which structural template to use? *J. Med. Chem.* **2009,** *52*, 5207-5216.

73. Nowak, M.; Kolaczkowski, M.; Pawlowski, M.; Bojarski, A. J. Homology modeling of the serotonin 5-HT_{1A} receptor using automated docking of bioactive compounds with defined geometry. *J. Med. Chem.* **2006,** *49*, 205-214.

74. Tehan, B. G.; Lloyd, E. J.; Wong, M. G.; Chalmers, D. K. Analysis of agonism by dopamine at the dopaminergic D_2 G-protein coupled receptor based on comparative modelling of rhodopsin. *Mol. Simul.* **2002,** *28*, 865-888.

75. Chambers, J. J.; Nichols, D. E. A homology-based model of the human 5-HT_{2A} receptor derived from an *in silico* activated G-protein coupled receptor. *J. Comput.-Aided Mol. Des.* **2002,** *16*, 511-520.

76. Dezi, C.; Brea, J.; Alvarado, M.; Ravina, E.; Masaguer, C. F.; Loza, M. I.; Sanz, F.; Pastor, M. Multistructure 3D-QSAR studies on a series of conformationally constrained butyrophenones docked into a new homology model of the 5-HT_{2A} receptor. *J. Med. Chem.* **2007,** *50*, 3242-3255.

77. Fano, A.; Ritchie, D. W.; Carrieri, A. Modeling the structural basis of human CCR5 chemokine receptor function: From homology model building and molecular dynamics validation to agonist and antagonist docking. *J. Chem. Inf. Model.* **2006,** *46*, 1223-1235.

78. Bissantz, C.; Schalon, C.; Guba, W.; Stahl, M. Focused library design in GPCR projects on the example of 5-HT_{2C} agonists: Comparison of structure-based virtual screening with ligand-based search methods. *Proteins* **2005,** *61*, 938-952.

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

80. Varady, J.; Wu, X.; Fang, X.; Min, J.; Hu, Z.; Levant, B.; Wang, S. Molecular modeling of the three-dimensional structure of dopamine 3 (D_3) subtype receptor: Discovery of novel and potent D_3 ligands through a hybrid pharmacophore- and structure-based database searching approach. *J. Med. Chem.* **2003,** *46*, 4377-4392.

81. Rivail, L.; Chipot, C.; Maigret, B.; Bestel, I.; Sicsic, S.; Tarek, M. Large-scale molecular dynamics of a G protein-coupled receptor, the human 5-HT_4 serotonin receptor, in a lipid bilayer. *J. Mol. Struct.-Theochem* **2007,** *817*, 19-26.

82. Farce, A.; Dilly, S.; Yous, S.; Berthelot, P.; Chavatte, P. Homology modelling of the serotoninergic 5-HT_{2C} receptor. *J. Enzyme Inhib. Med. Chem.* **2006,** *21*, 285 - 292.

83. Marco, E.; Foucaud, M.; Langer, I.; Escrieut, C.; Tikhonova, I. G.; Fourmy, D. Mechanism of activation of a G protein-coupled receptor, the human cholecystokinin-2 receptor. *J. Biol. Chem.* **2007,** *282*, 28779-28790.

84. Kiss, R.; Kiss, B.; Könczöl, Á.; Szalai, F.; Jelinek, I.; László, V.; Noszál, B.; Falus, A.; Keserű, G. M. Discovery of novel human histamine H_4 receptor ligands by large-scale structure-based virtual screening. *J. Med. Chem.* **2008,** *51*, 3145-3153.

85. Ivanov, A. A.; Fricks, I.; Kendall Harden, T.; Jacobson, K. A. Molecular dynamics simulation of the P2Y14 receptor. Ligand docking and identification of a putative binding site of the distal hexose moiety. *Bioorg. Med. Chem. Lett.* **2007,** *17*, 761-766.

86. Kim, S. K.; Gao, Z. G.; VanRompaey, P.; Gross, A. S.; Chen, A.; VanCalenbergh, S.; Jacobson, K. A. Modeling the adenosine receptors: Comparison of the binding domains of A_{2A} agonists and antagonists. *J. Med. Chem.* **2003,** *46*, 4847-4859.

87. Xhaard, H.; Rantanen, V. V.; Nyronen, T.; Johnson, M. S. Molecular evolution of adrenoceptors and dopamine receptors: implications for the binding of catecholamines. *J. Med. Chem.* **2006,** *49*, 1706-1719.

88. Zhang, Y.; Sham, Y. Y.; Rajamani, R.; Gao, J.; Portoghese, P. S. Homology modeling and molecular dynamics simulations of the mu opioid receptor in a membrane-aqueous system. *ChemBioChem* **2005,** *6*, 853-859.

89. Levoin, N.; Calmels, T.; Krief, S. p.; Danvy, D.; Berrebi-Bertrand, I.; Lecomte, J.-M.; Schwartz, J.-C.; Capet, M. Homology model versus X-ray structure in receptor-based drug design: A retrospective analysis with the dopamine D3 receptor. *ACS Med. Chem. Lett.* **2011,** *2*, 293-297.

90. Cavasotto, C. N.; Phatak, S. S. Homology modeling in drug discovery: Current trends and applications. *Drug Discov. Today* **2009,** *14*, 676-683.

91. Archer, E.; Maigret, B.; Escrieut, C.; Pradayrol, L.; Fourmy, D. Rhodopsin crystal: New template yielding realistic models of G-protein-coupled receptors? *Trends Pharmacol. Sci.* **2003,** *24*, 36-40.

92. Bissantz, C.; Bernard, P.; Hibert, M.; Rognan, D. Protein-based virtual screening of chemical databases. II. Are homology models of G-protein coupled receptors suitable targets? *Proteins* **2003,** *50*, 5-25.

93. de Graaf, C.; Foata, N.; Engkvist, O.; Rognan, D. Molecular modeling of the second extracellular loop of G-protein coupled receptors and its implication on structure-based virtual screening. *Proteins* **2008,** *71*, 599-620.

94. Costanzi, S. On the applicability of GPCR homology models to computer-aided drug discovery: a comparison between *in silico* and crystal structures of the β_2-adrenergic receptor. *J. Med. Chem.* **2008,** *51*, 2907-2914.

95. Goldfeld, D. A.; Zhu, K.; Beuming, T.; Friesner, R. A. Successful prediction of the intra- and extracellular loops of four G-protein-coupled receptors. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, doi:10.1073/pnas.1016951108

96. Cavasotto, C. N.; Orry, A. J. W.; Murgolo, N. J.; Czarniecki, M. F.; Kocsi, S. A.; Hawes, B. E.; Neill, K. A.; Hine, H.; Burton, M. S.; Voigt, J. H.; Abagyan, R. A.; Bayne, M. L.; Monsma, F. J. Discovery of novel chemotypes to a G-protein-coupled receptor through ligand-steered homology modeling and structure-based virtual screening. *J. Med. Chem.* **2008,** *51*, 581-588.

97. Evers, A.; Klebe, G. Ligand-supported homology modeling of G-protein-coupled receptor sites: Models sufficient for successful virtual screening. *Angew. Chem., Int. Ed. Engl.* **2004,** *43*, 248-251.

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

99. Bissantz, C.; Logean, A.; Rognan, D. High-throughput modeling of human G-protein coupled receptors: Amino acid sequence alignment, three-dimensional model

building, and receptor library screening. *J. Chem. Inf. Comput. Sci.* **2004,** *44*, 1162-1176.

100. Yuzlenko, O.; Kieć-Kononowicz, K. Molecular modeling of A_1 and A_2 adenosine receptors: comparison of rhodopsin- and β_2-adrenergic-based homology models through the docking studies. *J. Comput. Chem.* **2009,** *30*, 14-32.

101. Selent, J.; López, L.; Sanz, F.; Pastor, M. Multi-receptor binding profile of clozapine and olanzapine: A structural study based on the new β_2 adrenergic receptor template. *ChemMedChem* **2008,** *3*, 1194-1198.

102. Bruno, A.; Guadix, A. E.; Costantino, G. Molecular dynamics simulation of the heterodimeric mGluR2/5HT_{2A} complex. An atomistic resolution study of a potential new target in psychiatric conditions. *J. Chem. Inf. Model.* **2009,** *49*, 1602-1616.

103. Bruno, A.; Beato, C.; Costantino, G. Molecular dynamics simulations and docking studies on 3D models of the heterodimeric and homodimeric 5-HT_{2A} receptor subtype. *Future Med. Chem.* **2011,** *3*, 665-681.

104. Pellissier, L. P.; Sallander, J.; Campillo, M.; Gaven, F.; Queffeulou, E.; Pillot, M.; Dumuis, A.; Claeysen, S.; Bockaert, J.; Pardo, L. Conformational toggle switches implicated in basal constitutive and agonist-induced activated states of 5-hydroxytryptamine-4 receptors. *Mol. Pharmacol.* **2009,** *75*, 982-990.

105. Tan, K.; Pogozheva, I. D.; Yeo, G. S. H.; Hadaschik, D.; Keogh, J. M.; Haskell-Leuvano, C.; O'Rahilly, S.; Mosberg, H. I.; Farooqi, I. S. Functional characterization and structural modeling of obesity associated mutations in the melanocortin 4 receptor. *Endocrinology* **2009,** *150*, 114-125.

106. Lim, H. D.; Jongejan, A.; Bakker, R. A.; Haaksma, E.; de Esch, I. J. P.; Leurs, R. Phenylalanine 169 in the second extracellular loop of the human histamine H_4 receptor is responsible for the difference in agonist binding between human and mouse H_4 receptors. *J. Pharmacol. Exp. Ther.* **2008,** *327*, 88-96.

107. Straßer, A.; Wittmann, H.-J.; Seifert, R. Ligand-specific contribution of the N terminus and E2-loop to pharmacological properties of the histamine H_1-receptor. *J. Pharmacol. Exp. Ther.* **2008,** *326*, 783-791.

108. Kneissl, B.; Leonhardt, B.; Hildebrandt, A.; Tautermann, C. S. Revisiting automated G-protein coupled receptor modeling: the benefit of additional template structures for a neurokinin-1 receptor model. *J. Med. Chem.* **2009,** *52*, 3166-3173.

109. Dong, M.; Lam, P. C. H.; Pinon, D. I.; Abagyan, R.; Miller, L. J. Elucidation of the molecular basis of cholecystokinin peptide docking to its receptor using site-specific intrinsic photoaffinity labeling and molecular modeling. *Biochemistry* **2009,** *48*, 5303-5312.

110. Ivanov, A. A.; Barak, D.; Jacobson, K. A. Evaluation of homology modeling of G-protein-coupled receptors in light of the A_{2A} adenosine receptor crystallographic structure. *J. Med. Chem.* **2009,** *52*, 3284-3292.

111. Li, G.; Haney, K. M.; Kellogg, G. E.; Zhang, Y. Comparative docking study of anibamine as the first natural product CCR5 antagonist in CCR5 homology models. *J. Chem. Inf. Model.* **2009,** *49*, 120-132.

112. Dong, M.; Lam, P. C.-H.; Pinon, D. I.; Sexton, P. M.; Abagyan, R.; Miller, L. J. Spatial approximation between secretin residue five and the third extracellular loop of its receptor provides new insight into the molecular basis of natural agonist binding. *Mol. Pharmacol.* **2008,** *74*, 413-422.

113. Sudandiradoss, C.; Priya Doss, C. G.; Rajasekaran, R.; Ramanathan, K.; Purohit, R.; Sethumadhavan, R. Investigations on the interactions of scorpion neurotoxins with the predicted structure of D_1 dopamine receptor by protein-protein docking method. A bioinformatics approach. *C. R. Biol.* **2008,** *331*, 489-499.

114. Valant, C.; Gregory, K. J.; Hall, N. E.; Scammells, P. J.; Lew, M. J.; Sexton, P. M.; Christopoulos, A. A novel mechanism of G protein-coupled receptor functional selectivity. Muscarinic partial agonist MCN-A-343 as a bitopic orthosteric/allosteric ligand. *J. Biol. Chem.* **2008,** *283*, 29312-29321.

115. Tanrikulu, Y.; Proschak, E.; Werner, T.; Geppert, T.; Todoroff, N.; Klenner, A.; Kottke, T.; Sander, K.; Schneider, E.; Seifert, R.; Stark, H.; Clark, T.; Schneider, G. Homology model adjustment and ligand screening with a pseudoreceptor of the human histamine H_4 receptor. *ChemMedChem* **2009,** *4*, 820-827.

116. Ko, H.; Das, A.; Carter, R. L.; Fricks, I. P.; Zhou, Y.; Ivanov, A. A.; Melman, A.; Joshi, B. V.; Kováč, P.; Hajduch, J.; Kirk, K. L.; Harden, T. K.; Jacobson, K. A. Molecular recognition in the $P2Y_{14}$ receptor: Probing the structurally permissive terminal sugar moiety of uridine-5'-diphosphoglucose. *Bioorg. Med. Chem.* **2009,** *17*, 5298-5311.

117. Ehrlich, K.; Gotz, A.; Bollinger, S.; Tschammer, N.; Bettinetti, L.; Harterich, S.; Hubner, H.; Lanig, H.; Gmeiner, P. Dopamine D_2, D_3, and D_4 selective phenylpiperazines as molecular probes to explore the origins of subtype specific receptor binding. *J. Med. Chem.* **2009,** *52*, 4923-4935.

118. Shah, J. R.; Mosier, P. D.; Roth, B. L.; Kellogg, G. E.; Westkaemper, R. B. Synthesis, structure-affinity relationships, and modeling of AMDA analogs at 5-HT_{2A} and H_1 receptors: structural factors contributing to selectivity. *Bioorg. Med. Chem.* **2009,** *17*, 6496-6504.

119. Pecic, S.; Makkar, P.; Chaudhary, S.; Reddy, B. V.; Navarro, H. A.; Harding, W. W. Affinity of aporphines for the human 5-HT_{2A} receptor: Insights from homology modeling and molecular docking studies. *Bioorg. Med. Chem.* **2010,** *18*, 5562-5575.

120. Zhao, Y.; Lu, X.; Yang, C.-y.; Huang, Z.; Fu, W.; Hou, T.; Zhang, J. Computational modeling toward understanding agonist binding on dopamine 3. *J. Chem. Inf. Model.* **2010,** *50*, 1633-1643.

121. López, L.; Selent, J.; Ortega, R.; Masaguer, C. F.; Domínguez, E.; Areias, F.; Brea, J.; Loza, M. I.; Sanz, F.; Pastor, M. Synthesis, 3D-QSAR, and structural modeling of

benzolactam derivatives with binding affinity for the D_2 and D_3 receptors. *ChemMedChem* **2010,** *5*, 1300-1317.

122. Li, M.; Fang, H.; Du, L.; Xia, L.; Wang, B. Computational studies of the binding site of α_{1A}-adrenoceptor antagonists. *J. Mol. Model.* **2008,** *14*, 957-966.

123. Sherbiny, F.; Schiedel, A.; Maaß, A.; Müller, C. Homology modelling of the human adenosine A_{2B} receptor based on X-ray structures of bovine rhodopsin, the β_2-adrenergic receptor and the human adenosine A_{2A} receptor. *J. Comput.-Aided Mol. Des.* **2009,** *23*, 807-828.

124. Shim, J.-Y. Transmembrane helical domain of the cannabinoid CB_1 receptor. *Biophys. J.* **2009,** *96*, 3251-3262.

125. Deml, K.-F.; Beermann, S.; Neumann, D.; Strasser, A.; Seifert, R. Interactions of histamine H_1-receptor agonists and antagonists with the human histamine H_4-receptor. *Mol. Pharmacol.* **2009,** *76*, 1019-1030.

126. El Firar, A.; Voisin, T.; Rouyer-Fessard, C.; Ostuni, M. A.; Couvineau, A.; Laburthe, M. Discovery of a functional immunoreceptor tyrosine-based switch motif in a 7-transmembrane-spanning receptor: role in the orexin receptor OX1R-driven apoptosis. *The FASEB Journal* **2009,** *23*, 4069-4080.

127. Gregory, K. J.; Hall, N. E.; Tobin, A. B.; Sexton, P. M.; Christopoulos, A. Identification of orthosteric and allosteric site mutations in M_2 muscarinic acetylcholine receptors that contribute to ligand-selective signaling bias. *J. Biol. Chem.* **2010,** *285*, 7459-7474.

128. Lim, H. D.; de Graaf, C.; Jiang, W.; Sadek, P.; McGovern, P. M.; Istyastono, E. P.; Bakker, R. A.; de Esch, I. J. P.; Thurmond, R. L.; Leurs, R. Molecular determinants of ligand binding to H_4R species variants. *Mol. Pharmacol.* **2010,** *77*, 734-743.

129. Sun, X.; Li, Y.; Li, W.; Xu, Z.; Tang, Y. Computational investigation of interactions between human H_2 receptor and its agonists. *J. Mol. Graph. Model.* **2011,** *29*, 693-701.

130. Wang, Q.; Mach, R. H.; Luedtke, R. R.; Reichert, D. E. Subtype selectivity of dopamine receptor ligands: Insights from structure and ligand-based methods. *J. Chem. Inf. Model.* **2010,** *50*, 1970-1985.

131. Silva, M.; Heim, R.; Strasser, A.; Elz, S.; Dove, S. Theoretical studies on the interaction of partial agonists with the 5-HT_{2A} receptor. *J. Comput.-Aided Mol. Des.* **2011,** *25*, 51-66.

132. Dilly, S.; Liégeois, J.-F. Interaction of clozapine and its nitrenium ion with rat D2 dopamine receptors: In vitro binding and computational study. *J. Comput.-Aided Mol. Des.* **2011,** *25*, 163-169.

133. Sakhteman, A.; Lahtela-Kakkonen, M.; Poso, A. Studying the catechol binding cavity in comparative models of human dopamine D_2 receptor. *J. Mol. Graph. Model.* **2011,** *29*, 685-692.

134. Ísberg, V.; Balle, T.; Sander, T.; Jørgensen, F. S.; Gloriam, D. E. G protein- and agonist-bound serotonin 5-HT_{2A} receptor model activated by steered molecular dynamics simulations. *J. Chem. Inf. Model.* **2011,** *51*, 315-325.

135. Ostopovici-Halip, L.; Curpan, R.; Mracec, M.; Bologa, C. G. Structural determinants of the alpha2 adrenoceptor subtype selectivity. *J. Mol. Graph. Model.* **2011,** *29*, 1030-1038.

136. Patny, A.; Desai, P. V.; Avery, M. A. Homology modeling of G-protein-coupled receptors and implications in drug design. *Curr. Med. Chem.* **2006,** *13*, 1667-1691.

137. Becker, O. M.; Marantz, Y.; Shacham, S.; Inbal, B.; Heifetz, A.; Kalid, O.; Bar-Haim, S.; Warshaviak, D.; Fichman, M.; Noiman, S. G protein-coupled receptors: *In silico* drug discovery in 3D. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** *101*, 11304-11309.

138. Peeters, M. C.; van Westen, G. J. P.; Li, Q.; Ijzerman, A. P. Importance of the extracellular loops in G protein-coupled receptors for ligand recognition and receptor activation. *Trends Pharmacol. Sci.* **2011,** *32*, 35-42.

139. Nikiforovich, G. V.; Taylor, C. M.; Marshall, G. R.; Baranski, T. J. Modeling the possible conformations of the extracellular loops in G-protein-coupled receptors. *Proteins* **2010,** *78*, 271-285.

140. Sellers, B. D.; Zhu, K.; Zhao, S.; Friesner, R. A.; Jacobson, M. P. Toward better refinement of comparative models: Predicting loops in inexact environments. *Proteins* **2008,** *72*, 959-971.

141. Mehler, E. L.; Hassan, S. A.; Kortagere, S.; Weinstein, H. Ab initio computational modeling of loops in G-protein-coupled receptors: Lessons from the crystal structure of rhodopsin. *Proteins* **2006,** *64*, 673-690.

142. Mirzadegan, T.; Benko, G.; Filipek, S.; Palczewski, K. Sequence analyses of G-protein-coupled receptors: similarities to rhodopsin. *Biochemistry* **2003,** *42*, 2759-2767.

143. Conner, A. C.; Barwell, J.; Poyner, D. R.; Wheatley, M. The use of site-directed mutagenesis to study GPCRs In *Receptor Signal Transduction Protocols. Methods in Molecular Biology.*, 2011; Vol. 746, pp 85-98.

144. Mansour, A.; Meng, F.; Meador-Woodruff, J. H.; Taylor, L. P.; Civelli, O.; Akil, H. Site-directed mutagenesis of the human dopamine D_2 receptor. *Eur. J. Pharmacol., Mol. Pharmacol. Sect.* **1992,** *227*, 205-214.

145. Cho, W.; Taylor, L. P.; Mansour, A.; Akil, H. Hydrophobic residues of the D_2 dopamine receptor are important for binding and signal transduction. *J. Neurochem.* **1995,** *65*, 2105-2115.

146. Taylor, L. P.; Mansour, A.; Akil, H. Hydrophobic residues of the D_2 dopamine receptor are important for binding and signal transduction. *J. Neurochem.* **1995,** *65*, 2105-2115.

147. Javitch, J. A.; Ballesteros, J. A.; Chen, J.; Chiappa, V.; Simpson, M. M. Electrostatic and aromatic microdomains within the binding-site crevice of the D2 receptor:

contributions of the second membrane-spanning segment. *Biochemistry* **1999,** *38*, 7961-7968.

148. Javitch, J. A.; Fu, D.; Chen, J. Residues in the fifth membrane-spanning segment of the dopamine D2 receptor exposed in the binding-site crevice. *Biochemistry* **1995,** *34*, 16433-16439.

149. Javitch, J. A.; Fu, D.; Chen, J.; Karlin, A. Mapping the binding site crevice of the dopamine D_2 receptor by the substituted-cysteine accessibility method. *Neuron* **1995,** *14*, 825-831.

150. Shi, L.; Javitch, J. A. The second extracellular loop of the dopamine D_2 receptor lines the binding-site crevice. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** *101*, 440-445.

151. Shi, L.; Simpson, M. M.; Ballesteros, J. A.; Javitch, J. A. The first transmembrane segment of the dopamine D2 receptor: Accessibility in the binding-site crevice and position in the transmembrane bundle. *Biochemistry* **2001,** *40*, 12339-12348.

152. Katritch, V.; Rueda, M.; Lam, P. C.-H.; Yeager, M.; Abagyan, R. GPCR 3D homology models for ligand screening: Lessons learned from blind predictions of adenosine A2a receptor complex. *Proteins* **2010,** *78*, 197-211.

153. Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **2006,** *49*, 534-553.

154. Schrödinger Suite 2010 Induced Fit Docking protocol;, Glide, version 5.6; Schrödinger, LLC: New York, NY, 2010: Prime, version 2.2; Schrödinger, LLC: New York, NY, 2010.

155. Vilar, S.; Karpiak, J.; Berk, B.; Costanzi, S. In silico analysis of the binding of agonists and blockers to the β_2-adrenergic receptor. *J. Mol. Graph. Model.* **2011,** *29*, 809-817.

156. Krystek Jr., S. R.; Kimura, S. R.; Tebben, A. J. Modeling and active site refinement for G protein-coupled receptors: Application to the β_2 adrenergic receptor. *J. Comput.-Aided Mol. Des.* **2006,** *20*, 463-470.

157. Pérez-Nueno, V. I.; Ritchie, D. W.; Rabal, O.; Pascual, R.; Borrell, J. I.; Teixido, J. Comparison of ligand-based and receptor-based virtual screening of HIV entry inhibitors for the CXCR4 and CCR5 receptors using 3D ligand shape matching and ligand−receptor docking. *J. Chem. Inf. Model.* **2008,** *48*, 509-533.

158. Vilar, S.; Ferino, G.; Phatak, S. S.; Berk, B.; Cavasotto, C. N.; Costanzi, S. Docking-based virtual screening for ligands of G protein-coupled receptors: Not only crystal structures but also in silico models. *J. Mol. Graph. Model.* **2011,** *29*, 614-623.

159. Irwin, J. J.; Shoichet, B. K. ZINC − a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2004,** *45*, 177-182.

160. Engel, S.; Skoumbourdis, A. P.; Childress, J.; Neumann, S.; Deschamps, J. R.; Thomas, C. J.; Colson, A.-O.; Costanzi, S.; Gershengorn, M. C. A virtual screen for diverse ligands: Discovery of selective G protein-coupled receptor antagonists. *J. Am. Chem. Soc.* **2008,** *130*, 5115-5123.

161. Michino, M.; Abola, E.; Brooks, C. L.; Dixon, J. S.; Moult, J.; Stevens, R. C. Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. *Nat. Rev. Drug Discovery* **2009,** *8*, 455-463.

162. Kufareva, I.; Rueda, M.; Katritch, V.; Stevens, R. C.; Abagyan, R. Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. *Structure* **2011,** *19*, 1108-1126.

163. Bouvier, M. Oligomerization of G-protein-coupled transmitter receptors. *Nat. Rev. Neurosci.* **2001,** *2*, 274-286.

164. Milligan, G.; Lopez-Gimenez, J.; Wilson, S.; Carrillo, J. J. Selectivity in the oligomerisation of G protein-coupled receptors. *Semin. Cell Dev. Biol.* **2004,** *15*, 263-268.

165. George, S. R.; O'Dowd, B. F.; Lee, S. P. G-Protein-coupled receptor oligomerization and its potential for drug discovery. *Nat. Rev. Drug Discovery* **2002,** *1*, 808-820.

166. Portoghese, P. S. From models to molecules: Opioid receptor dimers, bivalent ligands, and selective opioid receptor probes. *J. Med. Chem.* **2001,** *44*, 2259-2269.

167. Szidonya, L.; Cserző, M.; Hunyady, L. Dimerization and oligomerization of G-protein-coupled receptors: Debated structures with established and emerging functions. *J. Endocrinol.* **2008,** *196*, 435-453.

168. Soulier, J. L.; Russo, O.; Giner, M.; Rivail, L.; Berthouze, M.; Ongeri, S.; Maigret, B.; Fischmeister, R.; Lezoualc'h, F.; Sicsic, S.; Berque-Bestel, I. Design and synthesis of specific probes for human 5-HT_4 receptor dimerization studies. *J. Med. Chem.* **2005,** *48*, 6220-6228.

169. Jastrzebska, B.; Fotiadis, D.; Jang, G.-F.; Stenkamp, R. E.; Engel, A.; Palczewski, K. Functional and structural characterization of rhodopsin oligomers. *J. Biol. Chem.* **2006,** *281*, 11917-11922.

170. White, J. H.; Wise, A.; Main, M. J.; Green, A.; Fraser, N. J.; Disney, G. H.; Barnes, A. A.; Emson, P.; Foord, S. M.; Marshall, F. H. Heterodimerization is required for the formation of a functional $GABA_B$ receptor. *Nature* **1998,** *396*, 679-682.

171. Gurevich, V. V.; Gurevich, E. V. GPCR monomers and oligomers: it takes all kinds. *Trends Neurosci.* **2008,** *31*, 74-81.

172. Panetta, R.; Greenwood, M. T. Physiological relevance of GPCR oligomerization and its impact on drug discovery. *Drug Discov. Today* **2008,** *13*, 1059-1066.

173. Nemoto, W.; Toh, H. Prediction of interfaces for oligomerizations of G-protein coupled receptors. *Proteins* **2005,** *58*, 644-660.

174. Romano, C.; Yang, W.-L.; O'Malley, K. L. Metabotropic glutamate receptor 5 Is a disulfide-linked dimer. *J. Biol. Chem.* **1996,** *271*, 28612-28616.

175. Gurevich, V. V.; Gurevich, E. V. How and why do GPCRs dimerize? *Trends Pharmacol. Sci.* **2008,** *29*, 234-240.

176. Fotiadis, D.; Liang, Y.; Filipek, S.; Saperstein, D. A.; Engel, A.; Palczewski, K. Atomic-force microscopy: Rhodopsin dimers in native disc membranes. *Nature* **2003,** *421*, 127-128.

177. Palczewski, K. Oligomeric forms of G protein-coupled receptors (GPCRs). *Trends Biochem. Sci.* **2010,** *35*, 595-600.

178. Maggio, R.; Vogel, Z.; Wess, J. Coexpression studies with mutant muscarinic/adrenergic receptors provide evidence for intermolecular "cross-talk" between G-protein-linked receptors. *Proc. Natl. Acad. Sci. U. S. A.* **1993,** *90*, 3103-3107.

179. Nemoto, W.; Fukui, K.; Toh, H. GRIPDB - G protein coupled receptor interaction partners database. *J. Recept. Signal Transduction* **2011,** *31*, 199-205.

180. Mancia, F.; Assur, Z.; Herman, A. G.; Siegel, R.; Hendrickson, W. A. Ligand sensitivity in dimeric associations of the serotonin 5-HT_{2C} receptor. *EMBO Rep.* **2008,** *9*, 363-369.

181. Hern, J. A.; Baig, A. H.; Mashanov, G. I.; Birdsall, B.; Corrie, J. E. T.; Lazareno, S.; Molloy, J. E.; Birdsall, N. J. M. Formation and dissociation of M_1 muscarinic receptor dimers seen by total internal reflection fluorescence imaging of single molecules. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107*, 2693-2698.

182. Pellissier, L. P.; Barthet, G.; Gaven, F.; Cassier, E.; Trinquet, E.; Pin, J.-P.; Marin, P.; Dumuis, A.; Bockaert, J.; Banères, J.-L.; Claeysen, S. G protein activation by serotonin type 4 receptor dimers. *J. Biol. Chem.* **2011,** *286*, 9985-9997.

183. Berthouze, M.; Ayoub, M.; Russo, O.; Rivail, L.; Sicsic, S.; Fischmeister, R.; Berque-Bestel, I.; Jockers, R.; Lezoualc'h, F. Constitutive dimerization of human serotonin 5-HT_4 receptors in living cells. *FEBS Lett.* **2005,** *579*, 2973-2980.

184. Maggio, R.; Barbier, P.; Colelli, A.; Salvadori, F.; Demontis, G.; Corsini, G. U. G protein-linked receptors: Pharmacological evidence for the formation of heterodimers. *J. Pharmacol. Exp. Ther.* **1999,** *291*, 251-257.

185. Ginés, S.; Hillion, J.; Torvinen, M.; Le Crom, S.; Casadó, V.; Canela, E. I.; Rondin, S.; Lew, J. Y.; Watson, S.; Zoli, M.; Agnati, L. F.; Vernier, P.; Lluis, C.; Ferré, S.; Fuxe, K.; Franco, R. Dopamine D_1 and adenosine A_1 receptors form functionally interacting heteromeric complexes. *Proc. Natl. Acad. Sci. U. S. A.* **2000,** *97*, 8606-8611.

186. George, S. R.; Lee, S. P.; Varghese, G.; Zeman, P. R.; Seeman, P.; Ng, G. Y. K.; O'Dowd, B. F. A transmembrane domain-derived peptide inhibits D_1 dopamine receptor function without affecting receptor oligomerization. *J. Biol. Chem.* **1998,** *273*, 30244-30248.

187. Stanasila, L.; Perez, J.-B.; Vogel, H.; Cotecchia, S. Oligomerization of the α_{1a}- and α_{1b}-adrenergic receptor subtypes. *J. Biol. Chem.* **2003,** *278*, 40239-40251.

188. Rashid, A. J.; So, C. H.; Kong, M. M. C.; Furtak, T.; El-Ghundi, M.; Cheng, R.; O'Dowd, B. F.; George, S. R. D1–D2 dopamine receptor heterooligomers with

unique pharmacology are coupled to rapid activation of G_q/11 in the striatum. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104*, 654-659.

189. Canals, M.; Marcellino, D.; Fanelli, F.; Ciruela, F.; de Benedetti, P.; Goldberg, S. R.; Neve, K.; Fuxe, K.; Agnati, L. F.; Woods, A. S.; Ferre, S.; Lluis, C.; Bouvier, M.; Franco, R. Adenosine A_{2A}-dopamine D_2 receptor-receptor heteromerization. *J. Biol. Chem.* **2003,** *278*, 46741-46749.

190. Hillion, J.; Canals, M.; Torvinen, M.; Casadó, V.; Scott, R.; Terasmaa, A.; Hansson, A.; Watson, S.; Olah, M. E.; Mallol, J.; Canela, E. I.; Zoli, M.; Agnati, L. F.; Ibáñez, C. F.; Lluis, C.; Franco, R.; Ferré, S.; Fuxe, K. Coaggregation, cointernalization, and codesensitization of adenosine A_{2A} receptors and dopamine D_2 receptors. *J. Biol. Chem.* **2002,** *277*, 18091-18097.

191. Kearn, C. S.; Blake-Palmer, K.; Daniel, E.; Mackie, K.; Glass, M. Concurrent stimulation of cannabinoid CB1 and dopamine D2 receptors enhances heterodimer formation: A mechanism for receptor cross-talk? *Mol. Pharmacol.* **2005,** *67*, 1697-1704.

192. Xu, J.; He, J.; Castleberry, A. M.; Balasubramanian, S.; Lau, A. G.; Hall, R. A. Heterodimerization of α_{2A}- and β_1-adrenergic receptors. *J. Biol. Chem.* **2003,** *278*, 10770-10777.

193. Ng, G. Y. K.; O'Dowd, B. F.; Lee, S. P.; Chung, H. T.; Brann, M. R.; Seeman, P.; George, S. R. Dopamine D2 receptor dimers and receptor-blocking peptides. *Biochem. Biophys. Res. Commun.* **1996,** *227*, 200-204.

194. Lee, S. P.; O'Dowd, B. F.; Ng, G. Y. K.; Varghese, G.; Akil, H.; Mansour, A.; Nguyen, T.; George, S. R. Inhibition of cell surface expression by mutant receptors demonstrates that D2 dopamine receptors exist as oligomers in the cell. *Mol. Pharmacol.* **2000,** *58*, 120-128.

195. Guo, W.; Shi, L.; Javitch, J. A. The fourth transmembrane segment forms the interface of the dopamine D_2 receptor homodimer. *J. Biol. Chem.* **2003,** *278*, 4385-4388.

196. Lee, S. P.; O'Dowd, B. F.; Rajaram, R. D.; Nguyen, T.; George, S. R. D2 dopamine receptor homodimerization is mediated by multiple sites of interaction, including an intermolecular interaction involving transmembrane domain 4. *Biochemistry* **2003,** *42*, 11023-11031.

197. Fonseca, J. M.; Lambert, N. A. Instability of a class A G protein-coupled receptor oligomer interface. *Mol. Pharmacol.* **2009,** *75*, 1296-1299.

198. Guo, W.; Urizar, E.; Kralikova, M.; Mobarec, J. C.; Shi, L.; Filizola, M.; Javitch, J. A. Dopamine D2 receptors form higher order oligomers at physiological expression levels. *EMBO J.* **2008,** *27*, 2293-2304.

199. Lavoie, C.; Mercier, J.-F.; Salahpour, A.; Umapathy, D.; Breit, A.; Villeneuve, L.-R.; Zhu, W.-Z.; Xiao, R.-P.; Lakatta, E. G.; Bouvier, M.; Hébert, T. E. β_1/β_2-Adrenergic

receptor heterodimerization regulates β_2-adrenergic receptor internalization and ERK signaling efficacy. *J. Biol. Chem.* **2002,** *277*, 35402-35410.

200. Mercier, J.-F.; Salahpour, A.; Angers, S.; Breit, A.; Bouvier, M. Quantitative assessment of β_1- and β_2-adrenergic receptor homo- and heterodimerization by bioluminescence resonance energy transfer. *J. Biol. Chem.* **2002,** *277*, 44925-44931.

201. Scarselli, M.; Novi, F.; Schallmach, E.; Lin, R.; Baragli, A.; Colzi, A.; Griffon, N.; Corsini, G. U.; Sokoloff, P.; Levenson, R.; Vogel, Z.; Maggio, R. D_2/D_3 dopamine receptor heterodimers exhibit unique functional properties. *J. Biol. Chem.* **2001,** *276*, 30308-30314.

202. Baragli, A.; Alturaihi, H.; Watt, H. L.; Abdallah, A.; Kumar, U. Heterooligomerization of human dopamine receptor 2 and somatostatin receptor 2: Co-immunoprecipitation and fluorescence resonance energy transfer analysis. *Cell. Signalling* **2007,** *19*, 2304-2316.

203. Jordan, B. A.; Trapaidze, N.; Gomes, I.; Nivarthi, R.; Devi, L. A. Oligomerization of opioid receptors with β_2-adrenergic receptors: A role in trafficking and mitogen-activated protein kinase activation. *Proc. Natl. Acad. Sci. U. S. A.* **2001,** *98*, 343-348.

204. Rocheville, M.; Lange, D. C.; Kumar, U.; Patel, S. C.; Patel, R. C.; Patel, Y. C. Receptors for dopamine and somatostatin: Formation of hetero-oligomers with enhanced functional activity. *Science* **2000,** *288*, 154-157.

205. Karpa, K. D.; Lin, R.; Kabbani, N.; Levenson, R. The dopamine D3 receptor interacts with itself and the truncated D3 splice variant D3nf: D3-D3nf interaction causes mislocalization of D3 receptors. *Mol. Pharmacol.* **2000,** *58*, 677-683.

206. Hebert, T. E.; Moffett, S.; Morello, J.-P.; Loisel, T. P.; Bichet, D. G.; Barret, C.; Bouvier, M. A peptide derived from a β_2-adrenergic receptor transmembrane domain inhibits both receptor dimerization and activation. *J. Biol. Chem.* **1996,** *271*, 16384-16392.

207. Angers, S.; Salahpour, A.; Joly, E.; Hilairet, S.; Chelsky, D.; Dennis, M.; Bouvier, M. Detection of β_2-adrenergic receptor dimerization in living cells using bioluminescence resonance energy transfer (BRET). *Proc. Natl. Acad. Sci. U. S. A.* **2000,** *97*, 3684-3689.

208. Fukushima, Y.; Asano, T.; Saitoh, T.; Anai, M.; Funaki, M.; Ogihara, T.; Katagiri, H.; Matsuhashi, N.; Yazaki, Y.; Sugano, K. Oligomer formation of histamine H2 receptors expressed in Sf9 and COS7 cells. *FEBS Lett.* **1997,** *409*, 283-286.

209. Milligan, G.; Ramsay, D.; Pascal, G.; Carrillo, J. J. GPCR dimerisation. *Life Sci.* **2003,** *74*, 181-188.

210. Angers, S.; Salahpour, A.; Bouvier, M. Biochemical and biophysical demonstration of GPCR oligomerization in mammalian cells. *Life Sci.* **2001,** *68*, 2243-2250.

211. Ciruela, F.; Casadó, V.; Mallol, J.; Canela, E. I.; Lluis, C.; Franco, R. Immunological identification of A1 adenosine receptors in brain cortex. *J. Neurosci. Res.* **1995,** *42*, 818-828.

212. Kaupmann, K.; Malitschek, B.; Schuler, V.; Heid, J.; Froestl, W.; Beck, P.; Mosbacher, J.; Bischoff, S.; Kulik, A.; Shigemoto, R.; Karschin, A.; Bettler, B. $GABA_B$-receptor subtypes assemble into functional heteromeric complexes. *Nature* **1998,** *396*, 683-687.

213. Pfleger, K. D. G.; Eidne, K. A. Monitoring the formation of dynamic G-protein-coupled receptor-protein complexes in living cells. *Biochem. J.* **2005,** *385*, 625-637.

214. Ayoub, M. A.; Pfleger, K. D. G. Recent advances in bioluminescence resonance energy transfer technologies to study GPCR heteromerization. *Curr. Opin. Pharmacol.* **2010,** *10*, 44-52.

215. Rivero-Müller, A.; Chou, Y.-Y.; Ji, I.; Lajic, S.; Hanyaloglu, A. C.; Jonas, K.; Rahman, N.; Ji, T. H.; Huhtaniemi, I. Rescue of defective G protein–coupled receptor function in vivo by intermolecular cooperation. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107*, 2319-2324.

216. Maggio, R.; Scarselli, M.; Novi, F.; Corsini, G. U. Heterodimerization of G-protein-coupled receptors reveals an unexpected level of pharmacological diversity. *Med. Chem. Res.* **2004,** *13*, 25-33.

217. Rozenfeld, R.; Devi, L. A. Receptor heteromerization and drug discovery. *Trends Pharmacol. Sci.* **2010,** *31*, 124-130.

218. Smith, N. J.; Milligan, G. Allostery at G protein-coupled receptor homo- and heteromers: Uncharted pharmacological landscapes. *Pharmacol. Rev.* **2010,** *62*, 701-725.

219. Zhang, A.; Liu, Z.; Kan, Y. Receptor dimerization - rationale for the design of bivalent ligands. *Curr. Top. Med. Chem.* **2007,** *7*, 343-345.

220. Dalrymple, M. B.; Pfleger, K. D. G.; Eidne, K. A. G protein-coupled receptor dimers: Functional consequences, disease states and drug targets. *Pharmacol. Ther.* **2008,** *118*, 359-371.

221. Guo, W.; Shi, L.; Filizola, M.; Weinstein, H.; Javitch, J. A. Crosstalk in G protein-coupled receptors: Changes at the transmembrane homodimer interface determine activation. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102*, 17495-17500.

222. Filizola, M. Increasingly accurate dynamic molecular models of G-protein coupled receptor oligomers: Panacea or Pandora's box for novel drug discovery? *Life Sci.* **2010,** *86*, 590-597.

223. Lambert, N. A. GPCR Dimers Fall Apart. *Sci. Signal.* **2010,** *3*, pe12.

224. Berthouze, M.; Rivail, L.; Lucas, A.; Ayoub, M. A.; Russo, O.; Sicsic, S.; Fischmeister, R.; Berque-Bestel, I.; Jockers, R.; Lezoualc'h, F. Two transmembrane Cys residues are involved in 5-HT4 receptor dimerization. *Biochem. Biophys. Res. Commun.* **2007,** *356*, 642-647.

225. Lopez-Gimenez, J. F.; Canals, M.; Pediani, J. D.; Milligan, G. The α_{1b}-adrenoceptor exists as a higher-order oligomer: Effective oligomerization is required for receptor maturation, surface delivery, and function. *Mol. Pharmacol.* **2007,** *71*, 1015-1029.

226. Carrillo, J. J.; López-Giménez, J. F.; Milligan, G. Multiple interactions between transmembrane helices generate the oligomeric α_{1b}-adrenoceptor. *Mol. Pharmacol.* **2004,** *66*, 1123-1137.

227. Russo, O.; Berthouze, M.; Giner, M.; Soulier, J. L.; Rivail, L.; Sicsic, S.; Lezoualc'h, F.; Jockers, R.; Berque-Bestel, I. Synthesis of specific bivalent probes that functionally interact with 5-HT_4 receptor dimers. *J. Med. Chem.* **2007,** *50*, 4482-4492.

228. Gonzalez-Maeso, J.; Ang, R. L.; Yuen, T.; Chan, P.; Weisstaub, N. V.; Lopez-Gimenez, J. F.; Zhou, M.; Okawa, Y.; Callado, L. F.; Milligan, G.; Gingrich, J. A.; Filizola, M.; Meana, J. J.; Sealfon, S. C. Identification of a serotonin/glutamate receptor complex implicated in psychosis. *Nature* **2008,** *452*, 93-97.

229. Vakser, I. A. Evaluation of GRAMM low-resolution docking methodology on the hemagglutinin-antibody complex. *Proteins* **1997,** *Supplement 1*, 226–230.

230. Vakser, I. A. Protein docking for low-resolution structures. *Protein Eng.* **1995,** *8*, 371-378.

231. Johnston, J. M.; Aburi, M.; Provasi, D.; Bortolato, A.; Urizar, E.; Lambert, N. A.; Javitch, J. A.; Filizola, M. Making structural sense of dimerization interfaces of delta opioid receptor homodimers. *Biochemistry* **2011,** *50*, 1682-1690.

232. Liang, Y.; Fotiadis, D.; Filipek, S.; Saperstein, D. A.; Palczewski, K.; Engel, A. Organization of the G protein-coupled receptors rhodopsin and opsin in native membranes. *J. Biol. Chem.* **2003,** *278*, 21655-21662.

233. Kim, S.-K.; Jacobson, K. A. Computational prediction of homodimerization of the A_3 adenosine receptor. *J. Mol. Graph. Model.* **2006,** *25*, 549-561.

234. Gray, J. J.; Moughon, S.; Wang, C.; Schueler-Furman, O.; Kuhlman, B.; Rohl, C. A.; Baker, D. Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *J. Mol. Biol.* **2003,** *331*, 281-299.

235. Portoghese, P. S.; Larson, D. L.; Sayre, L. M.; Yim, C. B.; Ronsisvalle, G.; Tam, S. W.; Takemori, A. E. Opioid agonist and antagonist bivalent ligands. The relationship between spacer length and selectivity at multiple opioid receptors. *J. Med. Chem.* **1986,** *29*, 1855-1861.

236. Portoghese, P. S. Bivalent ligands and the message-address concept in the design of selective opioid receptor antagonists. *Trends Pharmacol. Sci.* **1989,** *10*, 230-235.

237. Bhushan, R. G.; Sharma, S. K.; Xie, Z.; Daniels, D. J.; Portoghese, P. S. A bivalent ligand (KDN-21) reveals spinal δ and κ opioid receptors are organized as heterodimers that give rise to δ_1 and κ_2 phenotypes. Selective targeting of δ-κ heterodimers. *J. Med. Chem.* **2004,** *47*, 2969-2972.

238. Daniels, D. J.; Kulkarni, A.; Xie, Z.; Bhushan, R. G.; Portoghese, P. S. A bivalent ligand (KDAN-18) containing δ-antagonist and κ-agonist pharmacophores bridges δ_2 and κ_1 opioid receptor phenotypes. *J. Med. Chem.* **2005,** *48*, 1713-1716.

239. Daniels, D. J.; Lenard, N. R.; Etienne, C. L.; Law, P.-Y.; Roerig, S. C.; Portoghese, P. S. Opioid-induced tolerance and dependence in mice is modulated by the distance between pharmacophores in a bivalent ligand series. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102*, 19208-19213.

240. Zhang, S.; Yekkirala, A.; Tang, Y.; Portoghese, P. S. A bivalent ligand (KMN-21) antagonist for μ/κ heterodimeric opioid receptors. *Bioorg. Med. Chem. Lett.* **2009,** *19*, 6978-6980.

241. Berque-Bestel, I.; Lezoualc'h, F.; Jockers, R. Bivalent ligands as specific pharmacological tools for G protein-coupled receptor dimers. *Curr. Drug Discovery Technol.* **2008,** *5*, 312-318.

242. Albizu, L.; Moreno, J. L.; González-Maeso, J.; Sealfon, S. C. Heteromerization of G protein-coupled receptors: Relevance to neurological disorders and neurotherapeutics. *CNS Neurol. Disord.: Drug Targets* **2010,** *9*, 636-650.

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

244. Shonberg, J.; Scammells, P. J.; Capuano, B. Design strategies for bivalent ligands targeting GPCRs. *ChemMedChem* **2011,** *6*, 963-974.

245. Portoghese, P. S.; Ronsisvalle, G.; Larson, D. L.; Yim, C. B.; Sayre, L. M.; Takemori, A. E. Opioid agonist and antagonist bivalent ligands as receptor probes. *Life Sci.* **1982,** *31*, 1283-1286.

246. Halazy, S. G-protein coupled receptors bivalent ligands and drug design. *Expert Opin. Ther. Pat.* **1999,** *9*, 431-446.

247. Morphy, R.; Kay, C.; Rankovic, Z. From magic bullets to designed multiple ligands. *Drug Discov. Today* **2004,** *9*, 641-651.

248. Kuhhorn, J.; Hubner, H.; Gmeiner, P. Bivalent dopamine D_2 receptor ligands: Synthesis and binding properties. *J. Med. Chem.* **2011,** *54*, 4896-4903.

249. Leopoldo, M.; Lacivita, E.; Colabufo, N. A.; Niso, M.; Berardi, F.; Perrone, R. Bivalent ligand approach on 4-[2-(3-methoxyphenyl)ethyl]-1-(2-methoxyphenyl)piperazine: Synthesis and binding affinities for 5-HT_7 and 5-HT_{1A} receptors. *Bioorg. Med. Chem.* **2007,** *15*, 5316-5321.

250. Perez, M.; Pauwels, P. J.; Fourrier, C.; Chopin, P.; Valentin, J.-P.; John, G. W.; Marien, M.; Halazy, S. Dimerization of sumatriptan as an efficient way to design a potent, centrally and orally active 5-HT_{1B} agonist. *Bioorg. Med. Chem. Lett.* **1998,** *8*, 675-680.

251. Halazy, S.; Perez, M.; Fourrier, C.; Pallard, I.; Pauwels, P. J.; Palmier, C.; John, G. W.; Valentin, J.-P.; Bonnafous, R.; Martinez, J. Serotonin dimers: Application of the bivalent ligand approach to the design of new potent and selective 5-$HT_{1B/1D}$ agonists. *J. Med. Chem.* **1996,** *39*, 4920-4927.

252. Soriano, A.; Ventura, R.; Molero, A.; Hoen, R.; Casadó, V.; Cortés, A.; Fanelli, F.; Albericio, F.; Lluís, C.; Franco, R.; Royo, M. Adenosine A_{2A} receptor-antagonist/dopamine D_2 receptor-agonist bivalent ligands as pharmacological tools to detect A_{2A}-D_2 receptor heteromers. *J. Med. Chem.* **2009,** *52*, 5590-5602.

253. Kizuka, H.; Hanson, R. N. β-Adrenoceptor antagonist activity of bivalent ligands. 1. Diamide analogs of practolol. *J. Med. Chem.* **1987,** *30*, 722-726.

254. Zhang, Y.; Gilliam, A.; Maitra, R.; Damaj, M. I.; Tajuba, J. M.; Seltzman, H. H.; Thomas, B. F. Synthesis and biological evaluation of bivalent ligands for the cannabinoid 1 receptor. *J. Med. Chem.* **2010,** *53*, 7048-7060.

255. Zheng, Y.; Akgün, E.; Harikumar, K. G.; Hopson, J.; Powers, M. D.; Lunzer, M. M.; Miller, L. J.; Portoghese, P. S. Induced association of μ opioid (MOP) and type 2 cholecystokinin (CCK_2) receptors by novel bivalent ligands. *J. Med. Chem.* **2009,** *52*, 247-258.

256. Tanaka, T.; Nomura, W.; Narumi, T.; Masuda, A.; Tamamura, H. Bivalent ligands of CXCR4 with rigid linkers for elucidation of the dimerization state in cells. *J. Am. Chem. Soc.* **2010,** *132*, 15899-15901.

257. Abadi, A. H.; Lankow, S.; Hoefgen, B.; Decker, M.; Kassack, M. U.; Lehmann, J. Dopamine/serotonin receptor ligands, part III [1]: Synthesis and biological activities of 7,7′-alkylene-bis-6,7,8,9,14,15-hexahydro-5*H*-benz[*d*]indolo[2, 3-*g*]azecines - application of the bivalent ligand approach to a novel type of dopamine receptor antagonist. *Arch. Pharm. (Weinheim)* **2002,** *335*, 367-373.

258. Bonger, K. M.; van den Berg, R. J. B. H. N.; Heitman, L. H.; Ijzerman, A. P.; Oosterom, J.; Timmers, C. M.; Overkleeft, H. S.; van der Marel, G. A. Synthesis and evaluation of homo-bivalent GnRHR ligands. *Bioorg. Med. Chem.* **2007,** *15*, 4841-4856.

259. Decker, M.; Fulton, B. S.; Zhang, B.; Knapp, B. I.; Bidlack, J. M.; Neumeyer, J. L. Univalent and bivalent ligands of butorphan: Characteristics of the linking chain determine the affinity and potency of such opioid ligands. *J. Med. Chem.* **2009,** *52*, 7389-7396.

260. Fulton, B. S.; Knapp, B. L.; Bidlack, J. M.; Neumeyer, J. L. Effect of linker substitution on the binding of butorphan univalent and bivalent ligands to opioid receptors. *Bioorg. Med. Chem. Lett.* **2010,** *20*, 1507-1509.

261. Peng, X.; Knapp, B. I.; Bidlack, J. M.; Neumeyer, J. L. Synthesis and preliminary *in vitro* investigation of bivalent ligands containing homo- and heterodimeric pharmacophores at μ, δ and κ opioid receptors. *J. Med. Chem.* **2006,** *49*, 256-262.

262. Peng, X.; Knapp, B. I.; Bidlack, J. M.; Neumeyer, J. L. Pharmacological properties of bivalent ligands containing butorphan linked to nalbuphine, naltrexone, and naloxone at μ, δ, and κ opioid receptors. *J. Med. Chem.* **2007,** *50*, 2254-2258.

263. Bonger, K. M.; van den Berg, R. J. B. H. N.; Knijnenbur, A. D.; Heitman, L. H.; van Koppen, C. J.; Timmers, C. M.; Overkleeft, H. S.; van der Marel, G. A. Discovery of

selective luteinizing hormone receptor agonists using the bivalent ligand method. *ChemMedChem* **2009,** *4*, 1189-1195.

264. Christopoulos, A.; Grant, M. K. O.; Ayoubzadeh, N.; Kim, O. N.; Sauerberg, P.; Jeppesen, L.; El-Fakahany, E. E. Synthesis and pharmacological evaluation of dimeric muscarinic acetylcholine receptor agonists. *J. Pharmacol. Exp. Ther.* **2001,** *298*, 1260-1268.

265. Melchiorre, C. Polymethylene tetramines: A new generation of selective muscarinic antagonists. *Trends Pharmacol. Sci.* **1988,** *9*, 216-220.

266. Erez, M.; Takemori, A. E.; Portoghese, P. S. Narcotic antagonistic potency of bivalent ligands which contain β-naltrexamine. Evidence for simultaneous occupation of proximal recognition sites. *J. Med. Chem.* **1982,** *25*, 847-849.

267. Keller, M.; Teng, S.; Bernhardt, G.; Buschauer, A. Bivalent argininamide-type neuropeptide Y Y_1 antagonists do not support the hypothesis of receptor dimerisation. *ChemMedChem* **2009,** *4*, 1733-1745.

268. ChemDraw Ultra, version 12.0; CambridgeSoft Co.: Cambridge, MA.

269. Rajeswaran, W. G.; Cao, Y.; Huang, X. P.; Wroblewski, M. E.; Colclough, T.; Lee, S.; Liu, F.; Nagy, P. I.; Ellis, J.; Levine, B. A.; Nocka, K. H.; Messer, W. S. Design, synthesis, and biological characterization of bivalent 1-methyl-1,2,5,6-tetrahydropyridyl-1,2,5-thiadiazole derivatives as selective muscarinic agonists. *J. Med. Chem.* **2001,** *44*, 4563-4576.

270. Karellas, P.; McNaughton, M.; Baker, S. P.; Scammells, P. J. Synthesis of bivalent β_2-adrenergic and adenosine A_1 receptor ligands. *J. Med. Chem.* **2008,** *51*, 6128-6137.

271. Han, D.; Holger Försterling, F.; Li, X.; Deschamps, J. R.; Parrish, D.; Cao, H.; Rallapalli, S.; Clayton, T.; Teng, Y.; Majumder, S.; Sankar, S.; Roth, B. L.; Sieghart, W.; Furtmuller, R.; Rowlett, J. K.; Weed, M. R.; Cook, J. M. A study of the structure-activity relationship of $GABA_A$-benzodiazepine receptor bivalent ligands by conformational analysis with low temperature NMR and X-ray analysis. *Bioorg. Med. Chem.* **2008,** *16*, 8853-8862.

272. Choi, S.-K.; Green, D.; Ho, A.; Klein, U.; Marquess, D.; Taylor, R.; Turner, S. D. Designing selective, high sffinity ligands of 5-HT_{1D} receptor by covalent dimerization of 5-HT_{1F} ligands derived from 4-fluoro-*N*-[3-(1-methyl-4-piperidinyl)-1*H*-indol-5-yl]benzamide. *J. Med. Chem.* **2008,** *51*, 3609-3616.

273. Cao, Y.; Zhang, M.; Wu, C.; Lee, S.; Wroblewski, M. E.; Whipple, T.; Nagy, P. I.; Takács-Novák, K.; Balázs, A.; Tőrös, S.; Messer, W. S. Synthesis and biological characterization of 1-methyl-1,2,5,6-tetrahydropyridyl-1,2,5-thiadiazole derivatives as muscarinic agonists for the treatment of neurological disorders. *J. Med. Chem.* **2003,** *46*, 4273-4286.

274. Bonger, K. M.; Hoogendoorn, S.; van Koppen, C. J.; Timmers, C. M.; Overkleeft, H. S.; van der Marel, G. A. Synthesis and pharmacological evaluation of dimeric

follicle-stimulating hormone receptor antagonists. *ChemMedChem* **2009,** *4*, 2098-2102.

275. LaFrate, A. L.; Carlson, K. E.; Katzenellenbogen, J. A. Steroidal bivalent ligands for the estrogen receptor: Design, synthesis, characterization and binding affinities. *Bioorg. Med. Chem.* **2009,** *17*, 3528-3535.

276. Decker, M.; Si, Y.-G.; Knapp, B. I.; Bidlack, J. M.; Neumeyer, J. L. Synthesis and opioid receptor binding affinities of 2-substituted and 3-aminomorphinans: Ligands for μ, κ, and δ opioid receptors. *J. Med. Chem.* **2009,** *53*, 402-418.

277. LeBoulluec, K. L.; Mattson, R. J.; Mahle, C. D.; McGovern, R. T.; Nowak, H. P.; Gentile, A. J. Bivalent indoles exhibiting serotonergic binding affinity. *Bioorg. Med. Chem. Lett.* **1995,** *5*, 123-126.

278. Huber, D.; Hubner, H.; Gmeiner, P. 1,1′-Disubstituted ferrocenes as molecular hinges in mono- and bivalent dopamine receptor ligands. *J. Med. Chem.* **2009,** *52*, 6860-6870.

279. Alessi, M. L.; Norman, A. I.; Knowlton, S. E.; Ho, D. L.; Greer, S. C. Helical and coil conformations of poly(ethylene glycol) in isobutyric acid and water. *Macromolecules* **2005,** *38*, 9333-9340.

280. Bobrovnik, S. A. The influence of rigid or flexible linkage between two ligands on the effective affinity and avidity for reversible interactions with bivalent receptors. *J. Mol. Recognit.* **2007,** *20*, 253-262.

281. Lezoualc'h, F.; Jockers, R.; Berque-Bestel, I. Multivalent-based drug design applied to serotonin 5-HT_4 receptor oligomers. *Curr. Pharm. Des.* **2009,** *15*, 719-729.

282. Wang, M.; Pei, L.; Fletcher, P.; Kapur, S.; Seeman, P.; Liu, F. Schizophrenia, amphetamine-induced sensitized state and acute amphetamine exposure all show a common alteration: Increased dopamine D2 receptor dimerization. *Mol. Brain* **2010,** *3*, 25.

283. Seeman, P. Dopamine receptors and the dopamine hypothesis of schizophrenia. *Synapse* **1987,** *1*, 133-152.

284. Seeman, P.; Niznik, H. Dopamine receptors and transporters in Parkinson's disease and schizophrenia. *The FASEB Journal* **1990,** *4*, 2737-2744.

285. Seeman, P. All roads to schizophrenia lead to dopamine supersensitivity and elevated dopamine $D2^{high}$ receptors. *CNS Neurosci. Ther.* **2011,** *17*, 118-132.

286. Kang, X.; Simpson, G. M. Clozapine: More side effects but still the best antipsychotic. *J. Clin. Psychiatry* **2010,** *71*, 982-983.

287. Hippius, H. The history of clozapine. *Psychopharmacology* **1989,** *99*, S3-S5.

288. De Oliveira, I. R.; Juruena, M. F. Treatment of psychosis: 30 years of progress. *J. Clin. Pharm. Ther.* **2006,** *31*, 523-534.

289. Gründer, G.; Hippius, H.; Carlsson, A. The 'atypicality' of antipsychotics: A concept re-examined and re-defined. *Nat. Rev. Drug Discovery* **2009,** *8*, 197-202.

290. Uetrecht, J.; Zahid, N.; Tehim, A.; Mim Fu, J.; Rakhit, S. Structural features associated with reactive metabolite formation in clozapine analogues. *Chem.-Biol. Interact.* **1997,** *104*, 117-129.

291. Uetrecht, J. P. Metabolism of clozapine by neutrophils: Possible implications for clozapine-induced agranulocytosis. *Drug Saf.* **1992,** *7*, 51-56.

292. Feldman, J. Clozapine and agranulocytosis. *Psychiatr. Serv.* **1996,** *47*, 1177-1178.

293. Capuano, B.; Crosby, I. T.; Lloyd, E. J.; Taylor, D. A. Synthesis and preliminary pharmacological evaluation of 4´-arylmethyl analogues of clozapine. I. The effect of aromatic substituents. *Aust. J. Chem.* **2002,** *55*, 565-576.

294. Capuano, B.; Crosby, I. T.; Lloyd, E. J.; Podloucka, A.; Taylor, D. A. Synthesis and preliminary pharmacological evaluation of 4´-arylalkyl analogues of clozapine. II. Effect of the nature and length of the linker. *Aust. J. Chem.* **2003,** *56*, 875-886.

295. Capuano, B.; Crosby, I. T.; Lloyd, E. J.; Podloucka, A.; Taylor, D. A. Synthesis and preliminary pharmacological evaluation of 4′-arylalkyl analogues of clozapine. IV. The effects of aromaticity and isosteric replacement. *Aust. J. Chem.* **2008,** *61*, 930-940.

296. Capuano, B.; Crosby, I. T.; Lloyd, E. J.; Neve, J. E.; Taylor, D. A. Aminimides as potential CNS acting agents. I. Design, synthesis, and receptor binding of 4'-aryl aminimide analogues of clozapine as prospective novel antipsychotics. *Aust. J. Chem.* **2007,** *60*, 673-684.

297. Capuano, B.; Crosby, I.; Forsyth, C.; McRobb, F.; Moudretski, V.; Taylor, D.; Vom, A.; Yuriev, E. New hybrids of clozapine and haloperidol and their isosteric analogues: Synthesis, X-ray crystallography, conformational analysis and preliminary pharmacological evaluation. *Struct. Chem.* **2010,** *21*, 613-628.

298. Sasikumar, T. K.; Burnett, D. A.; Zhang, H.; Smith-Torhan, A.; Fawzi, A.; Lachowicz, J. E. Hydrazides of clozapine: a new class of D_1 dopamine receptor subtype selective antagonists. *Bioorg. Med. Chem. Lett.* **2006,** *16*, 4543-4547.

299. Su, J.; Tang, H.; McKittrick, B. A.; Burnett, D. A.; Zhang, H.; Smith-Torhan, A.; Fawzi, A.; Lachowicz, J. Modification of the clozapine structure by parallel synthesis. *Bioorg. Med. Chem. Lett.* **2006,** *16*, 4548-4553.

300. Williams, D. P.; Pirmohamed, M.; Naisbitt, D. J.; Maggs, J. L.; Park, B. K. Neutrophil cytotoxicity of the chemically reactive metabolite(s) of clozapine: Possible role in agranulocytosis. *J. Pharmacol. Exp. Ther.* **1997,** *283*, 1375-1382.

301. Liu, Z. C.; Uetrecht, J. P. Clozapine is oxidized by activated human neutrophils to a reactive nitrenium ion that irreversibly binds to the cells. *J. Pharmacol. Exp. Ther.* **1995,** *275*, 1476-1483.

Chapter 2

Homology modeling and docking evaluation of aminergic GPCRs

GPCRs represent a significant drug target of great interest to industry and academia alike. As such, the detailed 3D structures of these receptors will assist in our understanding of their function and will be of great use to structure-based drug design campaigns.

Prior to 2007, the only high-resolution crystal structure for GPCRs was the structure of bovine rhodopsin. Whilst this structure was used extensively for structure-based drug design, it shares relatively low sequence homology to pharmaceutically relevant GPCRs; rhodopsin has a different method of ligand binding, and contains a closed orthosteric site. Thus, homology models based on the rhodopsin template often required extensive optimization to generate a model that could be used for drug design.

Breakthroughs in the determination of high resolution crystal structures of class A, non-rhodopsin GPCRs have reinvigorated the field of structure-based drug design for GPCRs. Not only can these structures be used for drug design purposes, such as the β_2AR crystal structure, but they now add to the number of templates from which homology models can be built.

At the time of this study, very few homology models based on the β_2AR crystal structure of the pharmaceutically relevant aminergic GPCRs were published, and to our knowledge, none were freely available. As a result, we developed and optimized homology models of nine aminergic GPCRs (5-HT_{1B}R, 5-HT_{2A}R, 5-HT_{2B}R, 5-HT_{2C}R, D_2R, D_3R, D_4R, H_1R, and M_1 mAChR) and evaluated the homology models using small scale virtual screening, which was first tested on the β_2AR and A_{2A}AR crystal structures. This chapter contains a published article (refer to Appendix 1 for supporting information and Appendix 2 for the multiple

sequence alignment), which describes the homology model optimization and virtual screening evaluations of the nine homology models of aminergic GPCRs.

McRobb, F. M.; Capuano, B.; Crosby, I. T.; Chalmers, D. K.; Yuriev, E. Homology modeling and docking evaluation of aminergic G protein-coupled receptors. *J. Chem. Inf. Model.* **2010,** *50*, 626-637

It should also be noted that this work was carried out prior to the determination of both the D_3R (PDB ID: 3PBL, released November 2010) and the H_1R (PDB ID: 3RZE, released June 2011).

2.1 Declaration

2.1.1 Declaration by candidate

In the case of Chapter 2, I declare that the nature and extent of my contribution to the work was the following:

Nature of contribution	**Contribution (%)**
Development and refinement of homology models, design and running of virtual screening experiments, analysis of results, and manuscript preparation	80

The following co-authors contributed to this work:

Name	**Nature of contribution**	**Contribution (%)***
Ben Capuano	Co-author of manuscript	
Ian T. Crosby	Co-author of manuscript	
David K. Chalmers	Design of experiments and co-author of manuscript	
Elizabeth Yuriev	Design of experiments and co -author of manuscript	

** Percentage contribution only shown for co-authors who were students at Monash University at the time of their contribution to this work.*

Candidate's signature: [redacted] Date: 05 / 09/ 2011

2.1.2 Declaration by co-authors

The undersigned hereby certify that:

1. The above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors;
2. The co-authors meet the criteria for authorship, in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
3. The co-authors take public responsibility for their respective part of the publication, except for the responsible author who accepts overall responsibility for the publication;
4. There are no other authors of the publication according to these criteria;
5. Potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
6. The original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

Location of data storage: Department of Medicinal Chemistry and Drug Action, Monash Institute of Pharmaceutical Sciences, 381 Royal Parade, Parkville, Victoria, Australia

Co-author signatures:

Signed: [redacted] Date: 05 / 09/ 2011

Ben Capuano

Signed: [redacted] Date: 05 / 09/ 2011

Ian T. Crosby

Signed: [redacted] Date: 05 / 09/ 2011

David K. Chalmers

Signed: [redacted] Date: 05 / 09/ 2011

Elizabeth Yuriev

626 J. Chem. Inf. Model. 2010, 50, 626–637

Homology Modeling and Docking Evaluation of Aminergic G Protein-Coupled Receptors

Fiona M. McRobb, Ben Capuano, Ian T. Crosby, David K. Chalmers,* and Elizabeth Yuriev*

Medicinal Chemistry and Drug Action, Monash Institute of Pharmaceutical Sciences, Monash University (Parkville Campus), 381 Royal Parade, Parkville, VIC 3052 Australia

Received November 15, 2009

We report the development of homology models of dopamine (D_2, D_3, and D_4), serotonin (5-HT_{1B}, 5-HT_{2A}, 5-HT_{2B}, and 5-HT_{2C}), histamine (H_1), and muscarinic (M_1) receptors, based on the high-resolution structure of the β_2-adrenergic receptor. The homology models were built and refined using Prime. We have addressed the required modeling of extracellular loop 2, which is often implicated in ligand binding. The orthosteric sites of the models were optimized using induced fit docking, to allow for side-chain flexibility, and the resulting receptor models have been evaluated using protein validation tools. Of the nine homology models developed, six models showed moderate to good enrichment in virtual screening experiments (5-HT_{2A}, 5-HT_{1B}, D_2, 5-HT_{2C}, D_3, and M_1). The 5-HT_{2A} receptor displayed the highest enrichment in virtual screening experiments with enrichment factors of 6.1, 6.9, and 5.9 at 2, 5, and 10%, respectively, of the screened database. However, three of the models require further refinement (5-HT_{2B}, D_4, and H_1), due to difficulties in modeling some of the binding site residues as well as the extracellular loop 2. Our effort also aims to supplement the limited number of tested G protein-coupled receptor homology models based on the β_2 crystal structure that are freely available to the research community.

INTRODUCTION

There is an ongoing need for improved treatments for the disease states of schizophrenia, Parkinson's disease, and obesity, as many of the existing treatments cause undesirable side effects. For example, existing antipsychotic agents treat symptoms by acting at the dopamine D_2 and the serotonin 5-HT_{2A} receptors, however, undesired affinity at the serotonin 5-HT_{2C} and the histamine H_1 receptors has been implicated in weight gain.[1] We aim to use a structure-based approach to assist in the design of compounds that are potent at particular receptor subtypes, with reduced affinity for receptors believed to be responsible for undesirable effects. To aid our design process, we require receptor models (crystal structures or homology models), which can be used to predict the bound conformation and affinity of G protein-coupled receptor (GPCR) ligands.

Until late 2007, the only high-resolution structures of GPCRs available to inform the design of new drugs were of bovine rhodopsin,[2−4] and GPCR homology models were necessarily based on this template.[5−13] Recently, there has been a steady stream of GPCR crystal structures reported (which we hope will continue and grow). Structures of the human β_2-adrenergic receptor (β_2),[14−17] squid rhodopsin,[18,19] turkey β_1-adrenergic receptor (β_1),[20] bovine opsin,[21] bovine opsin bound to the carboxyl terminus of the Gα-subunit,[22] and human adenosine A_{2A} receptor (A_{2A})[23] have been solved, giving a wealth of structural information and revealing important structural differences between rhodopsin and other class A GPCRs, particularly in the orientation and positions of the transmembrane helices (TM) and in the structure of the loop regions.[16,24,25] The key features of these crystal structures are detailed in the Supporting Information (Table S1). The liganded structures have confirmed that the orthosteric binding site is in a similar position to that of the *cis*-retinal site in rhodopsin,[16,25] but compared to rhodopsin, the binding sites of β_1, β_2, and A_{2A} are large and open to the extracellular space. They also reveal a significantly different conformation of extracellular loop (ECL) 2, including a previously unseen helix and, in the β_1 and β_2 receptors, a second disulfide bond in ECL2, in addition to the previously identified conserved disulfide bond between ECL2 and TM3.[16,20] The recently crystallized receptors have higher homology with most other class A GPCRs than with rhodopsin and are better templates for the development of homology models (Table S2, Supporting Information).[26,27]

Since the publication of the β_2 receptor structure, a number of new GPCR homology models based on this template have been reported. A summary of the models listing their templates, modeling programs, and evaluation methods is given in Table 1. A number of automatically generated models are also present in online databases including the Protein Model Portal (http://proteinmodelportal.org)[45] and ModBase (http://modbase.compbio.ucsf.edu).[46] In developing their models, researchers have used a wide variety of alignment tools and homology modeling software including: MODELLER,[47] Sybyl,[48] Prime,[49] and ICM.[50] In most cases, these models have been built based on the 2RH1 crystal structure, but some used a combination of templates[35] (e.g., 2RH1 and 3EML). Accurate prediction of the loops, particularly those surrounding the orthosteric binding site, remains one of the more difficult aspects in GPCR homology modeling, as highlighted in the recent GPCR Dock 2008 experiment.[24] It is important to model these loops, particularly ECL2, as they are often implicated directly or indirectly

* Corresponding authors. Telephone: +61 3 9903 9110 (D.K.C.) and +61 3 9903 9611 (E.Y.). Fax: +61 3 9903 9582 (D.K.C.) and +61 3 9903 9582 (E.Y.). E-mail: David.Chalmers@pharm.monash.edu.au. (D.K.C.) and Elizabeth.Yuriev@pharm.monash.edu.au (E.Y.).

10.1021/ci900444q
Published on Web 03/01/2010

Table 1. Summary of Recently Reported GPCR Homology Models Based on β_2 Receptor Structure

receptor(s)	template	modeling programs	ECL2 modeling method	evaluation	docking program	ref
A_1 and A_{2A}	2RH1	CAChe	disulfide constraint	docking	CAChe	28
5-HT_{1A}, 5-HT_{2A}, 5-HT_{2B}, 5-HT_{2C}, 5-HT_6, 5-HT_7, α_1, α_2, D_2, D_3, D_4, H_1, M_1, M_4	2RH1	MODELLER	disulfide constraint	docking	GOLD	29
5-HT_4	2RH1	–	–	–	–	30
MC_4	2RH1	QUANTA	–	–	QUANTA	31
H_4	2RH1	MOE	disulfide and other residue constraints	–	–	32
H_1	2RH1	SYBYL	derived from 2RH1 template (partial)	–	Manual	33
NK_1	2RH1	MODELLER	derived from 2RH1 template	virtual screening	Glide	34
CCK_1	2RH1, 3EML	ICM	–	docking	ICM	35
A_{2A}	2RH1	MOE	derived from 2RH1 template	docking	IFD, Glide MOE	36
CCR_5	2RH1	InsightII	derived from 2RH1 template	docking	GOLD	37
secretin	2RH1	ICM	–	docking	ICM	38
D_1	2R4R	Swiss-model/Deep view	–	docking	GRAMM	39
M_2	2RH1	ICM	derived from 2RH1 template (partial)	docking	ICM	40
H_4	2RH1	MODELLER	–	docking	–	41
$P2Y_{14}$	3EML	Prime	derived from 3EML template (partial)	docking	Glide	42
D_2, D_3, D_4	2RH1	MODELLER	derived from 2RH1 template (partial)	docking	AUTODOCK	43
5-HT_{2A}, H_1	2RH1	MODELLER	derived from 2RH1 template	docking	GOLD	44

in ligand binding.[51] However, as loop length increases, current methods struggle, and loops over six amino acids are generally not modeled well.[52] If the reported models contained loops, then they were generated using either template-based or de novo modeling approaches (Table 1). Often a combination of these two methods was used, with the crystal structure as a template for all loops except ECL2, due to its significance in relation to the orthosteric site. A number of different techniques have been used to generate the conformation of ECL2 including modeling completely[28,34,37] or partially[33,36,40,42] based on the template or building with constraints on certain residues in the template[29,32] (usually the conserved cysteine in ECL2). Several models were further optimized using molecular dynamics, generally in a lipid bilayer.[28−30,33,37,41]

GPCR homology modeling is essentially an uncertain process, and it is important that the quality of models is evaluated. Potential causes of error, in addition to fundamental differences in structure between the template and target, include incorrect alignment and difficulties in modeling the variable loop regions as well as the refinement procedures utilized. Often the capacity to evaluate models is limited, but several approaches can be used to assess the quality of homology models. The stereochemical quality of models can be evaluated by assessment tools, such as PROCHECK,[53] WHAT_CHECK,[54] and MolProbity.[55] Site-directed mutagenesis data indicates residues that are implicated in ligand binding. Thus, by relating models and docking results to site-directed mutagenesis data, an inference can be made whether the model is a realistic representation of ligand binding. Finally, the ability of the binding site to correctly dock ligands can be assessed by docking known active compounds using programs, such as GOLD,[56] Glide,[57] or ICM.[50]

It is a significant problem for the scientific community that many reported modeling studies do not evaluate (or do not report evaluation) of published homology models. Of the studies reported in Table 1, a significant fraction do not report any evaluation. Furthermore, only a limited number of these models are freely available for use and comparison, and in many cases, the modeling procedures are described without sufficient detail to be reproducible. Our effort also aims to supplement the limited number of evaluated GPCR homology models, based on the β_2 crystal structure, that are freely available to the research community.

Specifically, to further investigate the receptors implicated in the disease states of interest to us and those implicated in drug side effects, we have built a series of nine homology models using the recent crystal structure of the β_2-adrenergic receptor (2RH1).[16] These models have been refined using experimental data and evaluated by docking and virtual screening studies. Additionally, to validate our homology modeling procedure, a β_2 model was built using the A_{2A} crystal structure as a template, and an A_{2A} model was built using the β_2 template.

EXPERIMENTAL SECTION

GPCR residues are identified using the Ballesteros−Weinstein nomenclature,[58] except for loop regions where crystal structure numbering is used. Molecular modeling was performed principally using the Schrödinger software suite. Homology models were built with Prime 1.6[49] and manually refined in Maestro 8.0.[59] Ligand molecules were prepared using LigPrep,[60] and docking was carried out with Glide 5.0.[57] Default settings were used, unless stated otherwise.

Cognate Ligand Docking. Ligands, extracted from the crystal structures of respective complexes and processed using LigPrep (in order to change their conformation), were redocked into β_2 and the A_{2A} crystal structures using the Glide program in the standard precision (SP) and extra precision (XP) modes and the induced fit docking (IFD) protocol.[61] The docking site was defined as a box 28 × 28 × 28 Å, identified by the centroid of the cocrystallized ligand. When docking into the A_{2A} crystal structure in the presence of crystallized water molecules, those close to the ligand were retained (residue numbers 505, 519, 550, 559, 565, 567, 573, and 576).

Homology Modeling. Homology models were built using the β_2 (2RH1) crystal structure[14,16] as the template. The T4-lysozyme inserted into the β_2 receptor between Gln 231 and Ser 262 to assist crystallization was removed. As the T4-lysozyme replaced ICL3, which is distant from the binding site, ICL3 was not modeled in the receptors. When using the A_{2A} crystal structure as a template, the T4-lysozyme between residues Leu 208 and Arg 222 was similarly removed.

The sequences of the human dopamine, serotonin, α- and β-adrenergic, adenosine, histamine, muscarinic, and bovine rhodopsin receptors were obtained from the Universal Protein

Resource (http://www.uniprot.org/) and aligned using ClustalW.[62] The structure alignment was manually adjusted to remove gaps in helices. Highly conserved residues in each TM were anchored, and the models were generated in Prime. The loops were refined using the refine loops tool with the extended high-loop refinement procedure. Since ECL2 is a long loop (between 13 and 33 residues), it was refined separately, in four steps, refining 6−8 residues in each step (Table S3, Supporting Information). Homology models were inspected to ensure that the side chains of the conserved residues were aligned to the template. Side chains were manually adjusted using the rotamers tool, if required. MolProbity[55] and PROCHECK[53] were used to assess the quality of the models as well as the protein report tool. MolProbity was also used to assign protonation states for the receptors and to optimize the placement of hydrogen atoms.

Comparison of Models and Crystal Structures. The A_{2A} and β_2 homology models were compared to their corresponding crystal structures by calculating the root-mean-square deviation (rmsd) of the α-carbons in Sybyl.[48] Only the residues present in the homology models were compared to the crystal structures. Rmsd values were also calculated for heavy atoms within 6 Å of the cocrystallized ligand (β_2 residues: 82, 86, 109, 110, 113, 114, 115, 117, 118, 193, 195, 199, 200, 203, 204, 207, 208, 286, 289, 290, 293, 308, 312, and 316 and A_{2A} residues: 66, 84, 85, 88, 167, 168, 169, 174, 177, 181, 246, 249, 250, 253, 264, 265, 267, 270, 271, and 274, crystal structure numbering).

Binding Site Refinement. The side-chain positions of the ligand binding site residues in each model were refined by docking an appropriate ligand into the site using IFD, which allows receptor flexibility. The docking site was defined as a box 28 × 28 × 28 Å centered on Asp 3.32, Trp 6.48, Phe 6.52, and Tyr 7.43 for all aminergic GPCRs. Up to 50 poses per ligand were collected. For initial docking, the van der Waals (vdW) radii and the partial atomic charges of both the ligand and the receptor were scaled to 0.5 to reduce the effect of steric clashes and to "soften" the surfaces of both the protein and the ligand so that a wider variety of poses could be generated. Prime was then used to optimize residues within 5 Å of the ligand poses. Finally, ligands were redocked back into all new receptor conformations, using the default vdW radii and the charge scaling (1.0 receptor, 0.8 ligand). The final models were selected after multiple iterations of model construction and refinement.

Enrichment Studies. A set of reported antagonists active at each receptor was obtained from the GLIDA database (http://pharminfo.pharm.kyoto-u.ac.jp/services/glida/index.php).[63] Active compounds (Table S4, Supporting Information) were built in ChemDraw and converted to three-dimensional (3D) using LigPrep, which also assigned formal charges according to physiological pH (pH 7.4). A set of 1000 drug-like decoy compounds with an average molecular weight of 360 g/mol was obtained from Schrödinger (http://www.schrodinger.com).[64] The properties of the ligands were assessed using QikProp[65] (Table S5, Supporting Information). Using the previously prepared 3D ligand structures, the average shape Tanimoto score was measured using ROCS,[66] taking the ligand from flexible receptor docking as a reference. The 2D Tanimoto score was calculated using UNITY in Sybyl. The decoy set, enriched with respective active compounds, was docked into each receptor structure using Glide XP and ranked by GlideScore. The docking site was defined as a box 28 × 28 × 28 Å, and the center of the binding site was identified using the coordinates of the center of carazolol in β_2 (or ZM-241,385 for docking into the A_{2A} receptor). One pose per ligand was collected. Enrichment factors were calculated at 2, 5, and 10% of the total database screened, using the following equation:[67]

$$EF^{x\%} = (Hits_{sampled}/N_{sampled}) \div (Hits_{total}/N_{total})$$

Table 2. Rmsd Values for Cognate Ligand Docking into the β_2 and A_{2A} Crystal Structures

docking method	rmsd to crystal structure (Å)		
	β_2	A_{2A}[a]	A_{2A}[b]
SP	0.85	9.83	0.79
XP	0.54	8.43	1.09
IFD	1.13	2.67	3.01

[a] Without crystallographic water. [b] With eight crystallographic water molecules.

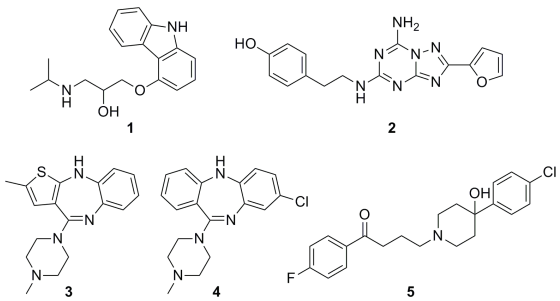


Figure 1. Structures of carazolol (**1**), ZM-241,385 (**2**), olanzapine (**3**), clozapine (**4**), and haloperidol (**5**).

RESULTS AND DISCUSSION

Cognate Ligand Docking. In a preliminary step, we investigated the ability of Glide to reproduce the bound poses of carazolol (**1**) in the β_2 and ZM-241,385 (**2**) in the A_{2A} crystal structures. We used the Glide SP and Glide XP methods, both of which use a rigid receptor, and IFD, which allows conformational changes within the receptor. The results are shown in Table 2.

Docking into the β_2 structure gave an rmsd of less than 2 Å in each case. In cognate ligand docking, an rmsd between a crystal structure and a docked pose of less than 2 Å is generally considered as a good result.[68] The A_{2A} crystal structure resolution is lower than that of the β_2 crystal structure, and the A_{2A} receptor does not make a strong ionic interaction with the ligand, such as the salt bridge observed in the aminergic GPCRs. Thus, docking into the A_{2A} crystal structure proved more challenging. Cognate docking into the A_{2A} receptor without crystallographic water molecules being present failed to reproduce the binding mode accurately with rmsd values all above 2 Å (Table 2). Only IFD was able to place the ligand in a similar conformation to the crystal structure, however, this caused a small conformational change to the binding site (rmsd 1.09 Å, residues within 6 Å of ligand). Due to the presence of water molecules interacting with the ligand in the crystal structure, we also docked into the A_{2A} receptor in the presence of eight

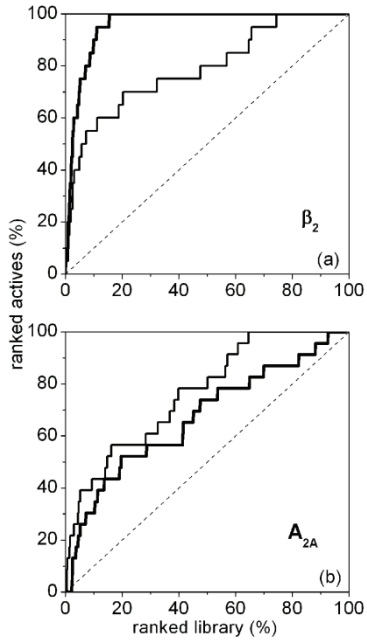


Figure 2. Enrichment plots for the (a) β_2 and (b) A_{2A} crystal structures. Bold line indicates virtual screening using Glide XP; Fine line indicates virtual screening using Glide SP; Dotted line indicates random.

crystallographic water molecules. This docking gave substantially better results for the SP and XP docking modes, with rmsd values less than 2 Å (Table 2). The furan and adenine rings superimposed well with the crystal structure, but as the substituted phenol was at the entrance of the binding site, its position proved more difficult to predict (Figure S1, Supporting Information). The results obtained with the crystallographic water molecules present indicate their importance in mediating protein−ligand interactions in the A_{2A} receptor, and their importance in being maintained for further docking studies.[69]

Virtual Screening Using the β_2 and A_{2A} Crystal Structures. To assess the ability of Glide to identify known GPCR antagonists in a library of drug-like decoys, we performed virtual screening experiments using the β_2 and A_{2A} crystal structures. To be useful in the prediction of active compounds, a docking program must be able to rank the known compounds higher than the decoy compounds. Accordingly, we docked sets of 20 (β_2) and 23 (A_{2A}) antagonists (Table S4, Supporting Information) and 1000 decoy compounds into each crystal structure. Figure 2 shows cumulative plots of the active compounds recovered against the library ranked by Glide Score.

Table 3 lists the enrichment factors and the maximum possible enrichment factor in each case. The maximum values are dependent on the total number of active compounds in each library. At the β_2 receptor, SP docking performed well, with enrichment factors of 12.1, 9.0, and 5.5 at 2, 5, and 10%, respectively, but the more computationally expensive XP docking was clearly superior with enrichment factors of 19.4, 14.0, and 9.0. XP docking also identified 95% of the active compounds in the decoy library within the top 20% of the library and placed all the ligands,

Table 3. Enrichment Factors for Virtual Screening with the β_2 and A_{2A} Crystal Structures (at x% of the Ranked Database Screened)

enrichment factor, %	β_2[a]		A_{2A}[b]	
	SP	XP	SP	XP
2	12.1	19.4	10.6	0.0
5	9.0	14.0	6.8	4.3
10	5.5	9.0	4.3	3.0

[a] Maximum enrichment factors for β_2: $EF^{2\%} = 48.6$, $EF^{5\%} = 20$, and $EF^{10\%} = 10$. [b] Maximum enrichment factors for A_{2A}: $EF^{2\%} = 44.5$, $EF^{5\%} = 19.7$, and $EF^{10\%} = 9.9$.

with the exception of the lowest ranked one, in similar conformations to the crystal structure of carazolol. This is a good result because not only did the Glide XP identify active compounds over the decoy compounds but also the docked poses conformed to experimental data. In one case, we were able to compare timolol in the 3D4S crystal structure to the docked pose of timolol obtained during virtual screening. An rmsd of 0.61 Å was obtained, indicating that our docking method was able to predict the binding mode of timolol very well.

In virtual screening, it is often found that large ligands rank higher due to a larger number of interactions with a target. This presents a challenge for the scoring function in order to identify lower molecular weight actives. In screening against the β_2 crystal structure, the active compounds had a lower molecular weight (300 g/mol) compared to that of the decoy ligands (360 g/mol) (Table S5, Supporting Information). Therefore, the enrichment factors obtained in this exercise are very encouraging and increase our confidence in the employed virtual screening procedure.

In the virtual screen against the A_{2A} receptor, the SP docking method performed moderately better than the XP throughout the virtual screening, with enrichment factors of 10.6, 6.8, and 4.3 at 2, 5, and 10%, respectively (Table 3). While enrichment was observed, overall the results were worse than those observed with the β_2 crystal structure. This is likely to be a result of both the different binding mode and site of the A_{2A} receptor in comparison to the β_2 receptor, as the adenosine ligand binds closer to the extracellular side of the receptor and does not make a strong ionic interaction with the receptor that is commonly observed in aminergic GPCRs. From the diverse range of ligand poses obtained from virtual screening, it is evident that the docking method in this instance may not be suitable for adenosine ligands or that, at least, it requires further optimization.

While we cannot directly compare our virtual screening results to others in the literature, due to different docking methods and decoy libraries, our virtual screening experiments with β_2 gave similar results to those obtained by de Graff et al. using Surflex and are substantially better than their GOLD results.[70] The results obtained were also comparable to the average enrichment factors obtained by Halgren et al. using Glide for virtual screening and using a number of different crystal structures.[71]

Development of GPCR Homology Models. We have used the 2RH1 crystal structure[14,16] of the β_2 GPCR as the template for homology modeling of the pharmaceutically significant D_2, D_3, D_4, 5-HT_{1B}, 5-HT_{2A}, 5-HT_{2B}, 5-HT_{2C}, M_1, H_1, and A_{2A} receptors. Of the currently available GPCR

crystal structures, 2RH1 has been obtained at the highest resolution (2.4 Å) and contains structural information for the extracellular side of the receptor, near the binding site. The β_1 structure would be a viable template, but due to the high similarity of the β_1 and β_2 crystal structures and the β_2 structure having the highest resolution, we did not use the β_1 structure as a template. We did not use the A_{2A} structure as a template because purinergic ligands bind to the A_{2A} receptor in a different orientation in the orthosteric site, compared to the adrenergic ligands, and adenosine receptors do not have the conserved Asp 3.32 residue, which is present in aminergic GPCRs. The A_{2A} structure was used for validation of the homology modeling procedure by building a model of β_2 based on this template.

Generating models for proteins with less than 30% overall homology to the template often means that the alignment can be unreliable.[72] However, GPCRs are a unique case, as the low sequence identity is compensated for by a high structural similarity, namely the 7 TM helices. Therefore, the key GPCR residues can be aligned to generate good quality homology models, particularly within the TM region.[73] The β_2 template shares a higher sequence identity with the receptors of interest, particularly the aminergic GPCRs, indicating that it is a better template than rhodopsin (Table S2, Supporting Information). The complete alignment for all homology models developed is provided in the Supporting Information.

To validate our homology modeling method, we built a model of the A_{2A} receptor based on the β_2 crystal structure and a model of the β_2 receptor based on the A_{2A} crystal structure. The rmsd (for the α-carbons) of the homology model to the crystal structure was 4.5 Å for the β_2 receptor, and the rmsd for the model and crystal structure for the A_{2A} receptor was 3.3 Å. The rmsd of binding site residues, within 6 Å of the ligand, for the β_2 model and the crystal structure was 3.0 Å and for the A_{2A} model and crystal structure was 3.1 Å.

The stereochemical quality of the models was evaluated using PROCHECK and MolProbity, with over 90% of backbone dihedral angles residing in the favored regions for all receptors. Any deviations in the Ramachandran plots[74,75] were investigated. The majority of deviations resided in the loop regions or away from the orthosteric site, and these deviations were left unchanged.

The significance of the ECLs in ligand binding, either directly or indirectly, is becoming increasingly apparent, from site-directed mutagenesis (e.g., Ile 184 in D_2),[76] docking, and most importantly, the recent crystal structures. In the β_2 receptor, the residues between the conserved Cys 191 in ECL2 and the top of TM5 play an essential role in shaping the binding site, maintaining the structure of the loop and interacting with the ligand (e.g., Phe 193).[14] Therefore, loop modeling of ECL2, in particular, is crucial.

Our homology models contain intracellular and extracellular loops based on the template structure (excluding ICL3 and termini), although we acknowledge that these regions have low reliability. In different receptors, these regions vary significantly in sequence, structure, and length (particularly for ECL2, which has between 11−33 amino acids). Therefore, we optimized the loop regions using the loop refinement tool in Prime, which is reported to work well (but becomes less reliable as the loop length increases).[52] The loop refinement procedure was able to successfully generate conformations for most loops. However, for a few models (5-HT_{1B}, 5-HT_{2C}, and D_2), we had to resort to the initial loop conformations for the longer loops, namely ECL2 and/or ECL3, as the loops obstructed the binding site.

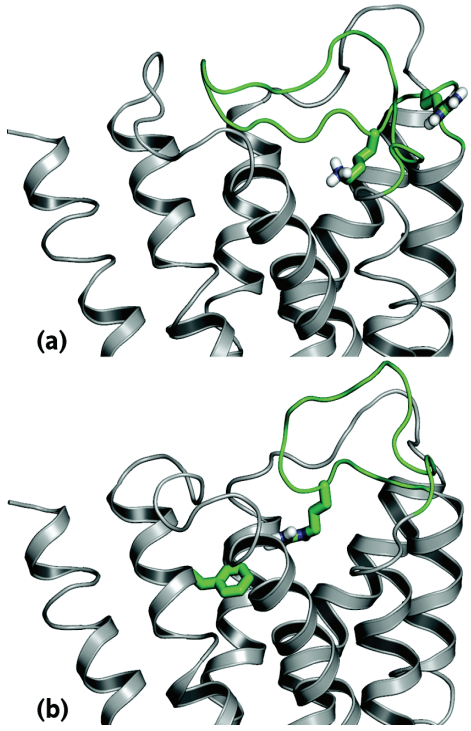


Figure 3. Problematic residues: (a) ECL2 residues Arg 211 and Lys 213 in the 5-HT_{2B} model and (b) Phe 2.61 and Arg 186 in the D_4 model (ECL2 highlighted in green). This image was created using PyMOL.[78]

The residues in ECL2 directly above the orthosteric binding site (between the cysteine in the conserved disulfide bond and the TM5 (residue 5.37)) play a significant role in defining the shape of the binding site. In the H_1, M_1, and 5-HT_{2B} receptors, this portion of the loop is longer than the template by two residues. In these receptors, it is likely that the architecture of ECL2 is different to that of the β_2 template, possibly with a helix directly above the binding site (between the disulfide bond and the TM5), or that TM5 is actually longer. We came to this conclusion because, in the case of 5-HT_{2B}, the extra residues include Arg 211 and Lys 213 (Figure 3a), which commonly make significant interactions with solvent and indicate the end of a helix in GPCRs.[77] In our models, we were unable to resolve this matter and have retained the length of TM5 from the template. In the H_1 and M_1 receptors, no such residues are present to indicate the potential start of TM5, so we again retained the length of TM5 from the β_2 crystal structure.

The conserved disulfide linkage between TM3 and ECL2 (e.g., from Cys 3.25 and Cys 191 in β_2) greatly restricts the position of ECL2. Difficulties were experienced when modeling ECL2 for the D_4 receptor, as the residue adjacent to the conserved Cys 185 was Arg 186 (Figure 3b), which protruded into the binding site and could interact with ligands in docking studies, altering the binding modes. Other problems encountered when modeling D_4 were with the

residue at position 2.61, which may be involved in subtype selectivity.[43] In the template, position 2.61 is a glycine, whereas it is a phenylalanine in D_4, which partially obstructs the entrance to the binding site. This restricts the size of the binding site, and there is no template to aid in the placement of this side chain. Similar situations were encountered with the H_1 and M_1 receptors, with large bulky residues in the models replacing smaller amino acids in the template. Further work is required to optimize these models.

Optimization of Ligand Binding Sites. The shapes of the binding sites in the homology models are likely to be influenced by the presence of carazolol in the β_2 crystal structure used as a template. We have used flexible receptor docking to moderate this bias and to also increase the volume of the binding pocket. In each receptor model, a known antagonist (Table S2, Supporting Information) was docked by IFD, generating multiple ligand−receptor structures. In each case, the shape and size of the binding site was visually inspected using surface rendering, ensuring that binding site residues were not obstructing the binding site. A single model complex was selected from the IFD structures on the basis of both the position of the ligand in the binding site and contacts observed between the ligand and the receptor and the comparison with site-directed mutagenesis. The models were also assessed using MolProbity to ensure that the flexible receptor docking did not introduce structural problems, such as steric clashes.

Our flexible receptor docking approach can be exemplified using the D_2 and 5-HT_{2A} models; upon docking olanzapine (**3**) into the D_2 homology model, several key interactions were observed that were consistent with site-directed mutagenesis data.[76,79−89] The ionic interaction between the protonated nitrogen of the ligand and the conserved Asp 3.32 (2.8 Å) and the hydrogen bonding between the secondary amine on olanzapine and Ser 5.42 (3.5 Å) were observed (Figure 4a). Many site-directed mutagenesis experiments have indicated the significance of the Asp 3.32 residue in aminergic GPCRs, due to its crucial interaction with the protonated nitrogen of the ligand.[85,90,91] TM5 residues in positions 5.42, 5.43, and 5.46, while not conserved throughout the aminergic GPCRs, are often involved in aminergic ligand−receptor interactions.[90] Olanzapine also made a substantial number of vdW contacts, particularly with the aromatic network (involving residues Trp 6.48, Phe 6.51, and Phe 6.52), and hydrophobic interactions with Ile 183 in ECL2, adjacent to the conserved Cys 182. The importance of the aromatic network has been identified by site-directed mutagenesis.[84,88,90,92]

Docking clozapine (**4**) into the 5-HT_{2A} receptor also displayed interactions consistent with site-directed mutagenesis.[91,92,94−99] Clozapine made an ionic interaction with Asp 3.32 (3.0 Å) (Figure 4b). In this instance, and unlike olanzapine in the D_2 receptor, clozapine failed to make a hydrogen bond with serine residues in TM5. Strong interactions with the aromatic network were observed, as was an interaction with Ala 230 in ECL2, adjacent to the conserved Cys 227. Similar results were obtained for the remaining homology models (Figure S2, Supporting Information).

Ligand Properties. Having built models of the D_2, D_3, D_4, 5-HT_{1B}, 5-HT_{2A}, 5-HT_{2B}, 5-HT_{2C}, M_1, and H_1 receptors, we utilized a virtual screening set of 20−57 known active compounds, specific for each receptor and subtype (Table

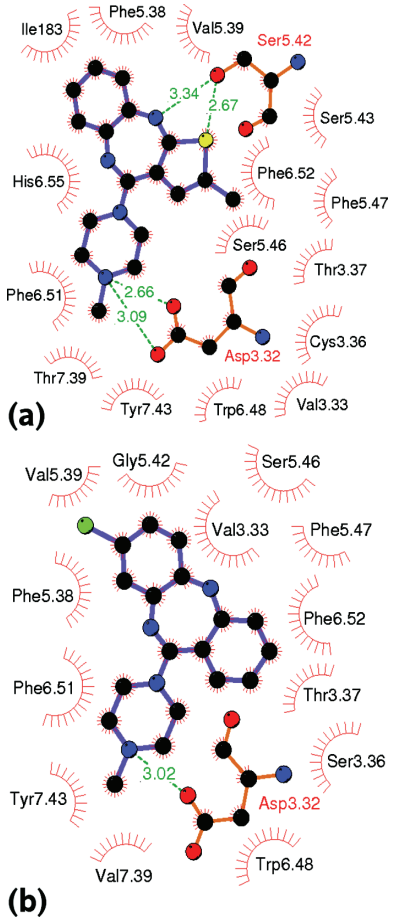


Figure 4. Schematic 2D plots of intermolecular interactions observed for flexible receptor docking: (a) olanzapine docked into the D_2 receptor and (b) clozapine docked into the 5-HT_{2A} receptor. Hydrogen atoms omitted for clarity. Nonbonded interactions are vdW, red spokes, and hydrogen bonds, dashed green lines. The plots were created with the program LIGPLOT.[93]

S4, Supporting Information), mixed with 1000 drug-like decoys obtained from Schrödinger,[64] to assess model quality in the region of the ligand binding site. To reduce the potential structural bias, the ligand used for binding site optimization (see above) was excluded from virtual screening experiments. The properties of the Schrödinger decoy library have been assessed by Friesner et al. and are believed to be representative of ligands of a pharmaceutical compound library that would be capable of competing with known actives in virtual screening.[64]

In this work, the properties of the decoy ligands and the active library compounds were predicted using QikProp, ROCS, and UNITY (Table S5, Supporting Information). Both the 2D and 3D Tanimoto calculations utilized the ligand from binding site refinement as the reference compound. The ROCS shape Tanimoto score was used to determine, on average, how similar the ligands were within each set and was found to be approximately 0.5 for each set, indicating that the sets are of similar diversity. An average 2D Tanimoto score was calculated using UNITY; the majority of the decoy ligands showed, on average, high diversity (0.2−0.3), with only the β_2 ligands displaying an average 2D diversity of 0.5. The properties investigated using QikProp included:

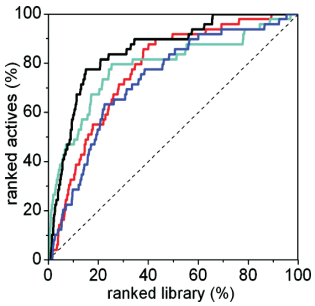


Figure 5. Enrichment plots for 5-HT_{2A} homology models: initial, red (model 1); after loop refinement, blue (model 2); after flexible receptor docking, cyan (model 3); and final, black (model 4).

molecular weight, number of rotatable bonds, polar surface area, calculated log *P*, number of hydrogen-bond donors and acceptors, and the solvent accessible volume. Generally, the properties of the active compounds were similar to those of the decoy library. The average molecular weight of the decoys was 360 g/mol, and the average molecular weight of all the active compounds was 354 g/mol. However, the average molecular weights for the β_2, H_1, and M_1 active compounds were smaller, on average, than the decoys (301, 328, and 320 g/mol, respectively). The average polar surface area of the active compounds was substantially lower than the decoys at 49 and 86.7 $Å^2$, respectively. The average number of hydrogen-bond donors in the active compounds was also lower than the decoys at 0.9 and 1.9, respectively. Of the decoy compounds, 49% contained positively charged amines, which is a characteristic known to be essential for activity for aminergic GPCRs.[90] Variations in the properties of the active compounds compared to those of the decoys were unavoidable, as the number of active compounds was limited.

Assessment of Homology Modeling Procedure by Virtual Screening. The 5-HT_{2A} homology model was used to assess our modeling procedure at different stages throughout the refinement process. Virtual screening was employed to dock a library of 1000 decoys and 49 active compounds into the 5-HT_{2A} model at four key stages during the homology modeling and the refinement process (Figure 5). First, homology model 1 (prior to any refinement) was used to screen the database. This model displayed relatively poor enrichment at 2 and 5% of the ranked database, however, 16 active compounds were recovered in the top 10% of the ranked database, showing modest enrichment of 3.3. Model 2 (after loop refinement) behaved similarly to model 1, with a slight improvement at the early stages of the screen. Model 2 also had a moderate enrichment of 2.9 at 10% of the ranked database, however only 14 actives were recovered at 10%. Model 3 (after flexible receptor docking) showed high early enrichment at 2% of the virtual screen, with an enrichment factor of 13.3. This model also displayed good enrichment at 5 and 10% of the ranked database screened (7.7 and 4.7, respectively), with 23 active compounds recovered in the top 10% of the ranked database. Model 4 (the final model), while marginally losing some very early enrichment, compared to model 3 showed the highest enrichment of 5.9 at 10% of the ranked database, recovering 29 out of the 49 active compounds. The area under the curve (AUC) was also calculated for the four enrichment plots to compare the four graphs (Table 4). Model 4 has the highest AUC value of 8463, indicating that this model displayed the best recovery of active compounds overall.

Table 4. AUC for the Enrichment Plots for the Initial, Intermediate, and Final Homology Models for the 5-HT_{2A} Receptor

model	AUC
initial (model 1)	7665.0
after loop refinement (model 2)	7279.0
after flexible receptor docking (model 3)	7837.6
final (model 4)	8463.1

Since we are interested in utilizing GPCR homology models for virtual screening purposes, we are aiming to achieve high enrichment within the top 10% of a virtual screen, as this is a common cutoff used to differentiate potential leads from inactive compounds. In this respect, the final model of the 5-HT_{2A} receptor shows a considerable improvement over the initial and intermediate models. This demonstrates the importance of such modeling steps as loop refinement and, particularly, binding site optimization.

Homology Model Evaluation by Virtual Screening. Figure 6 shows the virtual screening results as cumulative plots of the active compounds identified. Enrichment factors were calculated at 2, 5, and 10% of the ranked database screened (Table 5). The maximum possible enrichment factors for each screening run varied due to differences in the number of active compounds available for each receptor, reducing the utility of enrichment factors for direct comparison between models, particularly at 2% of the ranked database screened. However, we do report enrichment factors at 2% of the screen, as this describes early enrichment and is indicative of the success of the virtual screen. Enrichment factors at 5 and 10% of the screened database could be more easily used for more direct comparisons of the homology models. At 5% of the ranked database screened, we deemed that enrichment factors above 10 indicated high enrichment, above 5 indicated good enrichment, above 3 indicated moderate enrichment, and any enrichment factors less than or equal to 3 indicated poor enrichment. At 5% of the screen, three models showed good enrichment (5-HT_{2A}, 5-HT_{1B}, and D_2) and four models displayed moderate enrichment (M_1, 5-HT_{2C}, D_3, and 5-HT_{2B}), while two models showed little, if any, enrichment (D_4 and H_1). Similar results were also obtained at 10% of the ranked database screened. Sets of active compounds for receptors which contained more hydrogen-bond acceptors, on average, displayed slightly higher enrichment factors (e.g., 5-HT_{2A} and M_1) (Table S5, Supporting Information).

Of all homology models developed, 5-HT_{2A} performed the best with good enrichment factors obtained at 2, 5, and 10% of the screened database (Table 5). The curve tapers off after approximately 75% of the active compounds had been recovered (Figure 6b). Allowing receptor flexibility in virtual screening may improve the identification of diverse classes of active compounds, since subtle conformational changes in the protein may be required in order to accommodate different ligand types. However, this would significantly increase the computational cost. Binding modes were concordant with experimental data, with the earlier hits making more of the expected contacts, but generally, the majority of the key interactions were observed, as exemplified in Figure 7.

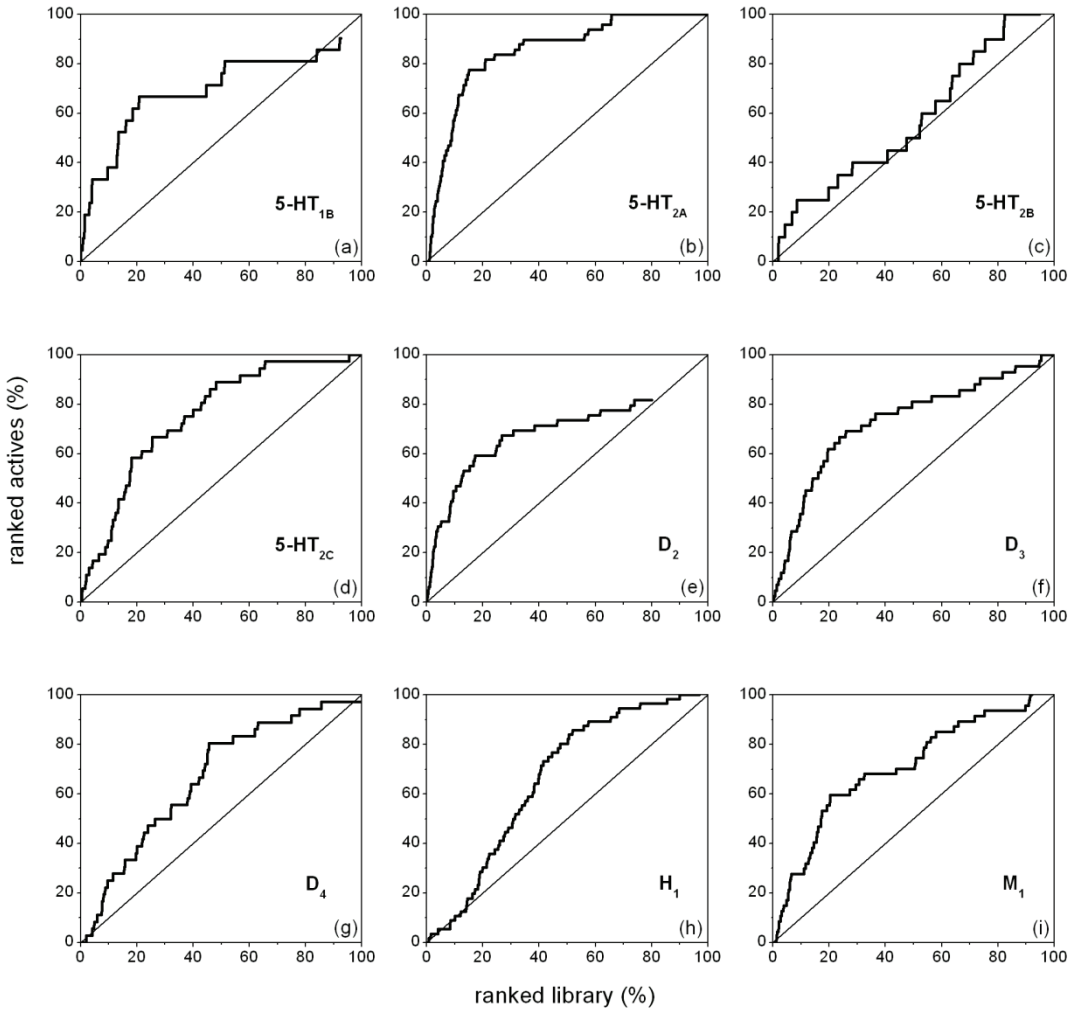


Figure 6. Enrichment plots for homology models (bold) with the percentage of the ranked library (*x*-axis) vs the ranked actives (*y*-axis). Fine line indicates random.

Table 5. Enrichment Factors for Homology Models and Maximum Enrichment Factors at *x*% of the Ranked Database Screened

	enrichment factor (at *x*% of the ranked database screened)					
receptor	2%	max 2%	5%	max 5%	10%	max 10%
5-HT_{2A}	6.1	21.4	6.9	19.8	5.9	10.0
5-HT_{1B}	9.3	48.6	6.5	19.6	3.8	9.9
D_2	7.1	21.4	6.1	19.8	4.5	10.0
M_1	3.2	22.3	3.4	19.8	2.8	10.0
5-HT_{2C}	5.3	28.0	3.2	19.9	2.4	10.0
D_3	4.6	24.3	3.2	19.7	3.5	9.9
5-HT_{2B}	4.9	48.6	3.0	20.0	2.5	10.0
D_4	1.4	28.8	1.7	19.9	2.5	10.0
H_1	1.7	18.9	1.1	18.9	1.1	10.0

The 5-HT_{1B} and D_2 models were similarly predictive. Both displayed high early enrichment at 5% of the ranked database (Table 5). The graphs tapered off after approximately 60−65% of the active compounds were retrieved (Figure 6a and e), with approximately 80% of the active compounds recovered in these screens. The compounds that were not identified in the virtual screen were often larger, bulkier compounds with a substituted tricyclic nucleus and an extended chain, indicating that the receptor conformation may need optimization in order to accommodate these ligands.

The 5-HT_{2B}, 5-HT_{2C}, D_3, and M_1 models all displayed moderate enrichment at 5 and 10% of the ranked database screened (Table 5). For the 5-HT_{2C}, D_3, and M_1 models, the majority of active compounds were identified in the virtual screen, but the number of active compounds identified tapers off after approximately 60% of actives were recovered. The docked poses generated for the active compounds retrieved later in the screen often lacked key interactions.

However, the enrichment plot for the 5-HT_{2B} model (Figure 6c) tapered off after approximately 25% of the active compounds were recovered, following which the identification of active compounds was sporadic and the curve showed little enrichment, as the curve tapered off outside the top 10% of the ranked library. This is not evident in the enrichment factors, as they were not calculated after 10% of the ranked database. The moderate virtual screening results for the 5-HT_{2B} receptor could be attributed to difficulties in

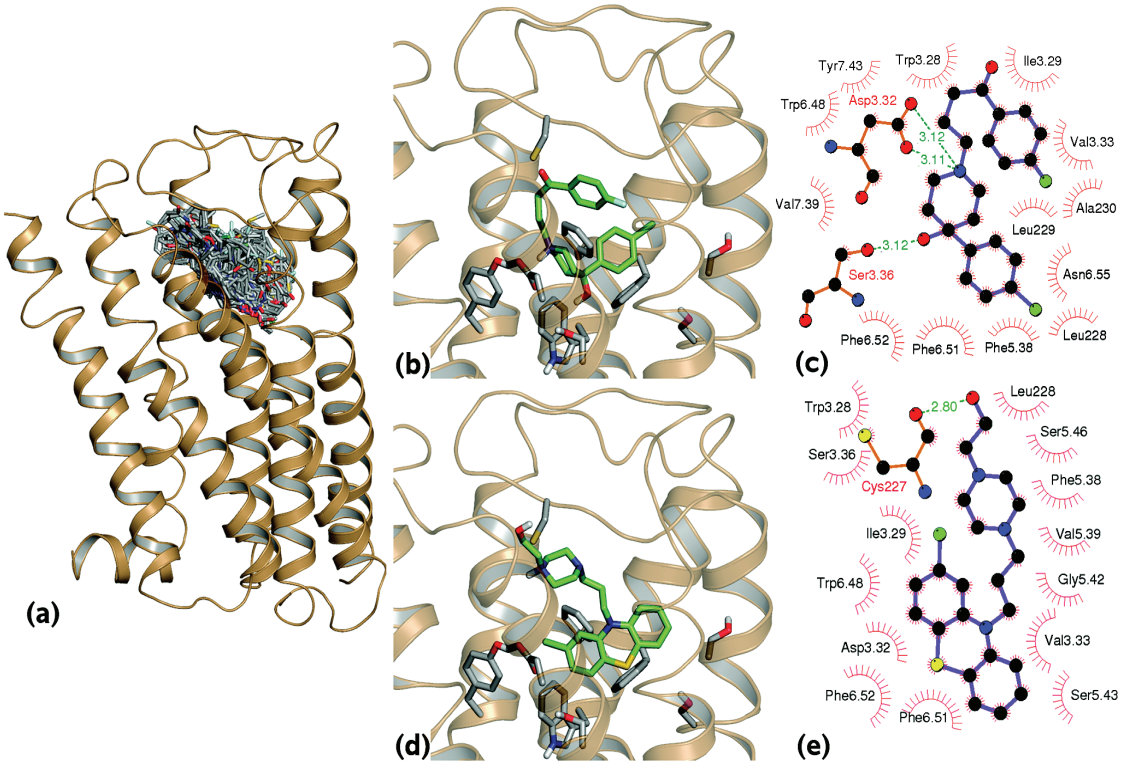


Figure 7. (a) All active compounds docked into the 5-HT_{2A} receptor during virtual screening. (b) A 3D image of an early hit, haloperidol, and (c) its corresponding schematic 2D plot showing intermolecular interactions. (d) A 3D image of a late hit, perphenazine, and (e) its corresponding schematic 2D plot showing intermolecular interactions. The 2D schematic plots were created with the program LIGPLOT.[93] Hydrogen atoms were omitted for clarity (nonbonded interactions; vdW, red spokes; hydrogen bonds, dashed greenlines). The 3D images were created using PyMOL.[78]

modeling ECL2, such as extra residues in the loop above the orthosteric site as well as bulky residues, such as Arg 211 and Lys 213 that protruded into the binding site from the loop (refer to Figure 3a). This model, particularly the loop region, will need further refinement.

The D_4 and H_1 homology models performed poorly in the virtual screens, showing little, if any, enrichment (Figure 6g and h). This is a result of difficulties experienced modeling ECL2 as well as optimizing the position of bulky residues in the binding site. The poses of the active compounds greatly varied, indicating additional refinement of the binding site, and ECL2 is required for these models.

GPCRs bind a diverse range of ligands, even at single receptors: for example, the D_2 receptor binds dibenzodiazepines (e.g., clozapine, **4**) and butyrophenones (e.g., haloperidol, **5**). Each ligand type is known to act at D_2 receptors, however, due the diversity in ligand shapes (as indicated by the average Tanimoto score in Table S5 of the Supporting Information), it is likely that the receptor undergoes conformational changes to accommodate these ligands. One receptor conformation may not, therefore, identify all the active compounds from a diverse set of ligands. During virtual screening, it was commonly observed that one class of ligand was favored over others, indicating that different receptor conformations may be required for extensive virtual screening studies.

To the best of our knowledge, the only study where a GPCR model (based, at least, partially on β_2) was evaluated by Glide-driven virtual screening is that of Kneissl et al.[34] These authors obtained an enrichment factor of 2.6 at 10% of the screened database for their model of the NK_1 receptor. The direct comparison of enrichment factors is not possible due to differences in the virtual libraries used. However, the combination of the enrichment factor values and the enrichment plots, presented by Kneissl et al., allows a cautious optimism with respect to our models. Five of our models delivered enrichment factors at 10% of the screened database greater than 2.6 and three, only marginally smaller. Furthermore, the enrichment curves presented here (except for 5-HT_{2B}, D_4 and H_1) have demonstrated better enrichment than the enrichment curves generated for NK_1.

CONCLUSIONS

We have developed a straightforward approach for the construction of aminergic GPCR homology models that utilizes Prime for initial model construction, induced fit docking for binding site optimization, and Glide for the evaluation by virtual screening of known ligands and decoys. The 5-HT_{2A} homology model was tested at four key stages during the refinement process to assess the developed protocol. Overall, each step in the refinement process advanced the models. Loop refinement improved the initial model slightly, whereas binding site optimization (by flexible receptor docking) dramatically increased the early enrichment. Using this protocol, we have developed a set of homology models (5-HT_{2A}, 5-HT_{1B}, D_2, 5-HT_{2C}, D_3, and M_1) which are able to distinguish active compounds from a set of drug-like decoys, with the 5-HT_{2A} receptor showing the

highest enrichment. A minority of the models developed using this method (5-HT_{2B}, D_4, and H_1) provided poorer results, which we believe is due to deficiencies in modeling extracellular loop 2. The homology models are available as Supporting Information so that researchers can use these structures and compare them to their own results. Further work is underway in our laboratory to investigate the refinement of these homology models using molecular dynamics in a solvated lipid bilayer.

ACKNOWLEDGMENT

F.M.M. is the recipient of an Australian Postgraduate Award (APA) scholarship. This work was supported by the Victorian Partnership for Advanced Computing (VPAC) and by the National Computational Infrastructure (NCI), which is supported by the Australian Commonwealth Government.

Supporting Information Available: A summary of recent GPCR crystal structures, sequence identity, the ligands used in flexible receptor docking and in virtual screening, length of ECL2 for each receptor and the ECL2 sections used in the loop refinement protocol, list of the active compounds docked during virtual screening, average ligand properties, cognate ligand docking results, schematic 2D plots of intermolecular interactions from flexible receptor docking, multiple sequence alignment file, and PDB files of homology models. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Kirk, S.; Glazebrook, J.; Grayson, B.; Neill, J.; Reynolds, G. Olanzapine-induced weight gain in the rat: role of 5-HT2C and histamine H1 receptors. *Psychopharmacology (Heidelberg, Ger.)* **2009**, *207*, 119–125.

(2) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Trong, I. L.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. Crystal structure of rhodopsin: a G protein-coupled receptor. *Science* **2000**, *289*, 739–745.

(3) Okada, T.; Fujiyoshi, Y.; Silow, M.; Navarro, J.; Landau, E. M.; Shichida, Y. Functional role of internal water molecules in rhodopsin revealed by x-ray crystallography. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5982–5987.

(4) Okada, T.; Sugihara, M.; Bondar, A.-N.; Elstner, M.; Entel, P.; Buss, V. The retinal conformation and its environment in rhodopsin in light of a new 2.2 crystal structure. *J. Mol. Biol.* **2004**, *342*, 571–583.

(5) Kalani, M. Y. S.; Vaidehi, N.; Hall, S. E.; Trabanino, R. J.; Freddolino, P. L.; Kalani, M. A.; Floriano, W. B.; Kam, V. W. T.; Goddard, W. A. III. The predicted 3D structure of the human D_2 dopamine receptor and the binding site and binding affinities for agonists and antagonists. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 3815–3820.

(6) Tehan, B. G.; Lloyd, E. J.; Wong, M. G.; Chalmers, D. K. Analysis of agonism by dopamine at the dopaminergic D_2 G-protein coupled receptor based on comparative modelling of rhodopsin. *Mol. Simulat.* **2002**, *28*, 865–888.

(7) Boeckler, F.; Lanig, H.; Gmeiner, P. Modeling the similarity and divergence of dopamine D_2-like receptors and identification of validated ligand-receptor complexes. *J. Med. Chem.* **2005**, *48*, 694–709.

(8) Ortore, G.; Tuccinardi, T.; Bertini, S.; Martinelli, A. A theoretical study to investigate D2DAR/D4DAR selectivity: receptor modeling and molecular docking of dopaminergic ligands. *J. Med. Chem.* **2006**, *49*, 1397–1407.

(9) Bissantz, C.; Schalon, C.; Guba, W.; Stahl, M. Focused library design in GPCR projects on the example of 5-HT_{2C} agonists: comparison of structure-based virtual screening with ligand-based search methods. *Proteins* **2005**, *61*, 938–952.

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

(11) Xhaard, H.; Rantanen, V. V.; Nyronen, T.; Johnson, M. S. Molecular evolution of adrenoceptors and dopamine receptors: implications for the binding of catecholamines. *J. Med. Chem.* **2006**, *49*, 1706–1719.

(12) Kiss, R.; Noszál, B.; Rácz, Á.; Falus, A.; Erős, D.; Keserú, G. M. Binding mode analysis and enrichment studies on homology models of the human histamine H4 receptor. *Eur. J. Med. Chem.* **2008**, *43*, 1059–1070.

(13) Evers, A.; Hessler, G.; Matter, H.; Klabunde, T. Virtual screening of biogenic amine-binding G-protein coupled receptors: comparative evaluation of protein- and ligand-based virtual screening protocols. *J. Med. Chem.* **2005**, *48*, 5448–5465.

(14) Rosenbaum, D. M.; Cherezov, V.; Hanson, M. A.; Rasmussen, S. G. F.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Yao, X.-J.; Weis, W. I.; Stevens, R. C.; Kobilka, B. K. GPCR engineering yields high-resolution structural insights into β_2-adrenergic receptor function. *Science* **2007**, *318*, 1266–1273.

(15) Rasmussen, S. G. F.; Choi, H.-J.; Rosenbaum, D. M.; Kobilka, T. S.; Thian, F. S.; Edwards, P. C.; Burghammer, M.; Ratnala, V. R. P.; Sanishvili, R.; Fischetti, R. F.; Schertler, G. F. X.; Weis, W. I.; Kobilka, B. K. Crystal structure of the human $\beta2$ adrenergic G-protein-coupled receptor. *Nature* **2007**, *450*, 383–387.

(16) Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G. F.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.; Stevens, R. C. High-resolution crystal structure of an engineered human β_2-adrenergic G protein coupled receptor. *Science* **2007**, *318*, 1258–1265.

(17) Hanson, M. A.; Cherezov, V.; Griffith, M. T.; Roth, C. B.; Jaakola, V.-P.; Chien, E. Y. T.; Velasquez, J.; Kuhn, P.; Stevens, R. C. A specific cholesterol binding site Is established by the 2.8 structure of the human β_2-adrenergic receptor. *Structure* **2008**, *16*, 897–905.

(18) Murakami, M.; Kouyama, T. Crystal structure of squid rhodopsin. *Nature* **2008**, *453*, 363–367.

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

(20) Warne, T.; Serrano-Vega, M. J.; Baker, J. G.; Moukhametzianov, R.; Edwards, P. C.; Henderson, R.; Leslie, A. G. W.; Tate, C. G.; Schertler, G. F. X. Structure of a β_1-adrenergic G-protein-coupled receptor. *Nature* **2008**, *454*, 486–491.

(21) Park, J. H.; Scheerer, P.; Hofmann, K. P.; Choe, H.-W.; Ernst, O. P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **2008**, *454*, 183–187.

(22) Scheerer, P.; Park, J. H.; Hildebrand, P. W.; Kim, Y. J.; Krausz, N.; Choe, H.-W.; Hofmann, K. P.; Ernst, O. P. Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **2008**, *455*, 497–502.

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

(24) Michino, M.; Abola, E.; Brooks, C. L.; Dixon, J. S.; Moult, J.; Stevens, R. C. Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. *Nat. Rev. Drug Discov.* **2009**, *8*, 455–463.

(25) Topiol, S.; Sabio, M. X-ray structure breakthroughs in the GPCR transmembrane region. *Biochem. Pharmacol.* **2009**, *78*, 11–20.

(26) Kobilka, B.; Schertler, G. F. X. New G-protein-coupled receptor crystal structures: insights and limitations. *Trends Pharmacol. Sci.* **2008**, *29*, 79–83.

(27) Mobarec, J. C.; Sanchez, R.; Filizola, M. Modern homology modeling of G-protein coupled receptors: which structural template to use. *J. Med. Chem.* **2009**, *52*, 5207–5216.

(28) Yuzlenko, O.; Kieæ-Kononowicz, K. Molecular modeling of A_1 and A_2 adenosine receptors: comparison of rhodopsin- and β_2-adrenergic-based homology models through the docking studies. *J. Comput. Chem.* **2009**, *30*, 14–32.

(29) Selent, J.; López, L.; Sanz, F.; Pastor, M. Multi-receptor binding profile of clozapine and olanzapine: a structural study based on the new β_2 adrenergic receptor template. *ChemMedChem* **2008**, *3*, 1194–1198.

(30) Pellissier, L. P.; Sallander, J.; Campillo, M.; Gaven, F.; Queffeulou, E.; Pillot, M.; Dumuis, A.; Claeysen, S.; Bockaert, J.; Pardo, L. Conformational toggle switches implicated in basal constitutive and agonist-induced activated states of 5-hydroxytryptamine-4 receptors. *Mol. Pharmacol.* **2009**, *75*, 982–990.

(31) Tan, K.; Pogozheva, I. D.; Yeo, G. S. H.; Hadaschik, D.; Keogh, J. M.; Haskell-Leuvano, C.; O'Rahilly, S.; Mosberg, H. I.; Farooqi, I. S. Functional characterization and structural modeling of obesity associated mutations in the melanocortin 4 receptor. *Endocrinology* **2009**, *150*, 114–125.

(32) Lim, H. D.; Jongejan, A.; Bakker, R. A.; Haaksma, E.; de Esch, I. J. P.; Leurs, R. Phenylalanine 169 in the second extracellular loop of the human histamine H_4 receptor is responsible for the difference in agonist binding between human and mouse H_4 receptors. *J. Pharmacol. Exp. Ther.* **2008**, *327*, 88–96.

(33) Straβer, A.; Wittmann, H.-J.; Seifert, R. Ligand-specific contribution of the N terminus and E2-loop to pharmacological properties of the histamine H_1-receptor. *J. Pharmacol. Exp. Ther.* **2008**, *326*, 783–791.
(34) Kneissl, B.; Leonhardt, B.; Hildebrandt, A.; Tautermann, C. S. Revisiting automated G-protein coupled receptor modeling: the benefit of additional template structures for a neurokinin-1 receptor model. *J. Med. Chem.* **2009**, *52*, 3166–3173.
(35) Dong, M.; Lam, P. C. H.; Pinon, D. I.; Abagyan, R.; Miller, L. J. Elucidation of the molecular basis of cholecystokinin peptide docking to its receptor using site-specific intrinsic photoaffinity labeling and molecular modeling. *Biochemistry* **2009**, *48*, 5303–5312.
(36) Ivanov, A. A.; Barak, D.; Jacobson, K. A. Evaluation of homology modeling of G-protein-coupled receptors in light of the A_{2A} adenosine receptor crystallographic structure. *J. Med. Chem.* **2009**, *52*, 3284–3292.
(37) Li, G.; Haney, K. M.; Kellogg, G. E.; Zhang, Y. Comparative docking study of anibamine as the first natural product CCR5 antagonist in CCR5 homology models. *J. Chem. Inf. Model.* **2009**, *49*, 120–132.
(38) Dong, M.; Lam, P. C.-H.; Pinon, D. I.; Sexton, P. M.; Abagyan, R.; Miller, L. J. Spatial approximation between secretin residue five and the third extracellular loop of its receptor provides new insight into the molecular basis of natural agonist binding. *Mol. Pharmacol.* **2008**, *74*, 413–422.
(39) Sudandiradoss, C.; Priya Doss, C. G.; Rajasekaran, R.; Ramanathan, K.; Purohit, R.; Sethumadhavan, R. Investigations on the interactions of scorpion neurotoxins with the predicted structure of D_1 dopamine receptor by protein-protein docking method. A bioinformatics approach. *C. R. Biol.* **2008**, *331*, 489–499.
(40) Valant, C.; Gregory, K. J.; Hall, N. E.; Scammells, P. J.; Lew, M. J.; Sexton, P. M.; Christopoulos, A. A novel mechanism of G protein-coupled receptor functional selectivity. *J. Biol. Chem.* **2008**, *283*, 29312–29321.
(41) Tanrikulu, Y.; Proschak, E.; Werner, T.; Geppert, T.; Todoroff, N.; Klenner, A.; Kottke, T.; Sander, K.; Schneider, E.; Seifert, R.; Stark, H.; Clark, T.; Schneider, G. Homology model adjustment and ligand screening with a pseudoreceptor of the human histamine H_4 receptor. *ChemMedChem* **2009**, *4*, 820–827.
(42) Ko, H.; Das, A.; Carter, R. L.; Fricks, I. P.; Zhou, Y.; Ivanov, A. A.; Melman, A.; Joshi, B. V.; Kováč, P.; Hajduch, J.; Kirk, K. L.; Harden, T. K.; Jacobson, K. A. Molecular recognition in the $P2Y_{14}$ receptor: Probing the structurally permissive terminal sugar moiety of uridine-5′-diphosphoglucose. *Bioorg. Med. Chem.* **2009**, *17*, 5298–5311.
(43) Ehrlich, K.; Götz, A.; Bollinger, S.; Tschammer, N.; Bettinetti, L.; Häbner, H.; Lanig, H.; Gmeiner, P. Dopamine D_2, D_3, and D_4 selective phenylpiperazines as molecular probes to explore the origins of subtype specific receptor binding. *J. Med. Chem.* **2009**, *52*, 4923–4935.
(44) Shah, J. R.; Mosier, P. D.; Roth, B. L.; Kellogg, G. E.; Westkaemper, R. B. Synthesis, structure-affinity relationships, and modeling of AMDA analogs at $5\text{-}HT_{2A}$ and H_1 receptors: structural factors contributing to selectivity. *Bioorg. Med. Chem.* **2009**, *17*, 6496–6504.
(45) Arnold, K.; Kiefer, F.; Kopp, J.; Battey, J.; Podvinec, M.; Westbrook, J.; Berman, H.; Bordoli, L.; Schwede, T. The protein model portal. *J. Struct. Funct. Genomics* **2009**, *10*, 1–8.
(46) Pieper, U.; Eswar, N.; Braberg, H.; Madhusudhan, M. S.; Davis, F. P.; Stuart, A. C.; Mirkovic, N.; Rossi, A.; Marti-Renom, M. A.; Fiser, A.; Webb, B.; Greenblatt, D.; Huang, C. C.; Ferrin, T. E.; Sali, A. MODBASE, a database of annotated comparative protein structure models, and associated resources. *Nucleic Acids Res.* **2004**, *32*, D217-D222.
(47) Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815.
(48) *Sybyl-X*, version 1.0; Tripos: St. Louis, MO, 2009.
(49) *Prime*, version 1.6; Schrödinger, LLC: New York, NY, 2007.
(50) Abagyan, R.; Totrov, M.; Kuznetsov, D. ICM - a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem.* **1994**, *15*, 488–506.
(51) Nygaard, R.; Frimurer, T. M.; Holst, B.; Rosenkilde, M. M.; Schwartz, T. W. Ligand binding and micro-switches in 7TM receptor structures. *Trends Pharmacol. Sci.* **2009**, *30*, 249–259.
(52) Dalton, J. A. R.; Jackson, R. M. An evaluation of automated homology modelling methods at low target template sequence similarity. *Bioinformatics* **2007**, *23*, 1901–1908.
(53) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **1993**, *26*, 283–291.
(54) Hooft, R. W. W.; Vriend, G.; Sander, C.; Abola, E. E. Errors in protein structures. *Nature* **1996**, *381*, 272.
(55) Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B., III; Snoeyink, J.; Richardson, J. S.; Richardson, D. C. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **2007**, *35*, W375–383.
(56) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
(57) *Glide*, version 5.0; Schrödinger, LLC: New York, NY, 2008.
(58) Ballesteros, J. A.; Weinstein, H.; Stuart, C. S. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. In. *Methods Neurosci., Academic Press:* **1995**, *25*, 366–428.
(59) *Maestro*, version 8.0; Schrödinger, LLC: New York, NY, 2007.
(60) *LigPrep*, version 2.2; Schrödinger, LLC: New York, NY, 2005.
(61) *Schrödinger Suite 2008 Induced Fit Docking protocol*; Schrödinger, LLC: New York, NY, 2005.
(62) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **1994**, *22*, 4673–4680.
(63) Okuno, Y.; Tamon, A.; Yabuuchi, H.; Niijima, S.; Minowa, Y.; Tonomura, K.; Kunimoto, R.; Feng, C. GLIDA: GPCR ligand database for chemical genomics drug discovery database and tools update. *Nucleic Acids Res.* **2008**, *36*, D907–912.
(64) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
(65) *QikProp*; version 3.1; Schrödinger, LLC: New York, NY, 2008.
(66) *ROCS*, version 2.3.1; OpenEye Scientific Software Inc.: Santa Fe, New Mexico, 2007.
(67) Pearlman, D. A.; Charifson, P. S. Improved scoring of ligand-protein interactions using OWFEG free energy grids. *J. Med. Chem.* **2001**, *44*, 502–511.
(68) Jain, A. Bias, reporting, and sharing: computational evaluations of docking methods. *J. Comput. Aided Mol. Des.* **2008**, *22*, 201–212.
(69) Huang, N.; Shoichet, B. K. Exploiting ordered waters in molecular docking. *J. Med. Chem.* **2008**, *51*, 4862–4865.
(70) de Graaf, C.; Rognan, D. Selective structure-based virtual screening for full and partial agonists of the $\beta 2$ adrenergic receptor. *J. Med. Chem.* **2008**, *51*, 4978–4985.
(71) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750–1759.
(72) Forster, M. J. Molecular modelling in structural biology. *Micron* **2002**, *33*, 365–384.
(73) Cavasotto, C. N.; Phatak, S. S. Homology modeling in drug discovery: current trends and applications. *Drug Discovery Today* **2009**, *14*, 676–683.
(74) Lovell, S. C.; Davis, I. W.; Arendall III, W. B.; de Bakker, P. I. W.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C. Structure validation by Cα geometry: ϕ, Ψ and $C\beta$ deviation. *Proteins* **2003**, *50*, 437–450.
(75) Ho, B. K.; Thomas, A.; Brasseur, R. Revisiting the Ramachandran plot: hard-sphere repulsion, electrostatics, and H-bonding in the α-helix. *Protein Sci.* **2003**, *12*, 2508–2522.
(76) Shi, L.; Javitch, J. A. The second extracellular loop of the dopamine D_2 receptor lines the binding-site crevice. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 440–445.
(77) Zhang, D.; Weinstein, H. Polarity conserved positions in transmembrane domains of G-protein coupled receptors and bacteriorhodopsin. *FEBS Lett.* **1994**, *337*, 207–212.
(78) DeLano, W. L. *The PyMOL molecular graphics system*, DeLano Scientific: Palo Alto, CA, 2002.
(79) Shi, L.; Simpson, M. M.; Ballesteros, J. A.; Javitch, J. A. The first transmembrane segment of the dopamine D2 receptor: accessibility in the binding-site crevice and position in the transmembrane bundle. *Biochemistry* **2001**, *40*, 12339–12348.
(80) Javitch, J. A.; Shi, L.; Simpson, M. M.; Chen, J.; Chiappa, V.; Visiers, I.; Weinstein, H.; Ballesteros, J. A. The fourth transmembrane segment of the dopamine D2 receptor: accessibility in the binding-site crevice and position in the transmembrane bundle. *Biochemistry* **2000**, *39*, 12190–12199.
(81) Javitch, J. A.; Ballesteros, J. A.; Chen, J.; Chiappa, V.; Simpson, M. M. Electrostatic and aromatic microdomains within the binding-site crevice of the D2 receptor: contributions of the second membrane-spanning segment. *Biochemistry* **1999**, *38*, 7961–7968.
(82) Fu, D.; Ballesteros, J. A.; Weinstein, H.; Chen, J.; Javitch, J. A. Residues in the seventh membrane-spanning segment of the dopamine D2 receptor accessible in the binding-site crevice. *Biochemistry* **1996**, *35*, 11278–11285.
(83) Javitch, J. A.; Fu, D.; Chen, J. Residues in the fifth membrane-spanning segment of the dopamine D2 receptor exposed in the binding-site crevice. *Biochemistry* **1995**, *34*, 16433–16439.

(84) Javitch, J. A.; Ballesteros, J. A.; Weinstein, H.; Chen, J. A cluster of aromatic residues in the sixth membrane-spanning segment of the dopamine D2 receptor is accessible in the binding-site crevice. *Biochemistry* **1998**, *37*, 998–1006.
(85) Javitch, J. A.; Fu, D.; Chen, J.; Karlin, A. Mapping the binding site crevice of the dopamine D2 receptor by the substituted-cysteine accessibility method. *Neuron* **1995**, *14*, 825–831.
(86) Mansour, A.; Meng, F.; Meador-Woodruff, J. H.; Taylor, L. P.; Civelli, O.; Akil, H. Site-directed mutagenesis of the human dopamine D2 receptor. *Eur. J. Pharm., Mol. Pharmacol. Sect.* **1992**, *227*, 205–214.
(87) Cox, B. A.; Henningsen, R. A.; Spanoyannis, A.; Neve, R. L.; Neve, K. A. Contributions of conserved serine residues to the interactions of ligands with dopamine D2 receptors. *J. Neurochem.* **1992**, *59*, 627–635.
(88) Taylor, L. P.; Mansour, A.; Akil, H. Hydrophobic residues of the D2 dopamine receptor are important for binding and signal transduction. *J. Neurochem.* **1995**, *65*, 2105–2115.
(89) Coley, C.; Woodward, R.; Johansson, A. M.; Strange, P. G.; Naylor, L. H. Effect of multiple serine/alanine mutations in the transmembrane spanning region V of the D2 dopamine receptor on ligand binding. *J. Neurochem.* **2000**, *74*, 358–366.
(90) Shi, L.; Javitch, J. A. The binding site of aminergic G-protein coupled receptors: the transmembrane segments and second extracellular loop. *Annu. Rev. Pharmacol. Toxicol.* **2002**, *42*, 437–467.
(91) Wang, C.; Gallaher, T.; Shih, J. Site-directed mutagenesis of the serotonin 5-hydroxytrypamine2 receptor: identification of amino acids necessary for ligand binding and receptor activation. *Mol. Pharmacol.* **1993**, *43*, 931–940.
(92) Roth, B. L.; Shoham, M.; Choudhary, M. S.; Khan, N. Identification of conserved aromatic residues essential for agonist binding and second messenger production at 5-hydroxytryptamine2A receptors. *Mol. Pharmacol.* **1997**, *52*, 259–266.
(93) Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **1995**, *8*, 127–134.
(94) Roth, B. L.; Willins, D. L.; Kristiansen, K.; Kroeze, W. K. 5-Hydroxytryptamine2-family receptors (5-hydroxytryptamine2A, 5-hydroxytryptamine2B, 5-hydroxytryptamine2C): where structure meets function. *Pharmacol. Ther.* **1998**, *79*, 231–257.
(95) Shapiro, D. A.; Kristiansen, K.; Kroeze, W. K.; Roth, B. L. Differential modes of agonist binding to 5-hydroxytryptamine2A serotonin receptors revealed by mutation and molecular modeling of conserved residues in transmembrane region 5. *Mol. Pharmacol.* **2000**, *58*, 877–886.
(96) Runyon, S. P.; Mosier, P. D.; Roth, B. L.; Glennon, R. A.; Westkaemper, R. B. Potential modes of interaction of 9-aminomethyl-9,10-dihydroanthracene (AMDA) derivatives with the 5-HT2A receptor: a ligand structure-affinity relationship, receptor mutagenesis and receptor modeling investigation. *J. Med. Chem.* **2008**, *51*, 6808–6828.
(97) Braden, M. R.; Nichols, D. E. Assessment of the roles of serines 5.43(239) and 5.46(242) for binding and potency of agonist ligands at the human serotonin 5-HT2A receptor. *Mol. Pharmacol.* **2007**, *72*, 1200–1209.
(98) Braden, M. R.; Parrish, J. C.; Naylor, J. C.; Nichols, D. E. Molecular interaction of serotonin 5-HT2A receptor residues Phe339(6.51) and Phe340(6.52) with superpotent *N*-benzyl phenethylamine agonists. *Mol. Pharmacol.* **2006**, *70*, 1956–1964.
(99) Choudhary, M.; Sachs, N.; Uluer, A.; Glennon, R.; Westkaemper, R.; Roth, B. Differential ergoline and ergopeptine binding to 5-hydroxytryptamine2A receptors: ergolines require an aromatic residue at position 340 for high affinity binding. *Mol. Pharmacol.* **1995**, *47*, 450–457.

CI900444Q

Chapter 3

Predicting the structure of the dopamine D_3 receptor: An evaluation of virtual screening approaches to GPCR modeling

The dopamine D_3 receptor (D_3R) belongs to the D_2-like family of dopamine receptors. The D_2-like dopamine receptors have been known to be implicated in CNS diseases such as Parkinson's disease and schizophrenia and represent a promising therapeutic target. However, a significant challenge when targeting these receptors, is the development of subtype selective compounds as the orthosteric site in the D_2, D_3 and D_4 receptors have high sequence homology. Hence, a greater understanding of the 3D structure of these receptors could assist in the design of more selective treatments.

In 2010, the high resolution crystal structure of the dopamine D_3 receptor (D_3R) was determined in complex with the small molecule eticlopride. Prior to its release an assessment, GPCR Dock 2010, was run to evaluate the status of the molecular modeling in the GPCR modeling community (the CXCR4 chemokine receptor was also part of this evaluation). Participants were asked to submit up to five models of the D_3R-eticlopride complex and rank them 1 to 5, with 1 being the most likely to represent the crystallized complex.

Participation in this assessment gave us an opportunity to thoroughly evaluate our modeling methods, already discussed in Chapter 2, and expand upon them. This chapter is included as an unpublished article, prepared and formatted for submission to the *Journal of Computer-Aided Molecular Design*. The multiple sequence alignment is in Appendix 2 and the supplementary material for this chapter appears in Appendix 3.

3.1 Declaration

3.1.1 Declaration by candidate

In the case of Chapter 3, I declare that the nature and extent of my contribution to the work was the following:

Nature of contribution	**Contribution (%)**
Development and refinement of homology models, design and running of virtual screening experiments, analysis of results, and manuscript preparation	75

The following co-authors contributed to this work:

Name	**Nature of contribution**	**Contribution (%)***
Kimberley C. McLean	Design of experiments	
Mark Agostino	Data analysis	5
Ian T. Crosby	Co-author of manuscript	
Ben Capuano	Co-author of manuscript	
David K. Chalmers	Design of experiments and co-author of manuscript	
Elizabeth Yuriev	Design of experiments and co -author of manuscript	

** Percentage contribution only shown for co-authors who were students at Monash University at the time of their contribution to this work.*

Candidate's signature: [redacted] Date: 05 / 09/ 2011

3.1.2 Declaration by co-authors

The undersigned hereby certify that:

1. The above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors;
2. The co-authors meet the criteria for authorship, in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
3. The co-authors take public responsibility for their respective part of the publication, except for the responsible author who accepts overall responsibility for the publication;
4. There are no other authors of the publication according to these criteria;
5. Potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
6. The original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

Location of data storage: Department of Medicinal Chemistry and Drug Action, Monash Institute of Pharmaceutical Sciences, 381 Royal Parade, Parkville, Victoria, Australia

Co-author signatures:

Signed: [redacted] Date: 05 / 09/ 2011

Kimberley C. McLean

Signed: [redacted] Date: 05 / 09/ 2011

Mark Agostino

Signed: [redacted] Date: 05 / 09/ 2011

Ian T. Crosby

Signed: [redacted] Date: 05 / 09/ 2011

Ben Capuano

Signed: [redacted] Date: 05 / 09/ 2011

David K. Chalmers

Signed: [redacted] Date: 05 / 09/ 2011

Elizabeth Yuriev

Predicting the structure of the dopamine D_3 receptor: An evaluation of virtual screening approaches to GPCR modeling

Fiona M. McRobb, Kimberley C. McLean, Mark Agostino, Ian T. Crosby, Ben Capuano, Elizabeth Yuriev and David K. Chalmers**

Medicinal Chemistry and Drug Action, Monash Institute of Pharmaceutical Sciences, Monash University (Parkville Campus), 381 Royal Parade, Parkville, VIC 3052 Australia

*To whom correspondence should be addressed. E.Y. Phone: [redacted]. E-mail: [redacted] D.K.C. Phone: [redacted]. E-mail: [redacted]

Abstract

G protein-coupled receptors (GPCRs) are important protein targets in drug discovery. Whilst significant advances are being made in the crystallography of GPCRs, we are still reliant on homology models for structural information about the large majority of these receptors. Herein, we report the development of dopamine D_3 receptor (D_3R) homology models, based on the crystal structure of the β_2-adrenergic receptor in complex with the small molecule eticlopride.

Prior to the release of the D_3R-eticlopride crystal structure (PDB: 3PBL), we prepared and submitted five candidate structures of the D_3R-eticlopride complex to the recent GPCR Dock 2010 assessment. Initial homology models were built using Prime. The receptor binding sites were optimized using Induced Fit Docking, generating 200 candidate

structures, which were evaluated using a number of approaches including virtual screening and visual assessment.

Following the release of the structure, we have evaluated our models against the experimental result. The five candidate structures gave reasonably good predictions of the binding mode of eticlopride (rmsd 2.33 to 3.98 Å) and the D_3R binding site (rmsd 3.2 to 3.6 Å). Only one of the candidate structures, Model 2, produced good enrichment in virtual screening ($EF^{10\%}$ 5.4) and only Model 5 replicated the intramolecular interactions in eticlopride that were observed in the D_3R crystal structure. In the D_3R crystal structure, the placement of a key binding site residue, His 6.55, is maintained by two hydrogen bonding interactions. These interactions are not present in Models 1-5, which influenced the binding mode of eticlopride in these structures.

Based on the analysis of crystal structures of aminergic GPCRs, the placement of several key residues is maintained by a network of hydrogen bonds, which can be broken or altered during flexible receptor docking. Thus, we propose that omitting key residues such as Asp 3.32, Trp 6.48 and His 6.55 from binding site optimization generates models that typically perform well in virtual screening. By using a diverse set of D_3R ligands for flexible receptor docking, we determined that benzamide ligands were the best ligand class to use for binding site optimization in the prediction of the D_3R-eticlopride complex. We have also demonstrated the importance of evaluating the candidate structures with different methods.

This work demonstrates that flexible receptor docking is a useful way to optimize model binding sites for structure-based drug design. We demonstrate that multiple models of a receptor should be considered in structure-based drug design. GPCR Dock 2010 has been an important tool for the assessment of our current techniques, as has the comparison of our D_3R models to the D_3R crystal structure.

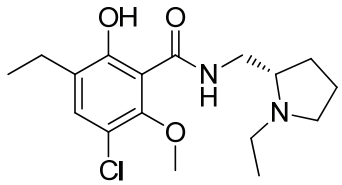
Introduction

The dopamine D_3 receptor (D_3R) is a G protein-coupled receptor (GPCR), which is implicated in many disease states including Parkinson's disease, schizophrenia, depression and drug addiction [1-6]. Until recently, no experimentally determined structures of any dopamine receptors, including the D_3R, were available and our knowledge of the three dimensional structure of the D_3R was derived from homology models [7-13]. Since 2007, advances in crystallization techniques have enabled a number of GPCR crystal structures to be solved [14-29], producing a wealth of structural information about GPCRs and invigorating the field of GPCR modeling by providing new and improved template options [30,31].

The release of new GPCR crystal structures presents an opportunity to evaluate the methods available for modeling GPCRs, as well as the suitability of new structures as templates. In 2008, a community-wide critical assessment of GPCR modeling and docking methods (GPCR Dock 2008) was run by a team from the Scripps Research Institute [32]. Participants were asked to generate models of the adenosine A_{2A} receptor ($A_{2A}R$) before the crystal structure was released [25]. One of the findings of that comparative study was that the accurate prediction of the loops, particularly extracellular loop 2 (ECL2), remained one of the more difficult aspects of GPCR homology modeling.

In anticipation of the release of several new GPCR complexes, a similar assessment was conducted in June-July 2010 (GPCR Dock 2010); "to evaluate the current status and uncover new areas of needed development in GPCR computational biology" [33]. Participants were asked to generate models of the D_3R in complex with eticlopride [28] (Figure 1) and of the CXCR4 chemokine receptor complex with a small molecule antagonist and with a cyclic peptide antagonist [34]. Unsurprisingly, in both GPCR Dock 2008 and 2010, modeling the more flexible regions of the receptors proved to be the most challenging. Additionally, it has also been noted that a significant number of the best performing models

in GPCR Dock 2010 were selected manually, using knowledge of the ligand-receptor interactions, indicating that the current scoring functions require further development [35].



(*S*)-eticlopride

Figure 1: The structure of (*S*)-eticlopride.

We have previously developed a protocol for homology modeling of GPCRs, which incorporates binding site optimization by flexible receptor docking and model evaluation using virtual screening. Using this protocol, we have developed homology models of the dopamine (D_2, D_3 and D_4), serotonin (5-HT_{1B}, 5-HT_{2A}, 5-HT_{2B} and 5-HT_{2C}), histamine (H_1) and muscarinic (M_1) receptors [9], based on the high resolution crystal structure of the β_2-adrenergic receptor (β_2AR) [14]. Of these nine homology models, six showed moderate to good enrichment in virtual screening experiments, including that of the D_3R. Using this protocol, we participated in the GPCR Dock 2010 assessment to predict the D_3R-eticlopride complex. In this work, we have further developed our modeling method to include multiple receptor models, which are evaluated by virtual screening and assessed in respect to site-directed mutagenesis data [36-42]. This study describes the continuing development of our GPCR modeling protocol, our experiences in the GPCR Dock 2010 assessment, and the comparison of our D_3R models submitted for the assessment with the later released crystal structure.

Experimental

Molecular modeling was performed principally using Schrödinger Suite 2010 through the Maestro interface. Default settings were used for all programs, unless stated otherwise. Homology models were built with Prime 2.2 [43], ligand molecules were prepared using LigPrep 2.4 [44]. Docking was carried out with Glide 5.6 [45-48]. Flexible receptor docking

was performed using the Induced Fit Docking (IFD) workflow implemented by Schrödinger [49,50]. GPCR residues are identified using the Ballesteros-Weinstein nomenclature [51], except for loop regions, where the crystal structure numbering is used.

Ligand preparation. A set of 43 D_3R antagonists, the same as used by us previously [9], were chosen as active compounds in flexible receptor docking and virtual screening (Figure S1, Supplementary Material). These structures were prepared using LigPrep, which also assigned formal charges according to physiological pH (pH 7.4). A set of 1,000 drug-like decoy compounds with an average molecular weight of 360 g mol^{-1} was obtained from Schrödinger (http://www.schrodinger.com) [46]. Benzamide and naphthamide compounds, including eticlopride, were docked as the (*S*)-isomers. Ligand properties were previously calculated [9] with QikProp [52] and have been included in the Supplementary Material (Table S1).

Homology modeling. The sequence of the D_3R was obtained from the Center for Membrane Protein Structure Determination [33]. A previously developed multiple sequence alignment was used (Supplementary Material) [9]. Homology models of the dopamine D_3R were built using the β_2AR crystal structure [14] as the template (PDB: 2RH1), including the L119W point mutation [9]. The T4-lysozyme was removed and ICL3 was not modeled. When generating homology models within Prime, the highly conserved residues were anchored (constraints were applied to the sequence alignment). The quality of the models was assessed using MolProbity [53] and the *Protein Report* tool within Maestro.

Flexible receptor docking. Binding site optimization was achieved via flexible receptor docking. Multiple GPCR binding site conformations were generated using either the homology models or, later, the D_3R crystal structure as the input structure. The docking site was defined as a cubic 28 Å box, centered on the centroid of Asp 3.32, Trp 6.48, Phe 6.52, and Tyr 7.43 residues. Selected antagonists – amisulpride, clozapine, eticlopride, haloperidol, melperone, metoclopramide, nafadotride, olanzapine, raclopride, sulpiride,

ziprasidone (Figure 2) – were docked using the default IFD protocol. In some cases residues Asp 3.32, Trp 6.48 or His 6.55 were excluded from the binding site optimization (see below). Either Glide SP or XP (standard or extra precision) was used for docking in the IFD workflow (see below). Up to 20 IFD complexes were collected per ligand. This method generated our "candidate structures".

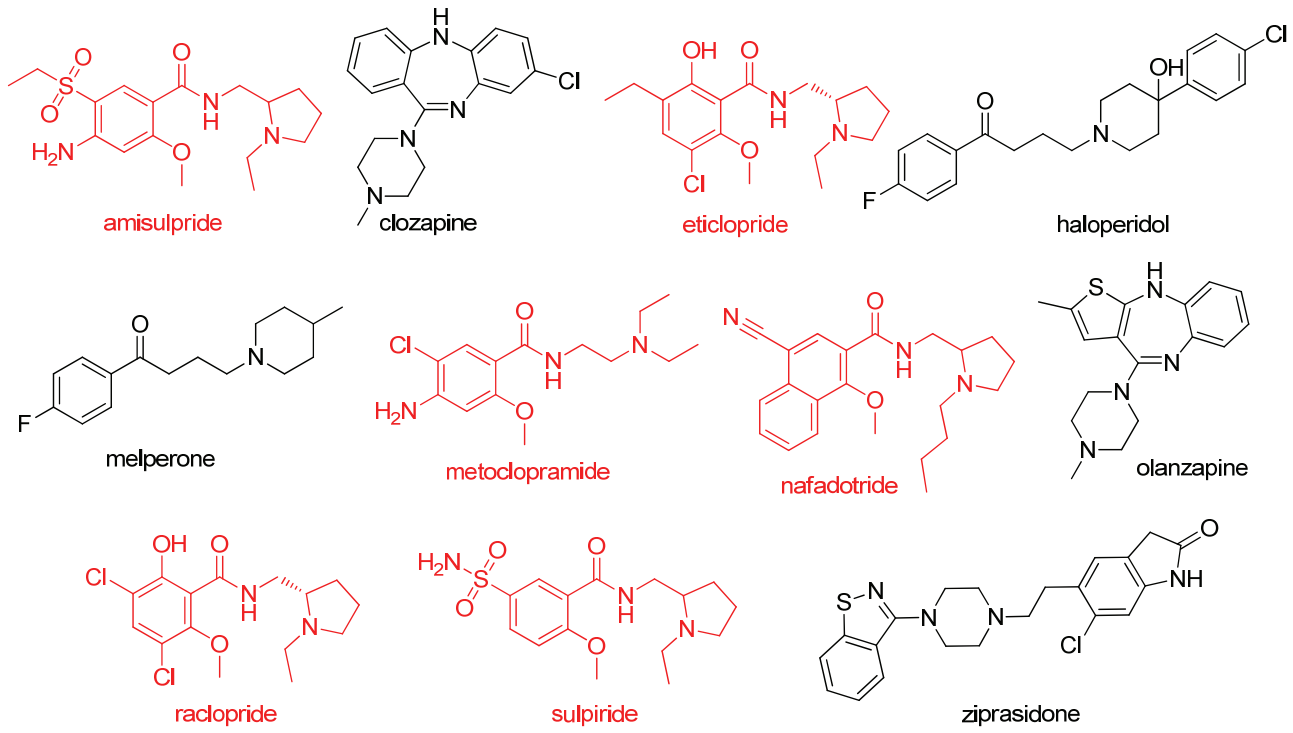


Figure 2: D_3R active compounds docked into each model during flexible receptor docking (compounds highlighted in red are the benzamide and naphthamide compounds used in re-docking experiments).

Virtual screening. The decoy set, enriched with active compounds, was docked into the candidate structures using Glide XP (based on previously developed protocol [9]) and the resulting complexes were ranked by GlideScore. The docking site was defined as a cubic 28 Å box and centered on the centroid of carazolol in complex with β_2AR [14]. One pose per ligand was collected.

Analysis of candidate structures. The quality of models was assessed using enrichment factors, ROC curves, area under the curve (AUC) and GlideScore values. The conformation of key binding site residues such as Asp 3.32 and Trp 6.48 were also considered.

Enrichment factors (EF) were calculated at 2%, 5% and 10% of the total database screened (top 21, 52 and 104 ranked compounds respectively), using the following equation [54]:

$$EF^{x\%} = (Hits_{sampled}/N_{sampled}) \div (Hits_{total}/N_{total})$$

Comparison of homology models to the crystal structure of D_3R. Receptor-ligand complexes were aligned to the crystal structure (3PBL, chain A) using the PyMOL [55] *align* function. Root mean square deviations (rmsd) were calculated using the *Superposition* tool in Maestro. For binding site comparisons, residues within 5 Å of the center of the binding site were considered. The backbone atoms of residues 32-56, 63-92, 100-133, 147-171, 188-216, 322-354, 364-400 (crystal structure numbering) were compared for transmembrane comparison.

Conformational analysis of eticlopride and cognate ligand docking. A conformational analysis of eticlopride was performed using MacroModel [56]. The conformational search employed the systematic torsional sampling method, Systematic Pseudo-Monte Carlo (SPMC), with a maximum of 100,000 steps. Constraints were placed on the hydroxy and methoxy groups of eticlopride (force constant 5 kJ mol^{-1} $Å^{-2}$) to maintain the intramolecular hydrogen bonds. Structures were minimized using Polak–Ribiere conjugate gradient (PRCG) algorithm, with a maximum of 1,000 iterations and a convergence threshold of 0.1 kJ mol^{-1} $Å^{-1}$. Stereochemistry was maintained and the extended torsion sampling method was utilized. The energy window for saving structures was 210 kJ mol^{-1}. The resulting structures were clustered using an rmsd threshold of 0.5 Å. Conformers were docked rigidly into the D_3R crystal structure using Glide SP. The binding site was identified using the centroid of the co-crystallized ligand and the previously established box size.

Results and Discussion

Our two main aims of modeling the D_3R were to evaluate our GPCR modeling protocol by predicting the binding mode of eticlopride in complex with the D_3R, for GPCR Dock 2010, and to generate a model that could be useful for structure-based drug design, such as a

virtual screening campaign. At the time of the study, the only published crystal structures of aminergic GPCRs were those of the adrenergic receptors, turkey β_1AR (PDB: 2VT4 [22]) and human β_2AR (PDB: 2RH1 [14], 2R4R [15], 2R4S [15] and 3D4S [16]). Additionally, crystal structures of the A_{2A}R (PDB: 3EML) [25] and bovine rhodopsin (PDB: 1U19) [57] were available. Based on the multiple sequence alignment [9], two templates, turkey β_1AR and human β_2AR, had higher homology to the D_3R (Table S2, Supplementary Material). Whilst the β_2AR had slightly lower sequence identity than β_1AR (32% versus 36%), the β_2AR crystal structure (PDB: 2RH1) [14] was chosen as the template because of its higher resolution (2.4 Å versus 2.7 Å). Additionally, the use of multiple templates to build homology models was not investigated, as it had been previously shown that using multiple templates (of the five available crystal structures at the time; rhodopsin, squid rhodopsin, β_1AR, β_2AR, A_{2A}AR) did not lead to a significant improvement in homology models [30].

In homology models the binding site is usually in a very similar arrangement to the parent crystal structure and we find that they are biased towards the co-crystallized ligand, in this case carazolol. Carazolol is smaller than many of the D_3R ligands and as a result, we found that many do not fit into the binding site of our initial homology model. To compensate for this bias, we optimized the models by using flexible receptor docking to fit a diverse range of eleven D_3R antagonists (Figure 2) into the binding site. The flexible docking allowed the side chains within the binding site to move, generating multiple binding site conformations. Three different methods were used for binding site optimization: Glide SP without constraints; Glide SP with Asp 3.32 and Trp 6.48 excluded from binding site refinement; and Glide XP. This step generated our 200 “candidate structures”.

Our original intended approach was to select good D_3R homology models from our candidate structures based on good enrichment in virtual screening. The screening step used a database of 1,000 drug-like decoys and 43 known D_3R antagonists, which were docked into each of the 200 candidate structures using Glide XP. Previous virtual screening studies

using the β_2AR crystal structure (2RH1) demonstrated that Glide XP outperformed Glide SP, thus Glide XP was used for virtual screening studies [9]. Ligands were ranked by GlideScore and enrichment factors were calculated at 2, 5 and 10% of the ranked library. AUC values were also calculated from ROC curves (Table S3, Supplementary Material). However, we found this approach to be limiting, as the majority of the models showed marginal enrichment. We therefore needed an alternative approach to assess the models. In the course of this exercise, particularly during the selection of a D_3R-eticlopride complex, we found that other aspects, such as reasonable ligand binding mode and orientation of critical receptor residues, needed to be taken into consideration. Implementation of these considerations is generally reliant on visual inspection of the candidate structures at different stages through the refinement process and on human expertise and intervention. This protocol is described below.

All 200 candidate structures were assessed using the protocol shown in Figure 3. Models that failed at any point were discarded. In Stage 1, residues such as Asp 3.32 and Trp 6.48, which are known to play an important role in ligand binding [58,10,59], were inspected. The side chain of Trp 6.48 was required to be in a conformation similar to that in the crystal structure of the β_2AR and the side chain of Asp 3.32 needed to point towards the binding site. In addition, the majority of active ligands were required to have successfully docked within the orthosteric binding site, including the making the key salt bridge interaction to Asp 3.32.

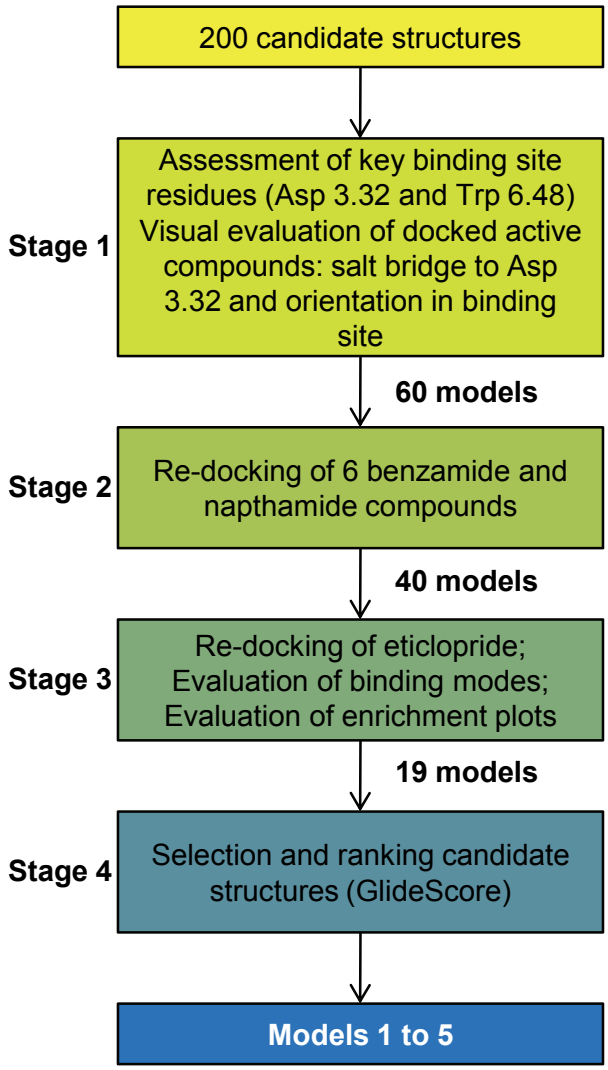


Figure 3: Method for selecting models.

Sixty structures remained after Stage 1, which were evaluated further. Stage 2 involved docking a set of six benzamide and naphthamide ligands (Figure 3) into the receptor models using Glide XP. Candidate structures where active compounds docked such that they made key interactions (i.e. salt bridge interaction to Asp 3.32 and van der Waals (vdW) interactions to residues Val 3.33, Trp 6.48, Phe 6.51 and Phe 6.52) were retained. This resulted in 40 candidate structures. In Stage 3, eticlopride was re-docked into the models using Glide XP. Eticlopride docked making the key interactions in 19 out of the 40 models, producing a set of diverse, yet plausible, binding modes. Based on ligand pose similarity, the number of models was reduced from 19 to five diverse candidate structures (Stage 4). These

final five candidate structures were ranked by GlideScore, which was used as the final basis for the ranking of the models that we submitted to the GPCR Dock 2010 assessment (Table 1).

Table 1: Enrichment factors (2, 5 and 10%), AUC and GlideScore for dopamine D_3R Models 1 to 5.

Model	$EF^{2\%}$	$EF^{5\%}$	$EF^{10\%}$	AUC	GlideScore (kcal mol^{-1})
1	4.73	2.81	1.89	0.743	-10.7
2	9.45	7.02	5.43	0.857	-9.7
3	0	0.94	1.18	0.600	-9.1
4	0	0.47	1.42	0.624	-8.4
5	0	1.4	2.13	0.575	-6.9

Figure 4 shows the bound orientation of eticlopride in Models 1-5 in conjunction with the corresponding enrichment plots. In these models we ensured that the ionic interaction between eticlopride and Asp 3.32 in D_3R was present. In Model 1, which was produced by IFD with eticlopride, a hydrogen bond was observed from the hydroxyl group of eticlopride to His 6.55 (Figure 4a). However, in virtual screening, this model produced only low enrichment (Figure 4b). In Model 2, which was produced by IFD with amisulpride, a hydrogen bond to His 6.55 was also present, but unlike Model 1, this bond was through the amide hydrogen of eticlopride (Figure 4c). Based on the binding modes of compounds in the available crystal structures, we believed that eticlopride was not in an ideal binding mode as it bound closer to ECL2 than the co-crystallized ligands, therefore we did not rank this as the top prediction of the D_3R-eticlopride complex although it gave the best enrichment, with an $EF^{10\%}$ of 5.4 (Figure 4d). In Model 3, which was produced by IFD with eticlopride and including constraints on the positions of Asp 3.32 and Trp 6.48, a hydrogen bond from the hydroxy group of eticlopride to His 6.55 was observed (Figure 4e). This model gave very little enrichment (Figure 4f). In Model 4, which was produced by IFD with eticlopride and

Glide XP, a different orientation of eticlopride in the binding site was seen, with the methoxy group pointing towards the extracellular side of the receptor (Figure 4g). This model gave very little enrichment in virtual screening (Figure 4h). Finally, in Model 5, which was produced by IFD with eticlopride and Glide XP, the binding mode did not have any additional intermolecular hydrogen bonding interactions (Figure 4i). However, in this model, eticlopride contained two intramolecular hydrogen bonds, the first between the hydroxy and the amide carbonyl oxygen and the second between the amide hydrogen and the methoxy group. These intramolecular interactions stabilize eticlopride in a planar conformation (Figure 5). The virtual screening enrichment for this model is also quite low (Figure 4j).

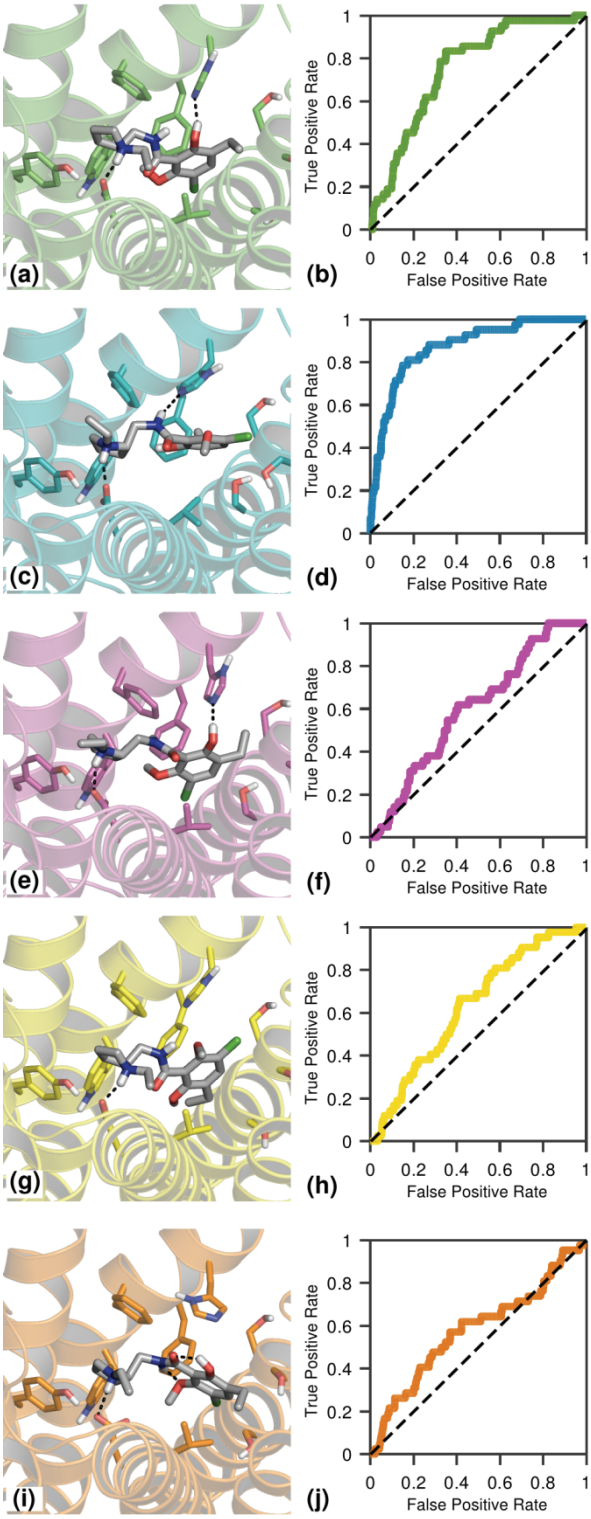


Figure 4: The binding modes of eticlopride (ECL2 is omitted for clarity) and virtual screening ROC curves for Model 1 (a and b), Model 2 (c and d), Model 3 (e and f), Model 4 (g and h) and Model 5 (i and j). The 3D images were created using PyMOL [55].

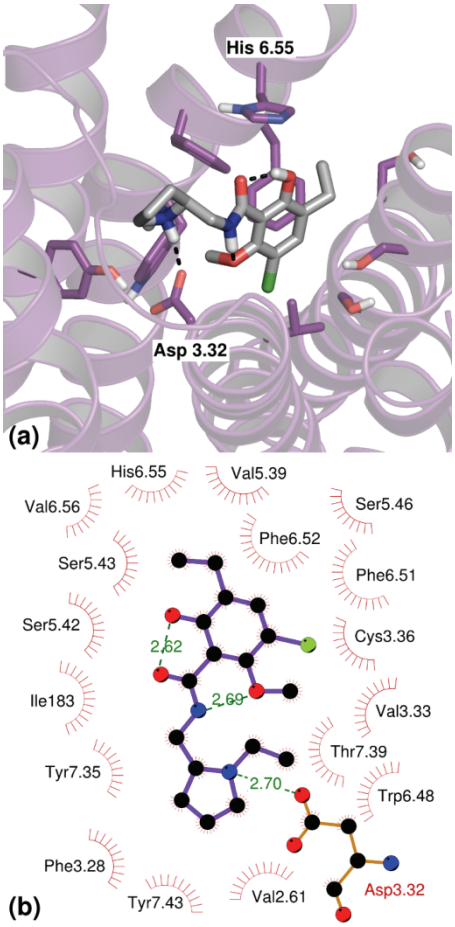


Figure 5: (a) The binding mode of eticlopride in the D_3R crystal structure, displaying polar interactions to the receptor and intramolecular ligand hydrogen bonds (ECL2 is omitted for clarity). The 3D image was created using PyMOL [55]. (b) A 2D schematic plot created with LIGPLOT [60]. Non-bonded interactions: red spokes. Hydrogen bonds: dashed green lines.

Comparison of the homology models with the D_3R crystal structure

The crystal structure of the dopamine D_3R complexed with eticlopride (PDB: 3PBL) [28] was released following the submission of the models to the GPCR Dock 2010. Figure 5a shows the conformation of eticlopride within the crystal structure. The structure confirmed that the ethyl-pyrrolidine moiety of eticlopride makes a salt bridge to Asp 3.32. It also revealed that the substituted benzamide ring binds within a hydrophobic pocket consisting of Val 3.33, Ser 5.42, Ser 5.43, Trp 6.48, Phe 6.51 and Phe 6.52, as well as Ile 183 from ECL2. Additionally, the bound conformation of eticlopride is maintained by two intramolecular

hydrogen bonding interactions (Figure 5b). These hydrogen bonds have also been observed in the crystal structure of eticlopride [61] and in QM calculations [62].

When the homology models were superimposed with the crystal structure, the overall structure of the proteins were seen to be quite similar, with backbone rmsd values for the transmembrane region of 1.36-1.37 Å (Table 2). Relative to the crystal structure, there were some movements in the TM helices of the models near the orthosteric site; TM6 moved the most, with an outward shift, TM5 had a minor outward shift and TM7 had a minor inward shift.

Comparison of the eticlopride binding mode in Models 1 to 5 with the D_3R-eticlopride crystal structure confirmed that the general orientation of eticlopride was correctly predicted. All models also reproduced a large majority (73-88 %) of ligand-protein contacts (Table 2 and Table S4, Supplementary Material). The most noticeable differences between the predicted and experimental structures were the predicted interactions between the ligand and transmembrane helix 4 (TM4) in the candidate structures, which were absent in the crystal structure. The rms deviations of the ligand range from 2.33-3.98 Å (Table 2). Model 3 was the closest to the crystal structure, with an rmsd of 2.33 Å. In three of our models (Models 1, 3 and 5), the orientation of the substituted aromatic ring was accurately predicted, with the hydroxyl group pointing towards the extracellular space. However, we had some difficulty in predicting the orientation of the pyrrolidine-ethyl moiety. Additionally, only Model 5 reproduced the two intramolecular hydrogen bonds in eticlopride present in the D_3R-eticlopride crystal structure.

Table 2: The prediction of close contacts (refer also to Table S4, Supplementary Material) and rms deviations between the top 5 models and the crystal structure 3PBL.

Model	% binding site contacts	rmsd (Å) to the crystal structure		
		ligand	binding site residues	transmembrane Cα
1	73	2.94	3.63	1.37
2	81	3.52	3.31	1.37
3	88	2.34	3.36	1.37
4	73	4.01	3.27	1.36
5	73	3.50	3.23	1.37

The difference between the orientations of binding site residues in the crystal structure and models were assessed by rmsd and ranged from 3.2 to 3.6 Å (Table 2). One of the key disparities between the candidate structures and the D_3R crystal structure was the placement of His 6.55, a residue which has been highlighted as significant by site-directed mutagenesis [10]. When a hydrogen bonding interaction between the ligand and His 6.55 was present, more favorable ligand poses and GlideScores often resulted. Prior to seeing the D_3R crystal structure, we interpreted this data as indicating that hydrogen bonding between the ligand and His 6.55 was favorable. However, the crystal structure shows that His 6.55 is stabilized in its conformation by hydrogen bonding to Tyr 7.35 and Ile 183 in ECL2 and, as a result, it can not make hydrogen bonds to a ligand in the binding site (Figure S2, Supplementary Material). This finding reinforces that site-directed mutagenesis data must be interpreted with care [63]. Additionally, because Ile 183 lies within ECL2, the positioning of the loop is paramount to predicting the His 6.55-Ile 183 interaction. The hydrogen bonding interaction between His 6.55 and Ile 183 was not observed in the candidate structures because ECL2 was too far away from His 6.55 (6-8 Å) to make a hydrogen bonding interaction (refer Figure S2, Supplementary Material). We believe that the difficulties in predicting the position of eticlopride in the binding site were, in part, due to the placement of the key

histidine residue on TM6 (His 6.55), which in the crystal structure is stabilized by hydrogen bonding to Tyr 7.35 and Ile 183 in ECL2.

Evaluation of Induced Fit Docking for binding site optimization

Our binding site optimization procedure uses Induced Fit Docking to generate candidate protein structures. In modeling the D_3R-eticlopride complex we initially allowed all residues within 5 Å of the center of the binding site to move. As GPCRs are inherently flexible, it is conceivable that even key binding site residues (e.g. Asp 3.32) may move to accommodate ligands of different shapes and sizes. However, we found that flexible receptor docking often moved key residues to unrealistic conformations; for example with Asp 3.32 no longer pointing into the orthosteric binding site. Analysis of the aminergic GPCR crystal structures [14-16,22], showed that in each case Asp 3.32 is stabilized by a hydrogen bonding interaction to Tyr 7.43 and thus it would be reasonable to expect that these residues remain in a relatively similar conformation in similar receptors. Therefore, in later work we omitted residue Asp 3.32 from binding site optimization. Prior to the release of the D_3R crystal structure, we also evaluated the conformations of the key binding site residue Trp 6.48 in the available crystal structures [14-16,22,25]. This residue was consistently found with χ_1 in a *gauche*+ conformation. Thus, we either omitted this residue from binding site optimization or ensured that it was in a similar conformation to the β_2AR template. By omitting Asp 3.32 and Trp 6.48 residues from binding site optimization, we ensured that the key residues were retained in conformations consistent with available crystal structures and for Asp 3.32 that the protein hydrogen bonding network was maintained. Crystal structures released since the D_3R crystal structure [17,19,20,23,28], also maintain similar conformations of Asp 3.32 and Trp 6.48, further supporting that omission of these residues from binding site optimization was a sound decision.

In contrast to our previously reported protocol, where only a single ligand was used for IFD [9], in this work we used eleven diverse D_3R antagonists in the flexible receptor

docking step, enabling us to determine the best class of molecule to use for binding site optimization of the D_3R-eticlopride complex. Quite reasonably, we found that two benzamide ligands, amisulpride and eticlopride, were the most useful in the prediction of the D_3R-eticlopride crystal structure. That is, all top five models resulted from using amisulpride or eticlopride for binding site refinement. However, it is interesting to note that the models generated using eticlopride itself for Induced Fit Docking (Models 1, 3, 4, and 5) showed marginal enrichment in virtual screening, which may limit their utility in structure-based drug design. In one instance (Model 2), amisulpride was used to generate a model that both gave good virtual screening results and a reasonable binding mode of eticlopride.

Virtual screening using the D_3R crystal structure

Once the D_3R crystal structure became available, we also evaluated our modeling methods using this structure. Firstly, we used cognate docking to assess the ability of Glide XP to reproduce the binding mode of eticlopride. The virtual screening protocol was evaluated by docking our library of active and decoy compounds. Enrichment factors and AUC values were compared with those obtained for Models 1 to 5, to assess if the crystal structure would be useful for structure-based drug design without any optimization. Finally, the binding site optimization protocol was evaluated using the D_3R crystal structure, using flexible receptor docking followed by virtual screening.

Cognate docking of eticlopride into D_3R crystal structure using Glide XP gave a marginally better ligand rmsd (1.80 Å) than Models 1 to 5 (2.33 to 3.98 Å). In the D_3R crystal structure eticlopride contains two internal hydrogen bonds, however, our docking procedure struggled to reproduce these interactions, particularly in the candidate structures (only Model 5 reproduced these interactions). As the 1.80 Å rmsd for cognate docking was relatively high, we wanted to assess if constraining the intramolecular hydrogen bonds would produce better cognate docking results. Accordingly we used a conformational search to generate a set of 480 eticlopride conformers, all containing the two internal hydrogen

bonds. The generated conformations were docked into the D_3R crystal structure using rigid ligand docking in Glide SP. The top ranked structure had an rmsd of 0.45 Å (Figure S3, Supplementary Material) suggesting that in this case Glide 5.6 does not correctly score internal hydrogen bonds.

Virtual screening was carried out using the established protocol, to assess if this method could identify active compounds from a database of decoys, using the D_3R crystal structure. However, like most of the candidate structures, virtual screening using the D_3R crystal structure produced marginal enrichment ($EF^{10\%}$ 1.89, Figure 6a, Table 3).

Based on the poor performance of the D_3R crystal structure in virtual screening, we wanted to investigate if the binding site optimization protocol developed for the candidate structures would improve the virtual screening results. Binding site optimization was carried out on the D_3R crystal structure using flexible receptor docking and based on our previous experience, Asp 3.32 and Trp 6.48 were omitted from binding site optimization. In some cases, His 6.55 was also omitted from flexible receptor docking, as this residue had influenced our models for GPCR Dock 2010. This step generated 54 candidate structures that were then submitted to virtual screening evaluation.

When residues Asp 3.32 and Trp 6.48 were omitted from binding site optimization, the best model in virtual screening (in terms of enrichment factors) was generated using olanzapine as the IFD ligand. This structure (3PBL_1) gave an $EF^{10\%}$ of 5.20, compared to the value of 1.89 for the crystal structure (Figure 6b, Table 3). When residues Asp 3.32, Trp 6.48 and His 6.55 were omitted from the binding site optimization, an improvement was observed in the early enrichment of the ranked database, with the best $EF^{2\%}$ of 8.27 for olanzapine (3PBL_2, Figure 6c, Table 3). Again, this model was a significant improvement upon the D_3R crystal structure. When the binding site optimized models were compared with the initial crystal structure, small movements were noted in residues Ile 183 and Phe 6.52 in both 3PBL_1 and 3PBL_2, as well as residue Cys 3.36 in 3PBL_1. These subtle

changes created a slightly larger binding site, thus allowing more of the active compounds to be identified earlier in virtual screening.

The poor performance of virtual screening using the crystal structure is likely to be due to the small size of the co-crystallized ligand in relation the majority of the D_3R antagonists. Thus, optimization of the binding site with a bulkier ligand allows for more of the active compounds to be appropriately accommodated in the binding site cavity.

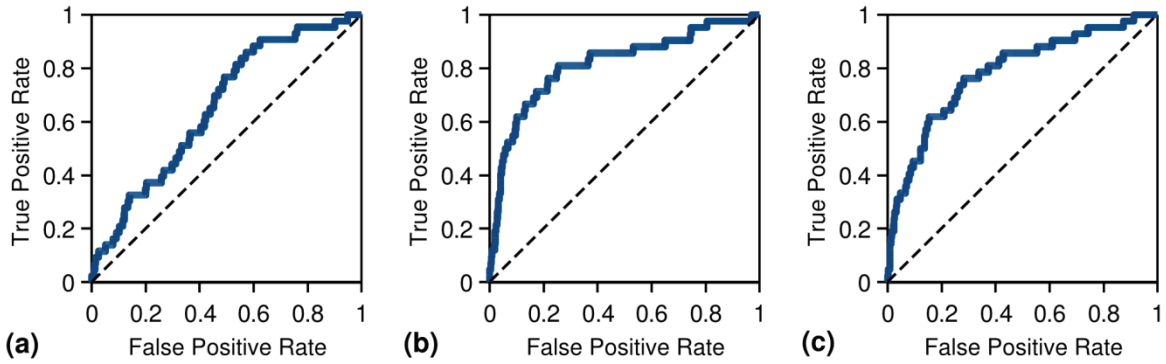


Figure 6: Enrichment plots for virtual screening into (a) the D_3R crystal structure, and the binding site-optimized D_3R crystal structures (b) 3PBL_1 (IFD fixing residues Asp 3.32, Trp 6.48) and (c) 3PBL_2 (IFD fixing residues Asp 3.32, Trp 6.48 and His 6.55).

Table 3: Enrichment factors (2, 5 and 10%) and AUC from virtual screening into the D_3R crystal structure. Models 3PBL_1 and 3PBL_2 were first prepared by binding site optimization using the D_3R crystal structure, fixing residues Asp 3.32, Trp 6.48 (residue His 6.55 was also fixed in 3PBL_2).

Model	**$EF^{2\%}$**	**$EF^{5\%}$**	**$EF^{10\%}$**	**AUC**
3PBL	4.73	2.34	1.89	0.65
3PBL_1	5.91	6.55	5.20	0.81
3PBL_2	8.27	6.09	4.25	0.78

The most accurate predictions of the D_3R-eticlopride complex during the GPCR Dock 2010 assessment were obtained by using a combination of techniques, particularly by including pharmacophore information, flexible receptor docking (automated or manual) and virtual screening evaluation [35]. The GPCR Dock assessment has reinforced the need to

use multiple techniques in GPCR modeling, especially taking into consideration the induced fit nature of ligand binding, as we have done with our binding site optimization protocol, as well as the placement of ECL2.

Conclusions

While significant advances are being made in the crystallography of GPCRs and the number of crystal structure are increasing, for the large majority of GPCRs we must still rely on homology models for structure-based drug design. Importantly, new GPCR crystal structures give us additional templates for homology modeling and also give us the opportunity to test our modeling methods. This work was undertaken to evaluate and improve approaches to homology modeling of GPCRs in general and specifically to generate models of D_3R that are useful for drug design.

Our modeling of the D_3R-eticlopride complex was completed in two stages. The initial studies were performed prior to the release of the D_3R crystal structure. Using our binding site optimization protocol, we developed five models that predicted the binding mode of eticlopride reasonably well, with ligand rmsd values between 2.33 and 3.98 Å. We found that the ligand used for Induced Fit Docking strongly influences the binding site that the closely related benzamides give the best results for the D_3R-eticlopride complex. This suggests that, in general, the 'known active' compounds used in virtual screening or Induced Fit Docking studies should be structurally similar to the compounds of interest. Additionally, we found that care needs to be taken when optimizing the receptor binding site that key residues are not distorted too far. We found that, for the D_3R, omitting key residues such as Asp 3.32, Trp 6.48 and His 6.55 from binding site optimization generates models that typically perform well in virtual screening.

Our ability to predict the binding mode of eticlopride was affected by two main factors; poor prediction of the docked eticlopride geometry and by difficulties in the placement of His 6.55, a key residue within the binding site. A lack of optimization of ECL2 and

misinterpretation of site-directed mutagenesis data ultimately influenced our ability to predict the D_3R-eticlopride complex. Although the solutions to all of these shortcomings are not straightforward, evaluation of the predicted ligand conformation using an independent method (e.g. using quantum mechanics calculations) should have identified the poor ligand geometry. We will implement additional checking in future work.

In the second part of this work we used flexible receptor docking to optimize the D_3R crystal structure, producing models that give good enrichment and will be useful for virtual screening studies.

GPCR Dock 2010 has been a valuable tool for the assessment of our current techniques. In this current study, as well as our previous work, we have developed a method that produces models that perform well in virtual screening and is applicable not only to optimizing the binding site of GPCRs, but other proteins with similar ligand induced fit effects.

Acknowledgements

F.M.M. and M.A. are recipients of an Australian Postgraduate Award (APA) scholarship. This work utilized computational resources and technical assistance provided by the Victorian Partnership for Advanced Computing (VPAC) and the Victorian Life Sciences Computation Initiative (VLSCI). K.C.M. was a VLSCI-sponsored student in the Undergraduate Research Opportunities Program. The authors thank the organizers of the Community wide GPCR Dock 2010 assessment.

Supplementary Material

Sequence identity for the D_3R. Virtual screening results for all 200 candidate homology models. A table of protein-ligand interactions for Models 1 to 5. A list of the D_3R active compounds used in docking. Figures displaying the placement of residue His 6.55 and

cognate rigid ligand docking of eticlopride. Multiple sequence alignment and PDB files for Models 1 to 5, 3PBL_1 and 3PBL_2.

References

1. Heidbreder CA, Newman AH (2010) Current perspectives on selective dopamine D_3 receptor antagonists as pharmacotherapeutics for addictions and related disorders. Ann N Y Acad Sci 1187:4-34.
2. Boeckler F, Gmeiner P (2007) Dopamine D3 receptor ligands - Recent advances in the control of subtype selectivity and intrinsic activity. Biochim Biophys Acta, Biomembr 1768:871-887.
3. Joyce JN (2001) Dopamine D3 receptor as a therapeutic target for antipsychotic and antiparkinsonian drugs. Pharmacol Ther 90:231-259.
4. Sokoloff P, Diaz J, Foll BL, Guillin O, Leriche L, Bezard E, Gross C (2006) The dopamine D_3 receptor: A therapeutic target for the treatment of neuropsychiatric disorders. CNS Neurol Disord: Drug Targets 5:25-43.
5. Pilla M, Perachon S, Sautel F, Garrido F, Mann A, Wermuth CG, Schwartz J-C, Everitt BJ, Sokoloff P (1999) Selective inhibition of cocaine-seeking behaviour by a partial dopamine D_3 receptor agonist. Nature 400:371-375.
6. Joyce JN, Millan MJ (2007) Dopamine D3 receptor agonists for protection and repair in Parkinson's disease. Curr Opin Pharmacol 7:100-105.
7. López L, Selent J, Ortega R, Masaguer CF, Domínguez E, Areias F, Brea J, Loza MI, Sanz F, Pastor M (2010) Synthesis, 3D-QSAR, and structural modeling of benzolactam derivatives with binding affinity for the D_2 and D_3 receptors. ChemMedChem 5:1300-1317.
8. Zhao Y, Lu X, Yang C-y, Huang Z, Fu W, Hou T, Zhang J (2010) Computational modeling toward understanding agonist binding on dopamine 3. J Chem Inf Model 50:1633-1643.
9. McRobb FM, Capuano B, Crosby IT, Chalmers DK, Yuriev E (2010) Homology modeling and docking evaluation of aminergic G protein-coupled receptors. J Chem Inf Model 50:626-637.
10. Ehrlich K, Gotz A, Bollinger S, Tschammer N, Bettinetti L, Harterich S, Hubner H, Lanig H, Gmeiner P (2009) Dopamine D_2, D_3, and D_4 selective phenylpiperazines as molecular probes to explore the origins of subtype specific receptor binding. J Med Chem 52:4923-4935.
11. Selent J, López L, Sanz F, Pastor M (2008) Multi-receptor binding profile of clozapine and olanzapine: A structural study based on the new β_2 adrenergic receptor template. ChemMedChem 3:1194-1198.

12. Kortagere S, Cheng S-Y, Antonio T, Zhen J, Reith MEA, Dutta AK (2011) Interaction of novel hybrid compounds with the D3 dopamine receptor: Site-directed mutagenesis and homology modeling studies. Biochem Pharmacol 81:157-163.

13. Wang Q, Mach RH, Luedtke RR, Reichert DE (2010) Subtype selectivity of dopamine receptor ligands: Insights from structure and ligand-based methods. J Chem Inf Model 50:1970-1985.

14. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SGF, Thian FS, Kobilka TS, Choi H-J, Kuhn P, Weis WI, Kobilka BK, Stevens RC (2007) High-resolution crystal structure of an engineered human β_2-adrenergic G protein coupled receptor. Science 318:1258-1265.

15. Rasmussen SGF, Choi H-J, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VRP, Sanishvili R, Fischetti RF, Schertler GFX, Weis WI, Kobilka BK (2007) Crystal structure of the human β2 adrenergic G-protein-coupled receptor. Nature 450:383-387.

16. Hanson MA, Cherezov V, Griffith MT, Roth CB, Jaakola V-P, Chien EYT, Velasquez J, Kuhn P, Stevens RC (2008) A specific cholesterol binding site Is established by the 2.8 Å structure of the human β_2-adrenergic receptor. Structure 16:897-905.

17. Wacker D, Fenalti G, Brown MA, Katritch V, Abagyan R, Cherezov V, Stevens RC (2010) Conserved binding mode of human β_2 adrenergic receptor inverse agonists and antagonist revealed by X-ray crystallography. J Am Chem Soc 132:11443-11445.

18. Bokoch MP, Zou Y, Rasmussen SGF, Liu CW, Nygaard R, Rosenbaum DM, Fung JJ, Choi H-J, Thian FS, Kobilka TS, Puglisi JD, Weis WI, Pardo L, Prosser RS, Mueller L, Kobilka BK (2010) Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. Nature 463:108-112.

19. Rasmussen SGF, Choi H-J, Fung JJ, Pardon E, Casarosa P, Chae PS, DeVree BT, Rosenbaum DM, Thian FS, Kobilka TS, Schnapp A, Konetzki I, Sunahara RK, Gellman SH, Pautsch A, Steyaert J, Weis WI, Kobilka BK (2011) Structure of a nanobody-stabilized active state of the β_2 adrenoceptor. Nature 469:175-180.

20. Rosenbaum DM, Zhang C, Lyons JA, Holl R, Aragao D, Arlow DH, Rasmussen SGF, Choi H-J, DeVree BT, Sunahara RK, Chae PS, Gellman SH, Dror RO, Shaw DE, Weis WI, Caffrey M, Gmeiner P, Kobilka BK (2011) Structure and function of an irreversible agonist-β_2 adrenoceptor complex. Nature 469:236-240.

21. Rasmussen SGF, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, Thian FS, Chae PS, Pardon E, Calinski D, Mathiesen JM, Shah STA, Lyons JA, Caffrey M, Gellman SH, Steyaert J, Skiniotis G, Weis WI, Sunahara RK, Kobilka BK (2011) Crystal structure of the β_2 adrenergic receptor-Gs protein complex. Nature advance online publication:DOI: 10.1038/nature10361.

22. Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, Henderson R, Leslie AGW, Tate CG, Schertler GFX (2008) Structure of a β_1-adrenergic G-protein-coupled receptor. Nature 454:486-491.

23. Warne T, Moukhametzianov R, Baker JG, Nehme R, Edwards PC, Leslie AGW, Schertler GFX, Tate CG (2011) The structural basis for agonist and partial agonist action on a β_1-adrenergic receptor. Nature 469:241-244.

24. Moukhametzianov R, Warne T, Edwards PC, Serrano-Vega MJ, Leslie AGW, Tate CG, Schertler GFX (2011) Two distinct conformations of helix 6 observed in antagonist-bound structures of a β_1-adrenergic receptor. Proc Natl Acad Sci U S A 108:8228-8232.

25. Jaakola V-P, Griffith MT, Hanson MA, Cherezov V, Chien EYT, Lane JR, Ijzerman AP, Stevens RC (2008) The 2.6 angstrom crystal structure of a human A_{2A} adenosine receptor bound to an antagonist. Science 322:1211-1217.

26. Xu F, Wu H, Katritch V, Han GW, Jacobson KA, Gao Z-G, Cherezov V, Stevens RC (2011) Structure of an agonist-bound human A_{2A} adenosine receptor. Science 332:322-327.

27. Lebon G, Warne T, Edwards PC, Bennett K, Langmead CJ, Leslie AGW, Tate CG (2011) Agonist-bound adenosine A_{2A} receptor structures reveal common features of GPCR activation. Nature 474:521-525.

28. Chien EYT, Liu W, Zhao Q, Katritch V, Won Han G, Hanson MA, Shi L, Newman AH, Javitch JA, Cherezov V, Stevens RC (2010) Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. Science 330:1091-1095.

29. Shimamura T, Shiroishi M, Weyand S, Tsujimoto H, Winter G, Katritch V, Abagyan R, Cherezov V, Liu W, Han GW, Kobayashi T, Stevens RC, Iwata S (2011) Structure of the human histamine H_1 receptor complex with doxepin. Nature 475:65-70.

30. Mobarec JC, Sanchez R, Filizola M (2009) Modern homology modeling of G-protein coupled receptors: Which structural template to use? J Med Chem 52:5207-5216.

31. Yarnitzky T, Levit A, Niv M (2010) Homology modeling of G-protein-coupled receptors with X-ray structures on the rise. Curr Opin Drug Discovery Dev 13:317-325.

32. Michino M, Abola E, Brooks CL, Dixon JS, Moult J, Stevens RC (2009) Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. Nat Rev Drug Discovery 8:455-463.

33. GPCR DOCK 2010. http://gpcr.scripps.edu/GPCRDock2010/index.html.

34. Wu B, Chien EYT, Mol CD, Fenalti G, Liu W, Katritch V, Abagyan R, Brooun A, Wells P, Bi FC, Hamel DJ, Kuhn P, Handel TM, Cherezov V, Stevens RC (2010) Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. Science 330:1066-1071.

35. Kufareva I, Rueda M, Katritch V, Stevens RC, Abagyan R (2011) Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. Structure 19:1108-1126.

36. Daniell SJ, Strange PG, Naylor LH (1994) Site-directed mutagenesis of Tyr417 in the rat D_2 dopamine receptor. Biochem Soc Trans 22:144S.

37. Sartania N, Strange PG (1999) Role of conserved serine residues in the interaction of agonists with D_3 dopamine receptors. J Neurochem 72:2621-2624.

38. Alberts GL, Pregenzer JF, Bin Im W (1998) Contributions of cysteine 114 of the human D3 dopamine receptor to ligand binding and sensitivity to external oxidizing agents. Br J Pharmacol 125:705-710.

39. Cho W, Taylor LP, Mansour A, Akil H (1995) Hydrophobic residues of the D_2 dopamine receptor are important for binding and signal transduction. J Neurochem 65:2105-2115.

40. Lundstrom K, Turpin MP, Large C, Robertson G, Thomas P, Lewell XQ (1998) Mapping of dopamine D_3 receptor binding site by pharmacological characterization of mutants expressed in CHO cells with the semliki forest virus system. J Recept Signal Transduction 18:133-150.

41. Mansour A, Meng F, Meador-Woodruff JH, Taylor LP, Civelli O, Akil H (1992) Site-directed mutagenesis of the human dopamine D_2 receptor. Eur J Pharmacol, Mol Pharmacol Sect 227:205-214.

42. Javitch JA, Fu D, Chen J, Karlin A (1995) Mapping the binding site crevice of the dopamine D_2 receptor by the substituted-cysteine accessibility method. Neuron 14:825-831.

43. Prime (2010). version 2.2 edn. Schrödinger, LLC, New York, NY

44. LigPrep (2010). version 2.4 edn. Schrödinger, LLC, New York, NY

45. Glide (2010). version 5.6 edn. Schrödinger, LLC, New York, NY

46. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 47:1739-1749.

47. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. J Med Chem 47:1750-1759.

48. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT (2006) Extra Precision Glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. J Med Chem 49:6177-6196.

49. Schrödinger Suite 2010 Induced Fit Docking protocol; . Glide, version 5.6; Schrödinger, LLC: New York, NY, 2010, Prime, version 2.2; Schrödinger, LLC: New York, NY, 2010

50. Sherman W, Day T, Jacobson MP, Friesner RA, Farid R (2006) Novel procedure for modeling ligand/receptor induced fit effects. J Med Chem 49:534-553.

51. Ballesteros JA, Weinstein H, Stuart CS (1995) Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in

G protein-coupled receptors. In: Methods in Neurosciences, vol 25. Academic Press, pp 366-428

52. QikProp (2008). version 3.1 edn. Schrödinger, LLC, New York, NY

53. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, III, Snoeyink J, Richardson JS, Richardson DC (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res 35:W375-383.

54. Pearlman DA, Charifson PS (2001) Improved scoring of ligand-protein interactions using OWFEG free energy grids. J Med Chem 44:502-511.

55. DeLano WL (2002) The PyMOL molecular graphics system. DeLano Scientific, Palo Alto, CA, USA

56. MacroModel (2010). version 9.8 edn. Schrödinger, LLC, New York, NY

57. Okada T, Sugihara M, Bondar A-N, Elstner M, Entel P, Buss V (2004) The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. J Mol Biol 342:571-583.

58. Dorfler M, Tschammer N, Hamperl K, Hubner H, Gmeiner P (2008) Novel D3 selective dopaminergics incorporating enyne units as nonaromatic catechol bioisosteres: Synthesis, bioactivity, and mutagenesis studies. J Med Chem 51:6829-6838.

59. Javitch JA, Ballesteros JA, Weinstein H, Chen J (1998) A cluster of aromatic residues in the sixth membrane-spanning segment of the dopamine D2 receptor is accessible in the binding-site crevice. Biochemistry 37:998-1006.

60. Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. Protein Eng 8:127-134.

61. Waagner A, Stensland B, Csoeregh I, De PT (1985) Molecular structure and absolute configuration of the hydrochloride of a novel dopamine receptor antagonist: 2*S*-(-)-5-chloro-3-ethyl-*N*-[(1-ethyl-2-pyrrolidinyl)methyl]-6-methoxysalicylamide. Acta Pharm Suec 22:101-110.

62. Saran A, Coutinho E (1994) Quantum mechanical calculations on dopamine D2-receptor antagonics: Conformation of remoxipride, eticlopride and NCQ115. Proc Indian Acad Sci, Chem Sci 106:149-161.

63. Katritch V, Rueda M, Lam PC-H, Yeager M, Abagyan R (2010) GPCR 3D homology models for ligand screening: Lessons learned from blind predictions of adenosine A2a receptor complex. Proteins 78:197-211.

Chapter 4

Homobivalent ligands of the atypical antipsychotic clozapine

Clozapine (**1**) is considered to be the leading atypical antipsychotic, however, its use is limited due to severe side effects, including a potentially fatal blood disorder, agranulocytosis. Clozapine is unparalleled in the management of treatment-resistant schizophrenia (patients who failed to respond to typical antipsychotic agents) and remains one of the best atypical antipsychotics.

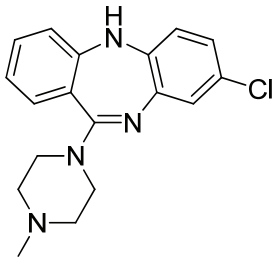


Figure 4.1 Structure of clozapine (**1**).

Clozapine exerts its therapeutic effect mostly by antagonism of central dopaminergic and serotonergic receptor systems belonging to the superfamily of GPCRs. There is increasing evidence that GPCRs act as dimers or higher order oligomers. GPCR dimers may represent a potentially novel pharmacological target, with unique signaling properties and could lead to more potent and selective compounds.

Bivalent ligands are compounds that consist of two pharmacophores covalently tethered by a spacer, and have been designed to improve the potency and selectivity of the original pharmacophore. More recently bivalent ligands have been used as tools to explore the concept of GPCR dimerization.

In the first section of Chapter 4, the design, synthesis and pharmacological evaluation of a series of bivalent ligands of clozapine is discussed. This is included as an unpublished

journal article, prepared and formatted for submission to the *Journal of Medicinal Chemistry*.

References for Section 4.2, the design, synthesis and pharmacological evaluation of homobivalent ligands of clozapine, appear directly following this manuscript. The compound numbering used throughout Chapter 4 is consistent with the compound numbering used in the manuscript in Section 4.2. Supporting information for Section 4.2 is located in Appendix 4.

Additionally, in Section 4.3, the modeling methods developed in Chapters 2 and 3 are used to build a homology model of the D_2R using the D_3R crystal structure as a template. This D_2R homology model was used to generate a number of models of the D_2R homodimer, which was then subjected to molecular dynamics simulations in a solvated phospholipid bilayer. The results obtained from this study were compared to the pharmacological results for the bivalent ligands developed in Section 4.2. References for Section 4.3 appear at the end of this chapter.

4.1 Declaration

4.1.1 Declaration by candidate

In the case of Chapter 4, I declare that the nature and extent of my contribution to the work was the following:

Nature of contribution	**Contribution (%)**
Synthesis, purification and analysis of all compounds and co-author of manuscript	60

The following co-authors contributed to this work:

Name	**Nature of contribution**	**Contribution (%)***
Elizabeth Yuriev	Co-author of manuscript	
Ian T. Crosby	Co-author of manuscript	
Arthur Christopoulos	Design of pharmacological experiments, data analysis and co-author of manuscript	
J. Robert Lane	Design of and conduct of pharmacological experiments, data analysis and co-author of manuscript	
Ben Capuano	Project design and co-author of manuscript	

** Percentage contribution only shown for co-authors who were students at Monash University at the time of their contribution to this work.*

Candidate's signature: [redacted] Date: 05 / 09/ 2011

4.1.2 Declaration by co-authors

The undersigned hereby certify that:

1. The above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors;
2. The co-authors meet the criteria for authorship, in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
3. The co-authors take public responsibility for their respective part of the publication, except for the responsible author who accepts overall responsibility for the publication;
4. There are no other authors of the publication according to these criteria;
5. Potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
6. The original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

Location of data storage: Department of Medicinal Chemistry and Drug Action, Monash Institute of Pharmaceutical Sciences, 381 Royal Parade, Parkville, Victoria, Australia

Co-author signatures:

Signed: [redacted] Date: 05 / 09/ 2011

Ian T. Crosby

Signed: [redacted] Date: 05 / 09/ 2011

Elizabeth Yuriev

Signed: [redacted] Date: 05 / 09/ 2011

Arthur Christopoulos

Signed: [redacted] Date: 05 / 09/ 2011

J. Robert Lane

Signed: [redacted] Date: 05 / 09/ 2011

Ben Capuano

4.2 Prepared manuscript

Homobivalent ligands of the atypical antipsychotic clozapine: Design, synthesis and pharmacological evaluation

Fiona M. McRobb,[†] Ian T. Crosby,[†] Elizabeth Yuriev,[†] Arthur Christopoulos,[‡] J. Robert Lane,[‡] and Ben Capuano[†]**

[†]Medicinal Chemistry and Drug Action, Monash Institute of Pharmaceutical Sciences, Monash University (Parkville Campus), 381 Royal Parade, Parkville, VIC 3052 Australia

[‡]Drug Discovery Biology, Monash Institute of Pharmaceutical Sciences, Monash University (Parkville campus), 381 Royal Parade, Parkville, VIC 3052 Australia

RECEIVED DATE

*To whom correspondence should be addressed.

B.C. Phone: [redacted] E-mail: [redacted]

J.R.L. Phone: [redacted] E-mail: [redacted]

Abstract

To date, all typical and atypical antipsychotics target the dopamine D_2 receptor. Clozapine represents the prototypical atypical antipsychotic, although it displays only moderate (sub-micromolar) affinity for the dopamine D_2 receptor. Herein, we present the design, synthesis and pharmacological evaluation of three series of homobivalent ligands of clozapine, differing in the length and nature of the spacer and the point of attachment to the pharmacophore. Attachment of the spacer at the N4′ position of clozapine yielded a series of homobivalent ligands that displayed the most promising affinity and activity for the dopamine D_2 receptor. A spacer length-dependent relationship with affinity or inhibitory potency was observed in both radioligand binding and functional studies. The 16 and 18 atom spacer bivalent ligands were the most active compounds, displaying low nanomolar affinity (1.41 and 1.35 nM) and a significant gain in affinity (75- and 79-fold, respectively) relative to the original pharmacophore, clozapine.

Introduction

Bivalent ligands are compounds that consist of two pharmacophores covalently tethered by an appropriate spacer.[1-4] A linking group joining the pharmacophore to the spacer can also be incorporated.[3,4] There are two general classes of bivalent ligands; homobivalent ligands, containing two identical pharmacophores and heterobivalent ligands, wherein the two pharmacophores are different. Most bivalent ligands have been developed with a dual aim: (*a*) to improve affinity, by providing additional interactions, and (*b*) to improve selectivity, if these additional interactions involve less conserved regions across a family of receptors.[2,3,5,6] More recently bivalent ligands have been used as tools to explore the concept of G protein-coupled receptor (GPCR) dimerization.

Three distinct hypotheses can explain why an increase in affinity is observed for a bivalent ligand, compared to the corresponding monovalent ligand.[7] The first possibility is that the local concentration of the pharmacophore is increased in the vicinity of the receptor binding site (because there are two pharmacophores covalently tethered), which increases the probability of a productive binding event. Secondly, that one pharmacophore of the bivalent ligand binds to the orthosteric site, whilst the second pharmacophore binds to a neighboring (allosteric) site within the same receptor. Ligands exploiting this mode of interaction have recently been termed bitopic or dualsteric ligands with several studies describing such ligands targeting muscarinic receptors or adenosine receptors.[8-10] The third, and most commonly favored possibility, is that the bivalent ligand binds to a dimeric complex of GPCRs, binding simultaneously at adjacent orthosteric sites. This binding event is thought to be a two-stage process where one pharmacophore of the bivalent ligand binds univalently to the receptor dimer, allowing the second pharmacophore to more readily associate with the adjacent protein of the dimer, thus leading to increased affinity and (potentially) selectivity.[7,11]

There is increasing evidence proposing that GPCRs act as dimeric or higher order oligomeric proteins, signifying the therapeutic potential of homo- and heterodimers as unique pharmacological targets.[12,13] Indeed, one way to elucidate the nature of GPCR homo- or heterodimers is to use bivalent ligands as pharmacological tools, to facilitate the determination of the distance between each binding site from each monomer.[2,7] Furthermore, provided a bivalent ligand binds simultaneously to two identical binding sites, the binding affinity should ideally be the product of the binding affinities of the two individual pharmacophores.[7] However, because all bivalent ligands exhibit two pharmacophores joined by a spacer, the nature, length, and flexibility of this latter structural feature itself can significantly influence the activity of the designed bivalent ligand. For instance, if the spacer is too short, the ligand cannot bridge both binding sites simultaneously. Furthermore, the rigidity or flexibility of the structure of the spacer can influence the behavior of the bivalent ligand.[14]

Much of the pioneering work describing bivalent ligands targeting GPCRs was led by the group of Portoghese, investigating bivalent ligands targeting opioid receptor subtypes.[1,2,5-7,15,16] For example, the tethering of the κ-selective antagonist pharmacophore 5′-guanidinonaltrindole to the δ-selective antagonist pharmacophore naltrindole, yielded a δ-κ opioid receptor heterodimer selective ligand with optimal in vitro and in vivo potency when a spacer length of 20-21 atoms was used.[17] Subsequently, homo- and heterobivalent ligands have also been developed to target a number of GPCRs including adenosine,[18-20] adrenergic,[19] cannabinoid,[21] dopamine,[18,22-24] muscarinic[25,26] and serotonin[27-29] receptors. Generally, for studies targeting dimeric GPCRs, the optimal spacer length described was in the range of 15 to 22 atoms. As compared to the appropriate monovalent compounds, these bivalent ligands often displayed increased affinity at the receptor under investigation ranging from a large 50-fold increase observed for bivalent ligands targeting the opioid receptor,[15] to

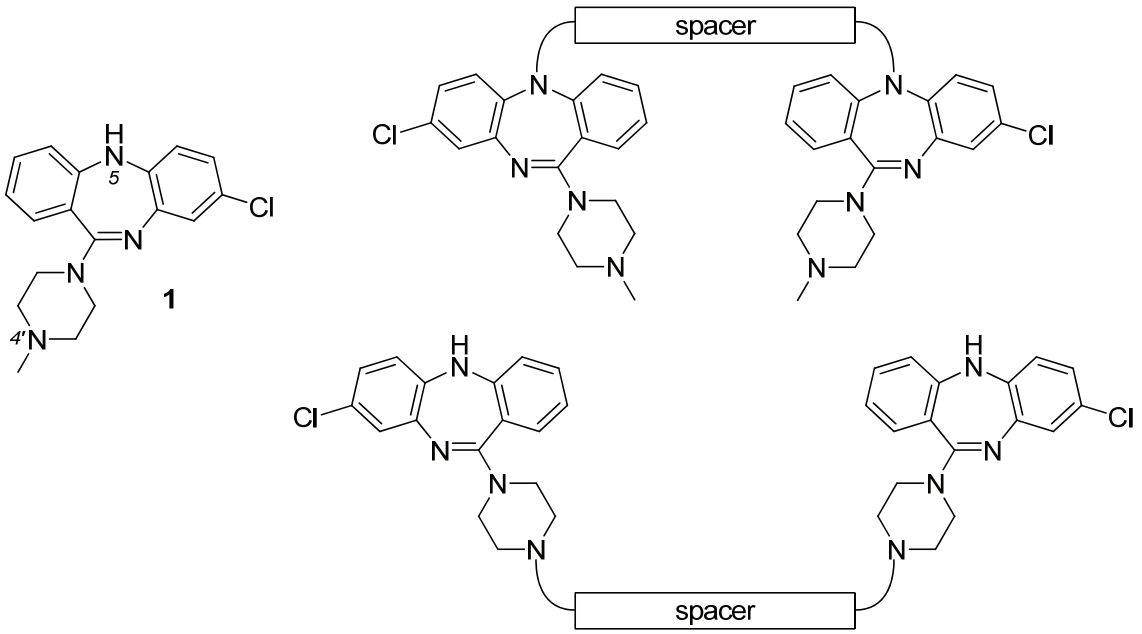
more modest increases in potency, such as the 4-fold increase observed for cannabinoid 1 receptor targeting bivalent ligands.[21]

The dopamine D_2 receptor (D_2R) is an established target for the treatment of disease states, such as schizophrenia and Parkinson's disease.[30-33] An increasing amount of evidence from both biochemical and biophysical studies, largely performed in heterologous cells systems, suggests that the D_2R may exist as a homomer with itself or form hetero-oligomeric complexes with other receptors such as the adenosine A_{2A} receptor ($A_{2A}AR$-D_2R),[34] somatostatin receptor type 5 (D_2R-$SSTR_5$),[35] serotonin 5-HT_{2A} receptor (5-$HT_{2A}R$-D_2R)[36] and the dopamine D_1 receptor (D_1R-D_2R).[37] These heteromers may represent novel therapeutic targets for the treatment of disease states in which the D_2R is implicated. Accordingly, a number of studies have explored bivalent ligands as a method to improve the affinity and selectivity of known pharmacophores against both D_2R receptor heteromers[18] and, most relevant to this study, homomers.[23,24] Abadi et al. developed a series of bivalent azecine derivatives, with the six carbon methylene spacer displaying the best, if moderate, activity.[22] More recently, Gmeiner and co-workers developed two series of bivalent ligands to target the D_2R; 1,1′-disubstituted ferrocenes,[23] and 1,4-disubstituted aromatic piperazines/piperidines.[24] Whilst the 1,4-disubstituted piperazines/piperidines bivalent ligands of varying spacer lengths displayed similar binding affinities, a Hill slope of 2, indicative of positive cooperativity, was observed for a spacer length of 22 atoms. As such, this ligand was proposed to bind simultaneously to two neighboring binding sites within a D_2R dimer.

Clozapine (**1**), a dopamine D_2R antagonist, is an atypical antipsychotic with unparalleled efficacy for the treatment of refractory schizophrenia.[38,39] However, the dibenzodiazepine structure of clozapine has been implicated in the potentially fatal blood disorder, agranulocytosis, which limits its use clinically.[40,41] This drug-induced dyscrasia is thought to result from the formation of a reactive nitrenium ion intermediate involving the N5 position

of clozapine (**1**, Chart 1).[42] Nevertheless, clozapine is a clinically effective antipsychotic that exerts its effect at a number of biogenic amine GPCRs such as dopamine, serotonin, histamine, adrenergic and muscarinic receptors, leading to a complex pharmacological profile.[43,44] Unlike the typical antipsychotics, such as haloperidol that have high affinity for D_2R, clozapine has a lower affinity, which has been proposed to be a result of the fast off-rate from the D_2R.[45] This work describes the design and synthesis of three series of homobivalent ligands of the atypical antipsychotic, clozapine (**1**), using two distinct attachment points and a series of simple dicarboxylic acid spacers. The main aim of this study was to determine if covalently tethering two molecules of clozapine would improve its affinity for the D_2R. Functional studies are reported for all compounds synthesized and promising compounds were further evaluated using radioligand binding studies.

Chart 1. Structure of clozapine (**1**) and general structures of homobivalent ligands of clozapine.



Ligand Design Rationale. Clozapine (**1**) is the prototype atypical antipsychotic,[39] and was an attractive pharmacophore for use in the design of homobivalent ligands to investigate the D_2R homodimer (Chart 1).

Both the N5 and the distal piperazine nitrogen (N4′) positions of clozapine were synthetically attractive points for the attachment of spacers for the preparation of homobivalent ligands. Clozapine analogues with attachments on the N4′ position[46-48] and the N5 position[49,50] have both been previously synthesized, and modifications at these positions were well tolerated.

By developing homobivalent ligands using the N5 attachment point, it may be possible that the drug-induced dyscrasia could be reduced or abolished.[42,51,52] However, directly acylating the N5 position of clozapine may affect the conformational and electronic properties of the tricyclic nucleus. Therefore, in addition to acylation at this position, converting the N5 position to the hydrazine functionality, so that the spacer attachment point was not directly attached to the tricyclic ring system was also investigated. Formation of a hydrazone at the N5 position has also been demonstrated to be well tolerated at this position.[50]

The other attachment point investigated was the distal piperazine nitrogen (N4′). The N4′ nitrogen is the ionizable nitrogen that interacts with the key aspartate residue on helix 3 (Asp $114^{3.32}$) at the entrance of the orthosteric binding site.[53] However, directly acylating at this position would significantly change the pK_a of the ionizable nitrogen and interfere with the critical electrostatic interaction with the receptor. Therefore a propylamine linker group between the ionizable nitrogen and the spacer was introduced. Previously, it has been shown that alkylation at the N4′ position of clozapine is well tolerated, as demonstrated by in vitro assays and in vivo behavioral models.[46-48]

In addition to the synthesis of homobivalent ligands, monovalent ligands were developed for each of the three pharmacophores, for comparison of pharmacological activity against the homobivalent ligands.

Simple dicarboxylic acids were selected as the spacers for the synthesis of homobivalent ligands of clozapine, as they possess the desired functionality to form a stable amide bond to the pharmacophore. These were used to determine the appropriate spacer length for the bivalent ligands. More complex dicarboxylic acids, incorporating heteroatom-rich functionalities, were also explored to improve any solubility issues that may arise from the inclusion of a polymethylene chain.

Results and Discussion

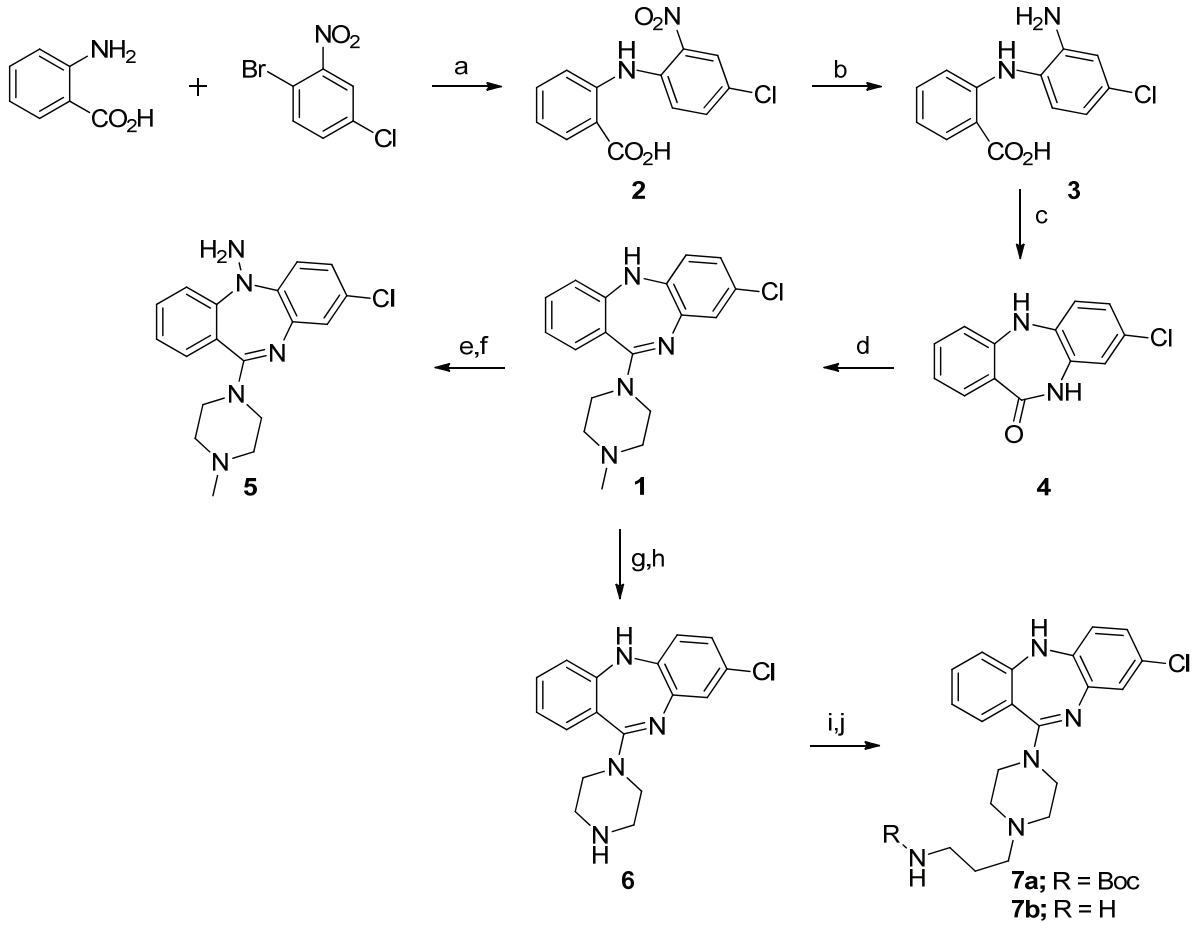
Chemical Synthesis. Three series of homobivalent ligands of clozapine were synthesized, using two distinct pharmacophore attachment points. The synthesis of the clozapine-based pharmacophores began with the preparation of clozapine (**1**) using a previously described procedure.[46,54]

Scheme 1 depicts the synthesis of clozapine (**1**), and key clozapine intermediates (**5** and **7b**) that were incorporated into the target homobivalent ligands. Coupling of anthranilic acid and commercially available 2-bromo-5-chloronitrobenzene under Ullmann reaction conditions produced the nitro acid (**2**). Subsequent reduction using sodium dithionite afforded the amino compound (**3**), which underwent a thermal cyclization under Dean-Stark conditions to yield the tricyclic lactam (**4**). Clozapine (**1**) was readily synthesized from **4** and *N*-methylpiperazine in the presence of the Lewis acid, titanium tetrachloride. All compounds were produced in good yields.

Following the procedure described by Su et al.,[50] clozapine (**1**) was converted to the clozapine hydrazine (**5**) intermediate, first by N-nitrosylation with isoamyl nitrite, followed by reduction with zinc metal in acetic acid to form **5** in moderate yields (48%). Clozapine

was also N-demethylated using α-chloroethyl chloroformate,[55] yielding *N*-desmethylclozapine (**6**), in respectable yield (69%). **6** was further alkylated with *tert*-butyl 3-bromopropylcarbamate in the presence of sodium iodide and *N*,*N*-diisopropylethylamine and subsequently deprotected (TFA), to yield the clozapine propylamine intermediate (**7b**).

Scheme 1. Synthesis of clozapine (**1**) and key clozapine intermediates; clozapine hydrazine (**5**), *N*-desmethylclozapine (**6**) and the clozapine propylamine intermediate (**7b**).

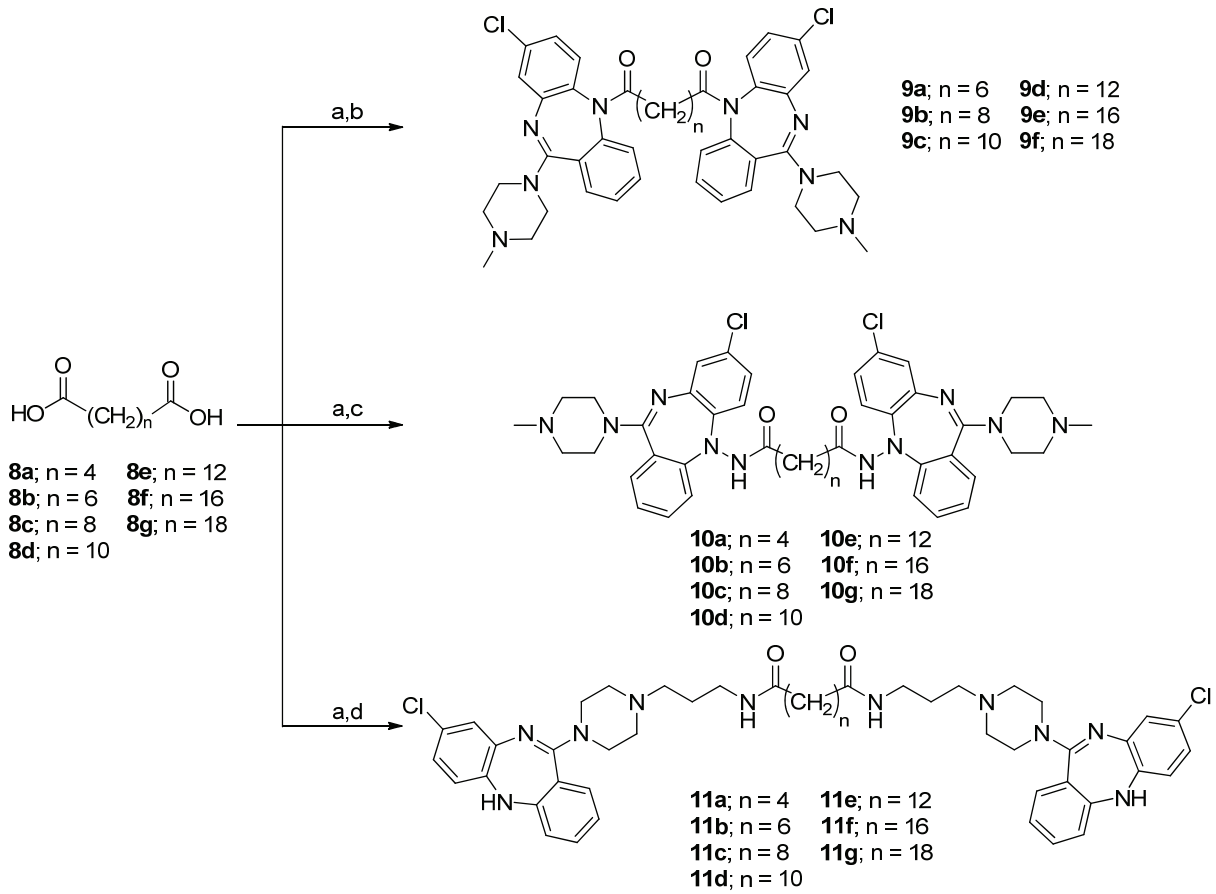


Reagents and conditions: (a) K_2CO_3, isoamyl alcohol, copper powder, reflux, 77%; (b) $Na_2S_2O_4$, aqueous ammonia, 80 °C, 74%; (c) *o*-xylene, reflux, 58%; (d) anisole, methyl piperazine, $TiCl_4$, 50-55 °C then reflux, 94%; (e) isoamyl nitrite, CH_2Cl_2; (f) zinc powder, HOAc, 15 °C, 48%; (g) 1,2-dichloroethane, α-chloroethyl chloroformate, 0 °C, then reflux; (h) CH_3OH, 50 °C, 69%; (i) *tert*-butyl 3-bromopropylcarbamate, NaI, DIPEA, CH_3CN, reflux, yields **7a**, 79%; (j) TFA, CH_2Cl_2, followed by base, 94%.

The dicarboxylic acids (**8a-g**) were converted to their corresponding diacid chlorides using oxalyl chloride and *N*,*N*-dimethylformamide and reacted, without further purification, with

1, **5** or **7b**, to yield the target homobivalent ligands as white (**9a-f**, 27-37%), off-white (**10a-g**, 38-79%) and yellow foams (**11a-g**, 37-66%) respectively, in moderate to good yields (Scheme 2).

Scheme 2. Synthesis of clozapine homobivalent ligands, using the clozapine pharmacophores **1**, **5** and **7b**.

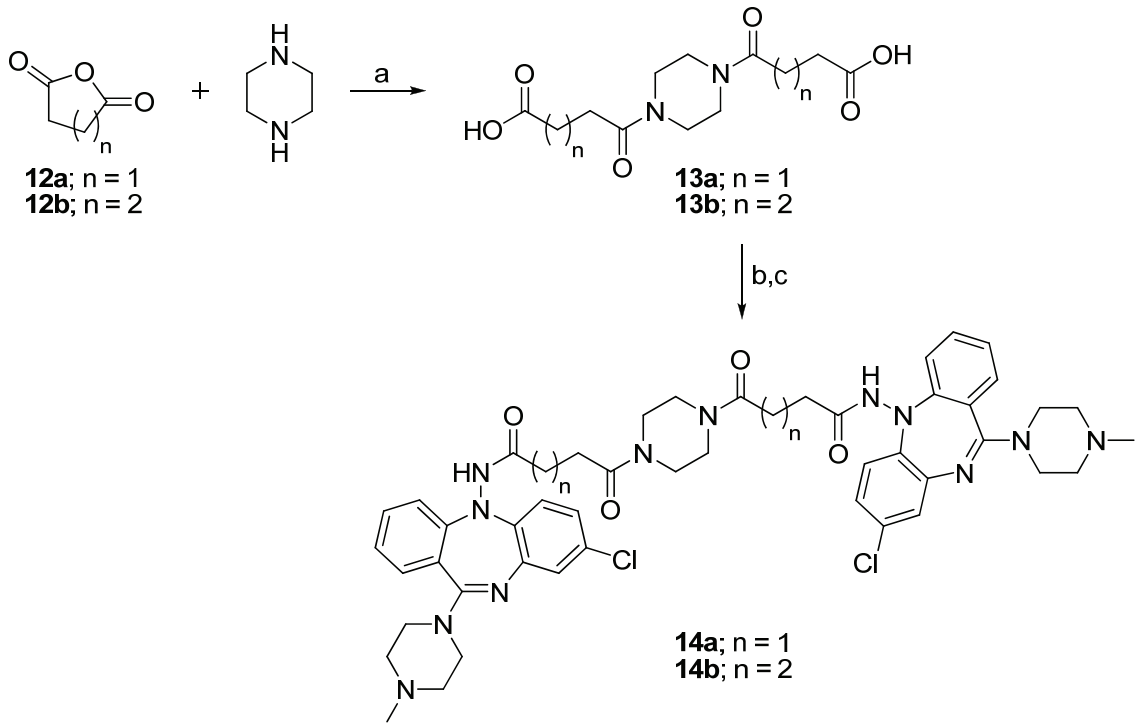


Reagents and conditions: (a) oxalyl chloride, DMF, CH_2Cl_2, 27-37%; (b) **1**, pyridine, DIPEA, CH_2Cl_2, 38-79%; (c) **5**, pyridine, DIPEA, CH_2Cl_2; (d) **7b**, pyridine, DIPEA or K_2CO_3, CH_2Cl_2, 37-66%.

In addition to the simple dicarboxylic acid spacers, more complex *N,N'*-disubstituted piperazinyloxocarboxylic acids (**13a-b**, Scheme 3) and dioxodioic acid (**15**, Scheme 4) spacers, with two spacer lengths of 12 atom and 14 atoms were synthesized. By developing spacers with additional functionalities, we were aiming to tune the hydrophobicity of the spacer by the incorporation of additional heteroatoms. **13a** and **13b** afforded white

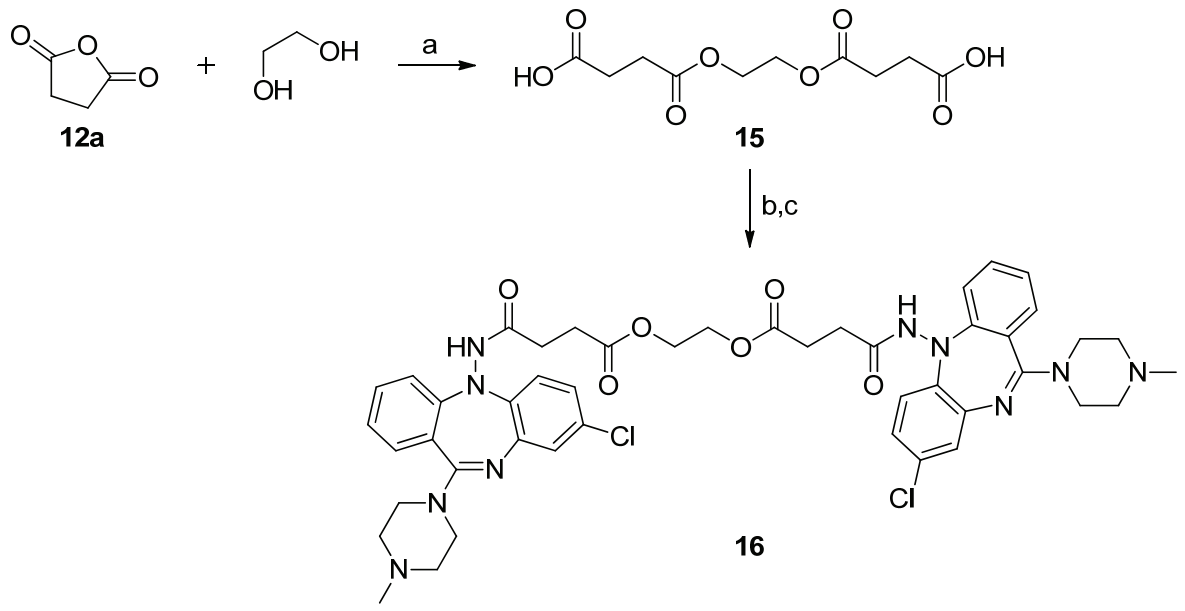
microcrystalline solids[56] by reacting piperazine with two equivalents of the desired cyclic anhydride at reflux, in good yields (62-82%). **15** was synthesized by heating ethylene glycol and two equivalents of succinic anhydride under Dean-Stark conditions and isolated in moderate yield (27%).[57] These spacers were converted to their corresponding diacid chlorides using oxalyl chloride and *N,N*-dimethylformamide and, without further purification, subsequently reacted with **5** to yield the corresponding homobivalent ligands as off-white foams, in moderate yields (24-36%). Interestingly, the target compounds **14a-b** and their corresponding precursors **13a-b**, displayed the existence of a mixture of cisoid and transoid amide rotamers by NMR spectroscopy.[56,57] Further investigation into varying spacer lengths was abandoned due to difficulties in synthesizing longer spacers, generally resulting in polymerization. Examples of these homobivalent ligands were only synthesized for the clozapine hydrazine pharmacophore (**5**), as a proof of concept.

Scheme 3. Synthesis of homobivalent ligands of **5**, containing more complex *N,N'*-disubstituted piperazinyloxocarboxylic acid spacers (**13a-b**).



Reagents and conditions: (a) toluene or 1,4-dioxane, reflux, 62% (**13a**), 82% (**13b**); (b) oxalyl chloride, DMF, CH_2Cl_2; (c) **5**, pyridine, CH_2Cl_2, 24% (**14a**), 36% (**14b**).

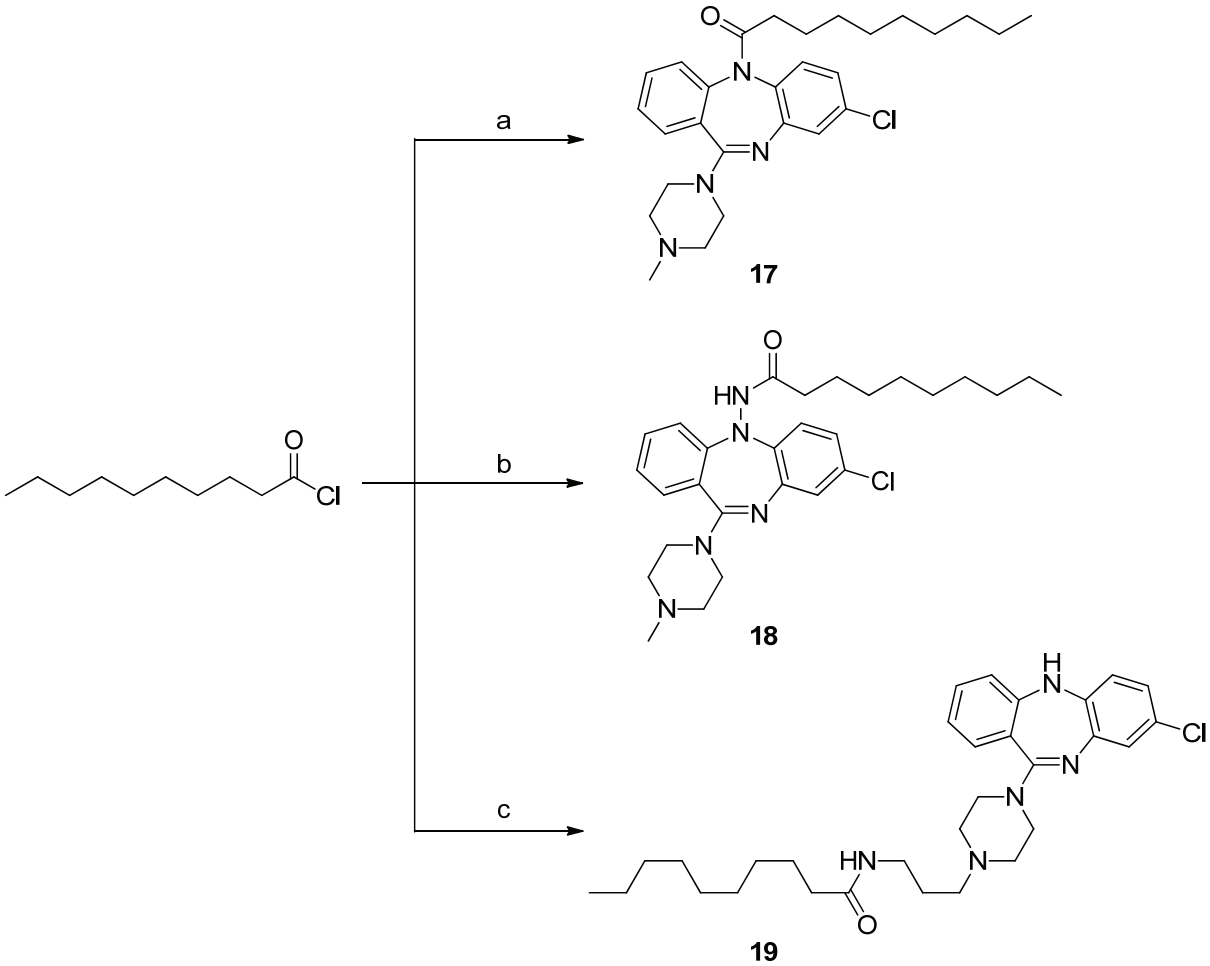
Scheme 4. Synthesis of homobivalent ligand of **5**, containing a dioxodioic acid spacer (**15**).



Reagents and conditions: (a) toluene, reflux, 27%; (b) oxalyl chloride, DMF, CH_2Cl_2; (c) **5**, pyridine, CH_2Cl_2, 46%.

Monovalent ligands were also synthesized for all three attachment points by the reaction of clozapine intermediates with decanoyl chloride in the presence of base, to yield the corresponding monovalent ligands (**17**, **18**, and **19**, Scheme 5) in moderate yields. These compounds were designed for comparative purposes in the pharmacological assays.

Scheme 5. Synthesis of monovalent ligands for the three pharmacophores (**1**, **5** and **7b**).



Reagents and conditions: (a) **1**, pyridine, DIPEA, CH_2Cl_2, 36%; (b) **5**, pyridine, CH_2Cl_2, 79%; (c) **7b**, pyridine, CH_2Cl_2, 63%.

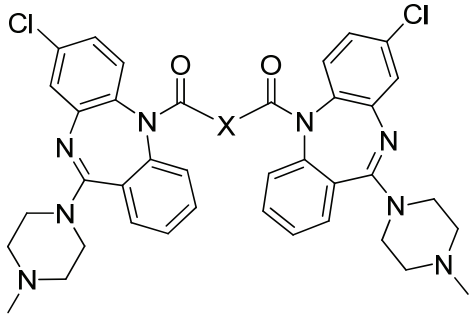
Functional Assays. The SureFire AlphaScreen™ pERK 1/2 assay kit (PerkinElmer/TGR BioSciences) was used to measure the ability of the endogenous agonist dopamine to stimulate phosphorylation of ERK 1/2 (T202/Y204) mediated by the activation of human dopamine D_{2L} receptor ($D_{2L}R$) stably expressed in a FlpIn CHO (Chinese hamster ovary) cell line. Dopamine behaved as an agonist with a pEC_{50} of 8.4 ± 0.08. To allow an estimation of the inhibitory potency of the homobivalent ligands, monovalent ligands and clozapine, we tested the ability of increasing concentrations of these ligands to antagonize an EC_{80} concentration (10 nM) of dopamine. IC_{50} values for all compounds were determined from the functional assay. Clozapine (**1**) displayed a sub-micromolar inhibitory potency with

a pIC_{50} of 6.69 ± 0.2 (206 nM). The clozapine N5 derivatives (**9a-f**, **17**, Table 1) displayed diminished antagonistic activity. Compound **17** showed some antagonistic activity (IC_{50} = 720 nM), which was approximately 4-fold less potent than clozapine (**1**). This result indicated that minor substitutions at this position could be tolerated, and was in agreement with other results of N5 acylated clozapine analogues.[58] However, homobivalent ligands with significantly larger substitutions at the N5 position due to the attachment of the second pharmacophore (**9a-f**), displayed negligible activity in the functional assay, which suggests that there is some degree of size limitation to the substitutions that can be made at this position.

A similar trend was observed for the hydrazide-linked homobivalent ligands (**10a-g**, Table 2), with marginal antagonistic activity, displaying at best, low micromolar activity in the functional assay. This observation was also applicable to the compounds containing the more complex dicarboxylic acid spacers (**14a-b**, **16**).

For the clozapine propylamine derivatives (**11a-g, 19**), a spacer length dependent effect (Figure 1, Table 3) was observed upon the inhibitory potency of this series of compounds. The monovalent ligand (**19**, 1.46 μM) was approximately 7-fold less active than clozapine (**1**, 206 nM). The 14 atom spacer (**11a**, 87 nM) showed notably more activity than clozapine and the monovalent ligand (**19**) (2.4-fold and 17-fold increase in potency, respectively). The highlight from the series was the 16 atom spacer homobivalent ligand (**11b**, 23 nM) exhibiting the best activity of all the compounds developed, being 9-fold more potent than clozapine in the functional assay. The 18 atom spacer (**11c**) exhibited slightly less activity (44 nM) compared to the 16 atom spacer, but was still 5-fold more active than clozapine. Beyond the 18 atom spacer, we observed a gradual, spacer length dependent, reduction in activity for the 20, 22, 26 and 28 atom spacers (**11d-g**) compared to the shorter homobivalent ligands.

Table 1. Inhibitory potency of the clozapine N5 homobivalent (**9a-f**) and monovalent (**17**) ligands to inhibit the effect of 10 nM dopamine in a SureFire AlphaScreen™ ERK1/2 phosphorylation assay using CHO cells expressing the D_2R. Data represents three separate experiments performed in duplicate.



Compound	Spacer length	Spacer type (X)	pIC_{50} ± SEM (IC_{50}, nM)
clozapine (**1**)	-	-	6.69 ± 0.20 (206)
17	-	-	6.11 ± 0.12 (776)
9a	8	$(CH_2)_6$	5.75 ± 0.22 (2,662)
9b	10	$(CH_2)_8$	< 5 (> 10,000)
9c	12	$(CH_2)_{10}$	< 5 (> 10,000)
9d	14	$(CH_2)_{12}$	< 5 (> 10,000)
9e	18	$(CH_2)_{16}$	< 5 (> 10,000)
9f	20	$(CH_2)_{18}$	< 5 (> 10,000)

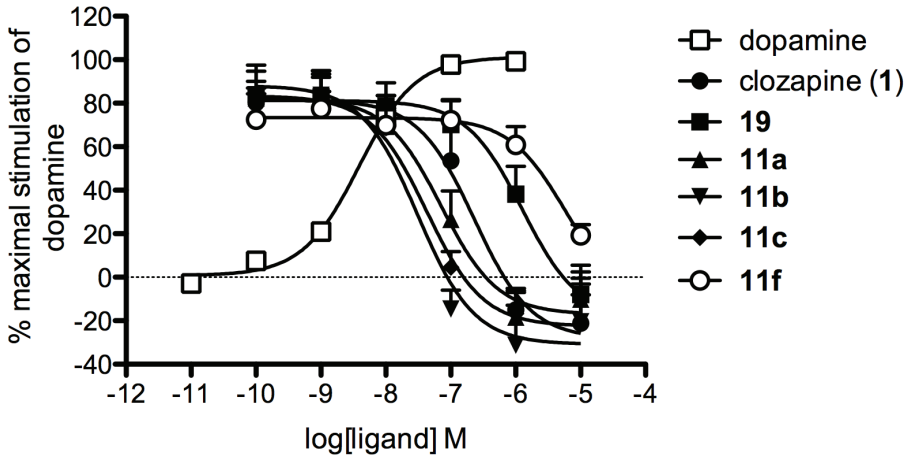
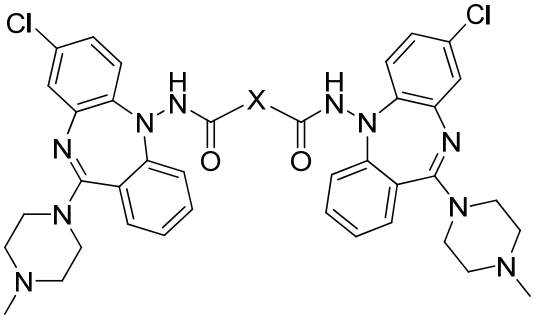


Figure 1. The ability of clozapine propylamine homobivalent (**11a-g**) and monovalent (**19**) ligands to inhibit the effect of 10 nM dopamine in a SureFire AlphaScreen™ ERK1/2 phosphorylation assay using CHO cells expressing the D_2R. Data represents three separate experiments performed in duplicate.

Table 2. Inhibitory potency of the clozapine hydrazide homobivalent (**10a-g**, **14a-b** and **16**) and monovalent (**18**) ligands to inhibit the effect of 10 nM dopamine in a SureFire AlphaScreen™ ERK1/2 phosphorylation assay using CHO cells expressing the D_2R. Data represents three separate experiments performed in duplicate.



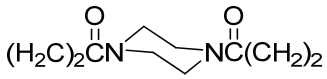
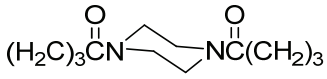
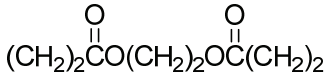
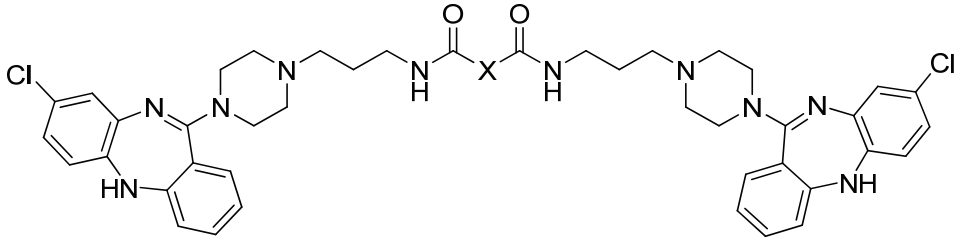
Compound	Spacer length	Spacer type (X)	pIC_{50} ± SEM (IC_{50}, nM)
clozapine (**1**)	-	-	6.69 ± 0.20 (206)
18	-	-	5.05 ± 0.03 (8,939)
10a	6	$(CH_2)_4$	5.69 ± 0.21 (2,078)
10b	8	$(CH_2)_6$	5.61 ± 0.13 (2,440)
10c	10	$(CH_2)_8$	< 5 (> 10,000)
10d	12	$(CH_2)_{10}$	< 5 (> 10,000)
10e	14	$(CH_2)_{12}$	< 5 (> 10,000)
10f	18	$(CH_2)_{16}$	< 5 (> 10,000)
10g	20	$(CH_2)_{18}$	< 5 (> 10,000)
14a	12		5.79 ± 0.22 (1,617)
14b	14		5.58 ± 0.20 (2,633)
16	12		5.62 ± 0.16 (2,412)

Table 3. Inhibitory potency of clozapine propylamine homobivalent (**11a-g**) and monovalent (**19**) ligands to inhibit the effect of 10 nM dopamine in a SureFire AlphaScreen™ ERK1/2 phosphorylation assay using CHO cells expressing the D_2R. Data represents three separate experiments performed in duplicate.



Compound	Spacer length	Spacer type (X)	pIC_{50} ± SEM (IC_{50}, nM)
clozapine (**1**)	-	-	6.69 ± 0.20 (206)
19	-	-	5.84 ± 0.16 (1,455)
11a	14	$(CH_2)_4$	7.06 ± 0.20 (87)
11b	16	$(CH_2)_6$	7.63 ± 0.20 (23)
11c	18	$(CH_2)_8$	7.35 ± 0.12 (44)
11d	20	$(CH_2)_{10}$	5.96 ± 0.12 (1,119)
11e	22	$(CH_2)_{12}$	4.97 ± 0.18 (11,000)
11f	26	$(CH_2)_{16}$	5.11 ± 0.14 (7,800)
11g	28	$(CH_2)_{18}$	< 5 (> 10,000)

Radioligand binding assays. As described above, the SureFire AlphaScreen™ pERK1/2 assay represents a useful assay to allow the screening of a range of both monovalent and bivalent clozapine derived ligands at the human $D_{2L}R$. To confirm their activity, the most promising clozapine bivalent ligands from the functional assay were further investigated by testing their ability to displace the radiolabeled antagonist [^{3}H]spiperone at the human $D_{2L}R$ expressed in FlpIn CHO cell membranes. The most active clozapine propylamine homobivalent ligands described earlier (**11a-c**), the corresponding monovalent ligand (**19**), clozapine (**1**), as well as a clozapine propylamine homobivalent ligand that displayed poor activity in the functional assay (**11f**) were evaluated (Figure 2, Table 4). The parent

compound clozapine (**1**) showed a similar pK_i (6.99 ±0.08, K_i = 106 nM) in this binding assay as compared to the inhibitory potency (pIC_{50}= 6.69± 0.2, IC_{50} = 206 nM) determined in the pERK1/2 functional assay. Furthermore, the order of inhibitory potency observed in the functional pERK1/2 assay was preserved in the radioligand binding assay; the homobivalent ligands with shorter spacers (**11a-c**) displaying significantly greater affinity than the homobivalent ligand with longer spacers (**11f**). Indeed, the bivalent ligands with shorter spacers (14-18 atoms, **11a-c**) displayed between 30 and 79 times greater affinity for the $D_{2L}R$ as compared to clozapine, with the two most active compounds demonstrating a low nanomolar affinity (**11b** and **11c**, 1.41 and 1.35 nM, respectively). However, one interesting discrepancy between the functional and radioligand binding data should be noted. In the [^{3}H]spiperone binding assay, the monovalent ligand (**19**) displayed markedly enhanced affinity (12-fold, 9.06 nM) compared to clozapine. By comparison in the functional assay, this compound displayed a 6-fold decrease in potency as compared to clozapine (**1**). Importantly, the most active compounds in this series (**11a-c**) still showed increases in affinity from 2.5- to 6.5-fold as compared to the monovalent compound. For all compounds tested, the inhibition curves had Hill slopes not significantly different from unity (Table 4).

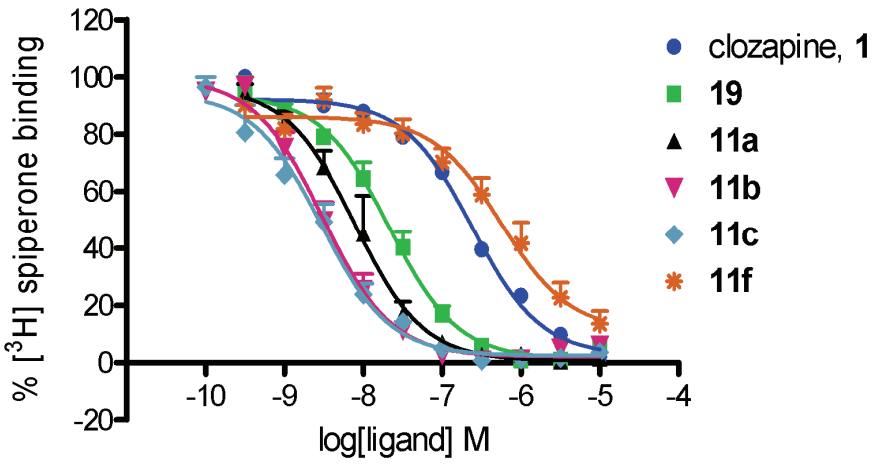


Figure 2. The ability of clozapine propylamine homobivalent (**11a-g**) and monovalent (**19**) ligands to inhibit the binding of the antagonist [^{3}H]spiperone at the D_2R expressed in FlpIN CHO cell membranes. Data represents three separate experiments performed in duplicate.

Table 4. The affinity of propylamine homobivalent (**11a-g**) and monovalent (**19**) ligands determined using competition binding experiments using the radiolabeled antagonist [^{3}H]spiperone at the D_2R expressed in FlpIN CHO cell membranes. Data represents three separate experiments performed in duplicate.

Compound	**Spacer length**	**p*K*$_i$ ± SEM**	***K*i ±SEM, nM**	**Hill slope**
clozapine (**1**)	-	6.99 ± 0.08	106 ± 20	0.91 ± 0.09
19	-	8.05 ± 0.06	9.06 ± 1.28	0.97 ± 0.12
11a	14	8.50 ± 0.14	3.56 ± 1.21	1.04 ± 0.19
11b	16	8.87 ± 0.09	1.41 ± 0.31	1.05 ± 0.06
11c	18	8.91 ± 0.14	1.35 ± 0.41	0.82 ± 0.10
11f	26	6.61 ± 0.13	269 ± 74.8	0.80 ± 0.20

The main aim of developing homobivalent ligands of clozapine was to assess the effect on affinity and functional activity of covalently tethering two clozapine pharmacophores by a spacer of a given length and type. Importantly, these homobivalent ligands were evaluated against clozapine, as well as the corresponding monovalent ligand, which incorporates a capped spacer. This allowed for the assessment of any activity or potency gains compared to clozapine, as well as using the monovalent ligand to evaluate the influence of the spacer group on binding affinity and activity. As postulated by Portoghese et al., if a bivalent ligand binds simultaneously to two identical binding sites, ideally the binding affinity should be the product of the binding affinities of the two individual pharmacophores.[7] Thus, gains in affinity or activity of the bivalent ligand can also be assessed by comparison to the original pharmacophore, in this case clozapine.

Of the three series of clozapine bivalent ligands, the clozapine propylamine series with the spacer attached from the N4′ distal piperazine nitrogen of clozapine, was identified as the preferred attachment point for the spacer. Importantly, all of the tested clozapine derivatives retained the antagonistic action of the parent compound. It is noteworthy that linking from

the ionizable piperazine nitrogen also generated bivalent ligands that displayed promising activity in another series of D_2R targeting homobivalent ligands.[24] This is likely to be a result of the positioning of the ionizable nitrogen, due to the formation of the key salt bridge with Asp $114^{3.32}$, at the entrance of the orthosteric binding site. For the clozapine propylamine bivalent ligands, it was interesting to note that affinity and activity was spacer-length dependent, with similar trends observed in both functional and radioligand binding studies. Both studies indicated the 16 and 18 atom spacers (**11b** and **11c**) were the most potent/active, followed by a gradual decrease in activity with increasing spacer length. These spacer lengths are within the 15 to 22 atom spacer length range identified in other bivalent ligand studies targeting GPCR dimers or oligomers, although they are shorter than the 22 atom spacer proposed by Gmeiner and co-workers for D_2R targeting homobivalent ligands.[24] However, in the study by Gmeiner and co-workers, the optimal compound in a 1,4-disubstituted aromatic piperazines/piperidines series was identified by a change in the Hill slope, unlike the clozapine propylamine bivalent ligand series (**11a-g**), in which a distinct spacer-length dependency for ligand affinity was observed.[24] Interestingly, although the steep Hill slopes observed for the 1,4-disubstituted aromatic piperazines/piperidines series is indicative of positive cooperativity between the two pharmacophores, this was not accompanied by any increase in affinity as compared to the monovalent compound. In contrast, for the clozapine propylamine bivalent ligands (**11b** and **11c**, 1.41 and 1.35 nM), the Hill slopes were at unity, yet potency gains of 6.4- and 6.7-fold relative to the monovalent ligand were observed (**19**, 9 nM), as well as a 75- and 79-fold increase in affinity compared to clozapine (**1**, 106 nM). If the spacer types of the two series are examined, they are quite similar, as they both contain an aliphatic spacer connected to the linking group by a propyl chain, using a triazine or an amide bond. However, in the case of the clozapine propylamine homobivalent ligands (**11a-g**), the linker is directly attached to the distal piperazine nitrogen in clozapine, where it is attached via a methoxy benzyl group

to the piperazine of the 1,4-disubstituted aromatic piperazines/piperidines series,[24] slightly increasing the rigidity of the compound. Now that an appropriate attachment point has been determined, as well as an approximate spacer length, it would be interesting to investigate the effect of more conformationally constrained spacers on affinity and functional activity. Other functionalized spacers, such as polyethylene glycols, could be incorporated to increase the aqueous solubility of the bivalent ligand. Additionally, different spacer attachment methods and linking groups could also be explored.

The monovalent ligand of the clozapine propylamine series displayed unique activity. Specifically, in the preliminary functional assay the monovalent ligand (**19**, 1.46 μM) was 7-fold less active than clozapine (**1**, 206 nM), yet in the radioligand binding assays the monovalent ligand (9 nM) displayed 12-fold greater activity than clozapine (106 nM). This underlines the need to pharmacologically evaluate compounds with both functional and binding assays. In a recent study using bitopic ligands to target the muscarinic M_2 AChR, Steinfeld et al. observed the seven-chain pharmacophore spacer itself also promoted an increase in compound affinity, highlighting the importance of including incremental fragments of novel bivalent or bitopic ligands in control experiments.[59] In the present study, the introduction of the second pharmacophore (the bivalent ligand), improved the affinity or inhibitory potency of the compound and implies that the presence of second pharmacophore engenders this gain of affinity or potency. Accordingly, both the gain in affinity and optimal spacer length observed for the clozapine propylamine bivalent series is consistent with the interaction of these ligands with a D_2R dimer. However, given that the spacers of the clozapine propylamine bivalent ligands (**11a-g**) are very flexible, it is plausible that other binding mechanisms, such as increasing the local concentration of the pharmacophore in the vicinity of the receptor binding site or binding to two topically distinct sites on one receptor, may also explain these observed gains in affinity. Although there is accumulating evidence to suggest that the D_2R receptor forms homo-oligomers, there remains sufficient controversy

regarding the ability of class A receptors to dimerize in vivo to give these other mechanisms due consideration. As such, these high affinity clozapine bivalent ligands (**11b** and **11c**) may represent useful pharmacological tools to investigate D_2R dimers in combination with appropriate biochemical or biophysical studies.

With respect to the significant affinity gains observed for the clozapine bivalent ligands (**11b** and **11c**, 75- and 79-fold) relative to the original pharmacophore, clozapine, it would also be of great interest to investigate how the covalent tethering of two clozapine pharmacophores affects the atypical nature of these compounds compared to clozapine.

Conclusions

We describe the design, synthesis and pharmacological evaluation of homobivalent ligands of the atypical antipsychotic, clozapine (**1**), differing in the nature and length of the spacer and point of attachment to the pharmacophore. The best attachment point for the synthesis of clozapine homobivalent ligands was identified as the N4′ position, which also incorporated a linking group between the ionizable nitrogen and the spacer. Both functional and binding assays revealed a spacer-length dependent effect for compounds **11a-g**, with the most active compounds (**11b** and **11c**) having spacer lengths of 16 and 18 atoms, respectively. These compounds displayed low nanomolar affinity (1.41 and 1.35 nM) and activity (23 and 44 nM respectively). Additionally, significant gains in affinity (75- and 79-fold) and functional activity (4.7- to 9-fold) relative to the original pharmacophore, clozapine. The clozapine propylamine bivalent ligands developed in this study could be of use to further elucidate the atypical nature of clozapine, as well as being used as pharmacological tools to investigate D_2R dimerization.

Experimental Methods

Chemistry General Experimental

All materials were reagent grade and purchased commercially from Sigma-Aldrich, Alfa Aesar, Tokyo Chemical Industry, AOKChem and Merck. Succinic anhydride and glutaric acid were recrystallized from chloroform and adipic acid was recrystallized from ethyl acetate prior to use. Ethyl acetate and hexane were redistilled prior to use. Dichloromethane was purified by pre-drying with calcium chloride and freshly distilling from calcium hydride prior to use.

Thin layer chromatography (TLC) was performed using Merck Silica Gel 60 F_{254} pre-coated plates (0.25 mm) and visualized by ultraviolet light, as well as staining with iodine or ninhydrin. Flash column chromatography used Merck Silica Gel 60, 230-400 mesh ASTM, following the method described by Still et al.[60] All compounds were pre-adsorbed onto coarse silica (70-230 mesh ASTM) prior to column chromatography, unless otherwise stated. Where gradient elution was utilized for column chromatography, the eluent was modified as detailed in the experimental, in 50-100 mL increments.

^{1}H NMR spectra were routinely recorded at 300.13 MHz using a Bruker Avance DPX-300 spectrometer or at 400.13 MHz using a Bruker Ultrashield-Avance III NMR spectrometer, using TOPSPIN v2.1 software, at 298 K, unless stated otherwise. Chemical shifts (δ_H) for all ^{1}H NMR spectra were reported in parts per million (ppm) using tetramethylsilane (TMS) as the internal standard (δ_H 0.00 ppm) in deuterated solvents, including chloroform ($CDCl_3$), d_6-dimethyl sulfoxide (d_6-DMSO), d_4-methanol (CD_3OD), d_6-acetone ($(CD_3)_2CO$) and deuterium oxide (D_2O), as indicated. The ^{1}H NMR spectra were reported as follows: chemical shift (δ), multiplicity, coupling constants (J) in Hertz (quoted to one decimal place ± 0.2 Hz), peak integration and assignment. In reporting the spectral data, the following abbreviations have been used: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad, app = apparent.

^{13}C NMR spectra were routinely recorded at 75.5 MHz using a Bruker Avance DPX-300 spectrometer or at 100.62 MHz using a Bruker Ultrashield-Avance III NMR spectrometer, using TOPSPIN v2.1 software, at 298 K, unless stated otherwise. Distortionless Enhancement by Polarization Transfer (DEPT) experiments were routinely used for ^{13}C NMR spectra. Chemical shifts (δ_C) for all ^{13}C NMR spectra were reported in parts per million (ppm), using the center of the solvent chemical shift as the reference: $CDCl_3$ (77.16), d_6-DMSO (39.52), CD_3OD (49.00) and d_6-acetone (29.84), as indicated.[61] ^{13}C NMR signals are assigned as: C_q = quaternary carbon, CH = methine carbon, CH_2 = methylene carbon and CH_3 = methyl carbon.

Melting points (mp) were determined using a Mettler Toledo MP50 melting point apparatus, except those marked (mp*), which were determined on a Kofler Hot Stage Micro-melting point apparatus, and are uncorrected.

Mass spectra were acquired in the positive and negative mode using an atmospheric pressure (ESI/APCI) ion source on a Micromass Platform II ESI/APCI single quadrupole mass spectrometer with sample management facilitated by an Agilent 1100 series HPLC system using MassLynx version 3.5 software. Liquid Chromatography Mass spectra (LCMS) were measured on an Agilent 6100 Series Single Quad LC/MS, Agilent 1200 Series HPLC. (Pump: 1200 Series G1311A Quaternary pump, Autosampler: 1200 Series G1329A Thermostatted Autosampler, Detector: 1200 Series G1314B Variable Wavelength Detector). Gradient takes 4 min to get to 100% ACN; maintain for 3 min and a further 3 min to return to the original 5% ACN.

High Resolution Mass Spectrometry analyses were obtained on a Waters Micromass LCT Premier XE Orthogonal Acceleration Time-of-Flight Mass Spectrometer coupled to an Alliance 2795 Separation Module using MassLynx version 4.1 software.

Analytical reverse-phase HPLC was carried out on a Waters Millenium 2690 system, fitted with a Phenomenex® Luna C8, 100 Å, 5 µm (50 x 4.60 mm I.D.) column. A binary

solvent system was used (solvent A: 0.1% TFA / H_2O, solvent B: 0.1% TFA / 19.9% H_2O / 80% acetonitrile), with UV detection at 214 nm. Method 1 used gradient elution beginning with 100% solvent A going to 20% solvent A / 80% solvent B, over 20 mins at a flow rate of 1 mL/min. Method 2 used gradient elution beginning with 80% solvent A / 20% solvent B going to 100% solvent B, over 20 mins at a flow rate of 1 mL/min. The purity of all tested compounds and key intermediates was determined to be > 95%.

Prior to pharmacological testing, all compounds were converted to their hydrochloride salts, using hydrogen chloride in diethyl ether (1.0 M).

Synthesis of spacers and clozapine pharmacophores

*2-(4-Chloro-2-nitroanilino)benzoic acid (**2**).* Following the procedure of Capuano et al.,[46] a mixture of 2-aminobenzoic acid (13.8 g, 101 mmol), 1-bromo-4-chloro-2-nitrobenzene (25.1 g, 106 mmol), anhydrous potassium carbonate (13.9 g, 101 mmol) and copper powder (0.50 g, 7.87 mmol) in isoamyl alcohol (200 mL), was heated at reflux for 4 h. The steam volatile components were removed by steam distillation, and acidification with aqueous hydrochloric acid (2 M) gave a precipitate that was collected by filtration. Recrystallization from aqueous ethanol gave **2** as red needles (22.7 g, 77.6 mmol, 77%), mp* 247-250 °C (lit.[46] 246-248 °C). ^{1}H NMR (300 MHz, d_6-acetone) δ 11.18 (s, 1H, NH), 8.16 (d, J = 2.5 Hz, 1H, H3′), 8.11 (d, J = 7.9 Hz, 1H, H6), 7.76 (d, J = 9.1 Hz, 1H, H6′), 7.65-7.51 (m, 3H, H3, H4, H5′), 7.14 (app t, J = 7.1 Hz, 1H, H5). ESI MS (m / z): 291.2 $[M - H]^-$.

*2-(2-Amino-4-chloroanilino)benzoic acid (**3**).* Following the procedure of Capuano et al.,[46] a mixture of **2** (10.1 g, 34.4 mmol) and aqueous ammonia (2 M, 250 mL) was warmed to 80 °C. Sodium dithionite (35.8 g, 205 mmol) was added portion-wise to the red colored solution until a color change to pale yellow was observed. Decolorizing charcoal was added and the solution was filtered whilst hot. The filtrate was adjusted to pH 4.5 using glacial acetic acid and the precipitate collected by filtration. Recrystallisation from methanol / water yielded **3** as pale tan micro-crystals (6.69 g, 25.5 mmol, 74%), mp* 193-195 °C (lit.[46] 198-

200 °C). ^{1}H NMR (300 MHz, d_6-DMSO) δ 8.97 (s, 1H, NH), 7.87 (dd, J = 8.0, 1.6 Hz, 1H, H6), 7.31 (ddd, J = 8.5, 7.0, 1.6 Hz, 1H, H4), 7.04 (d, J = 8.3 Hz, 1H, H6′), 6.83 (d, J = 2.4 Hz, 1H, H3′), 6.69 (ddd, J = 8.1, 7.1, 1.1 Hz, 1H, H5), 6.62-6.53 (m, 2H, H3, H5′), 5.21 (br s, 2H, NH_2). ESI MS (m / z): 263.2 $[M + H]^+$.

*8-Chloro-10,11-dihydro-5*H*-dibenzo[*b,e*][1,4]diazepin-11-one (**4**)*. Following the procedure of Capuano et al.,[46] a mixture of **3** (9.79 g, 37.3 mmol) and *o*-xylene (250 mL) was heated under Dean-Stark conditions for 96 h. The reaction mixture was then cooled, evaporated to dryness in vacuo and the resulting residue washed with hot aqueous ammonia (2 M, 2 × 50 mL). **4** was recrystallized from acetone / water yielding the title compound as off-white platelets (5.31 g, 21.7 mmol, 58%), mp* 234-236 °C (lit.[46] 232-233 °C). ^{1}H NMR (300 MHz, d_6-acetone) δ 9.00 (br s, 1H, H10), 7.84 (d, J = 7.9 Hz, 1H, H1), 7.36 (ddd, J = 8.0, 7.3, 1.7 Hz, 1H, H3), 7.30 (br s, 1H, H5), 7.15 (d, J = 2.2 Hz, 1H, H9), 7.10-6.91 (m, 4H, H2, H4, H6, H7). ESI MS (m / z): 245.4 $[M + H]^+$.

*8-Chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepine, clozapine (**1**)*. Following the procedure of Capuano et al.,[54] to a solution of *N*-methylpiperazine (4.53 mL, 40.8 mmol) in anhydrous anisole (20 mL) under nitrogen, was added titanium tetrachloride in toluene (1.0 M, 9 mL, 9.00 mmol). The mixture was warmed to 50-55 °C to which a hot solution of **4** (2.07 g, 8.46 mmol) in anhydrous anisole (170 mL) was added and heated at reflux overnight. TLC indicated unreacted **4** therefore additional *N*-methylpiperazine (4.53 mL, 40.8 mmol) was added to the reaction mixture, followed by titanium tetrachloride in toluene (1.0 M, 9 mL, 9.00 mmol). The reaction mixture was heated at reflux for a further 4 h, after which point the TLC indicated virtually complete consumption of starting material. The reaction was cooled, and concentrated in vacuo. The residue was partitioned between aqueous sodium hydroxide solution (1 M, 100 mL) and ethyl acetate (100 mL) and filtered. The organic layer was separated and the aqueous phase was extracted with ethyl acetate (3 × 50 mL). The combined organic extracts were washed with water (2 × 50 mL) and saturated

brine (50 mL), dried over anhydrous sodium sulfate and evaporated to dryness in vacuo. The resulting brown residue was purified by flash column chromatography (33% methanol / ethyl acetate) yielding a bright yellow foam, which was recrystallized from acetone / water to afford clozapine (**1**) (2.60 g, 7.94 mmol, 94%) as bright yellow prisms, mp* 183-184 °C (lit.[54] 181-183 °C). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.33-7.24 (m, 2H, H1, H3), 7.06 (d, *J* = 2.5 Hz, 1H, H9), 7.01 (ddd, *J* = 7.7, 7.4, 1.2 Hz, 1H, H2), 6.84-6.79 (m, 2H, H4, H7), 6.60 (d, *J* = 8.3 Hz, 1H, H6), 4.88 (s, 1H, NH), 3.46 (m, 4H, H2′, H6′), 2.54 (m, 4H, H3′, H5′), 2.34 (s, 3H, CH_3). ESI MS (*m / z*): 327.2 $[M + H]^+$.

*8-Chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-amine (**5**)*. Following the method of Su et al.,[50] a solution of **1** (1.25 g, 3.82 mmol) in dichloromethane (50 mL) and isoamyl nitrite (25 mL) was stirred at room temperature for 3 h. The solvent was evaporated in vacuo, yielding a pale yellow oil (8-chloro-11-(4-methylpiperazin-1-yl)-5-nitroso-5*H*-dibenzo[*b*,*e*][1,4]diazepine) that was carried on without further purification. A solution of the *N*-nitroso intermediate (1.36 g, 3.82 mmol) in glacial acetic acid (20 mL) was added dropwise to a suspension of zinc metal (6.25 g) in glacial acetic acid (100 mL) over 1 h at 10 to 15 °C. Additional zinc (0.5 g) was added periodically to maintain the green/yellow color of the solution. After 3 h, the solution was filtered and concentrated in vacuo yielding a pale brown residue. Water (25 mL) was added to the residue and the pH was adjusted to 11 using sodium hydroxide pellets. The aqueous layer was extracted with dichloromethane (3 × 50 mL), washed with water (50 mL) and saturated brine (50 mL), dried over anhydrous sodium sulfate and evaporated to dryness in vacuo. **5** was purified by flash column chromatography (50% methanol / ethyl acetate) and recrystallized from dichloromethane / hexane (0.635 g, 1.86 mmol, 48%) to yield **5** as beige colored crystals, mp 148-149 °C. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.76 (ddd, *J* = 8.2, 1.0, 1.0 Hz, 1H, H1/H4), 7.37-7.30 (m, 2H, H3/H2, H4/H1), 7.21 (d, *J* = 8.4 Hz, 1H, H6), 7.10 (ddd, *J* = 8.0, 6.1, 1.7 Hz, 1H, H2/H3), 6.86 (d, *J* = 2.2 Hz, 1H, H9), 6.78 (dd, *J* = 8.4, 2.3 Hz, 1H, H7), 4.39 (s, 2H, NH_2), 3.54 (m,

4H, H2′, H6′), 2.66 (m, 4H, H3′, H5′), 2.39 (s, 3H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$) δ 153.3 (C_q), 143.7 (C_q), 142.0 (C_q), 133.6 (C_q), 127.6 (CH), 127.0 (CH), 124.1 (C_q), 121.5 (CH), 120.0 (CH), 118.2 (CH), 116.8 (CH), 116.0 (C_q), 110.6 (CH), 54.9 (CH_2), 49.5 (CH_2), 46.4 (CH_3). ESI MS (*m / z*): 342.4 $[M + H]^+$.

*8-Chloro-11-(piperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepine (*N*-desmethylclozapine,* ***6****).* According to the procedure by Olofson, R. A. et al.,[55] clozapine (**1**, 2.00 g, 6.13 mmol) was dissolved in anhydrous 1,2-dicholoroethane (30 mL), under nitrogen at room temperature. The reaction mixture was cooled to 0 °C and α-chloroethyl chloroformate (2.64 mL, 24.5 mmol) was added dropwise to the stirred solution over 15 min, (resulting in a yellow precipitate), and maintained at 0 °C for an additional 15 min. The reaction was warmed to room temperature, giving a deep red solution, which was heated at reflux, under nitrogen, for 22 h. The reaction mixture was concentrated in vacuo, the oily brown residue dissolved in methanol (HPLC grade, 30 mL) and heated at 50 °C for 2 h, cooled and concentrated in vacuo. The oily brown residue was partitioned between ethyl acetate (50 mL) and aqueous hydrochloric acid (1 M, 50 mL). The aqueous layer was separated and adjusted to pH 10 using sodium hydroxide pellets. The aqueous layer was extracted with ethyl acetate (3 × 50 mL), the combined organic layers washed with water (2 × 50 mL) and saturated brine (50 mL), dried over anhydrous sodium sulfate, filtered and evaporated to dryness in vacuo. The yellow foam was purified by column chromatography (10% methanol / chloroform), yielding a yellow foam (**6**, 1.33 g, 4.23 mmol, 69%). ^1H NMR (400 MHz, $CDCl_3$) δ 7.26 (ddd, *J* = 7.8, 7.3, 1.6 Hz, 1H, H3), 7.24 (dd, *J* = 7.6, 1.5 Hz, 1H, H1), 7.06 (d, *J* = 2.4 Hz, 1H, H9), 6.99 (ddd, *J* = 7.8, 7.3, 1.1 Hz, 1H, H2), 6.83-6.76 (m, 2H, H4, H7), 6.60 (d, *J* = 8.3 Hz, 1H, H6), 5.06 (s, 1H, H5), 3.42 (m, 4H, H2′, H6′), 2.93 (m, 4H, H3′, H5′), 2.40 (s, 1H, H4′). ^{13}C NMR (101 MHz, $CDCl_3$) δ 163.1 (C_q), 152.8 (C_q), 141.8 (C_q), 140.6 (C_q), 131.9 (CH), 130.2 (CH), 128.9 (C_q), 126.7 (CH), 123.4 (C_q), 123.08 (CH), 123.06 (CH), 120.2 (CH), 120.1 (CH), 48.5 (CH_2), 45.9 (CH_2). LCMS (*m / z*): 313.1 $[M + H]^+$.

tert-*Butyl (3-bromopropyl)carbamate*. Di-*tert*-butyl dicarbonate (2.10 g, 9.64 mmol) in dichloromethane (3 mL) was added to a mixture of 3-bromopropylamine hydrobromide (2.01 g, 9.17 mmol) in dichloromethane (50 mL) according to the procedure by Zlatev et al.[62] Triethylamine (1.40 mL, 10.0 mmol) was added to the mixture, which was stirred at room temperature for 1 h. The reaction mixture was diluted with dichloromethane (40 mL), washed with aqueous sodium bicarbonate (saturated, 2 × 50 mL) and saturated brine (50 mL) dried over anhydrous sodium sulfate, filtered and concentrated in vacuo. The light brown oil was purified by column chromatography (10% ethyl acetate / hexane), yielding a colorless oil (1.97 g, 8.27 mmol, 90%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 4.68 (s, 1H, NH), 3.44 (t, J = 6.5 Hz, 2H, H3), 3.27 (app q, J = 6.4 Hz, 2H, H1), 2.05 (app p, J = 6.5 Hz, 2H, H2), 1.44 (s, 9H, $C(CH_3)_3$).

tert-*Butyl (3-(4-(8-chloro-5*H*-dibenzo[*b,e*][1,4]diazepin-11-yl)piperazin-1-yl)propyl) carbamate (**7a**)*. *N*-Desmethylclozapine (**6**) (1.02 g, 3.27 mmol), sodium iodide (0.482 g, 3.22 mmol) and *N,N*-diisopropylethylamine (670 µL, 3.85 mmol) were combined in acetonitrile (30 mL), under nitrogen. *tert*-Butyl (3-bromopropyl)carbamate (0.850 g, 3.57 mmol) was dissolved, under nitrogen, in acetonitrile (2 mL) and added to the stirred mixture. The reaction mixture was heated at reflux for 2 h after which time additional *tert*-butyl (3-bromopropyl)carbamate (0.400 g, 1.68 mmol) was added. After an additional 1 h, the reaction mixture was cooled to room temperature and concentrated in vacuo. The orange oil was purified by column chromatography (5% methanol / chloroform), yielding a yellow foam (1.22 g, 2.60 mmol, 79%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.29 (ddd, J = 7.9, 7.4, 1.6 Hz, 1H, H3″), 7.25 (m, 1H, H1″), 7.06 (d, J = 2.4 Hz, 1H, H9″), 7.01 (ddd, J = 7.8, 7.4, 1.1 Hz, 1H, H2″), 6.82 (m, 1H, H4″), 6.81 (dd, J = 8.3, 2.4 Hz, 1H, H7″), 6.61 (d, J = 8.3 Hz, 1H, H6″), 5.26 (br s, 1H, CONH), 4.92 (s, 1H, H5″), 3.45 (m, 4H, H3′, H5′), 3.20 (app q, J = 6.3 Hz, 2H, H1), 2.53 (m, 4H, H2′, H6′), 2.45 (t, J = 6.8 Hz, 2H, H3), 1.68 (app p, J = 6.7 Hz, 2H, H2), 1.44 (s, 9H, $C(CH_3)_3$). ^{13}C NMR (101 MHz, $CDCl_3$) δ 162.8 (C_q), 156.2 (C_q),

152.9 (C_q), 142.0 (C_q), 140.5 (C_q), 132.0 (CH), 130.4 (CH), 129.2 (C_q), 126.9 (CH), 123.6 (C_q), 123.18 (CH), 123.15 (CH), 120.2 (CH), 120.1 (CH), 79.1 (C_q), 57.0 (CH_2), 53.3 (CH_2), 47.4 (CH_2), 40.0 (CH_2), 28.6 (CH_3), 26.6 (CH_2). HPLC: t_R 8.01 min, >99% purity (Method 1). LCMS (*m / z*): 470.1 $[M + H]^+$. HRMS (*m / z*): $C_{25}H_{33}ClN_5O_2^+$ requires $[M + H]^+$ 470.2317; found 470.2330.

*3-(4-(8-Chloro-5*H*-dibenzo[*b,e*][1,4]diazepin-11-yl)piperazin-1-yl)propan-1-amine (**7b**)*. To a stirred solution of **7a** (1.22 g, 2.60 mmol) in dichloromethane (5 mL) at room temperature was added trifluoroacetic acid (5 mL, 65.3 mmol). The reaction mixture was stirred for 2 h before being diluted with dichloromethane (50 mL). Aqueous potassium carbonate (50 mL, 1 M) was added pipette-wise to the stirred mixture. The aqueous layer was separated and extracted with dichloromethane (3 × 50 mL), the combined organic layers were washed with water (2 × 50 mL) and saturated brine (50 mL), dried over anhydrous sodium sulfate, filtered and concentrated in vacuo to yield **7b** as a yellow foam (0.901 g, 2.44 mmol, 94%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.28 (ddd, *J* = 7.9, 7.4, 1.6 Hz, 1H, H3″), 7.25 (dd, *J* = 8.1, 1.6 Hz, 1H, H1″), 7.06 (d, *J* = 2.4 Hz, 1H, H9″), 7.00 (ddd, *J* = 7.8, 7.3, 1.1 Hz, 1H, H2″), 6.81 (m, 1H, H4″), 6.80 (dd, *J* = 8.3, 2.4 Hz, 1H, H7″), 6.60 (d, *J* = 8.3 Hz, 1H, H6″), 4.97 (s, 1H, H5″), 3.46 (m, 4H, H3′, H5′), 2.77 (m, 2H, H1), 2.54 (m, 4H, H2′, H6′), 2.45 (m, 2H, H3), 1.66 (app p, *J* = 7.0 Hz, 2H, H2), 1.51 (br s, 2H, NH_2). ^{13}C NMR (101 MHz, $CDCl_3$) δ 162.9 (C_q), 152.8 (C_q), 141.9 (C_q), 140.6 (C_q), 131.9 (CH), 130.4 (CH), 129.1 (C_q), 126.8 (CH), 123.5 (C_q), 123.09 (CH), 123.08 (CH), 120.2 (CH), 120.1 (CH), 56.6 (CH_2), 53.3 (CH_2), 47.4 (CH_2), 40.9 (CH_2), 30.5 (CH_2). HPLC: t_R 6.26 min, 95% purity (Method 2). LCMS (*m / z*): 370.1 $[M + H]^+$. HRMS (*m / z*): $C_{20}H_{25}ClN_5^+$ requires $[M + H]^+$ 370.1793; found 370.1788.

*Glutaric anhydride (**12b**)*. Glutaric acid (5.06 g, 38.3 mmol) and acetic anhydride (10 mL) were heated at reflux for 1h, cooled and concentrated in vacuo. The resulting oil was distilled affording **12b** (2.95 g, 25.8 mmol, 67%) as a clear oil, which crystallized upon

cooling, bp 118-122 °C at 1.5 mmHg (lit.[63] bp 138°C at 7 mmHg), mp 53.4-55.7 °C (lit.[64] mp 53-55 °C). ^{1}H NMR (400 MHz, $CDCl_3$) δ 2.76 (t, *J* = 6.7 Hz, 4H, H3, H5), 2.03 (app p, *J* = 6.7 Hz, 2H, H4). ^{13}C NMR (101 MHz, $CDCl_3$) δ 167.0 (C_q), 29.9 (CH_2), 16.3 (CH_2).

4,4'-(Piperazine-1,4-diyl)bis(4-oxobutanoic acid) ***(13a)***. Following the procedure of Liu et al.,[56] pulverized piperazine (0.508 g, 5.90 mmol) was dissolved in toluene (6 mL). Succinic anhydride (1.16 g, 11.6 mmol) was suspended in toluene (12 mL) and added to the stirred piperazine solution. This solution was heated at reflux for 28 h, then cooled and concentrated in vacuo, yielding a white powder. The product was recrystallized from ethanol, yielding **13a** as off-white crystals (1.02 g, 3.57 mmol, 62%), mp 158.7-160.7 °C (lit.[56] 156.6-158.5 °C). ^{1}H NMR (400 MHz, D_2O) A mixture of amide rotamers.[56] δ 3.72-3.56 (m, 8H, H2′, H3′, H5′, H6′), 2.80-2.72 (m, 4H, H2), 2.69-2.61 (m, 4H, H3). ^{13}C NMR (101 MHz, D_2O) A mixture of amide rotamers.[56] δ 177.4 (C_q), 172.98 (C_q), 172.94 (C_q), 44.9 (CH_2), 44.7 (CH_2), 41.6 (CH_2), 41.4 (CH_2), 29.1 (CH_2), 27.7 (CH_2), 27.6 (CH_2).

5,5'-(Piperazine-1,4-diyl)bis(5-oxopentanoic acid) ***(13b)***. Piperazine (0.247 g, 2.87 mmol) was dissolved in anhydrous dioxane (2 mL), with gentle warming. Glutaric anhydride (**12b**, 0.664 g, 5.82 mmol), suspended in anhydrous dioxane (2 mL), was added to the stirred piperazine solution and heated at reflux for 24 h then cooled to room temperature. The white precipitate was collected by filtration (**13b**, 0.741 g, 2.36 mmol, 82%), mp 157.0-158.1 °C. ^{1}H NMR (400 MHz, D_2O) A mixture of amide rotamers.[56] δ 3.70-3.55 (m, 8H, H2′, H3′, H5′, H6′), 2.53 (app t, *J* = 7.6 Hz, 4H, H2), 2.44 (t, *J* = 7.2 Hz, 4H, H4), 1.89 (app p, *J* = 7.3 Hz, 4H, H3). ^{13}C NMR (101 MHz, MeOD) A mixture of amide rotamers.[56] δ 176.9 (C_q), 173.7 (C_q), 173.6 (C_q), 46.6 (CH_2), 46.3 (CH_2), 42.8 (CH_2), 42.4 (CH_2), 34.0 (CH_2), 33.2 (CH_2), 33.1 (CH_2), 21.6 (CH_2).

4,4'-(Ethane-1,2-diylbis(oxy))bis(4-oxobutanoic acid) ***(15)***. Following the procedure of Asay et al.,[57] toluene (20 mL) and ethylene glycol (2.1 mL, 37.7 mmol) were combined and a suspension of succinic anhydride (7.58 g, 75.7 mmol) in toluene (20 mL) was added to the

stirred solution. The reaction was heated under Dean-Stark conditions overnight, cooled and then concentrated in vacuo. The product was recrystallized from ethyl acetate / hexane, yielding **15** as white crystals (5.45 g, 20.8 mmol, 27%), mp 75.8-76.4 °C (lit.[57] > 80 °C (decomposed)). ^{1}H NMR (400 MHz, D_2O) δ 4.36 (s, 4H, H1′, H2′), 2.70 (s, 8H, H2, H3). ^{13}C NMR (101 MHz, D_2O) δ 176.8 (C_q), 174.7 (C_q), 62.8 (CH_2), 28.8 (CH_2), 28.7 (CH_2).

Synthesis of bivalent ligands

General Procedure A: General procedure for preparation of clozapine bivalent ligands (9a-f). All glassware used in the following procedure was flame dried and cooled under nitrogen. The dicarboxylic acid (0.342 to 0.428 mmol) was suspended in dry dichloromethane (2 mL) at room temperature, under a nitrogen atmosphere. Oxalyl chloride (2.2 equivalents) and *N*,*N*-dimethylformamide (1 drop) were added to the solution, which was stirred for 1 h at room temperature. To this mixture was added a solution of clozapine (**1**, 1.8 equivalents) and pyridine (2.8 equivalents), and in some cases *N*,*N*-diisopropylethylamine (2.0 to 2.5 equivalents), in dry dichloromethane (2 mL). The reaction was monitored by TLC and, after 1 h, if a significant amount of **1** remained, an additional 0.5 equivalents of diacid chloride in dry dichloromethane (1 mL) was added to the stirred solution. The reaction mixture was stirred overnight at room temperature, under nitrogen. The reaction mixture was partitioned between dichloromethane (10 mL) and aqueous potassium carbonate (10 mL, 1 M). The organic layer was separated and the aqueous layer was extracted with dichloromethane (3 × 10 mL). The combined organic extracts were washed with water (2 × 10 mL) and saturated brine (10 mL), dried over anhydrous sodium sulfate, filtered and evaporated to dryness in vacuo, yielding a pale brown oil. The product was purified by pre-adsorption onto coarse silica, followed by flash column chromatography. Compounds were typically purified using 1-3 columns (refer to specific experimentals for details of eluent).

*1,8-Bis(8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl)octane-1,8-dione (**9a**).* 1,8-Octanedioic acid (0.060 g, 0.342 mmol) and oxalyl chloride were reacted, followed by the addition of **1** (0.209 g, 0.639 mmol) and pyridine (0.080 mL, 0.991 mmol) as per general procedure A. Additional 1,8-octanedioyl dichloride (0.037 g, 0.175 mmol) was added. Column chromatography conditions: column 1 (20% methanol / acetone, until clozapine eluted then 10% methanol / chloroform), column 2 (gradient elution: from 5% methanol / chloroform to 10% methanol / chloroform, increasing in 1% increments), column 3 (1% ammonia / 9 % methanol / chloroform). Yielding **9a** as a white foam (0.069 g, 0.087 mmol, 27%). ^{1}H NMR (400 MHz, $CDCl_3$, 320 K) δ 7.49 (ddd, J = 7.9, 7.0, 1.9 Hz, 2H, H3′), 7.41-7.30 (m, 6H, H1′, H2′, H4′), 7.14 (d, J = 2.4 Hz, 2H, H9′), 7.09 (d, J = 8.4 Hz, 2H, H6′), 6.95 (m, 2H, H7′), 3.70 (m, 4H, H2″a, H6″a), 3.47 (m, 4H, H2″b, H6″b), 2.46 (m, 4H, H3″a, H5″a), 2.36 (m, 4H, H3″b, H5″b), 2.32-2.20 (m, 8H, CH_3, H2a, H7a), 2.09 (m, 2H, H2b, H7b), 1.51 (m, 4H, H3, H6), 1.18 (m, 4H, H4, H5). ^{13}C NMR (101 MHz, $CDCl_3$, 320 K) δ 173.8 (C_q), 160.7 (C_q), 146.5 (C_q), 145.2 (C_q), 134.1 (C_q), 133.8 (C_q), 132.1 (CH), 129.2 (CH), 127.93 (CH), 127.87 (CH), 127.1 (C_q), 126.3 (2 × CH), 123.4 (CH), 55.0 (CH_2), 47.2 (CH_2), 46.19 (CH_3), 33.5 (CH_2), 29.0 (CH_2), 25.0 (CH_2). HPLC: t_R 11.76 min, 98% purity (Method 2). LCMS (m / z): 791.1 $[M + H]^+$, 396.2 $[M + 2H]^{2+}$. HRMS (m / z): $C_{44}H_{49}Cl_2N_8O_2^+$ requires $[M + H]^+$ 791.3350; found 791.3354.

General Procedure B: General procedure for preparation of hydrazide bivalent ligands (10a-g). All glassware used in the following procedure was flame dried and cooled under nitrogen. The diacid chloride was prepared using one of two methods. In method A, the dicarboxylic acid (14.8 mmol) was refluxed in neat thionyl chloride (10.0 mL, 138 mmol) for 90 min, then concentrated in vacuo before further use. In method B, the dicarboxylic acid (0.242 to 0.351 mmol) was suspended in dry dichloromethane (2 mL) at room temperature, under a nitrogen atmosphere. Oxalyl chloride (2.2 equivalents) and *N*,*N*-dimethylformamide (one drop) were added to the solution, which was stirred for 1 h at

room temperature. To the diacid chloride from Method A or B, was added a solution of 8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b*,*e*][1,4]diazepin-5-amine (**5**, 1.8 equivalents) and pyridine (2.8 equivalents), and in some cases *N*,*N*-diisopropylethylamine (2.5 equivalents), in dry dichloromethane (5 mL). The reaction was monitored by TLC and, after 1 h, if some **5** remained, an additional 0.5 equivalents of diacid chloride in dry dichloromethane (1 mL) was added to the reaction mixture and stirred overnight at room temperature. The reaction mixture was partitioned between dichloromethane (10 mL) and aqueous potassium carbonate (10 mL, 1 M). The organic layer was separated and the aqueous layer was extracted with dichloromethane (3 × 10 mL). The combined organic layers were washed with water (2 × 10 mL) and saturated brine (10 mL), dried over anhydrous sodium sulfate, filtered and evaporated to dryness in vacuo, yielding a pale brown oil. The product was purified by pre-adsorption onto coarse silica, followed by flash column chromatography. Compounds were typically purified using 1-3 columns (refer to specific experimentals for details of eluent).

N^{1},N^{6}*-Bis(8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl)hexanediamide (**10a**)*. Adipoyl dichloride was prepared according to Method A in general procedure B. A portion adipoyl dichloride (0.065 g, 0.357 mmol) in dry dichloromethane (5 mL) was added to **5** (0.227 g, 0.665 mmol) and pyridine (0.080 mL, 0.991 mmol) in dry dichloromethane (5 mL) according to general procedure B. Additional adipoyl dichloride (0.016 g, 0.090 mmol) was added. The crude product was recrystallized from dichloromethane / hexane to give **10a** as off-white crystals (0.209 g, 0.264 mmol, 79%), mp 185.3-186.4 °C. ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.46 (s, 2H, NH), 8.54 (d, *J* = 2.3 Hz, 2H, H9′), 7.78 (m, 2H, H1′/H4′), 7.44-7.37 (m, 4H, H3′/H2′, H4′/H1′), 7.36 (d, *J* = 8.6 Hz, 2H, H6′), 7.15 (m, 2H, H2′/H3′), 7.13 (dd, *J* = 8.5, 2.4 Hz, 2H, H7′), 3.57 (m, 8H, H2″, H6″), 2.68 (m, 8H, H3″, H5″), 2.39 (s, 6H, CH_3), 2.25 (m, 4H, H2, H5), 1.67 (m, 4H, H3, H4). ^{13}C NMR (75 MHz, $CDCl_3$) δ 170.6 (C_q), 153.9 (C_q), 141.9 (C_q), 133.8 (C_q), 132.8 (C_q), 128.3

(CH), 126.9 (C_q), 124.6 (CH), 123.7 (CH), 122.7 (CH), 121.9 (CH), 120.8 (CH), 116.3 (C_q), 110.7 (CH), 54.8 (CH_2), 49.4 (CH_2), 46.4 (CH_3), 37.8 (CH_2), 24.8 (CH_2). HPLC: t_R 9.72 min, >99% purity (Method 1). LCMS (*m / z*): 793.1 $[M + H]^+$, 397.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{42}H_{47}Cl_2N_{10}O_2^+$ requires $[M + H]^+$ 793.3255; found 793.3256.

*4,4'-(Piperazine-1,4-diyl)bis(*N*-(8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl)-4-oxobutanamide) (**14a**).* 4,4'-(Piperazine-1,4-diyl)bis(4-oxobutanoic acid) (**13a**, 0.090 g, 0.316 mmol) and oxalyl chloride were reacted, followed by the addition of **5** (0.191 g, 0.558 mmol) and pyridine (70 µL, 0.867 mmol) as per general experimental B. Additional 4,4'-(piperazine-1,4-diyl)bis(4-oxobutanoyl chloride) (0.052 g, 0.161 mmol) was added. Column chromatography conditions: 1% ammonia / 4% methanol / chloroform. Yielded **14a** as an off-white foam (0.063 g, 0.067 mmol, 24%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.70 (s, 2H, NH), 8.56 (s, 2H, H9″), 7.79 (ddd, *J* = 8.2, 0.9, 0.9 Hz, 2H, H1″/H4″), 7.42-7.36 (m, 4H, H3″/H2″, H4″/H1″), 7.36 (d, *J* = 8.5 Hz, 2H, H6″), 7.15 (m, 2H, H2″/H3″), 7.13 (dd, *J* = 8.5, 2.4 Hz, 2H, H7″), 3.59 (m, 8H, H2‴, H6‴), 3.55-3.46 (m, 4H, piperazine spacer), 3.42-3.36 (m, 4H, piperazine spacer), 2.69 (m, 8H, H3‴, H5‴), 2.64 (s, 8H, H2′, H3′), 2.40 (s, 6H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$) A mixture of amide rotamers.[56] δ 170.3 (C_q), 170.1 (C_q), 170.0 (C_q), 153.9 (C_q), 141.9 (C_q), 133.9 (C_q), 132.9 (C_q), 128.2 (CH), 126.9 (C_q), 124.8 (CH), 123.7 (CH), 122.6 (CH), 121.8 (CH), 120.7 (CH), 116.4 (C_q), 110.7 (CH), 54.8 (CH_2), 49.4 (CH_2), 46.3 (CH_3), 45.2 (CH_2), 45.0 (CH_2), 41.53 (CH_2), 41.45 (CH_2), 32.7 (CH_2), 32.6 (CH_2), 28.0 (CH_2). HPLC: t_R 7.22 min, 96% purity (Method 2). LCMS (*m / z*): 933.2 $[M + H]^+$, 467.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{48}H_{55}Cl_2N_{12}O_4^+$ requires $[M + H]^+$ 933.3841; found 933.3803.

*Ethane-1,2-diyl bis(4-((8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl)amino) -4-oxobutanoate) (**16**).* 4,4'-(Ethane-1,2-diylbis(oxy))bis(4-oxobutanoic acid) (**15**, 0.090 g, 0.343 mmol) and oxalyl chloride were reacted, followed by the addition of **5** (0.204 g, 0.596 mmol) and pyridine (76 µL, 0.942 mmol) as per general experimental B.

Column chromatography conditions: column 1 (0.5% ammonia / 2.5% methanol / chloroform), column 2 (gradient elution: from 2% methanol / chloroform to 5% methanol / chloroform, increasing in 1% increments). Yielded **16** as an off-white foam (0.124 g, 0.136 mmol, 46%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.64 (s, 2H, NH), 8.52 (d, *J* = 2.3 Hz, 2H, H9″), 7.78 (ddd, *J* = 8.3, 0.9, 0.9 Hz, 2H, H1″/H4″), 7.43-7.36 (m, 4H, H3″/H2″, H4″/H1″), 7.36 (d, *J* = 8.5 Hz, 2H, H6″), 7.14 (ddd, *J* = 8.1, 6.3, 1.8 Hz, 2H, H2″/H3″), 7.12 (dd, *J* = 8.5, 2.4 Hz, 2H, H7″), 4.16 (s, 4H, H1, H2), 3.58 (m, 8H, H2‴, H6‴), 2.72 (m, 8H, H3‴, H5‴), 2.66 (m, 4H, H2′/H3′), 2.57 (m, 4H, H3′/H2′), 2.42 (s, 6H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.3 (C_q), 169.4 (C_q), 153.7 (C_q), 141.8 (C_q), 133.6 (C_q), 132.7 (C_q), 128.3 (CH), 126.9 (C_q), 124.6 (CH), 123.7 (CH), 122.6 (CH), 121.8 (CH), 120.7 (CH), 116.3 (C_q), 110.6 (CH), 62.3 (CH_2), 54.6 (CH_2), 49.2 (CH_2), 46.1 (CH_3), 32.2 (CH_2), 28.9 (CH_2). HPLC: t_R 10.17 min, >99% purity (Method 1). LCMS (*m / z*): 909.1 $[M + H]^+$, 455.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{46}H_{51}Cl_2N_{10}O_6^+$ requires $[M + H]^+$ 909.3365; found 909.3328.

General Procedure C: General procedure for preparation of clozapine propylamine bivalent ligands (11a-g). All glassware used in the following procedure was flame dried and cooled under nitrogen. The dicarboxylic acid (0.169 to 0.184 mmol) was suspended in dry dichloromethane (2 mL) at room temperature, under a nitrogen atmosphere. Oxalyl chloride (2.2 equivalents) and *N*,*N*-dimethylformamide (one drop) were added to the solution, which was stirred for 1 h at room temperature. To this mixture was added a solution of **7b** (2.0 equivalents) and pyridine (2.8 to 6.8 equivalents), and in some cases *N*,*N*-diisopropylethylamine (2.5 equivalents) or anhydrous potassium carbonate (2.0 equivalents), in dry dichloromethane (3 mL). The reaction was monitored by TLC and, after 1 h, if any **7b** remained, an additional 0.5 equivalents of diacid chloride in dry dichloromethane (1 mL) was added to the solution and the reaction mixture stirred for an additional 2 to 5 h at room temperature. The reaction mixture was partitioned between ethyl acetate (30 mL) and aqueous sodium hydroxide (30 mL, 1 M). The organic layer was separated and the aqueous

layer was extracted with ethyl acetate (3 × 30 mL). The combined organic layers were washed with water (2 × 30 mL) and saturated brine (30 mL), dried over anhydrous sodium sulfate, filtered and evaporated to dryness in vacuo, yielding a yellow oil. The product was purified by pre-adsorption onto coarse silica, followed by flash column chromatography. Compounds were typically purified using 1-3 columns (refer to specific experimentals for details of eluent).

N^{1},N^{6}-*Bis(3-(4-(8-chloro-5*H*-dibenzo[*b,e*][1,4]diazepin-11-yl)piperazin-1-yl)propyl)adipamide (**11a**)*. 1,6-Hexanedioic acid (0.027 g, 0.184 mmol) and oxalyl chloride were reacted, followed by the addition of **7b** (0.130 g, 0.352 mmol) and pyridine (0.040 mL, 0.496 mmol) as per general procedure C. Additional 1,6-hexanedioyl dichloride (0.016 g, 0.089 mmol) was added. Column chromatography conditions: 10% methanol / chloroform. Yielded **11a** as a yellow foam (0.068 g, 0.080 mmol, 46%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.29 (m, 2H, H3‴), 7.24 (dd, *J* = 7.7, 1.4 Hz, 2H, H1‴), 7.08 (br t, *J* = 5.0 Hz, 2H, CONH), 7.05 (d, *J* = 2.4 Hz, 2H, H9‴), 7.00 (ddd, *J* = 7.6, 7.5, 1.0 Hz, 2H, H2‴), 6.86 (dd, *J* = 8.0, 0.6 Hz, 2H, H4‴), 6.82 (dd, *J* = 8.3, 2.4 Hz, 2H, H7‴), 6.66 (d, *J* = 8.3 Hz, 2H, H6‴), 5.26 (s, 2H, H5‴), 3.47 (m, 8H, H3″, H5″), 3.29 (td, *J* = 5.8, 5.8 Hz, 4H, H1′), 2.58 (m, 8H, H2″, H6″), 2.52 (t, *J* = 6.4 Hz, 4H, H3′), 2.17 (m, 4H, H2, H5), 1.69 (app p, *J* = 6.5 Hz, 4H, H2′), 1.64 (m, 4H, H3, H4). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.9 (C_q), 163.1 (C_q), 153.1 (C_q), 141.7 (C_q), 140.8 (C_q), 132.2 (CH), 130.3 (CH), 129.1 (C_q), 126.8 (CH), 123.5 (CH), 123.3 (C_q), 123.2 (CH), 120.4 (2 × CH), 57.2 (CH_2), 53.0 (CH_2), 47.3 (CH_2), 39.1 (CH_2), 36.5 (CH_2), 25.3 (CH_2), 25.2 (CH_2). HPLC: t_R 7.93 min, >99% purity (Method 1). LCMS (*m / z*): 849.2 $[M + H]^+$, 425.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{46}H_{55}Cl_2N_{10}O_2^+$ requires $[M + H]^+$ 849.3881; found 849.3881.

Synthesis of monovalent ligands

*1-(8-Chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl)decan-1-one (17).* **1** (0.103 g, 0.315 mmol) was dissolved in dry dichloromethane (3 mL), at room temperature, under a nitrogen atmosphere. Pyridine (37 µL, 0.459 mmol) and *N*,*N*-diisopropylethylamine (80 µL, 0.459 mmol) were added to the stirred solution followed by decanoyl chloride (127 µL, 0.612 mmol). After 2 h, the reaction mixture was diluted with dichloromethane (30 mL) and washed with aqueous hydrochloric acid (30 mL, 1 M). The organic layer was separated and the aqueous layer was further extracted with dichloromethane (2 × 30 mL). The combined organic layers were then washed with aqueous sodium hydroxide (3 × 30 mL), water (2 × 30 mL) and saturated brine (30 mL), dried over anhydrous sodium sulfate, filtered and concentrated in vacuo. The crude yellow oil was purified by flash column chromatography; column 1 (5% methanol / chloroform), column 2 (1% ammonia / 4% methanol / chloroform) to give the pure title compound **17** as a pale yellow oil (0.055 g, 0.114 mmol, 36%). ^{1}H NMR (400 MHz, $CDCl_3$, 320 K) δ 7.49 (ddd, *J* = 8.0, 7.0, 1.8 Hz, 1H, H3′), 7.40-7.30 (m, 3H, H1′, H2′, H4′), 7.15 (d, *J* = 2.4 Hz, 1H, H9′), 7.10 (d, *J* = 8.3 Hz, 1H, H6′), 6.96 (dd, *J* = 8.3, 2.3 Hz, 1H, H7′), 3.70 (m, 2H, H2″a, H6″a), 3.48 (m, 2H, H2″b, H6″b), 2.46 (m, 2H, H3″a, H5″a), 2.36 (m, 2H, H3″b, H5″b), 2.32-2.24 (m, 4H, NCH_3, H2a), 2.17 (m, 1H, H2b), 1.55 (m, 2H, H3), 1.33-1.17 (m, 12H, H4-H9), 0.87 (t, *J* = 7.0 Hz, 3H, H10). ^{13}C NMR (101 MHz, $CDCl_3$, 320 K) δ 173.7 (C_q), 160.5 (C_q), 146.3 (C_q), 145.0 (C_q), 133.9 (C_q), 133.6 (C_q), 131.9 (CH), 129.0 (CH), 127.7 (CH), 127.6 (CH), 126.9 (C_q), 126.1 (2 × CH), 123.1 (CH), 54.9 (CH_2), 46.9 (CH_2), 46.0 (CH_3), 33.5 (CH_2), 31.8 (CH_2), 29.3 (CH_2), 29.21 (CH_2), 29.19 (2 × CH_2), 25.1 (CH_2), 22.6 (CH_2), 14.0 (CH_3). HPLC: t_R 9.10 min, >99% purity (Method 2). LCMS (*m / z*): 481.2 $[M + H]^+$. HRMS (*m / z*): $C_{28}H_{38}ClN_4O^+$ requires $[M + H]^+$ 481.2729; found 481.2711.

N-*(8-Chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl)decanamide (18).* **5** (0.101 g, 0.295 mmol) was dissolved in dry dichloromethane (2 mL), at room

temperature, under a nitrogen atmosphere. Pyridine (70 µL, 0.869 mmol) was added to the stirred solution followed by decanoyl chloride (122 µL, 0.588 mmol). After 3 h, the reaction mixture was diluted with dichloromethane (30 mL) and worked up as described in the preparation of **17**. The crude brown oil was purified by flash column chromatography (2% methanol / chloroform) to give the pure title compound **18** as a pale brown oil (0.115 g, 0.232 mmol, 79%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.50 (s, 1H, NH), 8.59 (d, *J* = 2.3 Hz, 1H, H9′), 7.80 (ddd, *J* = 8.2, 0.9, 0.9 Hz, 1H, H1′/H4′), 7.44-7.37 (m, 2H, H3′/H2′, H4′/H1′), 7.36 (d, *J* = 8.5 Hz, 1H, H6′), 7.15 (m, 1H, H2′/H3′), 7.13 (dd, *J* = 8.5, 2.4 Hz, 1H, H7′), 3.57 (m, 4H, H2″, H6″), 2.68 (m, 4H, H3″, H5″), 2.41 (s, 3H, NCH_3), 2.27 (t, *J* = 7.5 Hz, 2H, H2), 1.63 (app p, *J* = 7.3 Hz, 2H, H3), 1.32-1.15 (m, 12H, H4-H9), 0.86 (t, *J* = 7.0 Hz, 3H, H10). ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.4 (C_q), 153.8 (C_q), 141.9 (C_q), 134.0 (C_q), 132.9 (C_q), 128.3 (CH), 126.9 (C_q), 124.7 (CH), 123.6 (CH), 122.7 (CH), 121.8 (CH), 120.7 (CH), 116.3 (C_q), 110.7 (CH), 54.8 (CH_2), 49.4 (CH_2), 46.4 (CH_3), 38.4 (CH_2), 31.9 (CH_2), 29.49 (CH_2), 29.45 (CH_2), 29.32 (CH_2), 29.25 (CH_2), 25.6 (CH_2), 22.7 (CH_2), 14.2 (CH_3). HPLC: t_R 9.17 min, 98% purity (Method 2). LCMS (*m / z*): 496.2 $[M + H]^+$. HRMS (*m / z*): $C_{28}H_{39}ClN_5O^+$ requires $[M + H]^+$ 496.2838; found 496.2831.

N-*(3-(4-(8-Chloro-5*H*-dibenzo[*b,e*][1,4]diazepin-11-yl)piperazin-1-yl)propyl)decanamide (**19**).* The clozapine propylamine (**7b**, 0.100 g, 0.269 mmol) was dissolved in dry dichloromethane (3 mL), at room temperature, under a nitrogen atmosphere. Pyridine (33 µL, 0.410 mmol) was added to the stirred solution followed by decanoyl chloride (71 µL, 0.342 mmol). After 1.5 h, additional decanoyl chloride was added (20 µL, 0.096 µmol). After 2 h, the reaction mixture was diluted with dichloromethane (30 mL) and worked up as described in the preparation of **17**. The crude yellow oil was purified by flash column chromatography (gradient elution: from 2% methanol / chloroform to 10% methanol / chloroform, increasing methanol in 2% increments) to give the pure title compound **19** as a yellow oil (0.089 g, 0.171 mmol, 63%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.30 (ddd, *J* = 7.9,

7.4, 1.6 Hz, 1H, H3‴), 7.26 (m, 1H, H1‴), 7.06 (d, *J* = 2.4 Hz, 1H, H9‴), 7.01 (ddd, *J* = 7.8, 7.4, 1.1 Hz, 1H, H2‴), 6.83 (ddd, *J* = 8.0, 1.1, 0.4 Hz, 1H, H4‴), 6.82 (dd, *J* = 8.3, 2.4 Hz, 1H, H7‴), 6.79 (br t, 1H, NH), 6.62 (dd, *J* = 8.4, 0.3 Hz, 1H, H6‴), 4.99 (s, 1H, H5‴), 3.45 (m, 4H, H3″, H5″), 3.35 (td, *J* = 6.0, 6.0 Hz, 2H, H1′), 2.55 (m, 4H, H2″, H6″), 2.51 (t, *J* = 6.3 Hz, 2H, H3′), 2.13 (m, 2H, H2), 1.70 (app p, *J* = 6.3 Hz, 2H, H2′), 1.61 (app p, *J* = 7.4 Hz, 2H, H3), 1.33-1.21 (m, 12H, H4-H9), 0.86 (t, *J* = 6.9 Hz, 3H, H10). ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.2 (C_q), 162.9 (C_q), 153.0 (C_q), 141.8 (C_q), 140.6 (C_q), 132.1 (CH), 130.3 (CH), 129.2 (C_q), 126.9 (CH), 123.4 (C_q), 123.4 (CH), 123.2 (CH), 120.3 (CH), 120.2 (CH), 57.7 (CH_2), 53.3 (CH_2), 47.5 (CH_2), 39.5 (CH_2), 37.2 (CH_2), 32.0 (CH_2), 29.6 (CH_2), 29.5 (CH_2), 29.5 (CH_2), 29.4 (CH_2), 26.0 (CH_2), 25.4 (CH_2), 22.8 (CH_2), 14.2 (CH_3). HPLC: t_R 7.29 min, >99% purity (Method 2). LCMS (*m / z*): 524.2 $[M + H]^+$. HRMS (*m / z*): $C_{30}H_{43}ClN_5O^+$ requires $[M + H]^+$ 524.3151; found 524.3149.

Biological Assays

Cell Culture

Chinese hamster ovary (CHO) FlpIn cells were stably transfected with the human D_2(long) dopamine receptor (D_2-CHOFlpIn). Cells were grown and maintained in DMEM containing 20mM HEPES, 5% fetal bovine serum and 200 µg/mL Hygromycin-B. Cells were maintained at 37 °C in a humidified incubator containing 5% CO_2, 95% O_2. For ERK1/2 phosphorylation assays, cells were seeded into 96-well silicon coated plates at a density of 50,000 cells/well. After 4 hours, cells were washed twice with PBS and then maintained in DMEM containing 20 mM HEPES for at least 16 hours before assaying.

Radioligand Binding Studies

When cells were approximately 90% confluent, they were harvested and centrifuged (300 *g*, 3 min). The resulting pellet was resuspended in assay buffer (20 mM HEPES, 6 mM $MgCl_2$, 1 mM EGTA, 1 mM EDTA; pH 7.4), and the centrifugation procedure repeated. The intact

cell pellet was then resuspended in assay buffer and homogenised using a Polytron homogeniser for three 10-second intervals on the maximum setting, with 30-second periods on ice between each burst. The homogenate was made up to 30 mL and centrifuged (1,000 *g*, 10 min, 25 °C), the pellet discarded and the supernatant recentrifuged at 30,000 *g* for 1 hour at 4 °C. The resulting pellet was resuspended in 5 mL assay buffer and the protein content determined using the method of Bradford.[65] The homogenate was then separated into 1 mL aliquots and stored frozen at -80 °C until required for radioligand binding assays. Membrane homogenates (5 μg/mL) were incubated in a 1 mL total volume of assay buffer containing ascorbic acid (0.1 %), BSA (0.1 %), [^{3}H]spiperone (0.1 nM) and a range of concentrations of ligand for 3 hours at 37 °C. Non-specific binding was defined using 10 μM butaclamol. Incubation was terminated by rapid filtration through Whatman GF/C filters using a Brandell cell harvestor (Gaithersburg, MD). Filters were washed three times with 3 mL aliquots of assay buffer and dried before the addition of 4 mL of scintillation cocktail (Ultima Gold; Packard Bioscience, Meriden, CT). Radioactivity was determined using scintillation counting.

ERK1/2 phosphorylation

Dose-response experiments in the absence or presence of ligand were performed at 37 °C in a 200 μL total volume of DMEM containing 20 mM HEPES and 0.1 % ascorbic acid. Dose-response stimulation or inhibition curves were generated by exposure of the cells antagonist ligand for 120 min and then dopamine for 5 min. Stimulation of cells was terminated by the removal of media and the addition of 100 μL of SureFire™ lysis buffer to each well. The plate was agitated for 1-2 min. A 4:1 v/v dilution of Lysate:SureFire™ activation buffer was made in a total volume of 50 μL. A 1:100:120 v/v dilution of AlphaScreen™ beads:activated lysate mixture:SureFire™ reaction buffer in an 11 μL total volume was then transferred to a white opaque 384-well Proxiplate™ in the dark. This plate was then incubated in the dark at

37 °C for 1.5 hours after which time the fluorescence signal was measured by a Fusion-™ plate reader (PerkinElmer), using standard AlphaScreen™ settings.

Data analysis

Data of radioligand binding experiments were analyzed using the non-linear regression curve fitting program Prism 5 (GraphPad, San Diego, CA, USA). For the displacement of [^{3}H]spiperone data was fit using a one site model with a variable Hill slope with the following the equation;

$$Y = \frac{(top - bottom)x^{n_H}}{x^{nH} + IC_{50}^{n_H}} \qquad (1)$$

where Y denotes the percent specific binding, top and bottom denote the maximal and minimal asymptotes respectively, x denotes the inhibitor potency (midpoint location) parameter and n_H denotes the Hill slope factor. Assuming simple competition, IC_{50} values were converted to K_i values using the Cheng-Prusoff equation.[66] In the functional ERK1/2 assay, agonist concentration response curves were fitted to the following four-parameter Hill equation using Prism 5;

$$response = \frac{(top - bottom)}{1 + (10^{\log EC_{50}} / x)^{n_H}} \qquad (2)$$

where top represents the maximal asymptote of the concentration response curves, bottom represents the lowest asymptote of the concentration-response curves, $\log EC_{50}$ represents the logarithm of the agonist EC_{50}, x represents the concentration of the agonist and n_H represents the Hill slope. To determine the inhibitory potency of the various monovalent and bivalent ligands data was fit to the following equation:

$$response = \frac{(top - bottom)}{(1 + 10^{(X - \log IC_{50})})} \qquad (3)$$

where top represents the maximal asymptote of the concentration response curves, bottom row represents the lowest asymptote of the concentration response curves, $\log IC_{50}$ represents the logarithm of the antagonist IC_{50}, x represents the concentration of the agonist and the

Hill slope is assumed to be unity. Data shown are the mean ± SEM of at least 3 separate experiments performed in duplicate.

Acknowledgements

F.M.M. is a recipient of an Australian Postgraduate Award (APA) scholarship. J.R.L. is a Monash University Larkins Fellow. A.C. is a Senior Research Fellow of the National Health and Medical Research Council of Australia (NHMRC). Portions of this work were funded by NHMRC Program Grant No. 519461 and Project Grant No. 1011920.

Supporting Information

Chemistry experimental and HPLC traces for monovalent and bivalent ligands. This material is available free of charge via the Internet at http://pubs.acs.org.

References

1. Portoghese, P. S.; Larson, D. L.; Sayre, L. M.; Yim, C. B.; Ronsisvalle, G.; Tam, S. W.; Takemori, A. E. Opioid agonist and antagonist bivalent ligands. The relationship between spacer length and selectivity at multiple opioid receptors. *J. Med. Chem.* **1986,** *29*, 1855-1861.
2. Portoghese, P. S. From models to molecules: Opioid receptor dimers, bivalent ligands, and selective opioid receptor probes. *J. Med. Chem.* **2001,** *44*, 2259-2269.
3. Berque-Bestel, I.; Lezoualc'h, F.; Jockers, R. Bivalent ligands as specific pharmacological tools for G protein-coupled receptor dimers. *Curr. Drug Discovery Technol.* **2008,** *5*, 312-318.
4. Shonberg, J.; Scammells, P. J.; Capuano, B. Design strategies for bivalent ligands targeting GPCRs. *ChemMedChem* **2011,** *6*, 963-974.
5. Bhushan, R. G.; Sharma, S. K.; Xie, Z.; Daniels, D. J.; Portoghese, P. S. A bivalent ligand (KDN-21) reveals spinal δ and κ opioid receptors are organized as heterodimers that give rise to δ_1 and κ_2 phenotypes. Selective targeting of δ-κ heterodimers. *J. Med. Chem.* **2004,** *47*, 2969-2972.
6. Daniels, D. J.; Kulkarni, A.; Xie, Z.; Bhushan, R. G.; Portoghese, P. S. A bivalent ligand (KDAN-18) containing δ-antagonist and κ-agonist pharmacophores bridges δ_2 and κ_1 opioid receptor phenotypes. *J. Med. Chem.* **2005,** *48*, 1713-1716.
7. Portoghese, P. S. Bivalent ligands and the message-address concept in the design of selective opioid receptor antagonists. *Trends Pharmacol. Sci.* **1989,** *10*, 230-235.
8. Mohr, K.; Tränkle, C.; Kostenis, E.; Barocelli, E.; De Amici, M.; Holzgrabe, U. Rational design of dualsteric GPCR ligands: Quests and promise. *Br. J. Pharmacol.* **2010,** *159*, 997-1008.
9. Narlawar, R.; Lane, J. R.; Doddareddy, M.; Lin, J.; Brussee, J.; IJzerman, A. P. Hybrid ortho/allosteric ligands for the adenosine A_1 receptor. *J. Med. Chem.* **2010,** *53*, 3028-3037.
10. Valant, C.; Sexton, P. M.; Christopoulos, A. Orthosteric/allosteric bitopic ligands: Going hybrid at GPCRs. *Mol. Interventions* **2009,** *9*, 125-135.
11. Morphy, R.; Kay, C.; Rankovic, Z. From magic bullets to designed multiple ligands. *Drug Discov. Today* **2004,** *9*, 641-651.
12. Rovira, X.; Pin, J.-P.; Giraldo, J. The asymmetric/symmetric activation of GPCR dimers as a possible mechanistic rationale for multiple signalling pathways. *Trends Pharmacol. Sci.* **2010,** *31*, 15-21.
13. Smith, N. J.; Milligan, G. Allostery at G protein-coupled receptor homo- and heteromers: Uncharted pharmacological landscapes. *Pharmacol. Rev.* **2010,** *62*, 701-725.

14. Bobrovnik, S. A. The influence of rigid or flexible linkage between two ligands on the effective affinity and avidity for reversible interactions with bivalent receptors. *J. Mol. Recognit.* **2007,** *20*, 253-262.

15. Daniels, D. J.; Lenard, N. R.; Etienne, C. L.; Law, P.-Y.; Roerig, S. C.; Portoghese, P. S. Opioid-induced tolerance and dependence in mice is modulated by the distance between pharmacophores in a bivalent ligand series. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102*, 19208-19213.

16. Zhang, S.; Yekkirala, A.; Tang, Y.; Portoghese, P. S. A bivalent ligand (KMN-21) antagonist for μ/κ heterodimeric opioid receptors. *Bioorg. Med. Chem. Lett.* **2009,** *19*, 6978-6980.

17. Waldhoer, M.; Fong, J.; Jones, R. M.; Lunzer, M. M.; Sharma, S. K.; Kostenis, E.; Portoghese, P. S.; Whistler, J. L. A heterodimer-selective agonist shows *in vivo* relevance of G protein-coupled receptor dimers. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102*, 9050-9055.

18. Soriano, A.; Ventura, R.; Molero, A.; Hoen, R.; Casadó, V.; Cortés, A.; Fanelli, F.; Albericio, F.; Lluís, C.; Franco, R.; Royo, M. Adenosine A_{2A} receptor-antagonist/dopamine D_2 receptor-agonist bivalent ligands as pharmacological tools to detect A_{2A}-D_2 receptor heteromers. *J. Med. Chem.* **2009,** *52*, 5590-5602.

19. Karellas, P.; McNaughton, M.; Baker, S. P.; Scammells, P. J. Synthesis of bivalent β_2-adrenergic and adenosine A_1 receptor ligands. *J. Med. Chem.* **2008,** *51*, 6128-6137.

20. Jacobson, K. A.; Xie, R.; Young, L.; Chang, L.; Liang, B. T. A novel pharmacological approach to treating cardiac ischemia. *J. Biol. Chem.* **2000,** *275*, 30272-30279.

21. Zhang, Y.; Gilliam, A.; Maitra, R.; Damaj, M. I.; Tajuba, J. M.; Seltzman, H. H.; Thomas, B. F. Synthesis and biological evaluation of bivalent ligands for the cannabinoid 1 receptor. *J. Med. Chem.* **2010,** *53*, 7048-7060.

22. Abadi, A. H.; Lankow, S.; Hoefgen, B.; Decker, M.; Kassack, M. U.; Lehmann, J. Dopamine/serotonin receptor ligands, part III [1]: Synthesis and biological activities of 7,7′-alkylene-bis-6,7,8,9,14,15-hexahydro-5*H*-benz[*d*]indolo[2, 3-*g*]azecines - application of the bivalent ligand approach to a novel type of dopamine receptor antagonist. *Arch. Pharm. (Weinheim)* **2002,** *335*, 367-373.

23. Huber, D.; Hubner, H.; Gmeiner, P. 1,1′-Disubstituted ferrocenes as molecular hinges in mono- and bivalent dopamine receptor ligands. *J. Med. Chem.* **2009,** *52*, 6860-6870.

24. Kuhhorn, J.; Hubner, H.; Gmeiner, P. Bivalent dopamine D_2 receptor ligands: Synthesis and binding properties. *J. Med. Chem.* **2011,** *54*, 4896-4903.

25. Melchiorre, C.; Angeli, P.; Lambrecht, G.; Mutschler, E.; Picchio, M. T.; Wess, J. Antimuscarinic action of methoctramine, a new cardioselective M-2 muscarinic receptor antagonist, alone and in combination with atropine and gallamine. *Eur. J. Pharmacol.* **1987,** *144*, 117-124.

26. Christopoulos, A.; Grant, M. K. O.; Ayoubzadeh, N.; Kim, O. N.; Sauerberg, P.; Jeppesen, L.; El-Fakahany, E. E. Synthesis and pharmacological evaluation of dimeric

muscarinic acetylcholine receptor agonists. *J. Pharmacol. Exp. Ther.* **2001,** *298*, 1260-1268.

27. Halazy, S.; Perez, M.; Fourrier, C.; Pallard, I.; Pauwels, P. J.; Palmier, C.; John, G. W.; Valentin, J.-P.; Bonnafous, R.; Martinez, J. Serotonin dimers: Application of the bivalent ligand approach to the design of new potent and selective 5-$HT_{1B/1D}$ agonists. *J. Med. Chem.* **1996,** *39*, 4920-4927.

28. Soulier, J. L.; Russo, O.; Giner, M.; Rivail, L.; Berthouze, M.; Ongeri, S.; Maigret, B.; Fischmeister, R.; Lezoualc'h, F.; Sicsic, S.; Berque-Bestel, I. Design and synthesis of specific probes for human 5-HT_4 receptor dimerization studies. *J. Med. Chem.* **2005,** *48*, 6220-6228.

29. Russo, O.; Berthouze, M.; Giner, M.; Soulier, J. L.; Rivail, L.; Sicsic, S.; Lezoualc'h, F.; Jockers, R.; Berque-Bestel, I. Synthesis of specific bivalent probes that functionally interact with 5-HT_4 receptor dimers. *J. Med. Chem.* **2007,** *50*, 4482-4492.

30. Seeman, P. Dopamine receptors and the dopamine hypothesis of schizophrenia. *Synapse* **1987,** *1*, 133-152.

31. Seeman, P.; Niznik, H. Dopamine receptors and transporters in Parkinson's disease and schizophrenia. *The FASEB Journal* **1990,** *4*, 2737-2744.

32. Seeman, P. Antiparkinson therapeutic potencies correlate with their affinities at dopamine $D2^{High}$ receptors. *Synapse* **2007,** *61*, 1013-1018.

33. Seeman, P. All roads to schizophrenia lead to dopamine supersensitivity and elevated dopamine $D2^{high}$ receptors. *CNS Neurosci. Ther.* **2011,** *17*, 118-132.

34. Canals, M.; Marcellino, D.; Fanelli, F.; Ciruela, F.; de Benedetti, P.; Goldberg, S. R.; Neve, K.; Fuxe, K.; Agnati, L. F.; Woods, A. S.; Ferre, S.; Lluis, C.; Bouvier, M.; Franco, R. Adenosine A_{2A}-dopamine D_2 receptor-receptor heteromerization. *J. Biol. Chem.* **2003,** *278*, 46741-46749.

35. Rocheville, M.; Lange, D. C.; Kumar, U.; Patel, S. C.; Patel, R. C.; Patel, Y. C. Receptors for dopamine and somatostatin: Formation of hetero-oligomers with enhanced functional activity. *Science* **2000,** *288*, 154-157.

36. Albizu, L.; Holloway, T.; González-Maeso, J.; Sealfon, S. C. Functional crosstalk and heteromerization of serotonin 5-HT_{2A} and dopamine D_2 receptors. *Neuropharmacology* **2011,** *61*, 770-777.

37. Rashid, A. J.; So, C. H.; Kong, M. M. C.; Furtak, T.; El-Ghundi, M.; Cheng, R.; O'Dowd, B. F.; George, S. R. D1–D2 dopamine receptor heterooligomers with unique pharmacology are coupled to rapid activation of G_q/11 in the striatum. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104*, 654-659.

38. Taylor, D. M.; Duncan-McConnell, D. Refractory schizophrenia and atypical antipsychotics. *J. Psychopharmacol.* **2000,** *14*, 409-418.

39. Kang, X.; Simpson, G. M. Clozapine: More side effects but still the best antipsychotic. *J. Clin. Psychiatry* **2010,** *71*, 982-983.

40. Feldman, J. Clozapine and agranulocytosis. *Psychiatr. Serv.* **1996,** *47*, 1177-1178.

41. Uetrecht, J.; Zahid, N.; Tehim, A.; Mim Fu, J.; Rakhit, S. Structural features associated with reactive metabolite formation in clozapine analogues. *Chem.-Biol. Interact.* **1997,** *104*, 117-129.

42. Uetrecht, J. P. Metabolism of clozapine by neutrophils: Possible implications for clozapine-induced agranulocytosis. *Drug Saf.* **1992,** *7*, 51-56.

43. Roth, B. L.; Sheffler, D. J.; Kroeze, W. K. Magic shotguns versus magic bullets: Selectively non-selective drugs for mood disorders and schizophrenia. *Nat. Rev. Drug Discovery* **2004,** *3*, 353-359.

44. Roth, B. L.; Sheffler, D.; Potkin, S. G. Atypical antipsychotic drug actions: Unitary or multiple mechanisms for 'atypicality'? *Clin. Neurosci. Res.* **2003,** *3*, 108-117.

45. Seeman, P.; Tallerico, T. Rapid release of antipsychotic drugs from dopamine D_2 receptors: An explanation for low receptor occupancy and early clinical relapse upon withdrawal of clozapine or quetiapine. *Am. J. Psychiatry* **1999,** *156*, 876-884.

46. Capuano, B.; Crosby, I. T.; Lloyd, E. J.; Taylor, D. A. Synthesis and preliminary pharmacological evaluation of 4´-arylmethyl analogues of clozapine. I. The effect of aromatic substituents. *Aust. J. Chem.* **2002,** *55*, 565-576.

47. Capuano, B.; Crosby, I. T.; Lloyd, E. J.; Podloucka, A.; Taylor, D. A. Synthesis and preliminary pharmacological evaluation of 4´-arylalkyl analogues of clozapine. II. Effect of the nature and length of the linker. *Aust. J. Chem.* **2003,** *56*, 875-886.

48. Capuano, B.; Crosby, I. T.; Lloyd, E. J.; Podloucka, A.; Taylor, D. A. Synthesis and preliminary pharmacological evaluation of 4′-arylalkyl analogues of clozapine. IV. The effects of aromaticity and isosteric replacement. *Aust. J. Chem.* **2008,** *61*, 930-940.

49. Sasikumar, T. K.; Burnett, D. A.; Zhang, H.; Smith-Torhan, A.; Fawzi, A.; Lachowicz, J. E. Hydrazides of clozapine: a new class of D_1 dopamine receptor subtype selective antagonists. *Bioorg. Med. Chem. Lett.* **2006,** *16*, 4543-4547.

50. Su, J.; Tang, H.; McKittrick, B. A.; Burnett, D. A.; Zhang, H.; Smith-Torhan, A.; Fawzi, A.; Lachowicz, J. Modification of the clozapine structure by parallel synthesis. *Bioorg. Med. Chem. Lett.* **2006,** *16*, 4548-4553.

51. Williams, D. P.; Pirmohamed, M.; Naisbitt, D. J.; Maggs, J. L.; Park, B. K. Neutrophil cytotoxicity of the chemically reactive metabolite(s) of clozapine: Possible role in agranulocytosis. *J. Pharmacol. Exp. Ther.* **1997,** *283*, 1375-1382.

52. Liu, Z. C.; Uetrecht, J. P. Clozapine is oxidized by activated human neutrophils to a reactive nitrenium ion that irreversibly binds to the cells. *J. Pharmacol. Exp. Ther.* **1995,** *275*, 1476-1483.

53. Mansour, A.; Meng, F.; Meador-Woodruff, J. H.; Taylor, L. P.; Civelli, O.; Akil, H. Site-directed mutagenesis of the human dopamine D_2 receptor. *Eur. J. Pharmacol., Mol. Pharmacol. Sect.* **1992,** *227*, 205-214.

54. Capuano, B.; Crosby, I. T.; Lloyd, E. J.; Neve, J. E.; Taylor, D. A. Aminimides as potential CNS acting agents. I. Design, synthesis, and receptor binding of 4'-aryl aminimide analogues of clozapine as prospective novel antipsychotics. *Aust. J. Chem.* **2007,** *60*, 673-684.

55. Olofson, R. A.; Martz, J. T.; Senet, J. P.; Piteau, M.; Malfroot, T. A new reagent for the selective, high-yield *N*-dealkylation of tertiary amines: Improved syntheses of naltrexone and nalbuphine. *J. Org. Chem.* **1984,** *49*, 2081-2082.

56. Liu, C.; Hudson, R. H. E.; Petersen, N. O. Convergent and sequential synthesis of dendritic, multivalent complexing agents. *Synthesis* **2002**, 1398-1406.

57. Asay, R. E.; Bradshaw, J. S.; Nielsen, S. F.; Thompson, M. D.; Snow, J. W.; Masihdas, D. R. K.; Izatt, R. M.; Christensen, J. J. The synthesis of novel macrocyclic multidentate compounds for dioxodioic acids. *J. Heterocycl. Chem.* **1977,** *14*, 85-90.

58. Baldessarini, R. J.; Campbell, A.; Webb, N. L.; Swindell, C. S.; Flood, J. G.; Shashoua, V. E.; Kula, N. S.; Hemamalini, S.; Bradley, M. O. Fatty acid derivatives of clozapine: Prolonged antidopaminergic activity of docosahexaenoylclozapine in the rat. *Neuropsychopharmacology* **2001,** *24*, 55-65.

59. Steinfeld, T.; Mammen, M.; Smith, J. A. M.; Wilson, R. D.; Jasper, J. R. A novel multivalent ligand that bridges the allosteric and orthosteric binding sites of the M_2 muscarinic receptor. *Mol. Pharmacol.* **2007,** *72*, 291-302.

60. Still, W. C.; Kahn, M.; Mitra, A. Rapid chromatographic technique for preparative separations with moderate resolution. *J. Org. Chem.* **1978,** *43*, 2923-2925.

61. Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. NMR chemical shifts of common laboratory solvents as trace impurities. *J. Org. Chem.* **1997,** *62*, 7512-7515.

62. Zlatev, I.; Giraut, A.; Morvan, F.; Herdewijn, P.; Vasseur, J.-J. δ-Di-carboxybutyl phosphoramidate of 2'-deoxycytidine-5'-monophosphate as substrate for DNA polymerization by HIV-1 reverse transcriptase. *Bioorg. Med. Chem.* **2009,** *17*, 7008-7014.

63. Cason, J.; Reist, E. J. Reactions of glutaryl dichloride with organometallic reagents. *J. Org. Chem.* **1958,** *23*, 1675-1679.

64. Cram, D. J.; Daeniker, H. U. Macro rings. V. Transannular effects in the 1,4-decamethylenebenzene series. *J. Am. Chem. Soc.* **1954,** *76*, 2743-2752.

65. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976,** *72*, 248-254.

66. Cheng, Y.-C.; Prusoff, W. H. Relationship between the inhibition constant (K_I) and the concentration of inhibitor which causes 50 per cent inhibition (I_{50}) of an enzymatic reaction. *Biochem. Pharmacol.* **1973,** *22*, 3099-3108.

4.3 Molecular modeling of the dopamine D_2 receptor homodimer

4.3.1 Introduction

There is increasing evidence that GPCRs can form dimers or higher order oligomers (Section 1.6)[1,2] and many researchers consider dimerization to be a common feature of the GPCR superfamily.[3,4] GPCR dimers may represent a novel therapeutic target, but current structural information regarding dimer formation is limited, consisting of atomic force microscopy[5] and transmission electron microscopy images of rhodopsin dimers,[6] as well as a crystal structure of the chemokine receptor CXCR4 homodimer (refer to Section 1.6.2).[7]

Biochemical techniques, such as site-directed mutagenesis studies, have also been used to investigate the dimerization interfaces of GPCRs. Dimerization interfaces for a number of aminergic GPCR dimers have been determined and these have been found to often involve transmembrane helices 1, 4 and 5. For example, cysteine cross-linking mutagenesis was used identify two dimerization interfaces in the D_2R homodimer; a TM1 interface, involving residues Tyr 1.34, Tyr 1.35, Leu 1.38 and Leu 1.41 in TM1 and residue Leu 438 in helix 8,[8] and a TM4 interface incorporating residues from six helical turns, including residues Arg 4.41, Val 4.44, Ile 4.48, Trp 4.50, Val 4.51, Leu 4.52, Phe 4.54, Thr 4.55, Ile 4.56, Cys 4.58, Pro 4.59, Leu 4.60, Leu 4.61 and Phe 4.62.[9,10] In other similar studies, a TM4 interface has also been implicated in dimerization interfaces of the 5-$HT_{2C}R$[11] 5-HT_4R[12] and $\alpha_{1B}AR$.[13,14] A TM1 interface has also been identified for $\alpha_{1B}AR$.[13,14] Additionally, cysteine residues on both TM3 and TM4 have been implicated in dimer formation for 5-HT_4R.[12]

Because there are very few experimentally determined structures of GPCR dimers (rhodopsin[5,6] and CXCR4[7]), molecular modeling has been used to develop a number of dimer models. Models of class A GPCR homodimers and heterodimers have been developed, including dopamine,[8,10] serotonin,[11,15-17] opioid[4,18] and adenosine[19] receptor dimers (refer to Section 1.6.6). Models of GPCR dimers have been developed using three broad approaches; protein-protein docking, experimentally determined structures (i.e. a

crystal structure or atomic force microscopy model) or the use of biochemical data (refer to Section 1.6.6). A few of these studies used the dimer models to determine the distance between the adjacent orthosteric binding sites.

Berque-Bestel and co-workers have developed a number of models of the 5-HT_4R homodimer[15,16] by using the protein-protein docking software GRAMM (global range molecular matching)[20,21] to predict dimerization interfaces. Their protein-protein docking predicted dimerization interfaces of TM2 / TM4 or TM4 / TM6,[16] with the minimum distance between adjacent orthosteric binding sites determined to be 22 Å.[15] Bruno et al. developed models of the 5-HT_{2A}R-mGluR2 heterodimer,[22] and the 5-HT_{2A}R homodimer[23] using protein-protein docking software, Rosetta++[24] in combination with visual comparison to atomic force microscopy model of rhodopsin dimer (PDB ID: 1N3M[25]). Gonzalez-Maeso et al. also built a model of the 5-HT_{2A}R-mGluR2 heterodimer[17] based on the atomic force microscopy model of rhodopsin.

Kim et al. built models of the A_3AR homodimer by superimposing protomers on to the atomic force microscopy model of the rhodopsin dimer (PDB ID; 1N3M[25]).[19] A number of symmetrical dimers with different contact interfaces were built (TM1 / TM2, TM2 / TM3, TM2 / TM4, TM3 / TM4, TM4 / TM5, TM5 / TM6, TM6 / TM7, and TM7 / TM1). Following molecular dynamics simulations, they identified TM4 / TM5 to be the most energetically favorable, followed by the TM1 / TM2 interface. This is consistent with the dimerization interfaces predicted for other class A GPCRs.

A model of the 5-HT_{2C}R homodimer was constructed by Mancia et al.[11] who used extensive cysteine cross-linking data to align the dimerization interfaces; namely TM1 and TM4 / TM5. In a similar manner, Filizola and co-workers, developed models of the D_2R homodimer[8,9] and the δ opioid receptor (DOR) homodimer[18] using cysteine cross-linking data to identify the dimerization interface and the guide the manual alignment of the two protomers. In an unpublished study by Daniels et al., the distance between adjacent

orthosteric sites for the μ opioid receptor homodimer has been reported to be approximately 27 Å.[4,26]

Of particular interest to our group, two models of the D_2R homodimer have been developed by Filizola and co-workers.[8,9] These dimer models used homology models of the D_2R monomeric structure built using either the rhodopsin (PDB ID: 1F88)[27] or β_2AR (PDB ID: 2RH1)[28] crystal structures as the modeling template. Dimers were assembled using mutagenesis data from cysteine cross-linking experiments to manually align the dimerization interfaces, where these studies indicated that dimerization or oligomerization occurred at the TM4 / TM5 or TM1 interface.[8-10]

As discussed in Section 1.7, Portoghese et al. have postulated that the two pharmacophores of a bivalent ligand can bind simultaneously to adjacent orthosteric sites of a GPCR dimer.[4,29] Manually docking a bivalent ligand into a dimer can be used to estimate the spacer length required to span the two orthosteric sites and assess whether there is any correlation with pharmacological results. Thus, in molecular modeling studies of the 5-HT_4R homodimer by Berque-Bestel and co-workers, a bivalent ligand with a 22 atom spacer that showed good biological activity was manually docked into the proposed GPCR dimer and spanned the distance between the two orthosteric sites.[15]

In Section 4.2, we report the design, synthesis and biological evaluation of clozapine homobivalent ligands with promising activity at the D_2R. This study aims to develop models of the D_2R homodimer, having varying dimerization interfaces, to determine the distance between the adjacent orthosteric sites and compare these results to the pharmacological data obtained for our clozapine homobivalent ligands. As such, we have constructed four models of the D_2R homodimer using methods similar to those used by Filizola and co-workers to develop models of the D_2R and DOR homodimers.[8-10,18] Cysteine cross-linking mutagenesis data was used to guide the manual alignment of the initial dimer model followed, in some cases, by protein-protein docking to optimize the dimer interface. We have further optimized

these four models using molecular dynamics simulations and determined the distance between the orthosteric sites. Finally, we have compared the results from the molecular dynamics simulations with the pharmacological data and manually docked a clozapine homobivalent ligand into one D_2R homodimer model.

4.3.2 Methods

Homology modeling and binding site optimization were performed using Schrödinger Suite 2011, through the Maestro interface.[30] Default settings were used for all programs, unless stated otherwise. The structure of clozapine was prepared using LigPrep 2.4.[31] Molecular dynamics simulations were performed using NAMD2 version 2.8.[32,33] Visual Molecular Dynamics (VMD) version 1.9, was used to visualize molecular dynamics trajectories.[34] Dimerization interfaces were analyzed using the DIMPLOT module of LIGPLOT.[35] GPCR residues are identified using the Ballesteros-Weinstein nomenclature,[36] except for loop regions, where the crystal structure numbering is used.

Bivalent ligands. To determine the maximum distances between ionizable piperazine nitrogens, extended models of the N4′ clozapine propylamine homobivalent ligands were built in Maestro.[30] Bivalent ligands were minimized using MacroModel,[37] using the OPLS 2005 force field and default settings. Distances were measured using the Maestro *measurement* tool.

Homology modeling. The D_2R homology model was built in Prime[38] using the crystal structure of D_3R (PDB ID: 3PBL)[39] as the template. Our previously developed multiple sequence alignment (Appendix 2), was used to align the D_2R to the D_3R structure. Highly conserved residues were anchored during model building (constraints were applied to the sequence alignment). ICL3 and the highly flexible *N*-terminus (residues 1 to 36) were not modeled.

Flexible receptor docking. Clozapine was docked into the D_2R model using Induced Fit Docking (IFD).[40,41] The orthosteric site was defined as a cubic 28 Å box, centered on the

centroid of Asp 3.32, Trp 6.48, Phe 6.52, and Tyr 7.43 residues. Residues Asp 3.32, Trp 6.48 and Tyr 7.43 were excluded from the binding site optimization and Glide SP (standard precision) was used for docking in the IFD workflow. Up to 20 complexes were collected. The top ranked D_2R-clozapine complex from IFD, according to the IFDScore (where the IFDScore is the sum of the GlideScore from the redocking step and 5% of the Prime energy score from the refinement step) was used as the protomer (monomeric unit) to build the D_2R homodimer.

Dimer building. Four models of the D_2R homodimer were built using one of two methods. In Method 1, cysteine cross-linking mutagenesis data[8-10] was used to guide the manual alignment of the D_2R protomers. Residues Tyr 1.34, Tyr 1.35, Leu 1.38 and Leu 1.41 in TM1 and residue Leu 438 in helix 8 have been implicated in the TM1 dimerization interface[8] and residues Arg 4.41, Val 4.44, Ile 4.48, Trp 4.50, Val 4.51, Leu 4.52, Phe 4.54, Thr 4.55, Ile 4.56, Cys 4.58, Pro 4.59, Leu 4.60, Leu 4.61, Phe 4.62 have been implicated in the TM4 dimerization interface.[9,10] In Method 2, the manually aligned models, built in Method 1, were submitted to protein-protein docking, using the RosettaDock server.[42] The top scoring model from RosettaDock, which also had a dimerization interface that was consistent with cysteine cross-linking data was selected.

Molecular dynamics model construction. The molecular dynamics model consisted of a homodimer surrounded by a solvated phospholipid bilayer. The homodimer models were embedded in a solvated 1-palmitoyl-2-oleoylphosphatidylcholine (POPC) lipid bilayer[43] (measuring 100 × 100 × 100 Å), using *Silico scripts*[44] to add lipids (bilayer_builder) and water (mol_solvate). A total of 100 lipid molecules were added per monolayer, which was then solvated with 16,000 water molecules.

Molecular dynamics. Molecular dynamics simulations were implemented using NAMD2. The protein was modeled using the CHARMM27 all-atom force field.[45,46] United-atom lipid

parameters were used.[43] The clozapine ligand parameters were developed by Yu Fang. Water was modeled using the TIP3P water model.[47]

Simulations used full periodic boundary conditions and rigid bonds and a 2 fs time step. Bonded interactions were calculated at every time step. Nonbonded interactions were calculated every 6 fs. The long-range electrostatic interactions were computed using the Particle Mesh Ewald (PME) algorithm[48] with a PME grid spacing of approximately 1 Å. Van der Waals (vdW) interactions were calculated up to an inter-atomic distance of 10 Å, then a switching function was used to smoothly bring the interaction energy to 0 between 10 and 12 Å.

Prior to MD simulations, each system was minimized for 1,000 steps using constraints on the protein and ligand. The lipid and water molecules were equilibrated for 10 ns with the protein and ligand fixed by constraints, followed by 10 ns with constraints on the protein backbone alone. The constrained equilibration was followed by 75-100 ns of unconstrained dynamics.

Manual docking of bivalent ligands. A snapshot of the D_2R dimer-clozapine complex, taken from the last frame of the simulation, was used as the 3D model to manually dock a bivalent ligand. Using Maestro, the longest spacer from the clozapine propylamine bivalent ligand series (**11g**) was used to join the two clozapine pharmacophores. The bivalent ligand was minimized using MacroModel[37] in the presence of the D_2R dimer, first using a steepest descents algorithm, followed by Polak–Ribiere conjugate gradient minimization.

4.3.3 Results and discussion

Three distinct proposals have been made to explain why an increase in affinity is sometimes observed for a bivalent ligand, as compared to the corresponding monovalent ligand.[29] The first binding hypothesis is that the incorporation of two pharmacophores in a single molecule increases the local concentration of the pharmacophore in the vicinity of the receptor binding site, thereby increasing the probability of a productive binding event. The

second binding hypothesis is that the bivalent ligand could act bitopically. Specifically, this occurs when one pharmacophore of the bivalent ligand binds to the orthosteric site whilst the second pharmacophore binds to a neighboring (allosteric) site within the same receptor. The third, and most commonly discussed binding hypothesis, is that a bivalent ligand binds to a GPCR dimer, with the two pharmacophores binding simultaneously at adjacent orthosteric sites. This binding event is thought to occur in two stages; one pharmacophore of the bivalent ligand binds univalently to the receptor dimer, allowing the second pharmacophore to more readily associate with the adjacent orthosteric binding site. This is proposed to explain the increased affinity and (potentially) selectivity observed for bivalent ligands.[29,49]

The clozapine homobivalent ligands shown in Figure 4.2 contain a spacer attached at the N4′ position of clozapine. As discussed in Section 4.2, these compounds have good binding affinity and functional activity at the D_2R. The main aim of this study was to investigate the third binding hypothesis for bivalent ligands and determine if these bivalent ligands could bind simultaneously to adjacent orthosteric sites of the D_2R homodimer. To address this, we have built a number of models of the D_2R homodimer and measured the approximate distances between the adjacent orthosteric sites in the model.

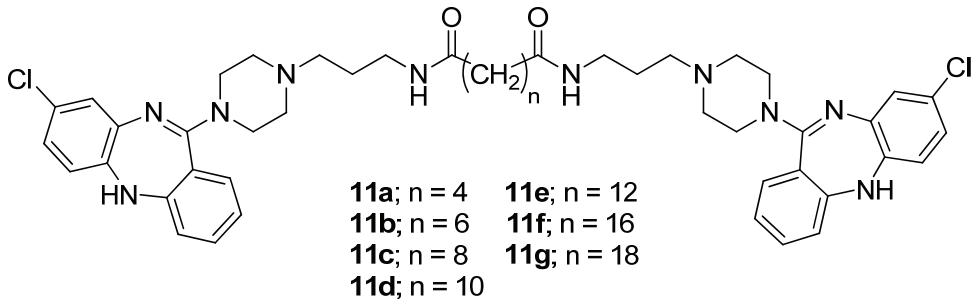


Figure 4.2 Structures of the clozapine propylamine homobivalent ligands (**11a-g**) that were synthesized and pharmacologically evaluated in Section 4.2.

4.3.3.1 Construction of D_2R homodimer models

A homology model of the monomeric D_2R, used as the protomer for the homodimers, was built using Prime[38] and optimized using Induced Fit Docking.[40] The recently solved crystal structure of D_3R (PDB ID: 3PBL)[39] is closely related to D_2R and was used as a template for the development of the D_2R homology model. The orthosteric site of the homology model was optimized by docking clozapine using flexible receptor docking.

Two different techniques, manual alignment and alignment using protein-protein docking, were investigated for the initial construction of D_2R homodimers. Manual alignment of the two D_2R protomers was performed simply by using local transformations in Maestro.[30] The cysteine cross-linking data was used to align the dimerization interface of the two proteins, similar to previously reported modeling of the D_2R homodimer.[8-10] Dimers containing TM1 (Model 1), TM4 (Model 2) or TM4 / TM5 (Model 3) interfaces were built. In the second approach, the RosettaDock server[42] was used to optimize the dimerization interfaces. However, we found that many of the dimers generated in protein-protein docking were poorly aligned with respect to the position of the phospholipid bilayer and only the TM1 interface structure yielded a dimer that could be used in further modeling studies (Model 4). This may be because protein-protein docking is generally used for water-soluble proteins, rather than membrane bound proteins. As a consequence, the position of the membrane in the protein-protein docking studies is not accounted for, resulting in poorly aligned models. The specific details of the four models (Models 1-4) of the D_2R homodimer generated are reported in Table 4.1, which details the modeling method and dimerization interface used to generate the initial dimer models, as well as the residues that form the dimerization interface at the end of the molecular dynamics simulations.

Table 4.1 The four models of the D_2R homodimer, including the modeling method, the dimerization interface and the residues at the dimerization interface (for both protomers) at the end of the molecular dynamics simulation.

Model	Building method	Dimer interface[a]	Residues at the dimer interface (last frame from MD)			
			Chain A		Chain B	
1	Manually aligned[b]	**TM1**	Tyr 1.35	Ile 7.51	Tyr 1.35	Trp 7.40
			Thr 1.37	Phe 7.56	Thr 1.37	Thr 7.55
			Leu 1.38	Ile 431[c]	Leu 1.38	Phe 7.56
			Leu 1.41	Arg 434[c]	Leu 1.41	Arg 434[c]
			Leu 1.42	Leu 438[c]	Val 1.45	Leu 438[c]
			Leu 7.41	Leu 441[c]	Phe 1.48	Leu 441[c]
			Val 7.44	His 442[c]	Val 7.33	His 442[c]
			Val 7.48		Leu 7.34	Cys 443[c]
2	Manually aligned[b]	**TM4**	Arg 145	Leu 4.61	Gln 66	Phe 4.54
			Arg 4.40	Phe 4.62	Tyr 2.41	Phe 4.62
			Phe 4.54		Ile 3.27	
3	Manually aligned[b]	**TM4/TM5**	Tyr 3.51	Val 4.51	Ile 3.48	Val 4.44
			Ala 3.55	Leu 4.52	Tyr 3.51	Val 4.51
			Met 138	Val 5.40	Ala 3.55	Thr 4.55
			Met 140	Tyr 5.41	Met 138	Pro 4.59
			Tyr 142	Ile 5.44	Tyr 142	Tyr 5.41
			Arg 145	Val 5.49	Asn 143	Val 5.45
			Tyr 146	Val 5.53	Tyr 146	Val 5.49
			Val 4.44	Val 5.57	Arg 4.40	Val 5.53
			Ile 4.48		Arg 4.41	Leu 5.56
4	Manually aligned[b] followed by protein-protein docking	**TM1**	Tyr 5.62	Ile 6.46	Leu 1.38	Phe 7.56
			Arg 5.66	Pro 6.50	Leu 1.41	Ile 431[c]
			Arg 5.67	Ile 6.53	Ala 7.47	Arg 434[c]
			Glu 368	Leu 7.41	Val 7.48	Leu 438[c]
			Gln 6.35	Phe 7.56	Ile 7.51	Cys 443[c]
			Gly 6.42		Thr 7.55	

[a]Residues used in protomer alignment (TM1 interface - Tyr 1.34, Tyr 1.35, Leu 1.38 and Leu 1.41, Leu 438 in helix 8[8]; TM4 and TM4/TM5 interface - Arg 4.41, Val 4.44, Ile 4.48, Trp 4.50, Val 4.51, Leu 4.52, Phe 4.54, Thr 4.55, Ile 4.56, Cys 4.58, Pro 4.59, Leu 4.60, Leu 4.61 and Phe 4.62[9,10]). [b]Guided by cysteine cross-linking mutagenesis data. [c]Helix 8.

4.3.3.2 Molecular dynamics simulations of the D_2R homodimer complexes

Having constructed four models of the D_2R homodimer (Models 1-4, Table 4.1), we used molecular dynamics simulations to optimize the dimerization interface. The models of the four D_2R homodimer complexes were embedded in a solvated, phospholipid bilayer containing clozapine in the orthosteric site of each monomer (Figure 4.3). Clozapine was used to prevent the binding site from collapsing and also because it was the pharmacophore used in the development of the homobivalent ligands. The simulations allowed the distances between the two protomers to be optimized, as the protomers could move over the course of the simulation to remove the close contacts that were present in the initial dimer complexes. Furthermore, provided that the simulation is run for a significant period of time, molecular dynamics simulations of GPCR dimers can be used to assess contacts at the dimerization interface. Specifically, to evaluate whether the contacts maintained compared to the biochemical data, if the dimerization interface changes over the course of the simulation, or if the protomers dissociate. The simulations were also used to monitor how the distance between adjacent orthosteric sites changed over the course of the simulation.

The membrane environment around the dimer was modeled using a 1-palmitoyl-2-oleoylphosphatidylcholine (POPC) united-atom lipid bilayer.[43] As large-scale molecular dynamics simulations are computationally intensive, the model can be simplified to speed up the simulation. Simplification of molecular dynamics simulations can include the use of united-atom force field parameters, where hydrogen atoms are modeled implicitly. The united-atom lipid parameters accelerated bilayer simulations by up to 50% in test studies.[43] Whilst some detail regarding the lipid bilayer is reduced using this technique, the detail of the protein and ligand were modeled in detail using an all-atom force field.

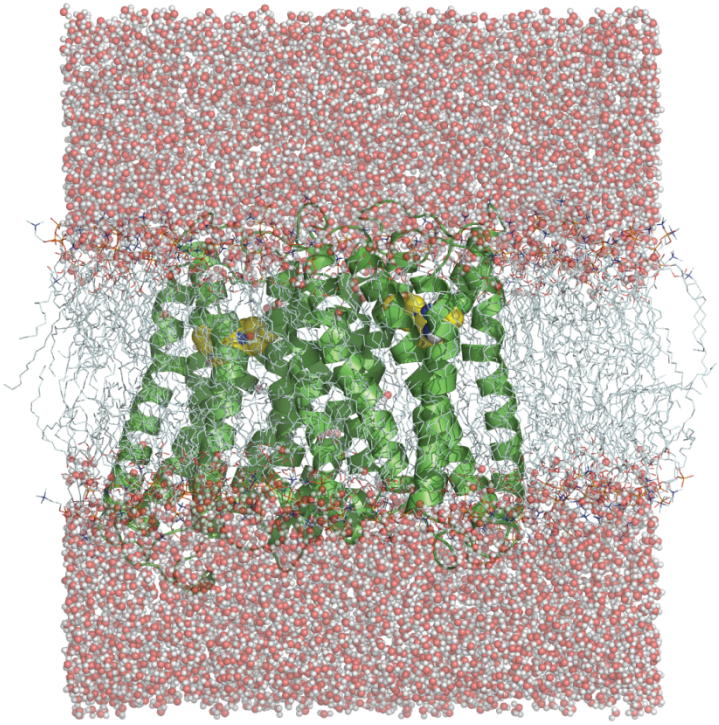


Figure 4.3 Homodimer of the D_2R in complex with clozapine (Model 1) embedded in a solvated phospholipid bilayer, (D_2R – green ribbons, clozapine – yellow spheres, water – red and white spheres, lipid – cyan sticks).

Molecular dynamics simulations were run using NAMD2.[32,33] Initial constraints were placed on the protein and ligand to allow the bilayer and water molecules to relax (10 ns). The constraints were released on everything except the protein backbone and the simulation was run for a further 10 ns. Finally, the systems were simulated, unconstrained, for a further 75-100 ns. Following the removal of all constraints from the protein, the protomers moved apart by approximately 2.5 to 4 Å (Figure 4.4 a, c, e and g) to reduce close contacts between the proteins at the dimer interface. Despite this movement, the protomers generally stayed in a similar alignment to the initial dimer model. Additionally, the clozapine ligands moved within the orthosteric sites over the course of the simulation, but each maintained the key salt bridge to Asp 3.32 (Figure 4.5).

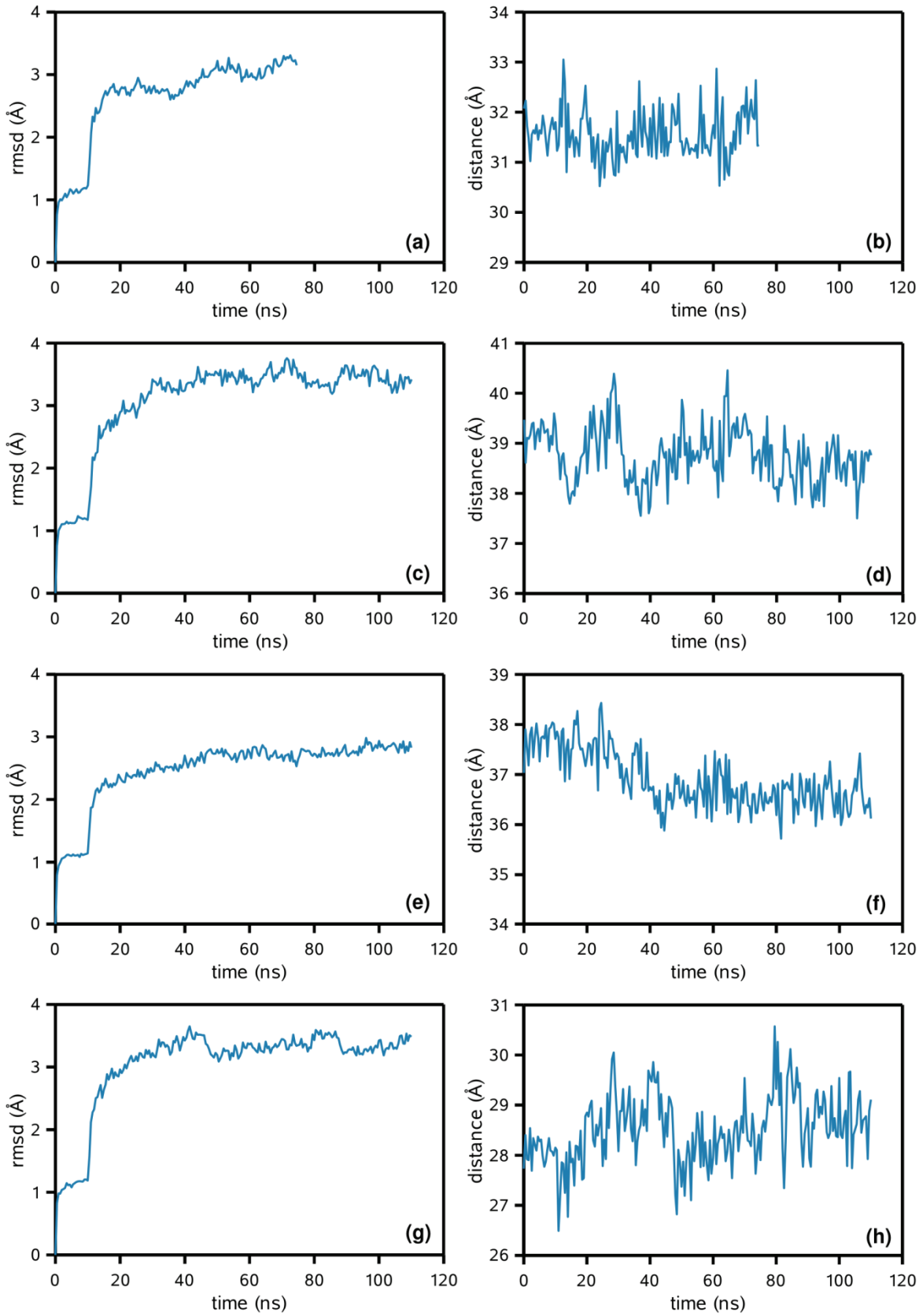


Figure 4.4 Rmsd of the protomers from the starting structure over the course of the simulation for; (a) Model 1, (c) Model 2, (e) Model 3 and (g) Model 4. Distance between the two ionizable nitrogens in the clozapine ligands over the course of the simulation for; (b) Model 1, (d) Model 2, (f) Model 3 and (h) Model 4. Note that for the first 10 ns there were constraints on the protein backbone, following which, the constraints were removed.

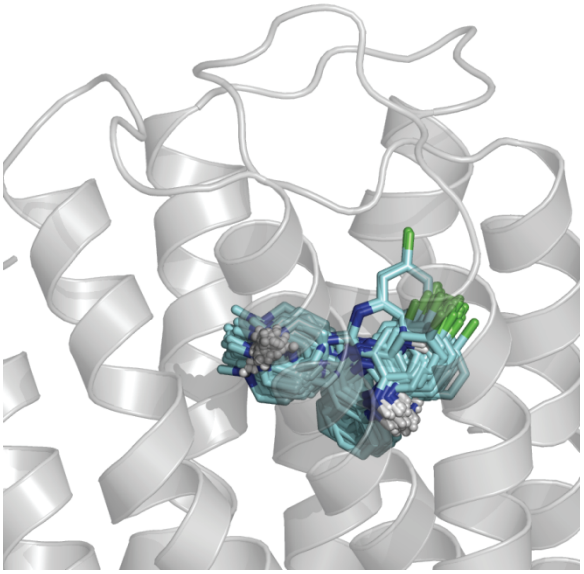


Figure 4.5 A representation of the movement of clozapine in the binding site of one protomer of the D_2R homodimer (Model 1) over the course of the molecular dynamics simulation (clozapine ligand – cyan sticks, D_2R – grey ribbon).

The models obtained from the final frames of the dynamics simulations were used to determine the distance between the clozapine ligands in adjacent orthosteric sites of the D_2R dimer (Figure 4.6). The distance between the ionizable nitrogens on the two clozapine ligands (Figure 4.4 b, d, f and h) can be used to estimate an appropriate spacer length required for a bivalent ligand to span the two orthosteric sites. For the two simulations with the TM1 dimerization interface (Models 1 and 4), different distances resulted. In Model 1, the distance between the ionizable nitrogens was 31-32 Å (Figure 4.4b and Figure 4.6a), whilst in Model 4 was 28-29.5 Å (Figure 4.4h and Figure 4.6d). The distance between the ionizable nitrogens for Models 2 and 3 were larger, 38-39 Å (Figure 4.4d and Figure 4.6b) and 36-37.5 Å (Figure 4.4f and Figure 4.6c) respectively. The differences found using alternate models of the same dimerization interface (i.e. TM1), demonstrate that small changes in the position of the two protomers can significantly alter the distance between the orthosteric binding sites.

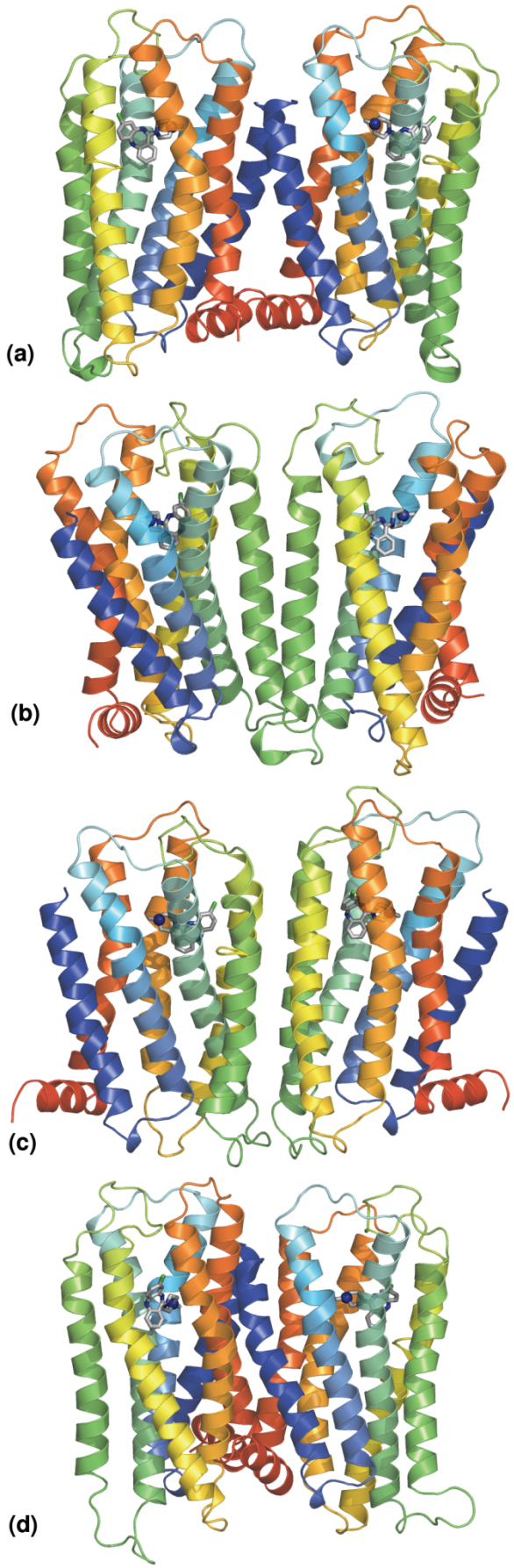


Figure 4.6 Dimer models from the last snapshot of each molecular dynamics simulation. (a) Model 1 – TM1 dimerization interface, (b) Model 2 – TM4 dimerization interface, (c) Model 3 – TM4 / TM5 dimerization interface and (d) Model 4 – TM1 dimerization interface. TM helix colors: TM1 – dark blue, TM2 – blue, TM3 – cyan, TM4 – green, TM5 – yellow, TM6 – orange, TM7 and helix 8 – red; clozapine – grey sticks with ionizable nitrogen highlighted as blue sphere.

4.3.3.3 Modeling the clozapine homobivalent ligands

Compounds **11b** and **11c** were found to be the most active clozapine homobivalent ligands (Section 4.2). The distances between the ionizable nitrogens of these ligands in their fully extended conformations were measured (Figure 4.7 and Table 4.2) and compared with the distances from molecular dynamics simulations to determine if the ligands could fit into Models 1-4.

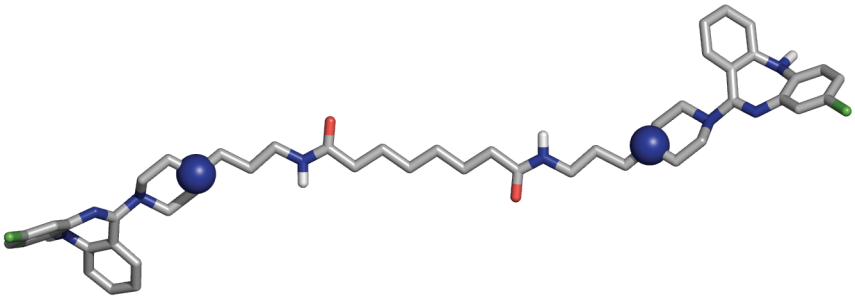


Figure 4.7 Clozapine homobivalent ligand **11b** in an extended conformation, with the two ionizable nitrogens displayed as spheres.

Table 4.2 Number of atoms in the spacer of the clozapine propylamine bivalent ligands (**11a-g**), IC_{50} values from functional assays (refer to Section 4.2 for details) and distances between the ionizable nitrogens in extended clozapine propylamine bivalent ligands.

Compound number	Number of atoms in spacer	IC_{50} (nM)	Length of extended spacer (Å)
11a	14	87	18.7
11b	16	23	21.2
11c	18	44	23.8
11d	20	1,119	26.7
11e	22	> 11,000	28.9
11f	26	7,800	33.9
11g	28	> 10,000	35.1

When the two distances were compared, we found that the results from the molecular dynamics simulations were not concordant with the biological results from Section 4.2, and thus at odds with the third bivalent ligand binding hypothesis (that the two pharmacophores bind simultaneously to adjacent binding sites). Specifically, the spacer lengths of the most

active bivalent ligands **11b** and **11c** were shorter (21.2-23.8 Å) than the distances measured in the molecular dynamics simulations (Models 1 and 4: 28-32 Å and Models 2 and 3: 36-39 Å). However, in this study, we have not investigated the first and second bivalent ligand binding hypotheses, which could also explain the increase affinity and activity of the clozapine homobivalent ligands. Additionally, the spacer that was used to join the two pharmacophores could also be involved in increasing the potency of the clozapine homobivalent ligands.

Whilst the results from the molecular dynamics simulations in this study are not concordant with the third bivalent ligand binding hypothesis, only a limited number of homodimer models (Models 1-4) and dimerization interfaces were investigated. As demonstrated with the homodimer models with a TM1 dimerization interface (Models 1 and 4), small changes in the alignment of the two protomers can alter the distance between the orthosteric binding sites. Thus, the third binding hypothesis for bivalent ligands cannot be conclusively ruled out at this stage.

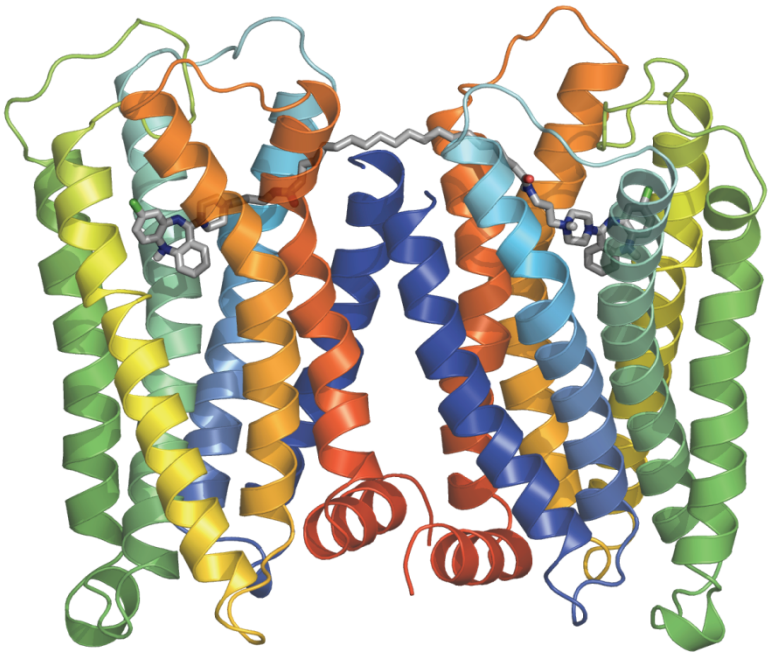


Figure 4.8 Dimer of D_2R (Model 1) in complex with clozapine homobivalent ligand, **11g** (shown as grey sticks).

To investigate the spacer length that would be required for a clozapine homobivalent ligand to bind simultaneously to a D_2R homodimer (if the third bivalent ligand binding hypothesis is true), we manually docked the longest bivalent ligand (**11g**) into Model 1, using the TM1 dimerization interface because this appears to be more plausible than the TM4 interface based on distance between ionizable nitrogens. Whilst compound **11g** displayed poor activity in pharmacological assays in Section 4.2, the spacer length of 28 atoms was long enough to simultaneously allow both pharmacophores to dock into adjacent orthosteric binding sites (Figure 4.8). However, it should be noted that, in Model 1, the spacer passes through a region that may potentially be occupied by the N-terminus residues 1-36, which were not modeled because there was no template available. The position of the N-terminus may affect the prediction of the dimerization interface, particularly those with a TM1 interface, as well as the placement of the bivalent ligand spacer.

4.3.4 Conclusions

Following our identification of the highly active clozapine homobivalent ligands **11b** and **11c**, we have developed four models of the D_2R homodimer, with each protomer in complex with the small molecule clozapine. The initial D_2R dimer models were manually aligned, guided by cysteine cross-linking mutagenesis data,[8-10] as well as using a protein-protein docking server, RosettaDock.[42] Each of the models was optimized using molecular dynamics simulations.

When the distances between the ionizable nitrogens in the molecular dynamics simulations were compared to the spacer lengths of the most active bivalent ligands identified in Section 4.2, we found that these results were not concordant with the third binding hypothesis for bivalent ligands (where both pharmacophores bind simultaneously to adjacent orthosteric binding sites). There are a number of reasons why the distances between the adjacent orthosteric sites in the molecular dynamics simulations and the spacer lengths may not correlate with the pharmacological data.

Firstly, the third binding hypothesis that a bivalent ligand binds simultaneously to two adjacent orthosteric sites may be incorrect. The increased activity of the bivalent ligand may also be explained by the bivalent ligand acting bitopically at a GPCR monomer; where one pharmacophore acts at the orthosteric site, whilst the other pharmacophore acts at an allosteric site (the second binding hypothesis). Alternatively, the covalent tethering of the two pharmacophores may simply increase the local concentration of the pharmacophores in the vicinity of the orthosteric binding site; increasing the likelihood of a favorable binding event (the first binding hypothesis). Currently, these binding hypotheses cannot be ruled out using the available pharmacology data and due to the flexibility of the spacers they are plausible explanations for the increased affinity and activity observed for bivalent ligands.

Secondly, in the current study, only contact dimers were investigated. An alternative proposition is that a domain-swapped dimer may occur; where one or more of the TM helices in one protomer exchanges with the corresponding TM helices of an adjacent protomer (Section 1.6.1, Figure 1.18). The rearrangement of helices in a domain-swapped dimer may allow the distance between the orthosteric sites to be smaller than those we have observed for the contact dimers.

Thirdly, this study has only investigated a limited number of dimer models. An increased number of dimer models, as well as other dimerization interfaces should be explored. For example, in the dimeric models of the 5-HT_4 receptor proposed by Russo et al.,[15] a minimal distance of 22 Å between adjacent orthosteric binding sites was identified, which is similar to the length of the most active clozapine bivalent ligands (**11b** and **11c**, 22-24 Å), however it involved a different dimerization interface (TM helices 2, 3 and 4). In addition, longer molecular dynamics simulations may assist in further optimization of the dimer interface. A bivalent ligand could also be incorporated into the molecular dynamics simulations, to evaluate the effect of the bivalent ligand spacer on GPCR dimerization.

Ultimately, a high-resolution structure of the D_2R homodimer is required to conclusively prove or disprove the third bivalent ligand hypothesis; that a bivalent ligand can bind simultaneously to both orthosteric sites of a GPCR dimer. By developing a number of models of the D_2R homodimer, we have provided useful insight into the distances between the adjacent binding sites of models of D_2R GPCR dimers. Finally, by comparing the models of the D_2R homodimers with the pharmacological results for the clozapine homobivalent ligands developed in Section 4.2, we were able to evaluate the third binding hypothesis of bivalent ligands.

References

1. Bouvier, M. Oligomerization of G-protein-coupled transmitter receptors. *Nat. Rev. Neurosci.* **2001,** *2*, 274-286.
2. Milligan, G.; Lopez-Gimenez, J.; Wilson, S.; Carrillo, J. J. Selectivity in the oligomerisation of G protein-coupled receptors. *Semin. Cell Dev. Biol.* **2004,** *15*, 263-268.
3. George, S. R.; O'Dowd, B. F.; Lee, S. P. G-Protein-coupled receptor oligomerization and its potential for drug discovery. *Nat. Rev. Drug Discovery* **2002,** *1*, 808-820.
4. Portoghese, P. S. From models to molecules: Opioid receptor dimers, bivalent ligands, and selective opioid receptor probes. *J. Med. Chem.* **2001,** *44*, 2259-2269.
5. Fotiadis, D.; Liang, Y.; Filipek, S.; Saperstein, D. A.; Engel, A.; Palczewski, K. Atomic-force microscopy: Rhodopsin dimers in native disc membranes. *Nature* **2003,** *421*, 127-128.
6. Jastrzebska, B.; Fotiadis, D.; Jang, G.-F.; Stenkamp, R. E.; Engel, A.; Palczewski, K. Functional and structural characterization of rhodopsin oligomers. *J. Biol. Chem.* **2006,** *281*, 11917-11922.
7. Wu, B.; Chien, E. Y. T.; Mol, C. D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.; Wells, P.; Bi, F. C.; Hamel, D. J.; Kuhn, P.; Handel, T. M.; Cherezov, V.; Stevens, R. C. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **2010,** *330*, 1066-1071.
8. Guo, W.; Urizar, E.; Kralikova, M.; Mobarec, J. C.; Shi, L.; Filizola, M.; Javitch, J. A. Dopamine D2 receptors form higher order oligomers at physiological expression levels. *EMBO J.* **2008,** *27*, 2293-2304.
9. Guo, W.; Shi, L.; Filizola, M.; Weinstein, H.; Javitch, J. A. Crosstalk in G protein-coupled receptors: Changes at the transmembrane homodimer interface determine activation. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102*, 17495-17500.
10. Guo, W.; Shi, L.; Javitch, J. A. The fourth transmembrane segment forms the interface of the dopamine D_2 receptor homodimer. *J. Biol. Chem.* **2003,** *278*, 4385-4388.
11. Mancia, F.; Assur, Z.; Herman, A. G.; Siegel, R.; Hendrickson, W. A. Ligand sensitivity in dimeric associations of the serotonin 5-HT_{2C} receptor. *EMBO Rep.* **2008,** *9*, 363-369.
12. Berthouze, M.; Rivail, L.; Lucas, A.; Ayoub, M. A.; Russo, O.; Sicsic, S.; Fischmeister, R.; Berque-Bestel, I.; Jockers, R.; Lezoualc'h, F. Two transmembrane Cys residues are involved in 5-HT4 receptor dimerization. *Biochem. Biophys. Res. Commun.* **2007,** *356*, 642-647.
13. Lopez-Gimenez, J. F.; Canals, M.; Pediani, J. D.; Milligan, G. The α_{1b}-adrenoceptor exists as a higher-order oligomer: Effective oligomerization is required for receptor maturation, surface delivery, and function. *Mol. Pharmacol.* **2007,** *71*, 1015-1029.

14. Carrillo, J. J.; López-Giménez, J. F.; Milligan, G. Multiple interactions between transmembrane helices generate the oligomeric α_{1b}-adrenoceptor. *Mol. Pharmacol.* **2004,** *66*, 1123-1137.

15. Russo, O.; Berthouze, M.; Giner, M.; Soulier, J. L.; Rivail, L.; Sicsic, S.; Lezoualc'h, F.; Jockers, R.; Berque-Bestel, I. Synthesis of specific bivalent probes that functionally interact with 5-HT_4 receptor dimers. *J. Med. Chem.* **2007,** *50*, 4482-4492.

16. Soulier, J. L.; Russo, O.; Giner, M.; Rivail, L.; Berthouze, M.; Ongeri, S.; Maigret, B.; Fischmeister, R.; Lezoualc'h, F.; Sicsic, S.; Berque-Bestel, I. Design and synthesis of specific probes for human 5-HT_4 receptor dimerization studies. *J. Med. Chem.* **2005,** *48*, 6220-6228.

17. Gonzalez-Maeso, J.; Ang, R. L.; Yuen, T.; Chan, P.; Weisstaub, N. V.; Lopez-Gimenez, J. F.; Zhou, M.; Okawa, Y.; Callado, L. F.; Milligan, G.; Gingrich, J. A.; Filizola, M.; Meana, J. J.; Sealfon, S. C. Identification of a serotonin/glutamate receptor complex implicated in psychosis. *Nature* **2008,** *452*, 93-97.

18. Johnston, J. M.; Aburi, M.; Provasi, D.; Bortolato, A.; Urizar, E.; Lambert, N. A.; Javitch, J. A.; Filizola, M. Making structural sense of dimerization interfaces of delta opioid receptor homodimers. *Biochemistry* **2011,** *50*, 1682-1690.

19. Kim, S.-K.; Jacobson, K. A. Computational prediction of homodimerization of the A_3 adenosine receptor. *J. Mol. Graph. Model.* **2006,** *25*, 549-561.

20. Vakser, I. A. Evaluation of GRAMM low-resolution docking methodology on the hemagglutinin-antibody complex. *Proteins* **1997,** *Supplement 1*, 226–230.

21. Vakser, I. A. Protein docking for low-resolution structures. *Protein Eng.* **1995,** *8*, 371-378.

22. Bruno, A.; Guadix, A. E.; Costantino, G. Molecular dynamics simulation of the heterodimeric mGluR2/5HT_{2A} complex. An atomistic resolution study of a potential new target in psychiatric conditions. *J. Chem. Inf. Model.* **2009,** *49*, 1602-1616.

23. Bruno, A.; Beato, C.; Costantino, G. Molecular dynamics simulations and docking studies on 3D models of the heterodimeric and homodimeric 5-HT_{2A} receptor subtype. *Future Med. Chem.* **2011,** *3*, 665-681.

24. Gray, J. J.; Moughon, S.; Wang, C.; Schueler-Furman, O.; Kuhlman, B.; Rohl, C. A.; Baker, D. Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *J. Mol. Biol.* **2003,** *331*, 281-299.

25. Liang, Y.; Fotiadis, D.; Filipek, S.; Saperstein, D. A.; Palczewski, K.; Engel, A. Organization of the G protein-coupled receptors rhodopsin and opsin in native membranes. *J. Biol. Chem.* **2003,** *278*, 21655-21662.

26. Daniels, D. J.; Poda, G.; Ferguson, D. M.; Portoghese, P. S. Unpublished data.

27. Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Trong, I. L.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M.

Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **2000,** *289*, 739-745.

28. Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G. F.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.; Stevens, R. C. High-resolution crystal structure of an engineered human β_2-adrenergic G protein coupled receptor. *Science* **2007,** *318*, 1258-1265.

29. Portoghese, P. S. Bivalent ligands and the message-address concept in the design of selective opioid receptor antagonists. *Trends Pharmacol. Sci.* **1989,** *10*, 230-235.

30. Maestro, version 9.2; Schrödinger, LLC: New York, NY, 2011.

31. LigPrep, version 2.4; Schrödinger, LLC: New York, NY, 2010.

32. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005,** *26*, 1781-1802.

33. Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. NAMD2: Greater scalability for parallel molecular dynamics. *J. Comput. Phys.* **1999,** *151*, 283-312.

34. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **1996,** *14*, 33-38.

35. Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **1995,** *8*, 127-134.

36. Ballesteros, J. A.; Weinstein, H.; Stuart, C. S. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. In *Methods Neurosci.*, Academic Press: 1995; Vol. 25, pp 366-428.

37. MacroModel, version 9.9; Schrödinger, LLC: New York, NY, 2011.

38. Prime, version 3.0; Schrödinger, LLC: New York, NY, 2011.

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

40. Schrödinger Suite 2011 Induced Fit Docking protocol;, Glide, version 5.7; Schrödinger, LLC: New York, NY, 2011: Prime, version 3.0; Schrödinger, LLC: New York, NY, 2011.

41. Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **2006,** *49*, 534-553.

42. Lyskov, S.; Gray, J. J. The RosettaDock server for local protein–protein docking. *Nucleic Acids Res.* **2008,** *36*, W233-W238.

43. Hénin, J. r.; Shinoda, W.; Klein, M. L. United-atom acyl chains for CHARMM phospholipids. *J. Phys. Chem. B* **2008,** *112*, 7008-7015.

44. Chalmers, D. K.; Roberts, B. P. Silico - a Perl molecular modelling toolkit, http://silico.sourceforge.net; 2011.

45. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983,** *4*, 187-217.

46. Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **2009,** *30*, 1545-1614.

47. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983,** *79*, 926-935.

48. Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An *N*.log(*N*) method for Ewald sums in large systems. *J. Chem. Phys.* **1993,** *98*, 10089-10092.

49. Morphy, R.; Kay, C.; Rankovic, Z. From magic bullets to designed multiple ligands. *Drug Discov. Today* **2004,** *9*, 641-651.

Chapter 5

Thesis outcomes and future work

G protein-coupled receptors play a key role in cell signaling pathways and, due to their highly druggable nature, they represent a significant pharmaceutical target. However, it is only in the last eleven years that high resolution X-ray crystal structures of GPCRs have been determined. Significant technological advances in the crystallization of membrane-bound proteins have allowed for the determination of a number of GPCR structures, albeit for a relatively small number of seven receptors out of 350 GPCR potential drug targets. From a drug design perspective, the determination of the structures of opsin, $A_{2A}AR$, β_1AR and β_2AR in their active states is particularly interesting, as comparison of the active and inactive states will greatly assist in the understanding of the mechanism of action of GPCRs.

For most of the past eleven years, the available high resolution crystal structures were limited to bovine rhodopsin. Whilst this structure was used extensively for the development of homology models of target GPCRs, there were significant limitations, such as the small size of the orthosteric binding site and the placement of ECL2, which closed off the orthosteric site. The solution of a number of high-resolution crystal structures of non-rhodopsin class A receptors reinvigorated the field of structure-based drug design for GPCRs. Several of these new structures have been used in large scale virtual screening campaigns, which have identified novel ligand chemotypes. Although the number of GPCR crystal structures is gradually increasing it is likely to be a slow process, because obtaining high resolution crystal structures of membrane-bound proteins can require years of work. In the meantime, homology models of pharmaceutically relevant GPCRs can be utilized for structure-based drug design. There are now a number of non-rhodopsin templates (β_2AR, β_1AR, $A_{2A}AR$, D_3R, CXCR4 and H_1R) that can be used for the development of homology

models. These share higher homology with many pharmaceutically relevant GPCR drug targets, which will assist in improving the quality of the models.

In this study, we have built and refined homology models of nine aminergic GPCRs (5-HT_{1B}R, 5-HT_{2A}R, 5-HT_{2B}R, 5-HT_{2C}R, D_2R, D_3R, D_4R, H_1R, M_1R), based on the high resolution crystal structure of β_2AR, which was the best available template for aminergic GPCR homology models at the time of this study (Chapter 2). Loop refinement procedures and flexible receptor docking were used to optimize the receptor models, particularly focusing on the orthosteric site. Small scale virtual screening, which was tested using the crystal structures of β_2AR and A_{2A}AR, was used to evaluate the homology models. Of the nine homology models developed, six showed moderate to good enrichment in virtual screening experiments (5-HT_{1B}R, 5-HT_{2A}R, 5-HT_{2C}R, D_2R, D_3R and M_1R), indicating that a number of these models, particularly the 5-HT_{2A}R structure, would provide a good starting point for structure-based drug design. Additionally, the 5-HT_{2A}R model was used to demonstrate improvements in virtual screening enrichment at each stage of the homology model refinement process. This work was published in the *Journal of Chemical Information and Modeling*,[1] and the final nine homology were made freely available.

Structure-based drug design for GPCRs is an ever evolving field that has been moving at a rapid pace over the last four years, particularly with the increase in the number of available high-resolution crystal structures. It will now be of great interest to develop homology models of aminergic GPCRs using one or more of the now available aminergic GPCR structures (β_1AR, β_2AR, D_3R H_1R) as templates, particularly for those of the same subtype, such as the dopamine receptors. Early work in this area is presented in Chapter 4, where modeling of D_2R using the D_3R crystal structure as a template is described. It would also be of interest to use the developed homology models for large scale virtual screening, with the incorporation of pharmacological testing to assess the top ranked virtual screening hits.

Of the GPCR crystal structures released over the past four years, the structure of D_3R in complex with eticlopride was of particular interest, as we had previously developed a model of the D_3R (Chapter 2). Prior to the release of the D_3R crystal structure, the GPCR Dock 2010[2] assessment was performed to evaluate the status of molecular modeling for GPCRs. Participants were required to submit up to five ranked models of the complex, with rank 1 being the model predicted to be the closest to the crystal structure complex. This was a further opportunity to evaluate the modeling methods established in Chapter 2, including refinement processes and small scale virtual screening. Chapter 3 discusses the generation of 200 models and the refinement and evaluation techniques used select the final five models, which were submitted to the GPCR Dock 2010 analysis. Additionally, these five structures are also compared to the D_3R eticlopride crystal structure.

Participation in GPCR Dock 2010 has been an invaluable experience for the evaluation of our GPCR modeling methods. Based on this assessment, particularly the benzamide ligand docking studies, an evaluation of the ligand conformation using an independent method (e.g. using quantum mechanics calculations), will be included as part of our future model assessment process. Additionally, in the latest version of Glide,[3] there is now an option to include 1,5- and 1,6-intramolecular hydrogen bonds in the scoring of docked ligand poses. Such systems were poorly predicted in our entries to GPCR Dock 2010. Additionally, our participation in the GPCR Dock 2010 assessment has encouraged us to assess the composition of the ligand libraries used in small scale virtual screening for model evaluation. To more rigorously evaluate homology models, a focused decoy library should be developed, containing decoys that more closely resemble the active compounds (i.e. for aminergic GPCRs they should all contain an ionizable nitrogen; of the current decoys, less than 50% had this property). Additionally, for targets such as the dopamine receptors that bind to a diverse range of ligands, small scale virtual screening libraries should be developed for a specific ligand class. For example, the dopamine receptors bind ligands such as

eticlopride (a benzamide compound), haloperidol (a butyrophenone) and clozapine (a dibenzodiazepine). Due to the induced fit nature of ligand binding, it may be unreasonable to develop an "all purpose" homology model that can identify compounds from these ligand classes equally well. By using small scale virtual screening libraries that are biased towards a particular ligand class, it may be possible to detect any bias in the homology models and thus classify them according to the types of ligands that preferentially dock. As a result, for large scale virtual screening studies, a number of homology models might be employed using a protein ensemble approach. Finally, other docking programs should also be evaluated for use in virtual screening studies.

In Chapter 4, we designed, synthesized and pharmacologically evaluated three series of clozapine homobivalent ligands, which differed in the nature and length of the spacer and point of attachment to the pharmacophore. The main aim of this work was to determine if covalently tethering pharmacophores would result in improved affinity at the D_2R compared to the initial pharmacophore, clozapine. In this study, we found that attaching the spacer from the distal piperazine nitrogen of clozapine (N4′) led to the best activity in both binding and functional assays for D_2R. In this case, a propylamine linker was also employed to move the spacer attachment point away from the N4′ and to minimize the disruption to the pK_a of the ionizable nitrogen. The most active compounds (**11b** and **11c**) had spacer lengths of 16 and 18 atoms, respectively. These bivalent ligands displayed low nanomolar receptor affinity (1.41 and 1.35 nM) and functional activity (23 and 44 nM respectively). A spacer-length dependent effect for compounds **11a-g** was observed in both functional and binding assays. Additionally, significant gains in affinity (75- and 79-fold) and functional activity (4.7- to 9-fold) were observed relative to the original pharmacophore, clozapine (106 nM in binding assays and 206 nM in functional assays).

Having identified an appropriate spacer attachment point and spacer length we now have the opportunity to optimize the spacer type and length of clozapine homobivalent ligands.

Proposed improvements include the introduction of more heteroatoms into the spacer to increase the aqueous solubility of the compounds, introducing more conformationally restricted linkers, as well as optimizing the linking method (i.e. is the amide bond in the correct position or can it be optimized?). A thorough method for the pharmacological evaluation of bivalent ligands has also been established, taking into consideration both the affinity and efficacy of these compounds. Given the multi-receptor binding profile of clozapine, it will also be of great interest to investigate the affinity and efficacy of these bivalent ligands at other GPCRs. In particular, other dopamine receptors and the serotonin, histamine and muscarinic receptors that are implicated in the mechanism of action and side effects of clozapine.

Chapter 4 (Section 4.3) describes the development of D_2R homodimer models which were developed using a homology model of D_2R based on the D_3R crystal structure template. The principal aim of this study was to investigate the bivalent ligand binding hypothesis; that both pharmacophores of a bivalent ligand bind simultaneously to adjacent orthosteric binding sites. Using the limited experimental cysteine-cross linking data, four models of D_2R homodimers were developed and subjected to molecular dynamics simulations in a solvated phospholipid bilayer. When the distances between adjacent orthosteric sites over the course of the molecular dynamics simulations were compared to the spacer lengths of our most active bivalent ligands, we found that these results were not concordant with the third bivalent ligand binding hypothesis. There are a number of possible explanations for this discrepancy. These include the possibility that the bivalent ligand may not bind according to the bivalent ligand binding hypothesis, that the dimers could be domain-swapped rather than contact dimers or only a limited number of dimerization interfaces were investigated in this study. Further models of GPCR dimers should be constructed using different modeling techniques, such as using the atomic force microscopy model of rhodopsin dimers to align the protomers or implementing coarse grain molecular

dynamics simulations to self-assemble the GPCR dimer in a membrane.[4] Additionally, longer molecular dynamics simulations may be useful to investigate large protein rearrangements.

A number of molecular models of GPCR monomers and dimers have been successfully built and evaluated in this study that will be useful for structure-based drug design. The bivalent ligands developed in this thesis display promising activity and affinity at the D_2R and will be valuable as pharmacological tools to investigate GPCR dimerization.

References

1. McRobb, F. M.; Capuano, B.; Crosby, I. T.; Chalmers, D. K.; Yuriev, E. Homology modeling and docking evaluation of aminergic G protein-coupled receptors. *J. Chem. Inf. Model.* **2010,** *50*, 626-637.
2. Kufareva, I.; Rueda, M.; Katritch, V.; Stevens, R. C.; Abagyan, R. Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. *Structure* **2011,** *19*, 1108-1126.
3. Glide, version 5.7; Schrödinger, LLC: New York, NY, 2011.
4. Stansfeld, P. J.; Sansom, M. S. P. From coarse grained to atomistic: A serial multiscale approach to membrane protein simulations. *J. Chem. Theory Comput.* **2011,** *7*, 1157-1166.

Appendices

Appendix 1: Supporting information for "Homology modeling and docking evaluation of aminergic GPCRs" (Chapter 2)

Table S1. Summary of recent GPCR crystal structures.

	PDB ID	Resolution (Å)	Binding site	Ligand	ECL2	Ref
Bovine rhodopsin	1U19[a]	2.2	closed	*cis*-retinal	β-sheet	1
β_2	2R4R[b]	3.4/3.7[c]	-	carazolol[b]	-	2
β_2	2R4S[b]	3.4/3.8[c]	-	carazolol[b]	-	2
β_2	3D4S	2.8	open	timolol	short helix	3
β_2	2RH1	2.4	open	carazolol	short helix	4-5
β_1	2VT4	2.7	open	cyanopindolol	short helix	6
Squid rhodopsin	2Z73	2.5	closed	*cis*-retinal	β-sheet	7
Squid rhodopsin	2ZIY	3.7	closed	*cis*-retinal	β-sheet	8
Opsin	3CAP	2.9	closed	-	β-sheet	9
Opsin	3DQB	2.7	closed	-	β-sheet	10
A_{2A}	3EML	2.6	open	ZM-241,385	short helix	11

[a] Highest resolution structure of rhodopsin chosen as example.

[b] The orthosteric site and ECLs were not resolved in this structure.

[c] 3.4 Å in the plane of the membrane and 3.7 Å (or 3.8 Å) perpendicular to the plane of the membrane.

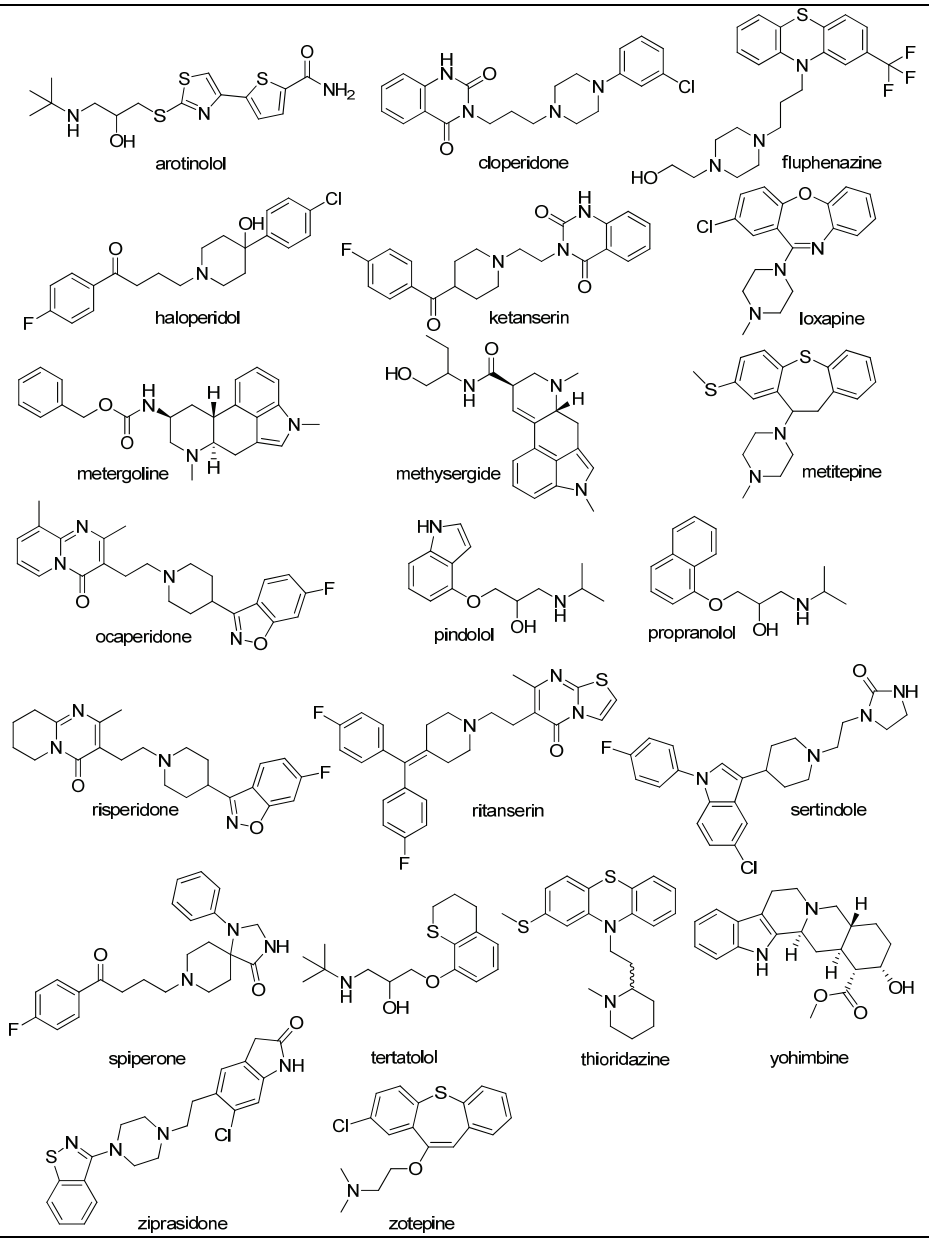
Table S2. Sequence identity between the GPCR targets and the rhodopsin, β_2 and A_{2A} templates measured over the transmembrane helical regions and the ligands used for IFD binding site refinement for each model.

Receptor	**% Homology**			**Ligand used for binding site refinement**
	rhodopsin	**β_2**	**A_{2A}**	
A_{2A}	25	34	100	-
β_2	23	100	34	-
5-HT_{1B}	23	41	35	cyanopindolol
5-HT_{2A}	22	40	30	clozapine
5-HT_{2B}	22	40	30	clozapine
5-HT_{2C}	23	42	30	olanzapine
D_2	26	41	37	olanzapine
D_3	28	38	33	clozapine
D_4	25	34	34	olanzapine
H_1	21	37	34	cetirizine
M_1	22	36	29	clozapine

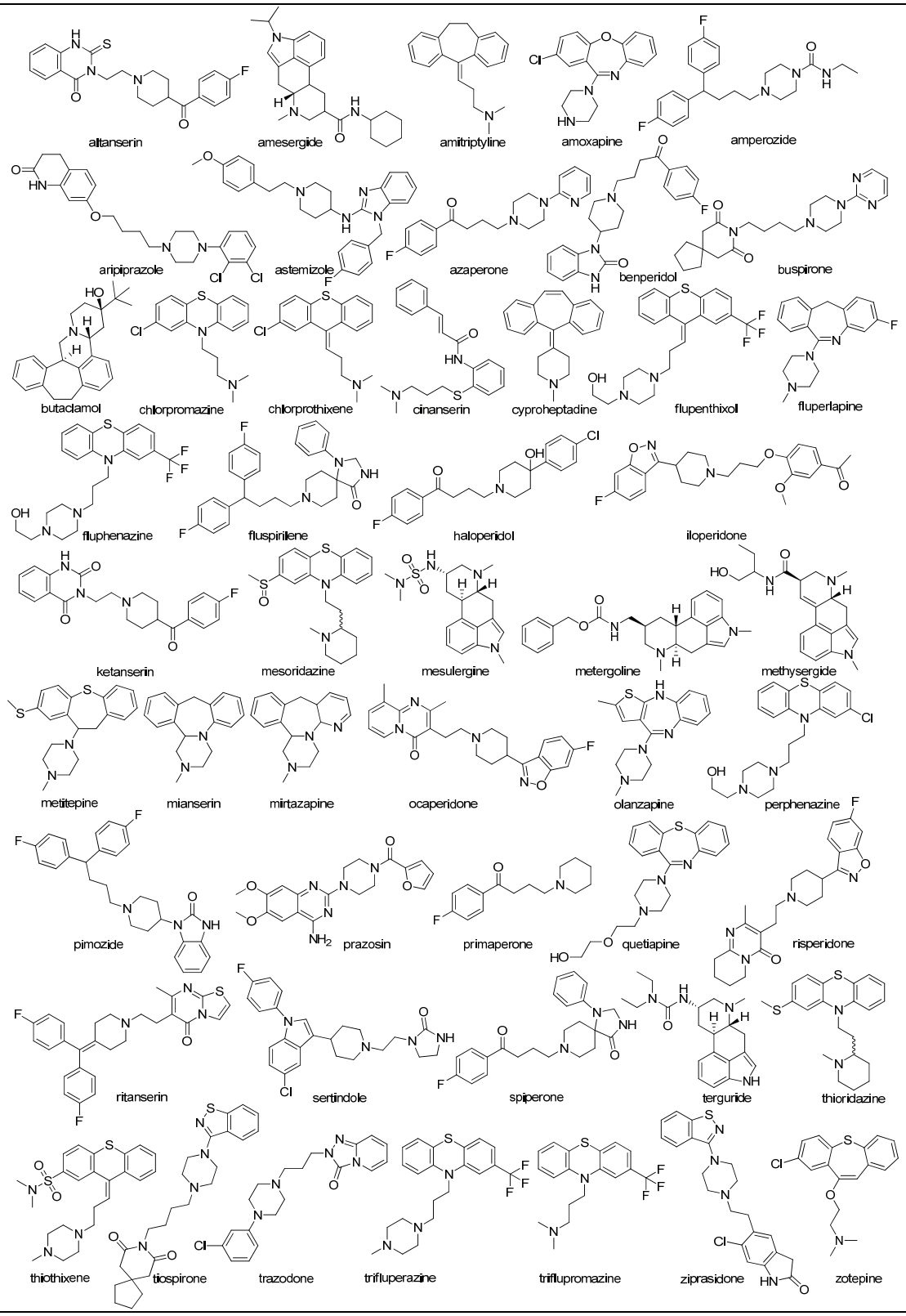
Table S3. Length of ECL2 for each receptor**.** The sections of ECL2 used in the loop refinement protocol for each model (receptor numbering).

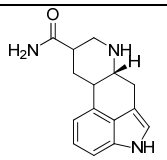
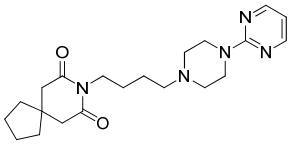
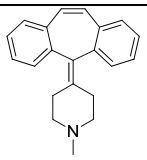
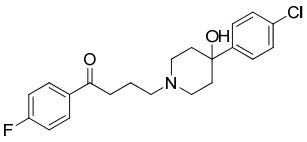
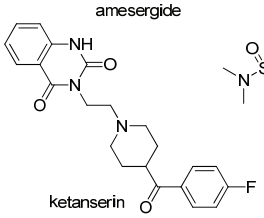
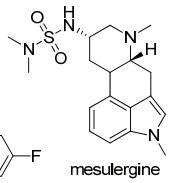
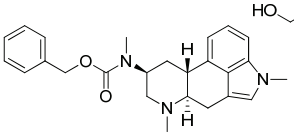
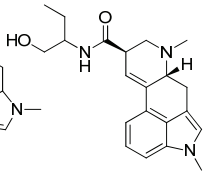
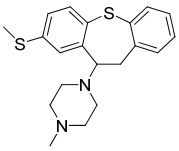
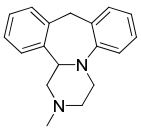
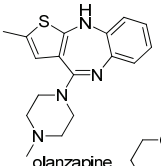
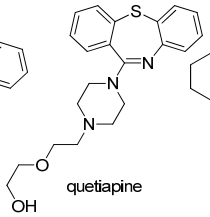
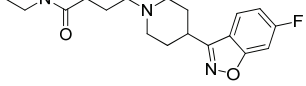
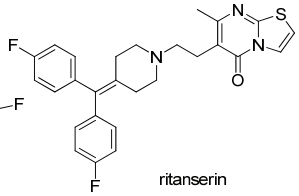
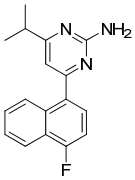
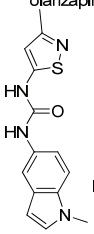
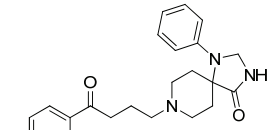
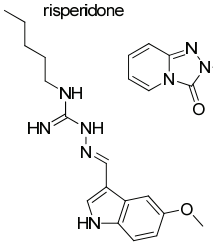
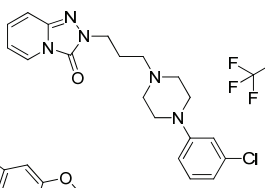
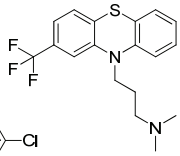
Receptor	**ECL2 length**	**Sections of ECL2 used in loop refinement**			
		1	**2**	**3**	**4**
5-HT_{1B}	17	189-193	194-199	200-205	193-200
5-HT_{2A}	19	214-219	220-224	225-231	219-225
5-HT_{2B}	21	194-200	201-207	208-214	200-208
5-HT_{2C}	20	193-199	200-205	206-212	199-206
A_{2A}	33	-	-	-	-
β_2	23	-	-	-	-
D_2	15	173-176	177-182	183-187	176-184
D_3	13	173-176	177-180	181-185	176-182
D_4	16	174-178	179-183	184-189	178-184
H_1	22	166-173	174-180	181-187	173-181
M_1	22	164-171	172-178	179-185	171-179

Table S4. List of all the active compounds docked into each model during virtual screening.

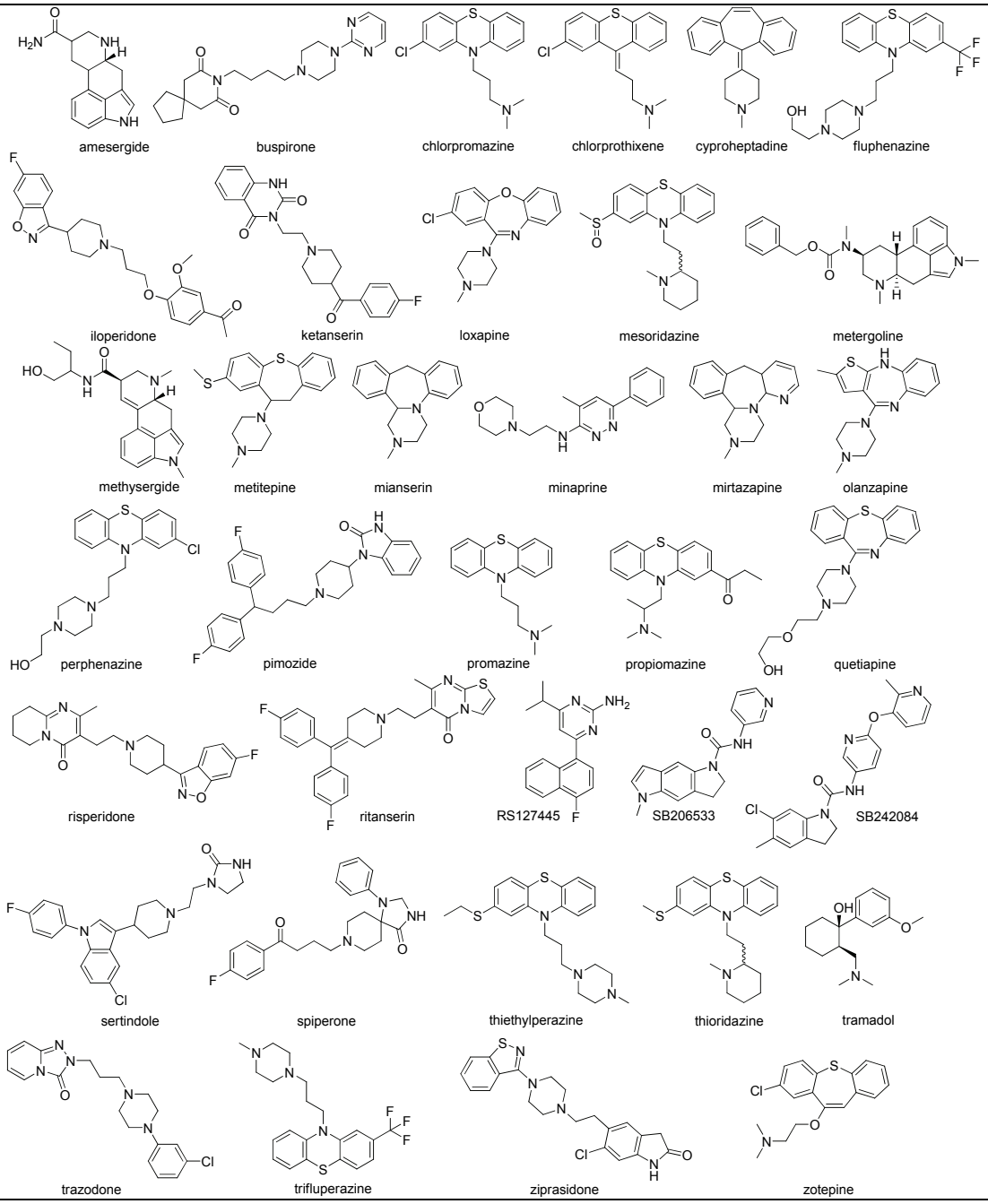
Receptor	**Actives**
5-HT_{1B}	

5-HT_{2A}

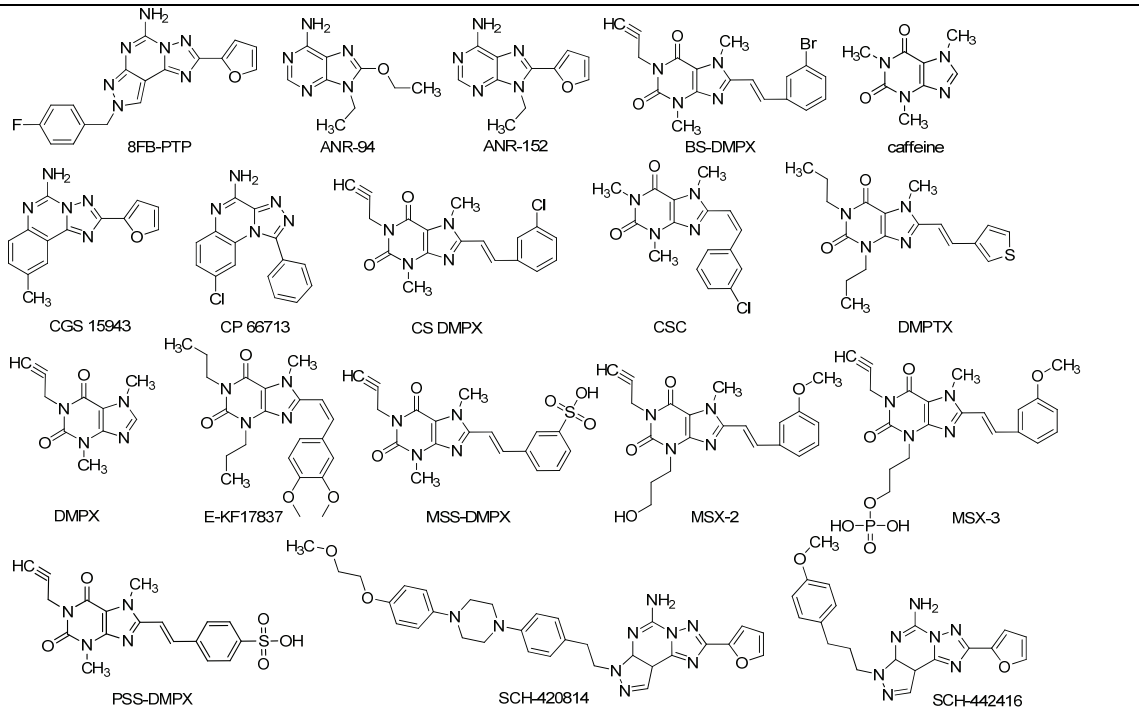

altanserin
amesergide
amitriptyline
amoxapine
amperozide
aripiprazole
astemizole
azaperone
benperidol
buspirone
butaclamol
chlorpromazine
chlorprothixene
cinanserin
cyproheptadine
flupenthixol
fluperlapine
fluphenazine
fluspirilene
haloperidol
iloperidone
ketanserin
mesoridazine
mesulergine
metergoline
methysergide
mettitepine
mianserin
mirtazapine
ocaperidone
olanzapine
perphenazine
pimozide
prazosin
primaperone
quetiapine
risperidone
ritanserin
sertindole
spiperone
terguride
thioridazine
thiothixene
tiospirone
trazodone
trifluperazine
triflupromazine
ziprasidone
zotepine

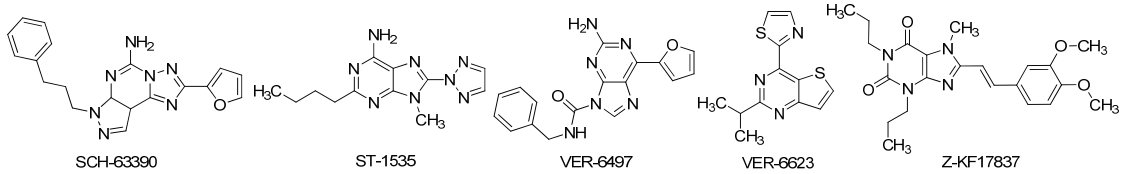
5-HT_{2B}	 amesergide	 buspirone	 cyproheptadine	 haloperidol		
	 ketanserin	 mesulergine	 metergoline	 methysergide	 metitepine	
	 mianserin	 olanzapine	 quetiapine	 risperidone	 ritanserin	
	 RS-127,445	 SB-204,741	 spiperone	 tegaserod	 trazodone	 triflupromazine

5-HT$_{2C}$

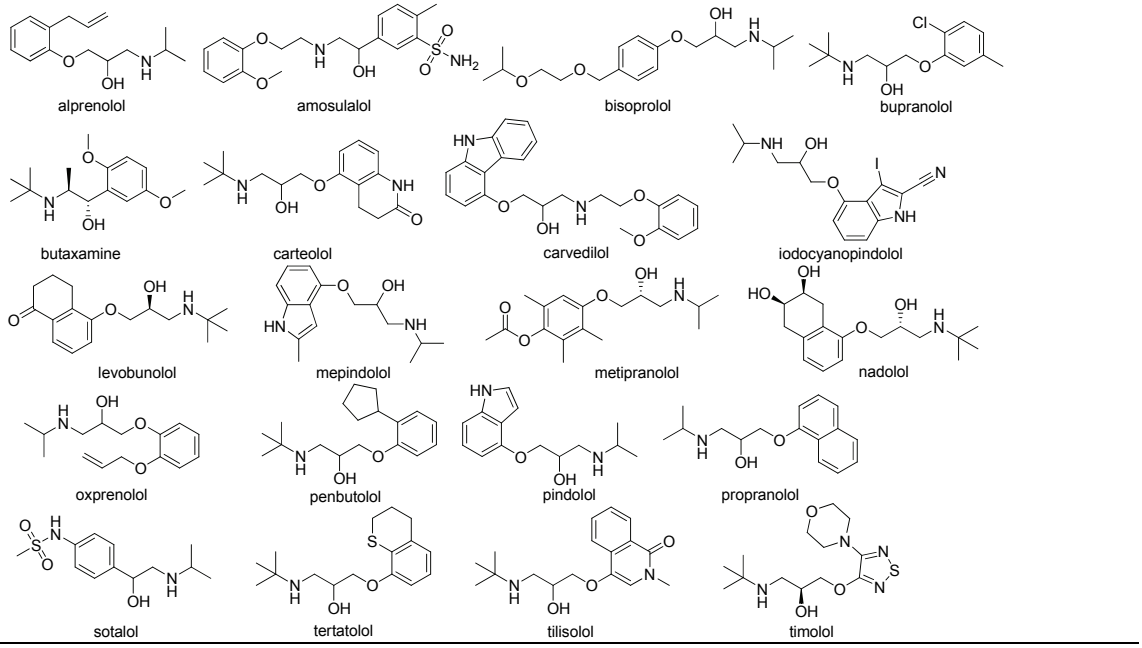

amesergide
buspirone
chlorpromazine
chlorprothixene
cyproheptadine
fluphenazine
iloperidone
ketanserin
loxapine
mesoridazine
metergoline
methysergide
metitepine
mianserin
minaprine
mirtazapine
olanzapine
perphenazine
pimozide
promazine
propiomazine
quetiapine
risperidone
ritanserin
RS127445
SB206533
SB242084
sertindole
spiperone
thiethylperazine
thioridazine
tramadol
trazodone
trifluperazine
ziprasidone
zotepine

A_{2A}

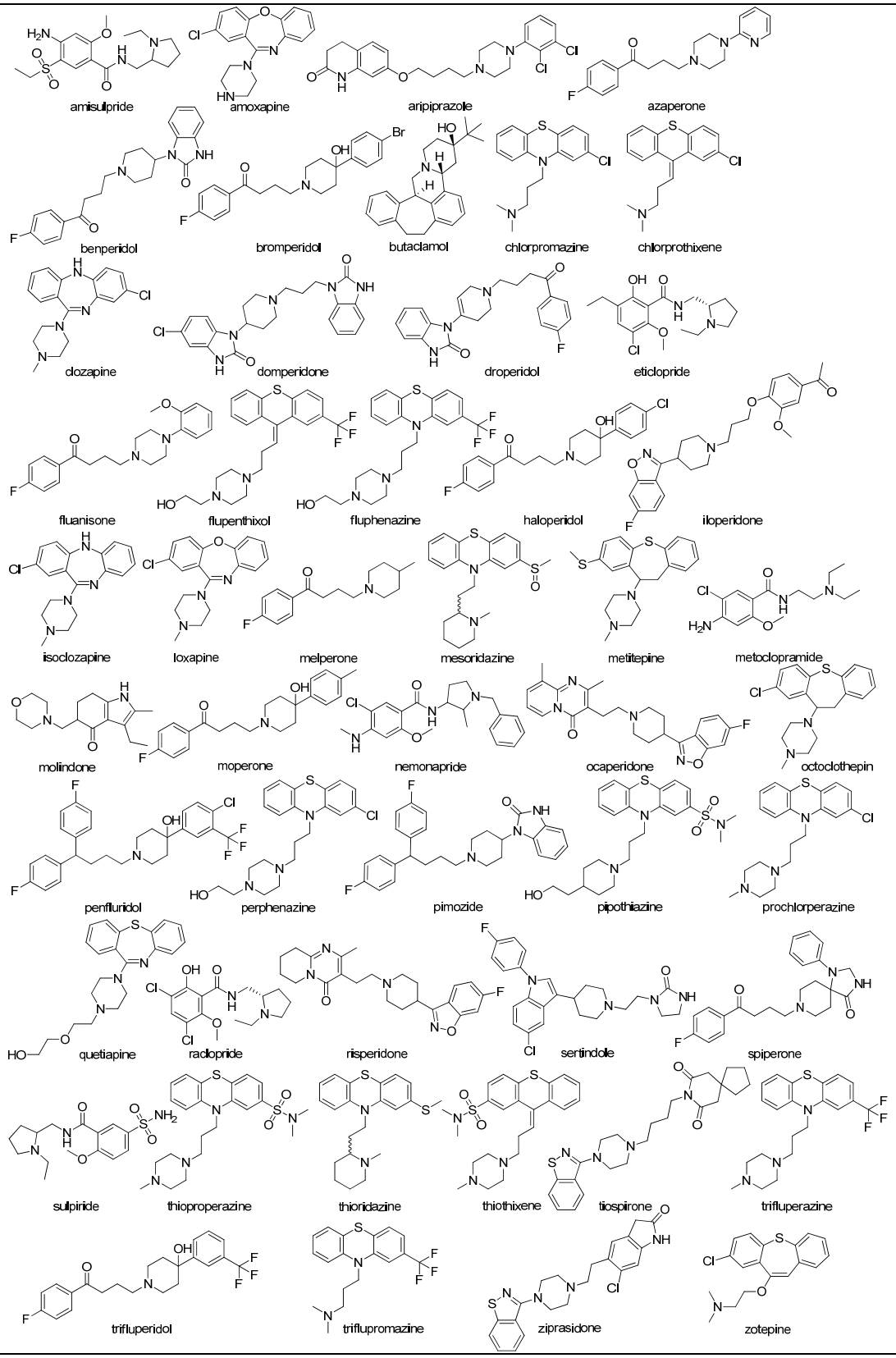

8FB-PTP
ANR-94
ANR-152
BS-DMPX
caffeine
CGS 15943
CP 66713
CS DMPX
CSC
DMPTX
DMPX
E-KF17837
MSS-DMPX
MSX-2
MSX-3
PSS-DMPX
SCH-420814
SCH-442416


SCH-63390
ST-1535
VER-6497
VER-6623
Z-KF17837

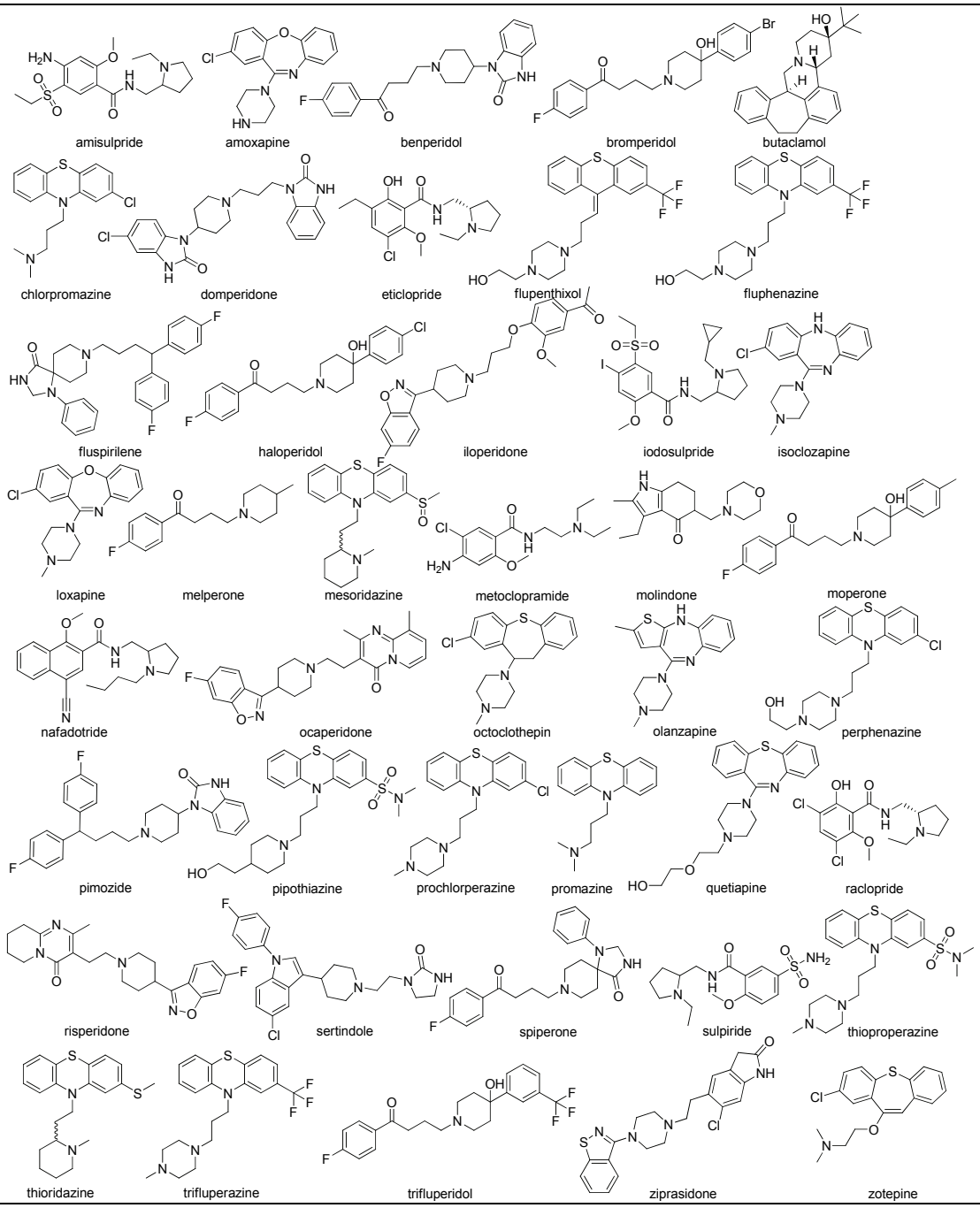
β_2

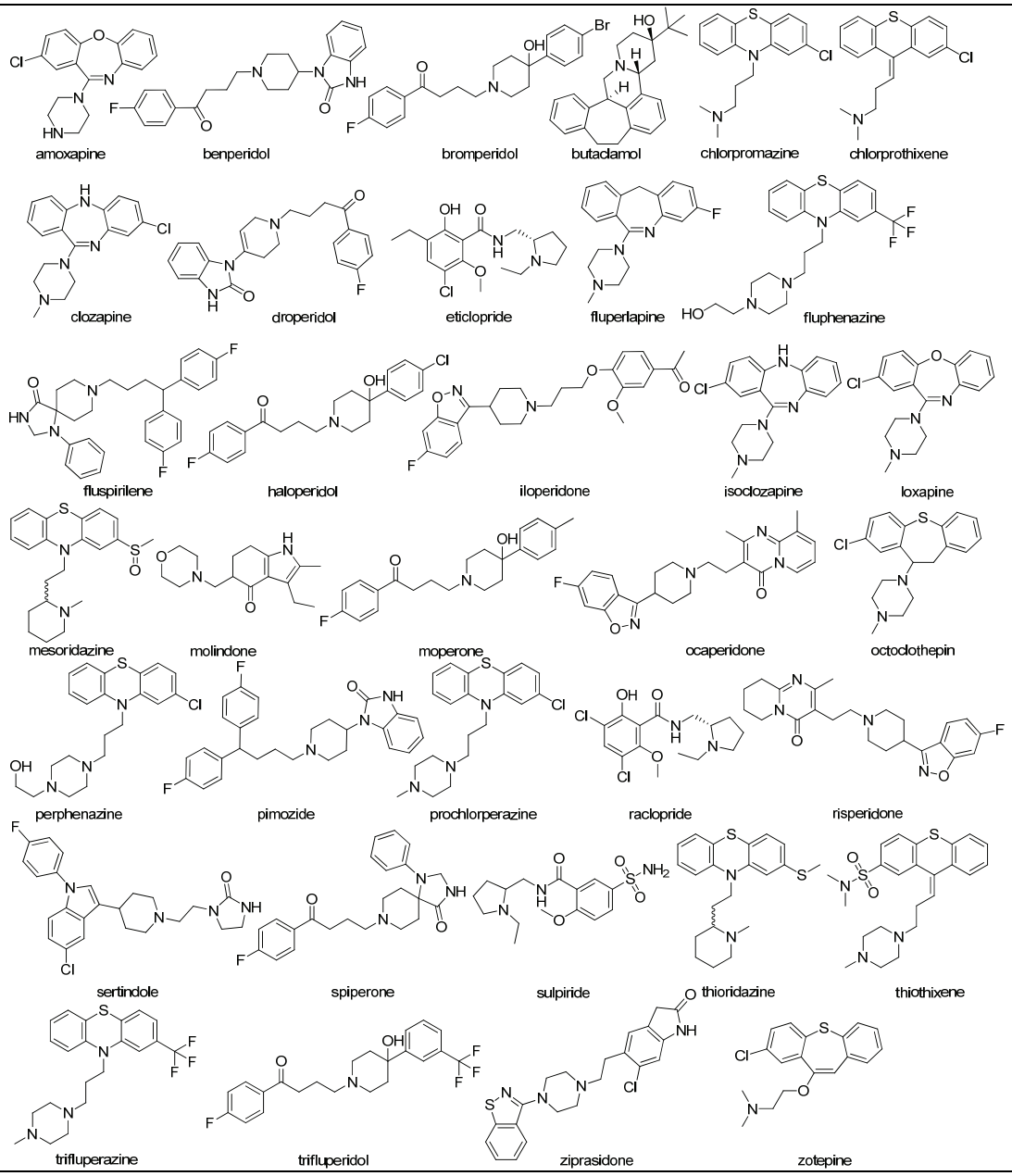

alprenolol
amosulalol
bisoprolol
bupranolol
butaxamine
carteolol
carvedilol
iodocyanopindolol
levobunolol
mepindolol
metipranolol
nadolol
oxprenolol
penbutolol
pindolol
propranolol
sotalol
tertatolol
tilisolol
timolol

D_2


amisulpride
amoxapine
aripiprazole
azaperone
benperidol
bromperidol
butaclamol
chlorpromazine
chlorprothixene
clozapine
domperidone
droperidol
eticlopride
fluanisone
flupenthixol
fluphenazine
haloperidol
iloperidone
isoclozapine
loxapine
melperone
mesoridazine
metitepine
metoclopramide
molindone
moperone
nemonapride
ocaperidone
octoclothepin
penfluridol
perphenazine
pimozide
pipothiazine
prochlorperazine
quetiapine
raclopride
risperidone
sertindole
spiperone
sulpiride
thioproperazine
thioridazine
thiothixene
tiospirone
trifluperazine
trifluperidol
triflupromazine
ziprasidone
zotepine

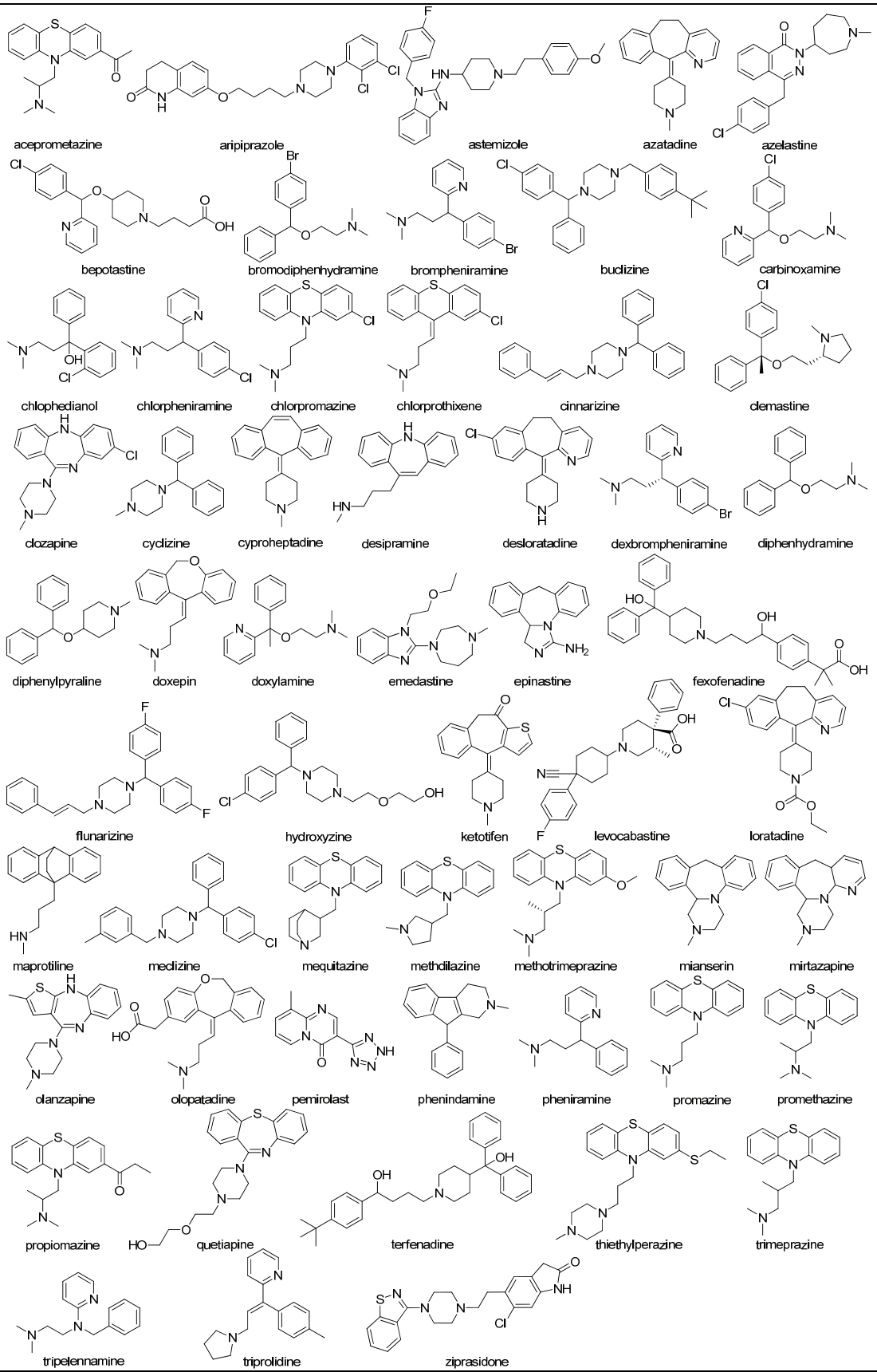
D_3

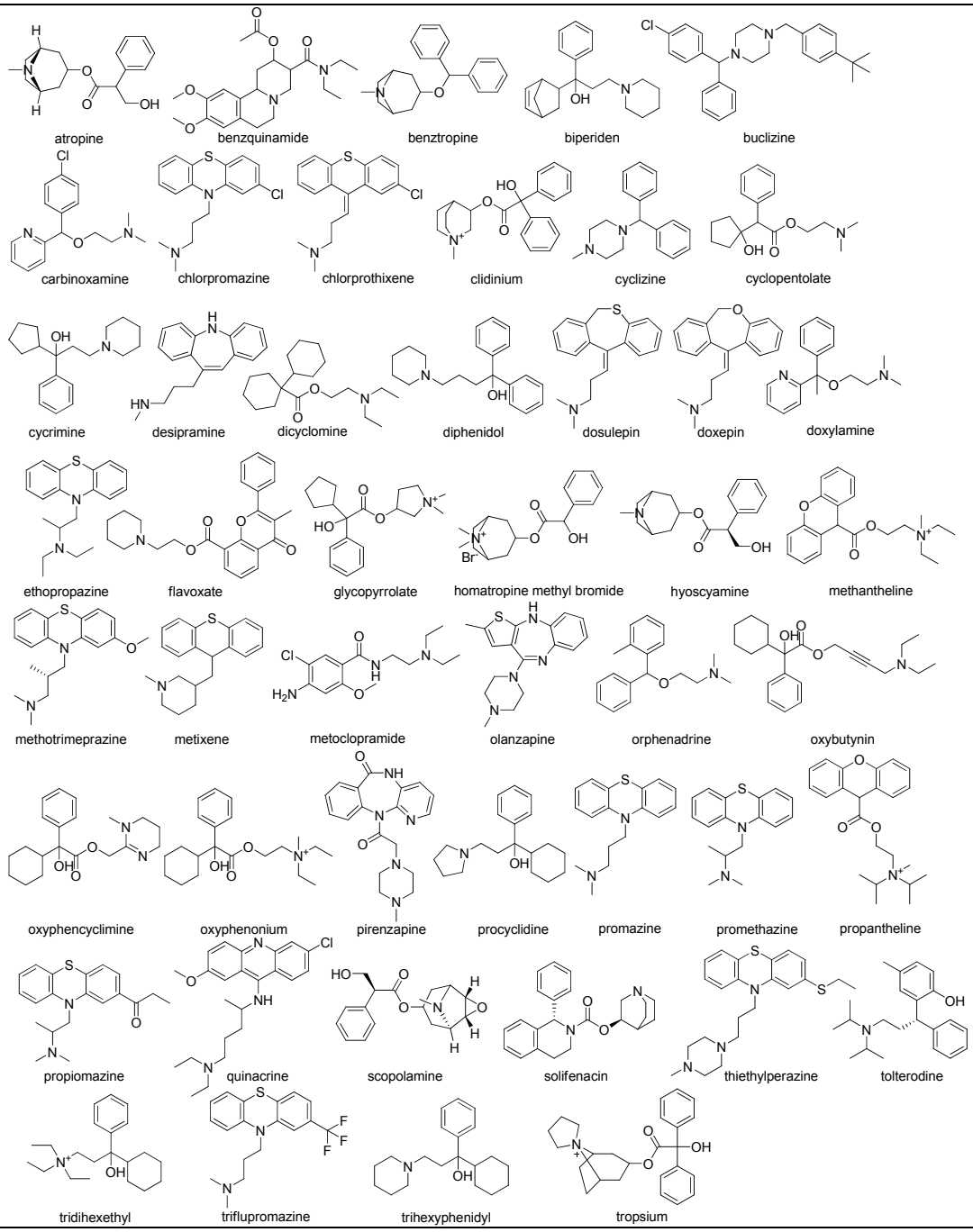

amisulpride
amoxapine
benperidol
bromperidol
butaclamol
chlorpromazine
domperidone
eticlopride
flupenthixol
fluphenazine
fluspirilene
haloperidol
iloperidone
iodosulpride
isoclozapine
loxapine
melperone
mesoridazine
metoclopramide
molindone
moperone
nafadotride
ocaperidone
octoclothepin
olanzapine
perphenazine
pimozide
pipothiazine
prochlorperazine
promazine
quetiapine
raclopride
risperidone
sertindole
spiperone
sulpiride
thioproperazine
thioridazine
trifluperazine
trifluperidol
ziprasidone
zotepine

D_4	

amoxapine
benperidol
bromperidol
butaclamol
chlorpromazine
chlorprothixene
clozapine
droperidol
eticlopride
fluperlapine
fluphenazine
fluspirilene
haloperidol
iloperidone
isoclozapine
loxapine
mesoridazine
molindone
moperone
ocaperidone
octoclothepin
perphenazine
pimozide
prochlorperazine
raclopride
risperidone
sertindole
spiperone
sulpiride
thioridazine
thiothixene
trifluperazine
trifluperidol
ziprasidone
zotepine

H_1


acepromazine
aripiprazole
astemizole
azatadine
azelastine
bepotastine
bromodiphenhydramine
brompheniramine
buclizine
carbinoxamine
chlophedianol
chlorpheniramine
chlorpromazine
chlorprothixene
cinnarizine
clemastine
clozapine
cyclizine
cyproheptadine
desipramine
desloratadine
dexbrompheniramine
diphenhydramine
diphenylpyraline
doxepin
doxylamine
emedastine
epinastine
fexofenadine
flunarizine
hydroxyzine
ketotifen
levocabastine
loratadine
maprotiline
meclizine
mequitazine
methdilazine
methotrimeprazine
mianserin
mirtazapine
olanzapine
olopatadine
pemirolast
phenindamine
pheniramine
promazine
promethazine
propiomazine
quetiapine
terfenadine
thiethylperazine
trimeprazine
tripelennamine
triprolidine
ziprasidone

M_1	

atropine
benzquinamide
benztropine
biperiden
buclizine
carbinoxamine
chlorpromazine
chlorprothixene
clidinium
cyclizine
cyclopentolate
cycrimine
desipramine
dicyclomine
diphenidol
dosulepin
doxepin
doxylamine
ethopropazine
flavoxate
glycopyrrolate
homatropine methyl bromide
hyoscyamine
methantheline
methotrimeprazine
metixene
metoclopramide
olanzapine
orphenadrine
oxybutynin
oxyphencyclimine
oxyphenonium
pirenzapine
procyclidine
promazine
promethazine
propantheline
propiomazine
quinacrine
scopolamine
solifenacin
thiethylperazine
tolterodine
tridihexethyl
triflupromazine
trihexyphenidyl
tropsium

Table S5. Average ligand properties calculated in QikProp[12] (average shape Tanimoto score, calculated in ROCS[13] and average 2D Tanimoto score calculated using UNITY in Sybyl[14]).

Property / Receptor	**MW (g/mol)**	**Rot. bonds**	**PSA (Å^2)**	**clogP**	**H-bond Donor**	**H-bond Accept.**	**Solvent Access. Volume (Å^3)**	**Shape Tanimoto Score[a]**	**2D Tanimoto Score[a]**
5-HT$_{1B}$	369	4.6	54.7	3.8	1.1	5.3	1185	0.52	0.29
5-HT$_{2A}$	376	4.1	44.0	4.2	0.6	5.1	1211	0.48	0.27
5-HT$_{2B}$	356	3.5	50.4	3.8	0.9	5.1	1162	0.48	0.26
5-HT$_{2C}$	362	3.6	39.9	4.1	0.5	4.8	1167	0.53	0.25
A$_{2A}$	344	5.0	88.5	2.9	1.3	6.3	1075	0.54	0.24
β$_2$	301	8.2	63.2	2.4	2.5	5.4	1055	0.54	0.53
D$_2$	378	4.6	46.9	4.1	0.7	5.4	1198	0.50	0.25
D$_3$	381	4.8	49.0	3.9	0.8	5.5	1193	0.51	0.30
D$_4$	375	4.0	44.1	4.1	0.7	5.1	1179	0.52	0.26
H$_1$	328	4.1	27.4	4.0	0.4	3.9	1095	0.61	0.33
M$_1$	320	4.9	29.5	4.0	0.4	4.0	1087	0.61	0.23
decoys[b]	360	5.5	86.7	3.1	1.9	5.8	1117	0.48	0.22

[a] Score of 0 indicates dissimilar ligands, maximum score of 1 indicates identical ligands.

[b] Schrödinger decoy library.[15]

Table S6. Enrichment factors for intermediate 5-HT_{2A} models at *x*% of the ranked database screened (maximum enrichment factors at 2% 21.4; 5% 19.8; 10% 10, of the ranked database screened).

Model	**Enrichment factor (at x% of the ranked database screened)**					
	2%		**5%**		**10%**	
	Enrichment factor	**Number of compounds recovered**	**Enrichment factor**	**Number of compounds recovered**	**Enrichment factor**	**Number of compounds recovered**
1 (Initial)	2.0	2	2.8	7	3.3	16
2 (After loop refinement)	3.1	3	2.4	6	2.9	14
3 (After IFD)	13.3	13	7.7	19	4.7	23
4 (Final)	6.1	6	6.9	17	5.9	29

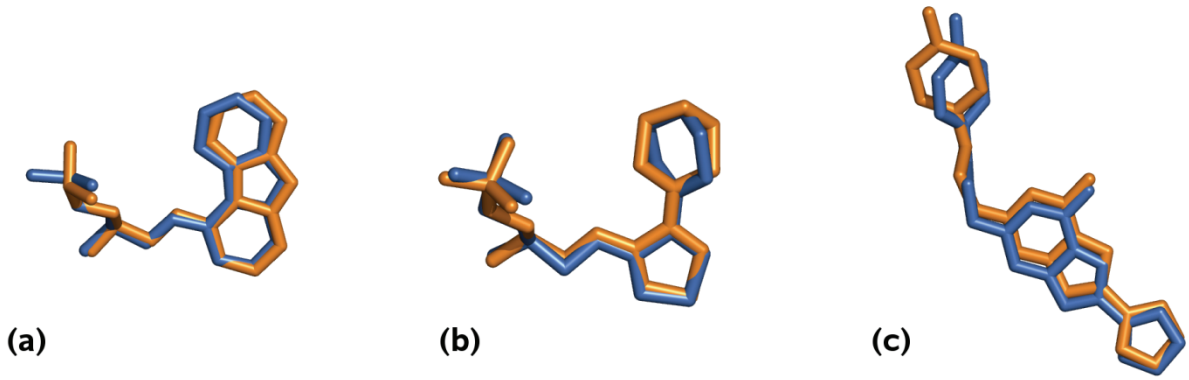


Figure S1. Cognate ligand docking results; crystal structure shown in blue, docked pose shown in orange (docking method in brackets). a) carazolol from 2RH1 crystal structure compared to docked pose of carazolol (Glide XP), b) timolol from 3D4S crystal structure compared to docked pose of timolol from virtual screening (Glide XP) and c) ZM-241,385 from 3EML crystal structure compared to docked pose of ZM-241,385 (Glide XP, with crystal structure water). This image was created in PyMOL.[16]

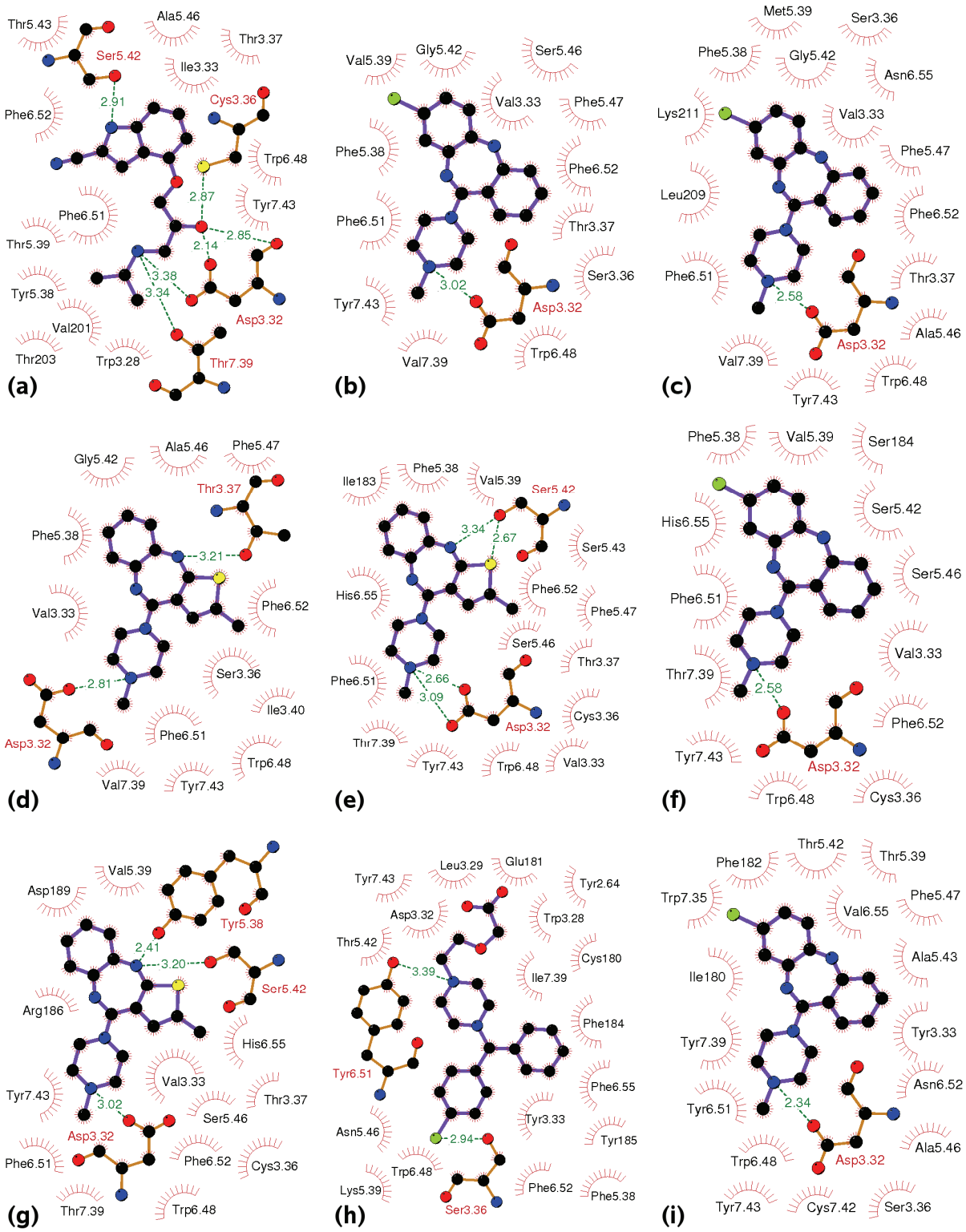


Figure S2. Schematic 2D plots of intermolecular interactions in the docked structures from IFD. a) cyanopindolol docked into the 5-HT_{1B} receptor, b) clozapine docked into the 5-HT_{2A} receptor, c) clozapine docked into the 5-HT_{2B} receptor, d) olanzapine docked into the 5-HT_{2C} receptor, e) olanzapine docked into the D_2 receptor, f) clozapine docked into the D_3 receptor, g) olanzapine docked into the D_4 receptor, h) cetirizine docked into the H_1 receptor, i) clozapine docked into the M_1 receptor. Hydrogen atoms omitted for clarity. Non-bonded interactions: vdW, red spokes; hydrogen bonds, dashed greenlines. Covalent bonds:

ligand, purple; protein, brown. Protein: side-chains are shown only for residues, to which a ligand is hydrogen bonded; a red single spoked arcs show residues involved only in vdW contact(s) with a ligand. Atoms: carbon, black; oxygen, red; nitrogen, blue; sulphur, yellow; chlorine; green. The plots were created with the program LIGPLOT.[17]

References

(1) Okada, T.; Sugihara, M.; Bondar, A.-N.; Elstner, M.; Entel, P.; Buss, V. The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. *J. Mol. Biol.* **2004,** *342*, 571-583.

(2) Rasmussen, S. G. F.; Choi, H.-J.; Rosenbaum, D. M.; Kobilka, T. S.; Thian, F. S.; Edwards, P. C.; Burghammer, M.; Ratnala, V. R. P.; Sanishvili, R.; Fischetti, R. F.; Schertler, G. F. X.; Weis, W. I.; Kobilka, B. K. Crystal structure of the human β2 adrenergic G-protein-coupled receptor. *Nature* **2007,** *450*, 383-387.

(3) Hanson, M. A.; Cherezov, V.; Griffith, M. T.; Roth, C. B.; Jaakola, V.-P.; Chien, E. Y. T.; Velasquez, J.; Kuhn, P.; Stevens, R. C. A specific cholesterol binding site Is established by the 2.8 Å structure of the human β_2-adrenergic receptor. *Structure* **2008,** *16*, 897-905.

(4) Rosenbaum, D. M.; Cherezov, V.; Hanson, M. A.; Rasmussen, S. G. F.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Yao, X.-J.; Weis, W. I.; Stevens, R. C.; Kobilka, B. K. GPCR engineering yields high-resolution structural insights into β_2-adrenergic receptor function. *Science* **2007,** *318*, 1266-1273.

(5) Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G. F.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.; Stevens, R. C. High-resolution crystal structure of an engineered human β_2-adrenergic G protein coupled receptor. *Science* **2007,** *318*, 1258-1265.

(6) Warne, T.; Serrano-Vega, M. J.; Baker, J. G.; Moukhametzianov, R.; Edwards, P. C.; Henderson, R.; Leslie, A. G. W.; Tate, C. G.; Schertler, G. F. X. Structure of a β_1-adrenergic G-protein-coupled receptor. *Nature* **2008,** *454*, 486-491.

(7) Murakami, M.; Kouyama, T. Crystal structure of squid rhodopsin. *Nature* **2008,** *453*, 363-367.

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

(9) Park, J. H.; Scheerer, P.; Hofmann, K. P.; Choe, H.-W.; Ernst, O. P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **2008,** *454*, 183-187.

(10) Scheerer, P.; Park, J. H.; Hildebrand, P. W.; Kim, Y. J.; Krausz, N.; Choe, H.-W.; Hofmann, K. P.; Ernst, O. P. Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **2008,** *455*, 497-502.

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

(12) *QikProp*, version 3.1; Schrödinger, LLC: New York, NY, 2008.

(13) *ROCS*, version 2.3.1; OpenEye Scientific Software Inc.: Santa Fe, New Mexico, 2007.

(14) *Sybyl-X*, version 1.0; Tripos: St. Louis, MO, 2009.

(15) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004,** *47*, 1739-1749.

(16) DeLano, W. L. *The PyMOL molecular graphics system*, DeLano Scientific: Palo Alto, CA, USA, 2002.

(17) Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **1995,** *8*, 127-134.

Appendix 2: Multiple sequence alignment (Clustal W 1.7)

NOTE: The most conserved residue in each transmembrane helix is highlighted in bold (corresponds to position X.50 in Ballesteros-Weinstein numbering). Approximate positions of transmembrane helices are highlighted with colored boxes (TM1 – purple, TM2 – dark blue, TM3 – blue, TM4 – green, TM5 – yellow, TM6 – orange, TM7 – red), however the start and end of the helices may vary slightly depending upon the specific receptor.

```
P07550_beta2  ---MGQ---------PGNGS------------AFLLAPNRSHA----PDHDVTQQRDEV-
P29274_A2A    ----------------MPIMGS-------------------------------------
P28222_5ht1b  ----MEEPGAQCAPPPPAGSET--------WVPQANLSSAPSQNCSAKDYIYQDS-ISL-
P28223_5ht2a  MDILCEENTSLSSTTNSLMQLNDDTRLYSNDFNSGEANTSDAFNWTVDSENRTNLSCEGC
P41595_5ht2b  -MALSYRVSELQSTIPEHILQSTFVHVIS-------SNWSGLQTESIPEEMKQIVEEQG-
P28335_5ht2c  MVNLRNAVHSFLVHLIGLLVWQCDISVSP-----VAAIVTDIFN-TSDG-GRFKFPDG--
P14416_D2     ------------MDPLNLSWYD--------DDLERQNWSRP--------FNGSDGKADR-
P35462_D3     --------------MASLSQLS--------SHL---NYTCG--------AENSTGASQA-
P21917_D4     ------------MGNRSTADAD--------GLLAGRGPAAG--------ASAGASAGLA-
P35367_H1     ------MSLPNSSCLLEDKMCE----------GNKTTMAS--------------------
P11229_M1     ---------MNTSAPPAVSPNI----------TVLAPGKG--------------------

P07550_beta2  -----------WVVGMG-IVMSLIVLAIVFGNVLVITAIAKFERLQTVTNYFITSLACAD
P29274_A2A    ------------------SVYITVELAIAVLAILGNVLVCWAVWLNSNLQNVTNYFVVSLAAAD
P28222_5ht1b  -----------PWKVLLVMLLALITLATTLSNAFVIATVYRTRKLHTPANYLIASLAVTD
P28223_5ht2a  LSPSCLSLLHLQEKNWSALLTAVVIILTIAGNILVIMAVSLEKKLQNATNYFLMSLAIAD
P41595_5ht2b  -----------NKLHWAALLILMVIIPTIGGNTLVILAVSLEKKLQYATNYFLMSLAVAD
P28335_5ht2c  ------------VQNWPALSIVIIIIMTIGGNILVIMAVSMEKKLHNATNYFLMSLAIAD
P14416_D2     -----------PHYNYYATLLTLLIAVIVFGNVLVCMAVSREKALQTTTNYLIVSLAVAD
P35462_D3     -----------RPHAYYALSYCALILAIVFGNGLVCMAVLKERALQTTTNYLVVSLAVAD
P21917_D4     -----------GQGAAALVGGVLLIGAVLAGNSLVCVSVATERALQTPTNSFIVSLAAAD
P35367_H1     -----------PQLMPLVVVLSTICLVTVGLNLLVLYAVRSERKLHTVGNLYIVSLSVAD
P11229_M1     ----------PWQVAFIGITTGLLSLATVTGNLLVLISFKVNTELKTVNNYFLLSLACAD

P07550_beta2  LVMGLAVVPFGAAHILMK-MWTFGNFWCEFWTSIDVLCVTASIETLCVIAVDRYFAITSP
P29274_A2A    IAVGVLAIPFAITISTGF---CAACHGCLFIACFVLVLTQSSIFSLLAIAIDRYIAIRIP
P28222_5ht1b  LLVSILVMPISTMYTVTG-RWTLGQVVCDFWLSSDITCCTASILHLCVIALDRYWAITDA
P28223_5ht2a  MLLGFLVMPVSMLTILYGYRWPLPSKLCAVWIYLDVLFSTASIMHLCAISLDRYVAIQNP
P41595_5ht2b  LLVGLFVMPIALLTIMFEAMWPLPLVLCPAWLFLDVLFSTASIMHLCAISVDRYIAIKKP
P28335_5ht2c  MLVGLLVMPLSLLAILYDYVWPLPRYLCPVWISLDVLFSTASIMHLCAISLDRYVAIRNP
P14416_D2     LLVATLVMPWVVYLEVVG-EWKFSRIHCDIFVTLDVMMCTASILNLCAISIDRYTAVAMP
P35462_D3     LLVATLVMPWVVYLEVTGGVWNFSRICCDVFVTLDVMMCTASILNLCAISIDRYTAVVMP
P21917_D4     LLLALLVLPLFVYSEVQGGAWLLSPRLCDALMAMDVMLCTASIFNLCAISVDRFVAVAVP
P35367_H1     LIVGAVVMPMNILYLLMS-KWSLGRPLCLFWLSMDYVASTASIFSVFILCIDRYRSVQQP
P11229_M1     LIIGTFSMNLYTTYLLMG-HWALGTLACDLWLALDYVASNASVMNLLLISFDRYFSVTRP
```

```
P07550_beta2  FKYQ---SLLTKNKARVIILMVWIVSGLTSFLPIQMHWYRATHQ-----------EAINC
P29274_A2A    LRYN---GLVTGTRAKGIIAICWVLSFAIGLTPMLGWN-------NCGQPKEGKNHSQGC
P28222_5ht1b  VEYS---AKRTPKRAAVMIALVWVFSISISLPPFF-WRQAK-AEEEVS------------
P28223_5ht2a  IHHS---RFNSRTKAFLKIIAVWTISVGISMPIPVFGLQDDSKVFKE-------------
P41595_5ht2b  IQAN---QYNSRATAFIKITVVWLISIGIAIPVPIKGIETDVDNPNN-------------
P28335_5ht2c  IEHS---RFNSRTKAIMKIAIVWAISIGVSVPIPVIGLRDEEKVFVNN------------
P14416_D2     MLYN--TRYSSKRRVTVMISIVWVLSFTISCPLLFGLNN---ADQNE-------------
P35462_D3     VHYQHGTGQSSCRRVALMITAVWVLAFAVSCPLLFGFNTT--GDPTV-------------
P21917_D4     LRYN---RQGGSRRQLLLIGATWLLSAAVAAPVLCGLNDVRGRDPAV-------------
P35367_H1     LRYL---KYRTKTRASATILGAWFLSFLWVIPILGWNHFMQQTSVRRE------------
P11229_M1     LSYR---AKRTPRRAALMIGLAWLVSFVLWAPAILFWQYLVGERTVLA------------

P07550_beta2  YANETCCDFFTN--QAYAIAS-SIVSFYVPLVIMVFVYSRVFQEAKRQLQKIDKSEGRFH
P29274_A2A    GEGQVACLFEDVVPMNYMVYFNFFACVLVPLLLMLGVYLRIFLAARRQLKQMESQPLPGE
P28222_5ht1b  -----ECVVNTDH-ILYTVYS-TVGAFYFPTLLLIALYGRIYVEARSRILKQTPNRTGKR
P28223_5ht2a  ----GSCLLADD---NFVLIG-SFVSFFIPLTIMVITYFLTIKSLQKEATLCVSDLGTRA
P41595_5ht2b  ----ITCVLTKERFGDFMLFG-SLAAFFTPLAIMIVTYFLTIHALQKKAYLVKNKPPQRL
P28335_5ht2c  ----TTCVLNDP---NFVLIG-SFVAFFIPLTIMVITYCLTIYVLRRQALMLLHGHTEEP
P14416_D2     ------CIIAN---PAFVVYS-SIVSFYVPFIVTLLVYIKIYIVLR-RRRKR------VN
P35462_D3     ------CSISN---PDFVIYS-SVVSFYLPFGVTVLVYARIYVVLKQRRRKR------IL
P21917_D4     ------CRLED---RDYVVYS-SVCSFFLPCPLMLLLYWATFRGLQRWEVARRAKLHGRA
P35367_H1     ----DKCETDFYDVTWFKVMT-AIINFYLPTLLMLWFYAKIYKAVRQHCQHRELINRSLP
P11229_M1     ----GQCYIQFLSQPIITFGT-AMAAFYLPVTVMCTLYWRIYRETENRARELAALQGSET

P07550_beta2  ----------------------------------VQNLSQVEQ-----------------
P29274_A2A    RARSTLQ-----------------------------------------------------
P28222_5ht1b  LTRAQ---------------------------LITDSPGSTSSVTSINSRVP-------D
P28223_5ht2a  KLASFSFLP-----------------------------QSSLSSEKLFQRSI--------
P41595_5ht2b  TWLTVSTVFQRDETPCSS----------PEKVAMLDGSRKDKALPNSGDETL--------
P28335_5ht2c  PGLSLDFLKCCKRNTAE---------------------EENSANPNQDQNAR--------
P14416_D2     TKRSSRAFRAHLRAP----------------LKGNCTHPEDMKLCTVIMKSNGSFPVNRR
P35462_D3     TRQNSQCNSVRPGFP----------------QQTLSPDPAHLEL-------KRYYSICQD
P21917_D4     PRRPSGPGPPSPTPP----------------APRLPQDPCGPDCAPPAPG-LPRGPCGPD
P35367_H1     SFSEIKLRPENPKGDAKK-------PGKESPWEVLKRKPKDAGGGSVLKSPSQTPKEMKS
P11229_M1     PGKG--------------GGSSSSSERSQPGAEGSPETPPGRCCRCCRAPRLLQAYSWKE
```

```
P07550_beta2  ------------------------------------------------------------
P29274_A2A    ------------------------------------------------------------
P28222_5ht1b  VPSES-GSPVYVNQVK-------------------------------------------
P28223_5ht2a  ------------------------------------------------------------
P41595_5ht2b  ------------------------------------------------------------
P28335_5ht2c  ------------------------------------------------------------
P14416_D2     RVEAARRAQELEMEMLSS---------------------TSPPERTRYSPIPPSHHQLTL
P35462_D3     TALGGPGFQERGGELKR-------------------------EEKTRNSLSPTIAPKLSL
P21917_D4     CAPAAPGLPPDPCGPDCA---------------------PPAPGLPQDPCGPDCAPPAPG
P35367_H1     PVVFSQEDDREVDKLYCFP--------LDIVHMQAAAEGSSRDYVAVNRSHGQLKTDEQG
P11229_M1     EEEED----EGSMESLTS-SEGEEPGSEVVIKMPMVDPEAQAPTKQPP--RSSPNTVKRP

P07550_beta2  ---------------------------------------------DGRTGHGLRRSSKFC
P29274_A2A    ------------------------------------------------------------
P28222_5ht1b  ---------------------------------------VRVS------DALLEKKKLMA
P28223_5ht2a  ------------------------------------------HREPGSYTGRRTMQSIS-
P41595_5ht2b  -----------------------------------------MRR-TSTIGKKSVQTIS-
P28335_5ht2c  -----------------------------------------RRKKKERRPRGTMQAIN-
P14416_D2     PDPSHHGLHSTPDSPAKPEKNGHAKDHPKIAKIFEIQTMPNGKTRTSLKTMSRR-KLSQQ
P35462_D3     ----------------------------------EVRKLSNGRLSTSLKLGPLQPRGVPL
P21917_D4     LPRGPCGPDCAPPAPGLPQDPCGPDCAPPAPGLPPDPCGSNCAPPDAVRAAALPPQTPPQ
P35367_H1     LNTHGASEISEDQMLGDSQSFSRTDSDTTTETAPGKGKLRSGSNTGLDYIKFTWKRLRSH
P11229_M1     TKK-----------------------------------GRDRAGKGQKPRGKEQLAKRKT

P07550_beta2  L----------KEHKALKTLGIIMGTFTLCWLPFFIVNIVHVIQD-----NLIRKEVYIL
P29274_A2A    -----------KEVHAAKSLAIIVGLFALCWLPLHIINCFTFFCPDCS---HAPLWLMYL
P28222_5ht1b  ----------ARERKATKTLGIILGAFIVCWLPFFIISLVMPICKDA---CWFHLAIFDF
P28223_5ht2a  -----------NEQKACKVLGIVFFLFVVMWCPFFITNIMAVICKES-CNEDVIGALLNV
P41595_5ht2b  -----------NEQRASKVLGIVFFLFLLMWCPFFITNITLVLCDS--CNQTTLQMLLEI
P28335_5ht2c  -----------NERKASKVLGIVFFVFLIMWCPFFITNILSVLCEKS-CNQKLMEKLLNV
P14416_D2     -----------KEKKATQMLAIVLGVFIICWLPFFITHILNIHCD-----CNIPPVLYSA
P35462_D3     -----------REKKATQMVAIVLGAFIVCWLPFFLTHVLNTHCQT----CHVSPELYSA
P21917_D4     TRRRRRAKITGRERKAMRVLPVVVGAFLLCWTPFFVVHITQALCPA----CSVPPRLVSA
P35367_H1     SRQYVSGLHMNRERKAAKQLGFIMAAFILCWIPYFIFFMVIAFCKN-----CCNEHLHMF
P11229_M1     FS-------LVKEKKAARTLSAILLAFILTWTPYNIMVLVSTFCKD-----CVPETLWEL
```

```
P07550_beta2  LNWIGYVNSGFNPLIY-CRSPDFRIAFQELLCLRRSSLK----AYGNGYSSNGNTGEQ--
P29274_A2A    AIVLSHTNSVVNPFIYAYRIREFRQTFRKIIRSHVLRQQEPFKAAGTSARVLAAHGSDGE
P28222_5ht1b  FTWLGYLNSLINPIIYTMSNEDFKQAFHKLIRFKCTS----------------------
P28223_5ht2a  FVWIGYLSSAVNPLVYTLFNKTYRSAFSRYIQCQYKENKKP-LQLILVNTIPALAYKSSQ
P41595_5ht2b  FVWIGYVSSGVNPLVYTLFNKTFRDAFGRYITCNYRATKSVKTLRKRSSKIYFRNPMAEN
P28335_5ht2c  FVWIGYVCSGINPLVYTLFNKIYRRAFSNYLRCNYKVEKKPPVRQIPRVAATALSGRELN
P14416_D2     FTWLGYVNSAVNPIIYTTFNIEFRKAFLKILHC---------------------------
P35462_D3     TTWLGYVNSALNPVIYTTFNIEFRKAFLKILSC---------------------------
P21917_D4     VTWLGYVNSALNPVIYTVFNAEFRNVFRKALRACC-------------------------
P35367_H1     TIWLGYINSTLNPLIYPLCNENFKKTFKRILHIRS-------------------------
P11229_M1     GYWLCYVNSTINPMCYALCNKAFRDTFRLLLLCRWDKRRWRKIPKRPGS----VHRTPSR

P07550_beta2  --SGYHVEQEKEN--------------------KLLCEDLPG-------------TEDFV
P29274_A2A    QVSLRLNGHPPGVWANGSAPHPERRPNGYALGLVSGGSAQESQGNTGLPDVELLSHELKG
P28222_5ht1b  ------------------------------------------------------------
P28223_5ht2a  LQMGQKK---------------------------------------------NSKQDAK
P41595_5ht2b  SKFFKKHGIRNGI-----------------------------------NPAMYQSPMRLR
P28335_5ht2c  VNIYRHT---------------------------------------------NEPVIEK
P14416_D2     ------------------------------------------------------------
P35462_D3     ------------------------------------------------------------
P21917_D4     ------------------------------------------------------------
P35367_H1     ------------------------------------------------------------
P11229_M1     QC----------------------------------------------------------

P07550_beta2  GHQGTVPSDNIDSQGRNCS---TNDSLL------
P29274_A2A    VCPEPPGLDDPLAQDGAGVS--------------
P28222_5ht1b  ----------------------------------
P28223_5ht2a  TTDNDCSMVALGKQHSEEASKDNSDGVNEKVSCV
P41595_5ht2b  SSTIQSSSIIL--LDTLLLTENEGDKTEEQVSYV
P28335_5ht2c  ASDNEPGIEMQ--VENLELPVNPSSVVSERISSV
P14416_D2     ----------------------------------
P35462_D3     ----------------------------------
P21917_D4     ----------------------------------
P35367_H1     ----------------------------------
P11229_M1     ----------------------------------
```

Appendix 3: Supplementary material for the manuscript "Predicting the structure of the dopamine D_3 receptor: An evaluation of virtual screening approaches to GPCR modeling" (Chapter 3)*

Table S1. Average ligand properties calculated in QikProp [1] as per method described in McRobb et al. [2]

Property/ Receptor	Molecular Weight (g/mol)	Rotatable Bonds	PSA ($Å^2$)	Calc logP	H-bond Donor	H-bond Acceptor	Solvent Accessible Volume ($Å^3$)
D_3 actives	381	4.8	49.0	3.9	0.8	5.5	1193
Decoys[a]	360	5.5	86.7	3.1	1.9	5.8	1117

[a] Schrödinger decoy library [3]

** Note: The referencing style for Appendix 3 is consistent with the prepared manuscript in Chapter 3, for submission to the* Journal of Computer-Aided Molecular Design.

Table S2. Sequence identity for the D_3R compared to the available templates at the time of model building (whole receptors, including loops – excluding C- and N-termini and ICL3).

Template	% Identities	% Positives	% Gaps
human β_2AR	32	51	6
turkey β_1AR	36	54	6
human A_{2A}R	27	45	11
bovine rhodopsin	23	40	7

Table S3. Virtual screening results for all 200 candidate structures; enrichment factors ($EF^{2\%}$, $EF^{5\%}$, $EF^{10\%}$) and area under the curve (AUC) (Models 1 to 5 highlighted in bold).

Model	$EF^{2\%}$	$EF^{5\%}$	$EF^{10\%}$	Enrichment AUC
1	**4.73**	**2.81**	**1.89**	**0.74**
2	**9.45**	**7.02**	**5.43**	**0.86**
3	**0**	**0.94**	**1.18**	**0.6**
4	**0**	**0.47**	**1.42**	**0.62**
5	**0**	**1.4**	**2.13**	**0.58**
6	0	1.87	1.18	0.61
7	5.91	4.21	4.25	0.76
8	0	0.94	1.42	0.5
9	1.18	1.4	1.18	0.66
10	1.18	1.87	1.89	0.66
11	1.18	2.34	2.13	0.7
12	3.54	4.68	4.02	0.71
13	8.27	6.09	3.07	0.62
14	5.91	6.09	3.54	0.62
15	0	2.81	3.54	0.69
16	7.09	3.74	2.6	0.6
17	1.18	0.94	0.95	0.64
18	2.36	3.28	2.13	0.69
19	2.36	2.34	2.36	0.69
20	7.09	4.68	3.78	0.71
21	0	2.34	3.07	0.75
22	1.18	2.81	3.07	0.71
23	5.91	6.55	4.49	0.8
24	1.18	2.81	3.54	0.71
25	2.36	3.28	2.36	0.58
26	1.18	0.94	2.13	0.63
27	0	2.34	2.36	0.64
28	1.18	2.34	2.13	0.58
29	5.91	3.28	3.54	0.66
30	1.18	0.94	1.42	0.67
31	1.18	0.94	2.13	0.62
32	3.54	2.81	2.13	0.65
33	2.36	3.28	3.54	0.65
34	2.36	2.81	2.84	0.6
35	1.18	1.4	3.31	0.69
36	4.73	3.74	2.36	0.69
37	2.36	1.4	2.13	0.61
38	1.18	0.94	1.42	0.63
39	0	1.87	2.36	0.68
40	0	0	0.47	0.51
41	0	0	1.42	0.58
42	0	0.47	0.71	0.56
43	1.18	0.94	0.95	0.56
44	0	0.94	1.18	0.6

45	4.73	1.87	1.65	0.54
46	0	0.94	0.95	0.61
47	1.18	0.94	1.42	0.65
48	1.18	0.94	0.95	0.51
49	0	0.94	1.42	0.61
50	1.18	0.47	1.18	0.57
51	1.18	0.94	0.71	0.5
52	1.18	1.87	1.65	0.61
53	2.36	1.4	1.65	0.64
54	1.18	1.4	1.89	0.57
55	1.18	2.81	2.36	0.58
56	0	1.87	1.42	0.68
57	3.54	2.81	3.07	0.71
58	2.36	3.28	2.6	0.63
59	1.18	0.94	1.18	0.68
60	1.18	2.34	2.36	0.63
61	3.54	4.21	3.07	0.62
62	1.18	2.34	1.89	0.58
63	0	1.87	1.65	0.62
64	0	0.47	1.65	0.69
65	1.18	2.34	1.89	0.64
66	0	0.94	1.18	0.58
67	2.36	2.34	1.89	0.56
68	0	2.81	2.13	0.62
69	0	0.47	1.89	0.62
70	1.18	2.81	2.36	0.57
71	1.18	1.4	1.42	0.63
72	1.18	3.74	3.07	0.72
73	5.91	3.28	2.36	0.68
74	2.36	3.28	2.13	0.6
75	1.18	1.87	2.36	0.67
76	0	1.4	1.65	0.65
77	0	0.47	1.18	0.61
78	4.73	2.81	1.89	0.61
79	0	2.34	2.6	0.63
80	4.73	3.28	2.13	0.6
81	2.36	1.4	1.89	0.72
82	0	1.87	1.18	0.58
83	0	0.47	1.65	0.66
84	0	0.94	1.65	0.66
85	1.18	1.4	1.65	0.66
86	0	1.4	1.65	0.67
87	0	1.87	1.89	0.57
88	0	0.94	2.13	0.64
89	2.36	2.34	2.13	0.56
90	1.18	1.87	1.89	0.54
91	1.18	1.4	1.42	0.55
92	2.36	1.87	2.36	0.68
93	1.18	1.4	1.42	0.63

94	0	0	0.71	0.62
95	0	1.4	1.89	0.54
96	0	0	1.18	0.62
97	2.36	2.81	2.13	0.61
98	3.54	2.81	3.31	0.78
99	3.54	2.81	3.07	0.71
100	1.18	3.74	2.84	0.61
101	1.18	0.94	1.18	0.63
102	1.18	2.34	1.89	0.57
103	3.54	2.34	1.89	0.69
104	4.73	3.74	2.84	0.65
105	3.54	2.81	3.31	0.66
106	1.18	2.81	2.13	0.54
107	1.18	0.47	2.13	0.69
108	0	2.81	3.07	0.71
109	1.18	1.4	1.42	0.65
110	2.36	2.34	2.13	0.63
111	1.18	1.87	1.89	0.64
112	1.18	2.81	2.36	0.64
113	4.73	3.74	2.36	0.61
114	0	0.94	0.71	0.48
115	3.54	1.87	1.18	0.58
116	1.18	1.87	1.65	0.62
117	1.18	2.81	1.65	0.63
118	1.18	1.87	2.13	0.7
119	0	2.34	1.89	0.73
120	1.18	2.81	3.07	0.68
121	3.54	3.28	1.89	0.64
122	1.18	1.4	1.18	0.55
123	0	1.87	2.36	0.65
124	3.54	1.87	2.13	0.68
125	2.36	1.4	0.95	0.52
126	0	0.47	2.36	0.66
127	1.18	1.87	2.36	0.61
128	1.18	1.4	1.65	0.58
129	1.18	1.4	2.6	0.7
130	3.54	1.87	2.36	0.6
131	0	0	1.18	0.52
132	0	1.87	2.36	0.71
133	1.18	0.94	1.18	0.57
134	0	0	0.95	0.58
135	1.18	2.34	2.6	0.61
136	0	0	0.95	0.6
137	4.73	3.28	3.07	0.69
138	1.18	2.34	2.84	0.7
139	3.54	1.87	2.13	0.7
140	2.36	1.87	2.13	0.65
141	0	0.94	1.89	0.6
142	1.18	2.34	1.89	0.57

143	3.54	3.74	3.07	0.64
144	4.73	2.81	1.89	0.74
145	3.54	2.34	1.89	0.69
146	3.54	4.21	4.02	0.67
147	1.18	0.47	2.13	0.69
148	0	2.81	3.07	0.71
149	1.18	2.81	2.13	0.54
150	2.36	2.34	2.84	0.55
151	1.18	1.87	2.13	0.62
152	0	0.47	0.71	0.51
153	2.36	2.34	1.89	0.65
154	3.54	2.34	1.42	0.49
155	2.36	0.94	1.65	0.68
156	0	0.47	1.42	0.61
157	0	0.94	1.18	0.55
158	2.36	2.34	2.84	0.73
159	1.18	4.21	2.84	0.63
160	0	0.47	1.89	0.68
161	0	0.47	0.71	0.57
162	1.18	1.4	0.71	0.53
163	0	0.47	0.47	0.51
164	0	0.47	0.71	0.55
165	0	0	0.71	0.58
166	1.18	0.47	1.18	0.66
167	3.54	4.21	3.07	0.64
168	0	0	0.71	0.49
169	1.18	1.4	1.42	0.57
170	0	0.94	0.47	0.52
171	0	0	0	0.53
172	2.36	0.94	0.95	0.57
173	1.18	0.94	1.18	0.67
174	0	0.47	0.95	0.6
175	3.54	3.74	2.6	0.76
176	0	0.47	1.65	0.55
177	2.36	3.74	3.07	0.67
178	3.54	3.74	2.84	0.66
179	3.54	3.74	3.78	0.6
180	5.91	3.74	3.31	0.67
181	3.54	3.74	2.84	0.62
182	1.18	1.87	2.36	0.59
183	1.18	1.87	2.84	0.62
184	0	1.4	1.65	0.6
185	2.36	3.28	3.31	0.66
186	2.36	1.87	0.95	0.57
187	1.18	2.34	2.13	0.59
188	0	0.47	0.71	0.54
189	0	2.34	1.89	0.62
190	0	1.4	1.65	0.64
191	5.91	5.15	3.07	0.63

192	0	1.87	1.89	0.47
193	0	0.47	1.65	0.66
194	4.73	1.87	1.65	0.61
195	0	1.87	2.13	0.57
196	3.54	2.81	1.65	0.66
197	1.18	1.4	2.13	0.67
198	1.18	2.81	1.65	0.6
199	0	0	0.71	0.5
200	2.36	2.34	2.6	0.68

Table S4. Binding site residues within 5 Å of eticlopride in the D_3R crystal structure and top 5 models using crystal structure numbering (*Ballesteros-Weinstein numbering is shown in brackets for residues in the transmembrane region).

Binding site residues	**Structure**					
	3PBL	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Val 82 (2.57)*	X	X	X	X	X	X
Val 86 (2.61)	X	X	X	X	X	X
Leu 89 (2.64)	X		X	X		
Phe 106 (3.28)	X			X	X	X
Val 107 (3.29)	X	X	X	X	X	X
Thr 108 (3.30)					X	
Asp 110 (3.32)	X	X	X	X	X	X
Val 111 (3.33)	X	X	X	X	X	X
Cys 114 (3.36)	X	X	X	X	X	X
Thr 115 (3.37)	X	X		X	X	X
Ile 118 (3.40)		X			X	
Val 164 (4.56)						X
Ser 165 (4.57)						X
Leu 168 (4.60)		X	X			X
Leu 169 (4.61)		X	X	X	X	X
Ser 182	X		X	X		
Ile 183	X	X	X	X	X	X
Ser 184		X	X	X	X	X
Asn 185		X	X	X	X	X
Pro 186 (5.36)		X	X	X	X	X
Phe 188 (5.38)	X	X	X	X		X
Val 189 (5.39)	X	X	X	X	X	
Ile 190 (5.40)	X			X		
Ser 192 (5.42)	X	X	X	X	X	X
Ser 193 (5.43)	X	X	X	X	X	X
Val 194 (5.44)				X		
Ser 196 (5.46)	X	X	X	X	X	X
Phe 197 (5.47)		X	X	X	X	X
Phe 338 (6.44)					X	
Trp 342 (6.48)	X	X	X	X	X	X
Phe 345 (6.51)	X	X	X	X	X	X
Phe 346 (6.52)	X	X	X	X	X	X
His 349 (6.55)	X	X	X	X	X	X
Val 350 (6.56)	X					
Tyr 365 (7.35)	X		X			
Ser 366 (7.36)		X	X	X		
Thr 369 (7.39)	X	X	X	X	X	X
Gly 372 (7.42)	X					
Tyr 373 (7.43)	X	X	X	X	X	X

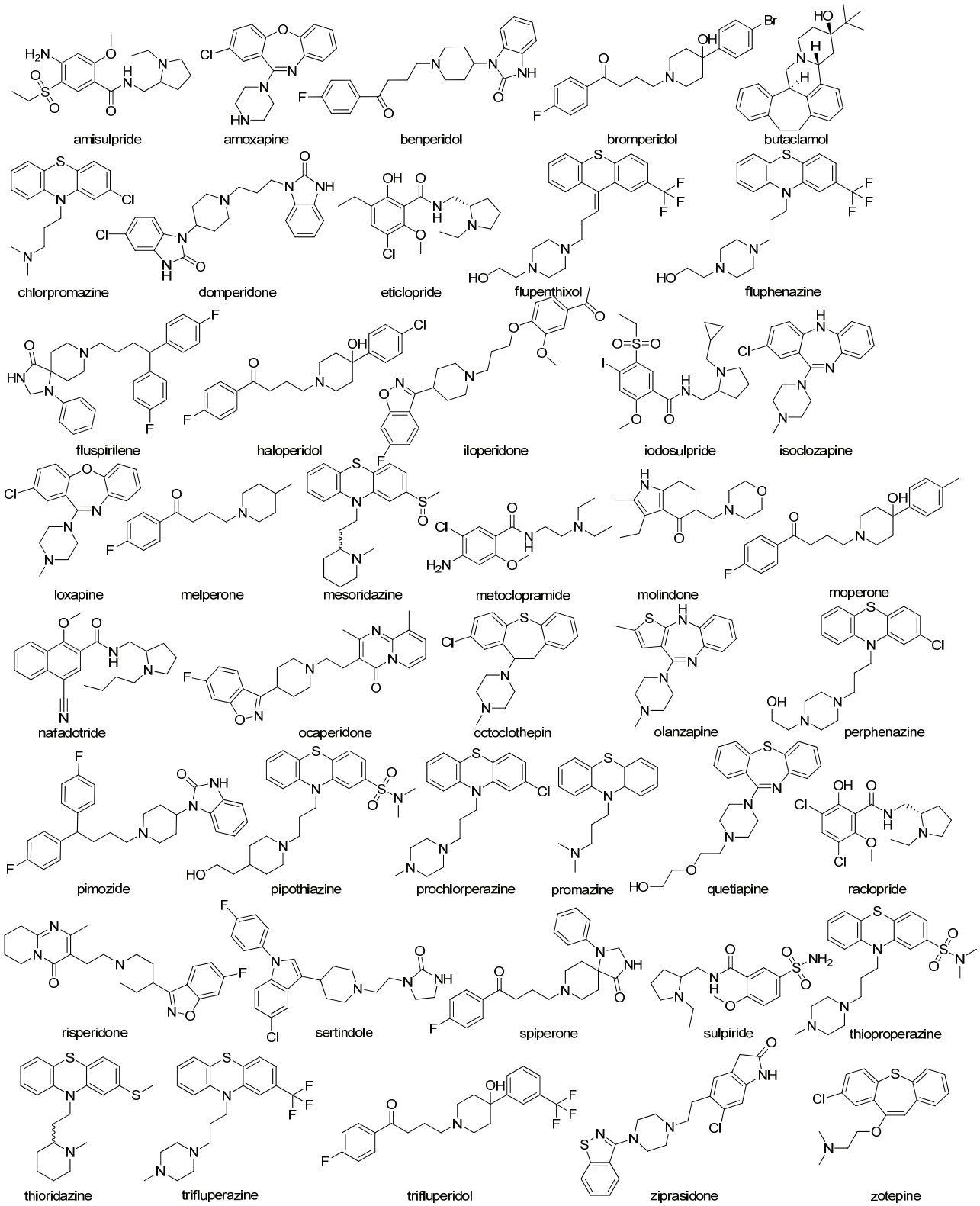


Figure S1. D_3R active compounds docked into each model during virtual screening.

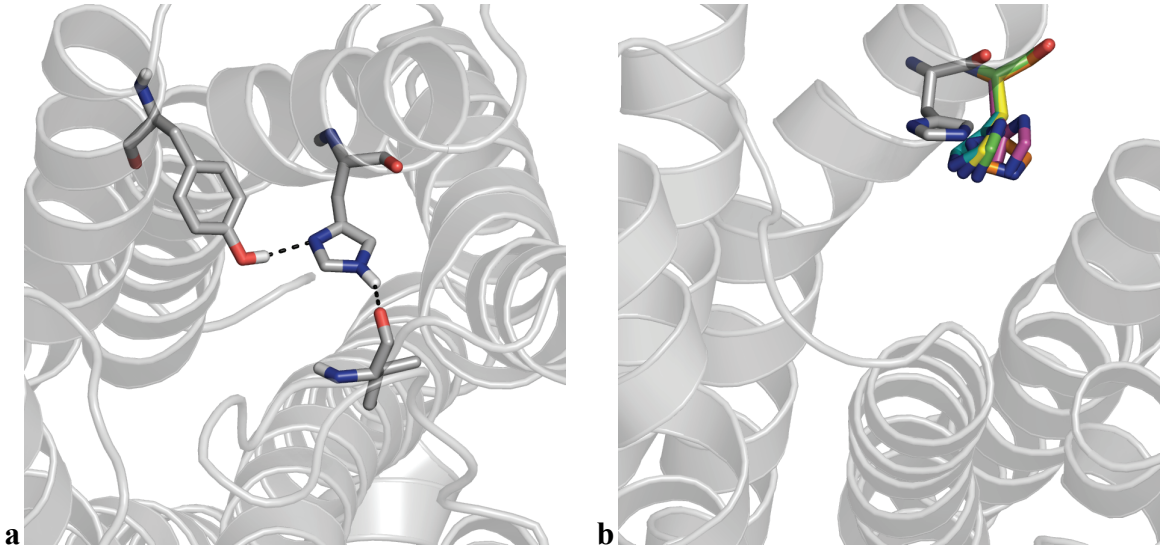


Figure S2. (a) His 6.55 in the eticlopride-D_3R crystal structure, displaying hydrogen bonds to Tyr 7.35 and Ile 183 and (b) placement of residue His 6.55 in crystal structure (grey), compared with Models 1-5 (color).

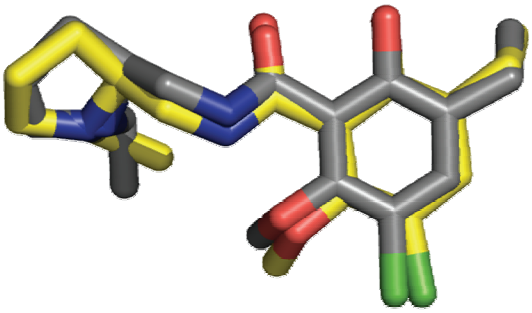


Figure S3. Cognate rigid ligand docking of eticlopride, after a conformational search (yellow), superimposed with crystal structure (grey).

References

1. QikProp (2008). version 3.1 edn. Schrödinger, LLC, New York, NY
2. McRobb FM, Capuano B, Crosby IT, Chalmers DK, Yuriev E (2010) Homology modeling and docking evaluation of aminergic G protein-coupled receptors. J Chem Inf Model 50:626-637.
3. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 47:1739-1749.

Appendix 4: Supporting information for "Homobivalent ligands of the atypical antipsychotic clozapine: Design, synthesis and pharmacological evaluation" (Chapter 4)*

** Note: The compound numbering for Appendix 4 is consistent with the prepared manuscript in Chapter 4.*

Experimental for the synthesis of bivalent ligands

Synthesis of clozapine N5 bivalent ligands

*1,10-Bis(8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl)decane-1,10-dione (**9b**)*. 1,10-Decanedioic acid (0.087 g, 0.428 mmol) and oxalyl chloride were reacted, followed by the addition of **1** (0.247 g, 0.757 mmol), *N*,*N*-diisopropylethylamine (145 µL, 0.832 mmol) and pyridine (0.095 mL, 1.18 mmol), as per general procedure A. Additional 1,10-decanedioyl dichloride (0.052 g, 0.216 mmol) was added. Column chromatography conditions: column 1 (5% methanol / acetone, until clozapine eluted then 10% methanol / chloroform), column 2 (gradient elution: from 5% methanol / chloroform to 10% methanol / chloroform, increasing in 1% increments), column 3 (1% ammonia / 9 % methanol / chloroform). Yielded **9b** as a white foam (0.115 g, 0.140 mmol, 37%). ^{1}H NMR (400 MHz, $CDCl_3$, 320 K) δ 7.49 (ddd, J = 8.0, 7.0, 1.9 Hz, 2H, H3′), 7.41-7.29 (m, 6H, H1′, H2′, H4′), 7.14 (d, J = 2.4 Hz, 2H, H9′), 7.10 (d, J = 8.2 Hz, 2H, H6′), 6.96 (m, 2H, H7′), 3.73 (m, 4H, H2″a, H6″a), 3.48 (m, 4H, H2″b, H6″b), 2.48 (m, 4H, H3″a, H5″a), 2.37 (m, 4H, H3″b, H5″b), 2.33-2.24 (m, 8H, CH_3, H2a, H9a), 2.13 (m, 2H, H2b, H9b), 1.52 (m, 4H, H3, H8), 1.23-1.12 (m, 8H, H4-H7). ^{13}C NMR (101 MHz, $CDCl_3$, 320 K) δ 173.9 (C_q), 160.6 (C_q), 146.5 (C_q), 145.2 (C_q), 134.1 (C_q), 133.8 (C_q), 132.1 (CH), 129.2 (CH), 127.92 (CH), 127.85 (CH), 127.0 (C_q), 126.3 (2 × CH), 123.3 (CH), 55.0 (CH_2), 47.0 (CH_2), 46.1 (CH_3), 33.7 (CH_2), 29.3 (CH_2), 29.2 (CH_2), 25.2 (CH_2). HPLC: t_R 12.40 min, 99% purity

(Method 2). LCMS (*m / z*): 819.1 $[M + H]^+$, 410.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{46}H_{53}Cl_2N_8O_2^+$ requires $[M + H]^+$ 819.3663; found 819.3694.

*1,12-Bis(8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl) dodecane-1,12-dione (**9c**)*. 1,12-Dodecanedioic acid (0.087 g, 0.376 mmol) and oxalyl chloride were reacted, followed by the addition of **1** (0.219 g, 0.671 mmol), *N*,*N*-diisopropylethylamine (163 µL, 0.936 mmol) and pyridine (0.085 mL, 1.05 mmol), as per general procedure A. Additional 1,12-dodecanedioyl dichloride (0.049 g, 0.182 mmol) was added. Column chromatography conditions: column 1 (5% methanol / acetone, until clozapine eluted then 10% methanol / chloroform), column 2 (gradient elution: from 5% methanol / chloroform to 10% methanol / chloroform, increasing in 1% increments), column 3 (1% ammonia / 9 % methanol / chloroform). Yielded **9c** as a white foam (0.080 g, 0.095 mmol, 28%). ^{1}H NMR (400 MHz, $CDCl_3$, 320 K) δ 7.50 (ddd, *J* = 8.0, 7.1, 1.9 Hz, 2H, H3′), 7.42-7.29 (m, 6H, H1′, H2′, H4′), 7.14 (d, *J* = 2.4 Hz, 2H, H9′), 7.10 (d, *J* = 8.3 Hz, 2H, H6′), 6.96 (dd, *J* = 8.3, 2.0 Hz, 2H, H7′), 3.72 (m, 4H, H2″a, H6″a), 3.48 (m, 4H, H2″b, H6″b), 2.47 (m, 4H, H3″a, H5″a), 2.36 (m, 4H, H3″b, H5″b), 2.32-2.23 (m, 8H, CH_3, H2a, H11a), 2.15 (m, 2H, H2b, H11b), 1.53 (m, 4H, H3, H10), 1.25-1.12 (m, 12H, H4-H9). ^{13}C NMR (101 MHz, $CDCl_3$, 320 K) δ 173.9 (C_q), 160.7 (C_q), 146.5 (C_q), 145.2 (C_q), 134.1 (C_q), 133.9 (C_q), 132.1 (CH), 129.2 (CH), 128.0 (CH), 127.9 (CH), 127.1 (C_q), 126.3 (2 × CH), 123.3 (CH), 55.0 (CH_2), 47.2 (CH_2), 46.2 (CH_3), 33.7 (CH_2), 29.5 (CH_2), 29.4 (2 × CH_2), 25.3 (CH_2). HPLC: t_R 13.04 min, 97% purity (Method 2). LCMS (*m / z*): 847.2 $[M + H]^+$, 424.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{48}H_{57}Cl_2N_8O_2^+$ requires $[M + H]^+$ 847.3976; found 847.4008.

*1,14-Bis(8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl) tetradecane-1,14-dione (**9d**)*. 1,14-Tetradecanedioic acid (0.089 g, 0.343 mmol) and oxalyl chloride were reacted, followed by the addition of **1** (0.205 g, 0.627 mmol), *N*,*N*-diisopropylethylamine (148 µL, 0.850 mmol) and pyridine (0.077 mL, 0.954 mmol), as per

general procedure A. Additional 1,14-tetradecanedioyl dichloride (0.046 g, 0.156 mmol) was added. Column chromatography conditions: column 1 (5% methanol / acetone, until clozapine eluted then 10% methanol / chloroform), column 2 (gradient elution: from 5% methanol / chloroform to 10% methanol / chloroform, increasing in 1% increments), column 3 (1% ammonia / 9 % methanol / chloroform). Yielded **9d** as a white foam (0.088 g, 0.101 mmol, 32%). ^{1}H NMR (400 MHz, $CDCl_3$, 320 K) δ 7.49 (ddd, *J* = 8.0, 7.1, 1.9 Hz, 2H, H3′), 7.42-7.29 (m, 6H, H1′, H2′, H4′), 7.14 (d, *J* = 2.4 Hz, 2H, H9′), 7.10 (d, *J* = 8.3 Hz, 2H, H6′), 6.96 (dd, *J* = 8.4, 2.1 Hz, 2H, H7′), 3.72 (m, 4H, H2″a, H6″a), 3.49 (m, 4H, H2″b, H6″b), 2.46 (m, 4H, H3″a, H5″a), 2.37 (m, 4H, H3″b, H5″b), 2.32-2.25 (m, 8H, CH_3, H2a, H13a), 2.16 (m, 2H, H2b, H13b), 1.54 (m, 4H, H3, H12), 1.28-1.13 (m, 16H, H4-H11). ^{13}C NMR (101 MHz, $CDCl_3$, 320 K) δ 174.0 (C_q), 160.7 (C_q), 146.5 (C_q), 145.2 (C_q), 134.1 (C_q), 133.8 (C_q), 132.1 (CH), 129.2 (CH), 128.0 (CH), 127.9 (CH), 127.1 (C_q), 126.3 (2 × CH), 123.3 (CH), 55.0 (CH_2), 47.2 (CH_2), 46.2 (CH_3), 33.7 (CH_2), 29.7 (CH_2), 29.6 (CH_2), 29.4 (2 × CH_2), 25.3 (CH_2). HPLC: t_R 8.69 min, 95% purity (Method 2). LCMS (*m / z*): 875.2 $[M + H]^+$, 438.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{50}H_{61}Cl_2N_8O_2^+$ requires $[M + H]^+$ 875.4289; found 875.4321.

*1,18-Bis(8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl) octadecane-1,18-dione (**9e**)*. 1,18-Octadecanedioic acid (0.118 g, 0.376 mmol) and oxalyl chloride were reacted, followed by the addition of **1** (0.218 g, 0.667 mmol), *N*,*N*-diisopropylethylamine (162 µL, 0.930 mmol) and pyridine (0.084 mL, 1.04 mmol), as per general procedure A. Additional 1,18-octadecanedioyl dichloride (0.059 g, 0.168 mmol) was added. Column chromatography conditions: column 1 (10% methanol / acetone, until clozapine eluted then 5% methanol / chloroform), column 2 (1% ammonia / 4 % methanol / chloroform). Yielded **9e** as a white foam (0.088 g, 0.094 mmol, 28%). ^{1}H NMR (400 MHz, $CDCl_3$, 320 K) δ 7.49 (ddd, *J* = 8.0, 7.2, 1.9 Hz, 2H, H3′), 7.42-7.30 (m, 6H, H1′, H2′, H4′), 7.15 (d, *J* = 2.3 Hz, 2H, H9′), 7.10 (d, *J* = 8.3 Hz, 2H, H6′), 6.96 (dd, *J* = 8.4, 2.1 Hz, 2H,

H7′), 3.73 (m, 4H, H2″a, H6″a), 3.49 (m, 4H, H2″b, H6″b), 2.47 (m, 4H, H3″a, H5″a), 2.37 (m, 4H, H3″b, H5″b), 2.33-2.24 (m, 8H, CH_3, H2a, H17a), 2.15 (m, 2H, H2b, H17b), 1.55 (m, 4H, H3, H16), 1.29-1.15 (m, 24H, H4-H15). ^{13}C NMR (101 MHz, $CDCl_3$, 320 K) δ 174.0 (C_q), 160.7 (C_q), 146.5 (C_q), 145.2 (C_q), 134.1 (C_q), 133.9 (C_q), 132.1 (CH), 129.2 (CH), 128.0 (CH), 127.9 (CH), 127.1 (C_q), 126.3 (2 × CH), 123.3 (CH), 55.0 (CH_2), 47.2 (CH_2), 46.2 (CH_3), 33.7 (CH_2), 29.83 (CH_2), 29.81 (CH_2), 29.77 (CH_2), 29.6 (CH_2), 29.43 (CH_2), 29.40 (CH_2), 25.3 (CH_2). HPLC: t_R 12.18 min, 99% purity (Method 1). LCMS (*m / z*): 931.2 $[M + H]^+$, 466.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{54}H_{69}Cl_2N_8O_2^+$ requires $[M + H]^+$ 931.4915; found 931.4922.

*1,20-Bis(8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl)icosane-1,20-dione (**9f**)*. 1,20-Icosanedioic acid (0.129 g, 0.377 mmol) and oxalyl chloride were reacted, followed by the addition of **1** (0.219 g, 0.671 mmol), *N*,*N*-diisopropylethylamine (160 µL, 0.919 mmol) and pyridine (0.084 mL, 1.04 mmol), as per general procedure A. Additional 1,20-icosanedioyl dichloride (0.068 g, 0.180 mmol) was added. Column chromatography conditions: column 1 (10% methanol / acetone, until clozapine eluted then 5% methanol / chloroform), column 2 (1% ammonia / 9 % methanol / chloroform), column 3 (gradient elution: from 5% methanol / chloroform to 10% methanol / chloroform, increasing in 1% increments). Yielded **9f** as a white foam (0.089 g, 0.093 mmol, 28%). ^{1}H NMR (400 MHz, $CDCl_3$, 320 K) δ 7.49 (ddd, *J* = 7.9, 7.0, 1.9 Hz, 2H, H3′), 7.41-7.29 (m, 6H, H1′, H2′, H4′), 7.15 (d, *J* = 2.4 Hz, 2H, H9′), 7.10 (d, *J* = 8.3 Hz, 2H, H6′), 6.96 (dd, *J* = 8.4, 2.2 Hz, 2H, H7′), 3.73 (m, 4H, H2″a, H6″a), 3.48 (m, 4H, H2″b, H6″b), 2.47 (m, 4H, H3″a, H5″a), 2.37 (m, 4H, H3″b, H5″b), 2.33-2.23 (m, 8H, CH_3, H2a, H19a), 2.16 (m, 2H, H2b, H19b), 1.54 (m, 4H, H3, H18), 1.29-1.14 (m, 28H, H4-H17). ^{13}C NMR (101 MHz, $CDCl_3$, 320 K) δ 174.0 (C_q), 160.6 (C_q), 146.5 (C_q), 145.2 (C_q), 134.1 (C_q), 133.9 (C_q), 132.1 (CH), 129.2 (CH), 128.0 (CH), 127.8 (CH), 127.1 (C_q), 126.3 (2 × CH), 123.3 (CH), 55.0 (CH_2), 47.1 (CH_2), 46.2 (CH_3), 33.7 (CH_2), 29.85 (CH_2), 29.84 (CH_2), 29.81 (CH_2), 29.77

(CH_2), 29.6 (CH_2), 29.42 (CH_2), 29.39 (CH_2), 25.3 (CH_2). HPLC: t_R 10.55 min, 97% purity (Method 2). LCMS (*m / z*): 959.3 $[M + H]^+$, 480.3 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{56}H_{73}C_{l2}N_8O_2^+$ requires $[M + H]^+$ 959.5228; found 959.5264.

Synthesis of clozapine hydrazide bivalent ligands

N^1,N^8*-Bis(8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl) octanediamide (**10b**)*. 1,8-Octanedioic acid (0.061 g, 0.351 mmol) and oxalyl chloride were reacted using Method B, followed by the addition of **5** (0.210 g, 0.616 mmol) and pyridine (0.080 mL, 0.991 mmol), as per general procedure B. Additional 1,8-octanedioyl dichloride (0.022 g, 0.103 mmol) was added. Column chromatography conditions: column 1 (gradient elution: from 2% methanol / chloroform to 10% methanol / chloroform, increasing in 2% increments), column 2 (gradient elution: from 2% methanol / chloroform to 5% methanol / chloroform, increasing in 1% increments). Yielded **10b** as an off-white foam (0.141 g, 0.171 mmol, 56%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.46 (s, 2H, NH), 8.56 (d, *J* = 2.3 Hz, 2H, H9′), 7.78 (ddd, *J* = 8.3, 0.9, 0.9 Hz, 2H, H1′/H4′), 7.41-7.33 (m, 4H, H3′/H2′, H4′/H1′), 7.36 (d, *J* = 8.5 Hz, 2H, H6′), 7.13 (dd, *J* = 8.5, 2.4 Hz, 2H, H7′), 7.12 (m, 2H, H2′/H3′), 3.55 (m, 8H, H2″, H6″), 2.66 (m, 8H, H3″, H5″), 2.38 (s, 6H, CH_3), 2.21 (t, *J* = 7.4 Hz, 4H, H2, H7), 1.55 (app p, *J* = 6.9 Hz, 4H, H3, H6), 1.23 (m, 4H, H4, H5). ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.1 (C_q), 153.8 (C_q), 141.9 (C_q), 133.9 (C_q), 132.8 (C_q), 128.3 (CH), 126.9 (C_q), 124.7 (CH), 123.7 (CH), 122.7 (CH), 121.8 (CH), 120.7 (CH), 116.2 (C_q), 110.6 (CH), 54.7 (CH_2), 49.3 (CH_2), 46.3 (CH_3), 38.1 (CH_2), 29.0 (CH_2), 25.3 (CH_2). HPLC: t_R 9.98 min, 95% purity (Method 1). LCMS (*m / z*): 821.2 $[M + H]^+$, 411.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{44}H_{51}Cl_2N_{10}O_2^+$ requires $[M + H]^+$ 821.3568; found 821.3552.

N^1,N^{10}*-Bis(8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl) decanediamide (**10c**)*. 1,10-Decanedioic acid (0.050 g, 0.246 mmol) and oxalyl chloride were reacted using Method B, followed by the addition of **5** (0.154 g, 0.452 mmol) and pyridine (0.055 mL, 0.681 mmol), as per general procedure B. Additional 1,10-decanedioyl

dichloride (0.030 g, 0.124 mmol) was added. Column chromatography conditions: 1% ammonia / 4% methanol / chloroform. Yielded **10c** as an off-white foam (0.111 g, 0.130 mmol, 58%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.49 (s, 2H, NH), 8.57 (d, *J* = 2.3 Hz, 2H, H9′), 7.79 (ddd, *J* = 8.2, 0.9, 0.9 Hz, 2H, H1′/H4′), 7.43-7.35 (m, 4H, H3′/H2′, H4′/H1′), 7.36 (d, *J* = 8.5 Hz, 2H, H6′), 7.15 (m, 2H, H2′/H3′), 7.13 (dd, *J* = 8.5, 2.4 Hz, 2H, H7′), 3.55 (m, 8H, H2″, H6″), 2.67 (m, 8H, H3″, H5″), 2.39 (s, 6H, CH_3), 2.24 (t, *J* = 7.4 Hz, 4H, H2, H9), 1.58 (app p, *J* = 7.3 Hz, 4H, H3, H8), 1.26-1.09 (m, 8H, H4-H7). ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.3 (C_q), 153.8 (C_q), 141.9 (C_q), 133.9 (C_q), 132.9 (C_q), 128.3 (CH), 126.9 (C_q), 124.7 (CH), 123.7 (CH), 122.7 (CH), 121.9 (CH), 120.8 (CH), 116.3 (C_q), 110.7 (CH), 54.8 (CH_2), 49.4 (CH_2), 46.4 (CH_3), 38.2 (CH_2), 29.3 (CH_2), 29.2 (CH_2), 25.5 (CH_2). HPLC: t_R 10.45 min, >99% purity (Method 1). LCMS (*m / z*): 849.2 $[M + H]^+$, 425.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{46}H_{56}Cl_2N_{10}O_2^{2+}$ requires $[M + 2H]^{2+}$ 425.1977; found 425.1975.

N^{1},N^{12}*-Bis(8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl) dodecanediamide (**10d**)*. 1,12-Dodecanedioic acid (0.057 g, 0.248 mmol) and oxalyl chloride were reacted using Method B, followed by the addition of **5** (0.149 g, 0.436 mmol) and pyridine (0.055 mL, 0.681 mmol), as per general procedure B. Additional 1,12-dodecanedioyl dichloride (0.032 g, 0.122 mmol) was added. Column chromatography conditions: gradient elution: from 2% methanol / chloroform to 6% methanol / chloroform, increasing in 1% increments. Yielded **10d** as an off-white foam (0.133 g, 0.151 mmol, 69%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.45 (s, 2H, NH), 8.58 (d, *J* = 2.3 Hz, 2H, H9′), 7.79 (ddd, *J* = 8.3, 0.9, 0.9 Hz, 2H, H1′/H4′), 7.44-7.38 (m, 4H, H3′/H2′, H4′/H1′), 7.36 (d, *J* = 8.5 Hz, 2H, H6′), 7.16 (m, 2H, H2′/H3′), 7.14 (dd, *J* = 8.5, 2.4 Hz, 2H, H7′), 3.59 (m, 8H, H2″, H6″), 2.70 (m, 8H, H3″, H5″), 2.42 (s, 6H, CH_3), 2.25 (t, *J* = 7.4 Hz, 4H, H2, H11), 1.61 (app p, *J* = 7.3 Hz, 4H, H3, H10), 1.29-1.07 (m, 12H, H4-H9). ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.4 (C_q), 153.7 (C_q), 142.0 (C_q), 134.0 (C_q), 133.0 (C_q), 128.4 (CH), 127.0 (C_q), 124.8 (CH), 123.7 (CH), 122.8 (CH), 121.8 (CH), 120.8 (CH), 116.3 (C_q), 110.8 (CH), 54.7

(CH_2), 49.3 (CH_2), 46.3 (CH_3), 38.4 (CH_2), 29.49 (CH_2), 29.47 (CH_2), 29.3 (CH_2), 25.6 (CH_2). HPLC: t_R 10.93 min, 97% purity (Method 1). LCMS (*m / z*): 877.1 $[M + H]^+$, 439.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{48}H_{59}Cl_2N_{10}O_2^+$ requires $[M + H]^+$ 877.4194; found 877.4188.

N[1],N[14]*-Bis(8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl) tetradecanediamide (**10e**)*. 1,14-Tetradecanedioic acid (0.060 g, 0.232 mmol) and oxalyl chloride were reacted using Method B, followed by the addition of **5** (0.145 g, 0.425 mmol) and pyridine (0.055 mL, 0.681 mmol), as per general procedure B. Column chromatography conditions: column 1 (gradient elution: from 2% methanol / chloroform to 10% methanol / chloroform, increasing in 2% increments), column 2 (gradient elution: from 2% methanol / chloroform to 6% methanol / chloroform, increasing in 1% increments). Yielded **10e** as an off-white foam (0.137 g, 0.152 mmol, 71%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.50 (s, 2H, NH), 8.58 (d, *J* = 2.3 Hz, 2H, H9′), 7.79 (ddd, *J* = 8.3, 0.9, 0.9 Hz, 2H, H1′/H4′), 7.45-7.38 (m, 4H, H3′/H2′, H4′/H1′), 7.36 (d, *J* = 8.5 Hz, 2H, H6′), 7.15 (m, 2H, H2′/H3′), 7.13 (dd, *J* = 8.5, 2.4 Hz, 2H, H7′), 3.57 (m, 8H, H2″, H6″), 2.69 (m, 8H, H3″, H5″), 2.41 (s, 6H, CH_3), 2.26 (t, *J* = 7.4 Hz, 4H, H2, H13), 1.62 (app p, *J* = 7.3 Hz, 4H, H3, H12), 1.31-1.10 (m, 16H, H4-H11). ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.4 (C_q), 153.8 (C_q), 141.9 (C_q), 134.0 (C_q), 132.9 (C_q), 128.3 (CH), 126.9 (C_q), 124.7 (CH), 123.6 (CH), 122.7 (CH), 121.8 (CH), 120.8 (CH), 116.3 (C_q), 110.7 (CH), 54.8 (CH_2), 49.3 (CH_2), 46.3 (CH_3), 38.3 (CH_2), 29.6 (CH_2), 29.52 (CH_2), 29.46 (CH_2), 29.3 (CH_2), 25.6 (CH_2). HPLC: t_R 11.34 min, 95% purity (Method 1). LCMS (*m / z*): 905.2 $[M + H]^+$, 453.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{50}H_{63}Cl_2N_{10}O_2^+$ requires $[M + H]^+$ 905.4507; found 905.4495.

N[1],N[18]*-Bis(8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl) octadecanediamide (**10f**)*. 1,18-Octadecanedioic acid (0.078 g, 0.242 mmol) and oxalyl chloride were reacted using Method B, followed by the addition of **5** (0.149 g, 0.436 mmol), *N*,*N*-diisopropylethylamine (105 µL, 0.603 mmol) and pyridine (0.055 mL, 0.681 mmol), as per general procedure B. Additional 1,18-octadecanedioyl dichloride (0.032 g, 0.091 mmol)

was added. Column chromatography conditions: gradient elution: from 2% methanol / chloroform to 5% methanol / chloroform, increasing in 1% increments (using a preconditioned column with 0.5% ammonia / 1.5% methanol / chloroform). Yielded **10f** as an off-white foam (0.097 g, 0.101 mmol, 46%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.50 (s, 2H, NH), 8.58 (d, *J* = 2.3 Hz, 2H, H9′), 7.80 (ddd, *J* = 8.2, 0.9, 0.9 Hz, 2H, H1′/H4′), 7.44-7.38 (m, 4H, H3′/H2′, H4′/H1′), 7.37 (d, *J* = 8.5 Hz, 2H, H6′), 7.15 (m, 2H, H2′/H3′), 7.13 (dd, *J* = 8.5, 2.4 Hz, 2H, H7′), 3.56 (m, 8H, H2″, H6″), 2.68 (m, 8H, H3″, H5″), 2.40 (s, 6H, CH_3), 2.26 (t, *J* = 7.4 Hz, 4H, H2, H17), 1.63 (app p, *J* = 7.3 Hz, 4H, H3, H16), 1.31-1.13 (m, 24H, H4-H15). ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.4 (C_q), 153.8 (C_q), 142.0 (C_q), 134.0 (C_q), 133.0 (C_q), 128.3 (CH), 126.9 (C_q), 124.7 (CH), 123.7 (CH), 122.7 (CH), 121.9 (CH), 120.8 (CH), 116.3 (C_q), 110.7 (CH), 54.8 (CH_2), 49.4 (CH_2), 46.4 (CH_3), 38.4 (CH_2), 29.79 (CH_2), 29.77 (CH_2), 29.7 (CH_2), 29.6 (CH_2), 29.5 (CH_2), 29.3 (CH_2), 25.6 (CH_2). HPLC: t_R 10.09 min, >99% purity (Method 2). LCMS (*m / z*): 961.2 $[M + H]^+$, 481.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{54}H_{72}Cl_2N_{10}O_2^{2+}$ requires $[M + 2H]^{2+}$ 481.2603; found 481.2600.

N^{1},N^{20}-*Bis(8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl) icosanediamide (**10g**)*. 1,20-Icosanedioic acid (0.084 g, 0.246 mmol) and oxalyl chloride were reacted using Method B, followed by the addition of **5** (0.151 g, 0.441 mmol), *N*,*N*-diisopropylethylamine (105 µL, 0.603 mmol) and pyridine (0.055 mL, 0.681 mmol), as per general procedure B. Additional 1,20-icosanedioyl dichloride (0.047 g, 0.123 mmol) was added. Column chromatography conditions: column 1 (1% ammonia / 2% methanol / chloroform), column 2 (gradient elution: from 2% methanol / chloroform to 6% methanol / chloroform, increasing in 1% increments), column 3 (1% ammonia / 4 % methanol / chloroform). Yielded **10g** as an off-white foam (0.084 g, 0.085 mmol, 38%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.49 (s, 2H, NH), 8.58 (d, *J* = 2.3 Hz, 2H, H9′), 7.80 (ddd, *J* = 8.2, 0.9, 0.9 Hz, 2H, H1′/H4′), 7.45-7.38 (m, 4H, H3′/H2′, H4′/H1′), 7.36 (d, *J* = 8.5 Hz, 2H, H6′), 7.16 (m, 2H, H2′/H3′), 7.13 (dd, *J* = 8.5, 2.4 Hz, 2H, H7′), 3.58 (m, 8H, H2″, H6″), 2.69 (m, 8H,

H3″, H5″), 2.41 (s, 6H, CH_3), 2.26 (t, *J* = 7.4 Hz, 4H, H2, H19), 1.63 (app p, *J* = 7.3 Hz, 4H, H3, H18), 1.31-1.14 (m, 28H, H4-H17). ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.4 (C_q), 153.8 (C_q), 141.9 (C_q), 134.0 (C_q), 132.9 (C_q), 128.3 (CH), 126.9 (C_q), 124.7 (CH), 123.7 (CH), 122.7 (CH), 121.9 (CH), 120.8 (CH), 116.3 (C_q), 110.7 (CH), 54.8 (CH_2), 49.4 (CH_2), 46.4 (CH_3), 38.4 (CH_2), 29.81 (CH_2), 29.79 (CH_2), 29.76 (CH_2), 29.7 (CH_2), 29.6 (CH_2), 29.5 (CH_2), 29.3 (CH_2), 25.6 (CH_2). HPLC: t_R 10.14 min, 99% purity (Method 2). LCMS (*m / z*): 989.3 $[M + H]^+$, 495.3 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{56}H_{75}Cl_2N_{10}O_2^+$ requires $[M + H]^+$ 989.5446; found 989.5491.

*4,4'-(Piperazine-1,4-diyl)bis(*N*-(8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4] diazepin-5-yl)-4-oxobutanamide) (**14a**).* 4,4'-(Piperazine-1,4-diyl)bis(4-oxobutanoic acid) (**13a**, 0.090 g, 0.316 mmol) and oxalyl chloride were reacted using Method B, followed by the addition of **5** (0.191 g, 0.558 mmol) and pyridine (70 μL, 0.867 mmol), as per general experimental B. Additional 4,4'-(piperazine-1,4-diyl)bis(4-oxobutanoyl chloride) (0.052 g, 0.161 mmol) was added. Column chromatography conditions: 1% ammonia / 4% methanol / chloroform. Yielded **14a** as an off-white foam (0.063 g, 0.067 mmol, 24%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.70 (s, 2H, NH), 8.56 (s, 2H, H9″), 7.79 (ddd, *J* = 8.2, 0.9, 0.9 Hz, 2H, H1″/H4″), 7.42-7.36 (m, 4H, H3″/H2″, H4″/H1″), 7.36 (d, *J* = 8.5 Hz, 2H, H6″), 7.15 (m, 2H, H2″/H3″), 7.13 (dd, *J* = 8.5, 2.4 Hz, 2H, H7″), 3.59 (m, 8H, H2‴, H6‴), 3.55-3.46 (m, 4H, piperazine spacer), 3.42-3.36 (m, 4H, piperazine spacer), 2.69 (m, 8H, H3‴, H5‴), 2.64 (s, 8H, H2′, H3′), 2.40 (s, 6H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$) A mixture of amide rotamers.[10] δ 170.3 (C_q), 170.1 (C_q), 170.0 (C_q), 153.9 (C_q), 141.9 (C_q), 133.9 (C_q), 132.9 (C_q), 128.2 (CH), 126.9 (C_q), 124.8 (CH), 123.7 (CH), 122.6 (CH), 121.8 (CH), 120.7 (CH), 116.4 (C_q), 110.7 (CH), 54.8 (CH_2), 49.4 (CH_2), 46.3 (CH_3), 45.2 (CH_2), 45.0 (CH_2), 41.53 (CH_2), 41.45 (CH_2), 32.7 (CH_2), 32.6 (CH_2), 28.0 (CH_2). HPLC: t_R 7.22 min, 96% purity (Method 2). LCMS (*m / z*): 933.2 $[M + H]^+$, 467.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{48}H_{55}Cl_2N_{12}O_4^+$ requires $[M + H]^+$ 933.3841; found 933.3803.

*5,5'-(Piperazine-1,4-diyl)bis(*N*-(8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4] diazepin-5-yl)-5-oxopentanamide) (**14b**)*. 5,5'-(Piperazine-1,4-diyl)bis(5-oxopentanoic acid) (**13b**, 0.099 g, 0.315 mmol) and oxalyl chloride were reacted using Method B, followed by the addition of **5** (0.200 g, 0.584 mmol) and pyridine (72 μL, 0.892 mmol), as per general experimental B. Column chromatography conditions: column 1 (gradient elution: from 2% methanol / chloroform to 10% methanol / chloroform, increasing in 2% increments), column 2 (1% ammonia / 4% methanol / chloroform). Yielded **14b** as an off-white foam (0.101 g, 0.105 mmol, 36%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.66 (br s, 1H, NH), 9.63 (br s, 1H, NH), 8.55 (d, *J* = 2.4 Hz, 2H, H9″), 7.79 (ddd, *J* = 8.2, 0.9, 0.9 Hz, 2H, H1″/H4″), 7.45-7.37 (m, 4H, H3″/H2″, H4″/H1″), 7.38 (d, *J* = 8.6 Hz, 2H, H6″), 7.18-7.13 (m, 4H, H2″/H3″, H7″), 3.57 (m, 8H, H2‴, H6‴), 3.52-3.43 (m, 4H, piperazine spacer), 3.34-3.25 (m, 4H, piperazine spacer), 2.71 (m, 8H, H3‴, H5‴), 2.44-2.31 (m, 14H, CH_3, H2′, H4′), 1.98 (app p, *J* = 7.1 Hz, 4H, H3′). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.9 (C_q), 170.7 (C_q), 153.8 (C_q), 141.8 (C_q), 133.6 (C_q), 132.8 (C_q), 128.3 (CH), 127.0 (C_q), 124.6 (CH), 123.8 (CH), 122.7 (CH), 121.9 (CH), 120.8 (CH), 116.3 (C_q), 110.6 (CH), 54.7 (CH_2), 49.3 (CH_2), 46.3 (CH_3), 45.3 (CH_2), 45.1 (CH_2), 41.5 (CH_2), 41.3 (CH_2), 36.7 (CH_2), 32.0 (CH_2), 31.9 (CH_2), 20.6 (CH_2), 20.5 (CH_2). HPLC: t_R 9.72 min, 98% purity (Method 1). LCMS (*m / z*): 961.1 $[M + H]^+$, 481.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{50}H_{60}Cl_2N_{12}O_4^{2+}$ requires $[M + 2H]^{2+}$ 481.2113; found 481.2122.

*Ethane-1,2-diyl bis(4-((8-chloro-11-(4-methylpiperazin-1-yl)-5*H*-dibenzo[*b,e*][1,4]diazepin-5-yl)amino)-4-oxobutanoate) (**16**)*. 4,4'-(Ethane-1,2-diylbis(oxy))bis(4-oxobutanoic acid) (**15**, 0.090 g, 0.343 mmol) and oxalyl chloride were reacted using Method B, followed by the addition of **5** (0.204 g, 0.596 mmol) and pyridine (76 μL, 0.942 mmol) as per general experimental B. Column chromatography conditions: column 1 (0.5% ammonia / 2.5% methanol / chloroform), column 2 (gradient elution: from 2% methanol / chloroform to 5% methanol / chloroform, increasing in 1% increments). Yielded **16** as an off-white foam

(0.124 g, 0.136 mmol, 46%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.64 (s, 2H, NH), 8.52 (d, *J* = 2.3 Hz, 2H, H9″), 7.78 (ddd, *J* = 8.3, 0.9, 0.9 Hz, 2H, H1″/H4″), 7.43-7.36 (m, 4H, H3″/H2″, H4″/H1″), 7.36 (d, *J* = 8.5 Hz, 2H, H6″), 7.14 (ddd, *J* = 8.1, 6.3, 1.8 Hz, 2H, H2″/H3″), 7.12 (dd, *J* = 8.5, 2.4 Hz, 2H, H7″), 4.16 (s, 4H, H1, H2), 3.58 (m, 8H, H2‴, H6‴), 2.72 (m, 8H, H3‴, H5‴), 2.66 (m, 4H, H2′/H3′), 2.57 (m, 4H, H3′/H2′), 2.42 (s, 6H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.3 (C_q), 169.4 (C_q), 153.7 (C_q), 141.8 (C_q), 133.6 (C_q), 132.7 (C_q), 128.3 (CH), 126.9 (C_q), 124.6 (CH), 123.7 (CH), 122.6 (CH), 121.8 (CH), 120.7 (CH), 116.3 (C_q), 110.6 (CH), 62.3 (CH_2), 54.6 (CH_2), 49.2 (CH_2), 46.1 (CH_3), 32.2 (CH_2), 28.9 (CH_2). HPLC: t_R 10.17 min, >99% purity (Method 1). LCMS (*m / z*): 909.1 [M + H]$^{+}$, 455.2 [M + 2H]$^{2+}$. HRMS (*m / z*): $C_{46}H_{51}Cl_2N_{10}O_6^{+}$ requires [M + H]$^{+}$ 909.3365; found 909.3328.

Synthesis of clozapine propylamine bivalent ligands

N^{1},N^{8}*-Bis(3-(4-(8-chloro-5*H*-dibenzo[*b,e*][1,4]diazepin-11-yl)piperazin-1-yl)propyl) octanediamide (**11b**)*. 1,8-Octanedioic acid (0.030 g, 0.169 mmol) and oxalyl chloride were reacted, followed by the addition of **7b** (0.127 g, 0.343 mmol) and pyridine (0.039 mL, 0.483 mmol), as per general procedure C. Column chromatography conditions: 10% methanol / chloroform. Yielded **11b** as a yellow foam (0.062 g, 0.071 mmol, 42%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.30 (m, 2H, H3‴), 7.24 (dd, *J* = 7.7, 1.4 Hz, 2H, H1‴), 7.05 (d, *J* = 2.4 Hz, 2H, H9‴), 7.03-6.97 (m, 4H, H2‴, CONH), 6.85 (dd, *J* = 7.9, 0.8 Hz, 2H, H4‴), 6.82 (dd, *J* = 8.3, 2.4 Hz, 2H, H7‴), 6.65 (d, *J* = 8.3 Hz, 2H, H6‴), 5.21 (s, 2H, H5‴), 3.45 (s, 8H, H3″, H5″), 3.32 (td, *J* = 5.9, 5.9 Hz, 4H, H1′), 2.57 (s, 8H, H2″, H6″), 2.52 (t, *J* = 6.3 Hz, 4H, H3′), 2.11 (m, 4H, H2, H7), 1.69 (app p, *J* = 6.3 Hz, 4H, H2′), 1.59 (app p, *J* = 7.0 Hz, 4H, H3, H6), 1.31 (m, 4H, H4, H5). ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.2 (C_q), 163.1 (C_q), 153.1 (C_q), 141.8 (C_q), 140.8 (C_q), 132.2 (CH), 130.3 (CH), 129.0 (C_q), 126.8 (CH), 123.43 (CH), 123.36 (C_q), 123.2 (CH), 120.35 (CH), 120.33 (CH), 57.6 (CH_2), 53.2 (CH_2), 47.5 (CH_2), 39.4 (CH_2), 37.0 (CH_2), 29.1 (CH_2), 25.9 (CH_2), 25.2 (CH_2). HPLC: t_R 7.94 min,

>99% purity (Method 1). LCMS (*m / z*): 877.2 $[M + H]^+$, 439.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{48}H_{59}Cl_2N_{10}O_2^+$ requires $[M + H]^+$ 877.4194; found 877.4199.

N[1],N[10]*-Bis(3-(4-(8-chloro-5*H*-dibenzo[*b,e*][1,4]diazepin-11-yl)piperazin-1-yl) propyl)decanediamide (**11c**)*. 1,10-Decanedioic acid (0.036 g, 0.176 mmol) and oxalyl chloride were reacted, followed by the addition of **7b** (0.131 g, 0.354 mmol) and pyridine (0.040 mL, 0.496 mmol), as per general procedure C. Additional 1,10-decanedioyl dichloride (0.023 g, 0.094 mmol) was added. Column chromatography conditions: column 1 (10% methanol / chloroform), column 2 (1% ammonia / 4% methanol / chloroform), column 3 (10% methanol / chloroform). Yielded **11c** as a yellow foam (0.071 g, 0.078 mmol, 45%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.29 (ddd, *J* = 7.8, 7.3, 1.5 Hz, 2H, H3‴), 7.24 (dd, *J* = 7.7, 1.5 Hz, 2H, H1‴), 7.06 (d, *J* = 2.4 Hz, 2H, H9‴), 7.00 (ddd, *J* = 7.7, 7.4, 1.1 Hz, 2H, H2‴), 6.94 (br t, *J* = 5.0 Hz, 2H, CONH), 6.85 (dd, *J* = 8.0, 0.8 Hz, 2H, H4‴), 6.82 (dd, *J* = 8.3, 2.4 Hz, 2H, H7‴), 6.65 (d, *J* = 8.3 Hz, 2H, H6‴), 5.20 (s, 2H, H5‴), 3.43 (m, 8H, H3″, H5″), 3.35 (td, *J* = 5.9, 5.8 Hz, 4H, H1′), 2.55 (m, 8H, H2″, H6″), 2.51 (t, *J* = 6.3 Hz, 4H, H3′), 2.11 (m, 4H, H2, H9), 1.69 (app p, *J* = 6.2 Hz, 4H, H2′), 1.58 (app p, *J* = 7.1 Hz, 4H, H3, H8), 1.32-1.23 (m, 8H, H4-H7). ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.2 (C_q), 163.1 (C_q), 153.2 (C_q), 141.8 (C_q), 140.8 (C_q), 132.1 (CH), 130.3 (CH), 129.0 (C_q), 126.9 (CH), 123.5 (C_q), 123.4 (CH), 123.1 (CH), 120.32 (CH), 120.28 (CH), 57.8 (CH_2), 53.3 (CH_2), 47.6 (CH_2), 39.6 (CH_2), 37.1 (CH_2), 29.4 (CH_2), 29.3 (CH_2), 26.0 (CH_2), 25.3 (CH_2). HPLC: t_R 8.45 min, >99% purity (Method 1). LCMS (*m / z*): 905.2 $[M + H]^+$, 453.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{50}H_{63}Cl_2N_{10}O_2^+$ requires $[M + H]^+$ 905.4507; found 905.4550.

N[1],N[12]*-Bis(3-(4-(8-chloro-5*H*-dibenzo[*b,e*][1,4]diazepin-11-yl)piperazin-1-yl)propyl) dodecanediamide (**11d**)*. 1,12-Dodecanedioic acid (0.040 g, 0.173 mmol) and oxalyl chloride were reacted, followed by the addition of **7b** (0.131 g, 0.354 mmol) and pyridine (0.040 mL, 0.496 mmol), as per general procedure C. Additional 1,12-dodecanedioyl dichloride (0.021 g, 0.091 mmol) was added. Column chromatography conditions: column 1 (10% methanol /

chloroform), column 2 (1% ammonia / 4% methanol / chloroform), column 3 (10% methanol / chloroform). Yielded **11d** as a yellow foam (0.059 g, 0.063 mmol, 37%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.29 (ddd, *J* = 7.9, 7.3, 1.6 Hz, 2H, H3‴), 7.25 (dd, *J* = 7.7, 1.6 Hz, 2H, H1‴), 7.06 (d, *J* = 2.4 Hz, 2H, H9‴), 7.00 (ddd, *J* = 7.8, 7.4, 1.1 Hz, 2H, H2‴), 6.91 (br t, *J* = 4.9 Hz, 2H, CONH), 6.85 (dd, *J* = 8.0, 0.9 Hz, 2H, H4‴), 6.82 (dd, *J* = 8.3, 2.4 Hz, 2H, H7‴), 6.64 (d, *J* = 8.3 Hz, 2H, H6‴), 5.15 (s, 2H, H5‴), 3.44 (m, 8H, H3″, H5″), 3.36 (td, *J* = 6.0, 5.9 Hz, 4H, H1′), 2.55 (m, 8H, H2″, H6″), 2.52 (t, *J* = 6.3 Hz, 4H, H3′), 2.12 (m, 4H, H2, H11), 1.70 (app p, *J* = 6.3 Hz, 4H, H2′), 1.59 (app p, *J* = 7.4 Hz, 4H, H3, H10), 1.28-1.21 (m, 12H, H4-H9). ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.2 (C_q), 163.1 (C_q), 153.1 (C_q), 141.8 (C_q), 140.8 (C_q), 132.1 (CH), 130.3 (CH), 129.1 (C_q), 126.9 (CH), 123.43 (C_q), 123.39 (CH), 123.1 (CH), 120.32 (CH), 120.25 (CH), 57.8 (CH_2), 53.4 (CH_2), 47.5 (CH_2), 39.6 (CH_2), 37.2 (CH_2), 29.44 (CH_2), 29.40 (CH_2), 29.38 (CH_2), 26.0 (CH_2), 25.3 (CH_2). HPLC: t_R 6.32 min, >99% purity (Method 2). LCMS (*m / z*): 933.2 $[M + H]^+$, 467.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{52}H_{67}Cl_2N_{10}O_2^+$ requires $[M + H]^+$ 933.4820; found 933.4854.

N^{1},N^{14}*-Bis(3-(4-(8-chloro-5*H*-dibenzo[*b,e*][1,4]diazepin-11-yl)piperazin-1-yl)propyl) tetradecanediamide (**11e**)*. 1,14-Tetradecanedioic acid (0.048 g, 0.184 mmol) and oxalyl chloride were reacted, followed by the addition of **7b** (0.131 g, 0.354 mmol), anhydrous potassium carbonate (0.050 g, 0.362 mmol) and pyridine (0.100 mL, 1.24 mmol), as per general procedure C. Additional 1,14-tetradecanedioyl dichloride (0.027 g, 0.093 mmol) was added. Column chromatography conditions: column 1 (gradient elution: from 2% methanol / chloroform to 5% methanol / chloroform, increasing in 1% increments, using a column preconditioned with 1% ammonia / 2% methanol / chloroform), column 2 (10% methanol / chloroform). Yielded **11e** as a yellow foam (0.064 g, 0.066 mmol, 37%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.30 (m, 2H, H3‴), 7.25 (dd, *J* = 7.7, 1.4 Hz, 2H, H1‴), 7.06 (d, *J* = 2.4 Hz, 2H, H9‴), 7.00 (m, 2H, H2‴), 6.88-6.83 (m, 4H, CONH, H4‴), 6.82 (dd, *J* = 8.3, 2.4 Hz,

2H, H7‴), 6.64 (d, *J* = 8.3 Hz, 2H, H6‴), 5.11 (s, 2H, H5‴), 3.45 (m, 8H, H3″, H5″), 3.35 (td, *J* = 5.9, 5.9 Hz, 4H, H1′), 2.55 (m, 8H, H2″, H6″), 2.51 (t, *J* = 6.3 Hz, 4H, H3′), 2.13 (m, 4H, H2, H13), 1.69 (app p, *J* = 6.2 Hz, 4H, H2′), 1.60 (app p, *J* = 7.3 Hz, 4H, H3, H12), 1.32-1.16 (m, 16H, H4-H11). ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.2 (C_q), 163.0 (C_q), 153.1 (C_q), 141.8 (C_q), 140.7 (C_q), 132.1 (CH), 130.3 (CH), 129.1 (C_q), 126.9 (CH), 123.43 (C_q), 123.37 (CH), 123.1 (CH), 120.3 (CH), 120.2 (CH), 57.8 (CH_2), 53.4 (CH_2), 47.6 (CH_2), 39.5 (CH_2), 37.2 (CH_2), 29.6 (CH_2), 29.52 (CH_2), 29.45 (CH_2), 29.4 (CH_2), 26.0 (CH_2), 25.4 (CH_2). HPLC: t_R 6.66 min, >99% purity (Method 2). LCMS (*m / z*): 961.2 $[M + H]^+$, 481.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{54}H_{71}Cl_2N_{10}O_2^+$ requires $[M + H]^+$ 961.5133; found 961.5139.

N^{1},N^{18}*-Bis(3-(4-(8-chloro-5*H*-dibenzo[*b,e*][1,4]diazepin-11-yl)piperazin-1-yl)propyl) octadecanediamide (**11f**)*. 1,18-Octadecanedioic acid (0.057 g, 0.182 mmol) and oxalyl chloride were reacted, followed by the addition of **7b** (0.132 g, 0.357 mmol), anhydrous potassium carbonate (0.050 g, 0.362 mmol) and pyridine (0.100 mL, 1.24 mmol), as per general procedure C. Additional 1,18-octadecanedioyl dichloride (0.026 g, 0.073 mmol) was added. Column chromatography conditions: column 1 (gradient elution: from 2% methanol / chloroform to 5% methanol / chloroform, increasing in 1% increments, using a column preconditioned with 1% ammonia / 2% methanol / chloroform), column 2 (10% methanol / chloroform). Yielded **11f** as a yellow foam (0.074 g, 0.072 mmol, 41%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.29 (ddd, *J* = 7.9, 7.3, 1.6 Hz, 2H, H3‴), 7.25 (dd, *J* = 7.7, 1.5 Hz, 2H, H1‴), 7.06 (d, *J* = 2.4 Hz, 2H, H9‴), 7.00 (ddd, *J* = 7.7, 7.4, 1.1 Hz, 2H, H2‴), 6.86-6.82 (m, 4H, CONH, H4‴), 6.82 (dd, *J* = 8.3, 2.4 Hz, 2H, H7‴), 6.63 (d, *J* = 8.3 Hz, 2H, H6‴), 5.06 (s, 2H, H5‴), 3.45 (m, 8H, H3″, H5″), 3.35 (td, *J* = 5.9, 5.9 Hz, 4H, H1′), 2.55 (m, 8H, H2″, H6″), 2.50 (t, *J* = 6.3 Hz, 4H, H3′), 2.13 (m, 4H, H2, H17), 1.69 (app p, *J* = 6.3 Hz, 4H, H2′), 1.61 (app p, *J* = 7.4 Hz, 4H, H3, H16), 1.34-1.18 (m, 24H, H4-H15). ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.2 (C_q), 162.9 (C_q), 153.0 (C_q), 141.8 (C_q), 140.7 (C_q), 132.1 (CH), 130.3 (CH), 129.1 (C_q), 126.9 (CH), 123.42 (C_q), 123.36 (CH), 123.1 (CH), 120.3

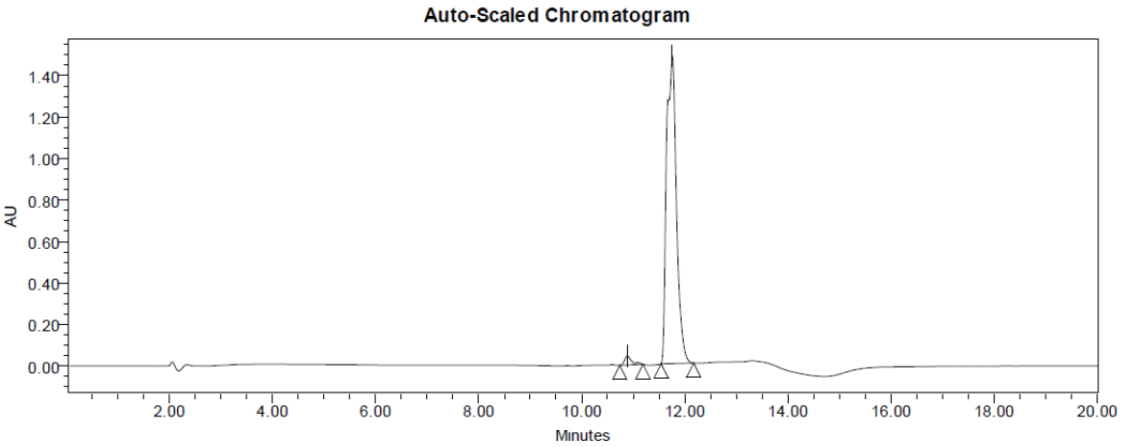
(CH), 120.2 (CH), 57.7 (CH_2), 53.3 (CH_2), 47.5 (CH_2), 39.5 (CH_2), 37.2 (CH_2), 29.72 (2 × CH_2), 29.68 (CH_2), 29.6 (CH_2), 29.52 (CH_2), 29.46 (CH_2), 26.0 (CH_2), 25.4 (CH_2). HPLC: t_R 7.64 min, 99% purity (Method 2). LCMS (*m / z*): 509.2 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{58}H_{79}Cl_2N_{10}O_2^+$ requires $[M + H]^+$ 1017.5759; found 1017.5789.

N^1,N^{20}*-Bis(3-(4-(8-chloro-5*H*-dibenzo[*b,e*][1,4]diazepin-11-yl)piperazin-1-yl)propyl) icosanediamide (**11g**)*. 1,20-Icosanedioic acid (0.062 g, 0.182 mmol) and oxalyl chloride were reacted, followed by the addition of **7b** (0.132 g, 0.357 mmol), *N*,*N*-diisopropylethylamine (78 µL, 0.448 mmol) and pyridine (0.040 mL, 0.0496 mmol), as per general procedure C. Additional 1,20-icosanedioyl dichloride (0.025 g, 0.067 mmol) was added. Column chromatography conditions: gradient elution: from 2% methanol / chloroform to 5% methanol / chloroform, increasing in 1% increments (using a column preconditioned with 1% ammonia / 2% methanol / chloroform). Yielded **11g** as a yellow foam (0.123 g, 0.118 mmol, 66%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.28 (ddd, *J* = 7.9, 7.4, 1.6 Hz, 2H, H3‴), 7.24 (dd, *J* = 7.8, 1.5 Hz, 2H, H1‴), 7.06 (d, *J* = 2.4 Hz, 2H, H9‴), 7.00 (ddd, *J* = 7.7, 7.4, 1.1 Hz, 2H, H2‴), 6.88-6.83 (m, 4H, CONH, H4‴), 6.81 (dd, *J* = 8.3, 2.4 Hz, 2H, H7‴), 6.63 (d, *J* = 8.3 Hz, 2H, H6‴), 5.13 (s, 2H, H5‴), 3.45 (m, 8H, H3″, H5″), 3.35 (td, *J* = 5.9, 5.9 Hz, 4H, H1′), 2.54 (m, 8H, H2″, H6″), 2.49 (t, *J* = 6.4 Hz, 4H, H3′), 2.13 (m, 4H, H2, H19), 1.69 (app p, *J* = 6.3 Hz, 4H, H2′), 1.61 (app p, *J* = 7.3 Hz, 4H, H3, H18), 1.34-1.17 (m, 28H, H4-H17). ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.2 (C_q), 162.9 (C_q), 153.0 (C_q), 141.8 (C_q), 140.7 (C_q), 132.1 (CH), 130.3 (CH), 129.0 (C_q), 126.8 (CH), 123.4 (C_q), 123.3 (CH), 123.1 (CH), 120.3 (CH), 120.2 (CH), 57.6 (CH_2), 53.3 (CH_2), 47.5 (CH_2), 39.4 (CH_2), 37.1 (CH_2), 29.72 (CH_2), 29.70 (2 × CH_2), 29.66 (CH_2), 29.6 (CH_2), 29.5 (CH_2), 29.4 (CH_2), 26.0 (CH_2), 25.4 (CH_2). HPLC: t_R 8.06 min, >99% purity (Method 2). LCMS (*m / z*): 523.4 $[M + 2H]^{2+}$. HRMS (*m / z*): $C_{60}H_{83}Cl_2N_{10}O_2^+$ requires $[M + H]^+$ 1045.6072; found 1045.6104.

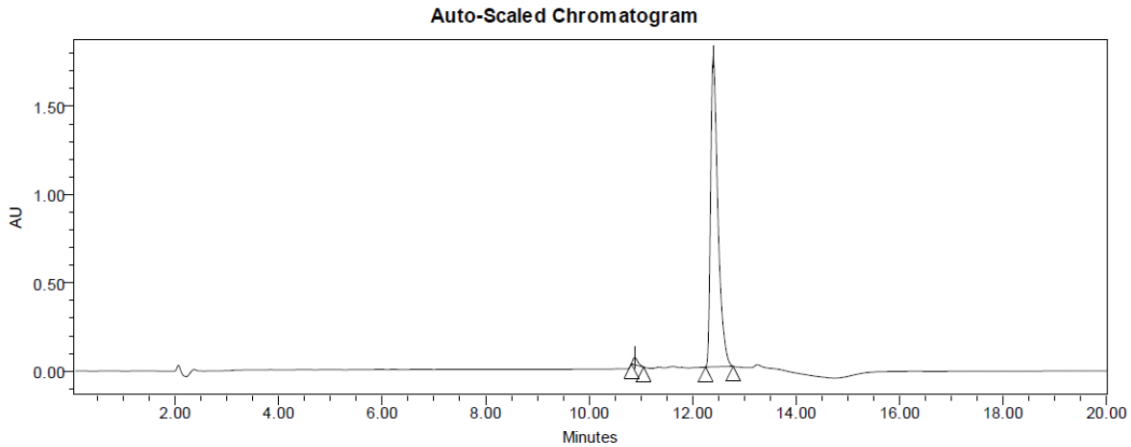
HPLC traces

HPLC traces for clozapine N5 bivalent ligands

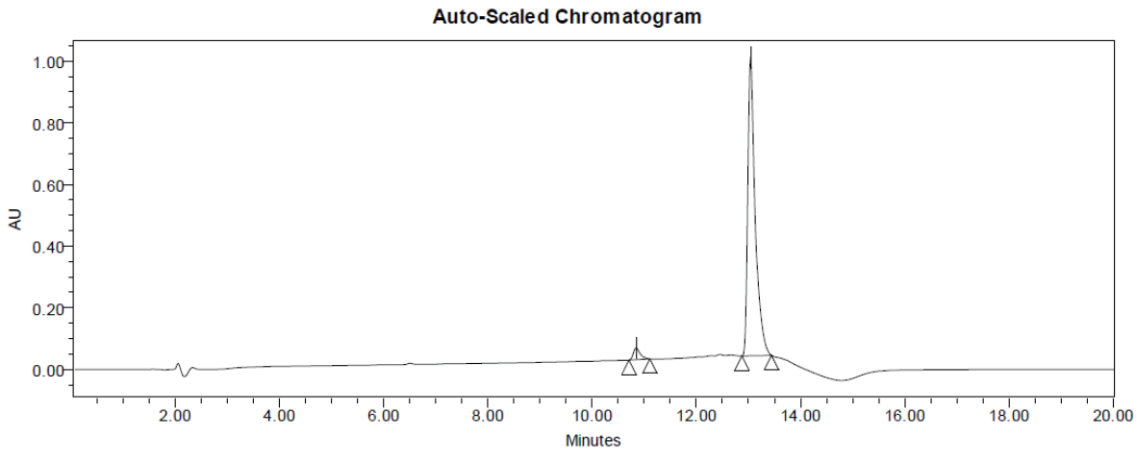
Compound **9a**



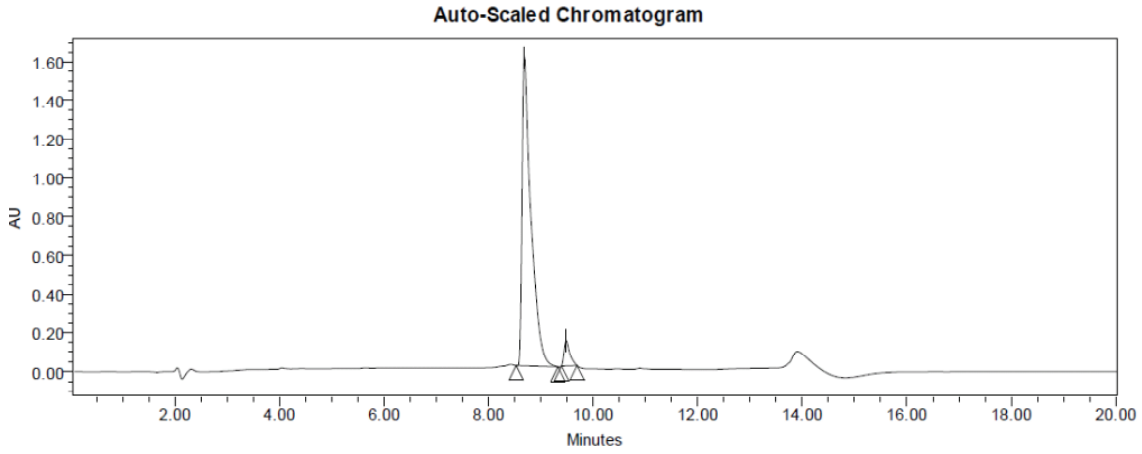
Compound **9b**



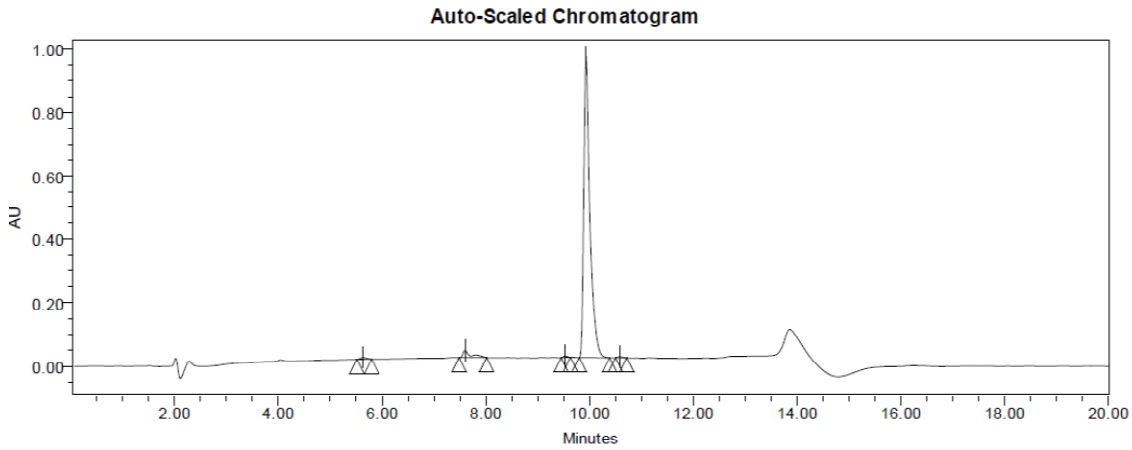
Compound **9c**



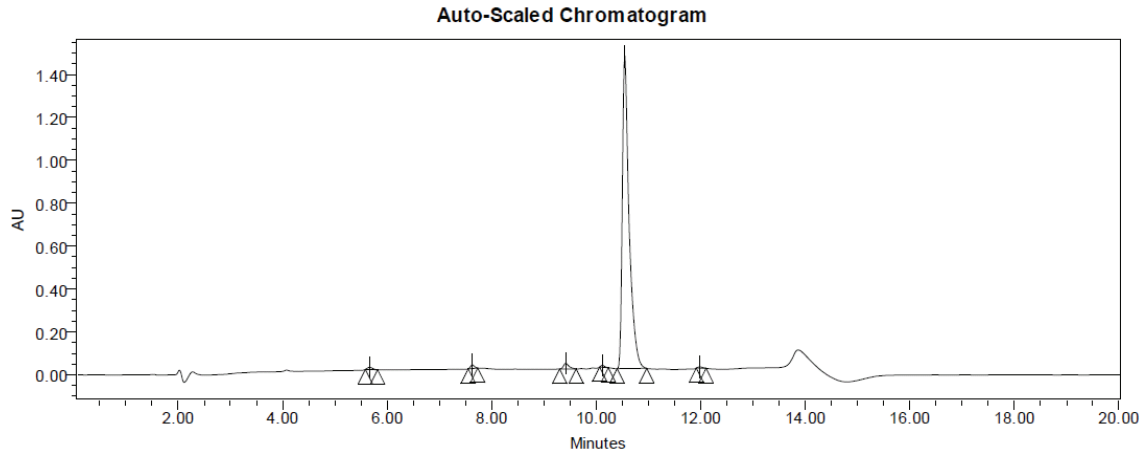
Compound **9d**



Compound **9e**

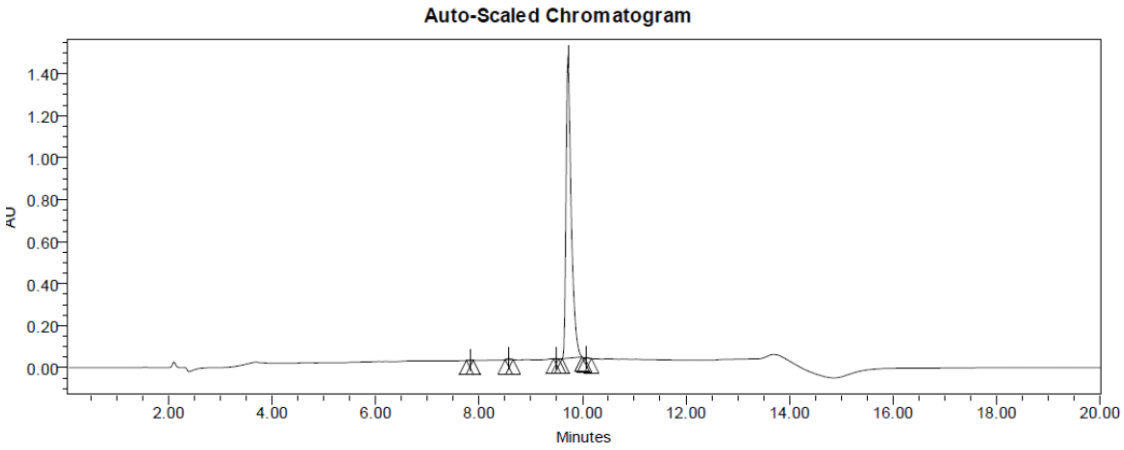


Compound **9f**

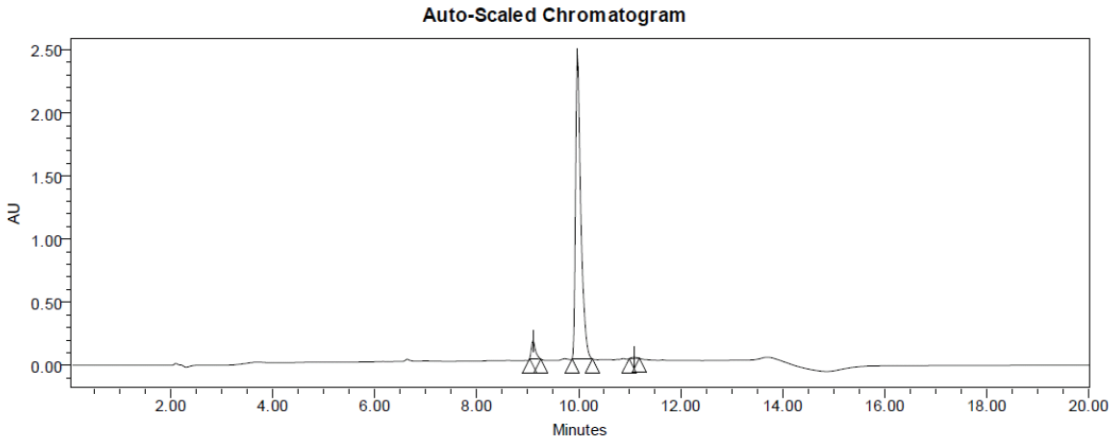


HPLC traces for clozapine hydrazide bivalent ligands

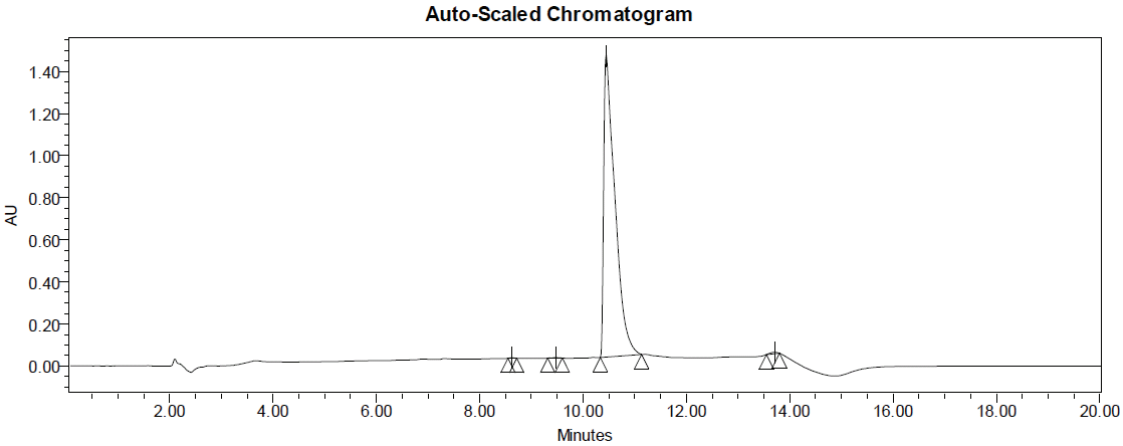
Compound **10a**



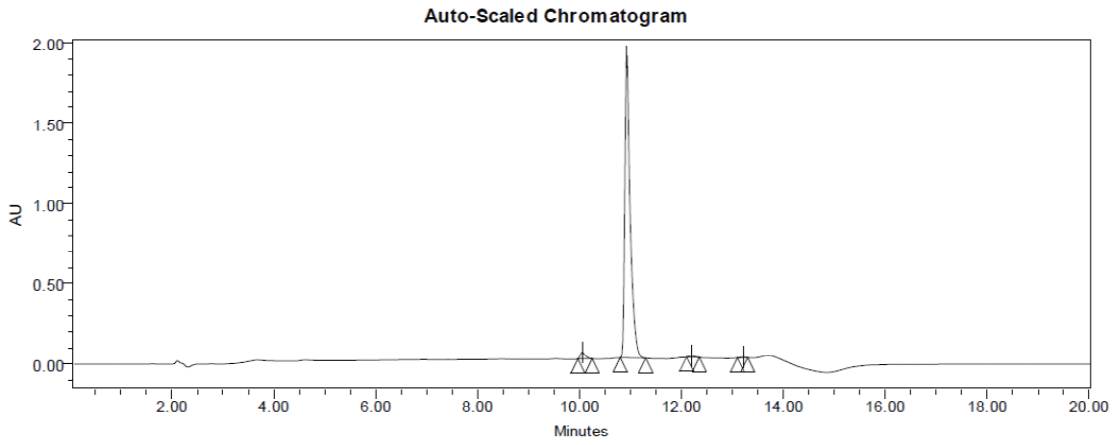
Compound **10b**



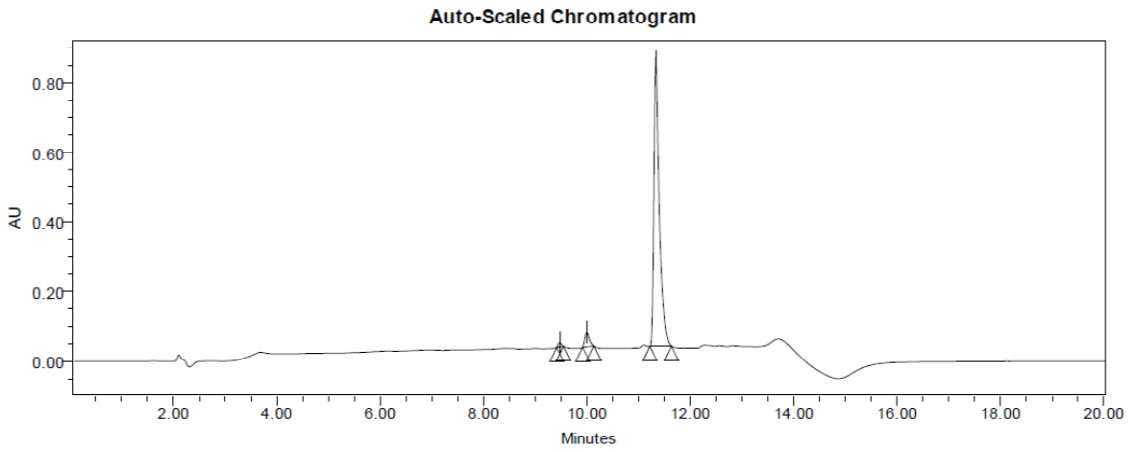
Compound **10c**



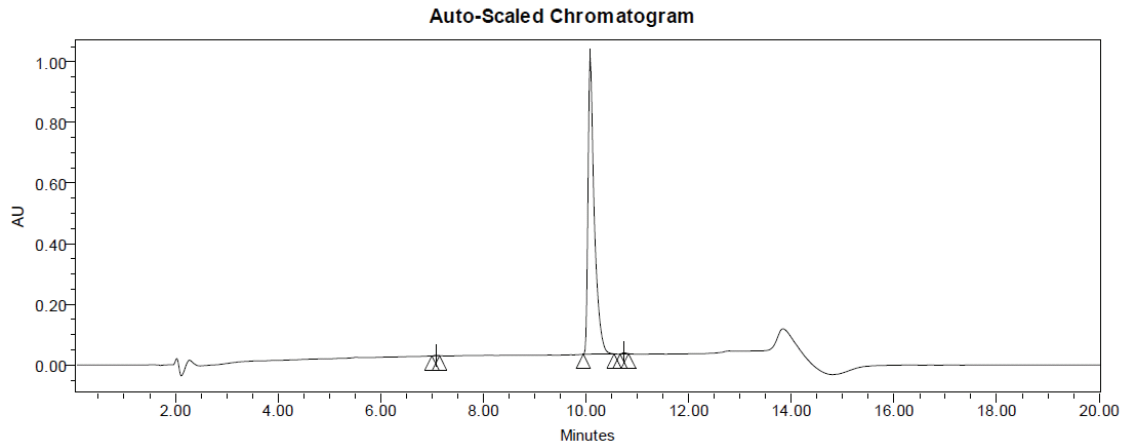
Compound **10d**



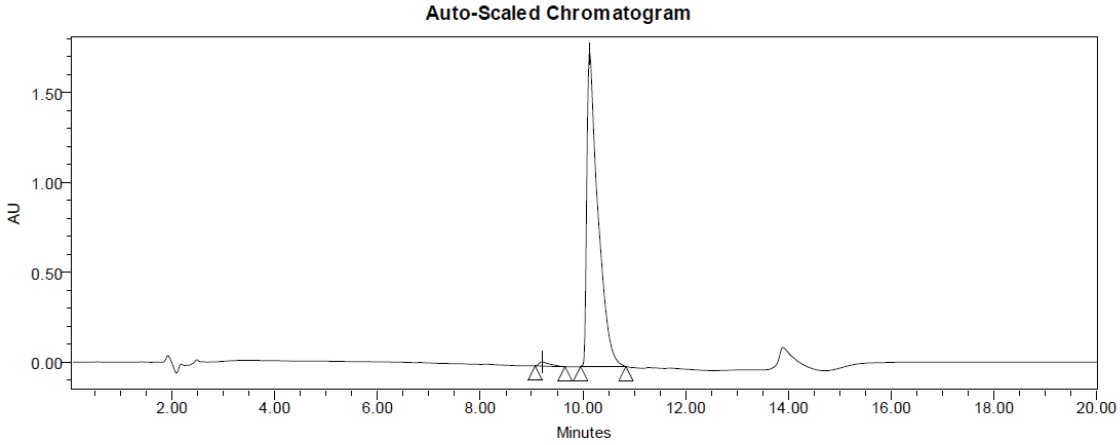
Compound **10e**



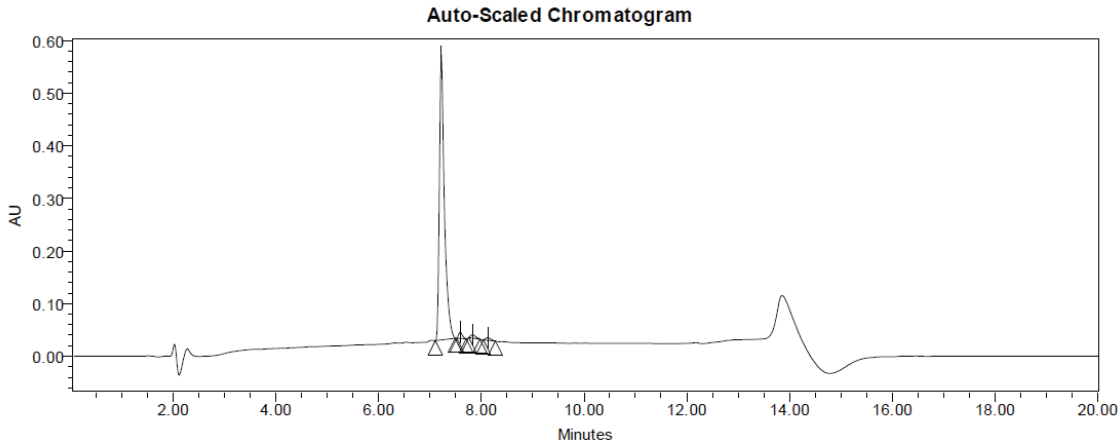
Compound **10f**



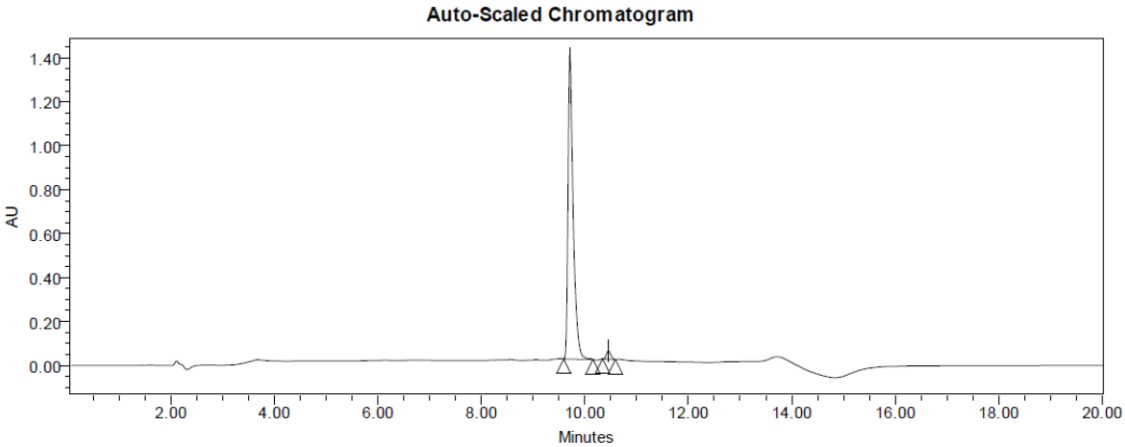
Compound **10g**



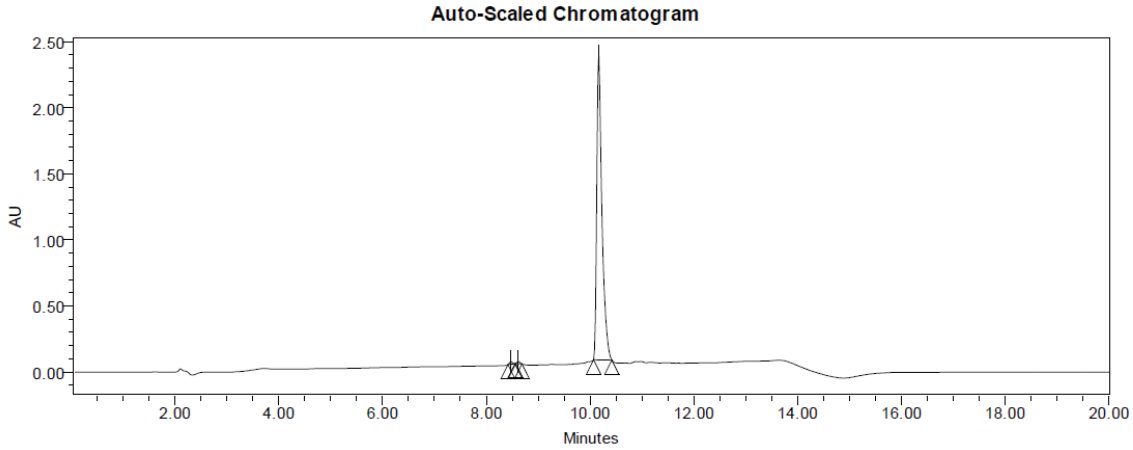
Compound **14a**



Compound **14b**

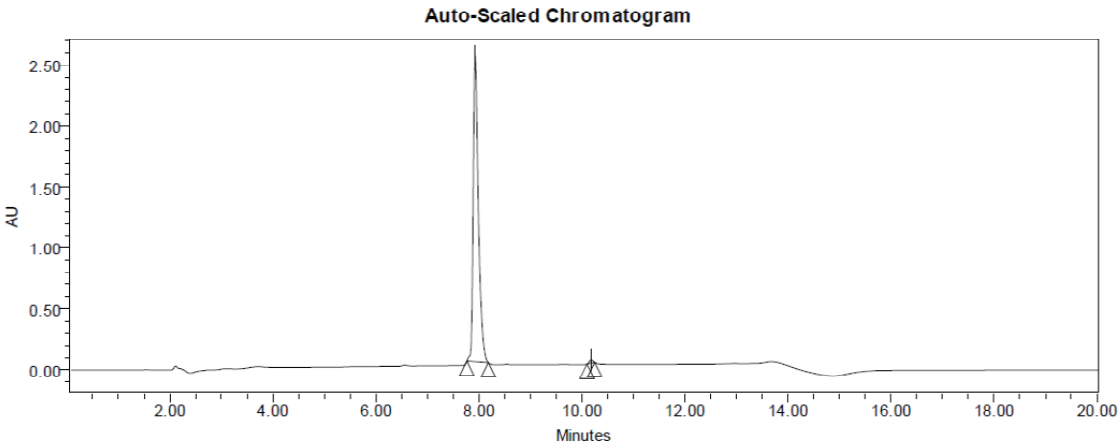


Compound **16**

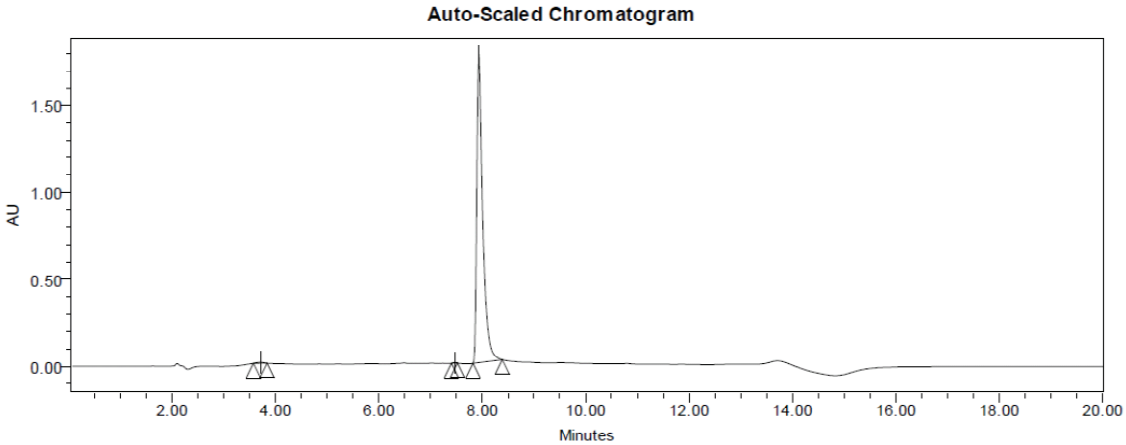


HPLC traces for clozapine hydrazide bivalent ligands

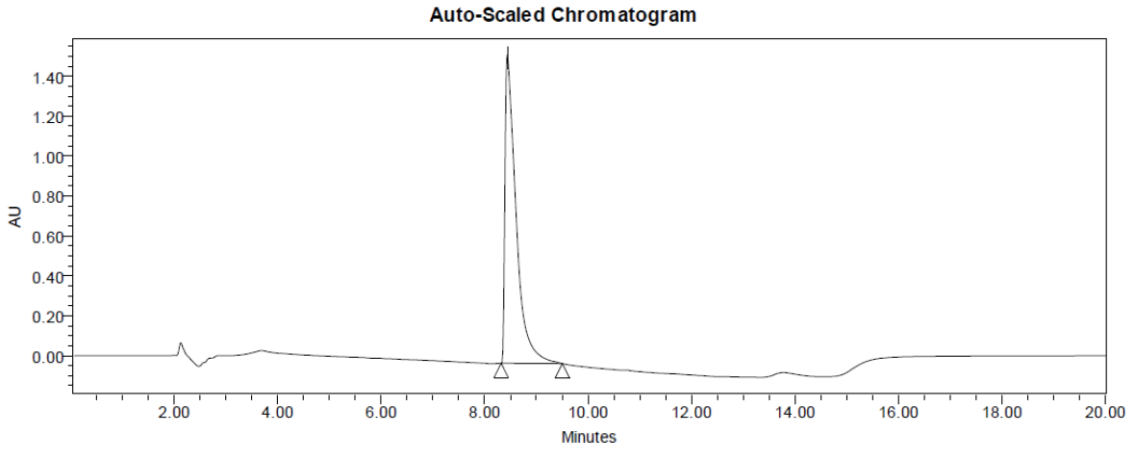
Compound **11a**



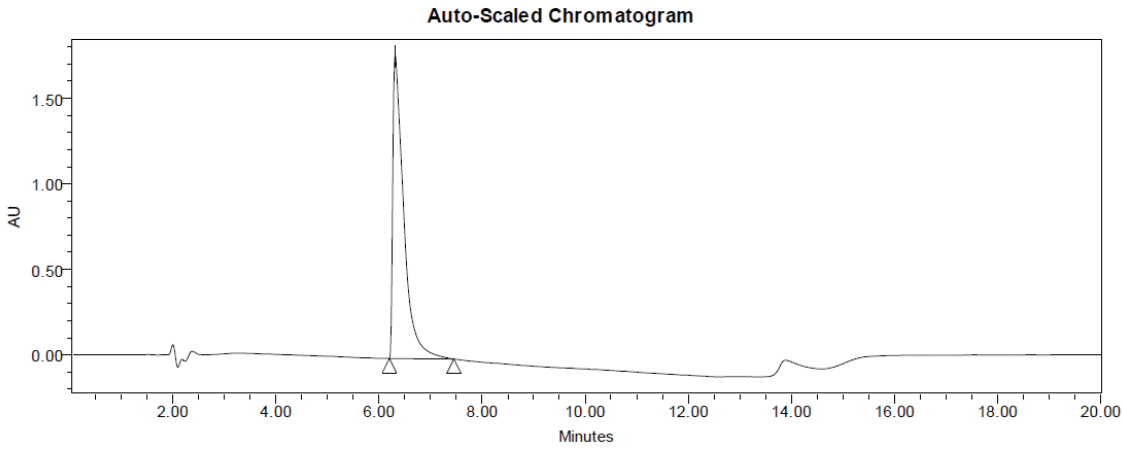
Compound **11b**



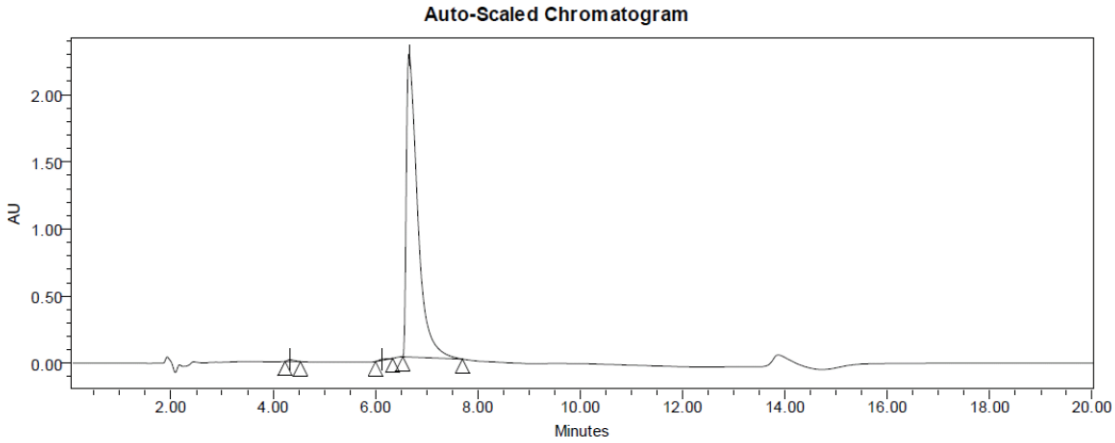
Compound **11c**



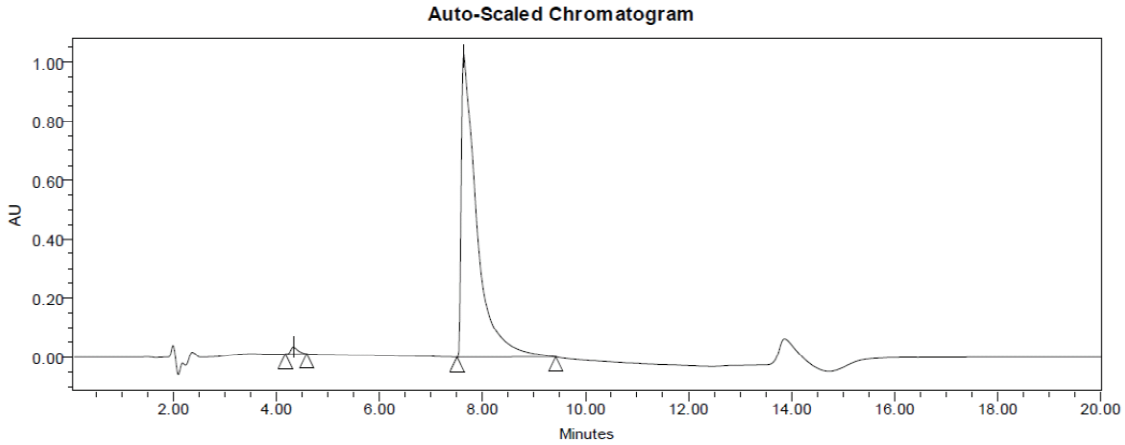
Compound **11d**



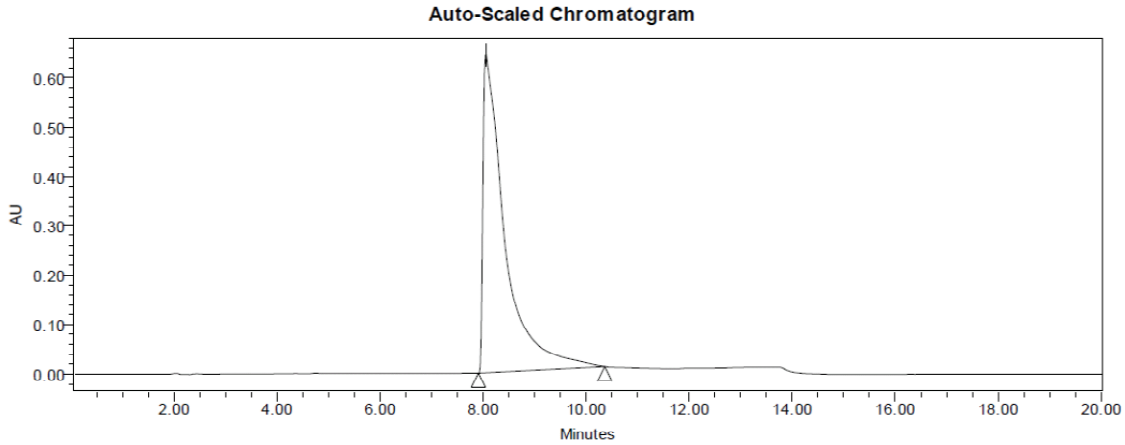
Compound **11e**



Compound **11f**

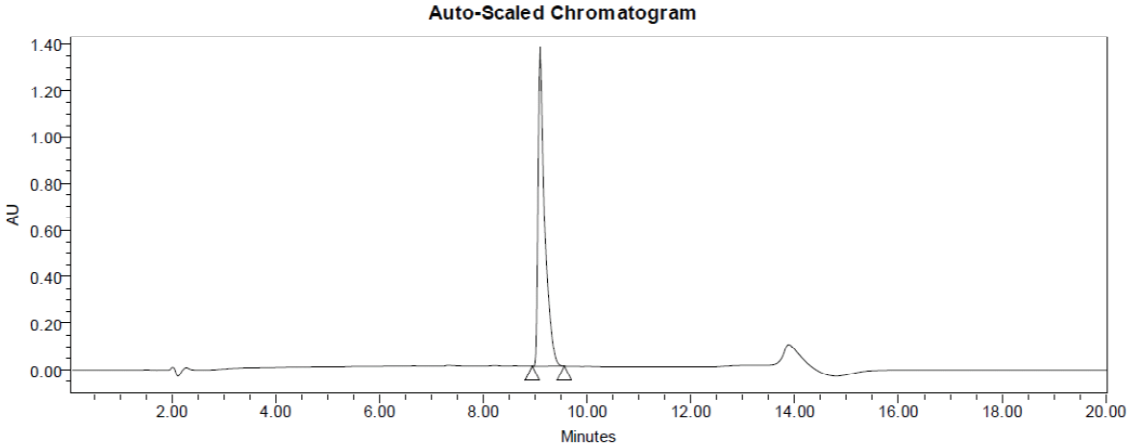


Compound **11g**

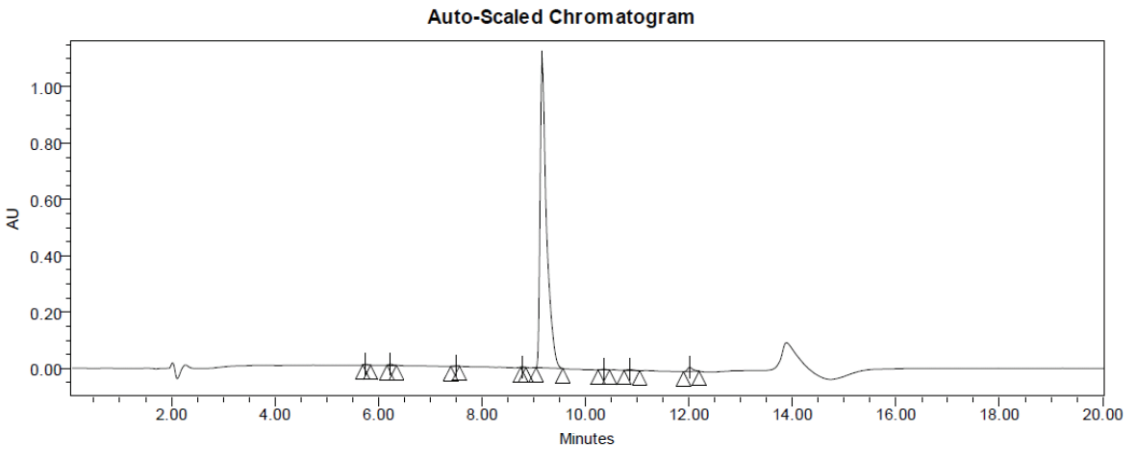


HPLC traces for monovalent ligands

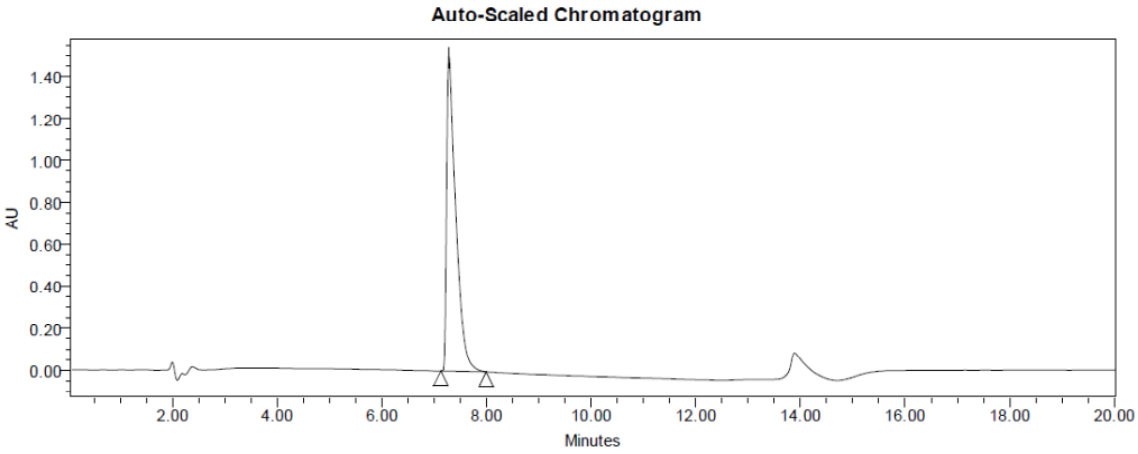
Compound **17**



Compound **18**



Compound **19**



References

1. Still, W. C.; Kahn, M.; Mitra, A. Rapid chromatographic technique for preparative separations with moderate resolution. *J. Org. Chem.* **1978,** *43*, 2923-2925.
2. Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. NMR chemical shifts of common laboratory solvents as trace impurities. *J. Org. Chem.* **1997,** *62*, 7512-7515.
3. Capuano, B.; Crosby, I. T.; Lloyd, E. J.; Taylor, D. A. Synthesis and preliminary pharmacological evaluation of 4´-arylmethyl analogues of clozapine. I. The effect of aromatic substituents. *Aust. J. Chem.* **2002,** *55*, 565-576.
4. Capuano, B.; Crosby, I. T.; Lloyd, E. J.; Neve, J. E.; Taylor, D. A. Aminimides as potential CNS acting agents. I. Design, synthesis, and receptor binding of 4'-aryl aminimide analogues of clozapine as prospective novel antipsychotics. *Aust. J. Chem.* **2007,** *60*, 673-684.
5. Su, J.; Tang, H.; McKittrick, B. A.; Burnett, D. A.; Zhang, H.; Smith-Torhan, A.; Fawzi, A.; Lachowicz, J. Modification of the clozapine structure by parallel synthesis. *Bioorg. Med. Chem. Lett.* **2006,** *16*, 4548-4553.
6. Olofson, R. A.; Martz, J. T.; Senet, J. P.; Piteau, M.; Malfroot, T. A new reagent for the selective, high-yield *N*-dealkylation of tertiary amines: improved syntheses of naltrexone and nalbuphine. *J. Org. Chem.* **1984,** *49*, 2081-2082.
7. Zlatev, I.; Giraut, A.; Morvan, F.; Herdewijn, P.; Vasseur, J.-J. d-Di-carboxybutyl phosphoramidate of 2'-deoxycytidine-5'-monophosphate as substrate for DNA polymerization by HIV-1 reverse transcriptase. *Bioorg. Med. Chem.* **2009,** *17*, 7008-7014.
8. Cason, J.; Reist, E. J. Reactions of glutaryl dichloride with organometallic reagents. *J. Org. Chem.* **1958,** *23*, 1675-1679.
9. Cram, D. J.; Daeniker, H. U. Macro rings. V. Transannular effects in the 1,4-decamethylenebenzene series. *J. Am. Chem. Soc.* **1954,** *76*, 2743-2752.
10. Liu, C.; Hudson, R. H. E.; Petersen, N. O. Convergent and sequential synthesis of dendritic, multivalent complexing agents. *Synthesis* **2002**, 1398-1406.
11. Asay, R. E.; Bradshaw, J. S.; Nielsen, S. F.; Thompson, M. D.; Snow, J. W.; Masihdas, D. R. K.; Izatt, R. M.; Christensen, J. J. The synthesis of novel macrocyclic multidentate compounds for dioxodioic acids. *J. Heterocycl. Chem.* **1977,** *14*, 85-90.

List of publications

McRobb, F. M.; Capuano, B.; Crosby, I. T.; Chalmers, D. K.; Yuriev, E. Homology modeling and docking evaluation of aminergic G protein-coupled receptors. *J. Chem. Inf. Model.* **2010,** *50*, 626-637.

Capuano, B.; Crosby, I. T.; McRobb, F. M.; Podloucka, A.; Taylor, D. A.; Vom, A.; Yuriev, E. The synthesis and preliminary pharmacological evaluation of a series of substituted 4′-phenoxypropyl analogues of the atypical antipsychotic clozapine. *Aust. J. Chem.* **2010,** *63*, 116-124.

Capuano, B.; Crosby, I.; Forsyth, C.; McRobb, F.; Moudretski, V.; Taylor, D.; Vom, A.; Yuriev, E. New hybrids of clozapine and haloperidol and their isosteric analogues: Synthesis, X-ray crystallography, conformational analysis and preliminary pharmacological evaluation. *Struct. Chem.* **2010,** *21*, 613-628.

Capuano, B.; Crosby, I. T.; McRobb, F. M.; Taylor, D. A.; Vom, A.; Blessing, W. W. JL13 has clozapine-like actions on thermoregulatory cutaneous blood flow in rats: Involvement of serotonin 5-HT_{1A} and 5-HT_{2A} receptor mechanisms. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* **2010,** *34*, 136-142.

McRobb, F. M.; McLean, K. C.; Agostino, M.; Crosby, I. T.; Capuano, B.; Yuriev, E.; Chalmers, D. K. Predicting the structure of the dopamine D_3 receptor: An evaluation of virtual screening approaches to GPCR modeling. **2011**, *prepared manuscript*.

McRobb, F. M.; Crosby, I. T.; Yuriev, E.; Christopoulos, A.; Lane, J. R.; Capuano, B. Homobivalent ligands of the atypical antipsychotic clozapine: Design, synthesis and pharmacological evaluation. **2011**, *prepared manuscript*.

List of conference presentations

McRobb, F. M.; Capuano, B.; Chalmers, D. K.; Crosby, I. T.; Roberts, B. P.; Yuriev, E. Homology modelling using the β_2-adrenergic receptor (Poster). 3D at the Cove, Gold Coast (2008).

McRobb, F. M.; Capuano, B.; Chalmers, D. K.; Crosby, I. T.; Roberts, B. P.; Shonberg, J.; Yuriev, E. Homology modelling of dopamine receptors (Poster). AMMA Young Modellers' Forum, Sydney (2008).

McRobb, F. M.; Capuano, B.; Chalmers, D. K.; Crosby, I. T.; Roberts, B. P.; Shonberg, J.; Yuriev, E. Homology modelling of the dopamine 2 receptor (Poster). 3rd Annual Postgraduate Symposium, Melbourne (2008).

McRobb, F. M.; Capuano, B.; Chalmers, D. K.; Crosby, I. T.; Yuriev, E. MIPS GPCR library: A library of validated homology models of pharmaceutically important GPCRs (Oral). AIMECS09, Cairns (2009).

McRobb, F. M.; Capuano, B.; Chalmers, D. K.; Crosby, I. T.; Yuriev, E. MIPS GPCR library of validated homology models (Poster). 4th Annual Postgraduate Symposium, Melbourne (2009).

McRobb, F. M.; Capuano, B.; Chalmers, D. K.; Crosby, I. T.; Yuriev, E. Homology modelling and docking evaluation of aminergic G protein-coupled receptors (Oral). Melbourne Meeting of Molecular Modellers, Melbourne (2010).

McRobb, F. M.; Capuano, B.; Chalmers, D. K.; Crosby, I. T.; Yuriev, E. Aminergic G protein-coupled receptors: Homology modeling and evaluation by virtual screening (Poster). EFMC-ISMC, Brussels (2010).

McRobb, F. M.; Capuano, B.; Chalmers, D. K.; Crosby, I. T.; Yuriev, E. Aminergic G protein-coupled receptors: Homology modeling and evaluation by virtual screening (Poster). 5th Annual Postgraduate Symposium, Melbourne (2010).

McRobb, F. M.; McLean, K. C.; Capuano, B.; Chalmers, D. K.; Crosby, I. T.; Yuriev, E. GPCR DOCK 2010 - predicting the structure of the D_3 dopamine receptor (Oral). 5th Annual Postgraduate Symposium, Melbourne (2010).

McRobb, F. M.; McLean, K. C.; Capuano, B.; Chalmers, D. K.; Crosby, I. T.; Yuriev, E. Predicting the structure of the dopamine D_3 receptor – our experience with GPCR DOCK 2010 (Oral). MM2010, Melbourne (2010).