DOCTOR OF PHILOSOPHY

Characterisation of protein-ligand interactions of glycine binding protein (GBP) as a model of the heteromeric glycine receptor

Jones, Mathew

Award date:
2021

A thesis submitted for the degree of Doctor of Philosophy

Characterising protein-ligand interactions of the glycine binding protein (GBP) as a model of the heteromeric glycine receptor

Mathew Jones

Supervisor:
Professor William N. Hunter

School of Life Sciences

University of Dundee

May 2021

Table of contents

Table of contents ..... **I**
List of figures ..... **III**
List of tables ..... **V**
Abbreviations ..... **VI**
Acknowledgements ..... **X**
Declaration ..... **XI**
Summary ..... **XII**
1. Introduction **1**
1.1 Pentameric ligand gated ion channels 1
1.2 Biophysical techniques for ligand screening 7
1.3 Complementary biophysical methods for the characterisation of protein-ligand interactions 10
1.4 DMSO, a good solvent or a complicating factor? 12
1.5 Computational ligand screening 12
1.6 Aims 13
2. Materials & methods **15**
2.1 Construct design 15
2.2 Cloning 16
2.3 Recombinant protein production 20
2.4 Biophysical techniques 26
2.5 Computational tools 33
2.6 Confocal microscopy 36
3. Assessment of three biophysical techniques used for screening potential ligands for pLGICs **37**
3.1 Introduction 37
3.2 Data from three biophysical assays 38
3.3 Comparison of ligand affinity measured by different techniques 44
3.4 Thermodynamic parameters utilised to understand binding 47
3.5 Widening the scope to other ligands of *Ac*AChBP 51
3.6 Comparison with published data 52
3.7 Consideration of alternative techniques to study *Ac*AChBP-ligand interactions 54
3.8 Summary 55
4. Characterisation of strychnine and N-methylbicuculline interacting with *Ac*AChBP and GBP **56**
4.1 Introduction 56
4.2 Deducing the binding affinity of strychnine with *Ac*AChBP and GBP using ITC and WF 56
4.3 Comparison of *Ac*AChBP: and GBP:strychnine crystal structures 59
4.4 Are interactions of *Ac*AChBP: and GBP:strychnine representative of the interactions in the physiologically relevant receptors, GlyR and nAChR? 60

4.5 Deducing the binding affinity of N-methylbicuculline with *Ac*AChBP and GBP by biophysical techniques 62

4.6 Crystal structure of GBP:N-methylbicuculline 64

4.7 Comparison of the GBP:N-methylbicuculline complex with the inhibitory pLGICs, GlyR and $GABA_AR$ 66

4.8 Low resolution crystal structure of a *Ac*AChBP:N-methylbicuculline complex 70

4.9 Summary 71

5. A screening campaign to identify new ligands for GlyR and nAChR, utilising the GBP and *Ac*AChBP surrogates 72

5.1 Introduction 72

5.2 NMR screening using an in-house compound library 73

5.3 Substructure analysis of the NMR hits seeking to determine similarities 74

5.4 Attempted measurement of compound binding and calculation of affinity 78

5.5 Computational docking of small molecules 78

5.6 Summary 105

6. Design of and attempted production of selected proteins 106

6.1 Construct design 106

6.2 Attempted protein production in baculovirus 110

6.3 Localisation at the membrane 110

6.4 Attempted protein production in lentivirus 114

6.5 Design of a HisR orthosteric binding site model 114

6.3 Summary 117

7. Conclusions and future work 118

8. References 121

A. Crystallographic statistics for structural complexes 137

B. Theoretical interaction parameters predicted by Molsoft 138

C. Preliminary work on the amino acid transporters, GlyT1C and PutP 140

C.1 Introduction 140

C.2 Materials and methods 140

C.3 Attempted production of GlyT 142

C.4 Cloning and attempted production of PutP 143

C.5 Summary 143

C.6 References 144

D. Publications 146

List of figures

1.1 Binding loops of the orthosteric site of GBP 3

1.2 Key orthosteric binding loops of GlyR structures superimposed 4

1.3 Subunit organisation and receptor function 5

1.4 Absorption and emission spectra of aromatic amino acids 11

2.1 Cloning strategy for the construction of GFP-tagged HisR 19

2.2 Examples of data used to inform protein purification 22

2.3 Effect of DMSO concentration (%) on GBP stability 31

2.4 Example setup for BLI 32

2.5 Chemical groups extensions screen 35

3.1 Models of both enantiomers of three ligands of interest bound to *Ac*AChBP 40

3.2 Ligand binding to *Ac*AChBP monitored by WF 41

3.3 ITC data for *Ac*AChBP binding various ligands 42

3.4 SPR data for *Ac*AChBP binding selected ligands 43

3.5 Optimal ITC binding curve for an enantiomeric mixture 45

3.6 The effect of c in an ITC experiment 46

3.7 (-)-Hosieine binding to *Ac*AChBP data monitored by ITC and WF 52

3.8 Schematic diagrams of ligand interactions with *Ac*AChBP and GBP 53

4.1 ITC data for strychnine binding GBP and *Ac*AChBP 57

4.2 Strychnine binding monitored by WF 58

4.3 Strychnine bound in the orthosteric site of GBP and *Ac*AChBP 60

4.4 Sequence alignment of the important binding loops 61

4.5 Differences between strychnine bound to GBP and GlyR 62

4.6 ITC data for N-methylbicuculline binding GBP and *Ac*AChBP 64

4.7 N-methylbicuculline binding monitored by WF 65

4.8 Representation of N-methylbicuculline bound to GBP 67

4.9 Differences between N-methylbicuculline bound to GBP and $GABA_AR$ 68

4.10 Superpositions of N-methylbicuculline and bicuculline 69

4.11 Structure of the $GABA_CR$ antagonist TPMPA 70

4.12 N-methylbicuculline binds differently in the *Ac*AChBP orthosteric binding site 71

List of figures

5.1 A comparison between four example ^{19}F spectra 75

5.2 NMR small molecule screen hits clustered by structural similarities 76

5.3 Binding of small molecule 1d monitored by WF 79

5.4 Docking controls 80

5.5 Images of nicotine bound in *Ac*AChBP-*apo* or *Ac*AChBP-nicotine 82

5.6 Selected hits from the NMR screen docked into the relevant target 85

5.7 Analysis of the top 50 Maybridge screening hits for GBP 87

5.8 Top ten GBP hits from the Maybridge screen docked to GBP 88

5.9 Analysis of the top 50 Maybridge screening hits for *Ac*AChBP:nicotine 91

5.10 Top ten *Ac*AChBP hits from the Maybridge screen docked to *Ac*AChBP:nicotine 92

5.11 Analysis of the top 50 Maybridge screening hits for *Ac*AChBP:*apo* 94

5.12 Top ten *Ac*AChBP hits from the Maybridge screen docked to *Ac*AChBP:*apo* 95

5.13 Binding of small molecule 3b to GBP monitored by WF 96

5.14 Sequence alignment of human nAChR α_7, *Ac*AChBP:nicotine and *Ac*AChBP:*apo* 99

5.15 Three ligands of nAChR, used as a comparison to compound hits 99

5.16 Sequence alignment of human GlyR α_1 and β with GBP 101

5.17 Three ligands of GlyR used as a comparison to small molecule hits 101

5.18 Extension of the top-ranking GBP small molecule, 3b 103

5.19 Extension of the top-ranking *Ac*AChBP:nicotine small molecule, 4b 104

5.20 Extension of the top-ranking *Ac*AChBP:*apo* small molecule, 5c 104

6.1 Sequence alignment of the main loops in the orthosteric binding site of pLGICs 106

6.2 GlyR $\beta^+\alpha^-$ interface model showing the loop C substitutions of the principal subunit 107

6.3 Sequence alignment of HisR1 (HisCl1) and HisR2 (HisCl2) 109

6.4 Example of an SDS-PAGE gel and Western blot for GlyR $\beta^+\alpha^-$ and HisR 111

6.5 Microscopy images of C-terminal GFP-tagged constructs 112

6.6 Microscopy images of N-terminal GFP-tagged constructs 113

6.7 Information used to guide the design of a histamine-binding protein (HisBP) 115

List of tables

1.1 Published structures of members of the pLGIC family 2

2.1 PCR solution set up 17

2.2 Methods for running a PCR 17

2.3 Protein production vectors 20

2.4 Buffers 26

2.5 WF experimental conditions 28

2.6 ITC experimental conditions 29

2.7 SPR experimental conditions 31

2.8 ICM parameters and descriptions 34

3.1 Binding data for the interaction of *Ac*AChBP with three ligands 39

3.2 Calculated parameters for computational docking of three *Ac*AChBP ligands 50

3.3 Binding data for the interaction of (-)-hosieine, strychnine and N-methylbicuculline 51

4.1 Binding data for the interaction of strychnine with *Ac*AChBP and GBP 57

4.2 Binding data for the interaction of N-methylbicuculline with *Ac*AChBP and GBP 63

5.1 Seven NMR compounds chosen to characterise further 77

6.1 Percentage sequence identity for HisR1 and HisR2 from different species 110

Abbreviations

The International System of Units (SI) and International Union of Pure and Applied Chemistry (IUPAC) abbreviations are used throughout. Amino acids are abbreviated using the standard one or three letter code and atoms are abbreviated to their elemental symbol.

5-HT_3R	Serotonin receptor
Ac	*Aplysia californica*
AChBP	Acetylcholine binding protein
AGF	Analytical gel filtration
AI	Autoinduction
Am	*Apis mellifera*
APF	Atomic property field
β-gal	β-galactosidase
BLI	Biolayer interferometry
BME	2-mercarptoethanol
CCP4	Collaborative computational project number 4
CNS	Central nervous system
Cryo-EM	Cryogenic electron microscopy
D_2O	Deuterium oxide
DDM	Dodecyl-β-D-maltopyranoside
Dm	*Drosophila melanogaster*
DMSO	Dimethyl sulfoxide
DSF	Differential scanning fluorimetry
Dsolv	Desolvation
EC_{50}	Effective concentration at half maximum
ECD	Extracellular domain
EDTA	Ethylenediaminetetraacetic acid
ECL	Enhanced chemiluminescence
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride

Abbreviations

ELIC	*Erwinia* ligand gated ion channel
ESKAPE	*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumanni*, *Pseudomonas aeruginosa* and *Enterobact*
FBS	Fetal bovine serum
$GABA_AR$	γ-Aminobutyric acid type-A receptor
$GABA_cR$	γ-Aminobutyric acid type-C receptor
GBP	Glycine binding protein
GDP	Guanosine-5'-diphosphate
GFP	Green fluorescent protein
GLIC	*Gloeobacter* ligand gated ion channel
GluCl	Glutamate receptor
GlyR	Glycine receptor
GlyT	Glycine transporter
GPCR	G-protein coupled receptor
GTP	Guanosine-5'-triphosphate
HBA	Hydrogen bond acceptor
HBD	Hydrogen bond donor
Hbond	Hydrogen bond
HEK293	Human embryonic kidney 293 cells
Hi5	*Trichoplusia ni*
HisBP	Histamine binding protein
HisR	Ionotropic histamine receptor
Hphob	Hydrophobic interaction
HPLC	High-performance liquid chromatography
Hs	*Homo sapiens*
IC_{50}	Inhibition concentration at half maximum
ICD	Intracellular domain
ICM	Internal coordinate modelling
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside

Abbreviations

ITC	Isothermal titration calorimetry
K_A	Association constant
K_D	Dissociation equilibrium constant
K_i	Inhibition constant
LB	Lysogeny broth
Ls	*Lymnaea stagnalis*
MC	Main chain
MW	Molecular weight
N	Stoichiometry
nAChR	Nicotinic acetylcholine receptor
NCS	Non-crystallographic symmetry
NHS	N-hydroxysuccinimide
NMR	Nuclear magnetic resonance spectroscopy
Nv	*Nasonia vitripennis*
Pa	*Pseudomonas aeruginosa*
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein data bank
PEI	Polyethylenimine
pEnvelope	Lentiviral envelope proteins plasmid
pLGIC	Pentameric ligand gated ion channel
pPackaging	Lentiviral packaging proteins plasmid
pTransfer	Lentiviral transfer plasmid
PutP	Proline transporter
PVDF	Polyvinylidene difluoride
QCIP	Quadruple resonance cryo inverse probe
QSAR	Quantitative structure-activity relationship
RMSD	Root-mean-square deviation
RotB	Rotatable bond
RSCC	Real space correlation coefficient

Abbreviations

RT	Room temperature
SASA	Solvent accessible surface area
SC	Side chain
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SE	Standard error
*Sf*9	*Spodoptera frugiperda*
SOC	Super optimal broth with catabolite repression
SPR	Surface plasmon resonance
SSM	Secondary structure method superimposition
^{t}Bu	Tert-butyl
Tc	*Tribolium castaneum*
TEV	Tobacco etch virus
TMD	Transmembrane domain
TPMPA	(1,2,5,6-Tetrahydropyridin-4-yl)methylphosphinic acid
UPGMA	Unweighted pair group method with arithmetic mean algorithm
VDW	Van der Waals interaction
VLS	Virtual ligand screening
waterLOGSY	Water ligand observed via gradient spectroscopy
WF	Tryptophan fluorescence quenching assay

Acknowledgements

I would like to thank Prof Bill Hunter for being such a good teacher and supervisor. I am glad I chose his project in 2017, thanks to it I have developed as a person and a scientist. Bill was there every step of the way, ready to give advice, both professionally and personally. I hope, by now, he now knows my name. I would also like to thank him for the supply of honey over the years and for being a good sport when either Wales or England beat Scotland in the rugby.

I would like to personally thank Dr Alice Dawson and Dr Thomas Eadsforth for their help and support over the years. They taught me everything I know and put up with my endless questions, some of which were very stupid, but they never told me so. They also offered advice along the way, primarily about what I should be drinking. They are definitely friends, not colleagues.

A big thanks goes out to the rest of the team, Alison Scott, Andy Shepherd, Avinash Punekar, Genady Pankov, Greg Stewart, Joel McMillan, Paul Fyfe, Pierre Petit, Samuel Davis and Sharon Shepherd. Even though some of us were from separate labs they would always offer to help where ever they could. They all made being in the lab an enjoyable experience with a ready supply of humour (primarily innuendo Fridays), and this is probably why we became good friends.

I would also like to thank all the technical specialists that have helped me along the way including, Prof Paul Davies, Dr Dan Fletcher, Dr Alan Prescott, Dr Andrew Orry and Sarah Niven. They selflessly gave up their time to teach me different techniques and offered help whenever I got stuck.

Thanks to my family and friends who have also supported me throughout my time in Dundee, through the highs and lows. Moral support was always provided through, chats, game nights, walks, dinners and lots of alcohol. It was all very much appreciated and will not be forgotten.

I would finally like to thank you for reading this thesis or at least giving it a go. If I was you, I would skip straight to the conclusions (Chapter 7) or just look at the pictures.

I will finish by telling you (and remind myself of) two things that I learnt during my PhD and have become principles I now live by. Firstly, if you do not ask, you do not get, this covers everything from getting help with a technique to going out for drinks after a hard day. Secondly, it is Ok to be beat as long as you get up and try again the next day. Sometimes science does not work, and you just have to stick with it.

Declaration

I hereby declare that the following thesis is based upon the results of investigations conducted by myself, and this thesis is of my own composition. This thesis has not in whole or part, been previously presented for a higher degree. Work other than my own is clearly indicated in the text by reference to the relevant researchers or their publications. Those people, who allowed me to refer to their unpublished observations or data, are acknowledged individually where they occur.

Mathew Jones

The work presented in this thesis is the work of Mathew Jones. Conditions of the relevant Ordnance and Regulations have been fulfilled.

Professor William N. Hunter

Summary

Pentameric ligand gated ion channels (pLGICs) have an essential role in mammals, insects and some prokaryotes. These membrane receptors convert chemical signals of neurotransmitters into electrical signals of an ion flow, mainly at synapses. Due to this function, the channels are involved in a breadth of biological systems including metabolism, sensory processing and cognition. All the members of the family share a similar structure. A receptor is made up of five subunits. Each subunit has three domains, the extracellular (ECD), transmembrane (TMD) and intracellular (ICD). Neurotransmitters bind in the orthosteric pocket formed at the interface of two ECDs. Ligand binding induces conformational changes that allow ions to flow through a pore generated by the TMD. The ICD is the site of cell-receptor communication.

This work focuses on the glycine receptor (GlyR), which has a role, among others, in chronic pain and whose malfunction leads to diseases such as epilepsy and hyperekplexia. In adults, the primary functional oligomer is a heteromeric arrangement of $\alpha_1\beta$, with other subtypes dispersed throughout the central nervous system (CNS). Current research has produced models of homomeric receptors, however, no structural data are available for the physiologically relevant heteromeric form.

An objective of this work was to characterise a new model for the ECD of the heteromeric GlyR, called glycine binding protein (GBP). GBP is based on *Aplysia californica* acetylcholine binding protein (*Ac*AChBP), a surrogate for the nicotinic acetylcholine receptor (nAChR). Other objectives were the production of another surrogate of a GlyR heteromer utilising models of the homomeric form. Also, the homologous histamine receptor (HisR) from *Apis mellifera* was investigated to understand how histamine binds and relate this binding to potentially mislabelled receptors in mammals. More detailed descriptions and experimental sections can be found in Chapter 1 and 2.

In Chapter 3, the biophysical techniques surface plasmon resonance (SPR), isothermal titration calorimetry (ITC) and tryptophan fluorescence (WF) were assessed utilising *Ac*AChBP and three of its ligands. Comparison of computational docking and thermodynamic parameters led to insights into ligand binding in *Ac*AChBP. Data showed good corroboration between the techniques and highlighted some limitations when considering small molecule interactions.

The archetypal ligands strychnine of GlyR and N-methylbicuculline of the $GABA_A$ receptor ($GABA_AR$) were utilised to investigate binding in GBP. Structural, computational and biophysical data were used for comparisons. Chapter 4 also presents a structure of the GBP:N-methylbicuculline complex at 2.4 Å and a preliminary model of the *Ac*AChBP:N-methylbicuculline complex at 2.9 Å. These structures emphasise

the promiscuous binding properties of N-methylbicuculline and give the first insight into its interaction with GlyR and nAChR. Data from the study of N-methylbicuculline suggest that it is promiscuous due to the flexibility of the orthosteric binding site of pLGIC members as well as the number of interactions the ligand can form, therefore a favourable orientation can always be adopted. Comparisons of structural sequence and binding data for both ligands suggested that GBP may be a good surrogate of the GlyR heteromer.

Utilising GBP and *Ac*AChBP as targets, two ligand screens were carried out; an NMR screen of 435 fluorinated compounds and a computational docking screen of 2,000 small molecules from Maybridge. The screens were conducted to discover new ligands of the proteins. Investigation of these ligands could further understanding of protein-ligand interactions with GBP and *Ac*AChBP, as well as the physiological receptors GlyR and nAChR. Similarities in chemical structure and binding poses of compounds that were shown to bind the target proteins were investigated. Validation of binding for selected small molecules was attempted. Two compounds were successfully validated by WF with K_D values of 237-393 μM. Exploiting published structures, the possibility to extend the small molecule scaffolds into empty pockets of the two proteins was examined. Extensions of three compounds were carried out suggesting potentialgroup additions that may improve binding interactions.

Prior to the initiation of this study, the crystal structures of homomeric GlyR α_1 and α_3 were reported. It was then considered possible to engineer the sequence to present a heteromeric binding site using a homomeric assembly. Utilising the homomeric channels, more sequence and structural conservation to the GlyR heteromer could be achieved. Therefore, a new surrogate was designed by substituting five residues of loop C of GlyR α_3. Loop C acts as a flexible loop that closes upon ligand binding leading to conformational changes, therefore is an important feature for the binding mechanism. Comparison of sequence and structural models of potential HisR constructs was also conducted, which led to the identification of AmHisR. Attempts were made to produce the proteins, however, problems were encountered. Chapter 6 details the approaches attempted to solubilise these membrane proteins. Due to there being no structural studies of HisR, binding models were constructed, and key residues were highlighted. Potentially, the same approach as for GBP could be employed using these models to produce an *Ac*AChBP-based surrogate for HisR.

1. Introduction

1.1 Pentameric ligand gated ion channels

1.1.1 pLGICs role in synaptic transmission

Pentameric ligand gated ion channels (pLGICs) are a large family of receptors located at neuronal synapses in vertebrates, invertebrates and are also present in a few prokaryotes. At synaptic junctions, they function as an intermediate translator between chemical and electrical signals for the continuation or inhibition of action potentials. There are two classes of pLGICs, dependent on ion selectivity, excitatory including nicotinic acetylcholine receptors (nAChR) and serotonin receptors (5-HT_3R) as well as inhibitory, which comprises $GABA_A$ receptors ($GABA_A$R) and glycine receptors (GlyR) (1).

Due to their location at the synapse, they are implicated in a breadth of human biological processes including motor control, sensory processing, sleep and pain (2). Owing to this involvement in key biological processes, when these receptors malfunction this leads to a variety of diseases (3). Malfunctions in nAChR and $GABA_A$R are linked with Alzheimer's (4,5), Parkinson's (6,7) and epilepsy (8,9), whereas defective GlyR function is linked with hyperekplexia (10) and chronic pain (11). These receptors are subject to the influence of some widely known compounds with examples including propofol and nicotine. The anaesthetic propofol acts as an allosteric modulator of both excitatory and inhibitory channels (12). Whereas nicotine is the cause of smoking addiction but is now also used to aid smoking cessation and targets nAChR (13). Therefore, a detailed understanding of pLGIC function and of how chemicals can modulate these receptors could inform early-stage drug discovery.

1.1.2 Structural insights of pLGICs

1.1.2.1 General architecture

Important insights into pLGIC structural configuration came from cryo-EM models of the nAChR from *Torpedo marmorata*. Unwin *et al.* served to define the overall architecture of pLGICs. Each subunit consists of three domains, the extracellular (ECD), transmembrane (TMD) and intracellular (ICD), arranged in a pseudo-pentameric motif of five subunits (14-16). Since then the amino acid sequences and the structures of many different members of the family have been elucidated (Table 1.1) and have been found to share similarities (17). The discovery of the prokaryote homologues, *Erwinia chrysanthemi* ligand gated ion channel (ELIC) (18) and *Gloeobacter violaceus* ligand gated ion channel (GLIC) (19), gave further structural insights and highlighted the importance of this family across organisms (20). Depending on receptor subtype, several different combinations can assemble as a pentamer. For example, $GABA_A$R has 19 genes linked to different subunit types (21). Both homomeric

and heteromeric receptors can be formed, as is the case in GlyR (one homomeric and four heteromeric (22)). These subtypes can function differently and are separated spatially and/or temporally (1). Structures of homomeric receptors were reported for the glutamate receptor (GluCl) (23), $GABA_AR$ β_3 (21), GlyR α_3 (24) and α_1 (25). The first heteromeric receptor of nAChR $\alpha_4\beta_2$ was published in 2016 (26) and is present in two stoichiometric forms (3α:2β or 2α:3β) (1). Further to this, structures of $GABA_AR$ heteromers were reported between 2017 and 2020 (27-33). Some of these models are important as they provide the first examples of tri-heteromeric receptors.

Organism	Protein	Technique	Resolution (Å)	Year (ref)
Lymnaea stagnalis	AChBP	X-ray	2.7	2001 (34)
Aplysia californica	AChBP	X-ray	2.2	2004 (35)
Torpedo marmorata	nAChR αβγδ	Cryo-EM	4.0	2005 (16)
Erwinia chrysanthemi	ELIC	X-ray	3.3	2008 (18)
Gloeobacter violaceus	GLIC	X-ray	3.1	2009 (19)
Caenorhabditis elegans	GluCl	X-ray	3.3	2011 (23)
Homo sapiens	$GABA_AR$ β_3	X-ray	3.0	2014 (21)
Mus musculus	5-HT_3R	X-ray	3.5	2014 (36)
Homo sapiens	GlyR α_3	X-ray	3.0	2015 (24)
Danio rerio	GlyR α_1	Cryo-EM	3.8	2015 (25)
Homo sapiens	nAChR $\alpha_4\beta_2$	X-ray	3.9	2016 (26)
Rattus norvegicus	$GABA_AR$ $\alpha_1\beta_1\gamma_{2S}$	Cryo-EM	3.8	2018 (31)

Table 1.1 Published structures of members of the pLGIC family. Structures are ordered by date. X-ray: X-ray crystallography and Cryo-EM: cryogenic electron microscopy.

1.1.2.2 Extracellular Domain

The ECD contains the orthosteric binding site at the interface between two subunits, where neurotransmitters interact with the receptor. The domain is generated by ten β-strands (β1-10) which generates a β-sandwich core (37). The study of acetylcholine binding protein (AChBP), a homologue of the ECD of pLGICs, has contributed to the understanding of the structure and function of this domain. AChBP is a small, soluble and highly stable protein found in molluscs, where it has a modulatory role in synaptic transmission. The structure of *Lymnaea stagnalis (Ls)* AChBP was first determined in 2001 (34) and later the orthologue from *Aplysia californica* (*Ac*AChBP) was reported (35). This protein, *Ac*AChBP, is used in this work. The two subunits involved in ligand binding are assigned as (+)-principal and (-)-complementary. The principal subunit donates three loops (A, B and C), then the complementary subunit donates four loops (D, E, F and G, Figure 1.1) to generate the orthosteric site (38).

Heteromeric channels offer distinctive binding sites within the pentamer. Due to the stoichiometry and arrangement of subunits a range of interfaces can be observed, generated by different binding site residues. For example, GlyR can be made up of α and β subunits, however, glycine preferentially

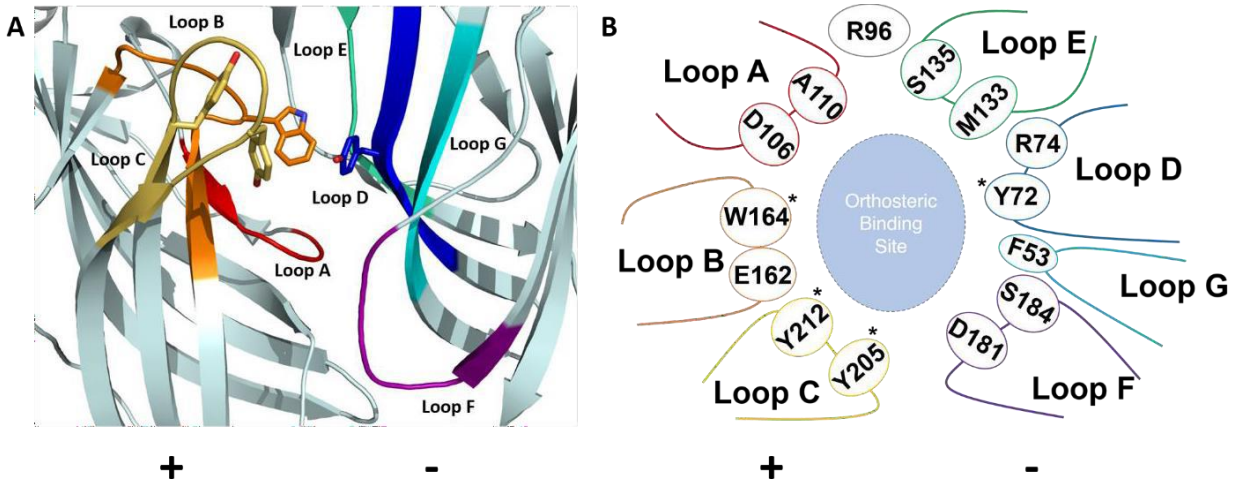


Figure 1.1 Binding loops of the orthosteric site of GBP. A) View of the binding site showing, loop A (red), loop B (orange), loop C (yellow) from the principal-(+) subunit and loop D (blue), loop E (green), loop F (purple), loop G (cyan) of the complementary-(-) subunit. Four residues of the aromatic cage are shown as sticks, defining the edges of the binding site. B) Schematic representation of the binding site, starred residues are responsible for the aromatic cage, shown in sticks.

binds at the $\beta^{+}\alpha^{-}$ site (39). Some key similarities occur in the orthosteric binding pockets of the pLGIC family. There is a motif of four conserved aromatic amino acids, which form interactions with the ligand. In *Ac*AChBP, these are Y72, W164, Y205 and Y212 (37). Another conserved feature is loop C, a flexible region able to adjust position during ligand binding (25). The closure of loop C is linked with global conformational changes of the ECD that may promote the binding of ligands in other pockets as well as transferring the signal to the TMD (29). When strychnine, an antagonist of GlyR, binds to the ECD, it blocks the closing of loop C and thus prevents any conformational changes from occurring (Figure 1.2) (25).

1.1.2.3 Transmembrane Domain

The TMD consists of four membrane-spanning α-helices (M1-4). The inner helix (M2) lines the vestibule which is punctuated by charged residues that are utilised for ion selection (40). The M1/M3 helices provide a barrier between M2 and the lipid environment. Helix M4 is the main site of interaction with the lipid bilayer (41).

The channels cycle through three states: resting, open and desensitised (Figure 1.3B). The resting state is present before ligand binding to the ECD. The ion pore is constricted and contains a conserved hydrophobic patch, which generates an energetic barrier preventing solvated ions from passing through the pore (42). Upon ligand binding, conformational changes of the ECD are communicated to the TMD via a covalent bond between β10 and M1, as well as non-covalent interactions between the β1-2 and β6-7 (Cys) loops with M2 and M3 (21,41). This leads to a twist and

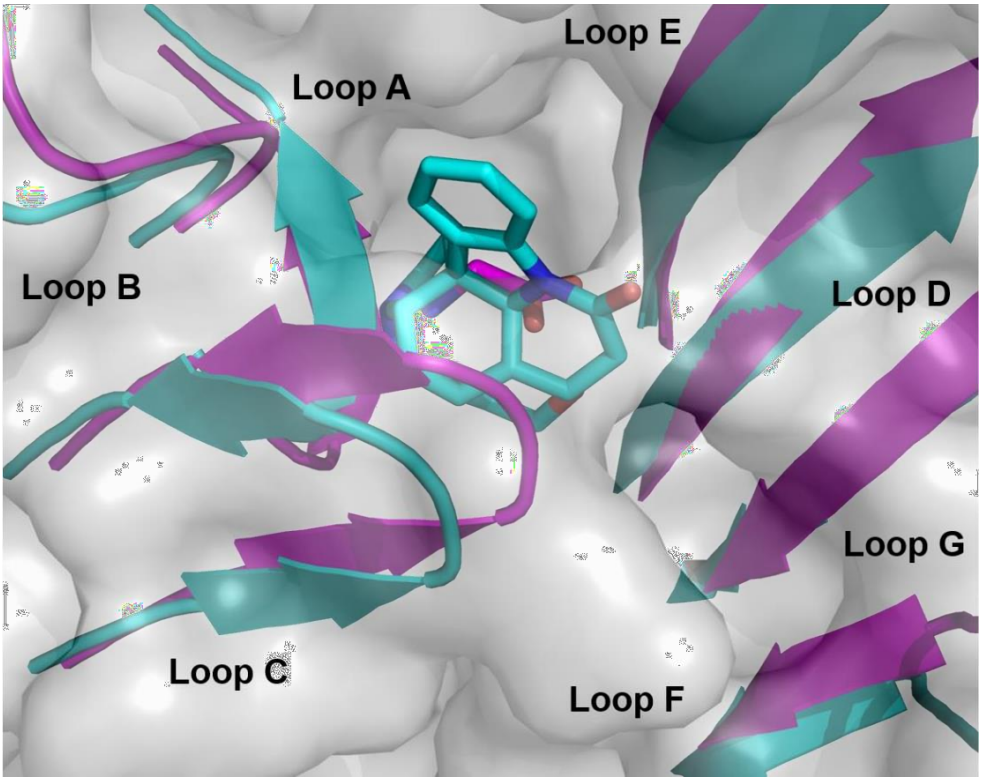


Figure 1.2 Key orthosteric binding loops of GlyR structures superimposed. Complexes with glycine (magenta, PDB code: 5VDH) and strychnine (cyan, PDB code: 5CFB) are compared. Strychnine and glycine are shown as sticks.

tilt of the pore-lining helix M2, opening the gating region and allowing solvated ion flow along the electrochemical gradient. This conformational change is a transition from the closed to open state. Ion selection is determined by the residues lining the pore. In nAChR, the residues of the channel are negatively charged, therefore select for cations whereas in inhibitory receptors, the residues are positively charged and select for Cl^-. The desensitisation state is finally adopted (41). This results in constriction of the pore towards the ICD preventing ion flow due to rearrangement of the TMD helices (17). This state also has an increased affinity for ligands (41). It is proposed that the receptor returns to a resting state by an untwisting of the ECD in an *apo* configuration (17).

Modification of lipid composition can reduce the agonist response or lead to slower desensitisation states (41,43,44). Several sites have been identified where lipids bind at the protein-lipid interface (annular) and between helices of the TMD (non-annular) (45). Structural evidence from nAChR models implies that specific lipids e.g. cholesterol do interact with these sites (46,47) and may affect the kinetics of transition state switching. However, more research is required before firm conclusions can be made about the role lipids have on pLGIC function (41).

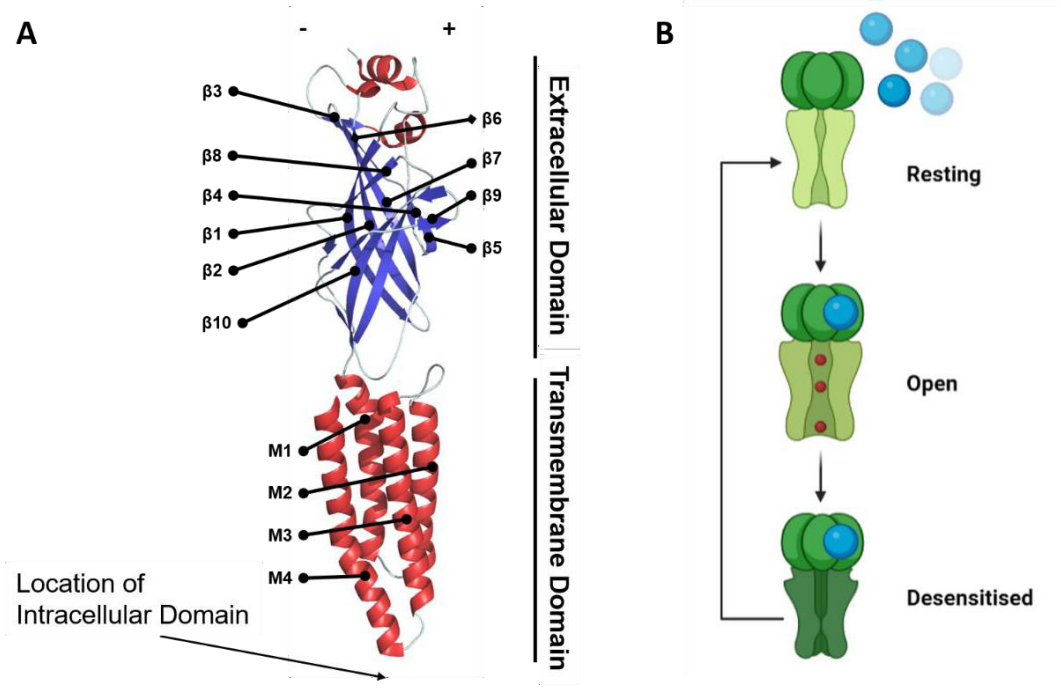


Figure 1.3 Subunit organisation and receptor function. A) The organisation of domains in an individual subunit. The β-strands (blue) that make up the ECD are labelled as are the helices (red) that form the TMD. The ICD was removed from the protein to aid crystallization (24). B) Schematic of the resting, open and desensitised states that a pLGIC proceeds through. The ECD and TMD are shown for each step. Ligands (blue sphere) and ions (red sphere) are also shown.

1.1.2.4 Intracellular domain

The ICD consists of a mostly disordered loop between the M3-M4 helices. The disordered regions have roles in membrane localisation and modulation (48). However, the initial and final sections of the loop are structured into helices which are involved in ion conductance (36).

As well as providing receptor-specific functions, the ICD provides a binding site for many proteins to interact with the channel, in order to modulate its function and downstream effectors. One example of an interacting protein is collybistin, which binds a proline-rich motif of the ICD and is a GTP/GDP exchange factor linked to the binding of gephyrin (49). Like collybistin, gephyrin is another ICD binder that leads to the anchoring and clustering of GlyRs, and some $GABA_A$Rs at the membrane (50).

1.1.3 Nicotinic acetylcholine receptors

There are two types of nAChRs, muscle and neuronal. The muscle nAChRs are located at the neuromuscular junction and can be made of five subunits in a stoichiometry of 2α, 1β, 1γ, 1δ or 1ε, similar to the electric organ in *T. marmorata*. The other nAChR subunits ($\alpha_{2\text{-}7,\,9,\,10}$ and $\beta_{2\text{-}4}$) were first identified from neuronal tissues, however, are located throughout the nervous and other systems (51). These receptors are implicated in a host of biological systems including learning and metabolism

(52). They are formed from α subunit homomers or α and β heteromers with stoichiometry affecting ligand binding, receptor function and downstream applications (53). For example, the α_5 subunit has no role in ligand binding (51,54), however, when partnered with the $\alpha_4\beta_2$ receptor, the predominant oligomer of the central nervous system (CNS) (55), the α_5 subunit affects receptor assembly and ligand-mediated upregulation (54,56). Also, the α_7 receptor, the other main oligomer of neuronal nAChRs, is primarily a homomer that is more permeable to Ca^{2+} ions than other nAChRs (57). This may seem trivial until one considers that a Ca^{2+} influx is linked to many secondary signalling cascades (58). Even though nAChRs are diverse, only a few structures have been published. This includes the first heteromeric structure of a mammalian pLGIC, α4β2, which provides insights into how heteromeric and homomeric receptors differ (26).

Due to the experimental challenges of producing full length pLGICs, AChBP has been utilised as a surrogate. AChBP, as previously mentioned (Chapter 1.1.2.2), shares homology with the ECD of pLGICs. It shares most sequence homology with the α subunits of nAChR (27% α_4 and α_7), therefore has been utilised in the investigation of ligand interactions with these receptors (59) and for early-stage drug discovery (60). Amino acid substitutions have been made to AChBP to generate better surrogates of nAChR oligomers (61-64) as well as to study 5-HT_3R (23% sequence identity) (65). During these investigations, many techniques using AChBP as a surrogate have been validated including tryptophan fluorescence (66), isothermal titration calorimetry (35), X-ray crystallography (34), surface plasmon resonance (67) and computational docking (68), providing a basis for future work utilising similar surrogate proteins.

1.1.4 Glycine receptors

GlyR has been linked to different biological processes including motor control and sensory signalling, therefore is important for the functioning of the human body (69). GlyR can be constructed from five different subunits $\alpha_{1\text{-}4}$ and β. As there are fewer subunits compared to nAChR, there are fewer possible oligomers of the receptor. In adults, the main form is an $\alpha_1\beta$ heteromer with either a 2:3 or 3:2 stoichiometry (39). Experiments utilising antibodies and atomic force microscopy suggest a 2α:3β ratio (70). The $\alpha_1\beta$ heteromer is therefore essential for the majority of GlyR function in the body and loss of either gene has been linked to impairment of reflex circuits resulting in hyperekplexia (22). In embryos, the main oligomer is a homomer of α_2 (71). It has been shown in some areas of the brain that, upon the switch from α_2 to $\alpha_1\beta$, α_2 is still maintained in combination with β (72). α_3 also forms heteromers with β and has been confirmed to localise in the retina and the dorsal horn of the spinal cord (22). Experiments on the later localisation have linked the $\alpha_3\beta$ receptor with chronic inflammatory pain and highlighted it as a potential therapeutic target (11). The final α subunit, α_4, is

a pseudogene with a stop codon before the fourth transmembrane helix (73). The receptor co-localises with β at various locations including the spinal cord and retina but its role is unknown (22,74). Structures of GlyR homomers have been reported (24,25,42), which give some insight into how GlyR functions. However, even though the heteromeric forms of GlyR are predominant and are essential, no structural data are available.

Dawson *et al.* reported a heteromeric GlyR surrogate, termed glycine binding protein (GBP), which utilises the similarity between *Ac*AChBP and the pLGIC family. The heteromeric $\beta^{+}\alpha^{-}$ interface is the site at which glycine binds (39), therefore it is of interest for drug discovery. To generate GBP nine substitutions of important binding residues in the orthosteric pocket were conducted and the interactions it makes with ligands were investigated. These residues were chosen based on sequence and structural alignments. The impact of the substitutions on protein stability was assessed by differential scanning fluorimetry. This was followed up by determining the structure and K_D of important ligands, such as strychnine, by X-ray crystallography and biophysical techniques (75), some of which are included in this thesis. GBP could also provide an option for the utilisation of different high throughput screening platforms to discover new ligands of the GlyR heteromer.

1.1.5 Histamine receptors

Histamine receptors in mammals were considered to be G-protein coupled receptors (GPCR) (76). However, evidence from rat showed fast inhibitory post synaptic potentials, which were blocked by the channel blocker picrotoxin (77). This was taken to infer the presence of an unknown histamine receptor or binding site. Data demonstrating that histamine binds and generates a current in $GABA_AR$ (78) were further confirmed when a structure of the molecule bound at a $\beta^{+}\beta^{-}$ homomeric interface of the human $GABA_AR$ was reported (21). Therefore, it may be possible that subunits of $GABA_AR$ may have been misassigned or have dual functionality. Insects possess histamine receptors (HisRs) which are members of the pLGIC family, these channels are important for sensory signalling (79). It appears that most insects studied possess two subtypes of the receptor, but it is still to be established if these form predominantly homomeric or heteromeric channels (80-82). These receptors are inhibitory and share sequence identity to the insect $GABA_AR$ (19%) (83).

1.2 Biophysical techniques for ligand screening

Nuclear magnetic resonance (NMR), surface plasmon resonance (SPR) and X-ray crystallography are commonly used to identify and characterise protein-ligand interactions. These biophysical techniques are also frequently used for screening compound libraries against protein targets (84-86).

1.2.1 NMR

This technique relies on the magnetic properties of certain isotopes (87), most commonly: ^{1}H, ^{13}C, ^{15}N and ^{19}F. These isotopes have a nuclei spin of ½, therefore can exist in one of two orientations (88). When an external magnetic field is applied to these nuclei, they align in one of two orientations, which have different energies. Slightly more nuclei will adopt the lower energy level, however, it is possible to excite these nuclei and shift them to the higher energy level. The oscillation between the two orientations occurs at a characteristic frequency that is equal to the difference between energy levels. When the number of nuclei in the higher energy level equals that of the lower level, no more excitation can occur, and the system becomes saturated. Higher energy nuclei can lose energy through T_1 (longitudinal) and T_2 (transverse) relaxation returning to the lower energy level. Detection of the energy difference between states leads to the determination of resonance frequency and identification of nucleus type (89). Measurement of individual nuclei oscillations are summed and recorded as a free induction decay (FID), which is Fourier transformed into the standard NMR spectrum. Three observable parameters are commonly reported from this spectrum, chemical shift, intensity/area under the curve and linewidth at half of the peak maximum, which identify the environment nuclei occupy, the population in that environment and the dynamic properties in solution, respectively. Differences in the local chemical environment lead to shielding or deshielding of nuclei by electrons of other atoms, which affects resonance frequency and signal location or chemical shift. Comparison of spectra, for example, with and without ligand present should lead to differences being observed for these three parameters, deducing if and potentially where a molecule binds (88).

NMR can be utilised in ligand screening via two approaches, protein- or ligand-observed (84). Both techniques require one component to be isotope labelled. For protein-observed NMR, the whole protein can be labelled during recombinant production by utilising ^{13}C and ^{15}N sources in the medium. Alternatively, site-specific labelling can be carried out (90). Once labelled, spectra with and without ligand are recorded. A change in one of the three observable measurements could indicate if binding occurs and in some cases the location (91). However, due to the labelling process, the amount of protein required and extensive optimisations, ligand-observed NMR is more commonly used for initial screening (92).

Two common approaches for ligand-observed NMR are saturation transfer difference and water ligand observed via gradient spectroscopy (waterLOGSY). In general, there should be a fast exchange of compounds between the free and bound states, and affinity is limited to 0.1-1000 μM (92). These two techniques require either the protein or solvent to be saturated by irradiation at a specific radio frequency, which is transferred to the compounds (93,94). Differences between initial and binding

spectra lead to the detection of changes in one of the three observable measurements. ^{19}F NMR further simplifies ligand-observed NMR, as fluorine atoms are not present in proteins, therefore can be utilized as a label (95). When a ^{19}F labelled ligand binds to a protein it adopts the slow tumble of the macromolecule resulting in fast relaxation and broad signals (larger linewidth). In comparison, unbound ligands tumble fast leading to slow relaxation and sharp peaks (smaller linewidth). Broadening of the peak (increase in linewidth) leads to the detection of binding (96). As well as linewidth, the chemical shift could change upon binding, but this is dependent on the location of the ^{19}F atom/s with respect to the protein (92). In this study, ^{19}F NMR was used to identify potential ligands for the proteins of interest.

1.2.2 X-ray crystallography

To investigate ligand binding, two approaches are generally used to form crystals for X-ray crystallography, co-crystallisation and ligand soaking (97). Co-crystallisation is when the ligand and protein are mixed in solution and most often incubated for a period of time before crystallisation trials are set up. Crystal soaking relies on the production of crystals of the protein. The protein crystals are soaked in a high concentration ligand solution (protein:ligand, 1:>2), depending on solubility, to saturate the binding sites (98). The ligand in the solution will diffuse through the pores and channels in the protein crystal and results in protein-ligand complexes. In some cases, high concentrations of compounds can be used, providing the potential to detect binding of weaker ligands that other methods may miss (84). Crystal formation is based on the theory of thermodynamic equilibrium. Due to vapour diffusion in the cell, which contains a reservoir and the protein-ligand drop, water transfers from the drop to the reservoir. This process supersaturates the protein within the drop, which is thermodynamically unstable. Protein molecules organise into a lattice and produce a crystal, which increases stability (99).

The scattering of X-rays by electrons in an ordered crystal gives rise to a structure-specific diffraction pattern. This diffraction pattern consists of discrete scattered X-ray beams, called reflections, which can be observed when Braggs law is obeyed. The scattering angles (the positions on the detector) and associated intensities of reflections are determined by the symmetry and three-dimensional electronic structure of the crystal. It is possible to calculate the distribution of electron density in the crystal with knowledge of the mathematical relationship between the diffraction pattern and the electronic structure. Interpretation of the electron density, in terms of atomic structure, then allows for a molecular model to be constructed (100,101). After data processing and refinement, protein-ligand interactions can be inferred. Conversely, X-ray crystallography does not give any information about affinity and therefore needs to be paired with other techniques (84).

1.2.3 Surface plasmon resonance (SPR)

SPR utilises p-polarised light, which is refracted by a glass prism directing the light on to the sensor chip. The electrons on the gold surface of the chip interact with the photons of light, generating surface plasmons. Generation of plasmons is dependent on absorption energy and results in a decrease of light intensity being reflected from the surface, at a specific (SPR) angle. Gold is generally used to coat the sensor surface as it produces an SPR signal and it is non-reactive to solutions used for the technique. Ligand binding to the protein on the biolayer causes a change in mass and the refractive index that leads to a change in SPR angle, which is detected. Changes in SPR angle are proportional to mass changes (102,103). Proteins can be immobilised on the surface by several affinity tag capture systems (104,105) and/or chemical processes (106) to prevent drift. Drift is caused by the protein detaching from the biolayer during experiments (107).

SPR is a high-throughput, sensitive technique that can be utilised to detect ligand binding, calculate K_D and measure binding kinetics during screening (84). Kinetic measurements of small molecules may require optimisation, however, K_D can still be calculated from the steady-state equilibrium (108). Due to its high-throughput properties, experiments can be optimised quickly, identifying problems such as off-target effects (102).

1.3 Complementary biophysical methods for the characterisation of protein-ligand interactions

After ligand screening, other biophysical techniques can be utilised to confirm binding and to gather additional data to improve the understanding of the binding interaction. These data can then be used to develop better compounds against the target.

1.3.1 Biolayer interferometry (BLI)

BLI is an optical technique similar to SPR, but it takes advantage of changes in the wavelength of the interference patterns rather than a change in the angle of reflection. White light is reflected from two layers, the reference and the sample layers. The sample or biolayer is the location at which protein is immobilised and where ligands can bind. A binding event causes a change in the density of this layer, which alters the optical properties leading to a wavelength change (109). The change in wavelength is directly proportional to the thickness of the biolayer (110). Changes are monitored continuously allowing kinetics to be investigated, as well as the affinity constant (111).

1.3.2 Tryptophan fluorescence quenching (WF)

The technique monitors the fluorescence of the aromatic amino acid tryptophan, which is sensitive to the environment it is present in. The other aromatic residues phenylalanine and tyrosine are less sensitive to the surrounding environment. In these residues (tryptophan>tyrosine, Figure 1.4),

excitation of electrons at 280 nm transitions them from a low to a high energy level, which is less stable. Energy absorbed by tyrosine is generally either quenched by its interactions with other residues or can be transferred to neighbouring tryptophans. Excitation is followed by a relaxation (high to low), which emits energy as fluorescence in a certain range (300-400 nm) (112-114). The introduction of a ligand changes the electronic environment of the protein leading to the formation of different interactions that affect the energy transfer and, generally, a quenching effect of the fluorescence is observed (115). As more ligands interact with more binding sites a larger quenching effect is observed until saturation occurs. Utilising changes in fluorescence, the affinity of a compound for a protein can be determined.

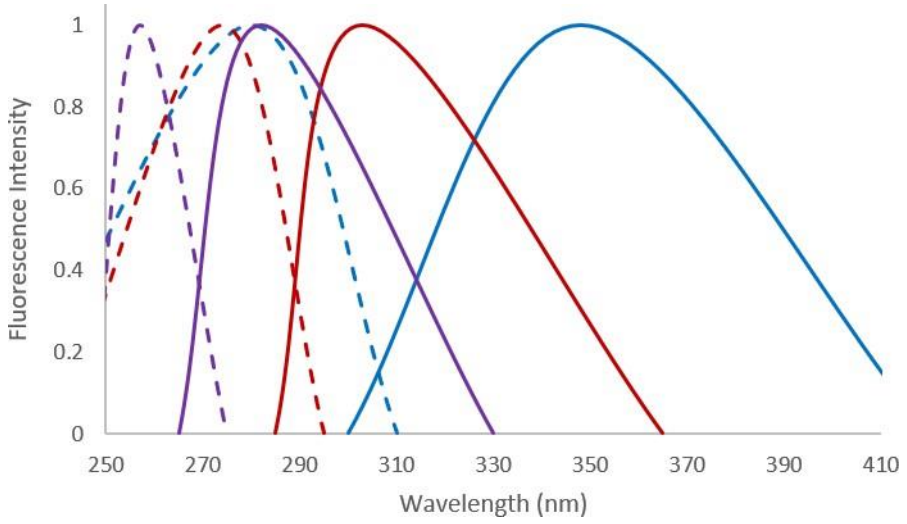


Figure 1.4 Absorption and emission spectra of aromatic amino acids. Example absorption (dashed line) and emission (solid line) spectra for tryptophan (blue), tyrosine (red) and phenylalanine (purple) residues when excited at 280 nm with white light.

1.3.3 Nano differential scanning fluorimetry (nanoDSF)/Differential scanning fluorimetry (DSF)

The intrinsic fluorescence of a protein can be monitored whilst increasing temperature to elucidate the stability of the sample. The fluorescence of the protein is monitored at two wavelengths, 330 and 350 nm corresponding to emission maxima for tyrosine and tryptophan residues (116). As in tryptophan fluorescence, as the electronic environment of these residues changes due to protein unfolding, so does the fluorescence (114). The ratio of the wavelengths is recorded over temperature and from this data the stability of the protein can be inferred. The presence of a ligand can alter the stability of a protein and can be quantified by the change in melting temperature. DSF works similarly except a dye is used to monitor unfolding. As a protein unfolds the hydrophobic core is revealed, therefore more dye binds to the protein. Water quenches the fluorescence of the dye, therefore binding of hydrophobic regions excludes water and increases fluorescence. Upon addition of ligands, the protein stability will be affected and the magnitude can be reported (117).

1.3.4 Isothermal titration calorimetry (ITC)

ITC measures the change in heat energy, either exothermic or endothermic, within a closed system. Two cells are present in the calorimeter, one is a reference and the other is a sample cell, both are maintained at a constant temperature. Upon ligand binding to the protein in the sample cell, it forms different interactions. Generating these interactions can lead to energy being released or taken up, which can lead to a temperature change of the cell. The instrument records the power, proportional to the energy, required to maintain the sample cell at the same temperature as the reference (118,119). If ligand binding is exothermic then the temperature in the sample cell increases. The supply of power decreases until the temperature of the sample and reference cells are equal. The trace of power usage gives rise to a negative peak. In the case of an endothermic process, a positive peak is noted since power increases to heat the sample cell to compensate for the drop in temperature. As binding sites become saturated the number of binding events and therefore energy change decreases, leading to smaller peaks. The peak areas represent a change in energy and the total change in enthalpy can be calculated as explained in Chapter 2.4.3 and the binding curve constructed. From the binding curve, five parameters can be determined, K_D, stoichiometry (N), change in Gibbs free energy, change in enthalpy and change in entropy (120). These parameters give additional information about the binding event and can be used to inform and improve compound development (121).

1.4 DMSO, a good solvent or a complicating factor?

DMSO is a highly effective solvent for organic molecules, therefore is widely used in ligand screening and characterisation. However, its use can lead to complications. Some compounds stored in DMSO have been shown to precipitate, aggregate or degrade (122-124), which generates false results when utilising such solutions in screens. The solvent can also act as a chaotropic agent and affect proteins by destabilising the structure, even at low concentrations (125,126). Disruption of protein structure may affect the interactions a ligand makes, therefore leading to false results in screening and inaccurate thermodynamic and kinetic parameters when characterising the binding interactions. To overcome issues with DMSO different approaches have been suggested, for example, designing more soluble compounds or utilising other solvents (126,127).

1.5 Computational ligand screening

Computational modelling can complement experimental techniques and be used to investigate a wide array of biological processes. For drug discovery, there are an array of computational methods that can be utilised to discover new ligands of target proteins (128), some of which are discussed below.

Molecular docking uses macromolecular structures as templates to dock compounds into and scores the binding based on the interactions they make. By varying the flexibility and restraints of the protein

and/or ligand in the docking programme different protein and ligand conformations can be accounted for (129). Flexible methods are the most computationally expensive, as all aspects of the protein and ligand are allowed to move. Semi-flexible docking allows for the ligand to move within the binding site, which can lead to optimisation of interactions. Sometimes, key amino acids are also allowed to be flexible to further improve binding orientations. Rigid docking prevents any protein or ligand movement (128,130).

Assessment of chemical properties of known ligands can be used to inform computational ligand screening. For these methods, the known ligands are assessed, and attempts made to identify key features, such as the number of hydrogen bonding groups (131,132). Pharmacophores can then be used to screen compound databases to identify common chemical groups and conformations (133). Quantitative structure-activity relationships (QSAR), calculate properties of a ligand that are linked to biological activity (128) and can also be used to search for common compounds in a database (134).

'Fragment-based' and *de novo* drug design have established a new paradigm in drug discovery (128). Small molecules are screened against targets, upon confirmation of binding these compounds are used as scaffolds to develop selective ligands with improved properties (135). Computational modelling allows for the investigation of libraries of different chemical groups and linkers, which may allow for the improvement of the identified scaffolds (136). These analyses can then be used to direct synthetic chemical modifications to be tested experimentally.

1.6 Aims

This research aims to advance our understanding of how selected members of the pLGIC family interact with and are modulated by ligands. In addition, this work aims to develop reagents and protocols that will support further research seeking new chemical modulators of pLGIC members in early-stage drug discovery.

Assessment of GBP as an appropriate model and surrogate of heteromeric GlyR ligand binding will be conducted. New models will also be generated, which may give insights into the ligand binding of other pLGICs. Appropriate methods for the characterisation of protein and ligands need to be identified and robust protocols developed. To address these aims the following objectives have been set:

- Assess the use of the biophysical techniques, WF, ITC and SPR for studying protein-ligand interactions with the surrogate proteins *Ac*AChBP and GBP
- Characterise GBP binding to strychnine in order to establish if this is an appropriate surrogate of GlyR

- Demonstrate the use of GBP in experimental and computational ligand screening to develop an understanding of key protein-ligand interactions and to discover new chemical matter
- Generate an improved heteromeric GlyR surrogate utilising the GlyR α_3 homomer that could further aid early-stage drug discovery
- Characterise interactions of N-methylbicuculline with *Ac*AChBP and GBP to investigate how this important compound may interact with selected members of the pLGIC family
- Attempt production of HisR to identify residues involved in histamine binding

Completion of these objectives should provide new models of ligand binding for various pLGICs and ligands, which will develop the understanding of ligand binding in the pLGIC family. Since the start of the project the field has advanced, for example, the publication of the first heteromeric $GABA_AR$ structure (Table 1.1). These developments were considered as the project progressed.

2. Materials & methods

2.1 Construct design

Vectors for the recombinant production of *Ac*AChBP and GBP were already available having been constructed by Dr. Alice Dawson (75). The sequence for *Ac*AChBP was taken from Uniprot (137) (Q8WSF8) but modified with two conservative substitutions: A60V and A135V, as previously reported (66). The GBP sequence is the same as for *Ac*AChBP with the following substitutions: T53F, Q74R, Y110A, I135S, G162E, S206CCP_KGTG. These substitutions provide a surrogate of the physiologically relevant $\beta^+\alpha^-$ orthosteric binding site of GlyR (75). These vectors, based on the pFastBac1 system (Invitrogen), encoded a product with a C-terminal tobacco etch virus (TEV) protease cleavage site followed by a His_6 tag. The His tag allows for purification by immobilised metal affinity chromatography (IMAC), after which the His tag can be removed by use of TEV protease (138).

To aid the design of other constructs, various sequence and structural alignments were made. The following sequences were extracted from the Uniprot database: *Ac*AChBP (Q8WSF8), *Homo sapiens* (*Hs*) GlyR α_1 (P23415), *Hs*GlyR α_2 (P23416), *Hs*GlyR α_3 (O75311), *Hs*GlyR β (P48167), *Hs*$GABA_AR$ β_1 (P18505), *Apis mellifera* (*Am*) HisR1 (Q0GQR2), *Am*HisR2 (A0A1B1QGD8), *Drosophila melanogaster* (*Dm*) HisR1 (Q9VGI0), *Dm*HisR2 (Q9VDU9), *Nasonia vitripennis* (*Nv*) HisR1 (D3UAF0), *Nv*HisR2 (D3UAF1), *Tribolium castaneum* (*Tc*) HisR1 (A8DMU7) and *Tc*HisR2 (A8DMU8). Sequence alignments were constructed in Jalview (139) using the Clustal Omega or Tcoffee algorithms. These alignments were used to investigate conservation between related sequences, particularly concentrating on the orthosteric binding site.

Experimentally determined structures of relevant proteins were used for comparison. These structures were retrieved from the Protein Data Bank (PDB) (140): GlyR α_1 (PDB code: 3JAF), GlyR α_3 (PDB code: 5TIN and 5CFB) and $GABA_AR$ β_3 (PDB code: 4COF). Structures were not available for HisR, therefore a homology model was built using the Phyre2 platform (141) and SymmDock (142). Structures were superimposed by the secondary structure method (SSM) in Coot (143), which aligns the C_α backbone atoms in multiple iterations. Calculations of root-mean-square deviation (RMSD) of C_α backbone alignments between subunits was carried out by the PDBeFold web server (143). In PyMol (The PyMol Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC), structures were aligned over multiple cycles. Important residues identified in the sequence alignments were examined further in the superimposed structures.

Protein sequences were also inputted into XtalPred web server (144) to investigate the properties of the proteins further. The server gives information about signal sequences, secondary structure, transmembrane and disordered domains as well as a propensity for crystallisation prediction. All these parameters are important considerations when designing constructs.

2.2 Cloning

2.2.1 General cloning methods

2.2.1.1 Transformation of vectors

Transformations of plasmids with competent cells were performed. Plasmid (1 µL, 10 ng) was added to a cell suspension (15-50 µL, depending on competency) and gently mixed. The cells were then incubated on ice for 30 m, next they were heat shocked at 42 °C for 45 s, followed by a further 5 m incubation on ice. SOC (Super Optimal broth with Catabolite repression) medium (150 µL) was added and the culture was incubated at 37 °C for 1 h. The cell suspension (150 µL) was gently mixed then plated out onto the relevant plate type. Plates were incubated for 24 h at 37 °C, unless otherwise stated, to allow for sufficient colony growth. All media and plates used in this study were sourced from the Media Kitchen, School of Life Sciences, University of Dundee.

2.2.1.2 PCR

PCR was used for the amplification of DNA and to introduce modifications. A solution of DNA, buffer, dNTPs, $MgSO_4$, DMSO, forward and reverse primers, water and polymerase was prepared (Table 2.1). The PCR was run in duplicate. A negative control without the addition of polymerase was also run. All PCR reactions (Table 2.2) were conducted on a Perkin Elmer GeneAmp PCR system 2400.

Primers contained overhangs, terminal residues that were not encoded in the original DNA sequence. Overhangs are necessary to allow the restriction digest enzymes to be able to bind the DNA and to cleave at the correct site (New England Biolabs [NEB], Cleavage Close to the End of DNA Fragments). DMSO is added to reactions to prevent DNA secondary structure formation, therefore improving yields (145).

Touchdown PCR (146) was sometimes employed for the construction and amplification of modified constructs. The reaction mix is the same for a basic PCR, however, the protocol is different (Table 2.2).

Colony PCR is an analytical approach that can be used to detect if colonies contain the plasmid/gene of interest, before sequencing (147). A different solution is prepared of DNA extract, GoTaq Green Master Mix, forward and reverse primers and water (Table 2.1). As this is a preliminary analysis only one reaction per colony was conducted and no controls were included. The method for a basic PCR reaction was run (Table 2.2). Analysis of the DNA gel led to the identification of colonies with the plasmid/gene of interest for use to miniprep (Qiagen kit), and to acquire enough sample for sequencing. All sequencing was carried out by the DNA Sequencing and Services facility, School of Life Sciences, University of Dundee.

	PCR (µL)	Touchdown PCR (µL)	Colony PCR (µL)
DNA (10 ng)	1	1	1
Hot Start Buffer (10x)	5	5	-
dNTPs (10 mM each)	5	5	-
$MgSO_4$ (25 mM)	2.5	2.5	-
DMSO (100%)	2.5	2.5	-
Forward primer (10 µM)	3	3	3
Reverse primer (10 µM)	3	3	3
KOD Hot Start polymerase (1 U μL^{-1})	1	1	-
GoTaq green master mix (2x)	-	-	25
Water	27	27	18
Total	50	50	50

Table 2.1 PCR solution set up.

	PCR	Touchdown PCR	Colony PCR
Initialise	94 °C for 5 m	94 °C for 3 m	94 °C for 5 m
Denature	94 °C for 30 s	94 °C for 30 s	94 °C for 30 s
Anneal	65 °C for 30 s	68-50 °C for 30 s	65 °C for 30 s
Extension	72 °C for 1 m kbp^{-1}	72 °C for 1 m kbp^{-1}	72 °C for 1 m kbp^{-1}
Final	72 °C for 10 m	72 °C for 5 m	72 °C for 10 m
Cycles	35	5 each for 68, 60, 55 °C 15 for 50 °C	35

Table 2.2 Methods for running a PCR.

2.2.1.3 TOPO cloning

TOPO blunt end ligation does not require digestion of the PCR products and can be ligated into a TOPO vector. This vector can then be sequenced using the M13 primers. The Zero Blunt TOPO PCR Cloning kit (Invitrogen, ThermoFisher) was utilised for this process.

2.2.1.4 Restriction digests

Restriction digestion uses enzymes that recognise certain DNA sequences and cleave at a specific site (148). The cleavage leaves an overhanging, single-stranded piece of DNA known as a 'sticky end'. DNA (1 µL, 1 µg) was mixed with FastDigest Green Buffer (ThermoFisher) and two of the relevant FastDigest restriction enzymes (1 µL each, ThermoFisher). The solution was mixed and left at room temperature (RT) for 15-30 m. Controls were also set up with one enzyme and no enzymes. After the digestion, the solutions were heated at 80 °C for 5 m to denature the enzymes and stop the digestion. Samples were then analysed/purified by DNA gel electrophoresis.

2.2.1.5 Ligations

Ligations take advantage of the complementary 'sticky ends' resulting from restriction digestion. A higher ratio of insert to vector is used to improve the ligation of the two, instead of just religating the vector closed. In a typical reaction, the vector (0.5 µL, 0.025 pmol) and insert (2 µL, 0.1 pmol) are mixed with the Clonables Ligation Premix (2.5 µL, 2x, Novagen, Merck). This is then left for 5 m at RT, after which it is transformed into NovaBlue cells (Novagen, Merck) and plated on lysogeny broth (LB) agar ampicillin plates.

2.2.1.6 DNA gel electrophoresis

Gel electrophoresis was used to analyse and purify DNA products, using a 1% (w/v) agarose gel with ethidium bromide (1 µL, 10 mg mL^{-1}, Sigma). Samples (40 µL) mixed with loading dye (6 µL, 0.25% [w/v] bromophenol blue, 0.25% [w/v] xylene cyanol FF, 30% [v/v] glycerol) were loaded into one of the eight wells of the gel and compared to a DNA ladder (Hyperladder 1 kb, Bioline). Colony PCR and digested samples already contained dye, therefore this addition was not required. The gel was then run for 45 m at 100 V, or until the dye front reached the end of the gel. Ethidium bromide can intercalate into the DNA and fluoresces under UV light (149), therefore the DNA bands can be observed using a UV illuminator (BDH) or a BioRad GelDoc EZ system. Bands representing the desired products were excised from the gel and the DNA purified using an extraction kit (Qiagen).

2.2.1.7 DNA concentration determination

DNA concentration determination was conducted on all final constructs, using a DeNovix spectrophotometer (A_{280}, Cambridge Bioscience) and was measured in µg $µL^{-1}$. The constructs were labelled, frozen and stored at -20 °C.

2.2.2 Preparation of baculovirus for expression in insect cells

A crystal structure of the human GlyR α_3 homomer was published (24). The gene encoding for the subunit was modified with the removal of the M3-M4 loop and replaced with an AGT tripeptide linker (23,24). The gene for this construct was ordered. The gene for the GlyR $\beta^+\alpha^-$ surrogate was identical to the previous construct except for five substitutions (H234Y, N236KG, K239Y, F240Y), designed to make a surrogate heteromeric $\beta^+\alpha^-$ orthosteric binding pocket. The gene for the full length *Am*HisR1 was ordered with a C-terminal TEV cleavage site and His_6 tag. All three genes, codon optimised for baculovirus expression, were ordered from GenScript in pFastBac1 vectors.

To generate the GFP-encoding constructs, different cloning strategies (Figure 2.1) were adopted utilising touchdown PCR (146). The GlyR $\beta^+\alpha^-$ surrogate and HisR plasmids were used as templates to design the new constructs. The final PCR products were ligated back into the pFastBac1 vector.

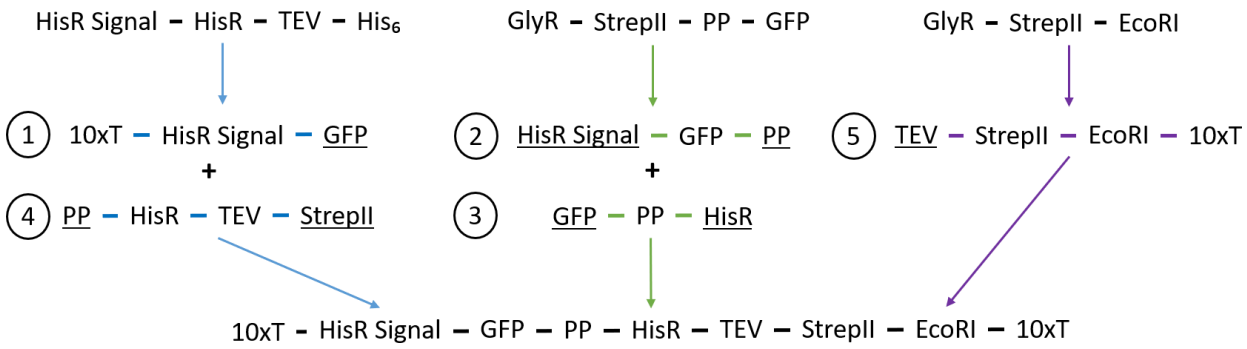


Figure 2.1 Cloning strategy for the construction of GFP-tagged HisR. Touchdown PCR was utilised. Three vectors were used (HisR, GlyRGFP and GlyR) to derive the relevant parts of the new vector. HisR signal is the signal sequence of the histamine receptor to target the receptor to the membrane and also contains the sequence for the BamHI restriction site. GFP is the green fluorescent protein for imaging. PP is the precision protease cleavage site to remove the GFP tag. HisR is the histamine receptor protein. TEV is the tobacco etch virus protease cleavage site to remove the StrepII tag. Underlined segments are overhangs and are used to stitch the gene fragments together in order 1-5. The 10xT are thymine overhangs at the N- and C-terminus to allow for restriction digest. The new construct was then digested using BamHI and EcoRI then inserted into a pFastBac1 vector.

Baculovirus production utilised the Bac-to-Bac system (Invitrogen). DH10Bac cells (Invitrogen) were transformed with the pFastBac1 vector containing the gene of interest. The culture was plated out onto insect cell plates (50 μg mL^{-1} kanamycin, 10 μg mL^{-1} tetracycline, 7 μg mL^{-1} gentamycin, 100 μg mL^{-1} X-gal and 40 μg mL^{-1} IPTG) and incubated for 3 d at 37 °C. Antibiotic selection is important, limiting growth to only a specific bacterium that contains both the viral DNA and the transformed pFastBac vector. Afterwards, a plate was checked for both blue and white colonies. X-gal can be cleaved by β-galactosidase (β-gal) into a molecule that forms an insoluble blue pigment (5,5'-dibromo-4,4'-dichloro-indigo). An inactive form of the enzyme (β-gal) is present in the host *Escherichia coli*. The bacmid (viral DNA) contains an α-peptide sequence, which when produced can generate an active form. If the gene of interest is transferred from the vector to the bacmid, this disrupts the α-peptide. The result is that white colonies are observed if the β-galactosidase remains inactive since no pigment is produced. This should signal that the gene of interest has been inserted.

A single white colony was spread onto a new insect cell plate and left for a further 3 d. The plate was checked again for blue and white colonies. A white colony streak was used to inoculate a culture of insect cell medium (500 mL). The culture was shaken at 200 rpm for 24 h at 37 °C, then centrifuged at 1,912 g for 15 m at 4 °C (Sigma 4K15 centrifuge, 11150 rotor). The resulting pellet was used to extract the bacmid (carried out by the DNA Sequencing and Services facility), which was stored at -20 °C until required.

2.2.3 Preparation of a lentivirus-based expression system for use in mammalian cells

A published method was utilised (150). The main safety concern is the formation of a virus that can self-replicate, therefore steps to ensure this does not happen were taken. Three plasmids are required to generate the lentivirus particles: transfer (pTransfer), packaging (pPackaging) and envelope (pEnvelope). Among these three plasmids are the minimal required viral elements necessary for protein production. As the components are separated, a replication-competent virus should not be formed. However, in the small number of cases a self-replicating virus occurs, a deletion within pTransfer inhibits replication after one cycle (150). The pTransfer (pHR-CMV-TetO2_3C-mVenus-Twin-Strep) vector was ordered from Addgene. The pPackaging and pEnvelope plasmids were acquired from the laboratory of Dr. Ignacio Moraga (University of Dundee). PCR was used to change the restriction sites of the HisR and GlyR surrogate genes to match that of the pTransfer vector. The vector and inserts were digested, ligated and transformed, then plated onto LB-agar ampicillin plates.

2.3 Recombinant protein production

2.3.1 General methods for the analysis of protein production and purification samples

2.3.1.1 SDS-PAGE

Samples (10 µL) were mixed with urea (10 µL, 6 M) and Laemmli sample buffer (10 µL, 2x, BioRad and 355 mM β-mercaptoethanol [BME]), before being heated at 70 °C (membrane proteins) or 100 °C (soluble proteins) for 4 m. Samples (13 µL) were loaded onto a polyacrylamide protein gel (Mini-PROTEAN TGX stain-free precast gels, BioRad) with a marker (8 µL, Precision Plus Protein Unstained

Protein	Expression System	Vector	Restriction Sites	Tags [(C) or (N) terminal]
***Ac*AChBP**	Baculovirus	pFastBac1	-	TEV-His_6 (C)
GBP	Baculovirus	pFastBac1	-	TEV-His_6 (C)
GlyR α_3	Baculovirus	pFastBac1	BamHI/EcoRI	StrepII (C)
GlyR α_3 surrogate	Baculovirus	pFastBac1	BamHI/EcoRI	StrepII (C)
GlyR α_3 surrogate	Baculovirus	pFastBac1	BamHI/EcoRI	StrepII-PP-GFP (C)
GlyR α_3 surrogate	Baculovirus	pFastBac1	BamHI/EcoRI	GFP-PP (N) StrepII (C)
GlyR α_3 surrogate	Lentivirus	pTransfer	EcoRI/XhoI	PP-mVenus-StrepII-StrepII (C)
HisR	Baculovirus	pFastBac1	BamHI/EcoRI	TEV-His_6 (C)
HisR	Baculovirus	pFastBac1	BamHI/EcoRI	GFP-PP (N) TEV-StrepII (C)
HisR	Lentivirus	pTransfer	EcoRI/XhoI	PP-mVenus-StrepII-StrepII (C)

Table 2.3 Protein production vectors.

Standards, BioRad), and run at 300 V for 17 m in TGS buffer (Table 2.4). If the gel was to be analysed by Western blotting a stained marker (8 µL, Prestained Protein Standards All Blue, Biorad) was also included. Gels were visualised with a BioRad GelDoc EZ system.

2.3.1.2 Western blot analysis

After visualisation, gels were transferred using the Trans-Blot Turbo Transfer System (BioRad) onto 0.2 µm polyvinylidene difluoride (PVDF) membranes (BioRad). The top left corner of each membrane was cut to define the orientation. The membrane was agitated with milk powder (5% [w/v]) in phosphate-buffered saline (PBS)-Tween (0.1% [v/v] Tween-20) for 1-24 h at 4 °C, to block non-specific binding sites. The primary antibody was added with further agitation for 1-24 h at 4 °C. The primary antibodies used were mouse anti-His (1:4,000, Abcam) and rabbit anti-Strep (1:10,000, Abcam). The membrane was then washed three times at RT with PBS-Tween. The secondary antibody was diluted in PBS-Tween and poured onto the membrane. Goat anti-mouse IgG (1:150,000, Abcam) and goat anti-rabbit IgG (1:50,000, Abcam) secondary antibodies were used. The membrane was agitated for 1 h at 4 °C. Again, the membrane was washed three times with PBS-Tween at RT. Enhanced chemiluminescence (ECL) Western blotting substrate (Pierce) was applied (2 mL $blot^{-1}$) to the membrane for 5 m, after which the excess was removed. The membrane was enclosed in a plastic wallet inside a dark box (Kodak). In a dark room, the film (Lumi-film chemiluminescent detection film, Sigma) was exposed to the membrane (0.5-10 m) and developed (Protec ECOMAX X-Ray Film Processor). The film and blot were overlaid, and the stained ladder intervals were marked.

2.3.1.3 Analytical gel filtration

Analytical gel filtration (AGF) was used to investigate the size and quaternary state of proteins. A Superdex 200 10/300 GL (GE Healthcare) column was equilibrated overnight in buffer then a sample (600 µL, 1 mg mL^{-1}) was loaded onto the column from a 0.5 mL loop. The sample was run for 1.5 column volumes with buffer. Due to the matrix within the column, large proteins/complexes run through faster than small proteins, which interact with the pores (151). The retention time of each peak was noted, and the molecular weight (MW) was extracted from a previously determined calibration curve. For GBP and *Ac*AChBP there are usually three peaks representing monomeric, pentameric and at least one other multimeric state (Figure 2.2).

2.3.1.4 Native gel analysis

Native-PAGE analysis was used to investigate the size of proteins in their native states (152). The native-PAGE gel was prepared as per the manufacturer's instructions (Novex), then a sample of protein (0.5 mg mL^{-1}) was loaded onto the gel with a marker (8 µL, NativeMark, Novex). The sample was run through the

gel for 90 m at 150 V. The gel was fixed (40% [v/v] methanol, 10% [v/v] acetic acid) and destained (8% [v/v] acetic acid), then visualised using a BioRad Gel Doc EZ Imager.

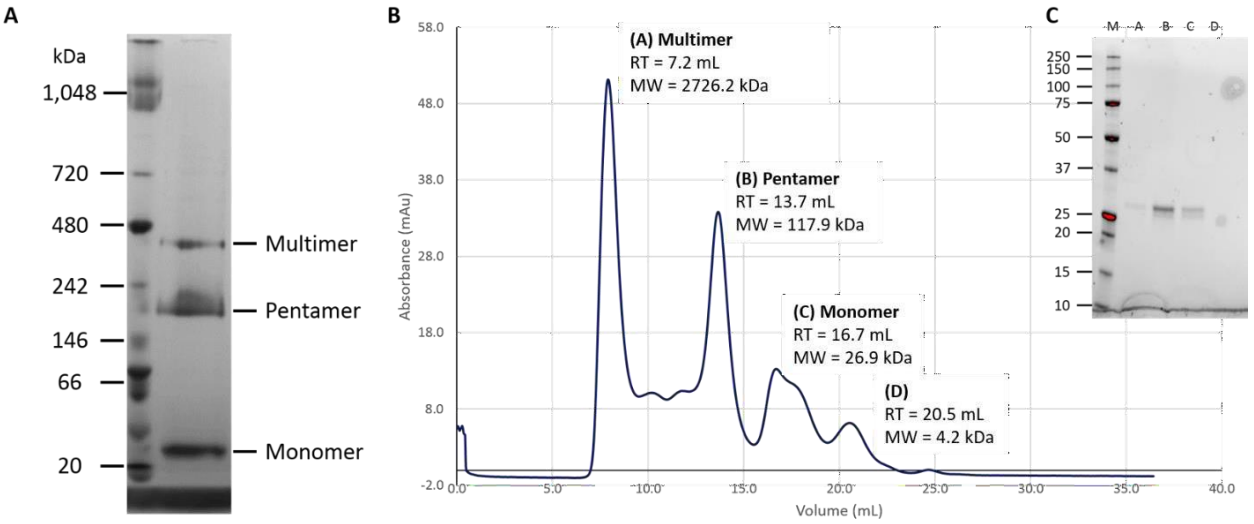


Figure 2.2 Examples of data used to inform protein purification. (A) A native-PAGE gel of GBP showing the presence of monomeric, pentameric and a higher-order form, possibly two pentamers. (B) A size exclusion chromatography profile showing the separation of different GBP assemblies: a multimer in the void (peak A), pentameric (peak B) and monomeric (peak C) forms with the retention times (RT) and estimated molecular weights (MW) shown for each peak. (C) SDS-PAGE gel relating to samples taken from the column in (B), showing that peaks A, B and C contain GBP. Fractions comprising peak B were pooled andused for binding assays and the crystallographic studies.

2.3.1.5 Protein concentration determination

Absorbance measurements at 280 nm were measured on a DeNovix spectrometer (Cambridge Bioscience). Determination of protein concentration was carried out using the Beer-Lambert law and the predicted molar extinction coefficients (ProtParam, (153)). Concentrated protein samples were kept at 4 °C if they were going to be used immediately. Samples were flash-frozen and stored long term at -80 °C.

2.3.2 The use of baculovirus for protein production

2.3.2.1 Virus production

All cell work was conducted in a tissue culture laminar flow cabinet, to reduce the risk of contamination. *Spodoptera frugiperda* (*Sf9*, Protein Production Team, University of Dundee) cells were plated out into a six-well plate in 1 mL of medium (Sf900 III, Gibco) at a density of $8x10^5$ cells mL^{-1} per well. A transfection mix consisting of 1 mL medium, 5 µL GeneJuice transfection reagent (Novagen) and 500 ng of the relevant bacmid, was pipetted dropwise into each well. Transfections were set up with the bacmid of interest, a positive control (dihydrofolate reductase bacmid) and negative control (no bacmid), each condition was carried out in duplicate. The plate was placed in a sealed container with a damp (sterile PBS, Gibco, Sigma) paper towel, to prevent wells from drying out. This container was then placed in a static

incubator at 27 °C. The following day, each well was topped up with dropwise addition of medium (1 mL), so as not to disturb the adherent cells. The samples were then left for 6 d at 27 °C.

The supernatant from each well was collected, this contained the virus (P0). Cultures of *Sf*9 cells (50 mL) were set up at a density of $20x10^5$ cells mL^{-1} in medium (supplemented with glutamine, penicillin and streptomycin), then the P0 virus (125 µL) was added from each of the three conditions (gene of interest, positive and negative controls), into separate flasks. This step was to amplify the virus. Cultures were left shaking at 135 rpm for 3-4 d at 26.5 °C. Three parameters were recorded from the cell counter (Cellometer Auto 1000, Nexcelom): cell density, diameter and viability. *Sf9* cells double every 24 h depending on the medium, therefore a potent virus should limit their growth and reduce the cell density. Upon infection the cells swell as the virus replicates, therefore cell diameter should increase (usually from 15 to 18 µm). Viability is calculated from the ratio of live to dead cells, therefore as cells die due to the release of viral particles the viability should decrease. Typical parameters before harvesting the virus are cell density $30\text{-}40x10^5$ cells mL^{-1}, diameter 18 µm and viability <96%. As the virus replicates mutations can occur, therefore the virus was collected early (90-95% viability) to ensure high quality. Transfection success can be determined by the length of time taken until virus collection. Cultures were centrifuged at 1,000 g for 10 m at 4 °C (Beckman Allegra X-12R centrifuge with a SX4750 rotor). The supernatant was collected and stored at 4 °C, this contains the virus (P1).

A final amplification step was conducted to maximise the number of high-quality virus particles present in the supernatant. Cultures of *Sf*9 cells (200 mL) were set up at a density of $20x10^5$ cells mL^{-1}, then the P1 virus (500 µL), containing the gene of interest, was added. Flasks were left at 26.5 °C for 3-4 d, shaking at 135 rpm. Cell density, diameter and viability were measured as previous. Cultures were centrifuged at 1,000 g for 10 m at 4 °C (Beckman Allegra X-12R centrifuge with a SX4750 rotor) then the supernatant was collected and stored at 4 °C. This supernatant contains the virus (P2).

2.3.2.2 Soluble protein production

For the production of both *Ac*AChBP and GBP, cultures of 500 mL *Trichoplusia ni* (Hi5, a gift from Dr. Ignacio Moraga, University of Dundee) cells were set up at a density of $20x10^5$ cells mL^{-1} in medium (Insect Xpress, Lonza, supplemented with glutamine, penicillin and streptomycin). Virus (5% [v/v], P2) was added to the flask and the culture was left shaking at 135 rpm at 26.5 °C for 2 d.

2.3.2.3 Soluble protein purification

Cultures were subjected to two centrifugation steps to remove whole cells and cell debris as *Ac*AChBP and GBP are secreted into the medium. Firstly, at 1,140 g for 10 m at 4 °C, then secondly at 2,910 g for 10 m at 4 °C (Beckman J6-MC centrifuge with a JS4.2 rotor). The supernatant was concentrated, and

buffer exchanged into buffer A1 (Table 2.4) with a Satorius Stedim Sartojet using a 10 kD membrane cassette (PESU). The solution was then passed through a 0.8/0.2 µm filter (Acrodisc, Pall), to remove contaminants. The filtered solution was then purified by IMAC (5 mL HisTrap Ni-NTA column, GE Healthcare) with an ÄKTA Purifier system (GE Healthcare) using Unicorn software. The column was washed with buffer A1 and then the non-specific binding proteins were eluted off the column with a 10-20% gradient of buffer B1 (Table 2.4, 50 and 100 mM imidazole). The protein eluted at 50% buffer B1 (Table 2.4, 250 mM imidazole) and was fractionated. Samples of the fractions were analysed by SDS-PAGE. Appropriate fractions identified as containing the protein of interest were pooled and concentrated (Macrosep Advance device with 10 kD cutoff, Pall), then buffer exchanged into buffer A1 (Table 2.4). Measurements of T_i were recorded for concentrated samples using a NanoTemper Tycho to check protein quality. This measurement can quantify batch-to-batch variation (see details in nanoDSF section later).

2.3.2.4 Attempts to produce and purify membrane-integral proteins

For the production of membrane proteins, cultures of *Sf9* (50-500 mL) cells were set up at a density of $20x10^5$ cells mL^{-1} in medium (Sf900 III, Gibco, supplemented with glutamine, penicillin and streptomycin). Virus (1-5% [v/v], P2) was added to the flask and the culture was left shaking at 135 rpm at 26.5 °C for 1-3 d.

Cells were pelleted at 2,554 g for 15 m at 4 °C (Beckman J6-MC centrifuge with a JS4.2 rotor). Pellets were kept on ice and then resuspended in buffer A2 (Table 2.4) with the addition of protease inhibitors (10 µL, 100x, Expedeon, Abcam) and DNase (100 µL, 10 mg mL^{-1} in 0.15 M NaCl, Sigma). Cells were either sonicated (10 s on, 10 s off for 10 cycles, Soniprep 150, MSE) on ice or broken open by the use of a chilled cell disruptor (30 kpsi, Constant Systems). Intact cells were removed by centrifugation at 10,000 g for 15 m at 4 °C (Beckman Avanti J25 with a JA-25.50 rotor). The supernatant was centrifuged at 45,000 g for 1 h at 4 °C (Beckman Optima L-100K ultracentrifuge with a 70 Ti rotor) to pellet the membranes. The supernatant containing the soluble fraction was removed. Membrane pellets were kept on ice and resuspended in buffer A2 (Table 2.4) supplemented with detergent by manual homogenisation. A range of detergents were screened for each construct at or above their critical micelle content (1-6% [w/v]). The suspension was then agitated for 3-24 h at 4 °C. A final ultracentrifugation was conducted at 45,000 g for 1 h at 4 °C (Beckman Optima L-100K ultracentrifuge with a 70 Ti rotor) and the supernatant was collected. Samples were taken at various stages and analysed by SDS-PAGE and Western blot.

2.3.3 Attempts to produce recombinant protein in mammalian cells

2.3.3.1 HEK293T cell culture

All lentiviral work was conducted in a category two laminar flow cabinet. A modified protocol (150) was used throughout the attempted protein production. HEK293T cells were grown in medium A (DMEM/F-12 nutrient mix [Gibco, ThermoFisher] supplemented with 10% fetal bovine serum [FBS, ThermoFisher]) at 37 °C with 5% CO_2 for 3 d. Medium was then removed and the cells were washed with PBS (5 mL, ThermoFisher) followed by incubation with trypsin (1 mL, 1x, ThermoFisher) for 10 m. Cells were resuspended and then diluted with medium A (5 mL). Then the culture was split by the addition of medium A (5 mL) to cell suspension (1 mL). Cells were left to adhere and incubated at 37 °C with 5% CO_2 for 3 d. Cultures were split every 3 d.

2.3.3.2 Lentivirus production

A transfection mix was made up containing pTransfer (3.3 µg), pEnvelope (3.3 µg), pPackaging (3.3 µg) and medium B (100 µL DMEM/F-12 supplemented with 1% non-essential amino acids solution [Gibco, ThermoFisher]). A separate solution of polyethylenimine (PEI, 25 µL, 1 mg mL^{-1}, Aldrich) in medium B (100 µL) was prepared and mixed. The PEI and transfection solutions were mixed and incubated at RT for 20 m. Medium A was removed from the cells and replaced with medium C (4 mL, DMEM/F-12 supplemented with 2% FBS) then 0.2 mL of the DNA/PEI mixture was added dropwise. The culture was incubated at 37 °C with 5% CO_2 for 3 d, after which the medium, containing the lentivirus, was collected. The virus was supplemented with fresh medium A (2 mL) and filtered (0.45 µm, Minisart, Satorius). Polybrene (6 µL, 10 mg mL^{-1}, Merck) was added to increase transfection efficiency and the virus was frozen and stored at -20 °C.

2.3.3.3 Membrane-integral protein production

A new culture of HEK293T cells was grown to 90% confluency in medium A (6 mL). The medium was removed, and the cells were washed with PBS (5 mL) then this was replaced with medium containing the lentivirus (6 mL). The culture was incubated at 37 °C with 5% CO_2 for 3 d, then the medium was removed and the cells were washed with PBS (5 mL, 3x). The cells were trypsinised (1 mL, 10 m) then medium A (5 mL) was added. The cells were expanded by the addition of cell suspension (2 mL) to medium A (10 mL) per flask. These transfected cultures should have formed stable cell lines, continually producing the protein of interest. Stable cell lines are produced when the gene of interest has been incorporated into the host cell genome. Cultures were centrifuged at 500 g for 10 m at RT (Beckman Allegra X-12R centrifuge with a SX4750 rotor). Pellets were incubated at RT with SDS (1% [w/v]) and samples were analysed by SDS-PAGE gel electrophoresis and Western blotting.

Buffer	Contents
A1	50 mM Tris-HCl pH 7.5, 250 mM NaCl
A2	50 mM Tris-HCl pH 8, 150 mM NaCl
B1	50 mM Tris-HCl pH 7.5, 250 mM NaCl, 500 mM imidazole
TGS	25 mM Tris, 192 mM glycine, 0.1% (w/v) SDS

Table 2.4 Buffers.

2.4 Biophysical techniques

2.4.1 X-ray crystallography

2.4.1.1 Crystal screening

Commercial screens were used initially to find conditions that supported crystal growth. Various screens were used including JCSG+ (Molecular Dimensions), PEGS (Qiagen) and Morpheus (Molecular Dimensions and dispensed by MRC Reagents and Services, University of Dundee). JCSG+ and PEGS are examples of sparse matrix screens, which sample conditions used previously to crystallise proteins. These conditions have been followed up and verified (154). The Morpheus screen is based on commonly ordered ligands from structures deposited in the PDB, these are then used as additives (155). Fresh protein preparations of *Ac*AChBP and GBP were concentrated immediately before use. To form co-crystals, ligand (5-10 mM) was incubated with the protein (1-8 mg mL^{-1}) for at least 30 m at 4 °C, before plate setup. N-methylbicuculline and small molecules identified from the NMR screen were utilised in these studies. MRC 1, 2 and 3 drop plates (Molecular Dimensions) were used for the screens. Reservoirs (50 µL) were either pipetted by hand or using a Matrix Hydra (Thermo Scientific) microdispenser. Protein (100 nL) and reservoir (100 nL) drops were pipetted by either a Phoenix DT (ARI) or Mosquito Xtal3 (TTP LabTech) crystallisation robots. Plates were sealed with clear adhesive covers (Molecular Dimensions) and stored at RT. Plates were checked every few days for crystal growth.

2.4.1.2 Optimisation of conditions for crystal growth

Screening identified conditions that supported crystal growth. Optimisation of these conditions by variation of different parameters including pH, temperature and composition of the drops led to the formation of better crystals. Reservoir solutions were made in-house and pipetted into plates (24-well, VDX with sealant, Hampton) by a Scorpion (ARI) robot. Drops were pipetted by hand, further varying the composition of the drops (protein:reservoir: 1:1, 2:1, 1:2). Again, plates were incubated at RT and checked regularly for crystal growth. In some cases, drops were seeded with crushed crystals. Seeding provides a nucleation site for crystal formation, which could initiate or speed up crystal growth (156).

2.4.1.3 Data collection and processing

Co-crystallisation proved most useful to generate samples of protein-ligand complexes for analyses, resulting in crystals that diffracted to a higher resolution, compared to soaking. In other cases, *apo*

protein crystals were grown and ligands were soaked in, to increase throughput. Trials of the most suitable ligand concentration were conducted. Crystals were taken out of drops using crystallisation loops, cryo-protected before being plunged into liquid nitrogen to freeze. Generally, cryo-protectants were utilised, these included paratone N and mother liquor with the addition of glycerol. Diffraction data were collected in-house on a rotating anode generator (Rigaku M007HF generator with Varimax Cu-VHF optics, Saturn 944HG+ CCD detector and AFC-11 4-axis partial χ goniometer), or from synchrotrons (Diamond Light Source, European Synchrotron Radiation Facility and Synchrotron Soleil).

Data were indexed and integrated with Mosflm (157) or XDS (158), then merged and scaled with Aimless (159). Sometimes the data were automatically processed using the AutoPROC (160) or DIALS (161) pipelines. The Matthews coefficient to estimate solvent content and the potential number of subunits in the asymmetric unit was determined using RuppWeb (162,163). The structures were solved by molecular replacement with Phaser (164), exploiting previously determined structures. All the software, unless otherwise stated, is available from the CCP4 suite (165).

2.4.1.4 Refinement

Various structures have been refined and two of these are reported in this thesis, GBP:N-methylbicuculline and *Ac*AChBP:N-methylbicuculline. Structures were refined using Coot (166) and Refmac5 (167). Refmac uses prior chemical knowledge of bond lengths and angles, among other parameters to refine and restrain protein structures and B-factors. Various weighting schemes were utilised throughout the refinement. These schemes define how stringent the restraints are. In the graphic program Coot, the electron density ($2F_O$-F_C) and difference density (F_O-F_C) maps were visualised, where F_O are the observed structure factors and F_C are the calculated structure factors. The $2F_O$-F_C map shows the electron density to which each residue of the protein should fit. The F_O-F_C map shows the difference density, which is a guide to where electrons (or atoms) should be added or removed. Due to the pentameric quaternary structure of AChBP, non-crystallographic symmetry (NCS) restraints were applied initially. Firstly, the backbone was investigated and any residues that did not fit into the density were deleted and where possible remodelled. Attempts to extend the model into unoccupied density were carried out, using knowledge of the protein and crystallisation conditions, for example, the presence of glycosylation, ligands and ions. Geometrical restraints for ligands were generated by the Grade server (grade.globalphasing.org). The dictionary and cif file generated were used to model ligands in Coot. Ions, glycosylation and ligands were modelled into density considering the surrounding chemical environment and restraints. Water molecules were introduced when only small features were left in the density, considering the peak height, shape and proximity to polar groups using hydrogen bonding distance criteria of 2.5-3.5 Å. When differences between the chains became apparent the NCS restraints were removed. Ramachandran outliers, indication that different rotamers might be present, density fit

and geometric analyses were used to inform decision making during the refinement process. B-factors are an indicator of the degree of flexibility and error of an atomic position in the model (168). Therefore, these were also monitored as a guide to infer if areas of the structures required attention. Each round of map and model inspection with Coot were followed by cycles of refinement with Refmac, with monitoring of the R_{factor} and R_{free} values. The refinement process was stopped when no further improvements to model geometry could be made, when the difference between the initial and modified R_{factor}/R_{free} values did not change significantly and when there were no significant features left in the difference density maps. All structural figures were prepared using PyMol and annotated in Microsoft PowerPoint. For the table of crystallographic statistics see Appendix A.

2.4.2 Tryptophan fluorescence quenching assay

Protein samples (10 µg mL^{-1}, 2 mL) were prepared in buffer A1 (Table 2.4), along with stock solutions of the ligand to be tested. The ligand stocks were made at 2000x the final concentration so that 1 µL could be added to the protein solution and this would not cause drastic dilution effects. Experiments were run on an LS55 PerkinElmer spectrophotometer, using FL WinLab 4.00.03 software. The detector sensitivity was set to 750 V, exciting the sample at 280 nm and measuring emissions between 300-400 nm with slit widths of 5 nm. Blanks of water and buffer were first run before every experiment to check background fluorescence. An initial check of intrinsic ligand fluorescence was also conducted at the highest ligand concentration. Protein samples were analysed to give a baseline fluorescence before additions of ligand were made. After each addition, the solution was mixed and reanalysed. Experiments were stopped when no change in absorbance was observed. The maximum absorbance for the initial protein was taken, then each titration point was measured at this wavelength (335-339 nm). Control experiments (buffer with the same concentration of DMSO into protein) were deducted and the percentage change was calculated in Microsoft Excel. Percentage change measurements were input into GraphPad Prism 7, for plotting and calculation of K_D using a non-linear regression model for ligand binding.

Protein	Ligand	Stock concentrations (mM)	Number of additions
*Ac*AChBP	(-)-Nicotine	0.2, 1, 4	5, 3, 4
*Ac*AChBP	(+/-)-Anatoxin	0.2, 1, 4	5, 3, 4
*Ac*AChBP	(+/-)-Anatabine	4, 12	10, 10
*Ac*AChBP	(-)-Hosieine	0.05, 0.2	8, 8
*Ac*AChBP	Strychnine	0.2	18
*Ac*AChBP	N-methylbicuculline	1, 2	10, 5
*Ac*AChBP	1d	50, 100	8, 6
GBP	Strychnine	20	20
GBP	N-methylbicuculline	4, 8	10, 5
GBP	1d	50, 100	8, 6
GBP	3b	200	10

Table 2.5 WF experimental conditions. Concentrations and additions of various ligands used in tryptophan fluorescence quenching assays and the relevant target protein.

2.4.3 Isothermal titration calorimetry

ITC was used to investigate *Ac*AChBP and GBP interactions with selected ligands. Experiments were conducted on a Malvern MicroCal PEAQ-ITC instrument at 25 °C, using the manufacturer's software. Buffer A1 was made fresh, filtered and degassed before making up stocks. Table 2.6 provides selected experimental conditions. Protein samples (350 µL) were made up in excess to allow for syringe error and to remove air in the chamber. Ligand stocks (500 µL) were made up in buffer A1, in excess to allow for pipetting error, and so that the same concentration of ligand was used in each follow up ITC experiment. Both protein and ligand samples were left at RT to equilibrate. The sample chamber and the syringe were washed with detergent and water, then the syringe was rinsed with methanol before each experiment. A protein sample (300 µL) was loaded into the cell, making sure to displace the air and not introduce bubbles, which can cause artefacts in the ITC traces. Ligand (100 µL) was pipetted into a micro-eppendorf, then the instrument loaded this into the syringe, making sure to remove air bubbles. The reference was set to 10 µcal s^{-1}, stir speed was set to 750 rpm, then the concentration of protein, concentration of ligand, number of injections, volume of each injection and spacing was set, depending on the experiment. Appropriate control experiments (buffer-buffer, buffer-protein and ligand-buffer) were carried out. The parameters K_A, N and ΔH were measured, then the K_D, change in Gibbs free energy (ΔG) and change in entropy (ΔS) were calculated. Mean data and standard errors are reported (n=3).

The change in heat energy (ΔQ) is shown as the area above/below the peaks, in the raw titration isotherm. In the software this energy is adjusted for moles of injectant, giving a ΔH value, in kcal mol^{-1}. The concentrations of ligand and protein per injection are also calculated, with knowledge of the exact cell volume and user set parameters. This considers dilution effects per injection. The molar ratio of ligand to protein is then calculated, per injection. The binding graph plots ΔH against the molar ratio, then the binding curve is modelled for the data, depending on the model selected (169).

Protein	[Protein] (µM)	Ligand	[Ligand] (µM)	Number	Injection Volume (µL)	Spacing (s)
*Ac*AChBP	25	(-)-Nicotine	250	18	2	180
*Ac*AChBP	25	(+/-)-Anatoxin	125	18	2	180
*Ac*AChBP	25	(+/-)-Anatabine	350	18	2	180
*Ac*AChBP	25	(-)-Hosieine	150	19	2	150
*Ac*AChBP	10	Strychnine	100	18	2	180
*Ac*AChBP	40	N-methylbicuculline	2000	13	3	180
GBP	25	Strychnine	500	18	2	180
GBP	40	N-methylbicuculline	2000	13	3	180

Table 2.6 ITC experimental conditions, ligand and protein concentrations used in experiments.

2.4.4 Surface plasmon resonance

SPR was used to investigate the association of *Ac*AChBP with (-)-nicotine, (+/-)-anatoxin and (+/-)-anatabine. Experiments were conducted on a Biacore T100 with T200 sensitivity upgrade and analysis software (GE Healthcare). A sensor chip (Ni-NTA series S, GE Healthcare) was loaded into the instrument and equilibrated with water then buffer A1. The chip was prepared by washing it with EDTA, buffer A1, $NiCl_2$ and buffer A1 again, then NHS (N-hydroxysuccinimide) and EDC (1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride) were mixed to prepare the sensor surface to further immobilise the protein and prevent drift during the experiment. NHS and EDC are used for the secondary coupling of protein to the sensor chip surface. *Ac*AChBP was loaded onto channel two and GBP was loaded onto channel four reaching a response level of 3,500 RU. Starting stocks of ligands were prepared in a 96 well plate then a two- or three-fold serial dilution was applied to give eight concentrations. Buffer A1 was also pipetted into some wells for start-up blank controls. DMSO (50% [v/v]) was used to remove the chance of ligand transfer between concentrations. If a ligand is observed as transferring after washing in DMSO it is classified as 'sticky', therefore may affect the interpretation of the results. To remove the need for a solvent correction, buffer A1 was supplemented with DMSO to the same concentration as in the 96 well plate. Sensorgrams were produced showing the effect of each concentration on the response over time. The end of each trace was taken to be steady-state equilibrium and recorded. These measurements were then plotted on a graph of response versus concentration and a curve of best fit was modelled. The K_D was read as half of the R_{max}. Due to the speed of association and dissociation, the kinetics of each binding event could not be measured. Further optimisation was underway to resolve these parameters, but due to time constraints, this was not completed.

Detection, in Biacore instruments, of ligand binding is via the SPR/resonance angle alteration. This change is converted into response units and is correlated with increasing ligand concentration. A raw sensorgram shows 1) the association curve, when molecules are binding to the protein continuously, 2) steady-state equilibrium when the association and dissociation rates are equal and 3) the dissociation curve, the removal of the ligand from the protein until the baseline is reached (170). For some molecules, the rate of association and dissociation are so fast that curves cannot be seen. In these instances, only steady-state equilibrium can be measured and the affinity calculated. The response at steady-state is measured and plotted against concentration, then a binding curve is modelled (171). The affinity is calculated for steady-state equilibrium, by the following equations.

Eq. 2.1 $R_{eq} = C\,R_{max}/K_D + C$, when $C = K_D$, $R_{eq} = R_{max}/2$

Protein	Ligand	Stock [Ligand] (µM)	Dilution factor	Number of dilutions
*Ac*AChBP	(-)-Nicotine	30	3	8
*Ac*AChBP	(+/-)-Anatoxin	60	3	8
*Ac*AChBP	(+/-)-Anatabine	100	3	8

Table 2.7 SPR experimental conditions. Stock ligand concentrations and dilutions for experiments.

2.4.5 Nano differential scanning fluorimetry

For analysis of *Ac*AChBP and GBP, solutions of protein (0.5 mg mL^{-1}, 20 µL) were diluted in buffer A1 and loaded into capillaries. The capillaries were then placed on the sample tray of the NanoTemper Tycho NT.6. The instrument checked for the presence of capillaries, calculated initial ratios and ramped the machine to the starting temperature of 35 °C. The fluorescence was monitored as the temperature increases from 35-95 °C, at a ramp rate of 30 °C m^{-1}. Upon protein unfolding, a sharp increase in fluorescence was observed, which eventually plateaus. The inflection temperature was measured at 50% of the fluorescence increase. For GBP, a T_i of ~89 °C should be reached. For *Ac*AChBP, a T_i should not be measured as it is greater than 95 °C, above the detection threshold of the instrument. Predicted values have been determined by previous nanoDSF and DSF experiments.

DMSO is used as a solvent for many of the ligands being studied. Since this molecule is known to complicate ligand binding studies, due to action as a chaotropic agent (172), the influence of DMSO on GBP stability was investigated. Various solutions (20 µL) of GBP (0.5 mg mL^{-1}) and a range of DMSO concentrations (0-10%) were set up. Screening identified that a DMSO concentration above 3% had a negative impact on GBP stability. A time course of 0-45 m, using 2% DMSO saw no change in protein stability. Therefore, in subsequent experiments the conditions were controlled to remain within this limit.

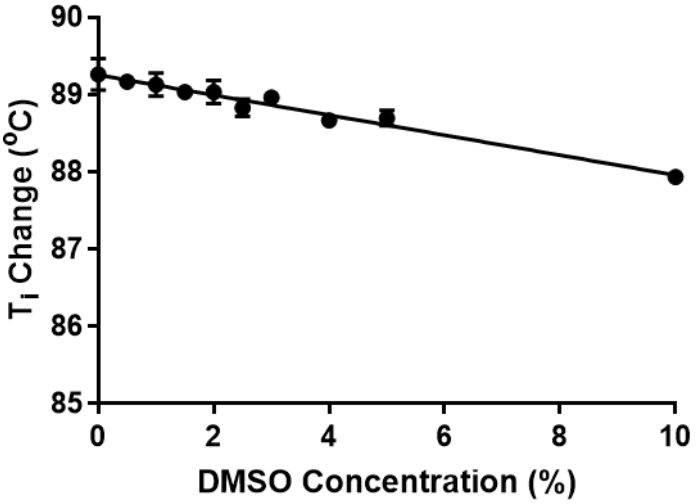


Figure 2.3 Effect of DMSO concentration (%) on GBP stability. Stability measured by a decreasing inflection temperature (T_i) by nanoDSF on a NanoTemper Tycho. The starting T_i for GBP is 89.1 °C and the low [DMSO] does not affect this dramatically. After 2% DMSO, the protein starts to destabilise, until 10% where a change of 1.2 °C is observed. A change of 0.5 °C is considered significant for this technique (173).

2.4.6 Biolayer interferometry

*Ac*AChBP or GBP samples (0.5 mg mL^{-1}, 300 µL) were set up in a 96 well plate with two wells of buffer (300 µL). The number of sensors used determines how many rows of protein and buffer are needed (an example is shown in Figure 2.4). Ligand stock solutions (400 µL) were made up and aliquoted into the end well of a fresh 96 well plate. A serial dilution series, over the previous five wells, was conducted. Each dilution was into buffer A1. The first six wells were filled with buffer A1 (300 µL), an example is shown in Figure 2.4. Biosensors (Ni-NTA, Fortebio) were incubated in buffer A1 for 10 m.

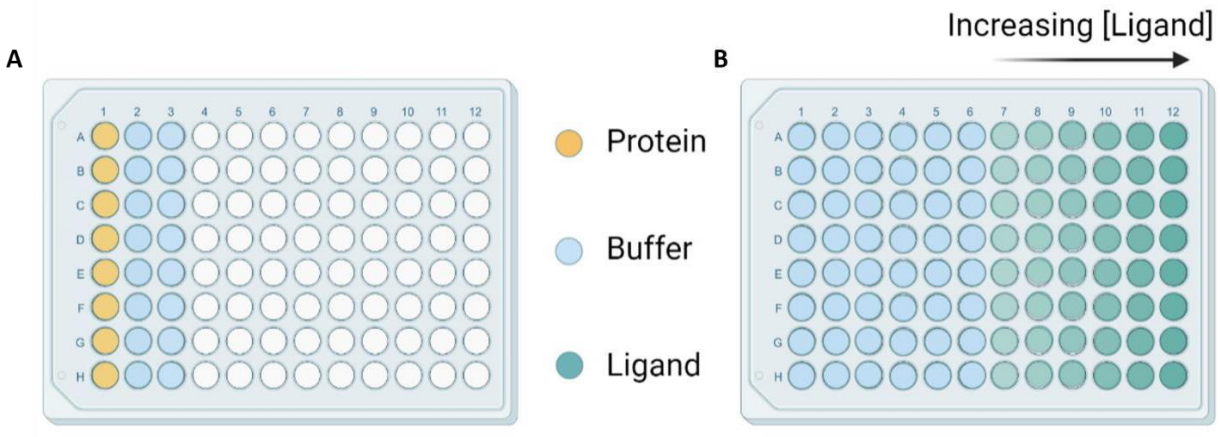


Figure 2.4 Example setup for BLI. A) The setup of the reagent plate for protein loading. Protein (orange, 0.5 mg mL^{-1}, 300 µL) and buffer A (blue, 300 µL) is pipetted into each labelled well. B) The setup of the sample plate for ligand binding. Buffer A (blue, 200 µL) is pipetted into the appropriate wells. The highest concentration of ligand (green) is made and pipetted into the last well (300 µL). For a 1 in 3 dilution, buffer A (200 µL) is pipetted into the other ligand wells (green), then 100 µL is transferred from the highest concentration into the next well and mixed. This is continued for the remaining four wells. Repeats are conducted in the other rows, three per ligand.

An Octet red (ForteBio) instrument was used. The sensors in buffer, the sample plate (ligands) and reagent plate (protein) were loaded and the temperature set to 25 °C. For protein loading, steps of baseline (60 s) in buffer, loading (600 s) in protein solutions and wash (600 s) in buffer were conducted. For the ligand binding assay, steps of baseline (60 s) in buffer, association (120 s) in ligand solution and dissociation (120 s) in buffer were included. The sensors started at the lowest ligand concentration and worked up to the highest concentration. The raw traces were analysed to remove artefacts before subtraction of control data. Controls were two-fold, baseline into buffer and non-specific binding to another protein (TEV protease), either both or the former controls were used. The data were then processed using a global fit method and 1:1 interaction modelling.

2.4.7 Nuclear magnetic resonance

For the NMR screen, GBP and *Ac*AChBP (500 µL, 10 µM) in buffer A1 supplemented with D_2O (10% [v/v]) for locking, were used. All NMR experiments were performed on a Bruker AVANCE III HD spectrometer operating at 500 MHz and fitted with a quadruple resonance cryo-inverse probe (QCIP). A predetermined

cocktail of 9-10 chemical species was added to the protein solutions; a total of 435 compounds. ^{19}F spectra were acquired in attempts to detect ligand binding. Strychnine (0.5 mM) was added in DMSO to the protein solution to displace bound ligands. ^{19}F spectra were reacquired. Spectra were analysed with TopSpin 4.0.1 (Bruker). Analysis of the two spectra led to the identification of potential ligands.

2.4.8 Differential scanning fluorimetry

Measurements on *Ac*AChBP and GBP were conducted on a Stratagene MX3005P QPCR machine (Agilent). Master solutions of protein (10 µM), Sypro orange dye (1:1000) and the ligand (varying concentrations) were made up. Samples (40 µL) were run in triplicate, with negative (no protein) and positive (protein but no ligand) controls. Data were collected utilising a Sypro appropriate filter (excitation: 490 nm, emission: 610 nm). The data were then analysed with Stratagene software, Microsoft Excel and GraphPad Prism 7.

2.5 Computational tools

2.5.1 AreaImol

The software is available through the CCP4 suite and was used for solvent accessible surface area (SASA) calculations. The structure files *Ac*AChBP:(-)-nicotine (PDB code: 5O87), *Ac*AChBP:(+)-anatoxin (PDB code: 6SH0), *Ac*AChBP:(-)-hosieine (PDB code: 6SGV), *Ac*AChBP:strychnine (PDB code: 5O8T), GBP:strychnine (PDB code: 5OBG), GBP:N-methybicuculline (PDB code: 5OBH) were downloaded from the PDB and inspected in Coot. AreaImol (174) was utilised to calculate the total SASA of the ligand in isolation and remaining SASA after binding. From this the percentage SASA lost upon binding was calculated. Due to the pentameric nature of the proteins, 5-10 copies of the ligand were utilised in calculations then the mean and standard error were reported.

2.5.2 Molsoft ICM

Molsoft ICM-Pro + VLS (v3.8-4a) (175,176) was used for docking simulations of various ligands. Structure files from the PDB were used and the core components of a single orthosteric binding pocket were isolated. PDB structural complexes used were: *Ac*AChBP:(-)-nicotine (PDB code: 5O87), *Ac*AChBP:(+)-anatoxin (PDB code: 6SH0), *Ac*AChBP:strychnine (PDB code: 5O8T), *Ac*AChBP:*apo* (PDB code: 2Y7Y), GBP:glycine (PDB code: 5OAN), GBP:strychnine (PDB code: 5OBG) and GBP:N-methylbicuculline (PDB code: 5OBH). In most cases, two interacting chains of the protein, the ligand and key water molecules were kept. The PDB file was converted into an ICM object. Ligands were moved out of the binding site so pockets could be identified using the 3D predict ICM Pocket Finder. The pockets identified were the orthosteric binding site defined by the aromatic cage. Using the protein object (with any key water

molecules present) and the previously identified pocket for the orthosteric binding site, receptor maps were generated. These are required for the docking of molecules to the binding site. Definition of key binding residues, the addition of hydrogen atoms, where possible, and calculation of charges were conducted. In the docking simulation, the receptor is treated as a rigid object, whereas the ligand is allowed to be flexible. The docking software screens all the potential conformations of a ligand and selects the most favourable, in terms of energy. These conformations are then docked to the pocket, ranked and the lowest energy conformation is reported. Ligands and libraries were screened against the pocket. For screening, the Maybridge "Ro3 fragment diversity set" of 2,000 compounds was downloaded (maybridge.com). The docking effort for single ligands and small sets (from NMR screening) was set to 10, however, this was decreased to 3 for docking of the Maybridge set. The thoroughness is the length of the simulation run and a setting of 1 is recommended for most scenarios (177). The use of a higher thoroughness means that the final docked conformation is more likely to be correct. Racemic sampling was conducted for compounds where appropriate.

Parameter	Description
ICM score	The ranking for a compound depending on how well it is predicted to bind; the lower the score the better
VLS score	Similar to the ICM score but accounts for the energy of ligand strain
Hbond	Predicted hydrogen bonding energy
Hphob	Predicted hydrophobic energy using the solvent accessible surface area
VDW	Predicted van der Waals energy, which considers interactions (including repulsion) of atoms over distance
RotB	Number of rotatable bonds
Dsolv	Predicted energy for the desolvation of hydrogen bond donors and acceptors

Table 2.8 ICM parameters and descriptions.

Important parameters reported from the docking simulation are shown in Table 2.8 with brief descriptions. The docking software reports hydrogen bond distances as the separation between the hydrogen and the acceptor group. The ICM (Internal Coordinate Modelling) score takes into account various weighted terms including, internal torsion, van der Waals, electrostatics, hydrogen bonding and hydrophobic interactions (178).

The 4D docking protocol was utilised to inform a decision as to which protein structures should be used to dock ligands and small molecules. 4D docking is used when conformational differences occur between different structures, for example, the closing of loop C in *Ac*AChBP. The 4D protocol makes a stack of different structures, with alterations in the binding site. The software then makes maps for each and trials docking of the ligands into each member of the stack. In the output, the identity of the structure that each ligand binds to is highlighted. In the case of the molecules tested this was the same for all of

them, therefore a basic docking simulation was conducted. For GBP, the complex with strychnine (PDB code: 5OBG) was used, whereas for *Ac*AChBP both the *apo* (PDB code: 2Y7Y) and nicotine complex (PDB code: 5O87) structures were utilised.

Images were generated using Molsoft ICM then exported and edited in Microsoft PowerPoint. One key point to note is how hydrogen bonds are represented in Molsoft figures. Hydrogen bonds are shown as coloured spheres, the colour indicates the length, long to short (red-blue) and the size indicates the strength, strong-weak (thick-thin).

When making alterations to the bound small molecules, for example when extending a molecule or testing if a group was necessary for binding, a different docking simulation was run. For trialling and removal of different groups, the specific chemical entity was changed and fitted into the pocket. This led to a virtual ligand screening (VLS) score, which rated if the change either improved binding or not. This new score was used as it also considered the effect of the additions on ligand strain. After selection of the relevant groups (Figure 2.5) and addition to the scaffold, the small molecules were redocked to the binding site. Different conformations of the modified molecule were calculated then a subset was screened against the pocket. Redocking was repeated until the VLS score did not change and a similar orientation of the compound was adopted. After these parameters had been met, the VLS score was recorded and compared to the original molecule.

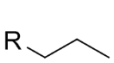
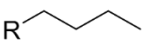
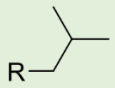
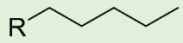
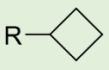
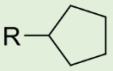
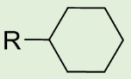
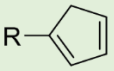
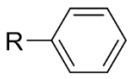
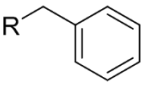
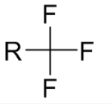
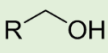
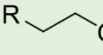
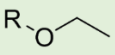
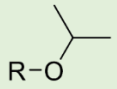
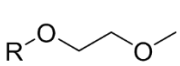
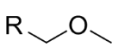
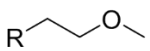
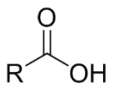
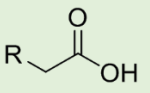
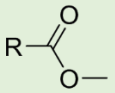
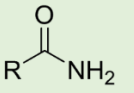
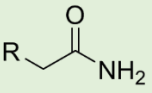
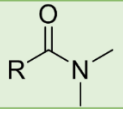
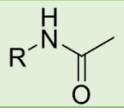
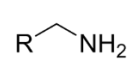
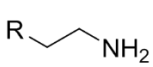
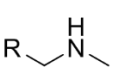
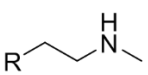
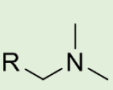
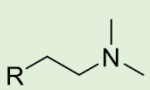
R—					
					
		R–F			R–Cl
R–OH					
					
					
$R–NH_2$					
					

Figure 2.5 Chemical groups extensions screen. The selection of chemical groups was screened when extending small molecule hits in Chapter 5.

2.5.3 Ligplot

The software (179) was downloaded from EMBL-EBI (https://www.ebi.ac.uk/thornton-srv/software/LigPlus/) and was used to make simplified interaction diagrams of structures from the PDB.

2.5.4 BioRender

Some figures included in Chapter 1 and 2 were made with the aid of BioRender (available through BioRender.com).

2.6 Confocal microscopy

GFP constructs were prepared and produced in insect cells, as detailed previously. Test and control cultures were set up. Samples were taken from cultures at 24 h time points. The samples were centrifuged at 1,000 g for 10 m at 4 °C (Beckman Allegra X-12R centrifuge with a SX4750 rotor) and the supernatant was discarded. Pellets were kept at 4 °C until they were required. Samples were resuspended in PBS and then imaged on a Zeiss LSM880 confocal microscope using the x63 Pan Apochromat objective (NA 1.4) by Dr. Alan Prescott (Microscopy Unit, University of Dundee). Excitation was with a 488 laser and emission collected from 493-598 nm.

3. Assessment of three biophysical techniques used for screening potential ligands for pLGICs

3.1 Introduction

A ligand screening campaign can serve two purposes. Firstly, it may detect binders and rule out non-binders. Secondly, it can be utilised to estimate affinity, so that follow up compounds can be prioritised for further study. The affinity of a compound for a target can be defined by the dissociation equilibrium constant (K_D). The constant represents the ratio between unbound and bound protein and ligand, when equilibrium is reached, in terms of concentration (180). Here, K_D will be utilised to assess three biophysical techniques: WF, ITC and SPR. The viability of the techniques to screen and characterise ligand binding with the pLGIC surrogates, *Ac*AChBP and GBP, will be evaluated to identify new chemical matter that can modulate these receptors.

The above techniques employ the measurement of different physical properties to study and quantify ligand binding. Excited electrons returning to a low energy state can emit fluorescence and this can be pronounced for the aromatic acid tryptophan (112,113,181). In our target proteins, a key tryptophan in the binding site should directly interact with ligands, resulting in a change to the environment of the Trp electronic structure, which can change the fluorescence associated with an electronic transition. Tyr and Phe residues also make a small contribution to the overall fluorescence signal (182). The physical property measured is a change in fluorescence parameters upon ligand binding (112,113,181).

In an ITC experiment two cells, one containing the sample and one a standard, are maintained at a constant temperature. The energy used to maintain the constant temperature is measured. When a ligand binds to a protein, in the sample cell, it generally results in either an exothermic or endothermic interaction and so causes a change in the temperature of the chamber. The change in power required to maintain a constant temperature in the cell is recorded. This change is proportional to the enthalpy change when the ligand binds to the protein. From this, the K_D, as well as a full thermodynamic profile of binding can be determined (120,183,184).

SPR measures the change in light wavelength or reflection angle, dependent on the instrument used. A change in either is caused by a physical property of the metal sensor surface, known as surface plasmon resonance. The outer electrons of the metal surface interact with the photons of light and are excited,

turning into surface plasmons. These electrons are sensitive to their environment, therefore a change in density on the chip surface can alter the local index of refraction, which changes the resonance condition of the surface plasmons. This change leads to variations in reflected light and detection (102,103,185).

3.2 Data from three biophysical assays

Three compounds: (-)-nicotine, (+)-anatoxin and (-)-anatabine are ligands of the nAChR that elicit a physiological response. Nicotine is present in cigarettes and linked to addiction by the modulation of nAChR (186). The (-)-enantiomer is the predominant naturally occurring and active form. Anatabine is a minor alkaloid present in tobacco with the (-)-enantiomer being the prevalent form (187). Anatoxin is a potent neurotoxin, which is produced by algae, its alternative name is 'rapid death factor' (188). The *Ac*AChBP surrogate provides a small, extremely stable platform to test the binding of some nicotinic receptor ligands (for more detail, see Chapter 1.1.3) (34). The stability of *Ac*AChBP has been assessed previously using DSF and nanoDSF with a T_m/T_i >95 °C, above the threshold of the instruments used. The methods described above were used to investigate the binding of these three ligands to *Ac*AChBP.

The relevant enantiomer (-)-nicotine was used, however, the commercially available anatoxin and anatabine samples are mixtures. Docking simulations suggest that both enantiomers of nicotine should bind similarly (ICM scores -26.3 to -28.5), however, (-)-nicotine is produced in excess naturally (187). Various techniques are available for separation of mixtures, for instance, high-performance liquid chromatography (HPLC) (189). However, trying to separate the enantiomers requires extensive optimisation of chiral selectors, column type, temperature and solvent contents (190,191). The use of racemic mixtures can be a problem when deducing the affinity of a compound for a target. This is because both enantiomers have the opportunity to interact with the protein, therefore contributing to or interfering with the parameters deduced. The extent to which they would interfere is dependent on their concentration and relative affinities.

A 50:50 mixture was present for (+/-)-anatoxin. Previous data suggest that the naturally occurring enantiomer, (+)-anatoxin is more active (K_i 0.09 µM with nAChR α_7) than the (-)-enantiomer (K_i 21 µM with nAChR α_7) (192,193). To investigate this further both enantiomers were docked to the *Ac*AChBP:(+)-anatoxin structure (PDB code: 6SH0). Docking simulations showed both enantiomers occupy a similar position in the binding pocket (Figure 3.1). However, the secondary amine is positioned differently (2 Å difference between the nitrogen atoms). The two predicted parameters, hydrogen bonding (-3.6 to

-4.2 kcal mol^{-1}) and hydrophobic interaction (-4.4 kcal mol^{-1}) energies are similar between the two enantiomers. Nonetheless, due to the change in position of the amine group, the van der Waals energy for (+)-anatoxin (-24.9 kcal mol^{-1}) is improved by 3.7 kcal mol^{-1} relative to (-)-anatoxin (-21.2 kcal mol^{-1}). Similarly, there is also a difference with the ICM scores with (+)-anatoxin scoring better (-28.3) than (-)-anatoxin (-24.6). A difference in energy and score may have occurred owing to the π-cation interaction that (+)-anatoxin makes with W164, which does not occur for (-)-anatoxin. Therefore, as noted in Table 3.1 and Figure 3.2-3.4 legends, the concentration for one enantiomer was used in analyses for (+)-anatoxin dependent experiments.

We assume that the commercial (+/-)-anatabine is a 50:50 enantiomeric mixture. Both enantiomers have similar potency (EC_{50} values) with the nAChR α_7 receptor, 50 µM for (+) and 70 µM for (-) (194). The separation of enantiomers was carried out by chiral-phase HPLC (194). Again, both enantiomers were docked to investigate if there were any differences in binding (Figure 3.1). These compounds were docked to the *Ac*AChBP:nicotine structure (75) due to chemical similarities between the ligands. There is minimal variation between the ICM scores (-22.5 to -25.4) and the three theoretical interaction energies, hydrogen bonding (-3.1 to -4.0 kcal mol^{-1}), hydrophobic (-4.6 kcal mol^{-1}) and van der Waals (-22.1 to -22.6 kcal mol^{-1}). The pyridine ring of (+)- and (-)-anatabine occupy identical positions forming a hydrogen bond with an important binding site water molecule that will be discussed later (Chapter 3.4.2). The secondary amine occupies a different position in the two forms, therefore the hydrogen atoms are also shifted (0.7 Å difference between the nitrogen atoms). Due to this difference, the (-)-enantiomer is predicted to make a main chain (MC) hydrogen bond with W164, which is not present for the (+)-enantiomer. The (+)-enantiomer may form other weaker hydrogen bonds which could make up for this loss of interaction. For the purpose of this study it was considered that both enantiomers may interact with *Ac*AChBP similarly, therefore a single concentration for both (+/-) was used in experiments.

Ligand	N (sites)	ΔG (kcal mol^{-1})	ΔH (kcal mol^{-1})	-TΔS (kcal mol^{-1})	K_{D1} (µM)	K_{D2} (µM)	K_{D3} (µM)
(-)-Nicotine	0.37 ± 0.03	-7.7 ± 0.1	-17.8 ± 0.9	10.0 ± 0.9	2.3 ± 0.2	0.41 ± 0.01	0.8 ± 0.1
(+)-Anatoxin	0.45 ± 0.02	-8.9 ± 0.1	-12.5 ± 0.2	3.6 ± 0.2	0.30 ± 0.03	0.20 ± 0.01	0.12 ± 0.02
(+/-)-Anatabine	0.4 ± 0.1	-7.5 ± 0.1	-4.5 ± 0.5	-2.9 ± 0.6	3.4 ± 0.5	8.5 ± 0.2	9.1 ± 1.3

Table 3.1 Binding data for the interaction of *Ac*AChBP with three ligands, (-)-nicotine, (+)-anatoxin and (+/-)-anatabine. Mean thermodynamic parameters are derived from ITC as well as dissociation constants: K_{D1} from ITC, K_{D2} determined from fluorescence measurements, K_{D3} from SPR. The standard error for each measurement is given (n=3). (+/-)-Anatoxin has a more active (+)-enantiomer determined from literature

data and comparison of the binding pose in docking simulations. Therefore, the concentration of the (+)-enantiomer was assumed at 50% of the mixture. (+/-)-Anatabine, is a mixture of two enantiomers with similar binding affinities, therefore the individual K_D values cannot be determined. The K_D values for this compound are an average of the two enantiomers.

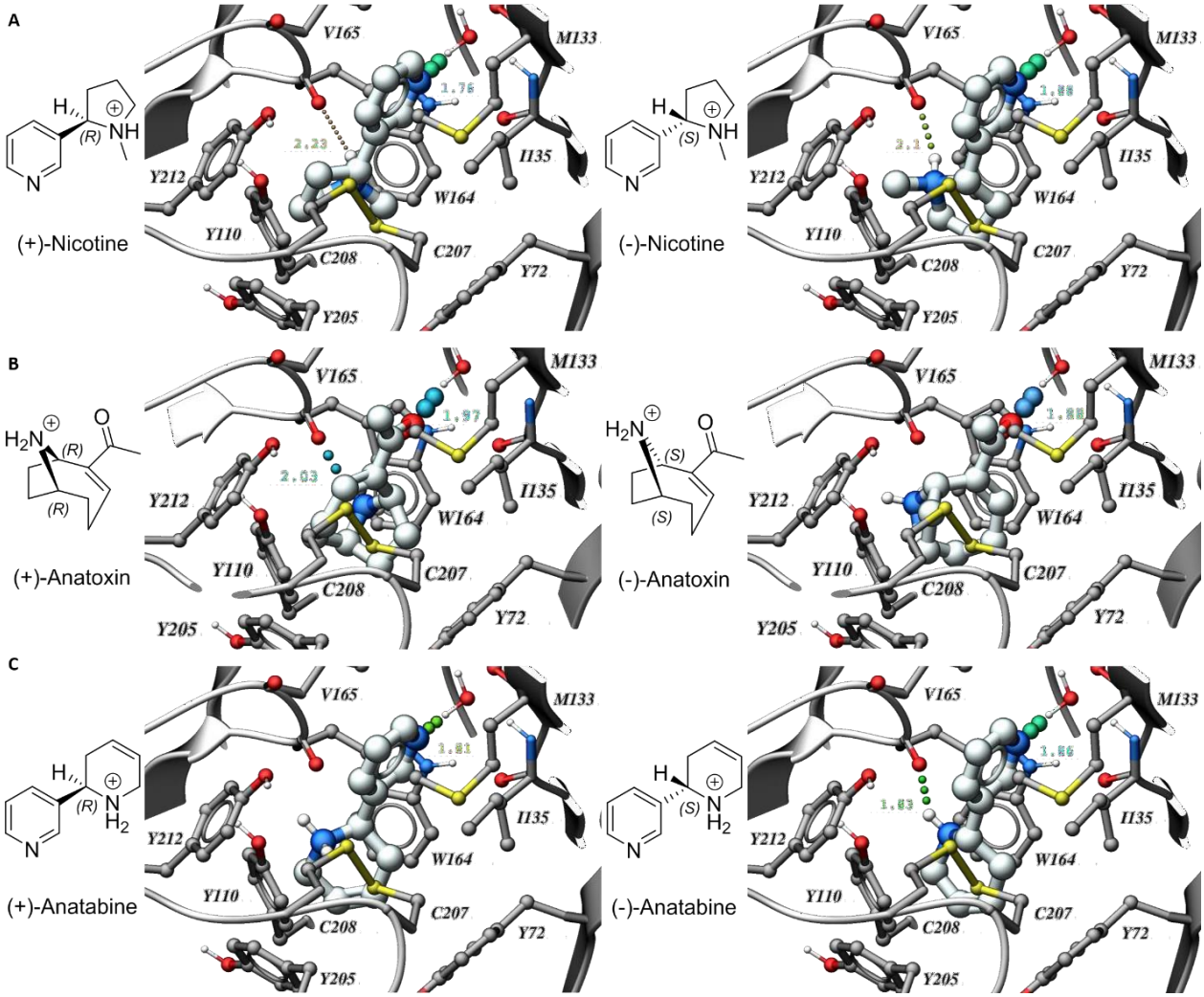


Figure 3.1 Models of both enantiomers of three ligands of interest bound to *Ac*AChBP. Parts of the protein backbone are shown as ribbons (light grey), important residues as sticks (grey), water molecules as spheres (red) and ligands as thick sticks (silver). Different atoms are coloured: oxygen (red), nitrogen (blue), sulfur (yellow) and hydrogen (white). Hydrogen bonding is shown (spheres), the colour represents long to short bond lengths (red-blue) and the sphere size shows strong to weak bonds (thick-thin). Residues are labelled in black and hydrogen bonds are labelled with their respective colours. The structure of each ligand is shown next to each image. (A) (+/-)-Nicotine docked into 5O87, (B) (+/-)-anatoxin docked into 6SH0 and (C) (+/-)-anatabine docked into 5O87.

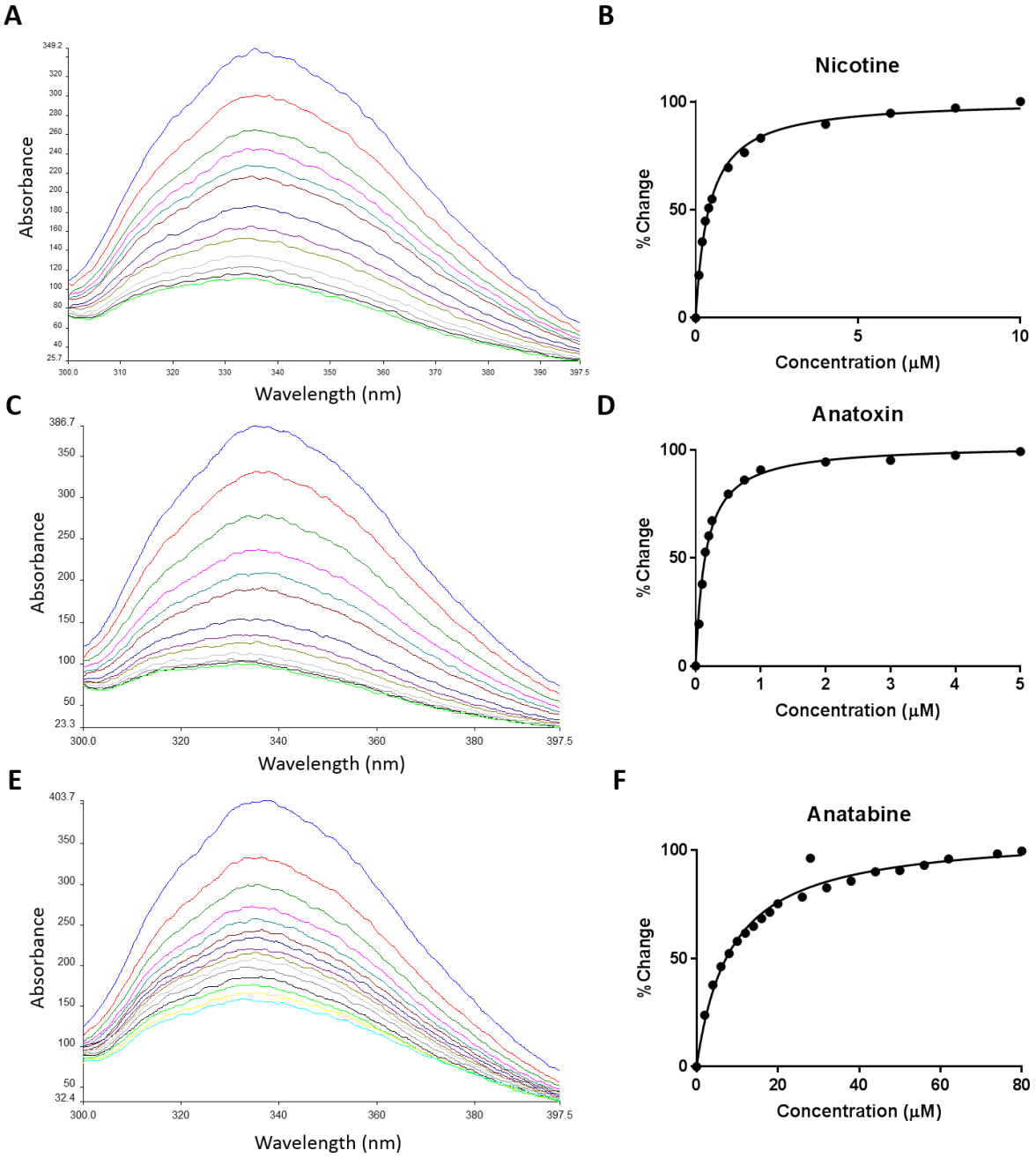


Figure 3.2 Ligand binding to *Ac*AChBP monitored by WF. Example traces obtained from titrations, starting with no ligand (blue) to maximum ligand concentration (cyan) and a plot of the mean percentage change versus concentration is shown for (-)-nicotine (A, B), (+)-anatoxin (C, D), and (+/-)-anatabine (E, F). Standard error bars are behind each data point (n=3). Due to the (+)-anatoxin enantiomer being more active, the experimental concentration used accounts for this. (+/-)-Anatabine is a mixture of two enantiomers with similar affinities.

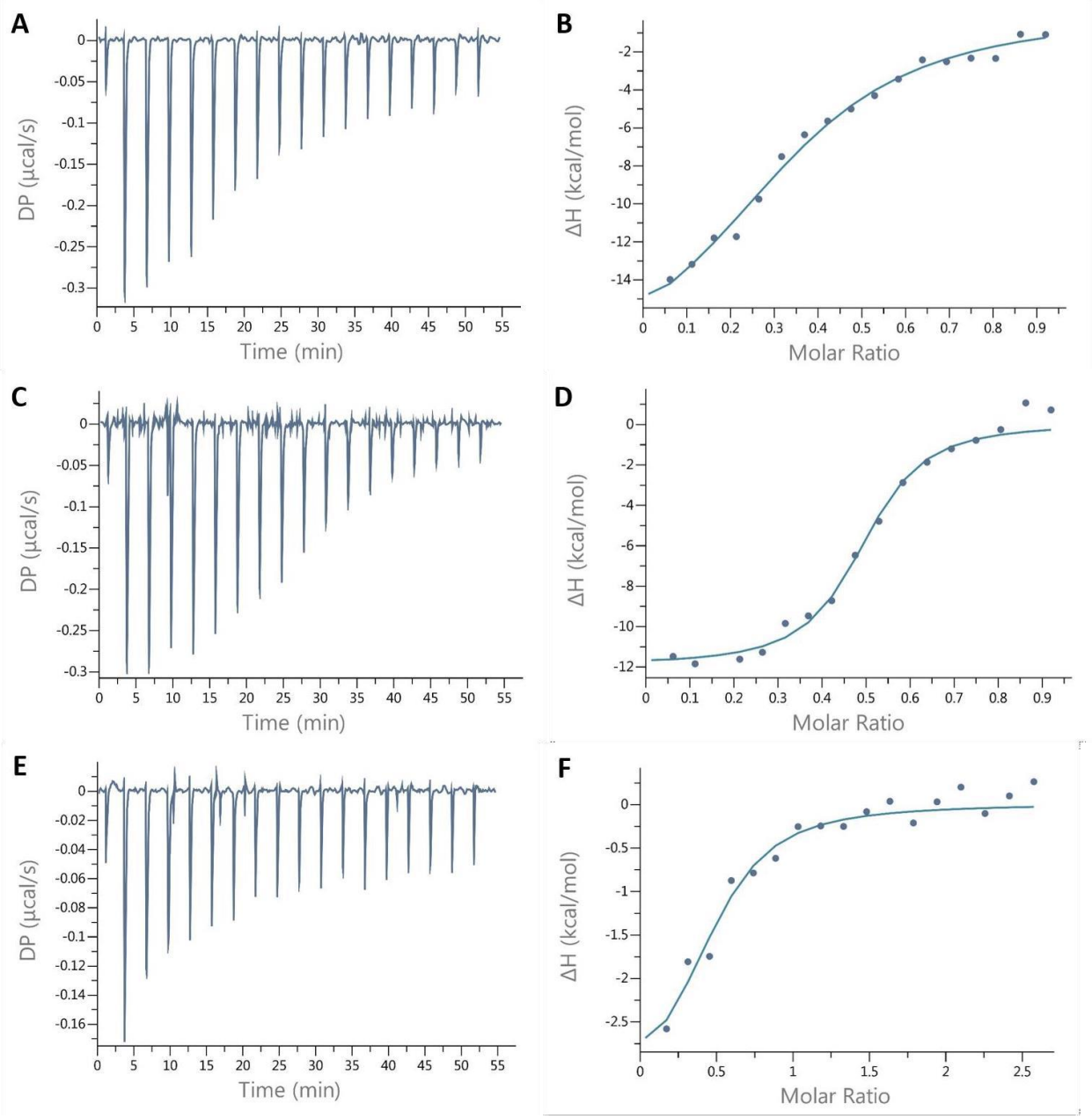


Figure 3.3 ITC data for *Ac*AChBP binding various ligands. Example isotherms and curves with the best fit are shown for (-)-nicotine (A, B), (+)-anatoxin (C, D) and (+/-)-anatabine (E, F), respectively. In the examples, a baseline deduction has been applied. All data have had the relevant controls (buffer-buffer, buffer-*Ac*AChBP, ligand-buffer) also deducted. Due to the (+)-anatoxin enantiomer being more active, the experimental concentration used accounts for this. (+/-)-Anatabine is a mixture of two enantiomers with similar affinities.

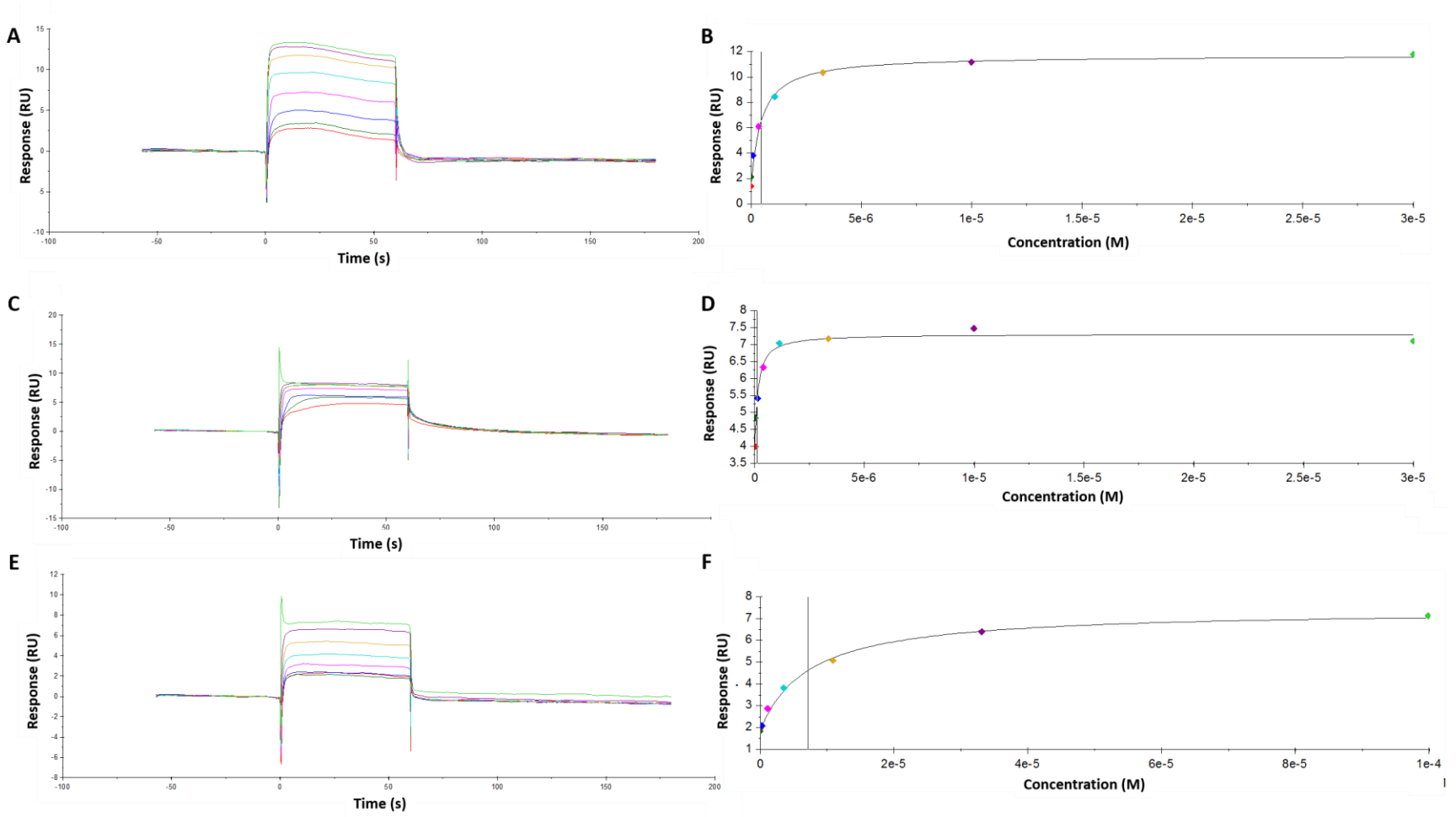


Figure 3.4 SPR data for *Ac*AChBP binding selected ligands. A sensorgram from an example titration and curve of best fit are shown for (-)-nicotine (A, B), (+)-anatoxin (C, D) and (+/-)-anatabine (E, F), respectively. The response value was taken at steady state, at the end of each trace for each concentration. The colour of each point relates to the sensorgram traces. The vertical black line indicates the point at which the K_D is measured. Due to the (+)-anatoxin enantiomer being more active, the experimental concentration used accounts for this. (+/-)-Anatabine is a mixture of two enantiomers with similar affinities, therefore the concentration of each is not considered.

3.3 Comparison of ligand affinity measured by different techniques

To investigate the ligand affinities of the three compounds selected as chemical probes, a WF assay was first utilised. A protocol for the assay had been established in the laboratory by Dr. Alice Dawson. The use of fluorescence measurements to monitor ligand binding takes advantage of a key residue (W164) in the binding site (for more detail, see Chapter 1.3.2 and 2.4.2). All compounds were initially tested for intrinsic fluorescence before a quenching assay was performed. (-)-Nicotine, an archetypal ligand of nAChRs, gave a K_D value of 0.41 ± 0.01 µM with *Ac*AChBP (Table 3.1). This value is consistent with others in the literature (K_D 0.25 µM by WF and 0.05 µM by ITC, (35,66)). Differences between the literature and measured values could have occurred due to differences in experimental setup for example buffers or equipment used to measure the fluorescent signal. (+)-Anatoxin gave a comparable K_D (0.20 ± 0.01 µM), reflecting its high affinity for the family of receptors (193). However, with less research conducted on this molecule, there are no literature values for comparison with *Ac*AChBP. Nonetheless, there is a K_D for the interaction with Torpedo nAChR of 0.1-0.2 µM measured by competitive inhibition assays, which is comparable (195). For (+/-)-anatabine, a lower affinity value was observed (8.5 ± 0.2 µM), again no literature K_D values are available for direct comparison. Other literature parameters will be discussed later (Chapter 3.6).

To elucidate the different K_D values for enantiomers present in a mixture, which could have given a combined affinity in tryptophan fluorescence, ITC was conducted. Fokkens and Klebe demonstrated that determination of individual enantiomeric binding affinity by ITC was possible for mixtures before resolution (196). This technique has also been used to calculate the affinity of different enantiomeric mixtures of serine protease inhibitors. For some examples two binding events could be seen, however, in other cases, it was more ambiguous. Plots from this example show two distinct binding curves if a racemic mixture can be separated (196). An optimal curve for such a mixture is shown in Figure 3.5.

Data from ITC experiments gave comparable results when compared to WF. For (-)-nicotine, ITC gave a K_D value of 2.3 ± 0.2 µM, compared to 0.41 ± 0.01 µM measured using the fluorescence-based assay. Similarly, a comparable value was measured by ITC for (+/-)-anatabine (3.4 ± 0.5 µM) compared to 8.5 ± 0.2 µM by the fluorescence assay. The K_D values for (+)-anatoxin are more similar, 0.30 ± 0.03 µM and 0.20 ± 0.01 µM by ITC and the fluorescence-based assay, respectively. Representative ITC binding curves for the three compounds are shown in Figure 3.3. An optimal curve is sigmoidal, giving information about the starting, mid and end-point of the protein-ligand interaction. The most comparable K_D values were measured for (+)-anatoxin, which has an optimal sigmoidal curve. Despite efforts to optimise ITC measurements of the other two ligands, an ideal sigmoidal curve could not be achieved which may contribute to variations between WF and ITC data, however, the data are of sufficient quality to draw conclusions.

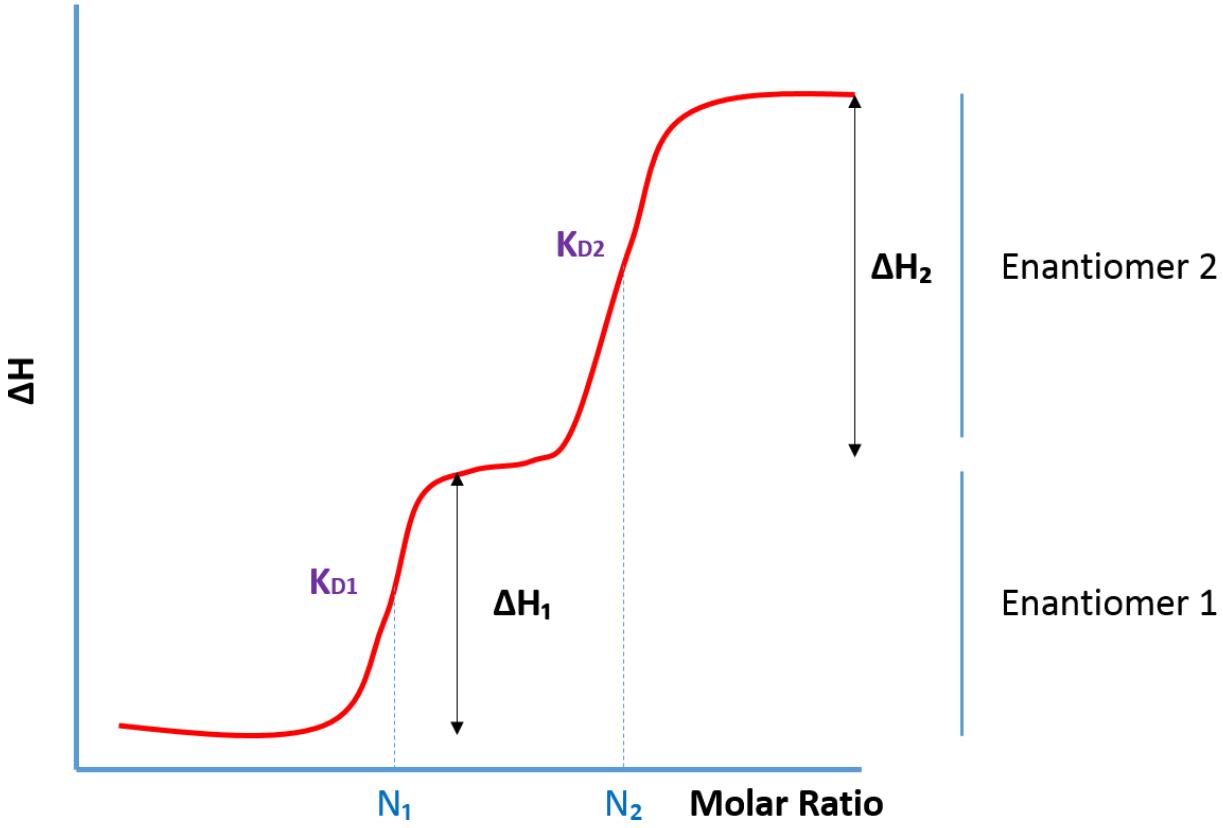


Figure 3.5 Optimal ITC binding curve for an enantiomeric mixture. The curve is plotted on axes of change of enthalpy against molar ratio. Two sigmoidal curves (red) are shown, one for each enantiomer. From each curve the thermodynamic parameters can be derived: change in enthalpy (ΔH, black), stoichiometry (N, blue) and affinity (K_D, purple).

The c value is a common reporter to whether an ITC experiment is reliable and if K_D can be measured. It represents the ratio between the concentration of the protein/titrand and either the association (K_A) or dissociation (K_D) constant, depending on which form of the equation is used (Eq. 3.1). K_A is derived from the ITC curve and is the inverse of K_D (197,198).

Eq. 3.1 $c = K_A [\text{Titrand}] = [\text{Titrand}] / K_D$

There are a variety of recommendations to what an ideal c value should be. It should be in the range of 1-1000 (169), but better ITC experiments are between 1-500 (199), with an ideal value of greater than 40 (200). The c value was used as an indicator to aid optimisation of ITC experiments with the three ligands and *Ac*AChBP. The c values for (-)-nicotine, (+)-anatoxin and (+/-)-anatabine are 11, 83 and 7, respectively using a protein concentration of 25 µM. All three values are within the ideal range, however, sigmoidal curves are not obtained for two ligands.

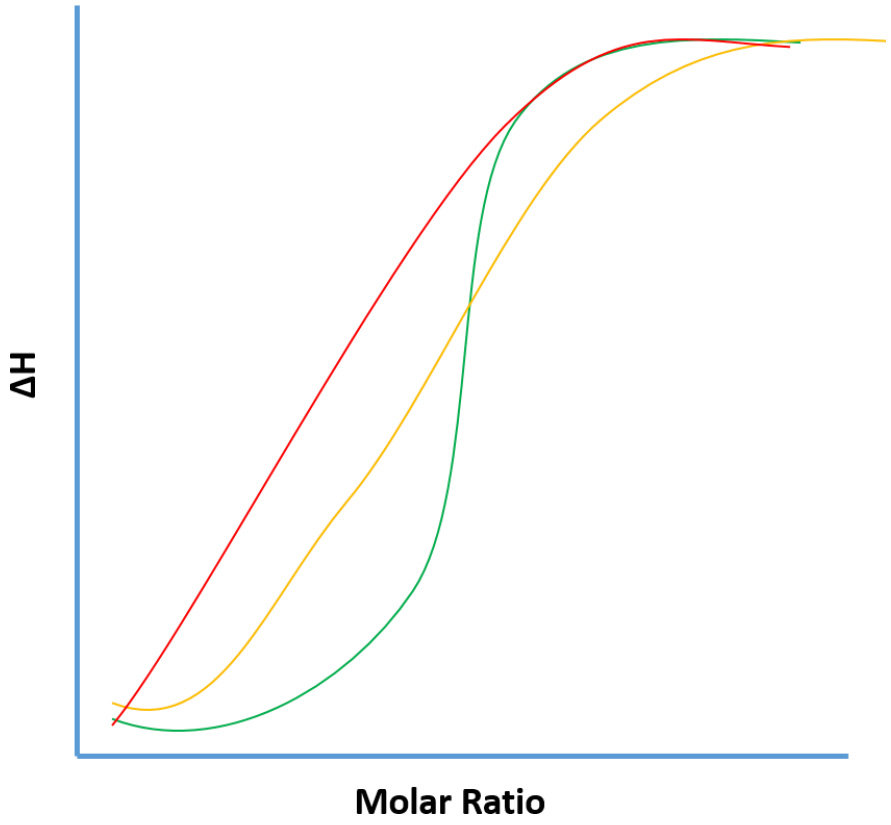


Figure 3.6 The effect of c in an ITC experiment. c has an effect on the formation of the binding curve. The curves are plotted on axes for change in enthalpy (ΔH) against molar ratio. The curves show the effect of increasing c value from low (c = 0.5, red) to medium (c = 5, orange) to an ideal value (c = 50, green). The c value is the ratio between the concentration of the protein and the dissociation constant ($c = [Protein]/K_D$).

Differences can be observed in the molar ratio determined for the three experiments (Figure 3.3). (+)-Anatoxin and (-)-nicotine have molar ratios close to one, whereas (+/-)-anatabine has a value above two. The molar ratio is determined by the concentration of ligand to protein in the sample cell. Therefore, binding site saturation may affect the formation of sigmoidal curves. Either binding sites are not saturated or are saturated too early. However, during optimisation problems were encountered when altering concentrations leading to decreased signal intensities, therefore no further improvement could be gained.

Despite testing a range of protein and ligand concentrations to investigate the binding of (+/-)-anatoxin and (+/-)-anatabine, elucidation of the effect of different enantiomers was not achieved. This supports previous observations stated above for the two compounds. The active enantiomer, (+)-anatoxin could lead to the binding curve and parameters observed from the (+/-)-anatoxin mixture. Due to the concentrations used, the binding of (-)-anatoxin may not be sufficient to record a heat change. A higher concentration may not lead to the elucidation of a (-)-anatoxin binding curve as all the binding sites would be filled by the more favourable (+)-enantiomer. For (+/-)-anatabine, only one binding curve was

recorded. This supports the idea that the two enantiomers are of similar affinity, therefore the K_D value will be an average. The ITC instrument used may not be able to detect the small differences between them.

SPR was employed as a third method. Due to the size of the compounds, the signal from binding was low. This coupled with the speed of association and dissociation meant that values for k_{on} and k_{off} could not be measured. This finding is consistent with using SPR and other small molecules (108). The dissociation constant (9.1 ± 1.3 μM) for (+/-)-anatabine is consistent with WF (8.5 ± 0.2 μM) and comparable to the ITC result (3.4 ± 0.5 μM). Consistency is also observed with WF and ITC for (+)-anatoxin SPR data (0.12 ± 0.02 μM). SPR results for (-)-nicotine (0.8 ± 0.1 μM) agree more closely with the K_D value from WF (0.41 ± 0.01 μM) than ITC (2.3 ± 0.2 μM). This corroboration of binding data further supports the idea that errors may be present in the ITC K_D values derived from non-optimal titration curves.

In summary, the data demonstrate that comparable values for the three techniques WF, ITC and SPR can be achieved. However, incomplete sigmoidal curves in the ITC experiments may have contributed to small K_D discrepancies between techniques. For small molecule screening, SPR is commonly used for hit identification due to its sensitivity and high throughput (84,201,202). Our data suggest that tryptophan fluorescence could offer an equally effective alternative for this purpose, depending on instrumentation for throughput and if the ligands are not fluorescent. ITC requires a longer timeframe to get similar results, due to the need for extensive optimisation. Also, ITC requires a large amount of material to run each experiment in comparison with the other two techniques. Due to these considerations, ITC is generally utilised in drug discovery after hit selection in order to improve their properties through the extra information that can be derived (121,203).

3.4 Thermodynamic parameters utilised to understand binding

A problem just discussed is incomplete ITC binding curves. The optimum curve is sigmoidal, from which the appropriate information can be derived. The value for K_D can be calculated from the gradient at the inflection point of the curve. Stoichiometry (N) is taken at half of the maximal response and is related to the molar ratio between ligand and protein. Change in enthalpy (ΔH) is measured from the bottom to the top of the curve (204). From these values, the change in Gibbs free energy (ΔG) can be calculated (Eq. 3.2).

Eq. 3.2 $\Delta G = -RT \ln K_D$

The calculated change in Gibbs free energy value, change in enthalpy and the constant temperature of the assay, can be used to calculate the change in entropy by rearranging an alternate form of the Gibbs equation ($\Delta G = \Delta H - T\Delta S$). Therefore, from a single curve five parameters can be extracted, which

provides details on the binding. Estimation of some measurements can occur for incomplete curves, however, these predicted values can still be used to provide details of the interactions.

3.4.1 The stoichiometry of ligand binding to *Ac*AChBP

The N value relates to how many molecules of a ligand are interacting with a protein entity. The simplest model is a 1:1 interaction. In the case of *Ac*AChBP there are five binding sites and if each site acts independently of the others, a 1:1 model will describe the ratio of ligand to a subunit. The ITC data indicated that N is 0.5 for (+)-anatoxin consistent with 2-3 sites being occupied. Both (-)-nicotine and (+/-)-anatabine have a lower N of 0.4, which suggests that less than half of the binding sites are occupied. The crystal structures of the *Ac*AChBP:nicotine (75) and the *Ac*AChBP:(+)-anatoxin complexes (PDB code: 6SH0) have all the orthosteric binding sites occupied with a single ligand. This discrepancy between the *Ac*AChBP:(-)-nicotine crystal structure and the ITC data may be due to the high concentrations of ligand present in the crystallisation conditions, about 40x that of the ITC experiments (75). Similar to the data presented, Celie *et al.* showed that an N of 3 (ligand/pentamer) could be achieved for *Ls*AChBP:nicotine ITC data (35). If this is taken per subunit a value of 0.6 is attained, which is consistent with the value measured (0.4). The reason why this is the case is still uncertain.

3.4.2 Reasons behind the differences in change of enthalpy for *Ac*AChBP ligands

Firstly, the difference in enthalpy changes due to binding for the three ligands to *Ac*AChBP will be considered. The value for ΔH comes from the energy released or utilised upon binding and is directly linked to what interactions are formed and/or disrupted, as well as changes in solvation. Enthalpy encompasses a more descriptive measurement, the heat of formation (ΔH_f), this describes the energy to form or break a bond. For a ligand binding to a protein, ΔH_f would describe the formation or breaking of an interaction (205). The energy of covalent bond formation, such as C-H, can be calculated; this is usually tens to hundreds of kcal mol^{-1}. However, for non-covalent bonds or interactions, as in a protein-ligand complex, much weaker energies are involved. For an electrostatic interaction, the energy can be calculated by Coulomb's law (Eq. 3.3).

Eq. 3.3 $E = k (q_1 q_2) / Dr$

For two equally but oppositely charged ions, electrostatic energies can range from 0-50 kcal mol^{-1}, depending on the distance. Hydrogen bonding is a specific type of electrostatic interaction. Primarily, energies range from 1 to 3 kcal mol^{-1}, dependent on distance. Van der Waals forces are caused by the movement of electrons in atoms generating dipoles, time-limited charges. Energies for this type of interaction are low (0.5-1 kcal mol^{-1}), which range over their effective distance of 3-8 Å (206).

The binding of all three ligands to *Ac*AChBP is an exothermic association. The largest ΔH value is observed for the interaction of *Ac*AChBP and (-)-nicotine (-17.8 ± 0.9 kcal mol^{-1}). A least-squares superposition of *Ac*AChBP complexes with (-)-nicotine and (+)-anatoxin show they are similar. The key residues of the binding site occupy similar positions in both structures, as do the two chemical entities. Both compounds make π-cation interactions with W164. Due to the pKa values, 7.9 and 9.4 for nicotine and anatoxin respectively, a MC hydrogen bond is formed with W164 as well. A hydrogen bond is also between the N of the pyridine ring with a key, well-ordered water molecule. In both structures, these water molecules have low B-factors of 23-24 $Å^2$. An ordered water molecule is consistently positioned here in AChBP complexes to donate a hydrogen bond to the ligand most notably in the *Ls*AChBP:nicotine structure (35). In the *Ac*AChBP complexes, the residues Y72, Y110, V125, M133, I135, W164, V165, Y205, C207, C208 and Y212 make van der Waals interactions with the ligand, contributing to binding (Figure 3.1). The only notable difference between the (-)-nicotine and (+)-anatoxin complexes lies with the key water-mediated hydrogen bond, previously mentioned. In the nicotine structure, this interaction links into a water-mediated hydrogen bond network through the protein, which will be highly stabilising, whereas the anatoxin structure lacks this hydrogen-bonding network throughout the protein. This may simply be due to the reduced resolution of the latter structure. (-)-Nicotine and (+)-anatoxin were docked into the binding sites of *Ac*AChBP from the respective complexes (Figure 3.1) and a variety of binding interactions were predicted. Upon comparison, both compounds have similar hydrogen bonding and hydrophobic energies, however, the values for (-)-nicotine are better than for (+)-anatoxin. (+)-Anatoxin has a better predicted van der Waals interaction energy (-24.9 kcal mol^{-1}) than (-)-nicotine (-23.5 kcal mol^{-1}), however, this does not explain the difference in enthalpy change. The desolvation energy predicted details the energy required to remove waters upon binding of the molecule. For (-)-nicotine the value of 5.6 kcal mol^{-1} is lower than that for (+)-anatoxin, 7.8 kcal mol^{-1}, which suggests that nicotine can be more readily desolvated and may require less energy to bind.

In the absence of an *Ac*AChBP:anatabine structure, the (-)-ligand was docked into the binding pocket of the (-)-nicotine complex (Figure 3.1). The modelled protein-ligand interactions are similar and ICM scores of all three ligands are comparable with values between -29 to -25 (ranked (-)-nicotine, (+)-anatoxin and (-)-anatabine, respectively). However, the estimated van der Waals interaction energy (-22.6 kcal mol^{-1}) is lower than for (-)-nicotine and (+)-anatoxin by 0.9 kcal mol^{-1} and 2.3 kcal mol^{-1}, respectively. Also, the desolvation energy (7.3 kcal mol^{-1}) is similar to (+)-anatoxin (7.8 kcal mol^{-1}). These two factors may be contributing to differences in enthalpy change upon binding.

	ICM Score	Hbond (kcal mol^{-1})	Hphob (kcal mol^{-1})	VDW (kcal mol^{-1})	Dsolv (kcal mol^{-1})
(-)-Nicotine	-28.5	-3.8	-4.9	-23.5	5.6
(+)-Anatoxin	-28.3	-3.6	-4.4	-24.9	7.8
(-)-Anatabine	-25.4	-4.0	-4.6	-22.6	7.3

Table 3.2 Calculated parameters for computational docking of three *Ac*AChBP ligands. The ligands (-)-nicotine and (-)-anatabine were docked into the *Ac*AChBP:(-)-nicotine structure (PDB code: 5O87), whereas (+)-anatoxin was docked into the *Ac*AChBP:(+)-anatoxin structure (PDB code: 6SH0). The ICM score indicates binding strength and the interaction energies for hydrogen bonding (Hbond), hydrophobic (Hphob), van der Waals (VDW) and desolvation (Dsolv) were predicted.

The SASA of (-)-nicotine, and (+)-anatoxin in bound and unbound states were calculated to investigate any correlation with binding properties. As no crystal structure for anatabine is available the calculation could not be performed. (-)-Nicotine and (+)-anatoxin have unbound SASAs of 333 Å^2 and 321 Å^2, respectively. There is already a difference in available surface to make contacts with the protein. This difference in surface area, 12 Å^2, relates to a potential hydrophobic free energy of 0.3 kcal mol^{-1} (207). Upon binding of (-)-nicotine to *Ac*AChBP there is a 59% (±0.2% SE) loss of SASA, compared to 56% (±0.2% SE) loss for (+)-anatoxin. If the solvent cannot access the surface of the ligand, this means that it is most likely buried in the protein and making contacts, therefore these data infer that nicotine has a greater area in contact with the protein. This indicates why the change in enthalpies may be different between the ligands. However, a 3% difference could also be caused by other factors including variation between binding pockets and ligand conformation/placement in the crystallographic structures.

3.4.3 Entropy changes upon ligand interactions with *Ac*AChBP

Entropy is the measure of disorder in a system. In the context of macromolecular structures and interactions, this involves contributions from displacing water molecules, conformational changes, increased or reduced flexibility of ligands or protein residues. A positive value means that disorder is increased and this is more favourable as it leads to a spontaneous interaction (205,208). Table 3.1 shows the interaction between *Ac*AChBP:(+/-)-anatabine has a negative entropy change value, in comparison to *Ac*AChBP:(+)-anatoxin and *Ac*AChBP:(-)-nicotine, which have a positive value. However, this value (-TΔS) is not the true entropy change (ΔS). The ΔS contributions of the compounds above are -0.03 ± 0.003, -0.01 ± 0.001 and 0.01 ± 0.002 kcal mol^{-1} K^{-1}, respectively for (-)-nicotine, (+)-anatoxin and (+/-)-anatabine. (+/-)-Anatabine has a favourable binding entropy change, nonetheless, all of these values are small in comparison to enthalpy changes, and therefore have minimal effect on binding. Consequently, it can be concluded that the binding of these ligands appears to be enthalpy driven.

3.5 Widening the scope to other ligands of *Ac*AChBP

With robust assays in place, there was an opportunity to investigate other compounds that are potential ligands and to identify similarities or differences. Three further compounds were selected for study. Hosiene A, derived from the roots of *Ormosia hosiei*, is a highly potent agonist of nAChRs (209,210). As such this small molecule should also interact with *Ac*AChBP. The two other compounds were strychnine, the archetypal ligand of GlyR, and N-methylbicuculline, an exemplary ligand of the $GABA_AR$ (Chapter 4). A new component was also brought forward for the study, the protein termed GBP, which is an engineered version of *Ac*AChBP designed to mimic aspects of a heteromeric GlyR (Chapter 1.1.4).

There is good agreement between the K_D values for (-)-hosieine, strychnine and N-methylbicuculline (Table 3.3) for the techniques and proteins tested. The only outlier is N-methylbicuculline with GBP, however, the difference between WF and ITC may be due to a secondary non-specific interaction of N-methylbicuculline to the protein (N 2.3 sites). All three of these compounds also have a MW greater than 300 Da, which is different from the small molecules previously discussed (Chapter 3.2) and may explain discrepancies when measuring binding and calculating affinity.

Ligand	Protein	N (sites)	ΔG (kcal mol^{-1})	ΔH (kcal mol^{-1})	-TΔS (kcal mol^{-1})	K_{D1} (μM)	K_{D2} (μM)
(-)-Hosieine	*Ac*AChBP	0.63 ± 0.03	-10.4 ± 0.2	-8.9 ± 0.1	-1.5 ± 0.3	0.025 ± 0.005	0.042 ± 0.004
Strychnine	*Ac*AChBP	0.5	-9.8 ± 0.2	-7.9 ± 0.4	-1.9 ± 0.5	0.07 ± 0.02	0.155 ± 0.007
N-methylbicuculline	*Ac*AChBP	0.83 ± 0.04	-7.4 ± 0.2	-4.0 ± 0.4	-3.3 ± 0.6	4.7 ± 1.6	1.24 ± 0.04
Strychnine	GBP	0.57 ± 0.03	-6.2 ± 0.1	-7.4 ± 1.0	1.3 ± 1.1	28.8 ± 3.2	26.6 ± 0.5
N-methylbicuculline	GBP	2.3 ± 0.1	-6.3 ± 0.2	-0.4 ± 0.1	-5.9 ± 0.3	29.6 ± 10.0	8.7 ± 0.5

Table 3.3 Binding data for the interaction of (-)-hosieine, strychnine and N-methylbicuculline. Mean thermodynamic parameters are given from ITC, as well as affinity values: K_{D1} from ITC and K_{D2} from the fluorescence-based assay. The parameters are for the interaction between *Ac*AChBP or GBP and the ligands. The standard error for each measurement is given (n=3). The raw traces and graphs can be seen in Chapter 4 for strychnine and N-methylbicuculline and in Figure 3.7 for (-)-hosieine.

One explanation to why these three compounds have a good agreement between ITC and WF data is that they are bigger (MW 230-367 Da) so may make more interactions (Figure 3.7), compared to (-)-nicotine and (+)-anatoxin. The structures for *Ac*AChBP:(-)-hosieine (PDB code: 6SGV), *Ac*AChBP:strychnine, GBP:strychnine (75) and GBP:N-methylbicuculline (211) complexes were available from the PDB. The percentage of SASA lost for each ligand was calculated using these structures. Comparing the starting SASA of (-)-hosieine (404 ± 0.3 Å^2 [n=10]), strychnine (451 ± 0.4 Å^2 for *Ac*AChBP [n=5] and 451 ± 0.5 Å^2 for GBP [n=4]) and N-methylbicuculline (527 ± 4.5 Å^2 [n=5]) it can be observed that these compounds are larger than (-)-nicotine (333 Å^2) and (+)-anatoxin (321 Å^2). However, due to the difference in size, flexibility and binding pose of these compounds, there is a higher chance for their

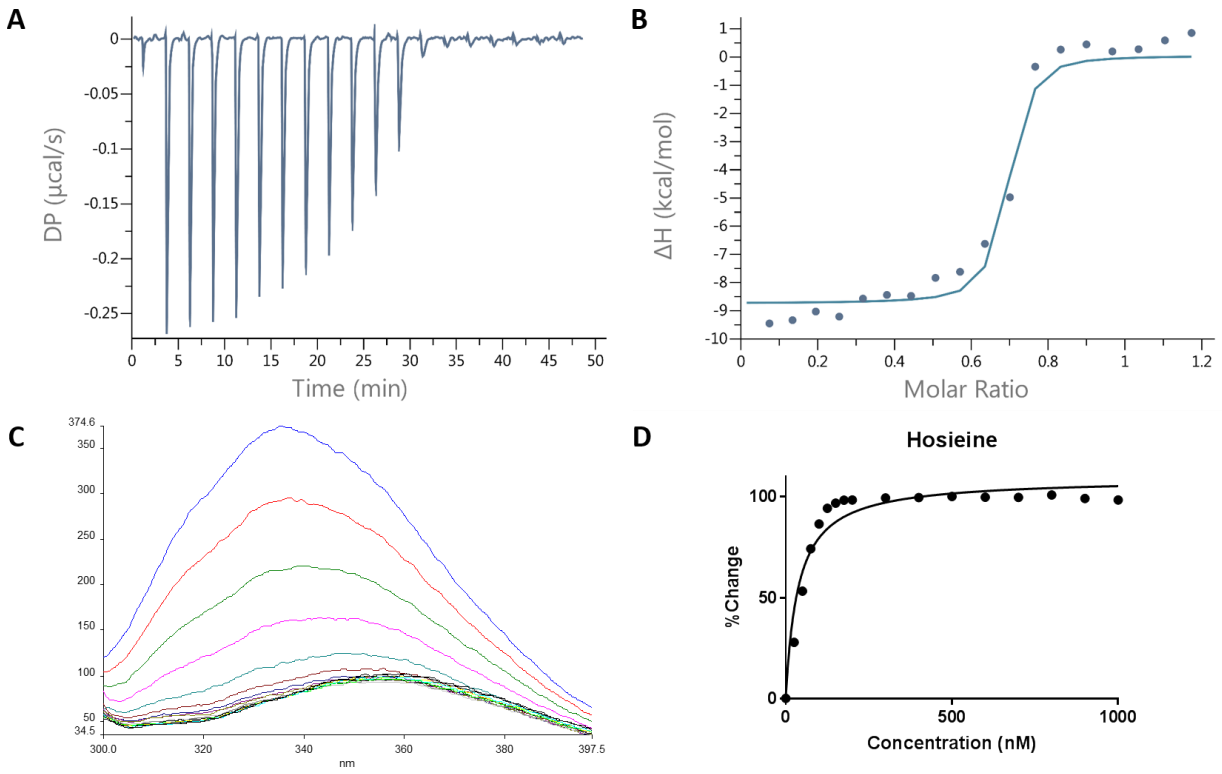


Figure 3.7 (-)-Hosieine binding to *Ac*AChBP data monitored by ITC (A, B) and WF (C, D). Graphs and traces are the same as those in Figure 3.2 and 3.4. Monitoring fluorescence both at constant and different wavelengths gave comparable K_D values for WF.

surface to not be in contact with the protein. This translates into lower values of percentage SASA lost upon binding (36-48%), in comparison with (-)-nicotine and (+)-anatoxin (56-59%). When accounting for the size of each compound and the SASA lost upon binding, all the compounds have a similar total contact area. Even though the total surface contact is similar for all the compounds, the number and strength of the interactions made by (-)-hosiene, strychnine and N-methylbicuculline may contribute towards better detection. For example, the binding of (-)-hosieine to *Ac*AChBP closely resembles the orientation of (-)-nicotine, however, the number of potential hydrogen bonds increases from two to four and the number of residues that interact by van der Waals interactions is increased by three (Figure 3.8).

3.6 Comparison with published data

Comparison of the data previously reported in the literature may validate further the use of the three biophysical techniques for the characterisation of *Ac*AChBP ligands. Unless stated otherwise data were extracted from CHEMBL (213). Due to the lack of K_D values, other parameters calculated from different assays were utilised. These parameters include IC_{50}, K_i and EC_{50} values. To calculate an IC_{50}, a displacement assay is conducted on physiologically relevant cells, which produce the receptor of interest or using recombinant protein. In this type of assay a known ligand, normally radiolabelled, is in competition with the ligand of interest for the binding site. The effect of competing out the labelled

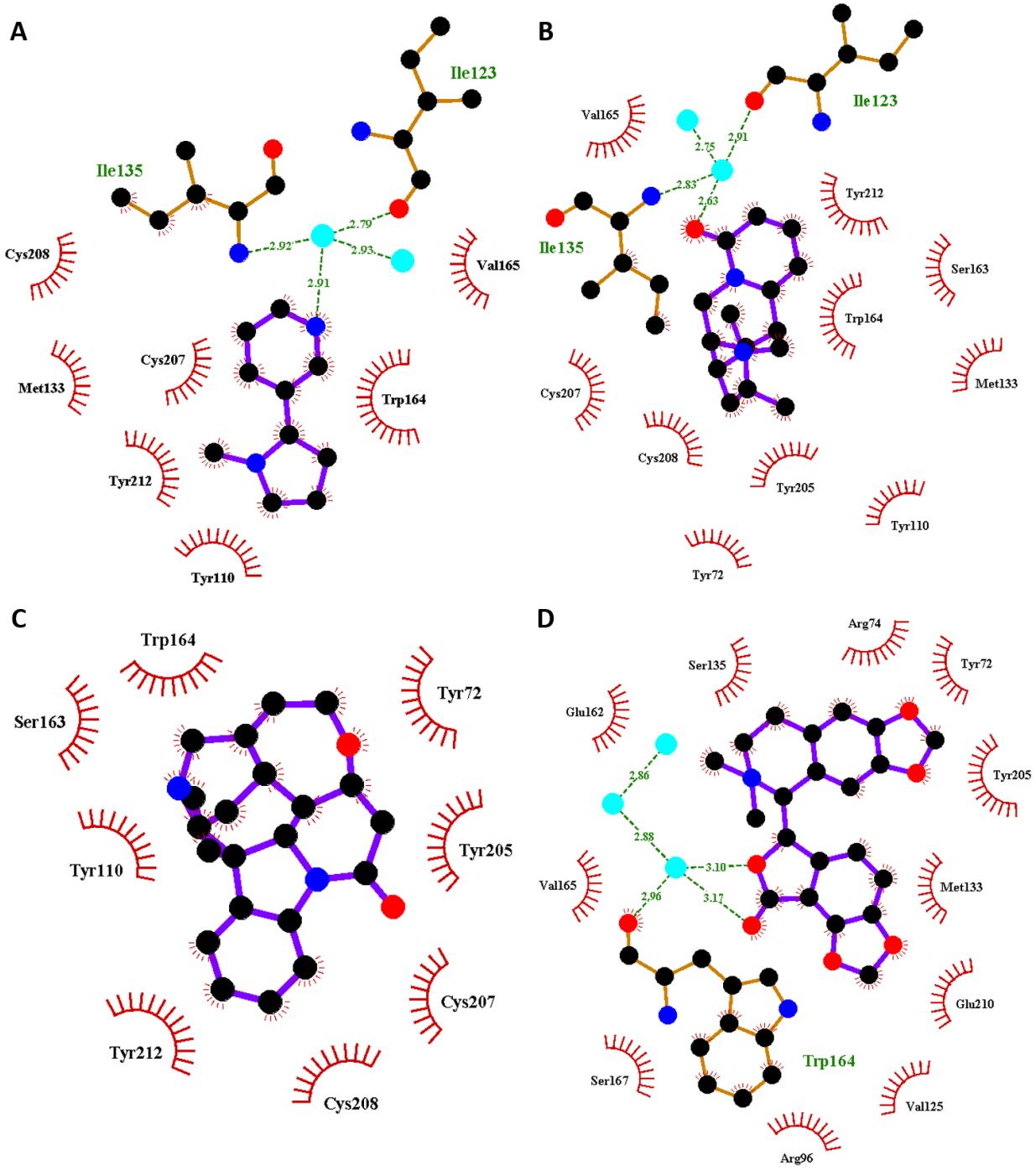


Figure 3.8 Schematic diagrams of ligand interactions with *Ac*AChBP and GBP. (A) *Ac*AChBP:(-)-nicotine, (B) *Ac*AChBP:(-)-hosieine, (C) *Ac*AChBP:strychnine and (D) GBP:N-methylbicuculline complexes. Ligands are shown in sticks with purple bonds, residues that participate in hydrogen bonding interactions are shown in sticks with orange bonds. Atoms are coloured: oxygen (red), nitrogen (blue) and carbon (black). Hydrogen bonds are shown as green dashed lines and atoms or residues that are involved in hydrophobic interactions are marked with red semi-circles. Water molecules appear as blue circles. The figures were made in LigPlus (212).

ligand is plotted against concentration (214). An IC_{50} is the concentration of an inhibitor where the response is reduced by half of the maximum. From this value, a K_i can then be calculated if the affinity of the labelled ligand is known (215). An EC_{50} is the effective concentration of a compound at half the maximal response (215). The EC_{50} is similar to IC_{50} and is dependent on the effect of the compound on the receptor.

As the archetypal ligand of nAChR, (-)-nicotine has a lot of binding data available, anatoxin has less and anatabine only has one source. Data relevant to the three ligands are dependent on the organism, receptor subtype and labelled ligand used. AChBP is a surrogate protein for nAChRs and shares the most similarity with the α_7 subtype (216). Due to this, data for AChBP (*Ls* and *Ac*) and the α_7 subtype are considered. IC_{50} values for (-)-nicotine range from 8-21 µM (217,218), whereas K_i values are between 0.2-4 µM (217,219-226) and values of 23-91 µM are determined for EC_{50} (227-230). All values for IC_{50} and K_i presented here have used a labelled α-bungarotoxin as the competitive ligand for the assays. Values of K_i for AChBP are similar between the two homologues (*Ac* and *Ls*), 0.08-0.6 µM (231-234). The K_D values (0.4-2.3 µM) from the three biophysical techniques, in comparison with this data, fit well with K_i values for α_7 and AChBP. This suggests that the affinity which was measured by all three techniques utilised is meaningful and representative of the literature data. The comparison also supports the use of displacement assays and the calculation of K_i values from the data, as a comparable technique. IC_{50} and K_i values for anatoxin competition with α-bungarotoxin at α_7 receptors ranged between 0.09-1.14 µM and 0.004 µM, respectively (226,235,236). These data demonstrate that anatoxin is of similar potency to nicotine. The K_D values determined (0.1-0.3 µM) fit with this observation and with the literature values. As previously stated, not much research has been conducted on the binding of anatabine, therefore only one study demonstrated an average EC_{50}, for both enantiomers, of 61 µM (194). This is not comparable with K_D values, as is observed for the comparison between nicotine EC_{50} and K_D.

3.7 Consideration of alternative techniques to study *Ac*AChBP-ligand interactions

In addition to the studies reported above, four other techniques to characterise ligand binding properties were considered and preliminary data generated. These methods are NMR, BLI, and two fluorescence-based approaches to study the influence of a ligand on protein stability.

NMR was used for the preliminary screening of a compound library against *Ac*AChBP, seeking to identify ligands (Chapter 5.2). The technique has also been developed further to quantify how strong a compound binds to the protein of interest (237). This method was being trialled, however, due to time restraints was not completed.

Similar to SPR, BLI uses biosensors and measures changes in interference patterns, caused by the binding of molecules (238). The technique was problematic for the analysis of molecules < 300 Da, as any changes were in the range of the background noise. A shift in the interference pattern is dependent on a density change on the biolayer, and due to the size of the compounds tested the signal may not be sufficient (239). Possibly due to the alternative method used in SPR to measure binding, this technique may be more effective at detecting the small changes occurring in intermolecular associations. One study analysed the binding of high affinity monoclonal antibodies to a target, on four biosensor platforms. Problems with noise and data reliability for BLI were reported in comparison with SPR (240). This study suggests experimental procedure, which may be linked to sensitivity leads to differences between SPR and BLI. Therefore, BLI may not be the most suitable method for screening small molecules, unlike SPR, and similar to ITC should be used for validation of improvements to drug-like compounds.

DSF and nanoDSF measure changes in fluorescence upon a temperature increase, which is an indicator of protein stability (117,241). Binding affinity could be deduced from these techniques by measuring the change in fluorescence upon increasing ligand concentrations. However, due to the stability of *Ac*AChBP (T_m and T_i > 95 °C), these methods are not suitable to measure the effects of ligand binding.

3.8 Summary

Three biophysical techniques, WF, ITC and SPR, have been used to measure the affinity of several ligands for *Ac*AChBP and GBP. The data presented display a level of agreement amongst these three techniques. Larger compounds showed good agreement between WF and ITC results, this may be due to additional interactions formed and their respective strength. Comparison of the recorded K_D values and literature parameters suggested there was a good agreement, particularly for nAChR α_7 and AChBP. Any of the three techniques could be appropriate to support compound screening and subsequent validation of potential ligands. Nonetheless, WF would be preferential due to small sample consumption and ease of use.

4. Characterisation of strychnine and N-methylbicuculline interacting with *Ac*AChBP and GBP

4.1 Introduction

Receptor-specific chemical tools have been used for the identification of CNS regions (242), as well as to assess the function of individual channel types (59). The most well-known chemical tools for the inhibitory receptors, GlyR and $GABA_AR$, are strychnine and bicuculline respectively.

The alkaloid strychnine is the archetypal antagonist of GlyR, with high affinity for homomeric (K_D 50 nM (243)) and heteromeric (K_i 210 nM (244)) receptors. In the GlyR homomeric receptor, it has been shown to inhibit pore opening by preventing the necessary conformational changes needed, upon ligand binding. Predominantly, it carries out this role by stopping the flexible loop C from closing (Chapter 1.1.2.3) (25).

Bicuculline is the exemplar ligand for $GABA_AR$. In the past, it was utilised to specifically recognise $GABA_ARs$, however, data have emerged suggesting it is a promiscuous ligand. Not only does it bind to $GABA_AR$ (IC_{50} 2.7 µM) (245), but it has been shown to interact with nAChR (IC_{50} 0.8 µM) (246), 5-HT_3R (IC_{50} 20 µM) (247), GlyR (IC_{50} 300 µM) (248) and a small-conductance calcium-activated potassium channel, which is located at the synapse (249). Two recent studies have provided structural data of the interaction between bicuculline and $GABA_AR$ (29,33), however, no data has been published for the other receptors.

The interactions of strychnine and N-methylbicuculline with *Ac*AChBP and GBP have been investigated by biophysical and structural techniques. These data provide new insights into the nature of binding to other receptor types. Data from these studies could also inform the use of GBP as a model for heteromeric GlyR. Work presented here contributed to two publications, Dawson *et al.* (75) and Jones *et al.* (211).

4.2 Deducing the binding affinity of strychnine with *Ac*AChBP and GBP using ITC and WF

The interaction between strychnine and both GBP and *Ac*AChBP was investigated using two biophysical techniques, WF and ITC. WF data were collected by exciting (280 nm) the aromatic residues, specifically a tryptophan (W164) in the orthosteric binding site and measuring emissions between 300-400 nm. Strychnine was titrated into the sample and the effect on fluorescence was measured. The measurements resulted in a K_D of 0.155 ± 0.007 µM for strychnine with *Ac*AChBP, which is consistent with previously reported data (0.04 and 0.02 µM by competitive binding assays and stopped-flow spectrofluorimetry, respectively (59,66)). After extensive optimisation of the WF experiments with GBP

a K_D of 26.6 ± 0.5 µM was calculated. This value indicates a lower affinity when compared to other data for homomeric GlyR α_1 (0.05 and 0.14 µM by radioligand saturation and ITC, respectively (250)) and α_3 (0.04 and 0.05 µM by SPR and ITC, respectively (24)). To further investigate the binding, ITC experiments were conducted. These took a longer time to optimise than WF assays, due to the need to balance the need for a sufficient heat change per titration point and saturation of binding sites. However, consistent data were recorded (Table 4.1).

Protein	N (sites)	ΔG (kcal mol^{-1})	ΔH (kcal mol^{-1})	-TΔS (kcal mol^{-1})	K_{D1} (µM)	K_{D2} (µM)
***Ac*AChBP**	0.5	-9.8 ± 0.2	-7.9 ± 0.4	-1.9 ± 0.5	0.07 ± 0.02	0.155 ± 0.007
GBP	0.57 ± 0.03	-6.2 ± 0.1	-7.4 ± 1.0	1.3 ± 1.1	28.8 ± 3.2	26.6 ± 0.5

Table 4.1 Binding data for the interaction of strychnine with *Ac*AChBP and GBP. Mean thermodynamic parameters are derived from ITC as well as dissociation constants: K_{D1} from ITC and K_{D2} determined from fluorescence measurements. The standard error for each measurement is given (n=3).

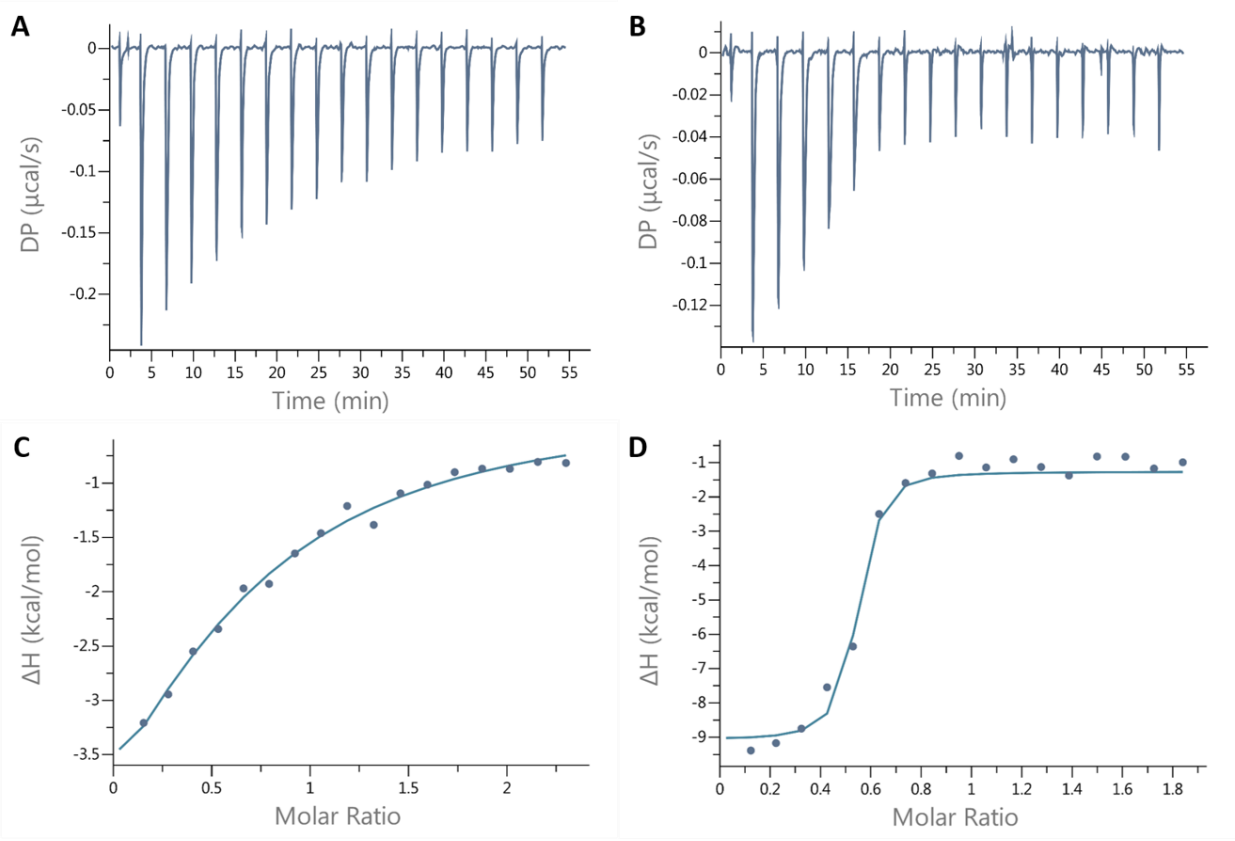


Figure 4.1 ITC data for strychnine binding GBP and *Ac*AChBP. Example isotherms (A, B) and curves with the best fit (C, D) are shown for GBP (A, C) and *Ac*AChBP (B, D). In the examples, a baseline deduction has been applied. All data have had the relevant controls (buffer-buffer, buffer-protein, ligand-buffer) deducted.

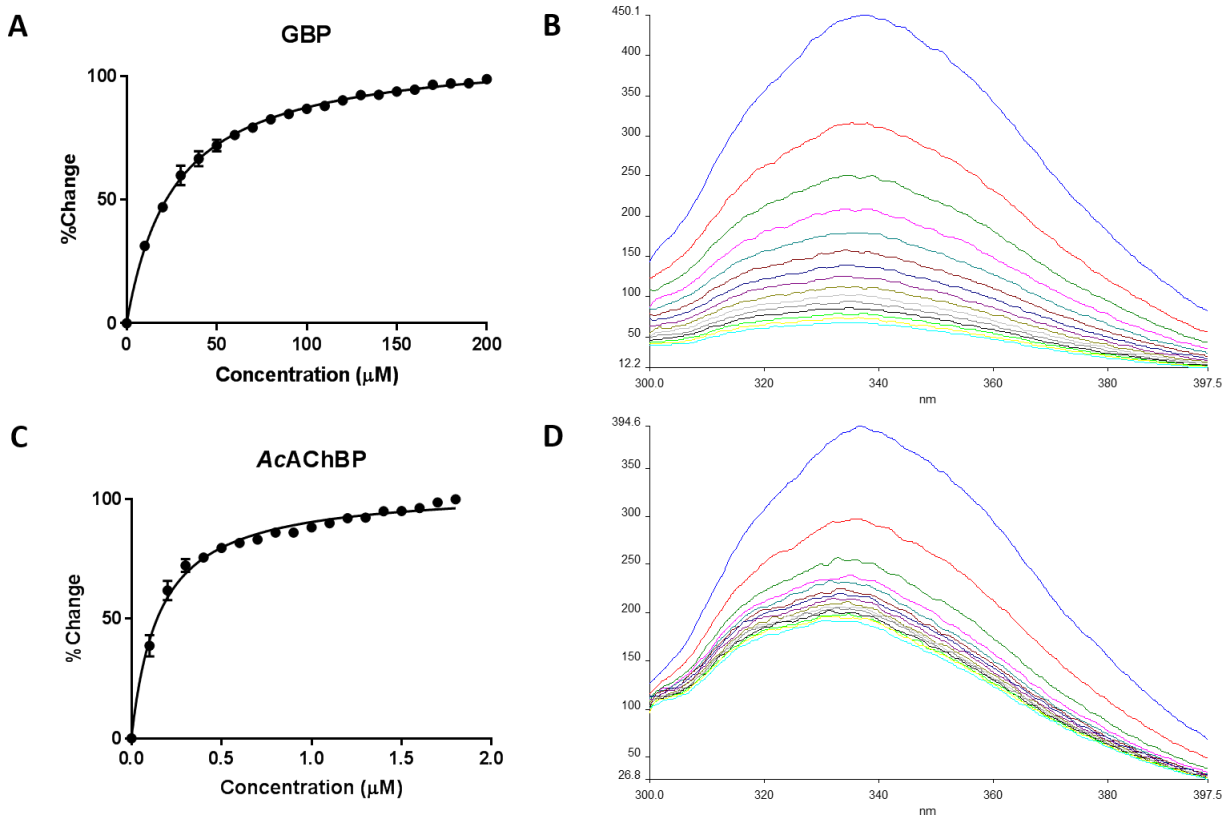


Figure 4.2 Strychnine binding monitored by WF. A plot of the mean percentage change versus concentration and example traces obtained from titrations, starting with no ligand (blue) to maximum ligand concentration (cyan) is shown for GBP (A, B) and *Ac*AChBP (C, D). Standard error bars are behind each data point (n=3).

The ITC derived K_D values are consistent with WF data. *Ac*AChBP has an ITC value of 0.07 ± 0.02 µM and WF K_D of 0.155 ± 0.007 µM. Whereas GBP has an ITC K_D of 28.8 ± 3.2 µM and a WF value of 26.6 ± 0.5 µM. This difference in K_D between the two proteins is most likely due to the substitutions made to *Ac*AChBP to form GBP which, potentially could disrupt binding interactions of strychnine with the protein.

As well as K_D, the ITC data provides a thermodynamic profile of binding. For both proteins, strychnine has an N of less than one. This suggests that not all binding sites are filled, 50% for *Ac*AChBP and 60% for GBP. This contrasts with ITC measurements made with GlyR α_3, which show an N of 1, therefore supporting a 1:1 interaction. However, these experiments were conducted in a different buffer (PBS compared to Tris) (24). The ΔH for strychnine binding to the two proteins is similar, therefore suggesting that they may make similar interactions. Due to substitutions in GBP, if similar interactions are being made this suggests that strychnine binding is driven by van der Waals interactions instead of specific hydrogen bonds or hydrophobic interactions with regions of the pocket. A difference between the proteins binding strychnine is the ΔS. *Ac*AChBP has a favourable ΔS, whereas GBP has an unfavourable change. When calculated, ΔS is 0.006 kcal mol^{-1} K^{-1} for *Ac*AChBP and -0.004 kcal mol^{-1} K^{-1} for GBP. One factor that contributes to entropy is SASA. As previously calculated (Chapter 3.5), the total SASA for

strychnine is about 450 Å^2, when binding to *Ac*AChBP it loses 36.3 ± 1.7% of the accessible surface, whereas binding to GBP it loses 44.1 ± 5.1%. This could explain the small variation in ΔS for ligand binding to the two proteins, however, other factors may also be involved. The difference observed between the change in enthalpy (-7.4 ± 1.1 kcal mol^{-1}) and entropy (1.3 ± 1.1 kcal mol^{-1}) for GBP:strychnine is comparable to that measured for GlyR α_3 by Huang *et al.*, ΔH of -16.3 kcal mol^{-1} and –TΔS of 6.4 kcal mol^{-1} (24).

4.3 Comparison of *Ac*AChBP: and GBP:strychnine crystal structures

Two structures of the *Ac*AChBP:strychnine complex have been reported (59,75). Superimposing these structures indicates a high degree of similarity (RMSD 0.5 Å). The ligand positions are preserved and residues in the orthosteric pocket display only slight differences in orientation, possibly due to flexibility of individual amino acids or different refinement strategies. Due to these similarities, only one structure was required for direct comparison to investigate *Ac*AChBP:strychnine binding. After inspecting the maps and due to similarities in resolution and chain completeness the in-house structure was chosen (PDB code: 5O8T). Strychnine binds surrounded by the aromatic cage residues Y72, W164, Y205 and Y212, which contribute van der Waals interactions (Figure 4.3). Other residues also contribute van der Waals interactions including T53, Q74, Y110, M133, I135, V165, D181, S184, S206, C207 and C208. One of the tertiary amines makes a hydrogen bond with the MC of W164. This amine is likely protonated, with a pKa of 8.3 (251). Due to the protonation of this amine, the resultant cation also forms an interaction with the π ring of the W164. This interaction is similar to that observed for other ligands of *Ac*AChBP including nicotine. The carbonyl of strychnine is solvent-accessible, which can lead to water-mediated MC interactions.

A comparison of the *Ac*AChBP:strychnine and GBP:strychnine complexes shows that, although positioned in the orthosteric site as expected, in most cases the ligand adopts different orientations. The five GBP:strychnine binding sites were superimposed to compare the position of the ligand and pick a representative chain. Chain B will be used as an exemplar for comparison as it is representative of the other strychnine orientations. In chain A there is a small difference in ligand orientation compared to chain B (RMSD 0.3 Å). One binding site of the GBP:strychnine complex has been modelled with two poses of strychnine. One of the ligands adopts an orientation similar to strychnine in *Ac*AChBP and the other is similar to the orientations found in the other binding sites. The similar strychnine orientation to *Ac*AChBP in GBP could be accommodated by a different rotamer of Y212, which is also present in this binding site and clashes with the example orientation. The exemplar orientation has rotated around 90°, compared to *Ac*AChBP:strychnine, and is now positioned with the indole group interacting with loop C and the ether directed towards S135. A reason for the different orientation would be the addition of polar residues (E162 and R74) directed into the pocket (75). These amino acid substitutions

would produce a clash with the orientation of strychnine as observed in the *Ac*AChBP:strychnine complex. Analysis of the GBP:strychnine complex and comparison with the *Ac*AChBP:strychnine structure indicates both conserved interactions (Y72, M133, W164, V165, Y205 and Y212) and different (R74, S135, E162, S163, Y166, S167, T208, E210 and P211) van der Waals interactions (Figure 4.3). Due to the change in orientation of strychnine, the tertiary amine no longer forms the MC hydrogen bond or π-cation interaction. Instead, this amine forms a water-mediated hydrogen bond with MC S163 and Y166. This interaction also links to other water molecules and residues setting up a stabilising hydrogen bonding network. Potentially, the loss of the π-cation interaction, which is important for binding to *Ac*AChBP (35), leads to the reduced affinity of strychnine for GBP.

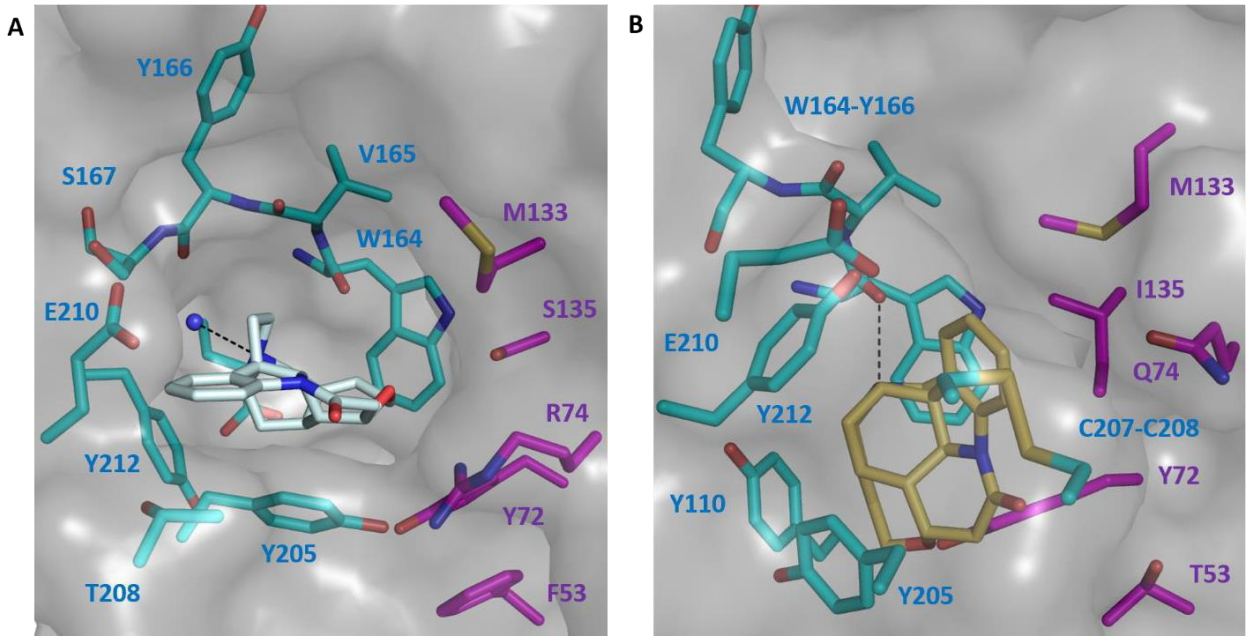


Figure 4.3 Strychnine bound in the orthosteric site of (A) GBP (PDB code: 5OBG) and (B) *Ac*AChBP (PDB code: 5O8T). Strychnine (silver for GBP and yellow for *Ac*AChBP) is shown bound surrounded by important residues of the principal-(+) (cyan) and complementary-(-) (magenta) subunits. Oxygen (red), nitrogen (blue) and sulfur (yellow) atoms are represented as well as a water molecule (blue sphere). Potential hydrogen bonds are shown as dashed lines.

4.4 Are interactions of *Ac*AChBP: and GBP:strychnine representative of the interactions in the physiologically relevant receptors, GlyR and nAChR?

The sequence alignment in Figure 4.4 was constructed to compare residue differences between *Ac*AChBP-nAChR and GBP-GlyR.

4.4.1 GlyR $\beta^{+}\alpha$

The sequence alignment of GBP, GlyR α_1 and GlyR β shows that all but one residue are conserved at the heteromeric interface. E210 on loop C is not conserved, however, removal of this amino acid in the native receptor should not affect binding as it does not contribute any specific interactions. Comparison of the

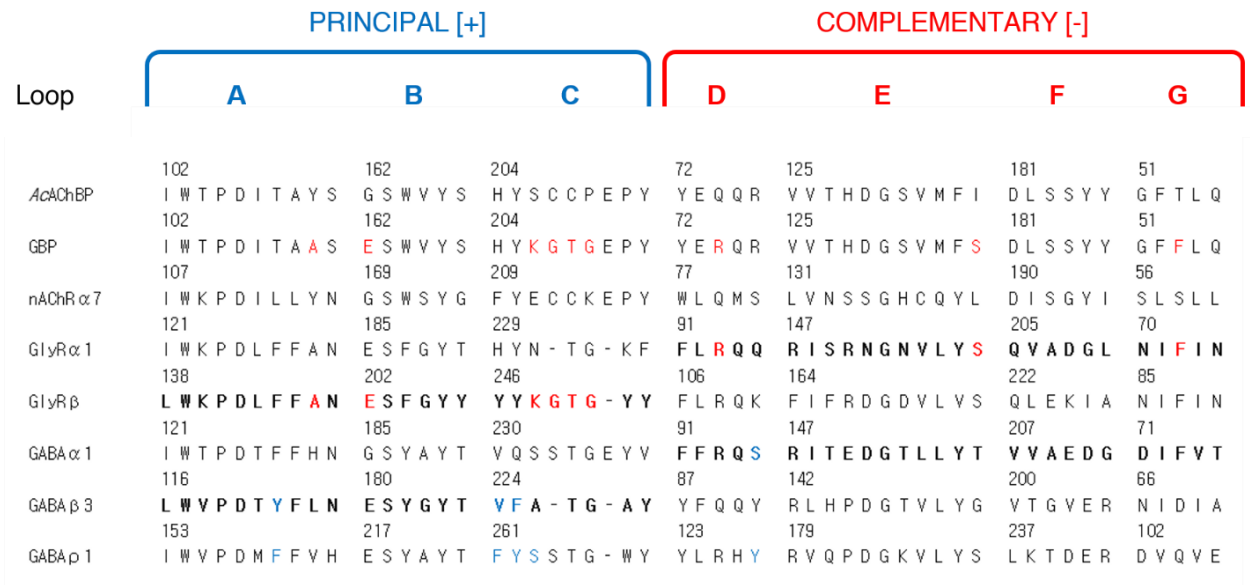


Figure 4.4 Sequence alignment of the important binding loops. The principal-(+) and complementary-(-) subunits of *Ac*AChBP, GBP, nAChR α_7, GlyR α_1, GlyR β, $GABA_AR$ α_1, $GABA_AR$ β_3 and $GABA_AR$ ρ_1 are compared. In red are residues of *Ac*AChBP which were substituted to form a heteromeric GlyR interface in GBP. The physiological interface of GlyR and $GABA_AR$ are shown in bold. In blue are residues that may prevent N-methylbicuculline binding $GABA_AR$ ρ_1.

principal subunit (loops A-C) binding site residues also demonstrates that loop C is the major difference between heteromeric ($\beta^+\alpha^-$) and homomeric interfaces ($\alpha^+\alpha^-$).

Two GlyR:strychnine complexes have been reported. These are the homomeric α_1 and α_3 receptors (24,25). The binding sites of these two homomeric structures are similar with slight alterations of residue orientation. Strychnine in these sites is also bound in a similar position. Comparing the structures of these homomers to GBP allows for the identification of conserved residue positions in the binding site, which supports the observations from the sequence alignment. However, even though there is a high conservation of residues, a different orientation of strychnine is adopted. This is emphasised by differences in affinity of strychnine with GBP (26.6-28.8 µM, Table 4.1) and the homomeric GlyR (0.04-0.14 µM (24,250)). One reason that may cause this change is due to residue differences, which make up the flexible loop C. GBP substitutions model the heteromeric GlyR C loop (KGTG) whereas the homomeric loop is different (NTG). This difference extends loop C of GBP, and the incorporation of a second glycine may lead to a more flexible loop able to accommodate the different orientation of strychnine observed. In the GlyR homomer structure, the C loop sits tightly closed against the pocket constraining the orientation of the ligand (Figure 4.5). Potentially, due to the flexibility of loop C in GBP, similar close interactions with strychnine cannot be formed possibly leading to the lower affinity observed.

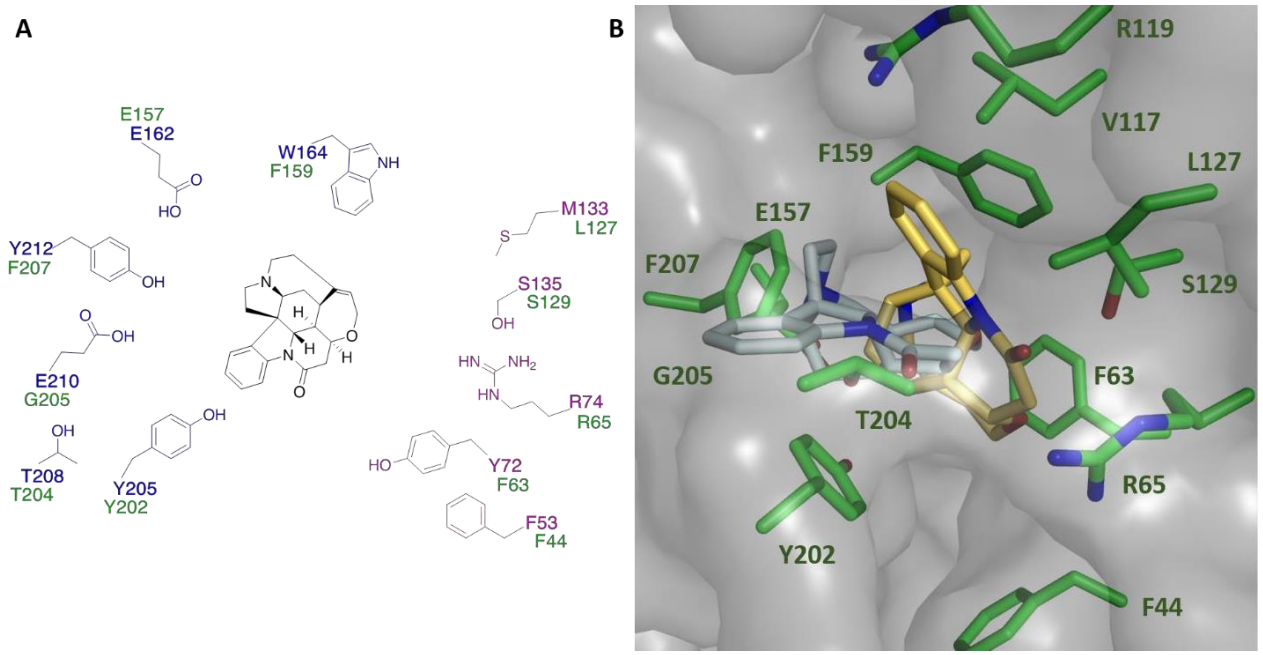


Figure 4.5 Differences between strychnine bound to GBP and GlyR. (A) A schematic of strychnine bound in the GBP binding pocket and the interactions it makes as seen in Figure 4.3A. The comparable residues in GlyR are shown in green. (B) Strychnine (yellow) bound in the homomeric GlyR (PDB code: 5CFB) orthosteric pocket. Residues (green) and the alternate orientation of strychnine in GBP (silver) are shown. Atoms are coloured as in Figure 4.3.

4.4.2 nAChR α_7

The sequences of *Ac*AChBP and nAChR α_7 were compared. As in GBP:strychnine binding, most of the interactions with the ligand are MC van der Waals, therefore substitution does not have a large effect. Important interactions are potentially maintained in the sequence of α_7, primarily the MC hydrogen bond and π-cation between strychnine and W164. No structures are available with strychnine bound to nAChRs for comparison. Therefore, based only on sequence data it can be suggested strychnine would bind in a similar orientation. However, data from competitive binding assays suggest that there is a difference between K_i values for *Ac*AChBP (0.04 µM, (59)) and nAChR α_7 (5.4-6.9 µM (252)). This suggests, similar to GBP and GlyR, other factors are in effect when strychnine binds to the native nAChR α_7.

4.5 Deducing the binding affinity of N-methylbicuculline with *Ac*AChBP and GBP by biophysical techniques

Bicuculline has a half-life of 45 minutes at physiological pH (253), therefore the more stable N-methylbicuculline was used in this study (254). N-methylbicuculline differs by the addition of a methyl group on N7, which generates a cationic species. Due to the pKa of bicuculline (14.7), the N7 position would be protonated, therefore a cation would also exist for the original compound. Firstly, WF was utilised to elucidate the affinity for N-methylbicuculline with GBP and *Ac*AChBP. Optimisation of each titration was conducted to find the ideal concentration of titrant, which differed between proteins. The

experiments resulted in K_D values of 1.24 ± 0.04 µM and 8.7 ± 0.5 µM for *Ac*AChBP and GBP respectively. Using these values to guide experimental design, ITC experiments were conducted. The optimisation was aided by monitoring the c value and maintaining it between 1-500, the optimal range (199). Due to the weak signal obtained from N-methylbicuculline titrations with GBP, extensive trialling of experimental conditions was required to obtain useable data. The final concentration of N-methylbicuculline used was 2 mM, which also gave a final DMSO concentration of 2%. GBP has been shown, by nanoDSF data, to be stable up until 2.5% DMSO (Chapter 2.4.5). Even after extensive optimisation, the data were not ideal due to relatively low enthalpic input (-0.4 ± 0.1 kcal mol^{-1}).

Protein	N (sites)	ΔG (kcal mol^{-1})	ΔH (kcal mol^{-1})	-TΔS (kcal mol^{-1})	K_{D1} (µM)	K_{D2} (µM)
***Ac*AChBP**	0.83 ± 0.04	-7.4 ± 0.2	-4.0 ± 0.4	-3.3 ± 0.6	4.7 ± 1.6	1.24 ± 0.04
GBP	2.3 ± 0.1	-6.3 ± 0.2	-0.4 ± 0.1	-5.9 ± 0.3	29.6 ± 10.0	8.7 ± 0.5

Table 4.2 Binding data for the interaction of N-methylbicuculline with *Ac*AChBP and GBP. Mean thermodynamic parameters are derived from ITC as well as dissociation constants: K_{D1} from ITC and K_{D2} determined from fluorescence measurements. The standard error for each measurement is given (n=3).

The interaction between GBP and N-methylbicuculline is entropically driven (-TΔS, -5.9 ± 0.3 kcal mol^{-1}) with a K_D of 29.6 ± 10.0 µM. For the interaction between *Ac*AChBP and the ligand, similar contributions from ΔH (-4.0 ± 0.4 kcal mol^{-1}) and -TΔS (-3.3 ± 0.6 kcal mol^{-1}) are noted and the K_D is 4.7 ± 1.6 µM. The K_D values determined by ITC are comparable with the values derived from WF. Calculation of the SASA lost for N-methylbicuculline binding to GBP is 43 ± 2%. As the AcAChBP:N-methylbicuculline has a less favourable change in entropy it could have a higher percentage of SASA lost leading to more interactions and an improved enthalpy change. However, a higher resolution structure of the *Ac*AChBP:N-methylbicuculline complex than reported below would be required to confirm this hypothesis. As well as the differences in affinity and thermodynamic parameters, there are differences with N, the molar ratio of ligand to target. N-methylbicuculline should bind in a 1:1 relationship, the ligand to the binding site, although for the two proteins this is different. For GBP N is around 2, which suggests that two molecules of N-methylbicuculline may bind per subunit, however, due to difficulties in data acquisition a full titration curve could not be determined, which means N may not be accurate. The possibilities exist that there is a second binding site, or that a second ligand binds at the orthosteric site, or aggregation of the ligand could be complicating the study. For example in the GBP:strychnine crystal structure, one pocket contained two molecules of the ligand (59). Conversely, *Ac*AChBP has an N of 0.8, which suggests that 80% of the binding sites are filled. This is consistent with other data reported for *Ac*AChBP (N of 0.6 for *Ls*AChBP:nicotine and 0.5 for *Ls*AChBP:carbamylcholine (35)) that show an N of 1 is not achieved for binding. Yet, there is no evidence of cooperativity or allosteric effects.

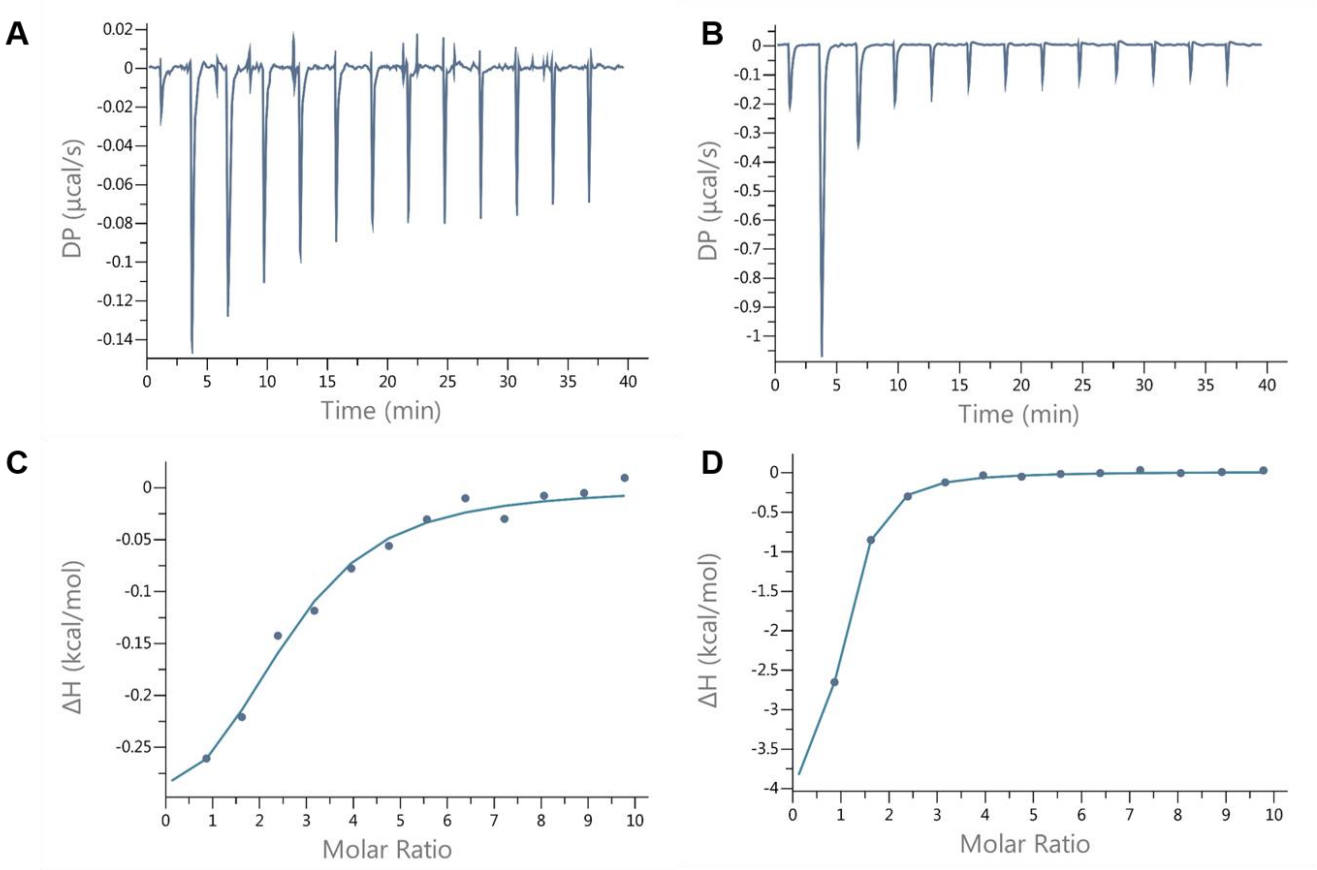


Figure 4.6 ITC data for N-methylbicuculline binding GBP and *Ac*AChBP. Example isotherms (A, B) and curves with the best fit (C, D) are shown for GBP (A, C) and *Ac*AChBP (B, D). In the examples, a baseline deduction has been applied. All data have had the relevant controls (buffer-buffer, buffer-protein, ligand-buffer) deducted.

4.6 Crystal structure of GBP:N-methylbicuculline

The crystal structure of the GBP:N-methylbicuculline complex was obtained at a resolution of 2.4 Å, with one molecule of N-methylbicuculline in each of the binding pockets. In chains B, C and E the residues K206GT and in chain D the residues G207T could not be modelled. This is due to diffuse density owing to the flexibility of loop C, which is in keeping with its function (25). In certain places, the density was clear enough to model the glycosylation of N91 (chains B, D and E) and place a well-ordered chloride ion (chains A-E, B-factor 37-58 $Å^2$). It was noted there was one orientation of the ligand, however, there were two slightly different poses. For brevity, the exemplar molecules will be considered here. Pose I comes from chain B, this has an average B-factor of 67 $Å^2$ and a real space correlation coefficient (RSCC) of 0.90. Pose II is from chain A with a B-factor of 97 $Å^2$ and a RSCC of 0.88. B-factors infer the degree of error in the placement of an atom, the higher the value the less certain, as is seen in more flexible regions. RSCCs are an indicator of ligand fit to electron density, 0.8 and above is considered a good fit, below 0.8 suggests a degree of caution be exercised when considering this part of the structure. From these values it can be inferred that both poses fit the electron density well, however, there is less order for pose II,

possibly due to more flexibility in that binding site, compared to pose I.

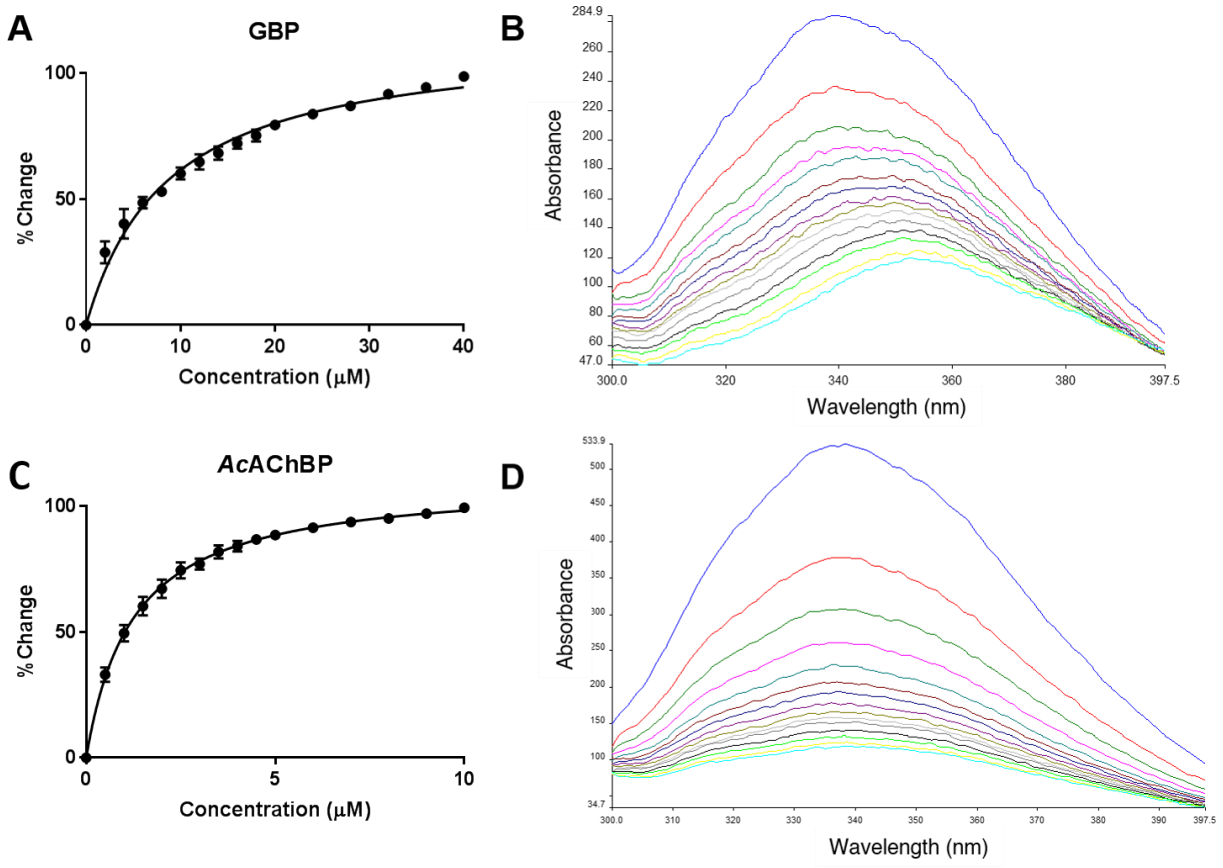


Figure 4.7 N-methylbicuculline binding monitored by WF. A plot of the mean percentage change versus concentration and example traces obtained from titrations, starting with no ligand (blue) to maximum ligand concentration (cyan) is shown for GBP (A, B) and *Ac*AChBP (C, D). Standard error bars are behind each data point (n=3).

4.6.1 Pose I

Pose I (Figure 4.8A) is adopted in three of the five binding sites. Due to the presence of six oxygen atoms on N-methylbicuculline there is the possibility of multiple hydrogen bonds. The isoquinoline motif has two oxygen atoms in the dioxolo group, one of these makes an interaction with Nε of R74 (3.3 Å), whereas the other is solvent accessible. The dioxolo group on the phathlide forms two hydrogen bonds, one with R96 (3.4 Å) and one with the protonated E210 (3.1 Å). The carbonyl and oxygen of the furan ring coordinate a well-ordered water molecule (B-factor 31 $Å^2$), this is then coordinated by the MC carbonyl of W164 and another well-ordered water (B-factor 31 $Å^2$). Superimposing this structure with *Ac*AChBP:nicotine complex indicated that the second water was conserved in both structures. A water-bridged network of hydrogen bonds is commonly reported with *Ac*AChBP structures.

A key interaction noted is the cation-π interaction with W164. This cation also sits in an electron-rich environment produced by Y72, E162, W164 and Y212. The cation-π interaction is a consistent feature of

other ligands that interact with *Ac*AChBP and is a major contributor to binding. Other residues including F53, Y72, V125, M133, S135, V165, Y166, S167, Y205 and Y212 all form van der Waals interactions with the ligand (Figure 4.8). Of these Y72, W164, Y205 and Y212 make up the aromatic cage, which is a defining feature of the pLGIC family.

4.6.2 Pose II

Pose II is slightly different to pose I with a rigid body rotation around N7. This may be due to a combination of crystallographic disorder in the structure and the flexible nature of the pLGIC binding site. There is conservation of van der Waals interactions in the two poses but details of hydrogen bonding change (Figure 4.8). For the isoquinoline motif, the hydrogen bond between the dioxolo group and R74 is maintained (3.2 Å), as is the exposure to the bulk solvent of the other oxygen atom. Due to the rotation, the carbonyl and furan oxygen atoms no longer interact with the well-ordered water, instead, the carbonyl forms a hydrogen bonding network with a dioxolo oxygen atom, Y212, V165 and a different water molecule (B-factor 42 $Å^2$). Y212 can now make this hydrogen bond due to a rotameric shift, like Y205, which accommodates the move of N-methylbicuculline. N7 of the isoquinoline motif is still maintained in a similar location as in pose I, which makes a cation-π interaction with W164. Van der Waals interactions are similar to those described for pose I.

4.7 Comparison of the GBP:N-methylbicuculline complex with the inhibitory pLGICs, GlyR and $GABA_AR$

4.7.1 GlyR

The model of GBP:N-methylbicuculline complex was compared to the model of the full length homomeric GlyR. After sequence and structural alignments, key residues involved in the binding of N-methylbicuculline in GBP are conserved in the GlyR α3 homomer (Figure 4.4). Of the twelve residues, five are strictly conserved (F53, R74, E162, S135 and Y205) and four are homologous (Y72, M133, W164 and Y212). This leaves three residues that are not conserved, firstly, E210 is a deletion in the sequence alignment, however, it contributes a hydrogen bond to the ligand in the complex. Nonetheless, a similar interaction could be formed with T208 in GlyR. R96E and V125R should be considered in tandem, the glutamic acid is pointing away from the binding site, therefore does not have a direct effect, however, it does form a hydrogen bond with the arginine, which pulls the basic residue into a position similar to that occupied by R96 and so could interact with N-methylbicuculline. This comparison demonstrates that GBP is consistent with the GlyR model and could be representative of N-methylbicuculline bindingin full length native heteromeric receptors.

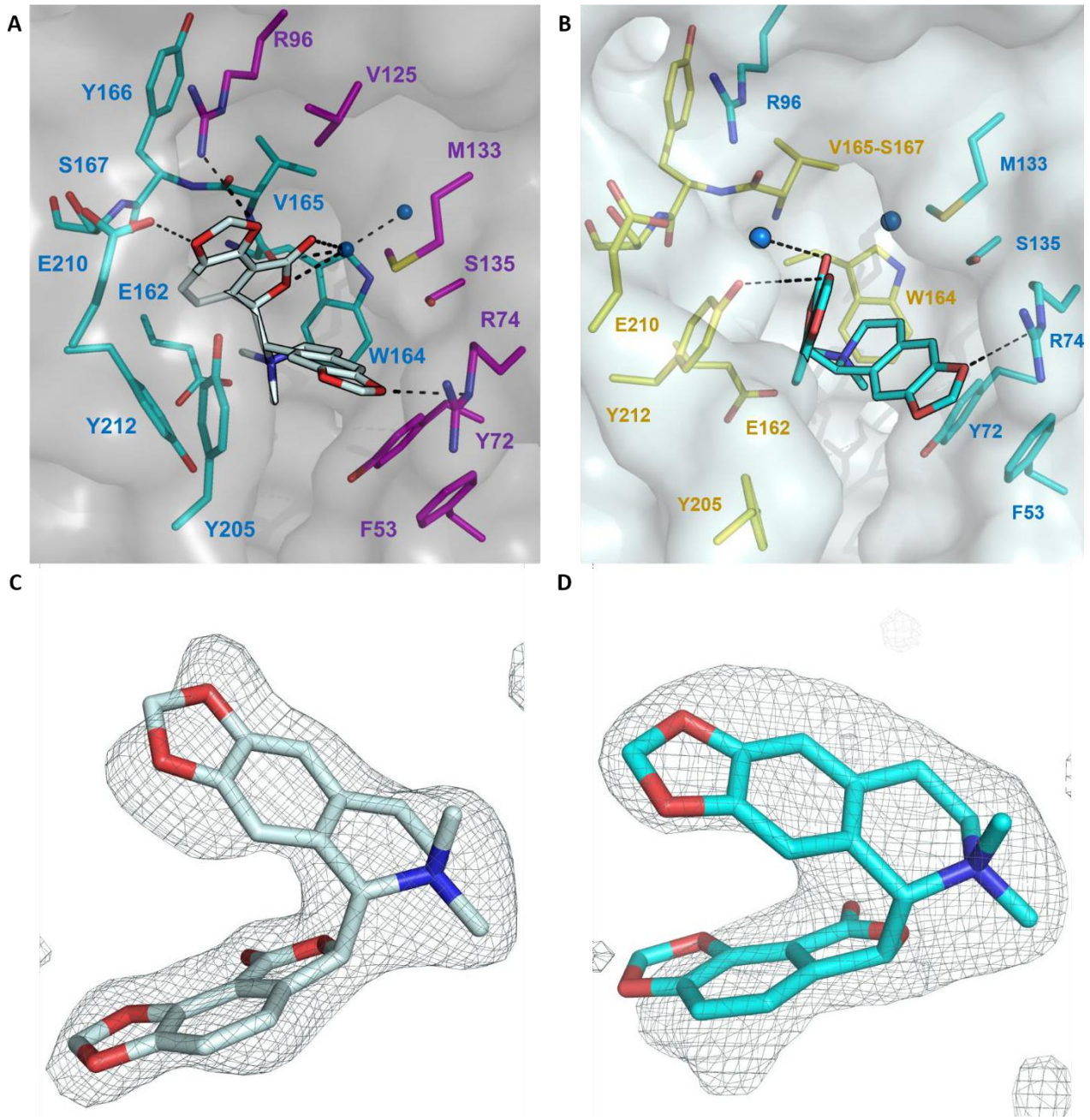


Figure 4.8 Representation of N-methylbicuculline bound to GBP. Two similar poses are observed in the orthosteric binding site (PDB code: 5OBH). (A) Pose I (silver) is in the binding site made up of the principal (cyan) and complementary (magenta) subunits. (B) Pose II (cyan) is in the binding site made up of the principal (yellow) and complementary (cyan) subunits. Atoms are depicted as in Figure 4.3. Polder (255) omit F_o-F_c maps of (C) pose I and (D) pose II are shown. Maps were calculated in Phenix (256) then normalised in Coot. The maps are contoured at 5σ. Atoms are depicted as in Figure 4.3.

4.7.2 $GABA_AR$

After comparing the GBP:N-methylbicuculline model with GlyR, it was decided to compare it to $GABA_AR$, which was previously thought of as the archetypal receptor for this ligand. Two structures show bicuculline binding to the heteromeric αβγ $GABA_AR$, by cryo-EM (29,33). Superposition of these structures themselves led to the observation that they are comparable (RMSD 0.5 Å). As such, only one receptor will be used for comparisons (PDB code: 6HUK). The structure shows one bicuculline molecule

bound in each of the two $\beta^{+}\alpha^{-}$ heteromeric binding sites (29). Upon inspecting the maps and calculation of the RSCCs of these two molecules (0.84 and 0.81), it was decided that they are representative of binding and comparable to the GBP:N-methylbicuculline complex.

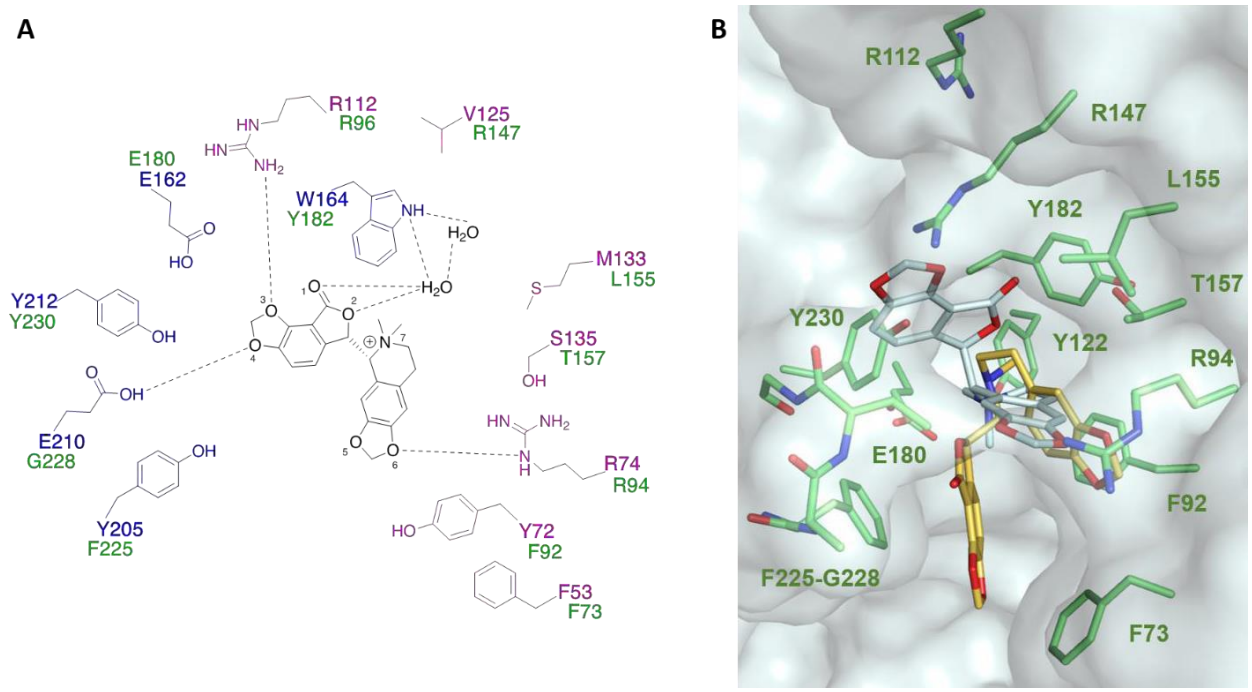


Figure 4.9 Differences between N-methylbicuculline bound to GBP and $GABA_AR$. (A) A schematic of N-methylbicuculline bound in the GBP binding pocket and the interactions it makes as seen in Figure 4.8A. The comparable residues in $GABA_AR$ are shown in green. (B) Strychnine (yellow) bound in the heteromeric $GABA_AR$ (PDB code: 6HUK) orthosteric pocket. Residues (green) and the alternate orientation of N-methylbicuculline in GBP (pose I, silver) are shown. Atoms are coloured as in Figure 4.3.

Firstly, a sequence and structural alignment were conducted. As with GlyR, this showed that many of the residues important for N-methylbicuculline binding in GBP are conserved in $GABA_AR$. Of the 13 residues, five are strictly conserved (F53, R74, R96, E162 and Y212) and five are homologous (Y72, M133, S135, W164 and Y205) (Figure 4.4). This leaves three residues that are not conserved (E210G, T108Y and V125R). Similar to GlyR, the substitution E210G (this is a deletion in the sequence alignment) does not affect the binding site. The T108Y, which does not have a role in GBP, adds extra aromaticity to the negatively charged pocket, in which N7 binds. A key difference between GBP and $GABA_AR$ is the V125R substitution, the arginine, in its current orientation forms a hydrogen bond with either the MC of G183 or the side chain (SC) of Y230 (Figure 4.9). This residue provides a steric block, preventing pose I of N-methylbicuculline binding GBP from being adopted in the $GABA_AR$ binding site. Due to this orientation not being feasible, another is adopted. This different orientation is established by a rotation, which maintains the isoquinoline motif in a similar position as in GBP. The position of N7 is also retained, within 1 Å, therefore conserving the cation-π interaction. Even though rotated, the dioxolo group of the isoquinoline motif forms a hydrogen bond with R94, which is analogous to R74 in GBP. In the new

orientation, a variety of van der Waals interactions are also established. No water molecules are included in the cryo-EM structure, therefore additional hydrogen bonds through water-mediation are not described. Due to the flexibility of the pLGIC ligand binding pockets and the number of interactions N-methylbicuculline can make, leads to its promiscuous activity across the protein family. Even though the orientations observed are different, N-methylbicuculline has a preferred conformation, as is seen by the overlay of pose I with pose II or the $GABA_AR$ bicuculline (Figure 4.10). In comparison, the overlay of pose I with the crystal structure from the CCDC database demonstrates the flexibility bicuculline can adopt.

The $GABA_AR$ structure and publications raise an issue in the use of bicuculline as a chemical tool for the study of pLGICs. The paper reports that N-methylbicuculline is used, however, the structure contains bicuculline. Due to differences in potency (257), it is important to know which compound was used so that comparisons can be conducted correctly.

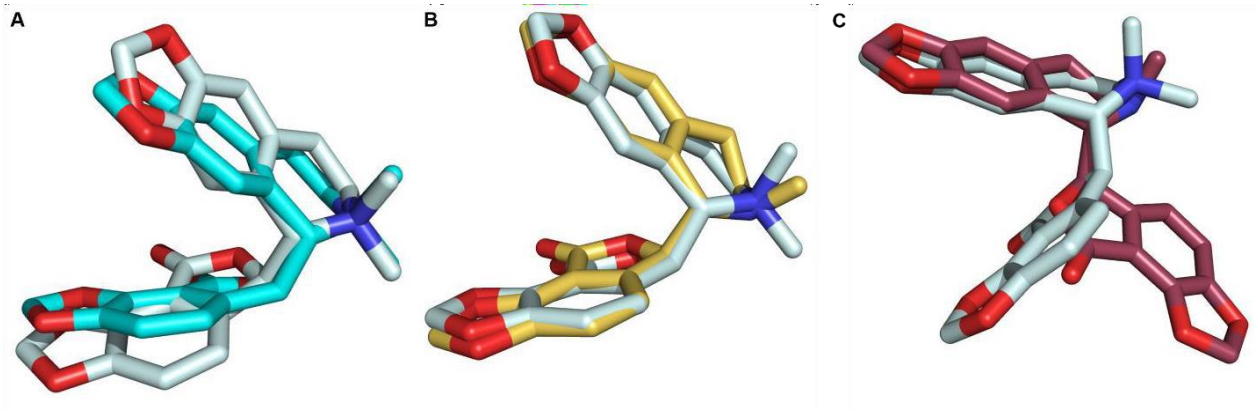


Figure 4.10 Superpositions of N-methylbicuculline and bicuculline. (A) A superposition of N-methylbicuculline pose I (silver) and pose II (cyan) from the complex with GBP. (B) Superposition of N-methylbicuculline pose I from GBP (silver) and bicuculline (yellow) from $GABA_AR$ (PDB code: 6HUK). (C) N-methylbicuculline pose I from GBP (silver) superposed on the single crystal structure of bicuculline from the CCDC database (code: Bicucul01, raspberry). For A and B the superpositions were calculated in Coot with the LSQ function using the GBP complex as the reference structure. In (C) the Superpose Ligands function in Coot was used.

4.7.3 GABA ρ receptors

$GABA_A$ ρ receptors, otherwise known as $GABA_C$ receptors ($GABA_CR$) are insensitive to bicuculline. Due to a lack of structural data on these proteins, insights into the reasons why rely on mutational studies. The substitutions Y127S, F159Y and F261YS_VF in $GABA_A$ ρ were key for the sensitivity of the receptor to bicuculline (258). These residues gave the largest increase in potency and by looking at the $GABA_AR$ structure the reasons can be rationalised. The Y127S substitution opens space behind R94 allowing movement in this residue and contributes van der Waals interactions to the binding site. The F159Y substitution adds a hydroxyl group to the binding site, which provides one of the interactions

with N7 of bicuculline and also forms a hydrogen bond to E202, which could stabilise this key residue, essential for GABA binding. Lastly, the tripeptide substitution to a dipeptide removes extra bulk from the binding site, allowing room for bicuculline to bind. In addition to these residues, from the analysis of the GBP:N-methylbicuculline structure and comparison to $GABA_AR$ models, the substitutions A229W and F73Q may also play a role. A229W, on loop C, adds extra bulk into the binding site therefore it may inhibit bicuculline binding. F73 is a conserved residue across GBP, GlyR and $GABA_AR$, the substitution to Gln would remove the aromaticity and may direct a polar group towards the binding site.

TPMPA (Figure 4.11) is a $GABA_CR$ specific antagonist. Comparison of this ligand to bicuculline suggests they share some chemical similarities. The pKa of TPMPA is 9.9 (213), therefore should be protonated upon binding. The protonated amine should make a π-cation interaction which is maintained by bicuculline and the proteins it binds. The phosphinic acid group (pKa 3.5 (213)) is perhaps then able to form hydrogen bonding networks. However, the size comparison between the two molecules is different, TPMPA (MW 161 Da) is a fraction of the size of bicuculline (MW 367 Da). This emphasises the effect of increased bulk, noted above, in the pocket as the main driver that prevents bicuculline binding to $GABA_CR$.

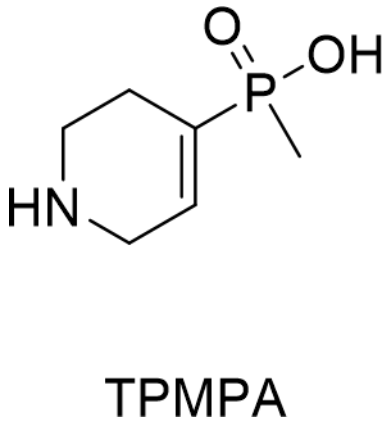


Figure 4.11 Structure of the $GABA_CR$ antagonist TPMPA.

4.8 Low resolution crystal structure of a *Ac*AChBP:N-methylbicuculline complex

A preliminary low resolution (2.9 Å) structure of the *Ac*AChBP:N-methylbicuculline complex was solved. Two pentamers make up the asymmetric unit with ten orthosteric sites. N-methylbicuculline was modelled into density observed in three of the pockets (associated with subunits A, D and J). The N-methylbicuculline molecule in chain D will be described as it possesses the lowest B-factor (61.5 $Å^2$) and best RSCC (0.88) (Figure 4.12), in comparison to molecules in chain A (B-factor 95.1 $Å^2$ and RSCC 0.78) and J (B-factor 73.9 $Å^2$ and RSCC 0.85). The ligand does not make any hydrogen bonds, instead, it forms a variety of van der Waals interactions with residues from both the principal (Y110, W164-Y166, Y205, C207, C208, E210 and Y212) and complementary (T53, Y72, Q74, R76, M133, I135 and D181) subunits (Figure 4.12A). Interestingly, the cation on N7 is positioned closer to Y110, instead of W164, which is

observed in the GBP:N-methylbicuculline complex. The cation sits in the electronegative pocket generated by Y110, W164 and Y212. One key difference between the GBP:N-methylbicuculline and *Ac*AChBP:N-methylbicuculline models is the amino acid that occupies the adjacent position to W164. In GBP, this residue is E162, whereas in *Ac*AChBP it is Y110. The different orientation adopted by N-methylbicuculline in the *Ac*AChBP model may be due to the repositioning of Y110 generating more space in the pocket, therefore the compound can bind deeper. A comparison between N-methylbicuculline bound to *Ac*AChBP and the other two proteins (GBP and $GABA_AR$) demonstrates a different orientation is adopted in each (Figure 4.12B). This clearly illustrates the promiscuous binding capabilities of N-methylbicuculline and the pLGIC family's binding site flexibility to accommodate it.

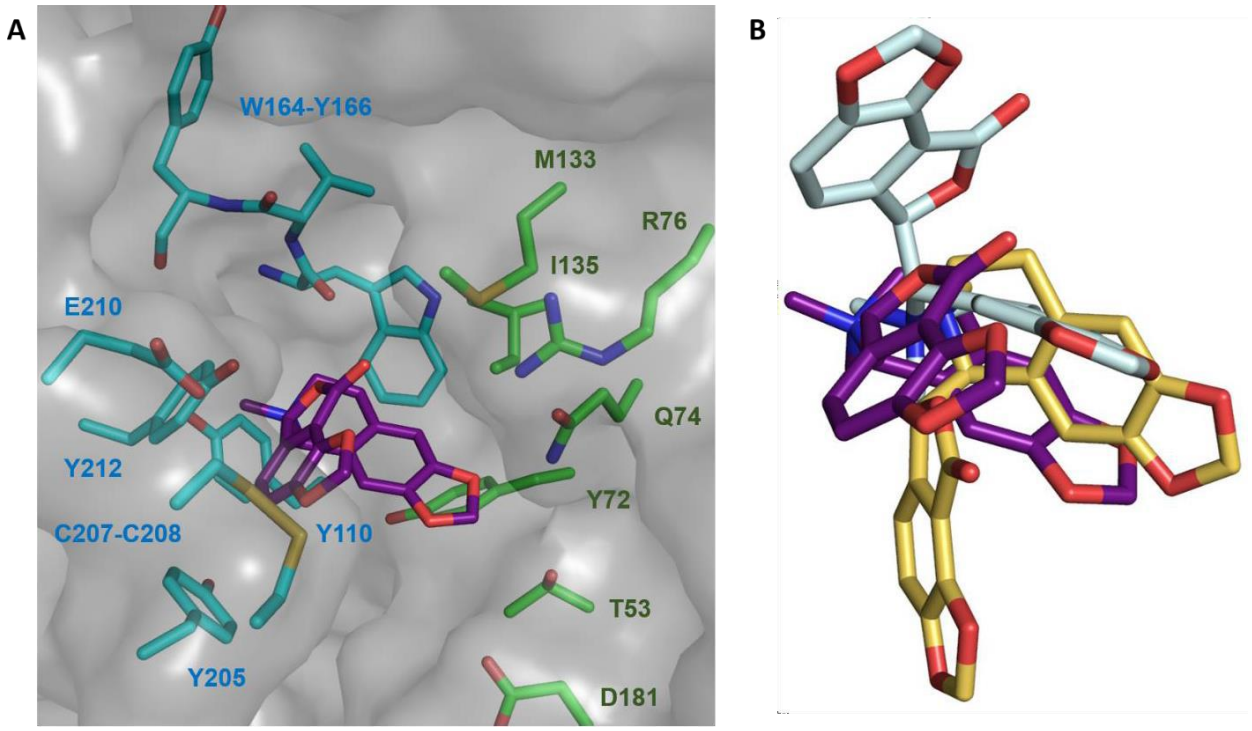


Figure 4.12 N-methylbicuculline binds differently in the *Ac*AChBP orthosteric binding site. (A) N-methylbicuculline (purple) is in the binding site made up of the principal (cyan) and complementary (green) subunits. (B) An overlay of the bicuculline molecules from *Ac*AChBP (purple), GBP pose I (silver) and $GABA_AR$ (yellow) after the principal subunits for each protein were superimposed (SSM) in Coot. Atoms are depicted as in Figure 4.3.

4.9 Summary

An analysis of the interactions between two key ligands of the inhibitory pLGICs and the model proteins, *Ac*AChBP and GBP, have been presented. The data suggest that different orientations of both strychnine and N-methylbicuculline can be adopted in the orthosteric binding sites of these proteins. In the absence of a GlyR heteromeric structure, GBP provides insights into how these ligands may interact with the receptor. Even though these compounds were utilised as receptor-specific chemical tools in the past they are promiscuous ligands due to the number of interactions they can form and the flexibility of the orthosteric binding site of the pLGIC family.

5. A screening campaign to identify new ligands for GlyR and nAChR, utilising the GBP and *Ac*AChBP surrogates

5.1 Introduction

Ligand screening is the starting point for many drug discovery projects where the objective is to find new chemical entities that interact with a target of interest. The number of compounds that can be screened has increased due to improvements in the detection of binding events and throughput (259). A chemical library can range from a few hundred compounds to a few million (85). Ruddikeit *et al.* published data on all the possible compounds that could be screened starting from one atom up to 17, using C, N, O, S and halogen atoms. In total their database consists of 166 billion compounds (260). From this example, the enormity of chemical space is observed and the potential for new ligands.

In the past, the rule of five was utilised as a set of parameters to select ligands to develop into drugs. These parameters are a list of chemical properties that are associated with increased absorption and permeation of orally bioavailable drugs and includes (261):

- MW of ≤ 500 Da
- Hydrogen bond donors (HBD) ≤ 5
- Hydrogen bond acceptors (HBA) ≤ 10
- A logP of ≤ 5

With the advent of new libraries of small molecules, a new paradigm was initiated in drug discovery that has led to the development of many new lead compounds (262). A new set of parameters were defined for these compounds following the rule of three (263):

- MW < 300 Da
- HBD ≤ 3
- HBA ≤ 3
- ClogP ≤ 3
- Rotatable bonds (RotB) ≤ 3
- Polar surface area ≤ 60

Due to the small size of the molecules, fewer interactions are predicted to be made, therefore more hits should be returned after the screening of a library. This outcome is linked with the increased ligand efficiency that small molecules possess in comparison to larger compounds thereby making them superior

starting points for extension (264). The small molecule libraries were also advantageous as more of chemical space could be sampled, in smaller collections (265).

Various biophysical techniques have been employed to conduct these screens, including SPR, X-ray crystallography and NMR (85,86,266). ^{1}H NMR allows for a flexible screening platform and is conducted in solution, therefore the native environment for binding interactions (93). It offers methods to detect ligand binding and for the calculation of affinity, by measuring the differences of one of three reported parameters, chemical shift, linewidth and intensity (88). However, as many compounds, proteins, buffers and reagents contain hydrogen, the complexity of a ^{1}H NMR spectrum limit its use (267). ^{19}F NMR, relative to ^{1}H, is a more recent screening platform, which is still being improved (268). Fluorine is not present in the repertoire of natural amino acids from which proteins are synthesised, therefore can be incorporated into potential ligands and so provides a label for NMR-based screening methods (269). The ^{19}F NMR spectra are therefore simpler than ^{1}H spectra due to fewer fluorine environments and the lack of coupling (267). Taking these factors into account ^{19}F NMR can provide a simpler platform for screening.

Another method for screening large compound libraries is computational docking (270). For docking, programmes mostly differ by what is defined as flexible (either the protein, ligand or both) and if a structure is available. Upon docking, the programmes can predict parameters, including how a compound binds to a target and what interactions it makes (271). Due to the speed and large quantity of data that docking can generate, dependent on computational resources, it can be utilised to inform and complement experimental methods (272).

Due to the essential role of GlyR and nAChR in human biology, modulation of these receptors is important to overcome relevant diseases, or for pain control and anaesthesia (228,273). The identification of new chemical matter that binds to GBP and *Ac*AChBP will be discussed, utilising both experimental and computational approaches. Using small molecule libraries, data from potential ligands will be collated to identify important ligand-protein interactions that could be used for compound development. These interactions will be compared to the physiologically relevant receptors to ascertain the validity of using GBP as a surrogate for GlyR and to provide information for nAChR ligand development.

5.2 NMR screening using an in-house compound library

The Drug Discovery Unit (University of Dundee) provided access to a library of 435 fluorine-containing compounds specifically assembled for screening by ^{19}F NMR. The library constituents were selected based on low mass, high solubility and potential for chemical modification (274). The library was screened against *Ac*AChBP and GBP. In these screens, we sought to identify new potential ligands, but the screening

also afforded the first opportunity to test this compound library and the ^{19}F NMR approach with two targets. Cocktails of nine to ten compounds were used, due to the simplicity of the ^{19}F spectra, which increased the throughput of the screen.

In ^{1}H NMR, large protein molecules tumble slowly (nanoseconds), resulting in fast relaxation and broad peaks in the spectrum (~100 Hz). However, as there are no fluorine atoms present on the protein these peaks are not detected in ^{19}F NMR. The small compounds, relative to the protein, tumble fast (picoseconds) leading to slow relaxation and sharp peaks (~1 Hz). If a potential ligand binds to the protein, it will adopt the relaxation of the protein leading to peak broadening (95). If the concentration of all the compounds in the cocktail are the same, the difference in peak intensity leads to the identification of binders. However, errors in compound concentration could occur due to evaporation of the solvent, precipitation of compounds and transfer of compounds from storage to the experimental setup. Owing to these potential errors, the peak intensities cannot be directly compared. Instead, competition assays can be carried out using a known ligand as a competitive binder. Upon the addition of this molecule to the NMR sample, it out competes any small molecule that may have bound, leading to an increase in peak intensity for the specific compound. For the described work in this thesis, strychnine was used, which has a known affinity for the two proteins, *Ac*AChBP (K_D 70-160 nM) and GBP (K_D 27-29 µM, Chapter 4.2). Using strychnine also allows for the identification of compounds that interact with the orthosteric binding site and not with other potential pockets. Compounds were screened at 30 µM, therefore an excess of strychnine was used (500 µM) to account for any molecules that had similar affinities.

In total 93 compounds were identified for the two proteins using the technique. Specifically, 21 and 59 unique compounds were identified for GBP and *Ac*AChBP respectively. An additional 13 compounds were identified as binding both proteins. An example of a potential compound binding is shown in Figure 5.1. This demonstrated that the fluorinated compound library could be screened by NMR and lead to hit identification.

5.3 Substructure analysis of the NMR hits seeking to determine similarities

To understand what type of compounds may bind to the proteins, and to seek out any common features a substructure analysis was conducted on the hits using the Molsoft ICM software (175,176). Compounds for the specific proteins were clustered based on similarities in their 2D structure by the unweighted pair

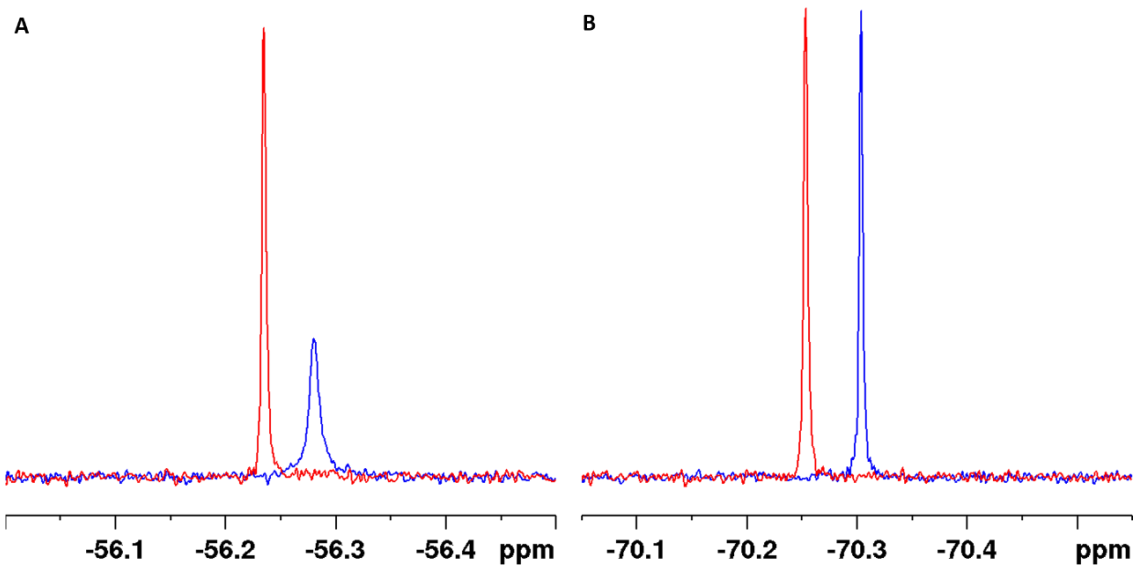
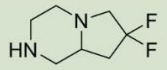
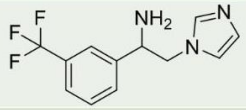
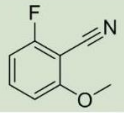
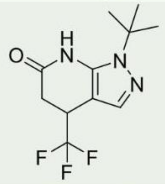
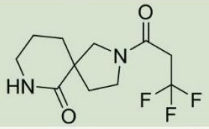
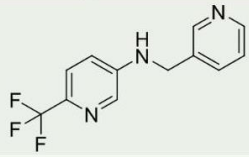
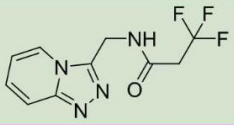


Figure 5.1 A comparison between four example ^{19}F spectra, with (red) and without (blue) strychnine, a competitive inhibitor of *Ac*AChBP and GBP. The spectra have been shifted so differences can be observed. A) Shows an increase in signal intensity of the red peak compared to the blue peak which shows small molecule binding in the orthosteric binding site of GBP. B) Shows no difference in intensity therefore no binding of the compound is inferred for *Ac*AChBP.

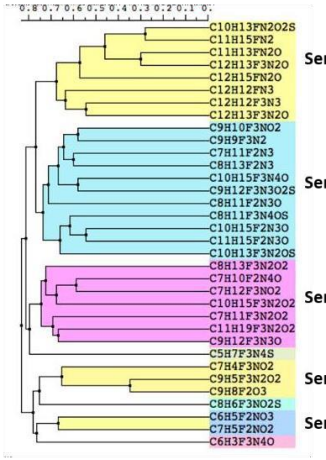
group method with arithmetic mean (UPGMA) algorithm. An arbitrary cluster threshold of 0.75 was used for both compound sets to limit the number of compound series produced, the distance between nodes indicates the level of similarity. Using the atomic property fields (APF) alignment method, the basic substructures could be identified. The APF tool identifies common physiochemical properties including lipophilicity or hydrogen bonding groups. It uses these properties to overlay the compounds by similarities, over many iterations (275). For the GBP hits, 34 compounds (21 unique and 13 non-specific) could be subdivided into five series with three unclassified molecules (Figure 5.2). *Ac*AChBP has a larger number of hits (59 unique and 13 non-specific) and these could be divided into seven series with four unclassified (Figure 5.2).

A range of potential substructures for the GBP hits were identified. All substructures consisted of a five- or six-membered ring, either aromatic or alkyl. Each ring possessed a hydrogen bond acceptor, donor or both as a substituent or incorporated into the ring itself. A comparison to *Ac*AChBP substructures found similar patterns, although slightly different groups. Perhaps due to conserved interactions with a few common residues shared between the protein binding sites, or adoption of similar orientations when bound. Differences in interactions are discussed later (Chapter 5.5.7.3), and these should be highlighted to develop an understanding of specificity.

A

Series	Pharmacophore	Example
1	Fused bicyclic ring, secondary amine	
2	6-membered aromatic ring, HBA	
3	6-membered ring, Methyl or CF_3, x2 HBA	
4	N,N-substituted 5-membered ring, HBD	
5	Alkyl ring, x3 HBA	
6	6-membered aromatic ring, N-substituted aromatic ring	
7	6-membered ring, amide	

B



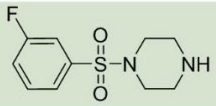
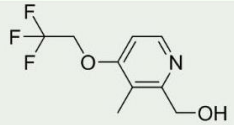
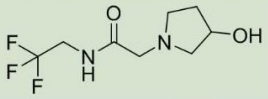
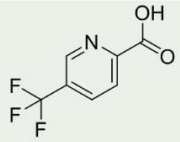
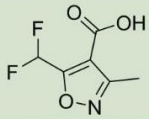
Series	Pharmacophore	Example
1	Phenyl, primary or secondary amine, tertiary amine	
2	N substituted aromatic ring, HBA	
3	N substituted ring, x2 HBA	
4	6-membered ring, carboxylic acid	
5	N-substituted aromatic ring, carboxylic acid, CF_2	

Figure 5.2 NMR small molecule screen hits clustered by structural similarities, using a generic cut-off of 0.75. The scale is indicative of compound similarity based on the common substructure. Each colour represents a different series, with the common structure of each detailed in the table and an example is given. A) *Ac*AChBP 72 hits and B) GBP 34 hits.

A follow-up BLI assay was trialled for the 93 compounds however, a sufficient signal to measure binding was not achieved. The assay was conducted using strychnine binding as a positive control and compounds binding to TEV protease as a negative control. After the first attempt, experiments were optimised utilising the control compound, strychnine. A second assay was conducted with the hit compounds, however, similar results were observed. Due to the lack of available material, no further experiments could be performed. Due to this reason, a subset of seven compounds were selected based on the substructure analysis and availability for further experimentation. Two compounds from two series and an unclassified molecule were chosen for *Ac*AChBP. Three compounds from different series were chosen for GBP. One compound that interacts with both proteins was also chosen, for comparative purposes (Table 5.1).

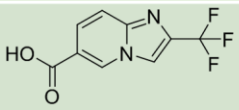
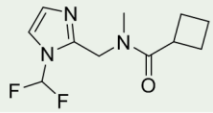
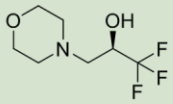
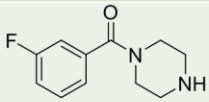
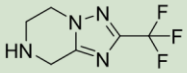
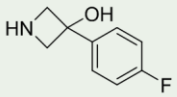
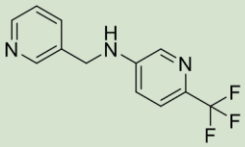
Compound	Structure	Protein	Series
1a		GBP	4
1b		GBP	2
1c		GBP	3
1d		GBP/AChBP	1/2
2a		AChBP	-
2b		AChBP	2
2c		AChBP	6

Table 5.1 Seven NMR compounds chosen to characterise further, showing the number, chemical structure, protein to which it is predicted to bind and the assigned compound series (Figure 5.2).

5.4 Attempted measurement of compound binding and calculation of affinity

Experiments were attempted to deduce the affinity of the compounds for their relevant proteins. Firstly, WF assays were considered. Due to the potential that the compounds themselves had intrinsic fluorescence each was tested (276). These data showed that all but one of the compounds were fluorescent in the relevant emission range (300-400 nm). For compound 1d, binding data were collected leading to the measurement of a K_D of 359 ± 4 μM and 393 ± 5 μM for *Ac*AChBP and GBP, respectively (Figure 5.3). Owing to the intrinsic fluorescence of the other compounds, this technique could not be used to measure their binding. SPR was also tested due to its common use in screening (185) however, a complete curve was not measured. Potentially, non-specific binding of the small molecules to the protein could have led to an incomplete curve due to changes in sensitivity or data analysis (277,278). Optimisation of these experiments was underway, testing new protein preparations and increasing the concentration of the compounds nonetheless, these were not completed due to time constraints. Initial ITC experiments demonstrated that standard methods, as conducted previously, were not able to elucidate K_D values of the small molecules. ITC has been shown to have a maximal threshold for the measurement of K_D values less than 250 μM (279).

Crystallisation trials were attempted with the seven selected hits from the NMR screen and the relevant proteins (GBP or *Ac*AChBP). Small crystals were obtained from sitting drop plates for multiple conditions. Initial diffraction data were collected and allowed for the selection of conditions for optimisation. However, due to time constraints optimisation was not completed and further data were not collected.

5.5 Computational docking of small molecules

Due to the ongoing optimisation of crystallisation of the protein-small molecule complexes, computational docking was conducted to investigate how they might bind to the relevant proteins. The docking could suggest information about the orientation of and interactions a compound would make when binding to the protein. This information could then be used in combination with experimental methods to develop molecules into drug-like compounds. However, these are only predictions and caution should be taken when interpreting the results.

5.5.1 Selecting GBP and *Ac*AChBP models for docking simulations

There are three GBP structures currently available in the PDB: complexes with glycine (resolution: 2.6 Å, PDB code: 5OAN), strychnine (resolution: 2.0 Å, PDB code: 5OBG) and N-methylbicuculline (resolution: 2.4 Å, PDB code: 5OBH). The Molsoft 4D docking tool was used to prepare a model of each structure and

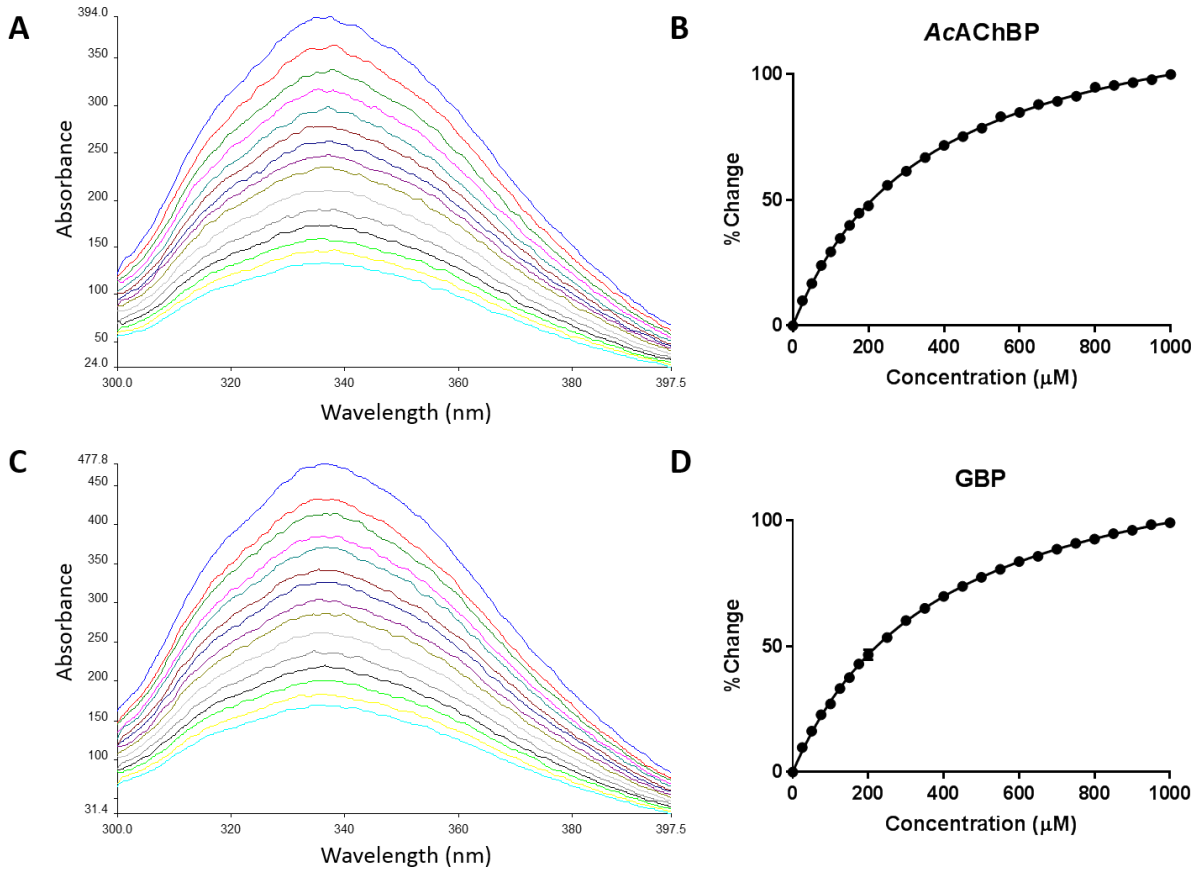


Figure 5.3 Binding of small molecule 1d monitored by WF. Example traces obtained from titrations, starting with no ligand (blue) to maximum ligand concentration (cyan) and a plot of the mean percentage change versus concentration is shown for *Ac*AChBP (A, B) and GBP (C, D). Standard error bars are behind each data point (n=3).

to align the binding sites; this accounts for conformational flexibility in the binding site ((280), Chapter 2.5.2). The compounds were then screened against each binding site in a single docking experiment. This identified the structure that forms the most favourable interactions with the compounds (280). Results from this docking suggested that all compounds preferentially bound to the GBP:strychnine structure. This structure was utilised for all further GBP computational docking experiments. Various structures were available for *Ac*AChBP with different ligands bound. However, 4D docking is most useful when structures with conformational changes are used (281), therefore three were chosen based on the ligand present, the resolution and analysis of the binding sites. The structures containing the compounds nicotine and strychnine were selected as they are well-documented ligands of nAChR. Also, these models would provide templates of both agonist and antagonist bound pockets, which differ in protein conformation. An *apo* structure was also deemed necessary as this could provide another alternate conformation of the pocket. Comparison of *apo* structures led to the identification of 2Y7Y as a good

model due to density fit and chain completeness. Again, 4D docking was conducted and two structures were selected based on what structure compounds preferentially bound: *Ac*AChBP:*apo* (resolution: 1.9 Å, PDB code: 2Y7Y) and *Ac*AChBP:nicotine complex (resolution: 2.2 Å, PDB code: 5O87). These two structures would allow for the deduction of differences in binding of a compound when interacting with an open, compared to a closed pocket.

5.5.2 Controls for docking simulations

Before initiating any docking simulations, the relevant ligands for the crystal structures utilised were re-docked. The compounds strychnine and nicotine were used for this purpose. For the GBP structure, no waters were included in docking as none contribute to strychnine binding. Upon binding strychnine to the pocket, an ICM score of -14.6 was achieved. This is the baseline score, which other compounds docked in this structure can be compared. A comparison between the re-docked strychnine molecule and the ligand in the original structure were in good agreement (Figure 5.4A). The ligand mostly makes van der Waals (VDW energy, -33.3 kcal mol^{-1}) interactions (Appendix B).

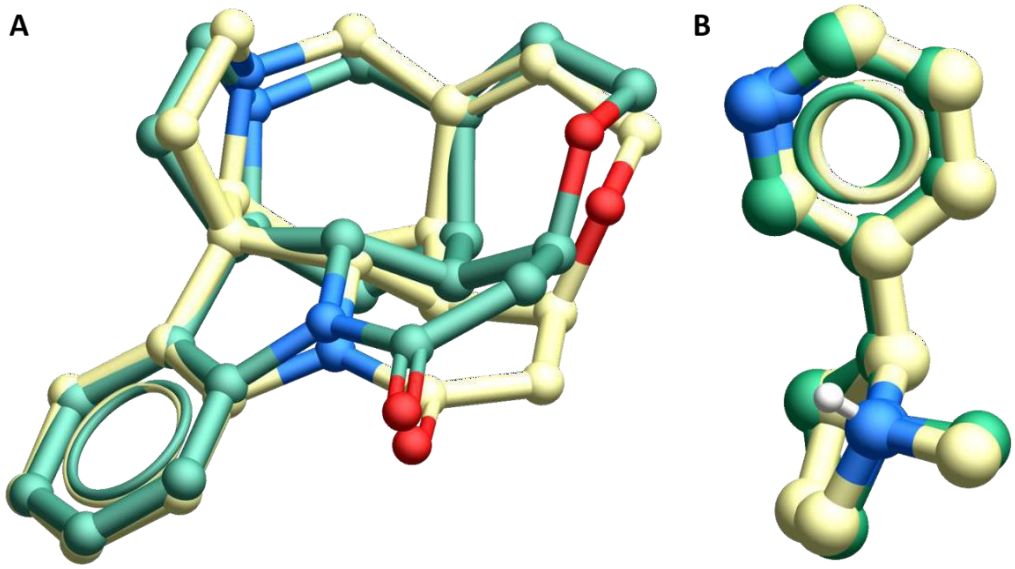


Figure 5.4 Docking controls. Images showing the consistency in original (yellow) and re-docked (green) compounds for docking simulation controls. A) GBP:strychnine (PDB code: 5OBG), largest difference 1.4 Å and B) *Ac*AChBP:nicotine (PDB code: 5O87), largest difference 0.3 Å.

Nicotine was docked into both the *Ac*AChBP structures. Docking into the *Ac*AChBP:*apo* structure was first conducted without any waters present. Nicotine was docked in a completely different orientation (Figure 5.5A) compared to the *Ac*AChBP:nicotine structure. A highly-ordered water is present in most structures and is important for ligand binding (75), as seen in Chapter 3.4.2. With this water present, the docked nicotine molecule was in a similar orientation as observed in the *Ac*AChBP:nicotine structure and an ICM

score of -23.6 was predicted. This water molecule was kept for both docking simulations. After re-docking nicotine into the *Ac*AChBP:nicotine structure, a comparison of the re-docked ligand and original nicotine molecule were in good agreement (Figure 5.4B). The ICM score for the re-docked nicotine was -28.4, which gives a baseline for future docking. There is a difference between the scores of nicotine binding to the two *Ac*AChBP structures potentially due to the interactions being made. In the *Ac*AChBP:nicotine structure, the pocket is smaller (189 $Å^3$), than *Ac*AChBP:*apo* (240 $Å^3$). The smaller pocket is probably due to the closing of the flexible loop C and may make more interactions as is observed in the difference between VDW energies (-16.4 to -23.5 kcal mol^{-1}, Appendix B).

5.5.3 Docking of the GBP NMR hits to predict binding interactions

The orthosteric pocket of GBP was determined to have a volume of 600 $Å^3$. Due to the conformation of the protein, the pocket extended beyond the cavity surrounded by the aromatic cage. For docking simulations, only the top part of the pocket, incorporating loops A-F, was utilised. In GlyR, strychnine is an antagonist and inhibits global conformational changes and pore opening (25). In a potentially similar manner, the binding pocket of GBP:strychnine should be an open conformation. Experimentally, upon small molecule binding the pocket may close, due to the size of the compounds, increasing the number and/or strength of interactions formed. This means that binding scores may be improved due to conformational changes not observed in the virtual screening. Protein flexibility was trialled, selecting five residues of the highly flexible loop C. However, the docking was computationally expensive, in comparison, to without protein flexibility. Due to this reason, this type of docking was not investigated further.

The hits identified as potential ligands for GBP (1a, 1b, 1c and 1d) are predicted to bind in two different orientations classified here as pose 1 (1a and 1b) and pose 2 (1c and 1d). The pose 2 compounds bind to one side of the binding pocket, which is made up of the residues of the principal subunit. From a comparison of the compounds bound in the orthosteric pocket (Figure 5.6A), similarities can be observed. The compounds consist of a six-membered ring (either piperazine or morpholine) and a fluorinated group; joined by an aliphatic linker with an oxygen-containing functional group. The six-membered ring occupies the top of the pocket making van der Waals interactions with T208 and E210, as well as MC of W164-S167. The fluorinated group is positioned at the bottom of the pocket between E162, Y205 and Y212, potentially making hydrophobic interactions. Due to the added flexibility of the linker of 1c, the hydroxyl group is predicted to make a weak hydrogen bond with S163 (2.1 Å). It is worth reiterating that there is a difference in hydrogen bonding distances between the docking models and crystal structures due to the

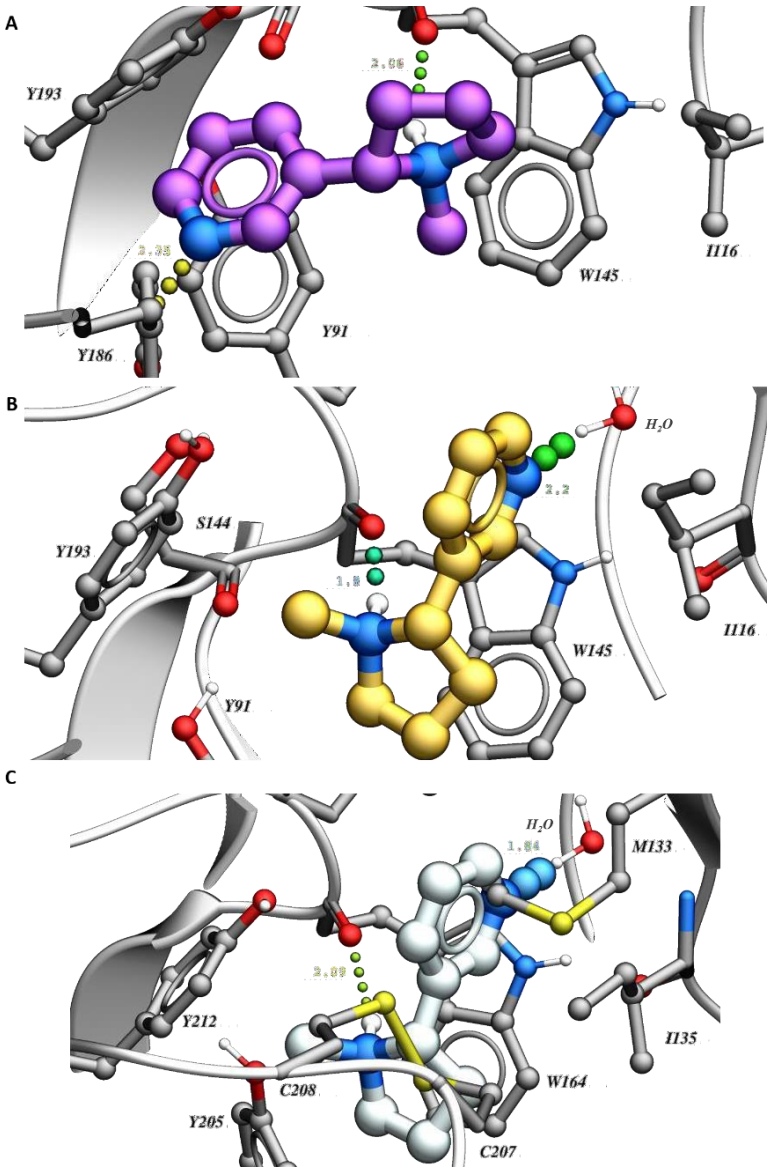


Figure 5.5 Images of nicotine bound in (A, B) *Ac*AChBP-*apo* (PDB code: 2Y7Y) or (C) *Ac*AChBP-nicotine (PDB code: 5O87). A) Shows nicotine binding in the *Ac*AChBP-*apo* pocket without the key water molecule present. Comparison of (B) and (C) shows that the re-docked nicotine molecule (yellow) with the water present in the *apo* pocket matches the original (silver) in the *Ac*AChBP-nicotine structure.

addition of hydrogen atoms (Chapter 2.5.2). The hydrogen bond with S163 contributes to the compounds hydrogen bonding (HBond) energy of -2.1 kcal mol^{-1}, as well as other weak interactions.

Compound 1d has a higher hydrophobic interaction energy (Hphob, -4.5 kcal mol^{-1}) than 1c (-3.9 kcal mol^{-1}), this is probably due to the fluoro-phenyl group. The major contributor for binding is van der Waals interactions for both compounds. 1d makes more interactions with a VDW energy of -21.5 kcal mol^{-1}, perhaps due to its planarity and rigidity, it can bind deeper in the pocket. As such 1d has a better ICM

score (-12.0) compared to 1c (-10.5) however, these are less than the score for strychnine (-14.6). This possibly suggests that these compounds are weaker binders. This difference between the ICM scores of the compounds and strychnine may be due to the greater size and the number of interactions strychnine makes, therefore increasing affinity. However, it has been noted that larger compounds have a low ligand efficiency in comparison to smaller molecules as not every atom plays a role in binding interactions (282).

The pose 1 compounds have the best binding scores (-16.2 to -16.9 for 1a and 1b), therefore would probably bind more efficiently to the protein. They also have a higher score than strychnine, so should make equally efficient interactions. Comparing the chemical structures, the two compounds are distinctly different. 1b is more flexible with a RotB score of 5, whereas 1c is planar with a RotB of 2. Even though there are chemical differences they bind to GBP in a similar orientation. They bind across the pocket making interactions with both the principal and complementary subunits. Both compounds are predicted to make a hydrogen bond with S135 however, these are estimated to be of different strengths. Possibly due to the potential to interact with more residues, the interaction scores are better for pose 1 compounds, in comparison with pose 2. Comparable to the pose 2 compounds, the fluorine groups of pose 1 compounds occupy a similar region between Y205 and Y212. As well as hydrogen bonding, the pose 1 compounds have an increased hydrophobic interaction. 1b has the highest Hphob energy (-5.7 kcal mol^{-1}). Perhaps this is due to the addition of extra hydrophobic groups such as the methyl and cyclobutane. A variety of van der Waals interactions are also made with R74, M133, W164-167, E162, Y205, T208, E210 and Y212, giving an average predicted energy of -20.0 kcal mol^{-1} for both pose 1 compounds.

5.5.4 Docking of *Ac*AChBP NMR hits to identify differences between *apo* and closed states

A closed ligand-bound and an *apo* structure were selected as templates for the docking calculations. A comparison between the docking outcomes to both structures may suggest if there are any major differences between the open and closed states.

Firstly, compounds (2a, 2b, 2c and 1d) were docked to the *Ac*AChBP:nicotine structure. All four compounds bind in a similar orientation to each other, close to W164. This is possibly due to the smaller pocket size and steric restrictions. Unlike GBP, the *Ac*AChBP binding pocket is restricted by the presence of Y110 and the Cys loop (C207-C208). The fluorine-containing groups occupy comparable positions between M133 and V165, possibly contributing to hydrophobic interactions. The lower portion of the compounds are either small or planar making interactions with Y110, Y205 and Y212. From the

predicted binding energies the main interactions are hydrophobic and van der Waals. All the ICM scores (-6.2 to -13.0) for the compounds are below that of nicotine (-28.4). This is potentially because the compounds appear to lack the capacity to form the hydrogen bonding contributions to binding that nicotine possesses. Also, these compounds are lacking the π-cation interaction, which has been highlighted in Chapter 3 as important for binding to *Ac*AChBP.

The same compounds were next docked into the *apo* structure. All four ligands were orientated differently, in comparison to the poses predicted using the *Ac*AChBP:nicotine structure as a template, most likely due to the extra space in the pocket. The extra space has possibly also led to fewer interactions being formed for most of the compounds, noted by the predicted energies. This also affects the ICM scores for 2a (-8.4) and 2b (-6.8), which have deteriorated. However, the scores for 1d and 2c have improved. The score for 1d (-16.1) is still lower than the value for nicotine (-23.6), but 2c (-22.9) is comparable. This compound (2c) is larger (MW 253 Da) than the others (MW 167-208 Da) and can spread out in the space provided. Owing to this, favourable interactions appear to have been made with Q36, Y53, Y91, W145, S165, Y186 and Y193. Hbond energy (-4.8 kcal mol^{-1}) contributes to the binding score with two optimal interactions that can be seen in Figure 5.6C. Hphob (-4.8 kcal mol^{-1}) and VDW (-21.2 kcal mol^{-1}) energies also contribute more, in comparison to nicotine binding (Appendix B).

5.5.5 Maybridge rule of three library screen to find new ligands

The availability of libraries containing large numbers of molecules, displaying diverse structures and the relative ease of computational docking calculations can inform structure-driven compound design approaches. The Maybridge rule of 3 diversity set of 2,000 compounds (Maybridge.com, (263)) is available in a relevant format and was deemed appropriate to computationally screen against GBP and *Ac*AChBP using the same approaches employed with the NMR hits. For each protein, all 2,000 compounds were screened at a thoroughness of 3, sampling racemic centres. Again, various conformations of each compound are deduced. A selection of the conformations are then screened against the pocket in different orientations and locations. The best scoring conformation, based on the interactions it makes is presented in the hit list (176). The same protein structures for the NMR hits were used for screening this library. After docking, the top 50 hits were ranked by ICM score and separated into series by substructure similarity. This was carried out to observe any related characteristics that bind these proteins. The top ten compounds were analysed further to understand how they may bind to the proteins.

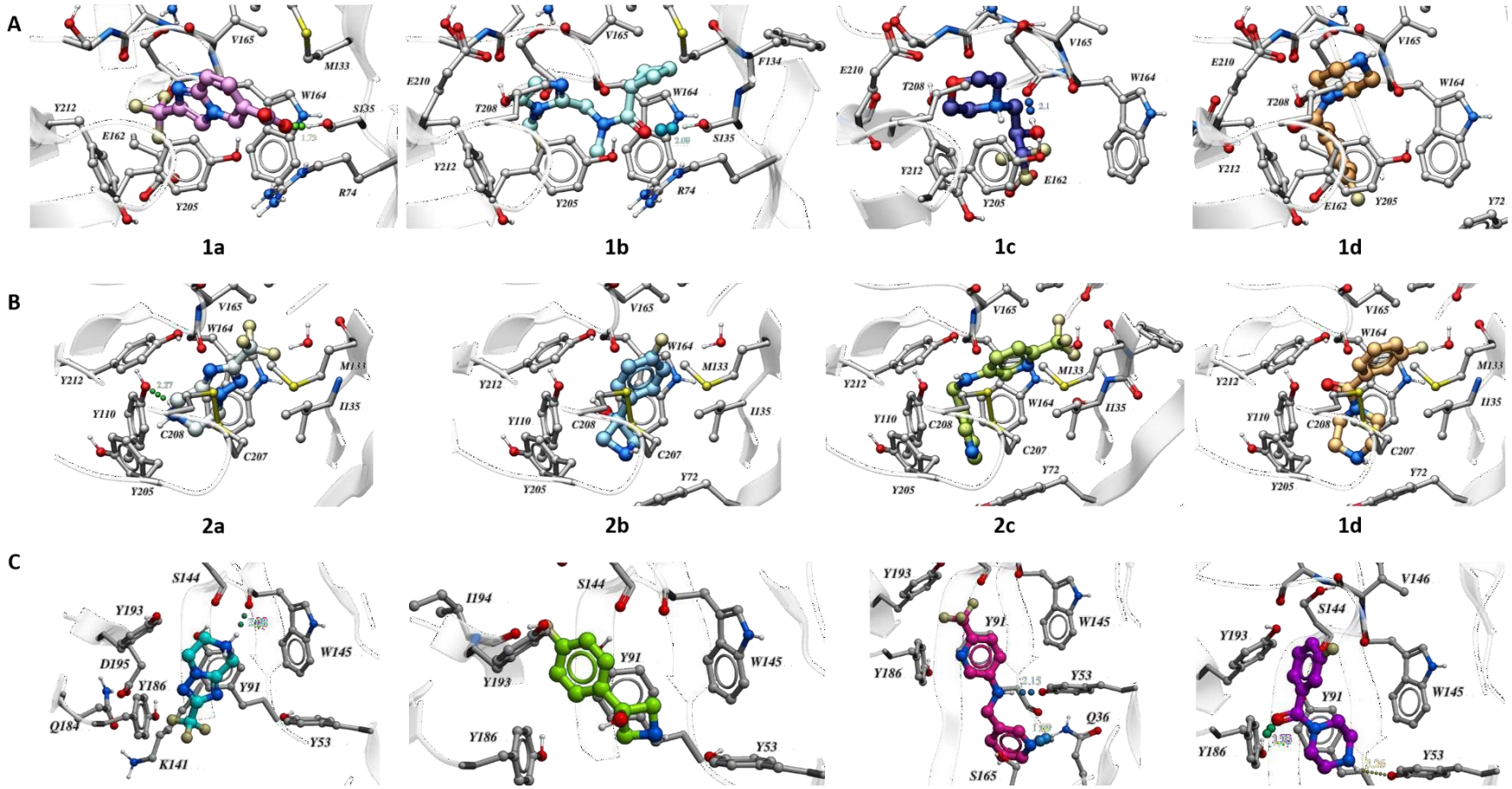


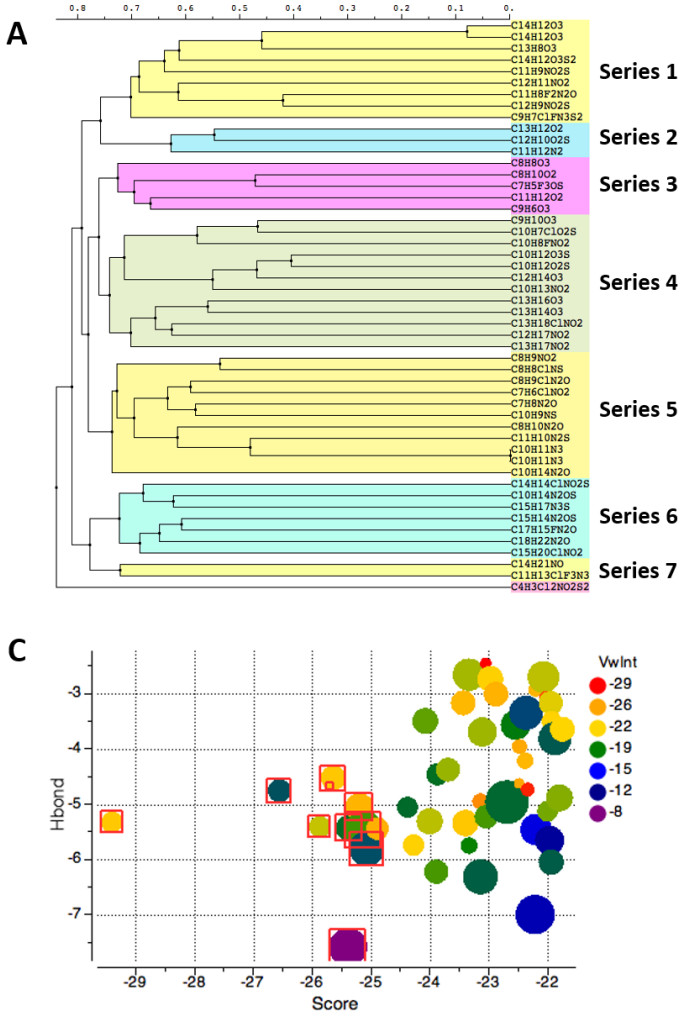
Figure 5.6 Selected hits from the NMR screen (Table 5.1) docked into the relevant target. A) GBP (PDB code: 5OBG), B) *Ac*AChBP:nicotine (PDB code: 5O87) and C) *Ac*AChBP:*apo* (PDB code: 2Y7Y). Small molecules (coloured, sticks) are shown bound in the orthosteric binding pocket of the protein (silver, ribbon). Residues (silver, sticks) important for binding are shown and labelled. Key atoms are coloured hydrogen (white), oxygen (red), nitrogen (blue), sulfur (yellow) and fluorine (pale yellow). Hydrogen bonds are shown as coloured spheres using size to indicate the predicted strength and the colour (red-blue) to indicate distance.

5.5.5.1 GBP Maybridge library hits – how do they bind?

The top 50 compounds predicted to bind GBP were classified into seven series by common substructure (Figure 5.7). One compound could not be classified into any series and does not share the common aromatic ring. There is a variation in ICM scores for the small molecules between -22 to -29 (Figure 5.7C), which are all higher than the ICM score for strychnine binding to the same protein (-14.6). This supports the use of small molecules in screening as they make specific, favourable interactions and may have a higher ligand efficiency in comparison to larger compounds. The top ten compounds (red boxes) are from different series ranging from 1-6 and the predicted binding energies are shown in Appendix B.

The common substructure for series 1 is shown in Figure 5.7. The three compounds of the top ten that are in this series are 3a, 3b and 3c. Two of the compounds (3a and 3b) are similar differing only in the position of the ether group. The fact that both these compounds were high-ranking hits may suggest that the common substructure is important for binding. A comparison of the predicted binding pose showed that they occupy a similar footprint in the binding site of GBP. As such they make comparable interactions, which are shown in their binding energies: Hbond (-5.3 to -5.4 kcal mol^{-1}) and Hphob (-4.8 kcal mol^{-1}). There is a difference in the VDW energy (-21.4 to -23.2 kcal mol^{-1}), which is possibly due to the difference in the ether group placement. This difference leads to a positional change of the phenyl ring and possibly increased interactions with the MC of V165-S167 for 3b. Due to this difference, the ICM scores are dissimilar (-29.4 for 3b and -25.9 for 3a) however, these are both still regarded as good hits. The third molecule (3c) also shares the presence of a carboxylic acid group. For the other two compounds, this group is predicted to interact with the R74 SC and is conserved by other top ten compounds. However, perhaps due to the rigidity (RotB 2, compared to 4 for 3a/b), 3c occupies a different position in the binding site. As such there are differences within the binding energies compared across the three compounds: Hbond (-4.7 to -5.4 kcal mol^{-1}) and VDW (-17.5 to -23.2 kcal mol^{-1}). Conceivably, the difference in orientation could cause a change in van der Waals interactions and lower the ICM score for 3c (-26.6).

As well as the common series substructure, the two compounds (3d and 3e) in series 2 also share other similarities. Both compounds consist of two phenol groups connected by a one-atom linker. The differences between these compounds are the position of one of the hydroxyl groups on a phenyl ring (*ortho* or *para*) and the linker atom (carbon or sulfur). The main difference is observed for the predicted VDW energy (-23.3 to -24.5 kcal mol^{-1}). This may be due to the additional interaction with R74 and/or due to the size of the linker atom (atomic radius S 1.8 Å and CH_2 2.0 Å). Even with these structural differences both compounds are positioned similarly in the binding site and have comparable ICM scores (-25.2 to -25.7).



B

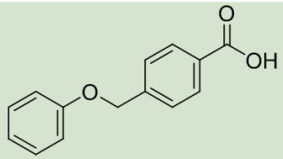
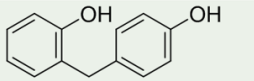
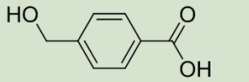
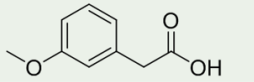
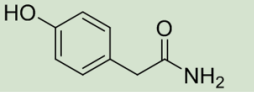
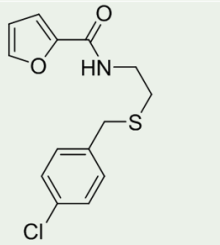
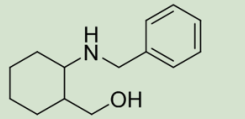
Series	Pharmacophore	Example
1	Phenyl, aromatic ring, HBD, x2 HBA	
2	Phenyl, aromatic ring, HBD	
3	Phenyl, hydroxyl	
4	Phenyl, two oxygen groups	
5	Phenyl, primary amine	
6	6-membered aromatic ring, x3 HBA	
7	6-membered aromatic ring, 6-membered ring	

Figure 5.7 Analysis of the top 50 Maybridge screening hits for GBP (PDB code: 5OBG). A) All 50 compounds clustered by common substructure (different colours), using a 0.75 cut off. B) Table of the common substructures of each series in (A) and an example. C) Comparison of ICM score (x-axis), hydrogen bonding (y-axis), van der Waals (colour) and hydrophobic interaction (size) energies for the top 50 hits. The top ten compounds are shown by red boxes.

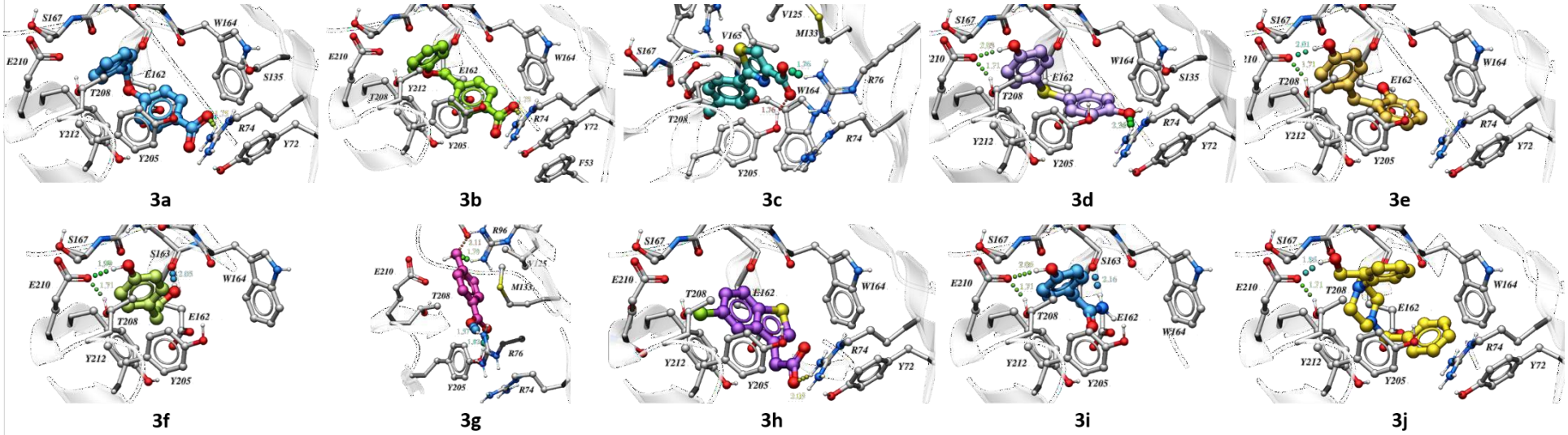


Figure 5.8 Top ten GBP hits from the Maybridge screen docked to GBP (PDB code: 5OBG). The same colour scheme and depiction as used in Figure 5.6 is used here with the addition of chlorine atoms (pale green).

Even though there is a similar substructure for series 3, the two compounds in the top ten bind in different positions. Due to their small size, they only bind predominantly to one subunit, either principal (3f) or complementary (3g). The two compounds make a variety of different interactions observed in the models and the predicted energies (Hbond -5.8 to -7.6 and VDW -8.4 to -17.3). Compound 3f is predicted to bind similar to series 2 compounds and makes an additional hydrogen bond with the MC of S163. Compound 3g is oriented in a similar position to the thiazole of 3c interacting with R76. However, 3g occupies a position higher in the pocket towards R96 and V125. The ICM scores for both compounds are comparable (within 0.3), which may suggest that increased hydrogen bonding energy of 3g may compensate for its low van der Waals energy.

The remaining three compounds are from three different series: 3h (series 4), 3i (series 5) and 3j (series 6). They are all predicted to occupy a similar position to series 1 compounds, and form interactions observed by other molecules of the top ten. Of note, 3j has the highest Hphob (-7.0 kcal mol^{-1}) and VDW (-25.4 kcal mol^{-1}) interaction energies. Similar to strychnine this is probably due to the size of the compound (MW 282 Da) in comparison to the others (MW 138-227 Da). However, again like strychnine, the extra size does not increase its ICM score (-25.7), which suggests low ligand efficiency.

Computational docking of the Maybridge library to GBP led to the identification of several small molecules that bind in the orthosteric site. Structural similarities between the compounds provide clues about which chemical features might be important for binding. Specific protein-ligand interactions were identified from comparisons of these compounds. This information could be coupled with other docking and experimental analyses to inform compound development.

5.5.5.2 Computational docking of the Maybridge library to *Ac*AChBP:nicotine

The top 50 hits for *Ac*AChBP:nicotine structure could be classified into six series with five compounds not fitting into any of the groups (Figure 5.9). The top ten compounds have an ICM score range of -23.0 to -29.6, which is comparable to the nicotine value of -28.4. These compounds are from four of the six series with one unclassified molecule.

All of the top ten compounds are mostly planar with limited flexibility (RotB 0-2). All compounds are predicted to bind in a similar position to nicotine in the *Ac*AChBP:nicotine pocket and make an interaction with the well-ordered water molecule (Chapter 3.4.2).

There are three compounds from series 1 in the top ten. As well as the common substructure, these specific examples all have an amine group that extends between M133 and V165. For two compounds (4a

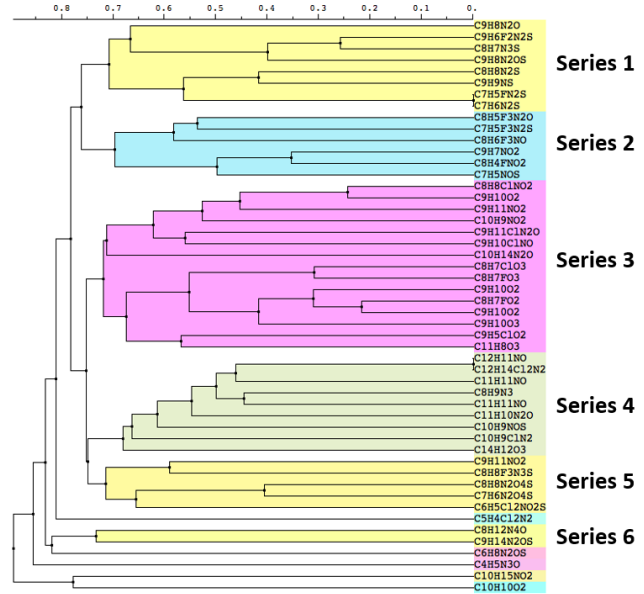
and 4b) this amine group makes the hydrogen bond with the key water molecule in the binding site. For 4c and 4b, the nitrogen atom in the ring would form a hydrogen bond with this water. For each of the three compounds, the phenyl group is positioned within the aromatic cage formed by Y72, W164, Y205 and Y212 (Chapter 1.1.2.2). Compounds 4a and 4c are similar when comparing ICM scores (-25.6 to -26.5), however, the main difference is the VDW energy (-23.5 to -25.0 kcal mol^{-1}). This is due to 4a extending towards Y110 and making more interactions. 4b has the highest ICM score (-29.6), possibly due to the highest Hbond energy (-9.3 kcal mol^{-1}). This is due to the variety of hydrogen bonds it makes to Y212, MC M133 and a potential bidentate interaction with the water molecule. 4d is the only compound from series 2 featured in the top ten. This compound shares many of the interactions previously discussed in series 1. As such it has a comparable ICM score (-25.4) to 4c (-25.6).

Series 3 comprises four of the top ten compounds. The substructure of these four compounds is almost identical thus they make similar interactions. The carbonyl substituent of all four molecules makes a strong (1.5-1.6 Å) hydrogen bond to the water molecule. A hydrogen bond is also potentially formed with Y212. The Hbond (-4.9 to -5.1 kcal mol^{-1}) and Hphob (-3.9 to -4.0 kcal mol^{-1}) energies, for the four compounds, are similar. Nonetheless, the VDW energy for compounds 4e and 4f (-24.1 to -24.9 kcal mol^{-1}) are better than the other two compounds, 4g and 4h (-21.4 to -21.6 kcal mol^{-1}). This suggests that either the ester and/or chlorine substituents for the two former compounds generate more favourable interactions.

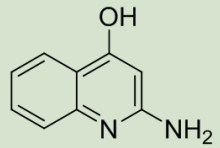
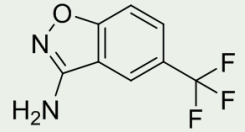
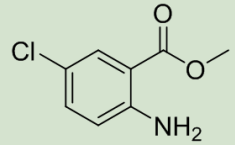
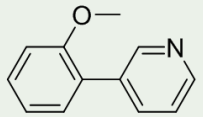
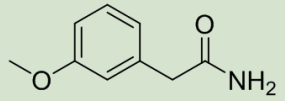
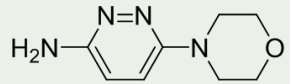
The remaining compounds are 4i (series 6) and 4j (unclassified). Both compounds form interactions already discussed (Figure 5.10). Of note is 4i, which is comparable to 4a, as it extends towards Y110 and makes more interactions leading to an increase of the VDW energy (-26.7 kcal mol^{-1}) and ICM score (-27.5) in comparison with other compounds (Appendix B).

In summary, the ten planar compounds docked as potential ligands into the *Ac*AChBP:nicotine complex all share common features as well as with nicotine, which suggests similar interactions with the template. One of the main interactions with the well-ordered water molecule that nicotine makes is conserved by all the compounds. However, it may prove difficult to treat these compounds as templates for chemical modification due to the small ligand binding pocket noted when nicotine is present.

A



B

Series	Pharmacophore	Example
1	Phenyl, pyrrole or pyridine	
2	Phenyl, secondary amine	
3	Phenyl, carbonyl	
4	6-membered ring, ring, x2 HBA	
5	Phenyl, EWG, primary amine	
6	Aromatic ring, primary amine, lipophilic substituent	

C

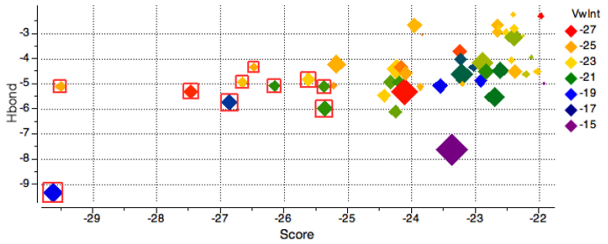


Figure 5.9 Analysis of the top 50 Maybridge screening hits for *Ac*AChBP:nicotine (PDB code: 5O87). A) All 50 compounds are clustered by common substructure (different colours), using a 0.75 cut off. B) Table of the common substructures of each series in (A) and an example. C) Comparison of ICM score (x-axis), hydrogen bonding (y-axis), van der Waals (colour) and hydrophobic interaction (size) energies for the top 50 hits. The top ten compounds are shown by red boxes. EWG stands for electron-withdrawing group.

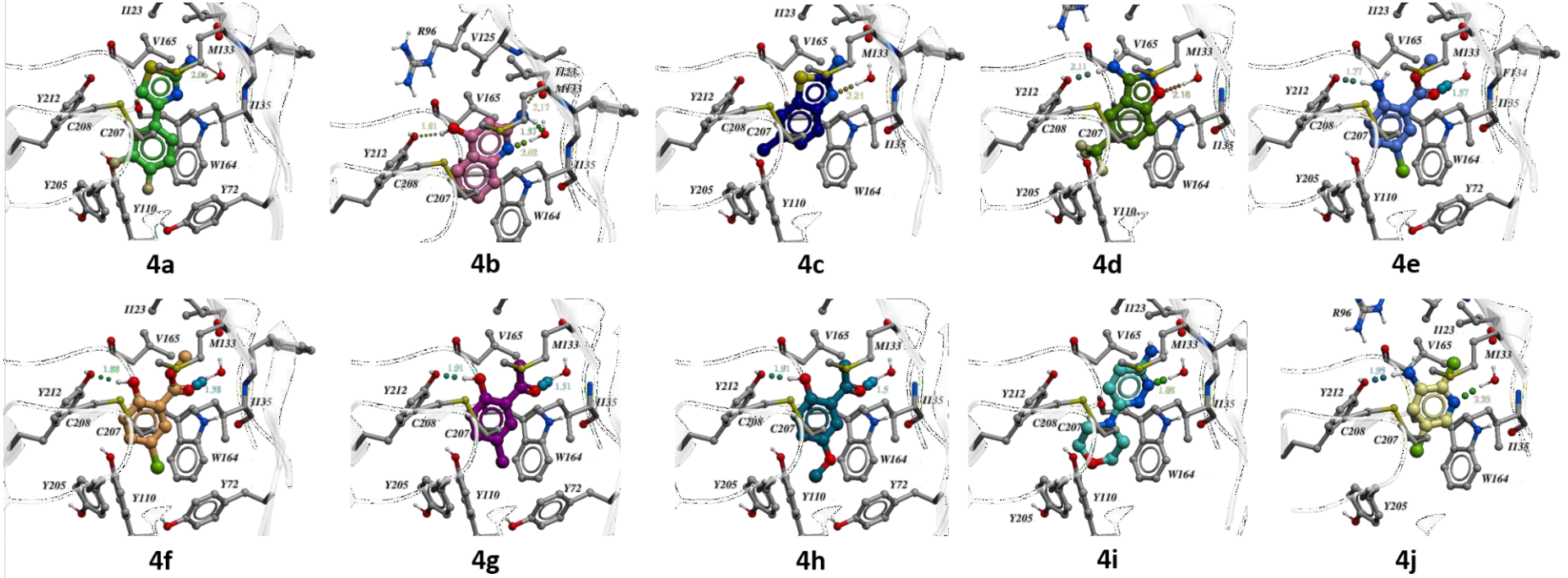


Figure 5.10 Top ten *Ac*AChBP hits from the Maybridge screen docked to *Ac*AChBP:nicotine (PDB code: 5O87) using Molsoft ICM. The same colour scheme and depiction as used in Figure 5.8 is used here.

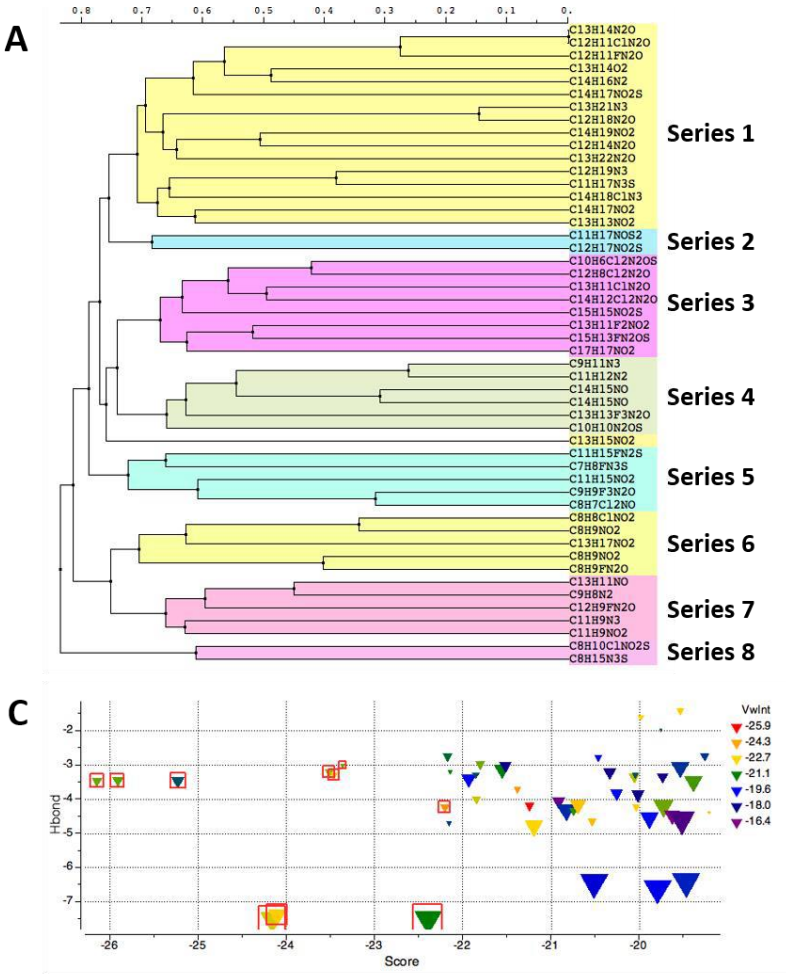
5.5.5.3 Different Maybridge compounds are predicted to bind *Ac*AChBP:*apo*

In comparison to hits for the *Ac*AChBP:nicotine complex, the top 50 hits for the *apo* structure were classified into eight series with one outlier (Figure 5.11). Like hits for GBP (RotB 2-4), there are flexible groups (RotB 2-6) within each of the series. The top ten ICM scores range from -22.2 to -26.2, which is comparable to the value for nicotine binding of -23.6. The top ten compounds, which are dissimilar to compounds that are bound to *Ac*AChBP:nicotine, come from four different series.

Four compounds from the top ten come from series 1 and three of them are the highest-ranking hits (ICM scores -25.2 to -26.2). These three compounds (5a, 5b and 5c) are also homologues with one change (CH_3, Cl or F). The fourth compound is the lowest ranking of the top ten. The three homologues occupy identical positions within the binding site, similar to nicotine. A strong hydrogen bond is made between the carbonyl and water (1.9-2.0 Å). The chemical differences between the compounds only affect the VDW energy (-20.5 to -21.8 kcal mol^{-1}). This is probably due to a change in the size of the substituents and interactions made; methyl (VDW radius 2.0 Å) and chlorine (VDW radius 1.8 Å) groups are larger than fluorine (VDW radius 1.5 Å) which is more similar to hydrogen (VDW radius 1.2 Å) (283). The other compound (5d) takes up a different position binding in a cavity between loop C and the principal subunit. Here, the VDW (-24.5 kcal mol^{-1}) and Hphob (-5.1 kcal mol^{-1}) energies are higher than the other three compounds (Appendix B). The Hbond energy is also higher for 5d (-4.2 kcal mol^{-1}). It makes one hydrogen bond between the nitrogen of the pyridine ring and Y186, but perhaps due to tight binding in the cavity, weaker interactions can be made. Even though it seems this molecule makes more interactions, the ICM score is lower (-22.2) possibly due to other factors considered to calculate the score e.g. desolvation energy.

The three compounds: 5e (series 3), 5f (series 3) and 5g (series 2) are predicted to bind in a similar position to nicotine and interact with the key water molecule. The compounds also extend towards Y91 forming additional interactions. Due to the similarities, all three compounds have comparable ICM scores (-23.4 to -23.5), Hbond (-3.0 to -3.3 kcal mol^{-1}), Hphob (-5.1 to -5.8 kcal mol^{-1}) and VDW (-21.8 to -23.4 kcal mol^{-1}) energies.

The final group of compounds to be discussed is series 6, with three compounds in the top ten. Similar to series 2, the compounds occupy the cavity between loop C and the principal subunit. Two compounds (5h and 5j) are homologues with a chlorine atom being the only difference. The acetamide substituent makes three MC hydrogen bonding interactions (S144, I152 and I194) leading to a similar Hbond energy (-7.5 kcal mol^{-1}). The presence of the chlorine improves the VDW (-21.1 to -22.4 kcal mol^{-1}) energy,



B

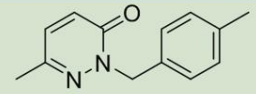
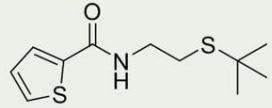
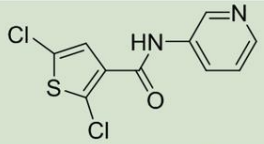
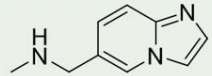
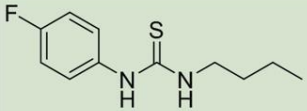
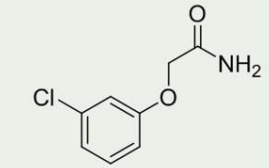
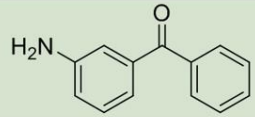
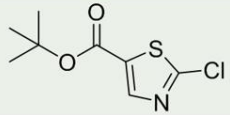
Series	Pharmacophore	Example
1	Aromatic ring, HBA	
2	5-membered aromatic ring, amide	
3	6-membered aromatic ring, aromatic ring, carbonyl	
4	6-membered ring, HBA, HBD	
5	Phenyl, secondary amine, CO or CS	
6	Phenyl, HBA, amide	
7	Phenyl, aromatic ring, HBA	
8	Thiazole, HBA, tertiary butyl	

Figure 5.11 Analysis of the top 50 Maybridge screening hits for *Ac*AChBP:apo (PDB code: 2Y7Y). A) All 50 compounds clustered by common substructure (different colours), using a 0.75 cut off. B) Table of the common substructures of each series in (A) and an example. C) Comparison of ICM score (x-axis), hydrogen bonding (y-axis), van der Waals (colour) and hydrophobic interaction (size) energies for the top 50 hits. The top ten compounds are shown by red boxes.

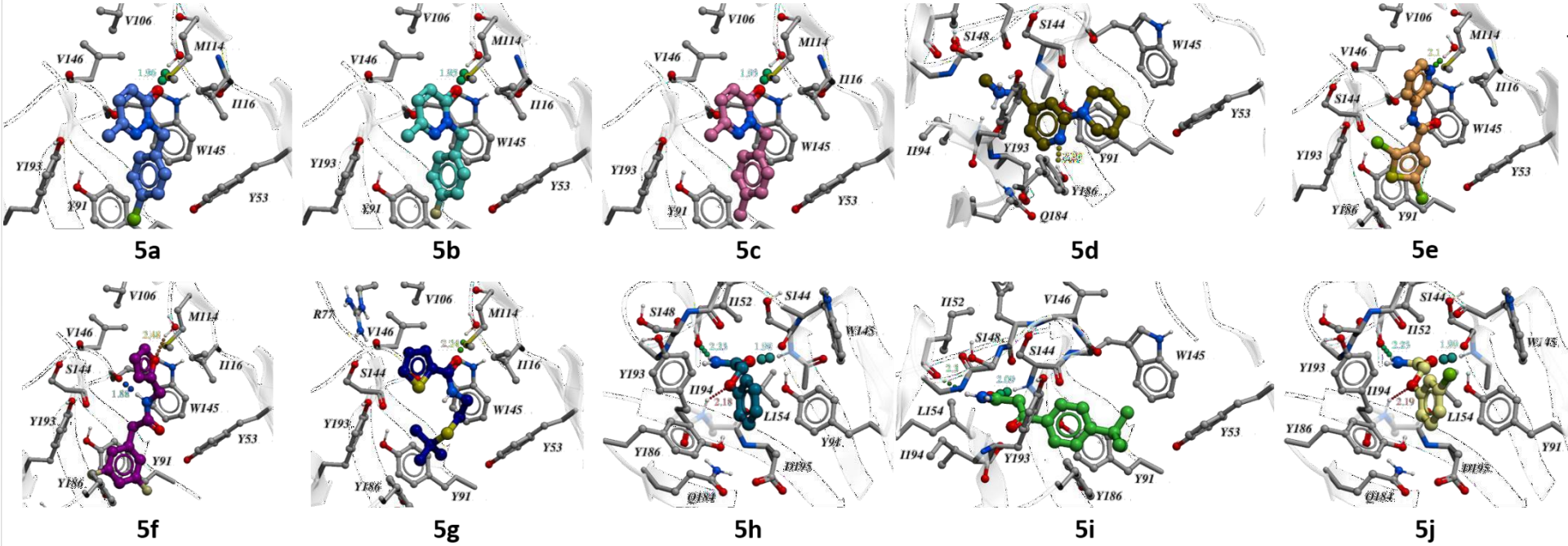


Figure 5.12 Top ten *Ac*AChBP hits from the Maybridge screen docked to *Ac*AChBP:*apo* (PDB code: 2Y7Y) using Molsoft ICM. The same colour scheme and depiction as used in Figure 5.8 is used here.

leading to an enhanced ICM score of -24.2, compared to -22.4. 5i has two differences, compared to 5h, a carbonyl instead of an ether and an additional ^{t}Bu substituent. Due to the addition of the carbonyl the Hbond energy decreases (-7.4 kcal mol^{-1}). However, the addition of the ^{t}Bu group improves the Hphob (-4.0 kcal mol^{-1}) and VDW (-23.0 kcal mol^{-1}) energies, which places 5i equivalent in ICM score to 5j (-24.1 to -24.2). Consequently, this suggests that the additional substituents on the phenyl ring are needed to improve binding.

From this set of compounds, there are two main sites of interaction in the *apo* binding pocket: 1) similar to nicotine in proximity to W145 and 2) in the cavity between loop C and the principal subunit. Similar to nicotine binding, the compounds that bind in position 1 make contacts with the key water molecule, after that most other interactions are hydrophobic or van der Waals. For position 2, the orientation of the ligand may be stabilised by a few hydrogen bonds with MC groups and close interaction with other residues. These observations coupled with those from the closed pocket screen could be used to inform ligand design against *Ac*AChBP and in the long term nAChRs.

5.5.6 Validation of computational screening via K_D determination

The highest ranking compounds, 3b and 4b, from the computational screen with GBP and *Ac*AChBP:nicotine were obtained to validate the results. Compounds from the screen with *Ac*AChBP:*apo* could not be acquired due to availability. Similar to the NMR compounds, WF was conducted to determine K_D values. The intrinsic fluorescence of each compound was first checked. 4b was identified as fluorescing in the relevant range, whereas 3b did not. WF experiments with 3b gave a K_D of 237 ± 5 µM (Figure 5.13). Due to time constraints, no other biophysical assays could be conducted. Comparing the K_D and ICM score for 3b may allow for the prediction of the affinity of other docked compounds i.e. similar ICM scores may relate to mid micromolar affinity.

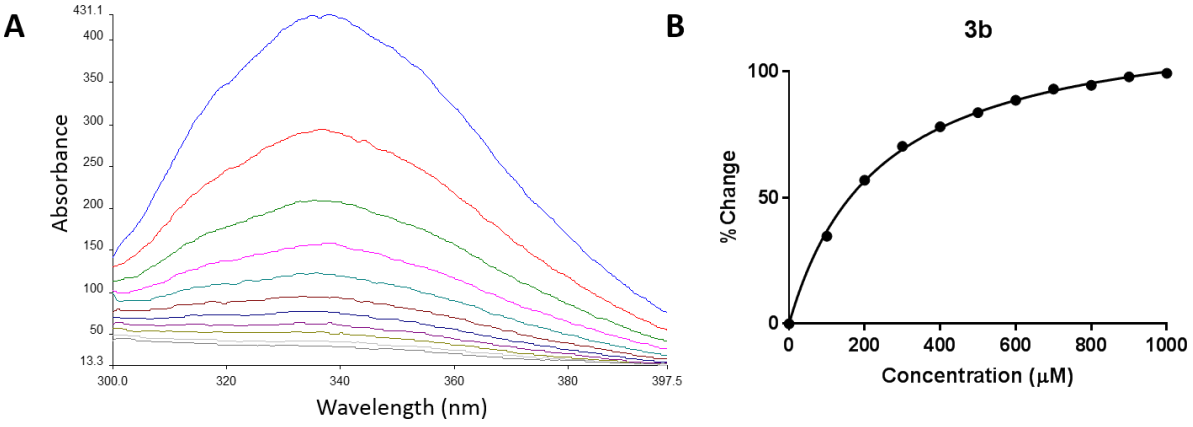


Figure 5.13 Binding of small molecule 3b to GBP monitored by WF. An example trace obtained from titrations, starting with no ligand (blue) to maximum ligand concentration (grey) (A) and a plot of the mean percentage change versus concentration (B) are shown. Standard error bars are behind each data point (n=3).

5.5.7 Are there differences between the hits for different proteins?

For GBP and *Ac*AChBP, 14 and 24 new small molecules have been identified respectively that potentially bind in the orthosteric pocket. A comparison of the molecules may elucidate structure-ligand binding relationships that could be combined with experimental methods to aid in the design of new drug-like compounds.

5.5.7.1 Comparison of NMR and Maybridge small molecule hits

As previously modelled in GBP NMR hits, there are two positions where compounds are predicted to bind, pose 1 and pose 2. This observation was also noted in the Maybridge top ten hits for GBP. Seven out of the ten compounds sit across the binding site in pose 1. Perhaps a key difference is the presence of a large phenyl group for the Maybridge compounds which may make more interactions with the principal subunit. In comparison, the NMR molecules have smaller groups, such as a CF_3 or imidazole, which could make fewer interactions. Possibly due to these interactions, the Maybridge ICM scores are better (-25.1 to -29.4) than the NMR compounds (-10.5 to -16.9). Pose 2 is adopted by compounds 1c, 1d, 3f and 3i, which make similar interactions. However, possibly due to the rigidity of the phenyl ring, in comparison to the flexibility of the morpholine/piperazine groups, the geometry is better to form a second hydrogen bond with E210 and increase van der Waals interactions. These interactions could explain why the ICM score is better for 3f/3i (-25.1) compared to 1c/1d (-10.5 to -12.0).

A comparison between the NMR and Maybridge hits derived from screening against the *Ac*AChBP:nicotine complex indicates that both sets occupy a comparable position. However, the NMR compounds score substantially lower (-6.2 to -13) than the Maybridge set (-25.4 to -29.6). One reason for this may be due to differences in interactions with the key water molecule. In the Maybridge set, every compound is predicted to make one or two hydrogen bonds with this water. Whereas, for the NMR compounds the fluorine group blocks any interaction.

Two binding positions have been identified for the Maybridge set of compounds interacting with *Ac*AChBP:*apo*. Most compounds are still driven by the interaction with the key water molecule. Perhaps due to the lack of this drive to make a hydrogen bond with the water molecule, the compounds from NMR make different interactions. Similarly, possibly due to the extra space and lack of constraints, both sets take advantage and form new interactions lower in the pocket, towards Y91 (Y110 of *Ac*AChBP:nicotine). This opens new areas of the pocket that could be targeted to increase affinity or specificity.

5.5.7.2 Comparison of *Ac*AChBP:nicotine and :*apo* hits

As noted above, all the small molecules bound in the *Ac*AChBP:nicotine structure are orientated similarly. This is due to the smaller pocket volume (189 $Å^3$) in comparison with *apo* (240 $Å^3$). As the pocket is smaller the number of van der Waals and hydrophobic interactions can increase. However, the potential to make hydrogen bonds is limited by space and orientation. Moving to the *apo* structure the opposite should happen, the potential for van der Waals and hydrophobic interactions decrease and hydrogen bonding opportunities increases. This is observed by the presence of two binding orientations in *apo* (previously described).

The nicotine-like orientation is similar between the *Ac*AChBP:*apo* and :nicotine structures. A hydrogen bond donor or acceptor on, or in, the ring of the hit compounds forms the interaction. In both cases, a carbonyl substituent could make a stronger and closer interaction than other groups. Another comparable interaction between the two sets of compounds is the position of a ring within the aromatic cage of the binding site. The aromatic cage provides van der Waals as well as charge-π and π-π interactions. Depending on the size of the ring, the strength of interactions with the aromatic cage will possibly be stronger in the *Ac*AChBP:nicotine structure due to proximity. In the *Ac*AChBP:nicotine structure, additional hydrogen bonds could be made with residues of loop C, interactions which are predicted to strengthen binding. Due to the conformation of loop C in the *apo* structure, these supplementary interactions, generally, cannot be established. The second orientation in the cavity between loop C and the principal subunit either cannot be adopted or is unfavourable in the *Ac*AChBP:nicotine binding site. This cavity may form similar interactions as in the *Ac*AChBP:nicotine pocket, due to size.

As discussed in Chapter 3.6, *Ac*AChBP provides a surrogate for the nAChR α_7 subunit (216). Figure 5.14 shows a sequence alignment between the two *Ac*AChBP structures used in docking and human nAChR α_7. The main binding residues, as discussed for compound binding, in *Ac*AChBP are conserved in the α_7 subunit, except for M133. In the α_7 receptor there is a substitution, M133Q however, as all interactions between the compounds and M133 are by van der Waals these should be maintained. Nonetheless, the hydrogen bonding capacity of glutamine needs to be considered as it may affect the binding of the compounds. Conservation of the key binding residues suggests that binding interactions should be maintained between the two proteins.

A comparison between known ligands and the compounds will allow for comments on the validity of the models generated and may give merit to the interactions observed. Three compounds are shown in Figure 5.15, the first two are agonists and the last is an antagonist (284). All three compounds have similarities, a carbonyl and an amine group. The carbonyl, similar to some of the compounds probably makes contacts with the binding site water molecule. This interaction may be bidentate as with

	PRINCIPAL [+]			COMPLEMENTARY [-]		
Loop	A	B	C	D	E	F
	107	169	210	77	131	190
nAChRα7	IWKPDILLYN	GSWSYG	YECCKEPY	WLQ	LVNSSGHCQYL	YI
	102	162	205	72	125	185
AChBP-nico	IWTPDITAYS	GSWVYS	YSCCPEPY	YEQ	VVTHDGSVMFI	YY
	83	143	186	53	106	166
AChBP-apo	IWTPDITAYS	GSWVYS	YSCCPEPY	YEQ	VVTHDGSVMFI	YY

Figure 5.14 Sequence alignment of human nAChR α_7, *Ac*AChBP:nicotine (PDB code: 5O87) and *Ac*AChBP:*apo* (PDB code: 2Y7Y). The alignment shows the main principal-(+) and complementary-(-) loops for binding.

compound 4b, due to the proximity of two hydrogen bonding groups. The core structure of antagonist SR 16594 is similar to compounds 4b, 4c and 4d. Possibly, the large 8-membered ring prevents the pocket from closing. If a similar group was attached to one of the similar three compounds a comparable effect may be observed.

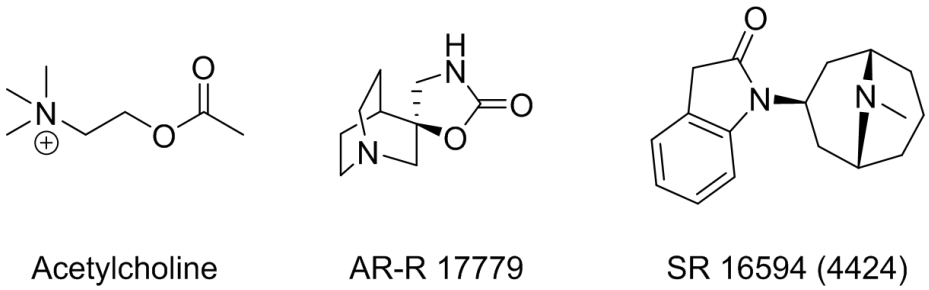


Figure 5.15 Three ligands of nAChR, used as a comparison to compound hits. Acetylcholine is the stereotypical ligand of the receptors, AR-R 17779 is an agonist and SR 16594 (4424) is an antagonist.

Numerous crystal structures of *Ac*AChBP:ligand complexes have been reported, which provides an extensive data set (285,286). Comparison between *Ac*AChBP:ligand structures may highlight the importance of different interactions. Tubocurarine is a neurotoxin that inhibits the activation of nicotinic receptors (59). Similarly, to the predicted orientation of some compounds in the *Ac*AChBP:*apo* structure, it occupies a position between the C loop and W145 however, does not make an interaction with the water molecule (PDB code: 2XYT). A K_D of 0.5 µM for the *Ac*AChBP:tubocurarine interaction (59), in comparison to nicotine binding (0.4-2.3 µM from Chapter 3.2) may suggest that the water-ligand interaction is not essential for binding. Several compounds in the *apo* pocket also do not make this interaction therefore, these may provide new scaffolds targeting other parts of the pocket. Likewise, tubocurarine, comparable to 2c, sits lower in the pocket making interactions with Q36 and S165. Using small groups or long linkers, to avoid antagonistic behaviour, this lower region could be exploited to aid in the selection and specificity of agonists. Cytisine is a plant alkaloid and has been used for smoking cessation (287). Similar to nicotine, it is a small molecule that interacts with W145 but unlike nicotine, it does not form an interaction with the stable water molecule (PDB code: 4BQT). The compounds in

the *Ac*AChBP:nicotine binding site occupy a similar position to cytisine however, they make interactions with the water molecule as well as loop C residues. This suggests that some of these compounds could have a similar affinity to cytisine with *Ac*AChBP (1.6 μM (287)).

5.5.7.3 Comparison of specific protein-compound interactions formed with GBP and *Ac*AChBP

The binding pocket of GBP is a hybrid, in terms of size, between *Ac*AChBP:nicotine and :*apo*. There is more space and fewer constraints, therefore a variety of interactions can form. This is partly due to the substitution of key residues, such as Y110A, but also partly because of the GBP:strychnine complex structure used for docking. As previously mentioned, strychnine is an antagonist of the native receptor preventing the closing of loop C (25) therefore, it will perhaps also prevent the closing in GBP.

Several residues in the orthosteric binding site are conserved between the two proteins, namely, Y72, W164, Y205 and Y212 which contribute to the aromatic cage. These residues provide general hydrophobic and van der Waals interactions. Protein specific substitutions such as Q74R, I135S and S206CCP_KGTG provide the specificity for the binding of these small molecules as demonstrated in previous sections. Possibly due to these substitutions, the compounds that bind to GBP are not predicted to display a similar orientation as the compounds that bind to *Ac*AChBP:nicotine.

A sequence comparison of GBP and GlyR has been discussed in Chapter 4.4.1 (Figure 5.16). The residues thought to be most important for compound binding in GBP are conserved in a $\beta^+\alpha^-$ GlyR binding site. There are three substitutions, W164F, R76Q and E210 deletion (E210del), which will be discussed. W164 provides van der Waals interactions with all the GBP ligands, therefore substituting this residue with Phe should not make much difference. The R76Q substitution is only important for the binding of two compounds 3g and 3c. This interaction may still be maintained with the substitution however, for 3g the other interacting residues, R96 and V165, are also not conserved. Consequently, 3g would probably not bind to GlyR. The final substitution is a deletion of E210. This has an impact on several compounds, which make a hydrogen bond with this residue including 3d, 3e, 3f, 3i and 3j. If these compounds were taken forward this interaction would have to be accounted for and it would have an impact on the binding score. However, without this residue (E210) a new pocket becomes available to build a molecule towards S167, which is conserved as a threonine in GlyR.

Various compounds bind to GlyR, including various amino acids, cannabinoids, ions and neurosteroids to name a few (288). Comparing GlyR ligands and the compounds identified will potentially show common chemical structures and binding modes, as was conducted with the *Ac*AChBP compounds. Figure 5.17 shows three competitive ligands of GlyR. The smallest of GlyR ligands are the amino acids including glycine and alanine (289). The two ligands, nipecotic acid and quinolinic acid are structurally similar to each other and are derivatives of piperidine and quinolone respectively (289,290). The

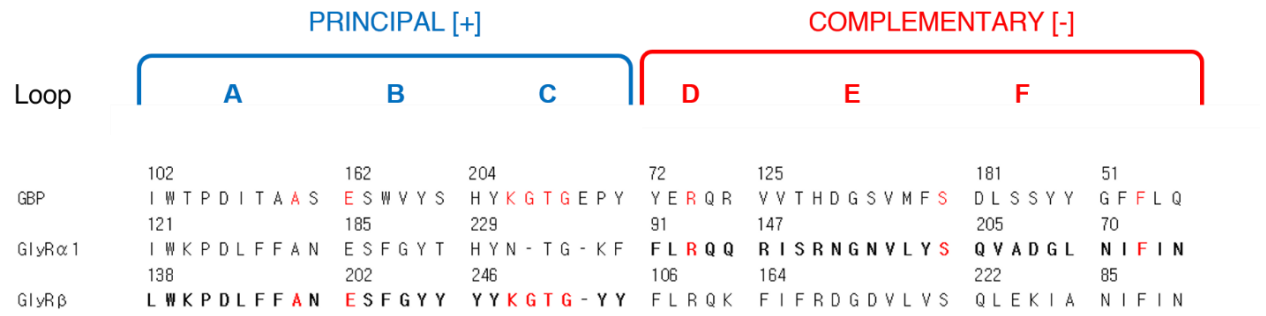


Figure 5.16 Sequence alignment of human GlyR α_1 and β with GBP (PDB code: 5OBG). The alignment shows the main principal-(+) and complementary-(-) loops for binding.

quinolinic acid derivative is comparable to 1a, two fused aromatic rings with a carboxylate and CF_3 group. Consequently, it would be expected that they would similarly interact with the target. The carboxylate group of nipecotic and quinolinic acids are also mirrored in some of the Maybridge GBP hits. This group can be related to the natural ligand, glycine. The GlyR α_3:glycine complex (PDB code: 5VDH) shows that the carboxylate of glycine interacts with R65 (291). This is comparable to R74 in GBP, which interacts with the carboxylate of some of the hit compounds. There is a difference in size between the natural ligand and the other two compounds in Figure 5.17. Nipecotic (MW 129 Da) and quinolinic (MW 167 Da) acid are smaller than strychnine (MW 334 Da), however, all three compounds are antagonists of GlyR (288). Similarities between the acids and the compounds identified from computational docking may lead to the conclusion that some of these molecules may also inhibit GlyR.

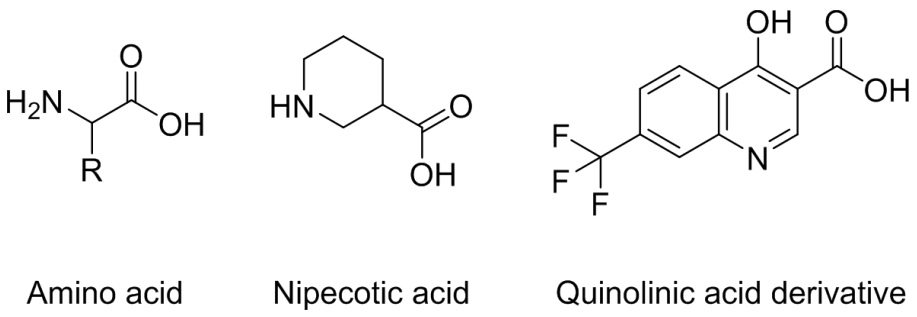


Figure 5.17 Three ligands of GlyR used as a comparison to small molecule hits. Glycine the physiologically relevant ligand is an amino acid. Other amino acids have also been shown to bind the receptor. Nipecotic acid is structurally similar to the amino acid and is based on piperidine. Quinolinic acid is based on quinolone and shares structural similarities with nipecotic acid.

To compare binding modes, structural data are required. For GBP, strychnine, N-methylbicuculline and glycine are the only ligands with structural data available in the PDB. As the structure of strychnine and N-methylbicuculline have been extensively discussed in Chapter 4 they will not be considered here. Glycine bound to GBP (PDB code: 5OAN) is positioned in the centre of the binding site within the aromatic cage and mainly makes interactions with the principal subunit (75). This is similar to a subset of compounds including 1d, 1c, 3f and 3i. The latter three compounds also make the same hydrogen bond, as glycine does, to MC W164. Without more ligands bound to GBP, further validation of

compound-protein interactions cannot be conducted. Other ligands bound to GlyR are accessible however, these bind in other sites than the orthosteric binding pocket. For example, AM-3607 is a candidate potentiator that binds between the extracellular subunits, neighbouring the orthosteric site (243).

5.5.8 Compound modification and extension to answer questions and improve binding

To address some of the issues raised above, example small molecules were modified to investigate the effect on binding. The top-ranking compound bound to each protein (GBP, *Ac*AChBP:nicotine and *Ac*AChBP:*apo*) were also modified and extended with various functional groups, to investigate any effect on predicted binding.

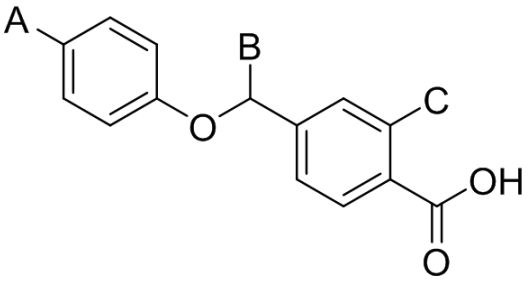
5.5.8.1 Is the GBP model complete?

A point previously mentioned, is the deletion of E210 in the heteromeric GlyR. This residue provides a hydrogen bond as well as van der Waals interactions with several compounds in GBP. The removal of the hydrogen bonding group on a compound was trialled. The 3i Maybridge compound was chosen, the hydroxyl was removed, and the molecule was re-docked to the GBP binding site. The removal of the group resulted in a Hbond energy decrease of 3 kcal mol^{-1} and the ICM score decreased from -25.1 to -18.6. Therefore, this suggests that the hydrogen bond with E210 was important for binding. Substitution of E210 with Gly or Ala led to a similar drop in ICM score however, the Hbond energy was maintained between -4.3 and -4.7 kcal mol^{-1}. Due to the replacement of E210 with smaller groups the compound was able to re-orientate forming an interaction with T208 instead. This analysis suggests that possibly the deletion of E210 from future docking experiments, and perhaps GBP itself, would make a more representative binding pocket of the $\beta^{+}\alpha^{-}$ GlyR orthosteric site.

5.5.8.2 Improving the binding score of a GBP hit

The top-ranking compound for GBP was 3b with an ICM score of -29.4. Looking at the binding pocket there were three potential areas to build into: 1) above towards R96, 2) out towards W164 and 3) back between E162 and Y212. There are many sites on 3b where extension could occur, therefore three were selected to exploit the pockets identified (Figure 5.18). A screen of 46 common groups per extension point was conducted. These common groups were taken from ICM and by analysis of common compound substituents (Chapter 2.5.2). The top three of each position are shown in Figure 5.18. A new term, VLS score, is used to investigate binding improvement, as this incorporates the effect of additions on ligand strain (Chapter 2.5.2). Ligand strain energy is calculated as the difference between the bound and free-state and can affect how a compound binds and the affinity to the target (292). Therefore, adding groups to a scaffold may increase the strain and cause problems with ligand binding.

A baseline VLS score of -24.9 was measured for a hydrogen substituent at each position. Extension from position A did not lead to a significant increase in score. However, building from positions B and C led to an improved score by 5.6 and 3.2 units. The highest-scoring group of each position was re-docked, optimised and position combinations were trialled (Chapter 2.5.2). From this analysis, it was identified that the ester (position A) and amine (position B) were enough to raise the ICM score to -29.4, compared to -24.2 of the optimised unaltered ligand.



	A	B	C
1	$-COCH_3$	$-NH_2$	$-CF_2$
2	-Ph	$-C_5H_6$	$-CONH_2$
3	$-OCH_3$	-OH	$-CH_2CONH_2$

Figure 5.18 Extension of the top-ranking GBP small molecule, 3b. Three positions (A, B and C) were chosen to exploit different pockets within the binding site (PDB code: 5OBG). The table shows the top three groups for each position that gave an improved binding score.

Interestingly, the incorporation of the cyclopentadiene at position B caused the ligand to take up a different orientation within the GBP binding site. This orientation was better than the original with a VLS score of -25.7 compared to -25.0. The addition formed a new hydrogen bond between the carboxylic acid and Y72, as well as R74. The change could have also increased the van der Waals interactions of the compound. This new compound could be further developed to manipulate binding interactions and improve predicted affinity.

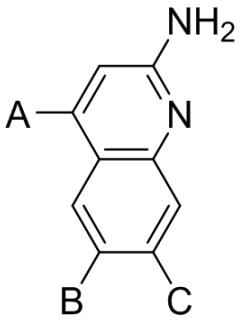
5.5.8.3 Identifying differences in binding between NMR and Maybridge hits with *Ac*AChBP

Previously, the potential problem of the fluorine atoms of the NMR compounds interfering with binding to *Ac*AChBP:nicotine was noted. To investigate this the CF_3 group of compound 2c was removed. Upon re-docking of the modified compound to the *Ac*AChBP:nicotine binding pocket, a substantial change was observed. The compound adopted a different pose with the pyridine ring placed between V165 and M133 with the formation of a hydrogen bond (1.6 Å) between the N-pyridine and water molecule, as seen for all the Maybridge compounds investigated. Additionally, due to the change in orientation, the second pyridine group was located in the centre of the aromatic cage and the amine linker made a further hydrogen bond with the MC of W164. The new orientation and interactions led to an improved ICM score of -21.0 (from -12.4).

Deletion of the CF_3 group of 2c and re-docking to *Ac*AChBP:*apo* was also conducted. This led to minimal orientation changes, however, a decrease in ICM score was observed (-18.9). This is perhaps due to the loss of the extra hydrophobic and van der Waals interactions that the group made.

5.5.8.4 Improving *Ac*AChBP binding scores by extending scaffolds

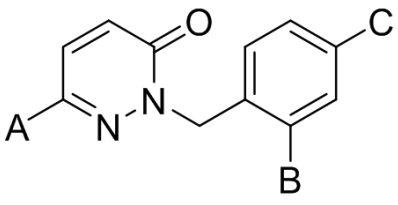
The top compound docked to *Ac*AChBP:nicotine is 4b with an ICM score of -29.6. Due to the small pocket and size of the compound, the extension of the scaffold is limited. There are two possible locations to extend the scaffold into: 1) from the hydroxyl towards R96 and 2) from the phenyl ring below loop C. Three positions of the ligand were investigated by screening the set of 46 common groups. A baseline value for position A of -17.1 and positions B/C of -22.4 were obtained. Removal of the hydroxyl, at position A, led to a decrease of the VLS score by 5.3, which suggests the hydrogen bond to Y212 is important for binding. The removal of the hydroxyl is also the reason why there is a difference in baseline scores for different positions. Again, the top three groups for each position are shown (Figure 5.19). The highest-ranking groups were re-docked with *Ac*AChBP:nicotine and it was found the highest-scoring combination was for positions B and C, with a VLS score of -42.3, an improvement of 19.8.



	A	B	C
1	$-CONH_2$	$-CONH_2$	$-C_5H_6$
2	$-NH_2$	$-CF_3$	-Ph
3	-OH	$-(CH_2)_2NH_2$	$-C_4H_7$

Figure 5.19 Extension of the top-ranking *Ac*AChBP:nicotine small molecule, 4b. Three positions (A, B and C) were chosen to exploit different pockets within the binding site (PDB code: 5O87). The table shows the top three groups for each position that gave an improved binding score.

Finally, the top-scoring small molecule for *Ac*AChBP:*apo*, 5c (ICM score -26.2), was investigated. Due to the open pocket, several large areas could be exploited by compound extension from the scaffold. Three positions were chosen to exploit three different pockets: 1) methyl toward Y193 (A), 2) phenyl alongside Y53 (B) and 3) methyl towards Y91 (C). Due to the open pocket, residues were far apart therefore, interactions could not be fully optimised as is seen by the low VLS scores. Also, owing to space within the pocket optimisation of modified compounds was harder, as many orientations were feasible. Eventually, the phenyl (position A) and amine (position C) were found to be optimal (Figure 5.20). These groups made additional interactions with the protein, including a hydrogen bond with Y53. This led to a VLS score of - 30.9, an improvement of 4.4.



	A	B	C
1	-Ph	$-CH_2Ph$	$-CH_2NHCH_3$
2	$-C_5H_6$	-Cl	$-CONH_2$
3	$-C_6H_{11}$	$-CF_3$	-Ph

Figure 5.20 Extension of the top-ranking *Ac*AChBP:*apo* small molecule, 5c. Three positions (A, B and C) were chosen to exploit different pockets within the binding site (PDB code: 2Y7Y). The table shows the top three groups for each position that gave an improved binding score.

5.6 Summary

Identification of several compounds that potentially bind to *Ac*AChBP or GBP have been found by experimental and computational approaches. Analyses of the orientations and binding interactions these compounds make with the relevant receptor have led to similarities and differences being recognised. The binding of two of the small molecules have been validated by WF. Sequence and structure comparisons suggest that many of the conclusions derived from the surrogate proteins can be translated to the native receptors, nAChR and GlyR, with some exceptions. Examples of potential compound extensions have been shown to highlight the possibility of developing lead compounds from these small molecule hits. However, these observations would need to be experimentally validated.

6. Design of and attempted production of selected proteins

6.1 Construct design

6.1.1 GlyR $\beta^+\alpha^-$

Except for GBP, there are no other models of the GlyR heteromeric orthosteric binding site, therefore a new construct was designed. The new construct represents the $\beta^+\alpha^-$ heteromeric interface of GlyR, which is the physiologically relevant site of glycine binding (39) but one formed by a homomeric assembly. The full length α_3 homomer, which has been produced previously by Huang *et al.* was used as a template (24). Comparison of *Ac*AChBP and human GlyR α_3 as a template suggests increased residue conservation in the orthosteric binding site and potentially fewer substitutions required to form a $\beta^+\alpha^-$ interface. A sequence alignment of human GlyR $\alpha_{1\text{-}3}$ and β was first used to identify residue differences between subunits (Figure 6.1).

	PRINCIPAL [+]			COMPLEMENTARY [-]			
Loop	A	B	C	D	E	F	G
	121	185	228	91	155	205	70
hGlyRa1	I W K P D L F F A N	E S F G Y T	K H Y N - T G K F	F L R Q Q	L Y S I R	Q V A D G L	N I F I N
	127	191	234	97	161	211	76
hGlyRa2	I W K P D L F F A N	E S F G Y T	K H Y N - T G K F	F L R Q Q	L Y S I R	Q V A E G L	N I F I N
	126	190	233	96	160	210	75
hGlyRa3	I W K P D L F F A N	E S F G Y T	K H Y N - T G K F	**F L R Q K**	**L Y S I R**	**Q V A E G L**	**N I F I N**
	138	202	245	106	172	222	85
hGlyRβ	**L W K P D L F F A N**	**E S F G Y T**	**K Y Y K G T G Y Y**	F L R Q K	L V S M R	Q L - E K I	N I F I N
	93	157	200	63	127	177	42
5TIN a3	I W K P D L F F A N	E S F G Y T	K H Y N - T G K F	F L R Q K	L Y S I R	Q V A E G L	N I F I N
	93	157	200	63	127	177	42
GlyR Het	L W K P D L F F A N	E S F G Y T	K Y Y K G T G Y Y	F L R Q K	L Y S I R	Q V A E G L	N I F I N

Figure 6.1 Sequence alignment of the main loops in the orthosteric binding site of pLGICs, A-C of the principal-(+) subunit and D-G of the complementary-(-) subunit. Sequences of GlyR $\alpha_{1\text{-}3}$ and β were retrieved from Uniprot. The sequence of the GlyR α_3 structure (PDB code: 5TIN) was also compared. The residues that make up the $\beta^+\alpha^-$ heteromeric interface are shown in bold. Highlighted residues show the required substitutions required to transform a principal-α_3 subunit into β.

Heteromeric $\beta^+\alpha^-$ receptors are formed by the only β subunit and either α_1 or α_3 (288). The sequence alignment indicates a high level of conservation between the α subunits when considering the relevant binding loops (D-G). Only two, conserved substitutions were identified between α_3 and α_1, E213D and K100Q. Therefore, model ligand binding within the GlyR $\beta^+\alpha^-$ should hold for interactions for both α receptor subtypes. To form a principal-(+) β subunit, only loops A-C need to be considered. From the alignment, it was proposed that six substitutions translate an α_3 principal subunit to β (I93L, H201Y,

N203KG, K206Y and F207Y). The conservative substitution, I93L, was not considered significant and was ignored. The major difference between α and β involves loop C, which, as noted elsewhere, is essential for the correct functioning of all pLGICs (25). The loop also provides many interactions, including two residues, which contribute to the aromatic cage, a feature of the pLGIC family (293). Similar to GBP, the C loop should be altered to potentially provide β subunit ligand interactions. Unlike GBP, the flanking tyrosine residues (Y246 and Y253) of GlyR β will also be included as these may provide additional interactions or stabilise the loops' conformation. Additionally, the N38Q substitution from the Huang *et al.* construct will be maintained to prevent glycosylation (24).

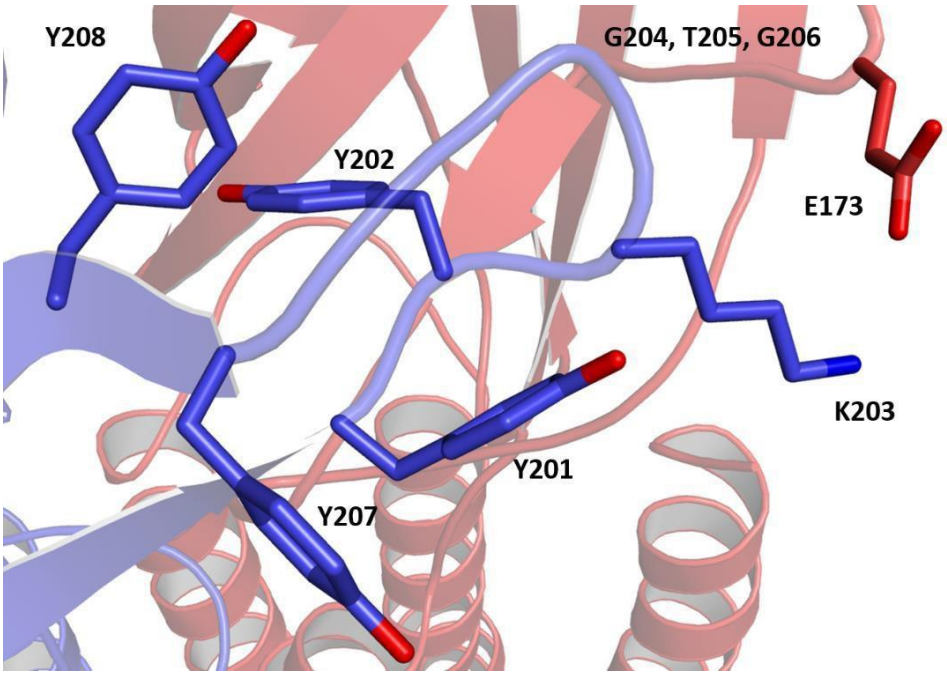


Figure 6.2 GlyR $\beta^{+}\alpha^{-}$ model showing the loop C substitutions (blue sticks) of the principal subunit (blue). The residue E173 (red sticks) from the complementary subunit (red) is shown as it may form a salt bridge with K203. Due to the size of the sidechains of G204 and G206, they are not shown. T205 is also not shown as this clashes with Y208. Oxygen (red), nitrogen (blue) and carbon (chain colour) atoms are shown.

A model of the GlyR $\beta^{+}\alpha^{-}$ was produced utilising the GlyR α_3 homomer protein sequence (24). The five substitutions (H201Y, N203KG, K206Y and F207Y) are potentially well-supported by the protein and should not affect stability. In the model, T205 and Y208 appear likely to clash, however, either due to a different rotamer of Y208 or loop flexibility this should not occur. The two aromatic residues, Y202 and Y208, often display as different rotamers in structural models as noted in the GBP:N-methylbicuculline complex (Chapter 4.5) and may form a stabilising hydrogen bonding interaction that maintains loop conformation. No other clashes with residues of the binding pocket are observed. The substitutions maintain Y202 and T205 in the orthosteric site but also add the SC of Y208. Even though T205 is

maintained, it occupies a different position. This appears to be due to the presence of an additional residue, G204 in the loop. K203 is positioned 3 Å away from the SC of E173, this distance could support a salt bridge. Such a strong interaction could stabilise the conformation of the loop. The Y201 and Y207 substitutions also point away from the binding pocket, nonetheless, these residues may contribute additional stabilising interactions in certain loop conformations.

6.1.2 HisR homomer

Two HisR subunits are commonly identified within different insect species, HisR1 (type-1) and HisR2 (type-2) (294). A sequence alignment of selected examples reported in Uniprot (137) is shown in Figure 6.3.

There is a high degree of conservation between different HisR receptor subtypes and species (Table 6.1). Some receptors have alternate sequences due to splicing (295), therefore there is a degree of uncertainty when utilising such sequences for protein production. Both type-1 and 2 receptors can form homomers with HisR1 channels being shown to be more sensitive to histamine (EC_{50} 3 µM) (296). The two subunits can also form heteromers (297), however, for initial protein production, only a single subunit will be investigated. Structural predictions identified that the type-1 HisRs were similar for most of the calculated parameters. Small differences were observed with the instability index and prediction of the number of transmembrane helices. The instability index calculation considers the number of specific dipeptides in a primary amino acid sequence that have a higher frequency in stable or unstable proteins. All proteins with a half-life >15 h had a score below 40 and were classified as stable, whereas unstable proteins with a half-life <5 h had a score greater than 40 (298). However, comparisons of HisRs demonstrated neither the number of transmembrane helices nor the instability index scores were significantly different. Due to these similarities, it was decided that any of the channels would be viable for protein production, the *A. mellifera* HisR1 protein was chosen to attempt production.

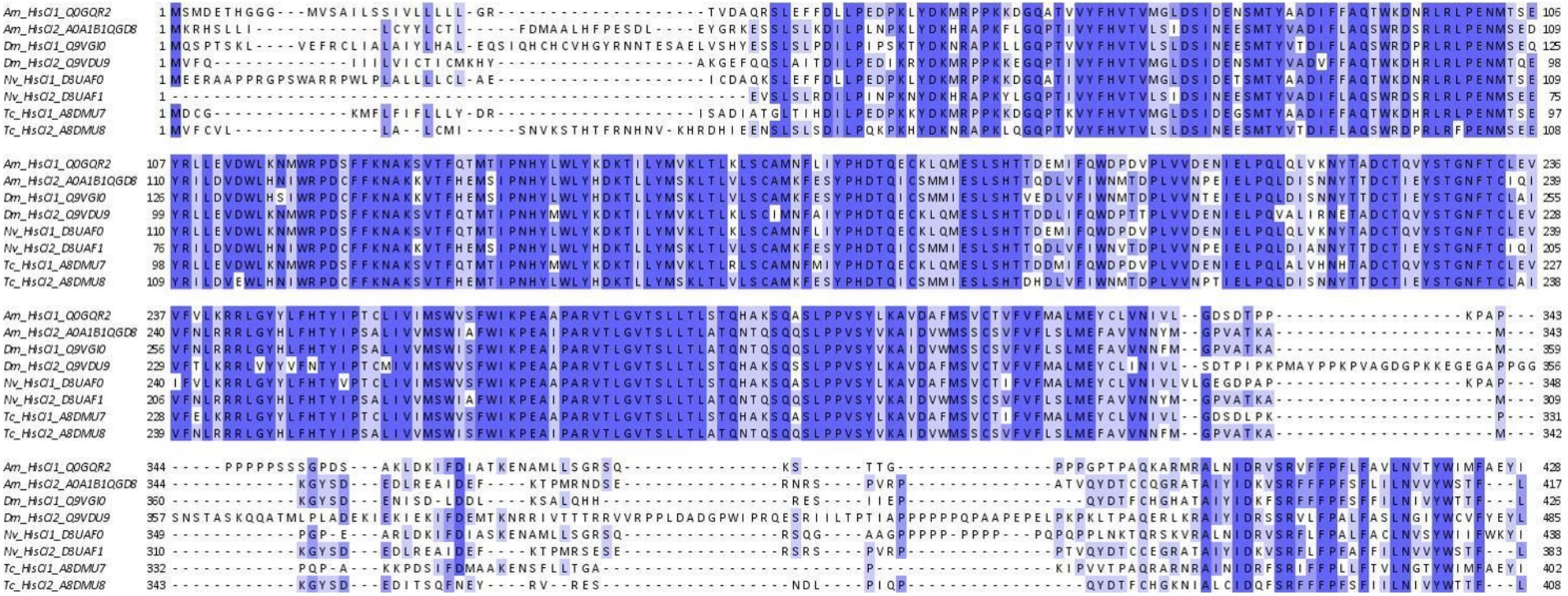


Figure 6.3 Sequence alignment of HisR1 (HisCl1) and HisR2 (HisCl2) from *D. melanogaster* (Dm), *A. mellifera* (Am), *N. vitripennis* (Nv) and *T. castaneum* (Tc).

%	Am1	Am2	Dm1	Dm2	Nv1	Nv2	Tc1	Tc2
Am1	-							
Am2	54	-						
Dm1	53	80	-					
Dm2	66	51	52	-				
Nv1	82	53	52	65	-			
Nv2	58	96	86	55	58	-		
Tc1	81	56	56	65	76	59	-	
Tc2	55	79	83	53	54	84	57	-

Table 6.1 Percentage sequence identity for HisR1 and HisR2 from different species *D. melanogaster* (Dm), *A. mellifera* (Am), *N. vitripennis* (Nv) and *T. castaneum* (Tc).

6.2 Attempted protein production in baculovirus

Attempts to produce both the GlyR $\beta^+\alpha^-$ and HisR proteins were carried out in insect cells, following the protocol used by Huang *et al.* to produce the homomeric GlyR α_3 (24). Fractions of insoluble, soluble and detergent-solubilised membrane proteins were analysed by Western blot utilising the relevant antibodies (Figure 6.4). The molecular weight of a single subunit of GlyR $\beta^+\alpha^-$ and HisR are 45.6 kDa and 50.3 kDa, respectively. No soluble or detergent-solubilised proteins were detected. Multiple bands were observed in insoluble fractions suggesting several forms of recombinant protein were present. No bands were detected in the negative control, which suggests that the bands in the experimental condition are due to the viral infection. Further experiments were carried out in an attempt to produce soluble protein. Such efforts included attempting to increase the production of proteins by altering baculovirus concentration, culture volume and incubation length, whilst at the same time attempts to solubilise the proteins including trialling different detergents, buffers and protocols. Analysis of different fractions throughout the experiments was conducted, however, the proteins remained insoluble.

Again, this raised questions about the production of a functional receptor and trafficking of it to the cell membrane in insect cells. If the proteins are produced and correctly trafficked, maybe there is an issue with the extraction of proteins from the membrane.

6.3 Localisation at the membrane

To examine if the receptors were being produced and trafficked to the cell membrane they were tagged with GFP. A preliminary experiment utilised GlyR $\beta^+\alpha^-$ model tagged with GFP at the C-terminus. A baculovirus expression system was used; infected and uninfected cells were visualised after 24 and 48 h (Figure 6.5).

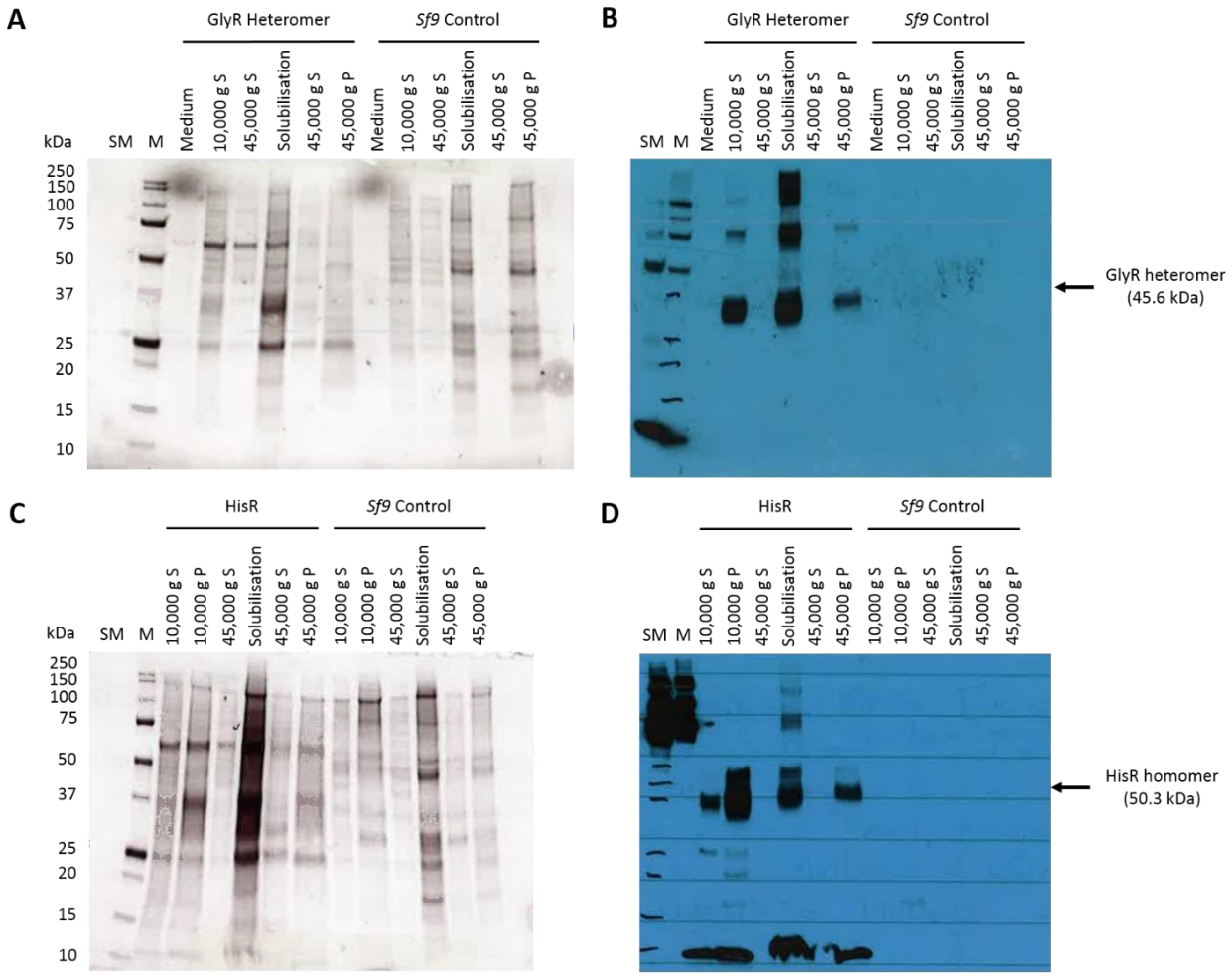


Figure 6.4 Example of an SDS-PAGE gel and Western blot for GlyR $\beta^{+}\alpha^{-}$ (A, B) and HisR (C, D). Stained (SM) and unstained (M) markers were used for size comparison. The intervals were marked on the Western blot from comparison with the SDS-PAGE gel for clarity. S and P are supernatant and pellet, respectively. In both cases, an *Sf9* negative control was conducted in tandem to detect background signals. An anti-His primary antibody was used for the detection of HisR, whereas an anti-Strep primary antibody was used to detect GlyR $\beta^{+}\alpha^{-}$.

Figure 6.5 shows that the infected cells are producing the protein compared to uninfected cells. By the difference in intensity, it can also be inferred that cells are still producing the protein after 48 h. However, this experiment also raised some questions. Due to the replication of the virus in the nucleus, the size of this organelle increases (299), therefore there is uncertainty about the localisation of the recombinant protein. Also, there was a risk that the GFP could have been cleaved from the protein, therefore the experiment may only show GFP localisation. Due to these reasons, a second experiment was set up. Both GlyR $\beta^{+}\alpha^{-}$ and HisR were N-terminally tagged with GFP and baculovirus' were produced. Experiments were run and visualised with the relevant controls, uninfected and GFP only infected cells.

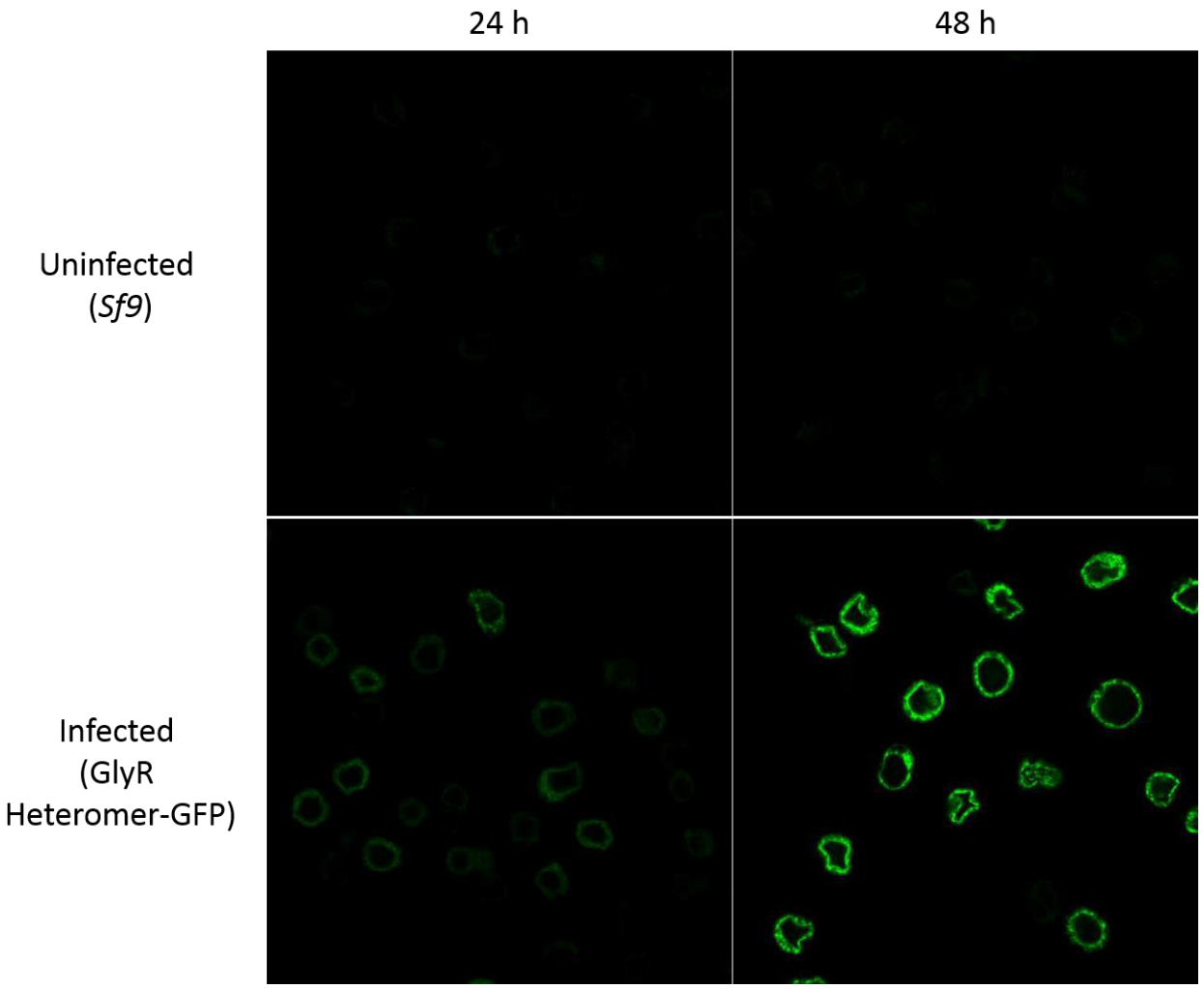


Figure 6.5 Microscopy images of C-terminal GFP-tagged constructs. The presence of GlyR $\beta^{+}\alpha^{-}$ model tagged with GFP at the periphery of infected *Sf9* insect cells. The signal intensity increases between 24-48 h demonstrating an increase in the protein produced. No background signal was observed for the uninfected controls.

The experiment showed no GFP signal in the uninfected cells (Figure 6.6), therefore there is no background signal. There is a clear difference between cells infected with the GFP only virus and the virus with GFP-tagged receptor. The GFP only cells have a widespread signal that occludes that of the nuclear stain in some cases. Whereas, the signal of the GFP-tagged receptor is localised to the membrane. Differences between signal intensity of Figure 6.5 and 6.6 are noted, this may be due to difficulty passing the GFP molecule through the membrane if it has already folded in the cytoplasm (300). These experiments suggest that the receptors are produced and localised to the membrane, therefore the issue may be due to inefficient extraction and solubilisation of the proteins during the purification process. However, membrane association does not guarantee correct folding, therefore such a sample may not be suitable for detergent extraction.

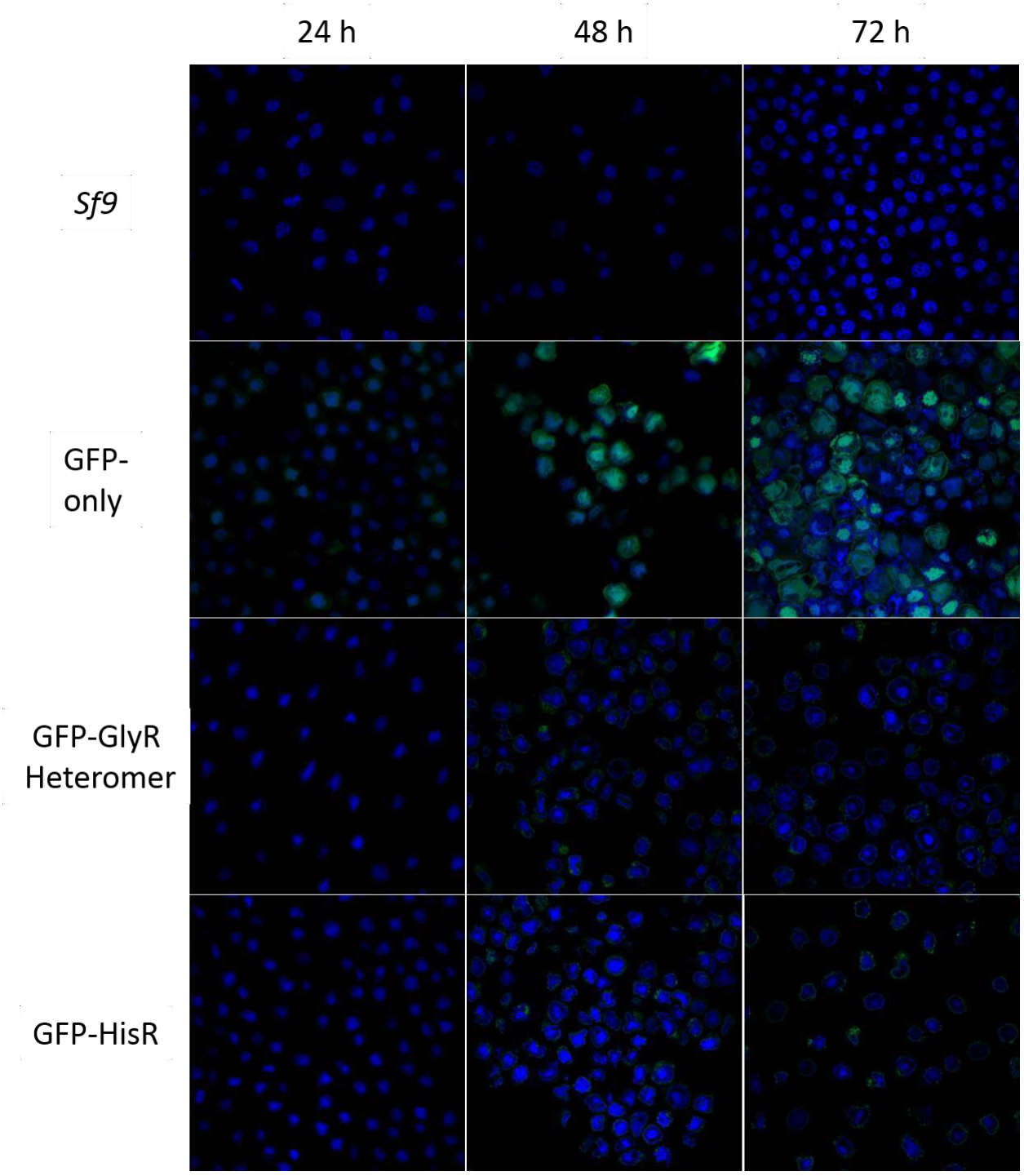


Figure 6.6 Microscopy images of N-terminal GFP-tagged constructs. Uninfected, GFP only, GFP-GlyR $\beta^{+}\alpha^{-}$ and GFP-HisR infected cells are shown. No GFP (green) signal is detected in the uninfected control. Widespread GFP signal is observed for the GFP only infected cells demonstrating that it is present in the cytoplasm. A lower GFP signal, located at the periphery of the cell is observed for both GFP-GlyR $\beta^{+}\alpha^{-}$ and GFP-HisR. This demonstrates that the receptors are most likely membrane-associated. In each condition, the nuclei (blue) were stained with Hoechst.

One possible explanation for the inability to extract the proteins from the membrane is detergent-resistant membranes. These membranes can occur in microdomains within the phospholipid bilayer and cannot be solubilised by detergents, even at high concentrations (301). These microdomains are known as lipid rafts and are subsections of the membrane that are comprised of different phospholipid components, compared to the rest of the membrane (302). It has been shown previously that members of the pLGIC family are located in detergent-resistant membranes that could not be solubilised with Triton X-100 (303). However, others have shown that detergent extraction of pLGICs, including GlyR, is possible (24,25). This suggests that either the most efficient extraction conditions have not been obtained or there are problems with protein production, in the case of the GlyR $\beta^+\alpha^-$ and HisR.

Another explanation is that during the extraction process aggregation of the protein may occur which leads to solubility issues. It has been suggested to limit exposure of membrane proteins to detergents as they may cause damage (29). To counter this, weaker detergents could be used or other methods such as a nanodisc protocol. However, less harsh detergents may extract the proteins with a lower efficiency, which generates other problems.

6.4 Attempted protein production in lentivirus

A different expression system was trialled to produce GlyR $\beta^+\alpha^-$ and HisR utilising the lentiviral system. Using the protocol by Elegheert *et al.*, previously used to produce $GABA_AR$ (150), an attempt to produce the receptors in HEK293 cells was conducted. Correct glycosylation may be essential for assembly and function as reported for $GABA_AR$ by Phulera *et al.* (31). After generating the stably-transfected cells, a preliminary analysis was carried out to determine receptor production. Soluble, detergent-solubilised and insoluble fractions were taken from cultures and analysed by Western blot. However, no bands were observed for either protein in any of the fractions.

6.5 Design of a HisR orthosteric binding site model

As HisR could not be detergent-solubilised, an *Ac*AChBP surrogate was proposed to study binding in these receptors. Based on the homology of HisR to GlyR α subunits and evidence of histamine binding to $GABA_AR$ (21), these proteins were used to infer relevant substitutions to form a histamine-binding orthosteric pocket. Histamine has a positive and negative modulatory effect on $GABA_AR$ and GlyR, respectively (304). A molecular model was also made using the sequence of *A. mellifera* HisR.

Histamine was docked into the orthosteric binding site of the HisR model (Figure 6.7B). The volume of the pocket is 285 $Å^3$, which is comparable to that in the *Ac*AChBP:nicotine complex (260 $Å^3$). The docking calculations suggest that the imidazole ligand forms a π-π interaction with Y142. One nitrogen atom of the ring and the primary amine make hydrogen bonds with the MC of the protein. The primary amine is

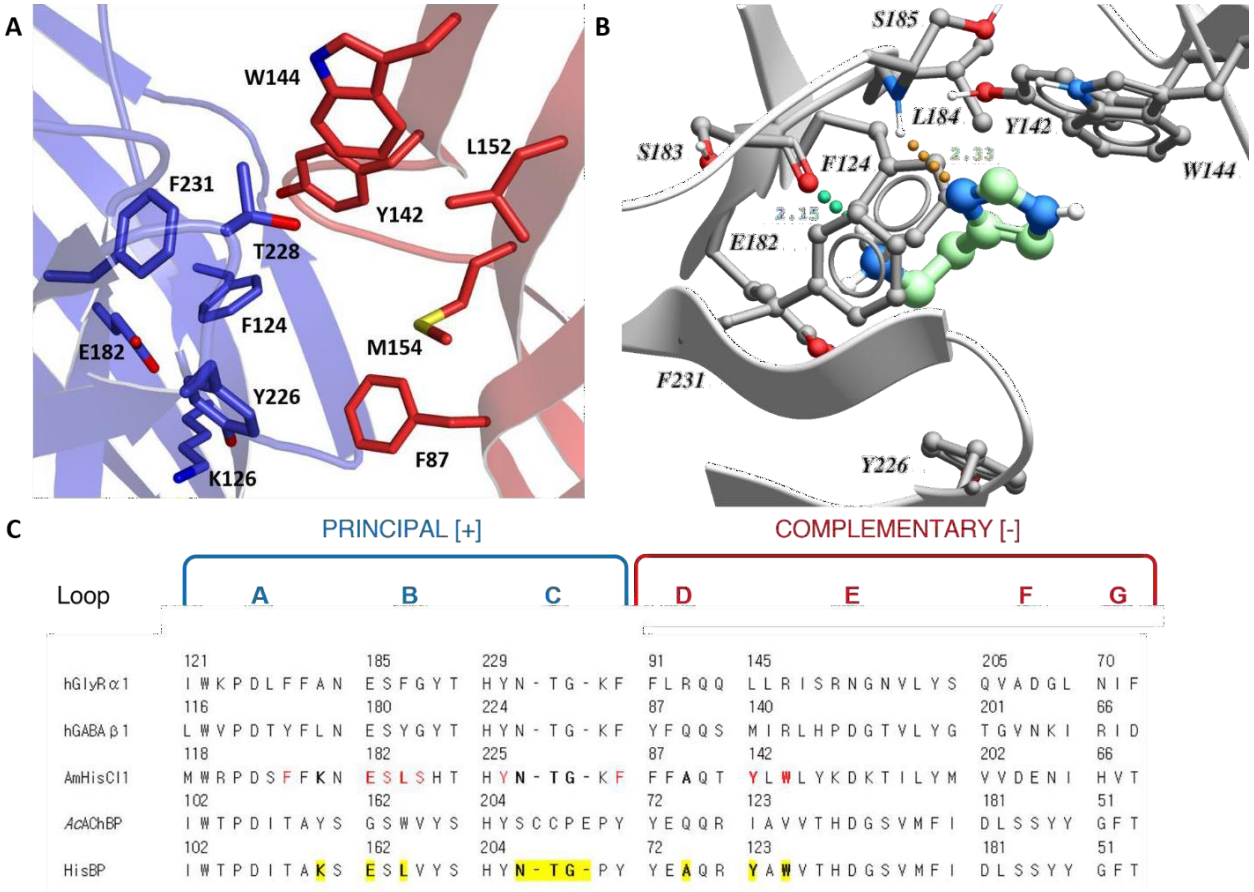


Figure 6.7 Information used to guide the design of a histamine-binding protein (HisBP). A) Model of the *A. mellifera* HisR generated by Phyre2 and SymmDock shows the major residue side chains (sticks), which contribute to the HisR orthosteric binding site. The principal (blue) and complementary (red) subunits are shown. Atoms are coloured as in Figure 6.2 with the addition of sulfur (yellow). B) Histamine (green carbon atoms) docked into the HisR model (grey). Residues involved in close interactions are shown as sticks. Two main chain hydrogen bonds (coloured spheres) are shown. C) Sequence alignment of the main loops that contribute to binding in the principal (A-C) and complementary (D-G) subunits. Sequences of human GlyR α_1, $GABA_AR$ β_1, *A. mellifera* HisR1 and *Ac*AChBP are compared. Residues in red are suggested to be involved in histamine binding in (B). Bold and highlighted residues are suggested substitutions to form HisBP.

likely protonated under physiological conditions (pKa 9.8), therefore can make a π-cation interaction with the SC of F231. The ligand may be involved in van der Waals interactions with the residues F124, W144, E182, S183, L184, S185 and Y226. Histamine does not make extensive van der Waals interactions (VDW energy -14.9 kcal mol^{-1}), leading to a low ICM binding score of -16.9, however, this is potentially due to its small size. Possibly, more interactions between histamine and HisR would be formed upon binding in solution, due to loop and residue flexibility, which is not accounted for in the docking simulation.

Using the sequence alignment, HisR and histamine-docking model, eleven substitutions are suggested to generate a histamine-binding protein (HisBP) from an *Ac*AChBP template. The two substitutions W164L and V125W should be considered together. W164L removes one of the main aromatic cage

residues in *Ac*AChBP and opens up a deeper part of the pocket. The aromaticity is partly replaced by the phenyl ring of V125W. Due to the large SC of tryptophan, the V125W substitution also blocks binding to the upper part of the pocket towards R96 in *Ac*AChBP. In the GBP:N-methylbicuculline complex, this arginine can provide additional interactions with ligands. These two substitutions are not conserved in GlyR or $GABA_AR$, therefore are specific for HisR. These changes alter the residues involved in binding. Due to the removal of W164, the substitution I123Y could position a Tyr residue in its stead, as is seen in the histamine-binding model (Y142, Figure 6.7) replacing the aromaticity lost from the W164 substitution. This change is also not observed in the other pLGICs compared.

The substitutions G162E and Y110K in *Ac*AChBP should be considered in tandem as well. To generate GBP these residues were also substituted, G162E and Y110A (75). G162E is maintained in both GlyR and $GABA_AR$ and directs an acidic residue into the binding site, which may aid in orientating ligands. Y110K removes the bulk of tyrosine, therefore allowing the G162E substitution. The lysine in the HisR model, together with F87, provides the bottom boundary of the binding site. This lysine could also make a salt bridge with Q224 of loop C. Loop C substitutions (S206CCPE_NTG) are conserved across $GABA_AR$ β, GlyR α and HisR and lead to the removal of the Cys loop in *Ac*AChBP. These substitutions also include the deletion of two residues in loop C decreasing loop flexibility. The final substitution Q74A would remove a polar group from the binding site, this is a HisR-specific alteration as a Gln is maintained in $GABA_AR$ and a different polar residue, Arg, is present in GlyR.

The suggested substitutions could form the basis of generating HisBP, however, some may not be tolerated by *Ac*AChBP. In particular, the W164L substitution, as the tryptophan provides interactions with all ligands that bind in the pocket and forms a link between two β strands. When attempting to generate GBP, Dawson *et al.* replaced W164 with Phe, however, this was not well-tolerated and the protein stability decreased (75). Nonetheless, the use of the I123Y substitution may negate this destabilisation.

Comparison of the data above and mutational studies investigating histamine binding in $GABA_AR$ suggest that the molecule interacts differently with the distinct receptors. One electrophysiology study suggested that the Q64A substitution of $GABA_AR$ β_3 decreases histamine potency (EC_{50} 480 μM), whereas Q64E increases it (EC_{50} 40 μM) compared to the unsubstituted form (EC_{50} 200 μM) (305). In HisR this residue is A89 therefore inferring that distinct binding conformations are adopted. This statement is supported by Thiel *et al.* who demonstrated that residues R120 (α_1), Y157, D163 (β_2), V175 and Q185 (β_3) are important for histamine binding to $GABA_AR$ (304). However, all but one of these residues are outside the binding pocket of HisR. The β_2 residue Y157 is not conserved in HisR, instead, the aromatic group is replaced by leucine (L184). The hypothesis that histamine may adopt separate conformations when bound to the receptors can only be proven by direct comparison of $GABA_AR$ and HisR structures.

Nonetheless, it is conceivable that a small molecule could bind multiple receptors in different positions as shown with N-methylbicuculline.

6.3 Summary

A full length human GlyR with a $\beta^{+}\alpha^{-}$ interface was designed as a surrogate of the physiologically relevant heteromeric receptor. A construct encoding *A. mellifera* HisR was prepared. An attempt to produce both of these pLGIC family members was carried out via baculovirus and lentivirus expression systems. Protein production was detected in the baculovirus but not the lentivirus system. Attempts to solubilise these proteins were unsuccessful. Microscopy data for GFP-tagged proteins suggested that the receptors are produced and trafficked to the membrane. This, therefore, suggests that there may be issues with folding or solubilisation with the detergents tested.

7. Conclusions and future work

Due to similarities with other pLGICs, which are proven drug targets, and its implication with various conditions, GlyR is attracting interest as a potential therapeutic target (273,306,307). Structural studies of human GlyR have been confined to the homomeric forms (24,25), possibly due to experimental challenges linked with producing a heteromeric receptor. These homomeric GlyR models represent a major advance to the field. However, a structural model of a physiologically relevant heteromeric binding site is sought to progress the search for and characterization of ligands to inform early-stage drug discovery. GBP, a surrogate of the heteromeric $\beta^+\alpha^-$ interface has been characterised. The data support that GBP is a good model of the physiologically relevant binding site of glycine and can be used for the discovery of novel ligands.

Data presented in Chapter 4 and 5 emphasised a difference between GBP and heteromeric GlyR, residue E210 in loop C. In the sequence alignment of GBP and the β principal subunit, E210 is not present. The structures of GBP:strychnine and GBP:N-methylbicuculline suggest that E210 makes mostly van der Waals interactions with the ligands. Computational docking of a Maybridge small molecule library to GBP suggested that some compounds are predicted to interact directly with E210 making close van der Waals interactions and hydrogen bonds. These interactions are not representative of the GlyR heteromer, therefore these small molecules may not bind to the receptor. Due to these reasons, a GBP E210 deletion mutant may be required. The deletion will shorten loop C but should not affect structural stability drastically as it is a surface residue (308). This alteration will make GBP more representative of the GlyR $\beta^+\alpha^-$ interface and may allow loop C to close more tightly/form closer interactions with ligands in the orthosteric binding site.

Three biophysical techniques SPR, ITC and WF were used to investigate interactions between *Ac*AChBP and ligands. A comparison of three ligands interacting with *Ac*AChBP revealed only small differences between K_D values from ITC and the other two techniques. These differences could have been caused by incomplete ITC titration curves. It was concluded that WF is the best technique due to its sensitivity, ease of use and accuracy. As well as providing comparable K_D values, using the techniques in tandem can give information about the thermodynamic and kinetic parameters of binding interactions. Kinetic parameters were not deduced for the three ligands, however, further optimisation of experiments could lead to binding curves being observed. Optimisation could be carried out by changing several variables including temperature, flow rate and concentration to allow calculation of k_{on} and k_{off}.

N-methylbicuculline is a promiscuous ligand of the pLGIC family. Structures presented in this work show N-methylbicuculline bound to GBP and *Ac*AChBP, surrogates for GlyR and nAChR. Structures of the ligand bound the $GABA_AR$ (29,33) have also been reported in the literature. These data support previously published binding parameters that demonstrate bicuculline binding to all members of the pLGIC family (245-248). Improved structural data are required for *Ac*AChBP:N-methylbicuculline to confirm the orientation as suggested by the preliminary complex.

GBP can be used for early-stage drug discovery. The NMR and computational ligand screens identified a variety of potential ligands for GBP and *Ac*AChBP. Two of these compounds have been validated by WF, thereby preliminary validating the use of the screening methods with GBP. Further validation of selected compounds from the two screens is required. Direct ITC methods did not result in binding curves for selected small molecules, however, a competition-based approach may give some indication of affinity (309).

Attempts to produce recombinant proteins corresponding to a GlyR $\beta^+\alpha^-$ surrogate and HisR were unsuccessful. The former construct was designed using GlyR α_3 as a template. Selection of a HisR construct was carried out following the generation of homology models and sequence comparisons. Baculovirus and lentivirus expression systems were tested. However, the proteins were never solubilised by detergent, raising questions about trafficking to the membrane and the purification protocol. Microscopy data demonstrated both proteins were membrane-associated but this cannot be taken as being correctly folded and incorporated into the membrane. Patch-clamp electrophysiology could be used to confirm if the two proteins are functional at the membrane, thereby answering the questions raised by microscopy. The technique can be used to detect ionic currents passing through a membrane and thus ion channel opening and closing (310). If successful, additional data could be gathered about the binding interactions of the two proteins.

An alternative route to investigate binding in HisR is to produce an *Ac*AChBP-based surrogate, as suggested in Chapter 6. A similar approach as used with GBP can be taken to examine the role of different amino acid substitutions on binding the native ligand, histamine. This data could then be compared with published results from $GABA_AR$ studies, which suggest a potential role of histamine in modulating receptor function (21). Upon completion, the HisR surrogate, HisBP, could be utilised as a model of the orthosteric binding site to investigate interactions with other ligands and for early-stage drug discovery targeting mammalian or insect HisRs.

Over the last three years several structures of the tri-heteromeric $GABA_AR$ have been reported (27,29,31,33). Further insights about $GABA_AR$ receptor function and modulation have been gained from these structures. They also provide models for drug discovery and the investigation of binding in

the physiologically relevant receptor. Presently, GBP is the only model of a heteromeric GlyR. It is hoped that a structure of a GlyR heteromer is imminent and will revolutionise the field as it has done for $GABA_AR$. Nonetheless, even with a full length model of a GlyR heteromer, GBP can play a role in initial investigations and screening of ligand binding, as a cheap and easy to use resource.

8. References

1. Pless, S. A., and Sivilotti, L. G. (2019) A tale of ligands big and small: An update on how pentameric ligand-gated ion channels interact with agonists and proteins. *Curr Opin Physiol* **2**, 19-26
2. Nemecz, A., Prevost, M. S., Menny, A., and Corringer, P. J. (2016) Emerging Molecular Mechanisms of Signal Transduction in Pentameric Ligand-Gated Ion Channels. *Neuron* **90**, 452-470
3. Lemoine, D., Jiang, R., Taly, A., Chataigneau, T., Specht, A., and Grutter, T. (2012) Ligand-gated ion channels: new insights into neurological disorders and ligand recognition. *Chem Rev* **112**, 6285-6318
4. Limon, A., Reyes-Ruiz, J. M., and Miledi, R. (2012) Loss of functional GABAA receptors in the Alzheimer diseased brain. *PNAS* **109**, 10071-10076
5. Hernandez, C. M., Kayed, R., Zheng, H., Sweatt, J. D., and Dineley, K. T. (2010) Loss of α7 nicotinic receptors enhances β-amyloid oligomer accumulation, exacerbating early-stage cognitive decline and septohippocampal pathology in a mouse model of Alzheimer's disease. *J Neurosci* **30**, 2442-2453
6. Perez, X. A., and Quik, M. (2011) Focus on α4β2* and α6β2* nAChRs for Parkinson's disease therapeutics. *Mol Cell Pharmacol* **3**, 1-6
7. Wang, T., Ye, X., Bian, W., Chen, Z., Du, J., Li, M., Zhou, P., Cui, H., Ding, Y.-Q., Qi, S., Liao, M., and Sun, C. (2020) Allopregnanolone modulates GABAAR-dependent CaMKIIδ3 and BDNF to protect SH-SY5Y cells against 6-OHDA-induced damage. *Front Cell Neurosci*. 10.3389/fncel.2019.00569
8. Villa, C., Colombo, G., Meneghini, S., Gotti, C., Moretti, M., Ferini-Strambi, L., Chisci, E., Giovannoni, R., Becchetti, A., and Combi, R. (2019) CHRNA2 and nocturnal frontal lobe epilepsy: identification and characterization of a novel loss of function mutation. *Front Mol Neurosci*. 10.3389/fnmol.2019.00017
9. Jones, D. M., Esmaeil, N., Maren, S., and Macdonald, R. L. (2002) Characterization of pharmacoresistance to benzodiazepines in the rat Li-pilocarpine model of status epilepticus. *Epilepsy Res* **50**, 301-312
10. Wang, C.-H., Hernandez, C. C., Wu, J., Zhou, N., Hsu, H.-Y., Shen, M.-L., Wang, Y.-C., Macdonald, R. L., and Wu, D. C. (2018) A missense mutation A384P associated with human hyperekplexia reveals a desensitization site of glycine receptors. *J Neurosci* **38**, 2818-2831
11. Harvey, R. J., Depner, U. B., Wassle, H., Ahmadi, S., Heindl, C., Reinold, H., Smart, T. G., Harvey, K., Schutz, B., Abo-Salem, O. M., Zimmer, A., Poisbeau, P., Welzl, H., Wolfer, D. P., Betz, H., Zeilhofer, H. U., and Muller, U. (2004) GlyR alpha3: an essential target for spinal PGE2-mediated inflammatory pain sensitization. *Science* **304**, 884-887
12. Lynagh, T., and Laube, B. (2014) Opposing effects of the anesthetic propofol at pentameric ligand-gated ion channels mediated by a common site. *J Neurosci* **34**, 2155-2159
13. Dwoskin, L. P., Smith, A. M., Wooters, T. E., Zhang, Z., Crooks, P. A., and Bardo, M. T. (2009) Nicotinic receptor-based therapeutics and candidates for smoking cessation. *Biochem Pharmacol* **78**, 732-743
14. Brisson, A., and Unwin, P. N. T. (1985) Quaternary structure of the acetylcholine receptor. *Nature* **315**, 474-477
15. Toyoshima, C., and Unwin, N. (1990) Three-dimensional structure of the acetylcholine receptor by cryoelectron microscopy and helical image reconstruction. *J Cell Biol* **111**, 2623-2635
16. Unwin, N. (2005) Refined structure of the nicotinic acetylcholine receptor at 4A resolution. *J Mol Biol* **346**, 967-989

17. Gielen, M., and Corringer, P. J. (2018) The dual-gate model for pentameric ligand-gated ion channels activation and desensitization. *J Physiol* **596**, 1873-1902
18. Hilf, R. J., and Dutzler, R. (2008) X-ray structure of a prokaryotic pentameric ligand-gated ion channel. *Nature* **452**, 375-379
19. Hilf, R. J., and Dutzler, R. (2009) Structure of a potentially open state of a proton-activated pentameric ligand-gated ion channel. *Nature* **457**, 115-118
20. Corringer, P. J., Baaden, M., Bocquet, N., Delarue, M., Dufresne, V., Nury, H., Prevost, M., and Van Renterghem, C. (2010) Atomic structure and dynamics of pentameric ligand-gated ion channels: new insight from bacterial homologues. *J Physiol* **588**, 565-572
21. Miller, P. S., and Aricescu, A. R. (2014) Crystal structure of a Human GABA(A) receptor. *Nature* **512**, 270-275
22. Lynch, J. W. (2009) Native glycine receptor subtypes and their physiological roles. *Neuropharmacology* **56**, 303-309
23. Hibbs, R. E., and Gouaux, E. (2011) Principles of activation and permeation in an anion-selective Cys-loop receptor. *Nature* **474**, 54-60
24. Huang, X., Chen, H., Michelsen, K., Schneider, S., and Shaffer, P. L. (2015) Crystal structure of human glycine receptor-alpha3 bound to antagonist strychnine. *Nature* **526**, 277-280
25. Du, J., Lu, W., Wu, S., Cheng, Y., and Gouaux, E. (2015) Glycine receptor mechanism elucidated by electron cryo-microscopy. *Nature* **526**, 224-229
26. Morales-Perez, C. L., Noviello, C. M., and Hibbs, R. E. (2016) X-ray structure of the humanα4β2 nicotinic receptor. *Nature* **538**, 411-415
27. Laverty, D., Desai, R., Uchański, T., Masiulis, S., Stec, W. J., Malinauskas, T., Zivanov, J., Pardon, E., Steyaert, J., Miller, K. W., and Aricescu, A. R. (2019) Cryo-EM structure of the human α1β3γ2 GABAA receptor in a lipid bilayer. *Nature*. 10.1038/s41586-018-0833-4
28. Liu, S., Xu, L., Guan, F., Liu, Y. T., Cui, Y., Zhang, Q., Zheng, X., Bi, G. Q., Zhou, Z. H., Zhang, X., and Ye, S. (2018) Cryo-EM structure of the human alpha 5 beta 3 GABAA receptor. *Cell Res* **28**, 958-961
29. Masiulis, S., Desai, R., Uchański, T., Serna Martin, I., Laverty, D., Karia, D., Malinauskas, T., Zivanov, J., Pardon, E., Kotecha, A., Steyaert, J., Miller, K. W., and Aricescu, A. R. (2019) GABAA receptor signalling mechanisms revealed by structural pharmacology. *Nature* **565**, 454-459
30. Miller, P. S., Scott, S., Masiulis, S., De Colibus, L., Pardon, E., Steyaert, J., and Aricescu, A. R. (2017) Structural basis for GABAA receptor potentiation by neurosteroids. *Nat Struct Mol Biol* **24**, 986-992
31. Phulera, S., Zhu, H., Yu, J., Claxton, D. P., Yoder, N., Yoshioka, C., and Gouaux, E. (2018) Cryo-EM structure of the benzodiazepine-sensitive α1β1γ2S tri-heteromeric GABAA receptor in complex with GABA. *eLife*. 10.7554/eLife.39383
32. Zhu, S., Noviello, C. M., Teng, J., Walsh Jr, R. M., Kim, J. J., and Hibbs, R. E. (2018) Structure of a human synaptic $GABA_A$ receptor. *Nature* **559**, 67-72
33. Kim, J. J., Gharpure, A., Teng, J., Zhuang, Y., Howard, R. J., Zhu, S., Noviello, C. M., Walsh Jr, R. M., Lindahl, E., and Hibbs, R. E. (2020) Shared structural mechanisms of general anaesthetics and benzodiazepines. *Nature* **585**, 303-308
34. Brejc, K., van Dijk, W. J., Klaassen, R. V., Schuurmans, M., van Der Oost, J., Smit, A. B., and Sixma, T. K. (2001) Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. *Nature* **411**, 269-276
35. Celie, P. H., van Rossum-Fikkert, S. E., van Dijk, W. J., Brejc, K., Smit, A. B., and Sixma, T. K. (2004) Nicotine and carbamylcholine binding to nicotinic acetylcholine receptors as studied in AChBP crystal structures. *Neuron* **41**, 907-914
36. Hassaine, G., Deluz, C., Grasso, L., Wyss, R., Tol, M. B., Hovius, R., Graff, A., Stahlberg, H., Tomizaki, T., Desmyter, A., Moreau, C., Li, X. D., Poitevin, F., Vogel, H., and Nury, H. (2014) X-ray structure of the mouse serotonin 5-HT3 receptor. *Nature* **512**, 276-281

37. Giastas, P., Zouridakis, M., and Tzartos, S. J. (2017) Understanding structure-function relationships of the human neuronal acetylcholine receptor: insights from the first crystal structures of neuronal subunits. *Br J Pharmacol*. 10.1111/bph.13838
38. Jones, A. K. (2018) Genomics, cys-loop ligand-gated ion channels and new targets for the control of insect pests and vectors. *Curr Opin Insect Sci* **30**, 1-7
39. Grudzinska, J., Schemm, R., Haeger, S., Nicke, A., Schmalzing, G., Betz, H., and Laube, B. (2005) The β subunit determines the ligand binding properties of synaptic glycine receptors. *Neuron* **45**, 727-739
40. Cymes, G. D., and Grosman, C. (2016) Identifying the elusive link between amino acid sequence and charge selectivity in pentameric ligand-gated ion channels. *PNAS* **113**, E7106-E7115
41. Thompson, M. J., and Baenziger, J. E. (2020) Structural basis for the modulation of pentameric ligand-gated ion channel function by lipids. *BBA-Biomembranes*. 10.1016/j.bbamem.2020.183304
42. Kumar, A., Basak, S., Rao, S., Gicheru, Y., Mayer, M. L., Sansom, M. S. P., and Chakrapani, S. (2020) Mechanisms of activation and desensitization of full-length glycine receptor in lipid nanodiscs. *Nat Commun*. 10.1038/s41467-020-17364-5
43. Addona, G. H., Sandermann Jr., H., Kloczewiak, M. A., and Miller, K. W. (2003) Low chemical specificity of the nicotinic acetylcholine receptor sterol activation site. *BBA-Biomembranes* **1609**, 177-182
44. Sunshine, C., and McNamee, M. G. (1992) Lipid modulation of nicotinic acetylcholine receptor function: the role of neutral and negatively charged lipids. *BBA-Biomembranes* **1108**, 240-246
45. Jones, O. T., and McNamee, M. G. (1988) Annular and nonannular binding sites for cholesterol associated with the nicotinic acetylcholine receptor1. *Biochemistry* **27**, 2364-2374
46. Gharpure, A., Teng, J., Zhuang, Y., Noviello, C. M., Walsh Jr., R. M., Cabuco, R., Howard, R. J., Zaveri, N. T., Lindahl, E., and Hibbs, R. E. (2019) Agonist selectivity and ion permeation in the α3β4 ganglionic nicotinic receptor. *Neuron* **104**, 501-511
47. Walsh Jr., R. M., Roh, S.-H., Gharpure, A., Morales-Perez, C. L., Teng, J., and Hibbs, R. E. (2018) Structural principles of distinct assemblies of the human α4β2 nicotinic receptor. *Nature* **557**, 261-265
48. Langlhofer, G., and Villmann, C. (2016) The intracellular loop of the glycine receptor: it's not all about the size. *Front Mol Neurosci*. 10.3389/fnmol.2016.00041
49. Breitinger, U., Weinlander, K., Pechmann, Y., Langlhofer, G., Enz, R., Becker, C.-M., Sticht, H., Kneussel, M., Villmann, C., and Breitinger, H.-G. (2020) A proline-rich motif in the large intracellular loop of the glycine receptor α1 subunit interacts with the Pleckstrin homology domain of collybistin. *J Adv Res*. 10.1016/j.jare.2020.09.009
50. Kasaragod, V. B., and Schindelin, H. (2018) Structure–function relationships of glycine and GABAA receptors and their interplay with the scaffolding protein gephyrin. *Front Mol Neurosci*. 10.3389/fnmol.2018.00317
51. Albuquerque, E. X., Pereira, E. F. R., Alkondon, M., and Rogers, S. W. (2009) Mammalian nicotinic acetylcholine receptors: from structure to function. *Physiol Rev* **89**, 73-120
52. Paterson, D., and Nordberg, A. (2000) Neuronal nicotinic receptors in the human brain. *Prog Neurobiol* **61**, 75-111
53. Karlin, A. (2002) Emerging structure of the Nicotinic Acetylcholine receptors. *Nat Rev Neurosci* **3**, 102-114
54. Ramirez-Latorre, J., Yu, C. R., Qu, X., Perin, F., Karlin, A., and Role, L. (1996) Functional contributions of a5 subunit to neuronal acetylcholine receptor channels *Nature* **380**, 347-351
55. Wonnacott, S. (1990) The paradox of nicotinic acetylcholine receptor upregulation by nicotine. *Trends Pharmacol Sci* **11**, 216-219
56. Mao, D., Perry, D. C., Yasuda, R. P., Wolfe, B. B., and Kellar, K. J. (2008) The α4β2α5 nicotinic cholinergic receptor in rat brain is resistant to up-regulation by nicotine in vivo. *J Neurochem* **104**, 446-456

57. Broide, R. S., and Leslie, F. M. (1999) The α7 nicotinic acetylcholine receptor in neuronal plasticity. *Mol Neurobiol* **20**, 1-16
58. Clapham, D. E. (2007) Calcium signalling. *Cell* **131**, P1047-1058
59. Brams, M., Pandya, A., Kuzmin, D., van Elk, R., Krijnen, L., Yakel, J. L., Tsetlin, V., Smit, A. B., and Ulens, C. (2011) A structural and mutagenic blueprint for molecular recognition of strychnine and d-tubocurarine by different cys-loop receptors. *PLoS Biol*. 10.1371/journal.pbio.1001034
60. Taylor, P., Talley, T. T., Radic, Z., Hansen, S. B., Hibbs, R. E., and Shi, J. (2007) Structure-guided drug design: conferring selectivity among neuronal nicotinic receptor and acetylcholine binding protein subtypes. *Biochem Pharmacol* **74**, 1164-1171
61. Abraham, N., Healy, M., Ragnarsson, L., Brust, A., Alewood, P. F., and Lewis, R. J. (2017) Structural mechanisms for α-conotoxin activity at the human α3β4 nicotinic acetylcholine receptor. *Sci Rep*. 10.1038/srep45466
62. Nemecz, A., and Taylor, P. (2011) Creating an α7 nicotinic acetylcholine recognition domain from the acetylcholine-binding protein: crystallographic and ligand selectivity analyses. *J Biol Chem* **286**, 42555-42565
63. Shahsavar, A., Ahring, P. K., Olsen, J. A., Krintel, C., Kastrup, J. S., Balle, T., and Gajhede, M. (2015) Acetylcholine-binding protein engineered to mimic the α4-α4 binding pocket in α4β2 nicotinic acetylcholine receptors reveals interface specific interactions important for binding and activity. *Mol Pharmacol* **88**, 697-707
64. Li, S.-X., Huang, S., Bren, N., Noridomi, K., Dellisanti, C. D., Sine, S. M., and Chen, L. (2011) Ligand-binding domain of an α7-nicotinic receptor chimera and its complex with agonist. *Nat Neurosci* **14**, 1253-1259
65. Kesters, D., Thompson, A. J., Brams, M., van Elk, R., Spurny, R., Geitmann, M., Villalgordo, J. M., Guskov, A., Danielson, U. H., Lummis, S. C., Smit, A. B., and Ulens, C. (2013) Structural basis of ligand recognition in 5-HT3 receptors. *EMBO Rep* **14**, 49-56
66. Hansen, S. B., Talley, T. T., Radic, Z., and Taylor, P. (2004) Structural and ligand recognition characteristics of an acetylcholine-binding protein from Aplysia californica. *J Biol Chem* **279**, 24197-24202
67. Edink, E., Rucktooa, P., Retra, K., Akdemir, A., Nahar, T., Zuiderveld, O., van Elk, R., Janssen, E., van Nierop, P., van Muijlwijk-Koezen, J., Smit, A. B., Sixma, T. K., Leurs, R., and de Esch, I. J. P. (2011) Fragment growing induces conformational changes in acetylcholine-binding protein: a structural and thermodynamic analysis. *J Am Chem Soc* **133**, 5363-5371
68. Ceron-Carrasco, J. P., Jacquemin, D., Graton, J., Thany, S., and Le Questel, J.-Y. (2013) New insights on the molecular recognition of imidacloprid with Aplysia californica AChBP: a computational study. *J Phys Chem B* **117**, 3944-3953
69. Rajendra, S., Lynch, J. W., and Schofield, P. R. (1997) The glycine receptor. *Pharmacol Ther* **73**, 121-146
70. Yang, Z., Taran, E., Webb, T. I., and Lynch, J. W. (2012) Stoichiometry and subunit arrangement of α1β glycine receptors as determined by atomic force microscopy. *Biochemistry* **51**, 5229-5231
71. Becker, C. M., Hoch, W., and Betz, H. (1988) Glycine receptor heterogeneity in rat spinal cord during postnatal development. *EMBO J* **7**, 3717-3726
72. Jonsson, S., Morud, J., Pickering, C., Adermark, L., Ericson, M., and Soderpalm, B. (2012) Changes in glycine receptor subunit expression in forebrain regions of the Wistar rat over development. *Brain Res* **1446**, 12-21
73. Simon, J., Wakimoto, H., Fujita, N., Lalande, M., and Barnard, E. A. (2004) Analysis of the set of GABAA receptor genes in the human genome. *J Bio Chem* **279**, 41422-41435
74. Heinze, L., Harvey, R. J., Haverkamp, S., and Wassle, H. (2007) Diversity of glycine receptors in the mouse retina: localization of the a4 subunit. *J Comp Neurol* **500**, 693-707
75. Dawson, A., Trumper, P., Souza, J. O. d., Parker, H., Jones, M. J., Hales, T. G., and Hunter, W. N. (2019) Engineering a surrogate human heteromeric α/β glycine receptor orthosteric site

exploiting the structural homology and stability of acetylcholine-binding protein. *IUCRJ* **6**, 1014-1023
76. Hough, L. B. (2001) Genomics meets histamine receptors: new subtypes, new receptors. *Mol Pharmacol* **59**, 415-419
77. Yang, Q. Z., and Hatton, G. I. (1994) Histamine mediates fast synaptic inhibition of rat supraoptic oxytocin neurons via chloride conductance activation. *Neuroscience* **61**, 955-964
78. Saras, A., Gisselmann, G., Vogt-Eisele, A. K., Erlkamp, K. S., Kletke, O., Pusch, H., and Hatt, H. (2008) Histamine action on vertebrate GABAA receptors: direct channel gating and potentiation of GABA responses. *J Biol Chem* **283**, 10470-10475
79. Sachse, S., Peele, P., Silbering, A. F., Guhmann, M., and Galizia, C. G. (2006) Role of histamine as a putative inhibitory transmitter in the honeybee antennal lobe. *Front Zool*. 10.1186/1742-9994-3-22
80. Jones, A. K., Bera, A. N., Lees, K. and Sattelle, D. B. (2010) The cys-loop ligand-gated ion channel gene superfamily of the parasitoid wasp, Nasonia vitripennis. *Heredity* **104**, 247-259
81. Jones, A. K., and Sattelle, D. B. (2007) The cys-loop ligand-gated ion channel gene superfamily of the red flour beetle, Tribolium castaneum. *BMC Genomics*. 10.1186/1471-2164-8-327
82. Jones, A. K. a. S., D. B. (2006) The cys-loop ligand-gated ion channel superfamily of the honeybee, Apis mellifera. *Invert Neurosci* **6**, 123-132
83. Gisselmann, G., Pusch, H., Hovemann, B. T., and Hatt, H. (2001) Two cDNAs coding for histamine-gated ion channels in D. melanogaster. *Nat Neurosci* **5**, 11-12
84. Robson-Tull, J. (2018) Biophysical screening in fragment-based drug design: a brief overview. *Biosci Horiz*. 10.1093/biohorizons/hzy015
85. Erlanson, D. A., Fesik, S. W., Hubbard, R. E., Jahnke, W., and Jhoti, H. (2016) Twenty years on: the impact of fragments on drug discovery. *Nat Rev Drug Discov* **15**, 605-619
86. Bartoli, S., Fincham, C. I., and Fattori, D. (2006) The fragment-approach: An update. *Drug Discov Today Technol* **3**, 425-431
87. Rule, G. S., and Hitchens, K. T. (2006) *Fundamentals of protein nmr spectroscopy*, Springer, Dordrecht, The Netherlands
88. Kleckner, I. R., and Foster, M. P. (2011) An introduction to NMR-based approaches for measuring protein dynamics. *Biochim Biophys Acta* **1814**, 942-968
89. Hore, P. J. (1995) *Nuclear Magnetic Resonance*, Oxford University Press, US
90. Hiroaki, H. (2013) Recent applications of isotopic labeling for protein NMR in drug discovery. *Expert Opin Drug Discov* **8**, 523-536
91. Ma, R., Wang, P., Wu, J., and Ruan, K. (2016) Process of fragment-based lead discovery—a perspective from NMR. *Molecules*. 10.3390/molecules21070854
92. Sugiki, T., Furuita, K., Fujiwara, T., and Kojima, C. (2018) Current NMR techniques for structure-based drug discovery. *Molecules*. 10.3390/molecules23010148
93. Viegas, A., Manso, J., Nobrega, F. L., and Cabrita, E. J. (2011) Saturation-transfer difference (STD) NMR: a simple and fast method for ligand screening and characterization of protein binding. *J Chem Educ* **88**, 990-994
94. Dalvit, C., Fogliatto, G., Stewart, A., Veronesi, M., and Stockman, B. (2001) WaterLOGSY as a method for primary NMR screening: practical aspects and range of applicability. *J Biomol NMR* **21**, 349-359
95. Tengel, T., Fex, T., Emtenäs, H., Almqvist, F., Sethson, I. and Kihlberg, J. (2004) Use of 19F NMR spectroscopy to screen chemical libraries for ligands that bind to proteins. *Org. Biomol. Chem.* **2**, 725-731
96. Becker, W., Bhattiprolu, K. C., Gubensak, N., and Zangger, K. (2018) Investigating protein–ligand interactions by solution nuclear magnetic resonance spectroscopy. *Chemphyschem* **19**, 895-906
97. Orita, M., Warizaya, M., Amano, Y., Ohno, K., and Niimi, T. (2009) Advances in fragment-based drug discovery platforms. *Expert Opin Drug Discov* **4**, 1125-1144

98. Hassell, A. M., An, G., Bledsoe, R. K., Bynum, J. M., Carter III, H. L., Deng, S.-J. J., Gampe, R. T., Grisard, T. E., Madauss, K. P., Nolte, R. T., Rocque, W. J., Wang, L., Weaver, K. L., Williams, S. P., Wisely, G. B., Xu, R., and Shewchuk, L. M. (2007) Crystallization of protein–ligand complexes. *Acta Cryst D Biol Crystallogr* **63**, 72-79
99. Rupp, B. (2010) *Biomolecular crystallography: principles, practice, and application to structural biology*, Garland Science, New York
100. Clegg, W. (1998) *Crystal structure determination*, Oxford University Press, Oxford, UK
101. Hammond, C. (1990) *Introduction to crystallography*, Oxford University Press Oxford, UK
102. Patching, S. G. (2014) Surface plasmon resonance spectroscopy for characterisation of membrane protein–ligand interactions and its potential for drug discovery. *BBA-Biomembranes* **1838**, 43-55
103. Tang, Y., Zeng, X., and Liang, J. (2010) Surface plasmon resonance: an introduction to a surface spectroscopy technique. *J Chem Educ* **87**, 742-746
104. Mozsolits, H., Thomas, W. G., and Aguilar, M.-I. (2003) Surface plasmon resonance spectroscopy in the study of membrane-mediated cell signalling. *J Pept Sci* **9**,77-89
105. Li, Y.-J., Bi, L.-J., Zhang, X.-E., Zhou, Y.-F., Zhang, J.-B., Chen, Y.-Y., Li, W., and Zhang, Z.-P. (2006) Reversible immobilization of proteins with streptavidin affinity tags on a surface plasmon resonance biosensor chip. *Anal Bioanal Chem* **386**, 1321-1326
106. Johnsson, B., Lofas, S., and Lindquist, G. (1991) Immobilization of proteins to a carboxymethyldextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. *Anal Biochem* **198**, 268-277
107. Huber, W., Perspicace, S., Kohler, J., Muller, F., and Schlatter, D. (2004) SPR-based interaction studies with small molecular weight ligands using hAGT fusion proteins. *Anal Biochem* **333**, 280-288
108. Shepherd, C. A., Hopkins, A. L., and Navratilova, I. (2014) Fragment screening by SPR and advanced application to GPCRs. *Prog Biophys Mol Biol* **116**, 113-123
109. Do, T., Ho, F., Heidecker, B., Witte, K., Chang, L., and Lerner, L. (2008) A rapid method for determining dynamic binding capacity of resins for the purification of proteins. *Protein Expr Purif* **60**, 147-150
110. Wallner, J., Lhota, G., Jeschek, D., Mader, A., and Vorauer-Uhl, K. (2013) Application of bio-layer interferometry for the analysis of protein/liposome interactions. *J Pharm Biomed* **72**, 150-154
111. Wartchow, C. A., Podlaski, F., Li, S., Rowan, K., Zhang, X., Mark, D., and Huang, K. S. (2011) Biosensor-based small molecule fragment screening with biolayer interferometry. *J Comput Aided Mol Des* **25**, 669-676
112. Ghisaidoobe, A. B. T., and Chung, S. J. (2014) Intrinsic tryptophan fluorescence in the detection and analysis of proteins: a focus on förster resonance energy transfer techniques. *Int J Mol Sci* **15**, 22518-22538
113. Matveeva, E. G., Morisseau, C., Goodrow, M. H., Mullin, C., and Hammock, B. D. (2009) Tryptophan fluorescence quenching by enzyme inhibitors as a tool for enzyme active site structure investigation: epoxide hydrolase. *Curr Pharm Biotechnol* **10**, 589-599
114. Moller, M., and Denicola, A. (2002) Protein tryptophan accessibility studied by fluorescence quenching. *Biochem Mol Biol Edu* **30**, 175-178
115. Chen, Y., and Barkley, M. D. (1998) Toward understanding tryptophan fluorescence in proteins. *Biochemistry* **37**, 9976-9982
116. Yoav, S., Stern, J., Salama-Alber, O., Frolow, F., Anbar, M., Karpol, A., Hadar, Y., Morag, E., and Bayer, E. A. (2019) Directed Evolution of Clostridium thermocellum beta-Glucosidase A Towards Enhanced Thermostability. *Int J Mol Sci*. 10.3390/ijms20194701
117. Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* **2**, 2212-2221

118. Perozzo, R., Folkers, G., and Scapozza, L. (2004) Thermodynamics of protein–ligand interactions: history, presence, and future aspects. *J Recept Signal Transduct* **24**, 1-52
119. Freyer, M. W., and Lewis, E. A. (2008) Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions. *Methods Cell Biol* **84**, 79-113
120. Ward, W. H. J., and Holdgate, G. A. (2001) Isothermal titration calorimetry in drug discovery. in *Progress in medicinal chemistry* (King, F. D., and Oxford, A. W. eds.), Elsevier. pp 309-376
121. Holdgate, G., Geschwindner, S., Breeze, A., Davies, G., Colclough, N., Temesi, D., and Ward, L. (2013) Biophysical methods in drug discovery from small molecule to pharmaceutical. in *Protein-ligand interactions. methods in molecular biology (methods and protocols)* (Williams, M., and Daviter, T. eds.), Humana Press, Totowa, NJ. pp 327-355
122. Kozikowski, B. A., Burt, T. M., Tirey, D. A., Williams, L. E., Kuzmak, B. R., Stanton, D. T., Morand, K. L., and Nelson, S. L. (2003) The effect of room-temperature storage on the stability of compounds in DMSO. *J Biomol Screen* **8**, 205-209
123. Kozikowski, B. A., Burt, T. M., Tirey, D. A., Williams, L. E., Kuzmak, B. R., Stanton, D. T., Morand, K. L., and Nelson, S. L. (2003) The effect of freeze/thaw cycles on the stability of compounds in DMSO. *J Biomol Screen* **8**, 210-215
124. Waybright, T. J., Britt, J. R., and McCloud, T. G. (2009) Overcoming problems of compound storage in DMSO: solvent and process alternatives. *J Biomol Screen* **14**, 708-715
125. Jackson, M., and Mantsch, H. H. (1991) Beware of proteins in DMSO. *Biochim Biophys Acta* **1078**, 231-235
126. Tjernberg, A., Markova, N., Griffiths, W. J., and Hallen, D. (2006) DMSO-related effects in protein characterization. *J Biomol Screen* **11**, 131-137
127. Di, L., Fish, P. V., and Mano, T. (2012) Bridging solubility between drug discovery and development. *Drug Discov Today* **17**, 486-495
128. Lin, X., Li, X., and Lin, X. (2020) A review on applications of computational methods in drug screening and design. *Molecules*. 10.3390/molecules25061375
129. Ferreira, L. G., Dos Santos, R. N., Oliva, G., and Andricopulo, A. D. (2015) Molecular docking and structure-based drug design strategies. *Molecules* **20**, 13384-13421
130. Sliwoski, G., Kothiwale, S., Meiler, J., and Lowe Jr., E. W. (2014) Computational methods in drug discovery. *Pharmacol Rev* **66**, 334-395
131. Schaller, D., Sribar, D., Noonan, T., Deng, L., Nguyen, T. N., Pach, S., Machalz, D., Bermudez, M., and Wolber, G. (2020) Next generation 3D pharmacophore modeling. *Wiley Interdiscip Rev Comput Mol Sci*. 10.1002/wcms.1468
132. Wermuth, C. G., Ganellin, C., Lindberg, P., and Mitscher, L. A. (1998) Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure Appl Chem* **70**, 1129-1143
133. Sun, H. (2008) Pharmacophore-based virtual screening. *Curr Med Chem* **15**, 1018-1024
134. Kumar, A., Rathi, E., and Kini, S. G. (2020) Identification of potential tumour-associated carbonic anhydrase isozyme IX inhibitors: atom-based 3D-QSAR modelling, pharmacophore-based virtual screening and molecular docking studies. *J Biomol Struct Dyn* **38**, 2156-2170
135. Erlanson, D. A. (2006) Fragment-based lead discovery: a chemical update. *Curr Opin Biotech* **17**, 643-652
136. Hartenfeller, M., and Schneider, G. (2010) De novo drug design. in *Chemoinformatics and computational chemical biology* (Bajorath, J. ed.), Humana Press, Totowa, NJ. pp 299-323
137. Pundir, S., Martin, M. J., and O'Donovan, C. (2017) Uniprot protein knowledgebase. *Methods Mol. Biol.* **1558**, 41-55
138. Booth, W. T., Schlachter, C. R., Pote, S., Ussin, N., Mank, N. J., Klapper, V., Offermann, L. R., Tang, C., Hurlburt, B. K., and Chruszcz, M. (2018) Impact of an N-terminal polyhistidine tag on protein thermal stability. *ACS Omega* **3**, 760-768

139. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-1191
140. Berman, H., Henrick, K., and Nakamura, H. (2003) Announcing the worldwide Protein Data Bank. *Nat Struct Mol Biol* **10**, 980
141. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845-858
142. Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., and Wolfson, H. J. (2005) PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res* **33**, W363-W367
143. Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Cryst D Biol Crystallogr* **60**, 2256-2268
144. Slabinski, L., Jaroszewski, L., Rychlewski, L., Wilson, I. A., Lesley, S. A., and Godzik, A. (2007) XtalPred: a web server for prediction of protein crystallizability. *Bioinformatics* **23**, 3403-3405
145. Jensen, M. A., Fukushima, M., and Davis, R. W. (2010) DMSO and betaine greatly improve amplification of GC-rich constructs in de novo synthesis. *Plos One*. 10.1371/journal.pone.0011024
146. Korbie, D. J., and Mattick, J. S. (2008) Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat Protoc* **3**, 1452-1456
147. Woodman, M. E. (2008) Direct PCR of intact bacteria (colony PCR). *Curr Protoco Microbiol* **9**, A.3D.1-A.3D.6
148. Roberts, R. J. (2005) How restriction enzymes became the workhorses of molecular biology. *Proc Natl Acad Sci USA* **102**, 5905-5908
149. Waring, M. J. (1965) Complex formation between ethidium bromide and nucleic acids. *J Mol Biol* **13**, 269-282
150. Elegheert, J., Behiels, E., Bishop, B., Scott, S., Woolley, R. E., Griffiths, S. C., Byrne, E. F. X., Chang, V. T., Stuart, D. I., Jones, E. Y., Siebold, C., and Aricescu, A. R. (2018) Lentiviral transduction of mamalian cells for fast, scalable and high-level production of soluble and membrane proteins. *Nat Protoc* **13**, 2991-3017
151. Blommel, P. G., Becker, K. J., Duvnjak, P., and Fox, B. G. (2008) Enhanced bacterial protein expression during auto-induction obtained by Alteration of lac repressor dosage and medium composition. *Biotechnol Prog* **23**, 585-598
152. Wittig, I., Braun, H., and Schagger, H. (2006) Blue native PAGE. *Nat Protoc* **1**, 418-428
153. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein identification and analysis tools on the ExPASy server. in *The Proteomics Protocols Handbook* (Walker, J. M. ed.), Humana Press Inc., Totowa, NJ. pp 571-607
154. Jancarik, J., and Kim, S.-H. (1991) Sparse matrix sampling: a screening method for crystallization of proteins. *J Appl Cryst* **24**, 409-411
155. Gorrec, F. (2009) The MORPHEUS protein crystallization screen. *J Appl Crystallogr* **42**, 1035-1042
156. Parambil, J. V., and Heng, J. Y. Y. (2017) Seeding in crystallisation. in *Engineering crystallography: from molecule to crystal to functional form* (Roberts, K., Docherty, R., and Tamura, R. eds.), Springer, Dordrecht. pp 235-245
157. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. W. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Cryst D Biol Crystallogr* **67**, 271-281
158. Kabsch, W. (2010) Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132
159. Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? *Acta Crystallogr D* **69**, 1204-1214
160. Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., and Bricogne, G. (2011) Data processing and analysis with the autoPROC toolbox. *Acta Cryst* **D67**, 293-302

161. Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I. D., Sauter, N. K., and Evans, G. (2018) DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr D Struct Biol* **74**, 85-97
162. Kantardjieff, K. A., and Rupp, B. (2003) Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci* **12**, 1865-1871
163. Matthews, B. W. (1968) Solvent content of protein crystals. *J Mol Biol* **33**, 491-497
164. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674
165. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242
166. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501
167. Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G. N. (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr D Biol Crystallogr* **60**, 2184-2195
168. Sun, Z., Liu, Q., Qu, G., Feng, Y., and Reetz, M. T. (2019) Utility of B-Factors in protein science: interpreting rigidity, flexibility, and internal motion and engineering thermostability. *Chem Rev* **119**, 1626-1665
169. Wiseman, T., Williston, S., Brandts, J. F., and Lin, L. N. (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal Biochem* **179**, 131-137
170. Englebienne, P., Van Hoonacker, A., and Verhas, M. (2003) Surface plasmon resonance: principles, methods and applications in biomedical sciences. *Spectroscopy* **17**, 255-273
171. GEHealthcare. (2010) Biacore T200 software handbook. AA Ed., General Electric Company, UK
172. Brayton, C. F. (1986) Dimethyl sulfoxide (DMSO): a review. *Cornell Vet* **76**, 61-90
173. Pankov, G. (2019) *Enabling early stage drug discovery targeting high molecular mass penicillin-binding proteins in Yersinia pestis.* Doctor of Philosophy, University of Dundee
174. Lee, B., and Richards, R. M. (1971) The interpretation of protein structures: estimation of static accessibility *J Mol Biol* **55**, 379-400
175. Abagyan, R. A., and Totrov, M. M. (1994) Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J.Mol. Biol.* **235**, 983-1002
176. Abagyan, R. A., Totrov, M. M. and Kuznetsov, D. N. (1994) ICM - a new method for protein modeling and design. Applications to docking and structure prediction from the distorted native conformation. *J. Comp. Chem.* **15**, 488-506
177. Abagyan, R., Orry, A., Raush, E., and Totrov, M. (2020) *Small molecule docking*, Molsoft LLC
178. Stjernschantz, E. (2002) *Evaluation of docking in ICM as a tool in structure-based drug design.* Masters, Uppsala University
179. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1996) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* **8**, 127-134
180. Ma, W., Yang, L., and He, L. (2018) Overview of the detection methods for equilibrium dissociation constant KD of drug-receptor interaction. *J Pharm Anal* **8**, 147-152
181. Callis, P. R. (2007) Exploring the electrostatic landscape of proteins with tryptophan fluorescence. in *Reviews in fluorescence* (Geddes, C. D. ed.), Springer, NY. pp 199-248
182. Gill, H. S. (2010) Evaluating the efficacy of tryptophan fluorescence and absorbance as a selection tool for identifying protein crystals. *Acta Crystallogr F* **66**, 364-372
183. Falconer, R. J., Penkova, A., Jelesarov, I., and Collins, B. M. (2010) Survey of the year 2008: applications of isothermal titration calorimetry. *J Mol Recognit* **23**, 395-413

184. Ladbury, J. E., and Chowdhry, B. Z. (1996) Sensing the heat: the application of isothermal titration calorimetry to thermodynamic studies of biomolecular interactions. *Chem Biol* **3**, 791-801
185. Nguyen, H. H., Park, J., Kang, S., and Kim, M. (2015) Surface plasmon resonance: a versatile technique for biosensor applications. *Sensors* **15**, 10481-10510
186. Feduccia, A. A., Chatterjee, S., and Bartlett, S. E. (2012) Neuronal nicotinic acetylcholine receptors: neuroplastic changes underlying alcohol and nicotine addictions. *Front Mol Neurosci*. 10.3389/fnmol.2012.00083
187. Armstrong, D. W., Wang, X., Lee, J., and Liu, Y. (1999) Enantiomeric composition of nornicotine, anatabine, and anabasine in tobacco. *Chirality* **11**, 82-84
188. Devlin, J. P., Edwards, O. E., Gorham, P. R., Hunter, N. R., Pike, R. K., and Stavric, B. (1977) Anatoxin-a, a toxic alkaloid from Anabaena flos-aquae. *Can J Chem* **5**, 1367-1371
189. Nguyen, L. A., He, H., and Pham-Huy, C. (2006) Chiral drugs: an overview. *Int J Biomed Sci* **2**, 85-100
190. Andersson, S., and Allenmark, S. G. (2002) Preparative chiral chromatographic resolution of enantiomers in drug discovery. *J Biochem Bioph Meth* **54**, 11-23
191. Toyo'oka, T. (2002) Resolution of chiral drugs by liquid chromatography based upon diastereomer formation with chiral derivatization reagents. *J Biochem Bioph Meth* **54**, 25-56
192. Macallan, D. R. E., Lunt, G. G., Wonnacott, S., Swanson, K. L., Rapoport, H., and Albuquerque, E. X. (1988) Methyllycaconitine and (+)-anatoxin-a differentiate between nicotinic receptors in vertebrate and invertebrate nervous systems. *Febs Lett* **226**, 357-363
193. Wonnacott, S., and Gallagher, T. (2006) The chemistry and pharmacology of anatoxin-a and related homotropanes with respect to nicotinic acetylcholine receptors. *Mar Drugs* **4**, 228-254
194. Xing, H., Keshwah, S., Rouchaud, A., and Kem, W. R. (2020) A pharmacological comparison of two isomeric nicotinic receptor agonists: the marine toxin isoanatabine and the tobacco alkaloid anatabine. *Mar Drugs*. 10.3390/md18020106
195. Aronstam, R. S., and Witkop, B. (1981) Anatoxin-a interactions with cholinergic synaptic molecules. *Proc Natl Acad Sci USA* **78**, 4639-4643
196. Fokkens, J., and Klebe, G. (2006) A simple protocol to estimate differences in protein binding affinity for enantiomers without prior resolution of racemates. *Angew Chem Int Ed* **45**, 985-989
197. Peters, W. B., Frasca, V., and Brown, R. K. (2009) Recent Developments in Isothermal Titration Calorimetry Label Free Screening *Comb Chem High Throughput Screen* **12**, 772-790
198. Thordarson, P. (2011) Determining association constants from titration experiments in supramolecular chemistry. *Chem Soc Rev* **40**, 1305-1323
199. Turnbull, W. B. (2005) Divided we fall? studying low affinity fragments of ligands by ITC. *MicroCal Application Note*, 2-3
200. Broecker, J., Vargas, C., and Keller, S. (2011) Revisiting the optimal c value for isothermal titration calorimetry. *Anal Biochem* **418**, 307-309
201. Andrews, S. P., Brown, G. A., and Christopher, J. A. (2014) Structure-based and fragment-based GPCR drug discovery. *ChemMedChem* **9**, 256-275
202. Schulz, M. N., and Hubbard, R. E. (2009) Recent progress in fragment-based lead discovery. *Curr Opin Pharmacol* **9**, 615-621
203. Freire, E. (2008) Do enthalpy and entropy distinguish first in class from best in class? *Drug Discov Today* **13**, 869-874
204. Malvern. (2015) *MicroCal PEAQ-ITC analysis software user manual*, Malvern Instruments, UK
205. Warn, J. R. W., and Peters, A. P. H. (1996) *Concise chemical thermodynamics*, 2nd ed., Chapman & Hall, London
206. Berg, J. M., Tymoczko, J. L., and Stryer, L. (2002) *Biochemistry*, 5th ed., W. H. Freeman and Co, New York
207. Chothia, C., and Janin, J. (1975) Principles of protein-protein recognition. *Nature* **256**, 705-708

208. Tame, J. R. H., O'Brien, R., and Ladbury, J. E. (1998) Isothermal titration calorimetry of biomolecules. in *Biocalorimetry applications of calorimetry in the biological sciences* (Ladbury, J. E., and Chowdhry, B. Z. eds.), John Wiley & Sons, England. pp 27-38
209. Huang, Y., Kong, K., and Wood, J. L. (2018) Total synthesis of (+)- and (±)-hosieine A. *Angew Chem Int Ed* **57**, 7664-7667
210. Pouny, I., Batut, M., Vendier, L., David, B., Yi, S., Sautel, F., Arimondo, P. B., and Massiot, G. (2014) Cytisine-like alkaloids from Ormosia hosiei Hemsl. & E.H. Wilson. *Phytochemistry* **107**, 97-101
211. Jones, M. J., Dawson, A., Hales, T. G., and Hunter, W. N. (2019) A structural rationale for N-methylbicuculline acting as a promiscuous competitive antagonist of inhibitory pentameric ligand-gated ion channels. *ChemBioChem* **21**, 1526-1533
212. Laskowski, R. A., and Swindells, M. B. (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* **51**, 2778-2786
213. Bento, A. P., Gaulton, A., Hersey, A., Bellis, L. J., Chambers, J., Davies, M., Kruger, F. A., Light, Y., Mak, L., McGlinchey, S., Nowotka, M., Papadatos, G., Santos, R., and Overington, J. P. (2014) The ChEMBL bioactivity database: an update. *Nucleic Acids Res* **42**, D1083-1090
214. Maguire, J. J., Kuc, R. E., and Davenport, A. P. (2012) Radioligand binding assays and their analysis. *Methods Mol Biol* **897**, 31-77
215. Marechal, E. (2011) Measuring bioactivity: Ki, IC50 and EC50. in *Chemogenomics and chemical genetics a user's introduction for biologists, chemists and informaticians* (Marechal, E., Roy, S., and Lafanechere, L. eds.), Springer, Berlin. pp 55-65
216. Ulens, C., Akdemir, A., Jongejan, A., van Elk, R., Bertrand, S., Perrakis, A., Leurs, R., Smit, A. B., Sixma, T. K., Bertrand, D., and de Esch, I. J. P. (2009) Use of acetylcholine binding protein in the search for novel α7 nicotinic receptor ligands. In silico docking, pharmacological screening, and X-ray analysis. *J Med Chem* **52**, 2372-2383
217. Faundez-Parraguez, M., Rabelo-Farias, N., Gonzalez-Gutierrez, J. P., Etcheverry-Berrios, A., Alzate-Morales, J., Adasme-Carreno, F., Varas, R., Bermudez, I., and Iturriaga-Vasquez, P. (2013) Neonicotinic analogues: selective antagonists for α4β2 nicotinic acetylcholine receptors. *Bioorg Med Chem* **21**, 2687-2694
218. Tomizawa, M., Lee, D. L., and Casida, J. E. (2000) Neonicotinoid insecticides: molecular features conferring selectivity for insect versus mammalian nicotinic receptors. *J Agric Food Chem* **48**, 6016-6024
219. Charton, Y., Guillonneau, C., Lockhart, B., Lestage, P., and Goldstein, S. (2008) Preparation and affinity profile of novel nicotinic ligands. *Bioorg Med Chem Lett* **18**, 2188-2193
220. Dallanoce, C., Magrone, P., Matera, C., Presti, L. L., Amici, M. D., Riganti, L., Clementi, F., Gotti, C., and Micheli, C. D. (2010) Synthesis of novel chiral Δ2-isoxazoline derivatives related to ABT-418 and estimation of their affinity at neuronal nicotinic acetylcholine receptor subtypes. *Eur J Med Chem* **45**, 5594-5601
221. Efanage, S. M. N., Tu, Z., von Hohenberg, K., Francesconi, L., Howell, R. C., Rampersad, M. V., Todaro, L. J., Papke, R. L., and HKung, M.-P. (2001) 2-(2-Piperidyl)- and 2-(2-pyrrolidyl)chromans as nicotine agonists: synthesis and preliminary pharmacological characterization. *J Med Chem* **44**, 4704-4715
222. Guandalini, L., Norcini, M., Varani, K., Pistolozzi, M., Gotti, C., Bazzicalupi, C., Martini, E., Dei, S., Manetti, D., Scapecchi, S., Teodori, E., Bertucci, C., Ghelardini, C., and Romanelli, M. N. (2007) Design, synthesis, and preliminary pharmacological evaluation of new quinoline derivatives as nicotinic ligands. *J Med Chem* **50**, 4993-5002
223. Lin, N.-H., Gunn, D. E., Ryther, K. B., Garvey, D. S., Donnelly-Roberts, D. L., Decker, M. W., Brioni, J. D., Buckley, M. J., Rodrigues, A. D., Marsh, K. G., Anderson, D. J., Buccafusco, J. J., Prendergast, M. A., Sullivan, J. P., Williams, M., Arneric, S. P., and Holladay, M. W. (1997) Structure−activity studies on 2-methyl-3-(2(S)-pyrrolidinylmethoxy)pyridine (ABT-089): an

orally bioavailable 3-pyridyl ether nicotinic acetylcholine receptor ligand with cognition-enhancing properties. *J Med Chem* **40**, 385-390

224. Matera, C., Pucci, L., Fiorentini, C., Fucile, S., Missale, C., Grazioso, G., Clementi, F., Zoli, M., Amici, M. D., Gotti, C., and Dallanoce, C. (2015) Bifunctional compounds targeting both D2 and non-α7 nACh receptors: Design, synthesis and pharmacological characterization. *Eur J Med Chem* **101**, 367-383
225. Pallavicini, M., Bolchi, C., Binda, M., Cillia, A., Clementi, F., Ferrara, R., Fumagalli, L., Gotti, C., Moretti, M., Pedretti, A., Vistoli, G., and Valoti, E. (2009) 5-(2-Pyrrolidinyl)oxazolidinones and 2-(2-pyrrolidinyl)benzodioxanes: Synthesis of all the stereoisomers and α4β2 nicotinic affinity. *Bioorg Med Chem Lett* **19**, 854-859
226. Tasso, B., Novelli, F., Sparatore, F., Fasoli, F., and Gotti, C. (2013) (+)-Laburnamine, a natural selective ligand and partial agonist for the α4β2 nicotinic receptor subtype. *J Nat Prod* **76**, 727-731
227. Bunnelle, W. H., Tietje, K. R., Frost, J. M., Peters, D., Ji, J., Li, T., Scanio, M. J. C., Shi, L., Anderson, D. J., Dyhring, T., Gronlien, J. H., Ween, H., Thorin-Hagene, K., and Meyer, M. D. (2009) Octahydropyrrolo[3,4-c]pyrrole: a diamine scaffold for construction of either α4β2 or α7-selective nicotinic acetylcholine receptor (nAChR) ligands. substitutions that switch subtype selectivity. *J Med Chem* **52**, 4126-4141
228. Holladay, M. W., Dart, M. J., and Lynch, J. K. (1997) Neuronal nicotinic acetylcholine receptors as targets for drug discovery. *J Med Chem* **40**, 4169-4194
229. Macor, J. E., Gurley, D., Lanthorn, T., Loch, J., Mack, R. A., Mullen, G., Tran, O., Wright, N., and Gordon, J. C. (2001) The 5-HT3 antagonist tropisetron (ICS 205-930) is a potent and selective α7 nicotinic receptor partial agonist. *Bioorg Med Chem* **11**, 319-321
230. Mazurov, A. A., Kombo, D. C., Hauser, T. A., Miao, L., Dull, G., Genus, J. F., Fedorov, N. B., Benson, L., Sidach, S., Xiao, Y., Hammond, P. S., James, J. W., Miller, C. H., and Yohannes, D. (2012) Discovery of (2S,3R)-N-[2-(pyridin-3-ylmethyl)-1-azabicyclo[2.2.2]oct-3-yl]benzo[b]furan-2-carboxamide (TC-5619), a selective α7 nicotinic acetylcholine receptor agonist, for the treatment of cognitive disorders. *J Med Chem* **55**, 9793-9809
231. Tomizawa, M., and Casida, J. E. (2011) Unique neonicotinoid binding conformations conferring selective receptor interactions. *J Agric Food Chem* **59**, 2825-2828
232. Tomizawa, M., Talley, T. T., Maltby, D., Durkin, K. A., Medzihradszky, K. F., Burlingame, A. L., Taylor, P., and Casida, J. E. (2007) Mapping the elusive neonicotinoid binding site. *PNAS* **104**, 9075-9080
233. Ussing, C. A., Hansen, C. P., Petersen, J. G., Jensen, A. A., Rohde, L. A. H., Ahring, P. K., Nielsen, E. O., Kastrup, J. S., Gajhede, M., Frølund, B., and Balle, T. (2013) Synthesis, pharmacology, and biostructural characterization of novel α4β2 nicotinic acetylcholine receptor agonists. *J Med Chem* **56**, 940-951
234. Zhang, H.-K., Eaton, J. B., Yu, L.-F., Nys, M., Mazzolari, A., van Elk, R., Smit, A. B., Alexandrov, V., Hanania, T., Sabath, E., Fedolak, A., Brunner, D., Lukas, R. J., Vistoli, G., Ulens, C., and Kozikowski, A. P. (2012) Insights into the structural determinants required for high-affinity binding of chiral cyclopropane-containing ligands to α4β2-nicotinic acetylcholine receptors: an integrated approach to behaviorally active nicotinic ligands. *J Med Chem* **55**, 8028-8037
235. Sharples, C. G., Karig, G., Simpson, G. L., Spencer, J. A., Wright, E., Millar, N. S., Wonnacott, S., and Gallagher, T. (2002) Synthesis and pharmacological characterization of novel analogues of the nicotinic acetylcholine receptor agonist (+/-)-UB-165. *J Med Chem* **45**, 3235-3245
236. Villeneuvea, G., Cécyre, D., Lejeune, H., Drouin, M., Lan, R., and Quirion, R. (2003) Rigidified acetylcholine mimics: conformational requirements for binding to neuronal nicotinic receptors. *Bioorg Med Chem* **13**, 3847-3851
237. Dalvit, C., Flocco, M., Knapp, S., Mostardini, M., Perego, R., Stockman, B. J., Veronesi, M., and Varasi, M. (2002) High-throughput NMR-based screening with competition binding experiments. *J Am Chem Soc* **124**, 7702-7709

238. Ciesielski, G. L., Hytonen, V. P., and Kaguni, L. S. (2016) Biolayer interferometry: a novel method to elucidate protein–protein and protein–DNA interactions in the mitochondrial DNA replisome. *Methods Mol Biol* **1351**, 223-231
239. Laigre, E., Goyard, D., Tiertant, C., Dejeu, J., and Renaudet, O. (2018) The study of multivalent carbohydrate–protein interactions by bio-layer interferometry. *Org Biomol Chem*. : 10.1039/c8ob01664j
240. Yang, D., Singh, A., Wu, H., and Kroe-Barrett, R. (2016) Comparison of biosensor platforms in the evaluation of high affinity antibody-antigen binding kinetics. *Anal Biochem* **508**, 78-96
241. Tullman, J., Christensen, M., Kelman, Z., and Marino, J. P. (2020) A ClpS-based N-terminal amino acid binding reagent with improved thermostability and selectivity. *Biochem Eng J*. 10.1016/j.bej.2019.107438
242. Zarbin, M. A., Wamsley, J. K., and Kuhar, M. J. (1981) Glycine receptor: light microscopic autoradiographic localization with [3H]strychnine. *J Neurosci* **1**, 532-547
243. Huang, X., Shaffer, P. L., Ayube, S., Bregman, H., Chen, H., Lehto, S. G., Luther, J. A., Matson, D. J., McDonough, S. I., Michelsen, K., Plant, M. H., Schneider, S., Simard, J. R., Teffera, Y., Yi, S., Zhang, M., DiMauro, E. F., and Gingras, J. (2017) Crystal structures of human glycine receptor α3 bound to a novel class of analgesic potentiators. *Nat Struct Mol Biol* **24**, 108-113
244. Jensen, A. A., Gharagozloo, P., Birdsall, N. J. M., and Zlotos, D. P. (2006) Pharmacological characterisation of strychnine and brucine analogues at glycine and α7 nicotinic acetylcholine receptors. *Eur J Pharmacol* **539**, 27-33
245. James, O. T., Livesey, M. R., Qiu, J., Dando, O., Bilican, B., Haghi, G., Rajan, R., Burr, K., Hardingham, G. E., Chandran, S., Kind, P. C., and Wyllie, D. J. A. (2014) Ionotropic GABA and glycine receptor subunit composition in human pluripotent stem cell-derived excitatory cortical neurones. *J Physiol* **592**, 4353-4363
246. Rothlin, C. V., Katz, E., Verbitsky, M., and Elgoyhen, B. (1999) The α9 nicotinic acetylcholine receptor shares pharmacological properties with type A γ-aminobutyric acid, glycine, and type 3 serotonin receptors. *Mol Pharmacol* **55**, 248-254
247. Sun, H., and Machu, T. K. (2000) Bicuculline antagonizes 5-HT and 3A a2 glycine receptors expressed in Xenopus oocytes. *Eur J Pharmacol* **391**, 243-249
248. Tokutomi, N., Kaneda, M., and Akaike, N. (1989) What confers specificity on glycine for its receptor site? *Br J Pharmacol* **97**, 353-360
249. Sah, P., and Faber, E. S. L. (2002) Channels underlying neuronal calcium-activated potassium currents. *Prog Neurobiol* **66**, 345-353
250. Wohri, A. B., Hillertz, P., Eriksson, P., Meuller, J., Dekker, N., and Snijder, A. (2013) Thermodynamic studies of ligand binding to the human homopentameric glycine receptor using isothermal titration calorimetry. *Mol Membr Biol* **30**, 169-183
251. (2014-2015) *CRC handbook of chemistry and physics*, 95th ed., CRC Press LLC, Boca Raton
252. Anand, R., Peng, X., and Lindstrom, J. (1993) Homomeric and native a7 acetylcholine receptors exhibit remarkably similar but non-identical pharmacological properties, suggesting that the native receptor is a heteromeric protein complex. *Febs Lett* **327**, 241-246
253. Olsen, R. W., Ban, M., and Miller, T. (1976) Studies on the neuropharmacological activity of bicuculline and related compounds. *Brain Res* **102**, 283-299
254. Krall, J., Balle, T., Krogsgaard-Larsen, N., Sorensen, T. E., Krogsgaard-Larsen, P., Kristiansen, U., and Frolund, B. (2015) GABAA receptor partial agonists and antagonists: structure, binding mode, and pharmacology. in *Diversity and functions of GABA receptors: a tribute to Hanns Möhler, part a* (Rudolph, U. ed.), Academic Press, USA. pp 201-227
255. Liebschner, D., Afonine, P. V., Moriarty, N. W., Poon, B. K., Sobolev, O. V., Terwilliger, T. C., and Adams, P. D. (2017) Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Crystallogr D Struct Biol* **73**, 148-157
256. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J.,Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R.

J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Struct Biol* **66**, 213-221
257. Pong, S. F., and Jr., L. T. G. (1972) N-methyl bicuculline, a convulsant more potent than bicuculline. *Brain Res* **42**, 486-490
258. Zhang, J., Xue, F., and Chang, Y. (2008) Structural determinants for antagonist pharmacology that distinguish the ρ_1 $GABA_C$ receptor from $GABA_A$ receptors. *Mol Pharmacol* **74**, 941-951
259. Renauld, J.-P., and Delsuc, M.-A. (2009) Biophysical techniques for ligand screening and drug design. *Curr Opin Pharmacol* **9**, 622-628
260. Ruddikeit, L., van Deursen, R., Blum, L. C., and Reymond, J.-L. (2012) Enumeration of 166 billion organic small molecules in the chemical universe database GDB-17. *J Chem Inf Model* **52**, 2864-2875
261. Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J. (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* **23**, 3-25
262. Jhoti, H., Cleasby, A., Verdonk, M., and Williams, G. (2007) Fragment-based screening using X-ray crystallography and NMR spectroscopy. *Curr Opin Chem Biol* **11**, 485-493
263. Congreve, M., Carr, R., Murray, C., and Jhoti, H. (2003) A 'Rule of Three' for fragment-based lead discovery. *Drug Discov Today* **8**, 876-877
264. Congreve, M., Murray, C. W., and Blundell, T. L. (2005) Keynote review: structural biology and drug discovery. *Drug Discov Today* **10**, 895-907
265. Hesterkamp, T., and Whittaker, M. (2008) Fragment-based activity space: smaller is better. *Curr Opin Chem Biol* **12**, 260-268
266. Siegal, G., Ab, E., and Schultz, J. (2007) Integration of fragment screening and library design. *Drug Discov Today* **12**, 1032-1039
267. Jordan, J. B., Poppe, L., Xia, X., Cheng, A. C., Sun, Y., Michelsen, K., Eastwood, H., Schnier, P. D., Nixey, T. and Zhong, W. (2011) Fragment based drug discovery: practical implementation based on 19F NMR spectroscopy. *J Med Chem* **55**, 678-687
268. Norton, R. S., Leung, E. W. W., Chandrashekaran, I. R., and MacRaild, C. A. (2016) Applications of 19F-NMR in fragment-based drug discovery. *Molecules*. 10.3390/molecules21070860
269. Dalvit, C., Fagerness, P. E., Hadden, D. T. A., Sarver, R. W., and Stockman, B. J. (2003) Fluorine-NMR experiments for high-throughput screening: theoretical aspects, practical considerations, and range of applicability. *J Am Chem Soc* **125**, 7696-7703
270. Banegas-Luna, A.-J., Ceron-Carrasco, J. P., and Perez-Sanchez, P. (2018) A review of ligand-based virtual screening web tools and screening algorithms in large molecular databases in the age of big data. *Future Med Chem* **10**, 2641-2658
271. Vaque, M., Ardevol, A., Blade, C., Salvado, M. J., Blay, M., Fernandez-Larrea, J., Arola, L., and Pujadas, G. (2008) Protein-ligand docking: a review of recent advances and future perspectives *Curr Pharm Anal* **4**, 1-19
272. Klebe, G. (2006) Virtual ligand screening: strategies, perspectives and limitations. *Drug Discov Today* **11**, 580-594
273. Lynch, J. W., Zhang, Y., Talwar, S., and Estrada-Mondragon, A. (2017) Glycine receptor drug discovery. *Adv Pharmacol* **79**, 225-253
274. Fletcher, D. (2020) Parameters used to select compounds for the DDU ^{19}F small molecule library.
275. Totrov, M. (2008) Atomic property fields: generalized 3D pharmacophoric potential for automated ligand superposition, pharmacophore elucidation and 3D QSAR. *Chem Biol Drug Des* **71**, 15-27
276. Zou, L., Harkey, M. R., and Henderson, G. L. (2002) Effects of intrinsic fluorescence and quenching on fluorescence-based screening of natural products. *Phytomedicine* **9**, 263-267

277. Visentin, J., Couzi, L., Dromer, C., Neau-Cransac, M., Guidicelli, G., Veniard, V., Nubret-le Coniat, K., Merville, P., Di Primo, C., and Taupin, J.-L. (2018) Overcoming non-specific binding to measure the active concentration and kinetics of serum anti-HLA antibodies by surface plasmon resonance. *Biosens Bioelectron* **117**, 191-200
278. Myszka, D. G., and Rich, R. L. (2000) Implementing surface plasmon resonance biosensors in drug discovery. *Pharmaceut Sci Tech* **3**, 310-317
279. Dolezal, O., Doughty, L., Hattarki, M. K., Fazio, V. J., Caradoc-Davies, T. T., Newman, J., and Peat, T. S. (2013) Fragment Screening for the Modelling Community: SPR, ITC, and Crystallography. *Aust J Chem* **66**, 1507-1517
280. Totrov, M., and Abagyan, R. (2008) Flexible ligand docking to multiple receptor conformations: a practical alternative. *Curr Opin Struct Biol* **18**, 178-184
281. Bottegoni, G., Kufareva, I., Totrov, M., and Abagyan, R. (2009) Four-dimensional docking: a fast and accurate account of discrete receptor flexibility in ligand docking. *J Med Chem* **52**, 397-406
282. Schultes, S., de Graaf, C., Haaksma, E. E. J., de Esch, I. J. P., Leurs, R., and Kramer, O. (2010) Ligand efficiency as a guide in fragment hit selection and optimization. *Drug Discov Today Technol* **7**, e152-e162
283. Smart, B. E. (2001) Fluorine substituent effects (on bioactivity). *J Fluor Chem* **109**, 3-11
284. Wonnacott, S. (2007) Nicotinic ACh receptors. *Tocris Rev*
285. Shahsavar, A., Gajhede, M., Kastrup, J. S., and Balle, T. (2016) Structural studies of nicotinic acetylcholine receptors: using acetylcholine-binding protein as a structural surrogate. *Basic Clin Pharmacol Toxicol* **118**, 399-407
286. Arias, H. R. (2012) Molecular interactions between ligands and nicotinic acetylcholine receptors revealed by studies with acetylcholine binding proteins. *J Thermodynam Cat*. 10.4172/2157-7544.1000116
287. Rucktooa, P., Haseler, C. A., van Elk, R., Smit, A. B., Gallagher, T., and Sixma, T. K. (2012) Structural characterization of binding mode of smoking cessation drugs to nicotinic acetylcholine receptors through study of ligand complexes with acetylcholine-binding protein. *J Bio Chem* **287**, 23283-23293
288. Laube, B., Maksay, G., Schemm, R., and Betz, H. (2002) Modulation of glycine receptor function: a novel approach for therapeutic intervention at inhibitory synapses? *Trends Pharmacol Sci* **23**, 519-527
289. Schmieden, V., and Betz, H. (1995) Pharmacology of the inhibitory glycine receptor: agonist and antagonist actions of amino acids and piperidine carboxylic acid compounds. *Mol Pharmacol* **48**, 919-927
290. Schmieden, V., Jezequel, S., and Betz, H. (1996) Novel antagonists of the inhibitory glycine receptor derived from quinolinic acid compounds. *Mol Pharmacol* **50**, 1200-1206
291. Huang, X., Chen, H., and Shaffer, P. L. (2017) Crystal structures of human GlyRα3 bound to ivermectin. *Structure* **25**, 945-950
292. Fu, Z., Li, X., and Merz Jr., K. M. (2011) Accurate assessment of the strain energy in a protein-bound drug using QM/MM x-ray refinement and converged quantum chemistry. *J Comput Chem* **32**, 2587-2597
293. Brams, M., Gay, E. A., Saez, J. C., Guskov, A., van Elk, R., van der Schors, R. C., Peigneur, S., Tytgat, J., Strelkov, S. V., Smit, A. B., Yakel, J. L., and Ulens, C. (2011) Crystal structures of a cysteine-modified mutant in loop D of acetylcholine-binding protein. *J Biol Chem* **286**, 4420-4428
294. Jones, A. K., Goven, D., Froger, J.-A., Bantz, A., and Raymond, V. (2021) The cys-loop ligand-gated ion channel gene superfamilies of the cockroaches Blattella germanica and Periplaneta americana. *Pest Manag Sci*. 10.1002/ps.6245
295. Yusein, S., Velikova, N., Kupenova, P., Hardie, R., Wolstenholme, A., and Semenov, E. (2008) Altered ivermectin pharmacology and defective visual system in Drosophila mutants for histamine receptor HCLB. *Invert Neurosci* **8**, 211-222

296. Zheng, Y., Hirschberg, B., Yuan, J., Wang, A. P., Hunt, D. C., Ludmerer, S. W., Schmatz, D. M., and Cully, D. F. (2002) Identification of two novel Drosophila melanogaster histamine-gated chloride channel subunits expressed in the eye. *J Bio Chem* **277**, 2000-2005
297. Gisselmann, G., Plonka, J., Pusch, H., and Hatt, H. (2004) Unusual functional properties of homo- and heteromultimeric histamine-gated chloride channels of Drosophila melanogaster: spontaneous currents and dual gating by GABA and histamine. *Neurosci Lett* **372**, 151-156
298. Guruprasad, K., Reddy, B. V. B., and Pandit, M. W. (1990) Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng Des Sel* **4**, 155-161
299. Ikeda, M., Hamajima, R., and Kobayashi, M. (2015) Baculoviruses: diversity, evolution and manipulation of insects. *Entomol Sci* **18**, 1-20
300. Rana, M. S., Wang, X., and Banerjee, A. (2018) An improved strategy for fluorescent tagging of membrane proteins for overexpression and purification in mammalian cells. *Biochemistry* **57**, 6741-6751
301. Brown, D. A., and Rose, J. K. (1992) Sorting of GPI-anchored proteins to glycolipid-enriched membrane subdomains during transport to the apical cell surface. *Cell* **68**, 533-544
302. Edidin, M. (2003) The state of lipid rafts: from model membranes to cells. *Annu Rev Biophys Biomol Struct* **32**, 257-283
303. Li, X., Serwanski, D. R., Miralles, C. P., Bahr, B. A., and De Blas, A. L. (2007) Two pools of Triton X-100-insoluble GABAA receptors are present in the brain, one associated to lipid rafts and another one to the postsynaptic GABAergic complex. *J Neurochem* **102**, 1329-1345
304. Thiel, U., Platt, S. J., Wolf, S., Hatt, H., and Gisselmann, G. (2015) Identification of amino acids involved in histamine potentiation of GABAA receptors. *Front Pharmacol*. 10.3389/fphar.2015.00106
305. Hoerbelt, P., Ramerstorfer, J., Ernst, M., Sieghart, W., Thomson, J. L., Hough, L. B., and Fleck, M. W. (2016) Mutagenesis and computational docking studies support the existence of a histamine binding site at the extracellular β3+β3− interface of homooligomeric β3 GABAA receptors. *Neuropharmacology* **108**, 252-263
306. Gilbert, D. F., Islam, R., Lynagh, T., Lynch, J. W., and Webb, T. I. (2009) High throughput techniques for discovering new glycine receptor modulators and their binding sites. *Front Mol Neurosci*. 10.3389/neuro.02.017.2009
307. Stead, C., Brown, A., Adams, C., Nickolls, S. J., Young, G., Kammonen, J., Pryde, D., and Cawkill, D. (2016) Identification of positive allosteric modulators of glycine receptors from a high-throughput ccreen using a fluorescent membrane potential assay. *J Biomol Screen* **21**, 1042-1053
308. Tokuriki, N., Stricher, F., Schymkowitz, J., Serrano, L., and Tawfik, D. S. (2007) The stability effects of protein mutations apear to be universally distributed. *J Mol Biol* **369**, 1318-1332
309. Ruhmann, E., Betz, M., Fricke, M., Heine, A., Schafer, A., and Klebe, G. (2015) Thermodynamic signatures of fragment binding: Validation of direct versus displacement ITC titrations. *BBA-Gen Subjects* **1850**, 647-656
310. Leyrer-Jackson, J. M., Olive, M. F., and Gipson, C. D. (2019) Whole-cell patch clamp electrophysiology to study ionotropic glutamatergic receptors and their roles in addiction. *Methods Mol Biol* **1941**, 107-135

A. Crystallographic statistics for structural complexes

Structure	GBP:N-methylbicuculline	*Ac*AChBP:N-methylbicuculline
Unit cell dimensions a, b, c (Å) α, β, γ (°)	 71.24, 132.24, 132.63 90.00, 90.00, 90.00	 79.46, 100.03, 100.22 89.55, 88.08, 66.92
Space group	$P2_12_12_1$	P1
Source	In-house	Synchrotron - Soleil
Wavelength (Å)	1.5418	0.9801
Subunits per asymmetric unit	5	10
Resolution range	46.82 – 2.40 (2.48-2.40)	48.42-2.90 (2.98-2.90)
Total number of reflections	306823 (27829)	187440 (13631)
Unique reflections	49817 (4510)	59954 (4566)
Redundancy	6.2 (6.2)	3.1 (3.0)
R_{merge}	0.130 (0.776)	0.127 (0.777)
Wilson B-factor ($Å^2$)	48.1	54.9
Completeness (%)	100 (100)	95.3 (97.5)
<I/σ(I)>	7.9 (1.8)	6.3 (1.4)
CC(1/2)	0.995 (0.826)	0.995 (0.644)
Refinement		
R_{work}/R_{free} (%)	20.29 / 23.61	27.72 / 30.86
Number of reflections for R_{work}/R_{free}	44310 / 2381	56915 / 2943
Protein residues	1025	2077
N-methylbicuculline molecules	5	3
Water molecules	404	-
Chloride ions	5	10
RMSDs		
Bonds (Å)	0.009	0.006
Angles (°)	1.42	1.44
Ramachandran plot		
Residues in favoured regions (%)	991 (98.4)	1865 (92.1)
Residues in allowed regions (%)	16 (1.6)	123 (6.1)
Residues in outlier regions (%)	-	38 (1.9)
Mean B-factors ($Å^2$)		
Protein atoms	43.5 / 40.3 / 45.5 / 52.3 / 46.0	66.9 / 69.3 / 60.0 / 56.6 / 66.6 / 64.7 / 65.3 / 69.3 /61.9 / 56.9
Water molecules	38.2	-
N-methylbicuculline	95.7 / 67.2 / 101.9 / 111.9 / 116.3	95.1 / 61.5 / 73.9
Chloride ions	37.5 / 37.1 / 49.1 / 56.4 / 58.4	59.5 / 55.9 / 46.9 / 39.6 / 55.3 / 58.2 / 61.3 / 58.4 / 41.1 / 40.9
PDB Code	5OBH	-

B. Theoretical interaction parameters predicted by Molsoft

Protein	Compound	Series	ICM score	Hbond (kcal mol^{-1})	Hphob (kcal mol^{-1})	VDW (kcal mol^{-1})
GBP	Strychnine	-	-14.6	0.0	-6.4	-33.3
GBP	1a	4	-16.2	-3.2	-4.1	-18.1
GBP	1b	2	-16.9	-2.7	-5.7	-21.8
GBP	1c	3	-10.5	-2.1	-3.9	-15.7
GBP	1d	1	-12.0	-0.4	-4.5	-21.5
*Ac*AChBP:nicotine	Nicotine	-	-28.4	-3.9	-4.8	-23.5
*Ac*AChBP:nicotine	1d	2	-6.2	-0.8	-4.9	-17.8
*Ac*AChBP:nicotine	2a	-	-12.9	-2.3	-4.0	-20.1
*Ac*AChBP:nicotine	2b	2	-13.0	-1.7	-4.1	-21.3
*Ac*AChBP:nicotine	2c	6	-12.4	-2.2	-5.5	-18.9
*Ac*AChBP:apo	Nicotine	-	-23.6	-3.9	-4.4	-16.4
*Ac*AChBP:apo	1d	2	-16.1	-3.4	-4.1	-16.5
*Ac*AChBP:apo	2a	-	-8.4	-1.7	-3.6	-14.0
*Ac*AChBP:apo	2b	2	-6.8	-1.5	-3.8	-16.1
*Ac*AChBP:apo	2c	6	-22.9	-4.8	-5.2	-21.2

Table B.1 The main docking outputs for simulations using the NMR hits with GBP (PDB code: 5OBG), *Ac*AChBP:nicotine (PDB code: 5O87) and *Ac*AChBP:*apo* (PDB code: 2Y7Y). For descriptions of the different parameters see Chapter 2.5.2. Strychnine and nicotine were re-docked to the relevant receptoras controls. The series relates to the common substructure in Figure 5.2 of Chapter 5.5.

Protein	Compound	Series	ICM score	Hbond (kcal mol^{-1})	Hphob (kcal mol^{-1})	VDW (kcal mol^{-1})
GBP	Strychnine	-	-14.6	0.0	-6.4	-33.3
GBP	3a	1	-25.9	-5.4	-4.8	-21.4
GBP	3b	1	-29.4	-5.3	-4.8	-23.2
GBP	3c	1	-26.6	-4.7	-4.6	-17.5
GBP	3d	2	-25.2	-5.0	-3.9	-24.5
GBP	3e	2	-25.7	-4.5	-4.4	-23.3
GBP	3f	3	-25.1	-5.8	-3.1	-17.3
GBP	3g	3	-25.4	-7.6	-2.7	-8.4
GBP	3h	4	-25.4	-5.4	-4.2	-18.2
GBP	3i	5	-25.1	-5.5	-2.7	-19.6
GBP	3j	6	-25.7	-4.7	-7.0	-25.4
*Ac*AChBP:nicotine	Nicotine	-	-28.4	-3.9	-4.8	-23.5
*Ac*AChBP:nicotine	4a	1	-26.5	-4.3	-4.3	-25.0
*Ac*AChBP:nicotine	4b	1	-29.6	-9.3	-3.1	-18.7
*Ac*AChBP:nicotine	4c	1	-25.6	-4.8	-3.7	-23.5
*Ac*AChBP:nicotine	4d	2	-25.4	-6.0	-3.5	-21.5
*Ac*AChBP:nicotine	4e	3	-29.5	-5.1	-4.0	-24.9
*Ac*AChBP:nicotine	4f	3	-26.7	-4.9	-4.0	-24.1
*Ac*AChBP:nicotine	4g	3	-26.2	-5.1	-3.9	-21.4
*Ac*AChBP:nicotine	4h	3	-25.4	-5.1	-3.9	-21.6
*Ac*AChBP:nicotine	4i	6	-27.5	-5.3	-3.7	-26.7
*Ac*AChBP:nicotine	4j	-	-26.9	-5.7	-3.5	-19.9
*Ac*AChBP:apo	Nicotine	-	-23.6	-3.9	-4.4	-16.4
*Ac*AChBP:apo	5a	1	-25.9	-3.5	-4.9	-21.7
*Ac*AChBP:apo	5b	1	-25.2	-3.5	-4.7	-20.5
*Ac*AChBP:apo	5c	1	-26.2	-3.5	-4.9	-21.8
*Ac*AChBP:apo	5d	1	-22.2	-4.2	-5.1	-24.5
*Ac*AChBP:apo	5e	3	-23.5	-3.2	-5.1	-22.1
*Ac*AChBP:apo	5f	3	-23.5	-3.3	-5.3	-23.4
*Ac*AChBP:apo	5g	2	-23.4	-3.0	-5.8	-21.8
*Ac*AChBP:apo	5h	6	-22.4	-7.5	-2.8	-21.1
*Ac*AChBP:apo	5i	6	-24.1	-7.4	-4.0	-23.0
*Ac*AChBP:apo	5j	6	-24.2	-7.5	-3.1	-22.4

Table B.2 The main docking outputs for simulations using the Maybridge hits with GBP (PDB code: 5OBG), *Ac*AChBP:nicotine (PDB code: 5O87) and *Ac*AChBP:*apo* (PDB code: 2Y7Y). For descriptions of the different parameters see Chapter 2.5.2. Strychnine and nicotine were re-docked to the relevant receptor as controls. The series relates to the common substructure in Figures 5.7, 5.9 and 5.11 of Chapter 5.5.

C. Preliminary work on the amino acid transporters, GlyT1C and PutP

C.1 Introduction

C.1.1 Human glycine transporters

At synaptic junctions, specific amino acid transporters recycle the neurotransmitters released into the cleft by transporting them into the presynaptic neuron and/or glial cells. Of relevance here are glycine transporters (GlyT1 and 2), GABA transporters and monoamine transporters, which transport serotonin (1). The reuptake of neurotransmitters is driven by a Na^+ electrochemical gradient, a counter Cl^- ion is also transported (2).

GlyTs are predicted to consist of twelve transmembrane helices, five intracellular and six extracellular loops (3). There are two transporters, GlyT1 can operate in both directions and is present in multiple cell and neuron types, whereas GlyT2 can only take up glycine into the cell for glycinergic neurons (4). Their essential role at the synapse means that mutations have been linked to disease including motor-sensory deficits and neuromotor disorder (3).

C.1.2 Bacterial proline transporter

Orthologues of the neurotransmitter transporters exist in prokaryotes and are used to transport amino acids. Amino acid transport in bacteria plays an essential role in metabolism (5), however, proline transport is involved in a wider range of biological processes including signalling and protection from oxidative stress (6,7). In some bacterium, proline is the only source of carbon and nitrogen for metabolism (8). ESKAPE pathogens are six bacterial classes (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumanni*, *Pseudomonas aeruginosa* and *Enterobacter*) that are on the rise in hospitals as well as being multidrug-resistant, therefore are of keen research interest (9). It has been shown that some ESKAPE pathogens possess the proline transporter, PutP, and mutants of the protein affect virulence (10,11), therefore the protein may be a candidate for target validation for early-stage drug discovery.

C.2 Materials and methods

C.2.1 Construct design

Uniprot sequences of *Hs*GlyT1C (P48067), *Hs*GlyT2 (Q9Y345) and *Pseudomonas aeruginosa* (*Pa*) PutP (Q9I5F5) were extracted. At the time of the work, no structures for any of the proteins existed, therefore structural models were made using Phyre2.

C.2.2 Cloning of GlyT1C for expression in insect cells

The gene for GlyT1C was ordered from Integrated DNA Technologies (IDT). This included an N-terminal Kozak sequence and a C-terminal TEV cleavage site with His_6 tag, in a pUCIDT-Amp vector (IDT). Flanking restriction sites (NdeI and XhoI) were also included. The gene was codon optimised for baculovirus expression. A PCR was conducted to amplify the gene. PCR products were analysed by gel electrophoresis and the desired product was isolated. The insert was then digested, to form 'sticky ends' and ligated into a pFastBac1 vector, which was then transformed into NovaBlue competent cells (Novagen) and plated onto a LB-agar ampicillin plate.

C.2.3 Cloning of PutP for expression in *E. coli*

The gene encoding PutP was ordered from IDT in a pUCIDT-Amp vector (IDT). Due to the potential requirement for using multiple vectors with different tags, restriction sites were not included. Codon sequences were optimised for expression in *E. coli*. A PCR was conducted to introduce the start and end restriction sites required to insert the gene into a modified pET21a vector (Novagen). The PCR product was ligated into TOPO (Invitrogen) and the integrity of the construct was confirmed by sequencing. PCR products were analysed by gel electrophoresis, then the relevant DNA was extracted and purified (Qiagen kit). The PCR amplified PutP gene and pET21a plasmid were digested with the appropriate restriction enzymes (Table C.1). The cut plasmid and insert were purified by gel electrophoresis and then extracted (Qiagen kit). The pET21a plasmid and PutP insert were ligated and transformed into DH5α (NEB) using the Clonables kit (Novagen, Merck) and plated onto a LB-agar ampicillin plate. The plate was incubated for 24 h, after which colonies were picked and grown in LB (20 mL, Media Kitchen, University of Dundee) at 37 °C with shaking at 200 rpm for 24 h. Each colony was also analysed by colony PCR. The cultures were centrifuged at 1,912 g for 15 m at 4 °C (Sigma 4K15 centrifuge, 11150 rotor), then the plasmid was extracted by miniprep (Qiagen kit). The purified plasmid was sent for sequencing to confirm that it was correct.

Protein	Expression System	Vector	Restriction Sites	Tags [(C) or (N) terminal]
GlyT1C	Baculovirus	pFastBac1	BamHI/EcoRI	TEV- His_8 (C)
PutP	*E. coli*	pET21a	NdeI/XhoI	His_6 (C)

Table C.1 Vectors used for protein production attempts of GlyT1C and PutP.

C.2.4 Attempt at recombinant production of PutP in *E. coli*

A fresh transformation of plasmid DNA (1 µL, 10 ng) into competent cells (15 µL, C43 [DE3]) was conducted and then the culture was plated out onto LB-agar ampicillin plates, which were incubated at 37 °C for 24 h. C43 (DE3) cells are widely used for the production of recombinant membrane proteins (12). A single colony was picked to start a small LB culture (25 mL), which was incubated at

37 °C for 24 h, shaking at 200 rpm. The starter culture was transferred to a flask containing auto-induction (AI) media (500 mL). The culture was grown for around 3 h at 37 °C with shaking at 200 rpm, until an optical density (Abs_{600}) of 0.6 was reached. The temperature was then lowered to 20 °C and the culture was maintained for 24 h with shaking at 200 rpm. AI media contains a limited amount of glucose (0.1% maximum), which is preferentially metabolised by the bacteria. When the glucose is depleted, they switch to consuming glycerol (1.2% maximum) then lactose (0.6% maximum). By-products of lactose release the lac repressor leading to the expression of genes encoding the protein of interest (13).

Cells were pelleted at 1,912 g for 15 m at 4 °C (Sigma 4K15 centrifuge, 11150 rotor) and the pellet was kept on ice. Cells were resuspended in buffer A1 (50 mM Bis-Tris-HCl pH 6, 150 mM NaCl, 0.1 mg mL^{-1} L-proline) with the addition of protease inhibitors (10 µL, 100x, Expedeon, Abcam) and DNase (100 µL, 10 mg mL^{-1} in 0.15 M NaCl, Sigma). The cells in the suspension were broken open by the use of a chilled cell disruptor (30 kpsi, Constant Systems). Intact cells were removed by centrifuging at 5,000 g for 15 m at 4 °C (Beckman Avanti J25 with a JA-25.50 rotor). The supernatant was then centrifuged again at 45,000 g for 1 h at 4 °C (Beckman Optima L-100K ultracentrifuge with a 70 Ti rotor) to pellet the membranes. The supernatant was removed, this contains the soluble fraction. The pellet was resuspended and manually homogenised in buffer A1 with the addition of lipids (1 mg mL^{-1}, SigmaAldrich) and n-dodecyl-β-D-maltopyranoside (DDM, 1% [w/v], Anatrace). The solution was centrifuged at 45,000 g for 1 h at 4 °C (Beckman Optima L-100K ultracentrifuge with a 70 Ti rotor). Samples were taken at each stage to analyse by SDS-PAGE and Western blot.

C.3 Attempted production of GlyT

Homology models were prepared for both GlyT1 and GlyT2 followed by sequence and structural alignments. Models used the *D. melanogaster* dopamine (47% sequence ID) (14) and human serotonin (43% sequence ID) (15) transporters as templates. There are different isoforms (A-C) for GlyT1, with C being the canonical version. The two proteins, GlyT1C and GlyT2, share a sequence identity of 44% and have good structural conservation (RMSD 0.6 Å). The model predicts there are three large loops (out of the eleven in total), these may have a role in function and localisation, two intracellular and one extracellular. In the models these loops are disordered. In addition, no structural or biophysical data are available for either transporter. Due to the lack of structure for the protein and its interesting role in biology, GlyT1C was taken forward for cloning. A tagged full length gene encoding GlyT1C was cloned successfully and baculovirus production was underway but due to time constraints, no further work was conducted on this protein. Recently, a crystal structure of GlyT at 3.4 Å was reported (20). This study could be used as a basis for further work.

C.4 Cloning and attempted production of PutP

A structural model of PutP was generated using the *Proteus mirabilis* Na^+-coupled sialic acid (20% sequence ID) (16) and *Vibrio parahaemolyticus* Na^+-coupled galactose (19% sequence ID) (17) transporters as templates. Using the model, structural predictions and evidence of full length *E. coli* PutP purification (18,19), it was decided the full length PutP protein should be trialled before making alterations to the sequence. The *putP* gene was cloned into a pET21a vector with a C-terminal His_6 tag and expressed utilising the *E. coli* C43 (DE3) strain for membrane protein production. A modified protocol for the production of the *E. coli* homologue was used (19) to produce PutP and analysis at different stages was conducted by SDS-PAGE and Western blot with an anti-His antibody (Figure C.1). Preliminary results showed the presence of several bands, one between 37-50 kDa and another around 75 kDa. The molecular weight of full length PutP is 55.5 kDa, which is similar in weight to that of the lower band, however, the results are ambiguous and more work is required. As this was only a preliminary test, no controls were conducted, therefore a negative control should be run in tandem to check for background signals in the Western blot. Mass spectrometry protein identification analysis could also be performed on the different bands to suggest if the protein has been produced and to identify the relevant band.

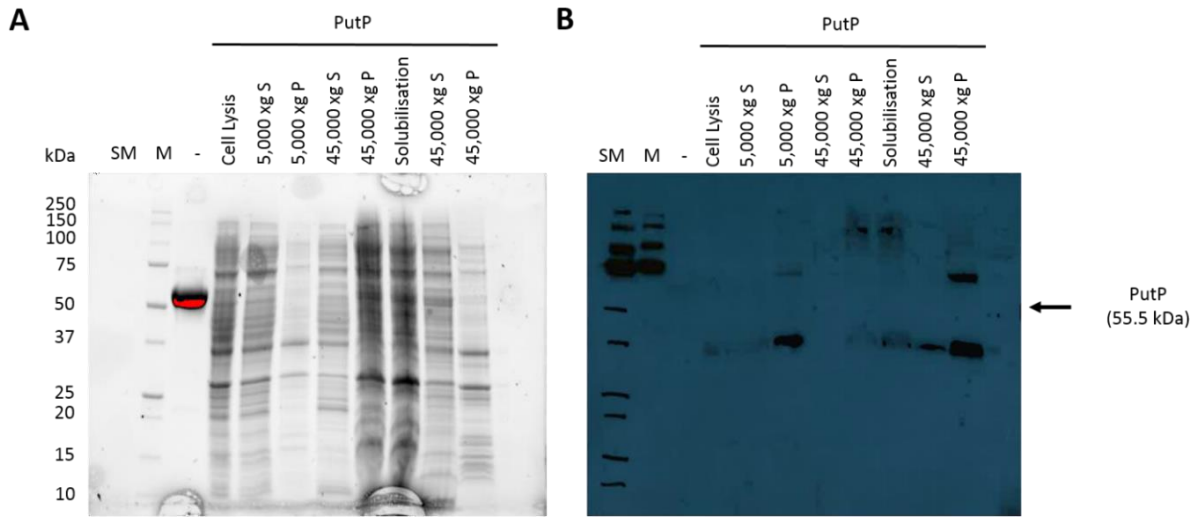


Figure C.1 Example of (A) SDS-PAGE protein gel and (B) Western blot for PutP. Stained (SM) and unstained (M) markers were used for size comparison. The intervals were marked on the Western blot from comparison with the SDS-PAGE gel for clarity. S and P are supernatant and pellet, respectively. A negative control (-) protein was run in tandem. An anti-His primary antibody was used for the detection of PutP.

C.5 Summary

GlyT1C was cloned and a baculovirus expression system was set up, however, due to time constraints protein production was not attempted. Preliminary data for PutP is inconclusive. However, more work is required to identify if it is PutP and to optimise the conditions.

C.6 References

1. Iversen, L. (2006) Neurotransmitter transporters and their impact on the development of psychopharmacology. *Br J Pharmacol* **147**, S82-S88
2. Focke, P., Wang, X., and Larsson, H. P. (2014) Neurotransmitter transporters: structure meets function. *Structure* **21**, 694-705
3. Eulenburg, V., Armsen, W., Betz, H., and Gomeza, J. (2005) Glycine transporters: essential regulators of neurotransmission. *Trends Biochem Sci* **30**, 325-333
4. Zeilhofer, H. U., Acuna, M. A., Gingras, J., and Yevenes, G. E. (2018) Glycine receptors and glycine transporters: targets for novel analgesics? *Cell Mol Life Sci* **75**, 447-465
5. Burkovski, A., and Kramer, R. (2002) Bacterial amino acid transport proteins: occurrence, functions, and significance for biotechnological applications. *Appl Microbiol Biotechnol* **58**, 265-274
6. Christgen, S. L., and Becker, D. F. (2019) Role of proline in pathogen and host interactions. *Antioxid Redox Signal* **30**, 683-709
7. Zheng, R., Feng, X., Wei, X., Pan, X., Liu, C., Song, R., Jin, Y., Bai, F., Jin, S., Wu, W., and Cheng, Z. (2018) PutA Is Required for Virulence and Regulated by PruR in Pseudomonas aeruginosa. *Front Microbiol*. 10.3389/fmicb.2018.00548
8. Nakada, Y., Nishijyo, T., and Itoh, Y. (2002) Divergent structure and regulatory mechanism of proline catabolic systems: characterization of the putAP proline catabolic operon of Pseudomonas aeruginosa PAO1 and Its regulation by PruR, an AraC/XylS family protein. *J Bacteriol* **184**, 5633-5640
9. Rice, L. B. (2008) Federal funding for the study of antimicrobial resistance in nosocomial pathogens: no ESKAPE. *J Infect Dis* **197**, 1079-1081
10. Bayer, A. S., Coulter, S. N., Stover, C. K., and Schwan, W. R. (1999) Impact of the high-affinity proline permease gene (putP) on the virulence of Staphylococcus aureus in experimental endocarditis. *Infect Immun* **67**, 740-744
11. Tielen, P., Rosin, N., Meyer, A.-K., Dohnt, K., Haddad, I., Jansch, L., Klein, J., Narten, M., Pommerenke, C., Scheer, M., Schobert, M., Schomburg, D., B., T., and Jahn, D. (2013) Regulatory and Metabolic Networks for the Adaptation of Pseudomonas aeruginosa Biofilms to Urinary Tract-Like Conditions. *PloS One*. 10.1371/journal.pone.0071845
12. Miroux, B., and Walker, J. E. (1996) Over-production of proteins in Escherichia coli: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* **260**, 289-298
13. Blommel, P. G., Becker, K. J., Duvnjak, P., and Fox, B. G. (2008) Enhanced bacterial protein expression during auto-induction obtained by Alteration of lac repressor dosage and medium composition. *Biotechnol Prog* **23**, 585-598
14. Penmatsa, A., Wang, K. H., and Gouaux, E. (2013) X-ray structure of dopamine transporter elucidates antidepressant mechanism. *Nature* **503**, 85-90
15. Coleman, J. A., Green, E. M., and Gouaux, E. (2016) X-ray structures and mechanism of the human serotonin transporter. *Nature* **532**, 334-339
16. Wahlgren, W. Y., Dunevall, E., North, R. A., Paz, A., Scalise, M., Bisignano, P., Bengtsson-Palme, J., Goyal, P., Claesson, E., Caing-Carlsson, R., Andersson, R., Beis, K., Nilsson, U. J., Farewell, A., Pochini, L., Indiveri, C., Grabe, M., Dobson, R. C. J., Abramson, J., Ramaswamy, S., and Friemann, R. (2018) Substrate-bound outward-open structure of a Na +-coupled sialic acid symporter reveals a new Na + site. *Nat Commun*. 10.1038/s41467-018-04045-7
17. Watanabe, A., Choe, S., Chaptal, V., Rosenberg, J. M., Wright, E. M., Grabe, M., and Abramson, J. (2010) The mechanism of sodium and substrate release from the binding pocket of vSGLT. *Nature* **468**, 988-991
18. Hanada, K., Yamato, I., and Anraku, Y. (1988) Purification and reconstitution of Escherichia coli proline carrier using a site specifically cleavable fusion protein. *J Bio Chem* **263**, 7181-7185

19. Chen, C.-C., and Wilson, T. H. (1986) Solubilization and functional reconstitution of the proline transport system of Escherichia coli. *J Biol Chem* **261**, 2599-2604
20. Shahsavar, A., Stohler, P., Bourenkov, G., Zimmermann, I., Siegrist, M., Guba, W., Pinard, E., Seeger, M. A., Schneider, T. R., Dawson, R. J. P., and Nissen, P. (2020) Structural insights into glycine reuptake inhibition. *bioRxiv*. 10.1101/2020.12.20.110478

D. Publications

Dawson, A., Trumper, P., Souza, J. O. d., Parker, H., Jones, M. J., Hales, T. G., and Hunter, W. N. (2019) Engineering a surrogate human heteromeric α/β glycine receptor orthosteric site exploiting the structural homology and stability of acetylcholine-binding protein. *IUCRJ* **6**, 1014-1023

Jones, M. J., Dawson, A., Hales, T. G., and Hunter, W. N. (2019) A structural rationale for N-methylbicuculline acting as a promiscuous competitive antagonist of inhibitory pentameric ligand-gated ion channels. *ChemBioChem* **21**, 1526-1533

IUCrJ
ISSN 2052-2525
BIOLOGYjMEDICINE

Engineering a surrogate human heteromeric α/β glycine receptor orthosteric site exploiting the structural homology and stability of acetylcholine-binding protein

Alice Dawson,[a] Paul Trumper,[a] Juliana Oliveira de Souza,[a] Holly Parker,[a] Mathew J. Jones,[a] Tim G. Hales[b] and William N. Hunter[a]*

[a]Division of Biological Chemistry and Drug Discovery, School of Life Sciences, University of Dundee, Dundee DD1 5EH, Scotland, and [b]Division of Systems Medicine, School of Medicine, Ninewells Hospital, University of Dundee, Dundee DD1 9SY, Scotland. *Correspondence e-mail: w.n.hunter@dundee.ac.uk

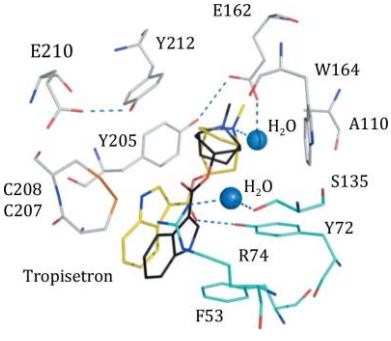
Received 29 March 2019
Accepted 9 August 2019

Edited by J. L. Smith, University of Michigan, USA

Keywords: acetylcholine-binding protein; crystal structures; glycine receptor; ligand-gated ion channel; nicotine; strychnine; tropisetron.

PDB references: acetylcholine-binding protein, wild type, strychnine complex, 5o8t; wild type, nicotine complex, 5o87; variant I, strychnine complex, 5oa0; variant II, HEPES complex, 5oad; variant II, tropisetron complex, 5oaj; variant II, strychnine complex, 5oal; variant III, glycine complex, 5oan; variant III, strychnine complex, 5obg

Supporting information: this article has supporting information at www.iucrj.org



Protein-engineering methods have been exploited to produce a surrogate system for the extracellular neurotransmitter-binding site of a heteromeric human ligand-gated ion channel, the glycine receptor. This approach circumvents two major issues: the inherent experimental difficulties in working with a membrane-bound ion channel and the complication that a heteromeric assembly is necessary to create a key, physiologically relevant binding site. Residues that form the orthosteric site in a highly stable ortholog, acetylcholine-binding protein, were selected for substitution. Recombinant proteins were prepared and characterized in stepwise fashion exploiting a range of biophysical techniques, including X-ray crystallography, married to the use of selected chemical probes. The decision making and development of the surrogate, which is termed a glycine-binding protein, are described, and comparisons are provided with wild-type and homomeric systems that establish features of molecular recognition in the binding site and the confidence that the system is suited for use in early-stage drug discovery targeting a heteromeric α//β glycine receptor.

1. Introduction

Pentameric ligand-gated ion channels (pLGICs) are important neurotransmitter receptors in the human central nervous system (CNS). The pLGIC family includes the γ-aminobutyric type A receptors ($GABA_A$Rs), nicotinic acetylcholine receptors (nAChRs), the 5-hydroxytryptamine type 3 receptor (5-HT_3R) and, of particular interest to us, glycine receptors (GlyRs). The proportionate activation of these receptors ensures a balance between neuronal excitation and inhibition (Corringer *et al.*, 2012; Lemoine *et al.*, 2012; Thompson *et al.*, 2010), and mutations that perturb the balance are associated with neurological and psychiatric disorders (Helbig *et al.*, 2008; Shiang *et al.*, 1993). The pharmacological relevance of pLGICs is well recognized, with members being targeted by anesthetics or drugs to treat anxiety as examples (Lemoine *et al.*, 2012; Olsen, 2018). The successful use of relatively few drugs against the large pLGIC family suggests future opportunities for drug discovery if an improved understanding of specific structure–activity relationships, appropriate chemical tools and techniques were available. However, there are inherent difficulties in targeting complex, multi-subunit membrane-bound ion channels for drug discovery. The presence of detergents can complicate compound screens, and multiple ligand-binding sites that vary depending on the conformational state of the ion channel are also problematic. To this we add the very

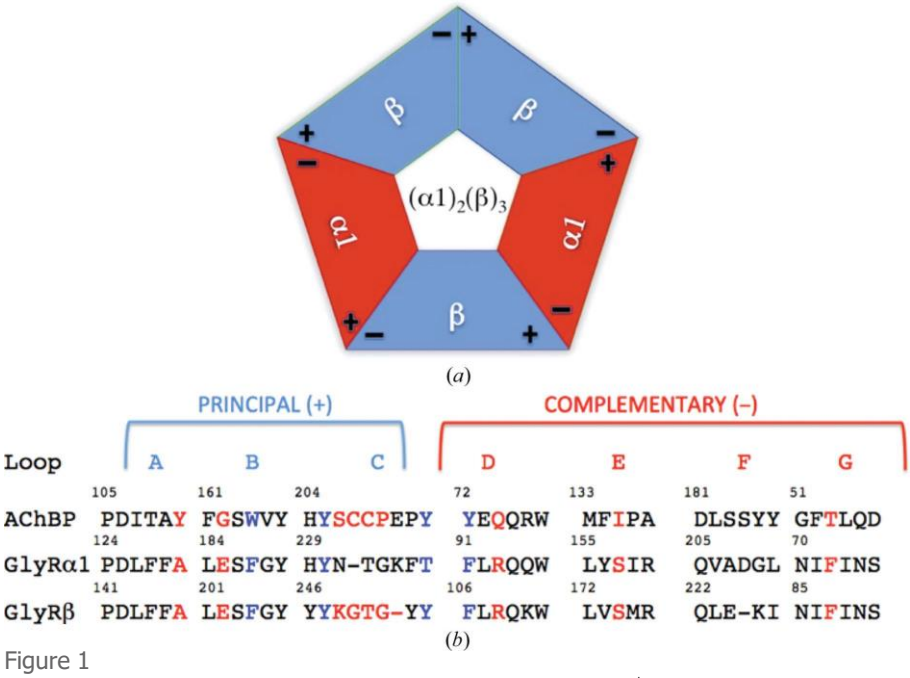


Figure 1
A schematic of a heteropentameric GlyR. The stoichiometry is $(a1)_2(/3)_3$, with the a1 subunit in red and the /3 subunit in cyan. Plus and minus symbols indicate the positions of the principal and complementary sides of the binding site, respectively. In this arrangement there are three types of binding site: two a1(+)//3(–), two a1(–)//3(+) and one /3(+)//3(–). (*b*) Comparison of the loop segments that create the orthosteric ligand-binding sites in *Ac*AChBP, human GlyR-a1 and GlyR-/3. The residues colored red indicate where amino-acid substitutions have been carried out to create GBP. The four residues colored blue contribute to the binding site but have not been changed owing to structural conservation.

significant complication that the overwhelming majority of human pLGICs of physiological and pharmacological relevance are heteromeric, with distinct subunit combinations that display unique biophysical and pharmacological profiles. These assemblies are unevenly distributed throughout the CNS and its periphery, and the heterogeneity provides an opportunity for the development of ligands with receptor-subtype specificity (Dutertre *et al.*, 2012; Shan *et al.*, 2012; Webb & Lynch, 2007). In large part, owing to difficulties in the recombinant protein production of heteromeric samples, structural studies are largely restricted to homomeric pLGICs, with a limited capacity to characterize the details of selectivity that can guide the development of selective chemical probes necessary to support fundamental studies or drug discovery. There have been modeling exercises (Bergmann *et al.*, 2013; Richter *et al.*, 2012) and very recently highly significant progress with studies of the heteromeric human a4/32 nAChR (Morales-Perez *et al.*, 2016; Walsh *et al.*, 2018) and a/3y $GABA_A$ receptors (Laverty *et al.*, 2019; Masiulis *et al.*, 2019; Phulera *et al.*, 2018; Zhu *et al.*, 2018).

Our interest is the GlyR subtype, a particularly appealing target for the development of novel muscle relaxants and the treatment of neuropathic pain (Burgos *et al.*, 2016; Imlach, 2017; Lynch, 2009; Lynch *et al.*, 2017). The prevalent forms of human GlyR are a1/3 heteropentamers with 2:3 or 3:2 stoichiometry (Durisic *et al.*, 2012; Grudzinska *et al.*, 2005; Yang *et al.*, 2012). Several lines of evidence support the existence of mammalian glycine receptors with the 2a1:3/3 stoichiometry. Firstly, mutagenesis experiments implicate residues in the /3 subunit in binding glycine to a1/3 receptors expressed in *Xenopus* oocytes (Grudzinska *et al.*, 2005). Secondly, experiments with concatenated a1-/3 tandem constructs demonstrated that functional receptors were only expressed with the additional inclusion of the /3 subunit, but not the a1 subunit, implicating 2a1:3/3. Thirdly, atomic force microscopy with epitope-tagged a1 and /3 subunits expressed in HEK293 cells indicated a 2a1:3/3 stoichiometry (Yang *et al.*, 2012). Mutations affecting key residues in the orthosteric agonist site at the a1(–)//3(+) interface affect the potency of both activation by glycine and inhibition by strychnine (Grudzinska *et al.*, 2005). We therefore set out to generate a high-fidelity surrogate of this a1(–)//3(+) orthosteric binding site using the 2a1:3/3 stoichiometry [Fig. 1(*a*)] by exploiting protein-engineering methods and the thermal stability of acetylcholine-binding protein from *Aplysia californica* (*Ac*AChBP). Acetylcholine-binding protein is a highly conserved ortholog of the pLGIC extracellular ligand-binding domain (ECD) with properties similar to nAChR (Lemoine *et al.*, 2012; Rucktooa *et al.*, 2012; Sauguet *et al.*, 2015; Shahsavar *et al.*, 2016; Sixma & Smit, 2003). Studies on *Ac*AChBP and the *Lymnaea stagnalis* protein (*Ls*AChBP) have defined the selective recognition of ligands and provided surrogates for the excitatory nAChR and 5-HT_3R ECDs (Kesters *et al.*, 2013; Price *et al.*, 2016). We outline comparative informatics that guided decision making and the characterization of the resulting proteins as we, in stepwise fashion, converted *Ac*AChBP to a glycine-binding protein (GBP) displaying an orthosteric site with the structural features of heteromeric human a1(–)//3(+) GlyR. Crystallographic and cryo-EM structures of homomeric human GlyR-a3 (Huang *et al.*, 2015; Huang, Chen *et al.*, 2017; Huang, Shaffer *et al.*, 2017) and zebrafish GlyR-a1 (Du *et al.*, 2015) provide templates that validate our approach.

2. Materials and methods

2.1. Site-directed mutagenesis and protein production

The amino-acid sequences corresponding to *Ac*AChBP (Q8WSF8) and human GlyR-a1 (P23415) and GlyR-/3 (P48167) were retrieved from UniProt (http://www.uniprot.org/). Our numbering scheme correlates with the full-length sequences in these entries. A series of models were prepared using *Phyre* (Kelley *et al.*, 2015). Sequences were aligned with *Clustal Omega* (Sievers & Higgins, 2014), and the structure of

*Ac*AChBP (for example PDB entry 2xys; Brams *et al.*, 2011) and homology models were inspected, and mutations were modeled in *Coot* (Emsley & Cowtan, 2004) to inform the design of substitutions. A stepwise approach was adopted, leading to assessments of which substitutions were important and tolerated, *i.e.* produced soluble, stable protein that was able to bind known ligands, taking into consideration ligand selectivity compared with the wild type (WT). The DNA encoding *Ac*AChBP, together with several other constructs, was purchased from GenScript. Site-directed mutagenesis was carried out and altered genes were ligated into the pFastBac system for secretion using the baculovirus/*Sf*9 system. Protein preparation followed published methods (see, for example, Hansen *et al.*, 2004) and included the use of affinity and size-exclusion chromatography.

2.2. Thermostability and ligand binding

Fluorescence-based screening by differential scanning fluorimetry (DSF; see, for example, Eadsforth *et al.*, 2012) was used to determine the melting temperature (T_m) values. An Mx3005P RT PCR system (Stratagene) was used to monitor protein unfolding by the increase in fluorescence of SYPRO Orange dye (Invitrogen). Assays were carried out in 40 µl volumes with proteins at around 10 µ*M* in 50 m*M* Tris–HCl, 250 m*M* NaCl pH 7.5 in 96-well RT PCR plates (ABgene). To investigate the influence of the chemical probe strychnine, 1 µl of strychnine dissolved in DMSO or buffer and then diluted with buffer was incubated with the protein solutions for 5 min prior to 71 cycles of 1°C temperature increments starting at 25°C. After each 1°C increase the sample was excited at 492 nm and fluorescence emission was recorded at 610 nm. The melting temperatures were plotted against a reference control sample of DMSO only. The strychnine concentration varied between 0.1 and 65 m*M*, with a requirement to limit the concentration of DMSO in the final mixture to <2.5%. Data are presented in Supplementary Table S1.

2.3. Isothermal titration calorimetry (ITC)

The interaction of strychnine with *Ac*AChBP and GBP was investigated using ITC. Measurements were carried out with a MicroCal PEAQ-ITC (Malvern Panalytical) at 25°C. The protein solutions (10 µ*M* *Ac*AChBP, 40 µ*M* GBP) were prepared by dialysis against buffer (50 m*M* Tris–HCl pH 7.5, 250 m*M* NaCl) at 4°C overnight. Strychnine solutions (concentrations of 100 and 500 µ*M*) were prepared in the same buffer. For the experiments, the initial injection of one 0.4 µl aliquot was followed by 17×2 µl injections at 3 min intervals. In each case the injection needle acted as a paddle, stirring the cell contents at 750 rev min^{-1}, and the reference was set at 10 µcal s^{-1}. Data were analyzed using the software supplied by the manufacturer to calculate K_d, ΔH, $-T\Delta S$, ΔG and N, which were derived from a one-binding-site model. Control measurements, buffer into buffer, strychnine into buffer and buffer into protein solutions, were used to determine an appropriate offset adjustment. Examples of the data, averaged parameters and standard errors derived from three titrations are presented in Supplementary Fig. S1.

2.4. Tryptophan fluorescence-quenching assay

Measurements were recorded using a PerkinElmer LS-55 spectrophotometer with the detector sensitivity set to 750 V. Stock solutions of 10 µg ml^{-1} GBP and *Ac*AChBP were prepared, along with two strychnine stock solutions of 100 µ*M* and 1 m*M* in the same buffer as used for the ITC experiments. The protein samples (2 ml) were excited at a wavelength of 280 nm, and emission was recorded between 300 and 400 nm. For GBP, aliquots of 20 µl of the 1 m*M* strychnine stock were used, followed by mixing. For AChBP, additions of 2 µl of the 100 µ*M* strychnine stock were made, followed by mixing. Experiments were carried out in triplicate and the percentage change in fluorescence was calculated. Data were analyzed using *Microsoft Excel* and *GraphPad Prism* 7. Examples of the data, parameters and standard errors derived from three titrations are presented in Supplementary Fig. S2.

2.5. Crystallographic analyses

Each protein sample (4 mg ml^{-1} in 50 m*M* Tris–HCl pH 7.5, 250 m*M* NaCl) was incubated with the appropriate ligand for 1 h before setting up crystallization trials using sitting-drop vapor diffusion with standard sparse-matrix screens. Initial conditions were identified and then optimized for each sample (Supplementary Table S2). Ultimately, this led to six distinct crystal forms. Crystals were harvested using a nylon loop, cryoprotected with reservoir solution adjusted to contain 30% ethane-1,2-diol or 30% glycerol and then flash-frozen in liquid N_2. Diffraction data were recorded in-house, using beamline I04-1 at Diamond Light Source or beamline ID23-1 at the European Synchrotron Radiation Facility. Images were indexed and integrated using *XDS* (Kabsch, 2010). The data were scaled using *AIMLESS* (Evans & Murshudov, 2013) from the *CCP*4 suite (Winn *et al.*, 2011) and the structures were solved by molecular replacement with *Phaser* (McCoy *et al.*, 2007). The initial model for molecular-replacement calculations was the wild-type structure (PDB entry 2xys). Multiple rounds of automated restrained refinement in *REFMAC*5 (Murshudov *et al.*, 2011) combined with electron-density and difference density map inspection and interpretation using *Coot* (Emsley & Cowtan, 2004) were carried out. Asn91 is glycosylated and *N*-acetyl-d-glucosamine was modeled onto several subunits. Whilst inspecting the different maps it was clear that additional ligands present in the crystallization mixture were ordered in the structures. These were assigned and refined satisfactorily as chloride, citrate, ethane-1,2-diol, isopropyl alcohol or phosphate. Water molecules and side-chain conformers were included in the models as appropriate. The asymmetric units of the different crystal forms contained either five, ten or 15 subunits, and strict noncrystallographic symmetry restraints were applied during most of the refinement and were relaxed towards the end of the process. Dictionaries of ligand restraints were assembled using *grade* (Smart *et al.*, 2014). Model geometry was assessed with

Table 1
Contributions of key residues in the orthosteric site of *Ac*AChBP and the human GlyR-α1(−)/β3(+) heteromeric site.

Residues in bold were substituted with the human equivalents to create glycine-binding protein (GBP).

*Ac*AChBP			GlyR-β3(+)	
Residue	Loop	Role	Residue	Comment
Tyr110	A	Aromatic lining of the site, with hydroxyl contribution	Ala146	Reduction in size, makes space for Glu202
Gly162	B	Adjacent to Tyr110	Glu202	Increase in size and introduces negative charge
Ser163	B	Hydroxyl forms a hydrogen bond to the Tyr166 amide to hold Trp164 and Val165 in place	Ser203	Strictly conserved
Trp164	B	Aromatic contribution to site, inter-subunit hydrogen bond to Ile135 carbonyl	Phe204	Conserved aromatic with slight reduction in bulk, no hydrogen bond
Val165	B	Hydrophobic contribution	Gly205	Reduction in size
Tyr205, Tyr212	C	Tyrosine pair contributes aromatic lining and hydroxyls to site	Tyr247, Tyr253	Strictly conserved
Ser206, Cys207, Cys208, Pro209	C	Disulfide contributes hydrophobic lining to site and restrains the loop conformation	Lys248, Gly249, Thr250, Gly251	Changes likely to give more conformational freedom to loop C

*Ac*AChBP			GlyR-α1(−)	
Residue	Loop	Role	Residue	Comment
Tyr72	D	Aromatic contribution	Phe91	Conserved
Gln74	D	Abuts Ile135, Met133	Arg93	Introduction of bulk and positive charge
Ile123, Ala124, Val125	E	Hydrophobic contributions from Val125 and Ala124 C[a]; Ile123 carbonyl directed into site	Leu145, Leu146, Arg147	Conserved Ile/Leu but increase in size for Leu146 and Arg147; aliphatic part of Arg147 side chain lines site
Met133	E	Hydrophobic lining to site, inter-subunit van der Waals interactions with loop C disulfide	Leu155	Conserved
Ile135	E	Hydrophobic lining	Ser157	Reduction in size allows space for Arg93, addition of a polar group in site
Thr53	G	Abuts Tyr72	Phe72	Increase in bulk and hydrophobicity to position Phe91
Gln55	G	Inter-subunit hydrogen bond to carbonyl serves to place Tyr110 and van der Waals interactions to position Tyr72 in site	Asn74	Conserved

MolProbity (Chen *et al.*, 2010) and the PDB validation tools. Figures were generated using *PyMOL* (Schro¨ dinger). Further details, including relevant statistics, are presented in Supplementary Table S3 and Fig. S3.

3. Results and discussion

3.1. Comparison of *Ac*AChBP and GlyR sequences to inform surrogate design

The alignment of the amino-acid sequences of *Ac*AChBP, human GlyR-α1 and human GlyR-β3 [Fig. 1(*b*)], and homology modeling together with published mutagenesis data (see, for example, Grudzinska *et al.*, 2005; Pless, Hanek *et al.*, 2011; Pless, Leung *et al.*, 2011; Yu *et al.*, 2014) on the effects of specific substitutions were used to guide the conversion of the orthosteric site of *Ac*AChBP to that of a human GlyR-α1(−)/β3(+) heteromer. The orthosteric site is constructed at the subunit–subunit interface by seven loop segments. Three of these loops (labeled A–C) form the (+) or principal side of the site and four (labeled D–G) form the (−) or complementary side [Fig. 1(*b*)]. Loop F was judged to be sufficiently distant from the orthosteric binding site to be ignored. Residues with side chains directed into the orthosteric site were marked for attention (Fig. 2, Table 1). In stepwise fashion, we generated baculovirus expression systems encoding genes for WT and altered versions of *Ac*AChBP, purified and characterized the recombinant proteins to understand the consequences of alterations in and around the orthosteric site. DSF allowed us

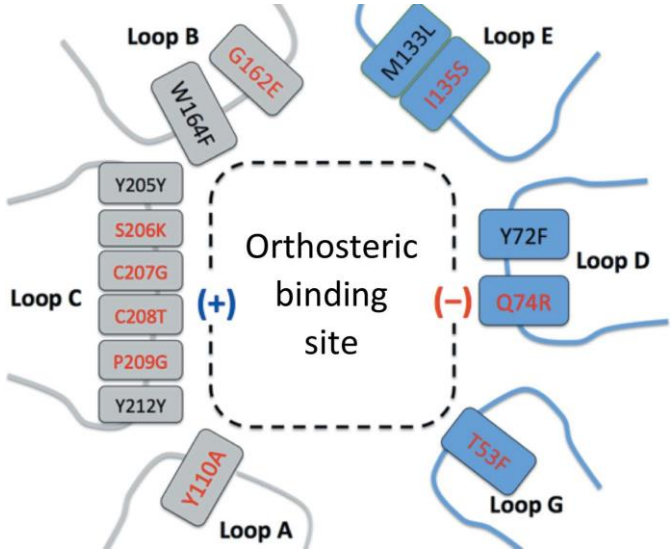


Figure 2
Schematic to describe the construction of and key residues in the orthosteric binding site of *Ac*AChBP and the corresponding amino acids in the human GlyR-α1(−)/β3(+) heteromeric site. Substitutions in red convert *Ac*AChBP into GBP.

to measure the changes in stability (T_m) as a consequence of amino-acid substitutions and ligand binding (Supplementary Table S1). Tryptophan fluorescence provided data relating to binding affinity, and crystallographic analyses of eight ligand complexes provided structural data (Supplementary Table S3, Figs. S2 and S3). Three well characterized pLGIC modulators, nicotine, tropisetron and strychnine, were used as chemical probes to provide control data.

3.2. Characterization of variants I and II

The *Ac*AChBP–nicotine complex crystal structure provided a check of the orthosteric binding site and direct comparison with the *Ls*AChBP complex (Celie *et al.*, 2004), and confirmed that the binding sites and protein–ligand interactions are highly conserved. The orthosteric site is a narrow hydrophobic cavity dominated by five aromatic residues on one side, a disulfide bond and four aliphatic residues on the other. Of 20 residues that contribute to this site (Table 1), 17 are conserved between *Ac*AChBP and *Ls*AChBP, with only three differences of note: Val125 in *Ac*AChBP changes to arginine, Thr53 to lysine and Gln55 to isoleucine. The Thr53/Gln55 combination contributes to the positioning of Tyr72, which interacts directly with ligands. In *Ls*AChBP, the aliphatic parts of the lysine and the isoleucine side chains help to position a tryptophan, which occupies the place of Tyr72 in *Ac*AChBP. In *Ac*AChBP, the side chain of Val125 contributes to a hydrophobic surface of the orthosteric site and also serves to position the side chain of Arg96, which participates in an inter-subunit salt bridge with Glu170. In *Ls*AChBP, the glutamate is conserved and the equivalent residues to Arg96 and Val125 are serine and arginine, respectively. The smaller serine side chain provides space for the arginine to occupy the same position to also form a salt bridge with the conserved glutamate, whilst the aliphatic component of the arginine essentially mimics the contributions of Val125 to the binding site.

Variant I incorporated five changes: T53F, Q74R, Y110A, I135S and W164F. The key observation from the structure of the variant I strychnine complex concerned the Y110A, W164F and I135S substitutions. Trp164 NE1 donates a hydrogen bond to the carbonyl of Ile135, thus linking two /3-strands from different subunits. Removal of the stabilizing interaction is likely to contribute to D.T_m of this variant (–40°C) compared with the wild-type protein. We also note a biphasic melting curve that may represent first dissociation of the pentamer followed by unfolding of the subunit. The electron density of the phenylalanine (Phe164) was poorly ordered, perhaps as the reduction in the side-chain size of an adjacent residue (Y110A) opened up one side of the binding site, allowing a greater degree of conformational freedom. This variant nevertheless retained the ability to bind strychnine, as revealed in the complex crystal structure and by a D.T_m of +20°C. The structure also indicted that the Y110A substitution created space to accommodate a G162E substitution (see later). Thr108 abuts Trp164 on adjacent /3-strands in *Ac*AChBP, and we reasoned that Trp164 could be retained since the residue equivalent to the adjacent Thr108 is Phe144 in human GlyR-/3 and the six-membered ring of the indole would replicate the Phe144/Phe204 combination in the human system. The retention of tryptophan also preserved the ability to exploit fluorescence measurements for binding studies.

Variant II therefore reverted back to Trp164, but with the inclusion of a G162E substitution. The residues now changed (Thr53, Gln74, Tyr110, Ile135 and Gly162) correspond to Phe72, Arg93, Ala146, Ser157 and Glu202 in the human GlyR-a1(–)//3(+) orthosteric site [Figs. 1(*b*) and 2]. The important aromatic residues Tyr72 (Phe91), Trp164 (Phe204), Tyr205 (Tyr247) and Tyr212 (Tyr253) are well conserved in the two systems [Figs. 1(*b*) and 2]. The T_m of 80°C for variant II is an increase of 25°C compared with variant I and is only 10°C lower compared with the WT protein. The incorporation of the inter-strand hydrogen bond between Trp164 and Ile135 is likely to support this recovery of thermal stability. Strychnine binding to variant II resulted in a D.T_m of +5°C.

Crystal structures of variant II with HEPES, tropisetron and strychnine revealed that the G162E substitution was accommodated with the structure essentially unperturbed compared with variant I, although now incorporating two charged residues (Arg74 and Glu162) to polarize the binding site such that the principal side is negatively charged and the complementary side is positive. The complex structure with HEPES showed this crystallization buffer component binding in two orientations in the orthosteric site in a similar fashion to that reported for *Ls*AChBP (Celie *et al.*, 2004). The variant II complex showed tropisetron [Fig. 3(*a*)] to be present in two of the ten orthosteric sites in the asymmetric unit, with the other sites being occupied by the cryoprotectant ethane-1,2-diol and the N-terminal histidine tails of symmetry-related molecules. Although the tropisetron occupies the same space, our interpretation of the electron density is that this modulator of GlyR (Yang *et al.*, 2007) displays two poses [Supplementary Figs. S3(*c*) and S3(*d*)]. One pose is similar to that observed in the WT *Ac*AChBP (Hibbs *et al.*, 2009) complex, whilst the other is rotated approximately 180° (Fig. 3). When bound to the WT protein, the tropane-bridged piperidine binds in the same position as the pyrrolidine moiety of nicotine, forming van der Waals interactions with the side chains of Tyr72 from one subunit and Tyr205, Tyr212 and Trp164 from the other subunit. The quaternary amine N1 donates a hydrogen bond to the carbonyl of Trp164 and the methyl substituent forms van der Waals interactions with Tyr110. A solvent-mediated hydrogen-bonding network links the tropisetron carbonyl to the carbonyl of Val165 and Tyr212 hydroxyl group of one subunit and the carbonyl groups of Ile106 and Met133 on the partner subunit. The ether/carbonyl link between the tropane and indole groups forms van der Waals contacts to the Cys207–Cys208 disulfide part of loop C on the (+) side and the side chain of Ile135 on the (–) side. The indole group is positioned with van der Waals contacts to Cys207 on one side and to the side chains of Tyr72, Gln74 and Met133 on another subunit. The indole N10 forms a hydrogen bond to a water molecule, which in turn interacts with Thr53 and Asp181 and other solvent molecules that form a network of hydrogen bonds in and around the binding site. The second pose is

influenced by the Y110A and G162E substitutions, which allow a solvent-bridged interaction between the glutamate and tropisetron N1. van der Waals interactions between the tropane and aromatic residues are maintained in both poses with minor adjustments of side chains. The T53F substitution and the placement of the aromatic group help to place the arginine from the Q74R substitution to participate in a cation–π stacking arrangement of the guanidinyl moiety and the indole system, pushing tropisetron over towards the disulfide linkage on loop C. A solvent-mediated link between Ser135 and Arg74 may also contribute to the placement of the guanidinyl moiety. The indole N10 is directed out towards bulk solvent, whilst the carbonyl group accepts a hydrogen bond donated from the side chain of Tyr72.

Of note is the observation that tropisetron can adopt two poses in the same binding site. It is not unusual to observe a statistical disorder in which two orientations of a ligand are present in the population of molecules in a crystal (see, for example, Khalaf *et al.*, 2014). The possibility exists that here also tropisetron adopts more than one orientation in the binding site, in effect a mixed population, but the electron-density maps suggest a dominant pose in each of the two binding sites that are occupied [Supplementary Figs. S3(*c*) and S3(*d*)]. This observation matches well with previous work on tropisetron and derivatives interacting with the 5-HT_3R that indicate that different binding orientations are possible (Ruepp *et al.*, 2017).

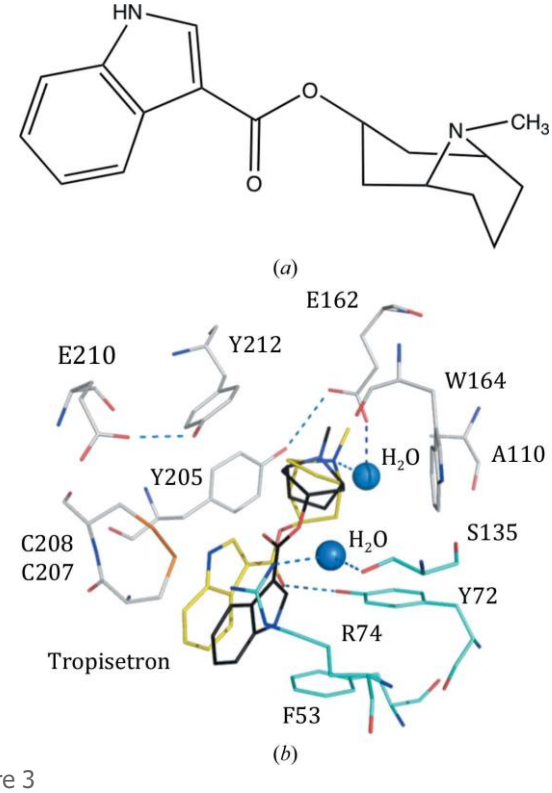


Figure 3
Tropisetron adopts two poses in the orthosteric site of variant II. (*a*) The chemical structure of tropisetron. (*b*) The interacting residues of variant II are shown with C positions colored white for the principal side and cyan for the complementary side, with one tropisetron pose (yellow C positions). Two water molecules discussed in the text are depicted as blue spheres; O and N positions are red and blue, respectively. Selected hydrogen-bonding interactions are shown as blue dashed lines. The second pose, which is common with that adopted in WT *Ac*AChBP (PDB entry 2wnc), is shown with black C atoms.

The K_d values for the binding of nicotine and tropisetron to *Ac*AChBP are 250 and 480 n*M* by monitoring intrinsic tryptophan fluorescence quenching with stopped-flow spectrofluorimetry (Hansen *et al.*, 2005). Comparable values were obtained with our tryptophan fluorescence measurements: 245 (±20) and 275 (±15) n*M*. This validated assay was applied to investigate how the substitutions might influence ligand affinity. Neither variant I nor variant II appeared to be able to bind nicotine. The combined Y110A and I135S substitutions may open up the binding site such that nicotine can longer bind in an optimal fashion. However, we were unable to co-crystallize these variants with glycine, neither did the fluorescence assay register any glycine binding.

3.3. Variant III is a glycine-binding protein

Variants I and II presented structural features consistent with site-directed mutagenesis and electrophysiological data that suggest interacting roles for specific residues (see, for example, Yu *et al.*, 2014). However, our structures also emphasized that accurately replicating the α(−)//β(+) heteromeric site required changes to the β(+) side loop C, where the major differences between α-form and β-form sequences occur (Figs. 1 and 2). Single-site substitutions were not obvious and we judged it necessary to make a major change, with four substitutions being incorporated (S206K, C207G, C208T and P209G). These substitutions had the potential to release the conformational restraint of the Cys207–Cys208 disulfide and, with two glycine residues now included, to increase the conformational mobility of the loop. Variant III was produced in recombinant form and characterized.

The substitutions did not have an adverse effect on the stability of the protein, with T_m values of 80 and 81°C noted for variants II and III, respectively (Supplementary Table S1). The binding of strychnine to variant III led to a small increase, +3°C, in ΔT_m, and when glycine was tested a ΔT_m of +2°C was observed. These changes are small and are unlikely to be significant. Attempts to observe an association between variant III and glycine using the fluorescence assay, ITC and biolayer interferometry failed to show any binding. This may be a consequence of testing a compound with such a low mass (about 75 Da). However, variant III co-crystallized with glycine and the ligand occupies four of the five orthosteric binding sites in the asymmetric unit. With this proof of binding we named variant III glycine-binding protein (GBP). The molecular packing in the crystal lattice of the GBP–glycine complex places a histidine from the affinity tag in the other site. The structure of wild-type *Ac*AChBP in complex with strychnine, the archetypal GlyR antagonist, reported here as a control (PDB entry 2xys; Brams *et al.*, 2011) provided a comparison for the GBP–strychnine complex, which we also crystallized. During our study, crystal structures of homomeric human GlyR-α3 complexes with glycine (PDB entry 5tin) and

strychnine (PDB entry 5cfb) became available (Huang, Chen *et al.*, 2017; Huang, Shaffer *et al.*, 2017), allowing direct comparisons.

The orthosteric site in the GBP–glycine structure is highly conserved with that of the WT GlyR homomer structures and variants I and II, with the notable exception of loop C, which now adopts a configuration that allows two tyrosine residues (Tyr205 and Tyr212) to contribute to the binding site (Fig. 4). One edge of the Tyr205 side chain helps to form one side of the binding site, with the hydroxyl group placed to donate a hydrogen bond to the carbonyl of Tyr166 (not shown), thereby linking two segments, and to accept a hydrogen bond from the glycine ligand. Tyr212 is aligned with Tyr205, creating a π-electron-rich region to interact with the glycine amino group. Solvent-mediated interactions link the Tyr212 hydroxyl with the carboxylate of Glu162 (not shown), which in turn interacts with the glycine amino group. The glycine is tucked between and participates in van der Waals interactions with the edge of Tyr205 and the face of the Trp164 indole. The glycine carboxylate is directed towards the Arg74 guanidinyl moiety, but the distances (2:3.8 Å) are too long to represent direct hydrogen-bonding interactions and a solvent-mediated association is noted.

In the structure of the homomeric GlyR-α3 glycine complex (Huang, Shaffer *et al.*, 2017), the glycine carboxylate accepts hydrogen bonds donated by the side chains of Arg65, Ser129 and Thr204, the latter on loop C. The amino group of the ligand makes a solvent-mediated interaction with Glu157 and a direct hydrogen bond to the carbonyl of Phe159. The glycine participates in van der Waals interactions with the side chains of Phe63 and Phe159, whilst the amino group occupies a π-electron-rich area between Phe159 and Phe207. A water molecule bridges this amino group to the carbonyl of Ser158 and the carboxylate of Glu157 with three hydrogen bonds. The major differences between the two structures and the key to forming a heteromeric site reside in loop C, the part of the binding site that distinguishes the α- and β-forms of the receptor. In GlyR-α3 the loop segment comprising residues 199–207 is in a closed conformation, whilst the equivalent residues 202–211 in GBP form a more open structure.

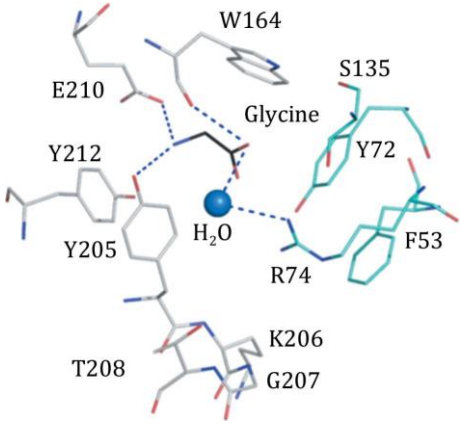


Figure 4
Glycine in an orthosteric site of GBP. A similar color scheme as shown in Fig. 3 is used, with glycine C positions in black.

However, in GBP the loop conformation places the side chains of the tyrosine pair (205 and 212) directed into the binding site, whereas in GlyR-α3 only one tyrosine, Tyr202, is thus positioned, forming a hydrogen bond with the conserved acidic Glu157.

The structure of the GBP–strychnine complex displays a well ordered ligand in three out of five binding sites per asymmetric unit. In two sites the density is less clear, the thermal parameters are elevated (Supplementary Table S3) and different strychnine orientations are noted. A previously published structure of the WT *Ac*AChBP–strychnine complex shows a single molecule occupying four of the five orthosteric sites in the asymmetric unit and one site with two ligands bound (Brams *et al.*, 2011). In both of these structures the molecular packing in the crystal lattice affects the conformation of loop C in one subunit and results in a more open binding site, providing room for ligands to adopt different orientations. We note also that this alkaloid displays a propensity to dimerize or aggregate at high concentrations (Reinscheid *et al.*, 2016). Our structure of the WT *Ac*AChBP–strychnine complex displays a single well ordered strychnine in all five binding sites per asymmetric unit and we confine our comparison to this binding pose (Fig. 5).

The orientation and the position of the alkaloid in both WT *Ac*AChBP structures is very similar. There is a single direct hydrogen bond between the alkaloid and the protein donated by the protonated tertiary amine to the carbonyl of Trp164. The pK_a of strychnine is approximately 8.3 (Haynes, 2015), hence the confidence that protonation has occurred. The potential hydrogen-bond acceptors on strychnine, ether and amide O atoms, are placed to interact with solvent and in so doing then form bridges to the protein. Extensive van der Waals interactions involving four tyrosine residues (110, 205, 212 and 72), Trp164, Met133, Ile135 and the Cys20–Cys208 disulfide are likely to explain the high affinity of strychnine for this binding site, with a reported K_i of 38 n*M* (Brams *et al.*, 2011) and K_d of 15 n*M* (Hansen *et al.*, 2004), the latter based on a stopped-flow spectrofluorimetry assay. The use of ITC and a radioligand saturation assay to characterize the interaction between strychnine and a recombinant homomeric GlyR-α1 system gave K_d values of 138 (± 55) and 52 (± 6) n*M*, respectively (Wo¨ hri *et al.*, 2013). The application of ITC and surface plasmon resonance methods with recombinant homomeric GlyR-α3 produced K_d values of 52 (±2) and 43 (±3) n*M*, respectively (Huang *et al.*, 2015). For comparative purposes we employed ITC to characterize the binding of strychnine to WT *Ac*AChBP and GBP (Supplementary Fig. S1). The resulting K_d for the interaction with WT *Ac*AChBP is 74.1 (±22.6) n*M* and that with GBP is 28.8 (±3.2) μ*M*. Corroboration of these data was sought using a fluorescence-based assay, where K_d values of 155 (±7) n*M* and 27 (±0.5) μ*M* were determined for WT *Ac*AChBP and GBP, respectively.

Strychnine is a promiscuous ligand that is active against different receptors. Electrophysiological assays indicate that the alkaloid, although the prototypical competitive antagonist of GlyR, also displays the same activity against some nAChR

subtypes, with IC_{50} values that range from 350 n*M* to around 40 m*M* (Albuquerque *et al.*, 1998; Garc´ıa-Colunga & Miledi, 1999; Jensen *et al.*, 2006). The affinity of strychnine interacting with *Ac*AChBP is consistent with the activity that this ligand displays against the orthologous binding sites presented by nAChRs. The substitutions that were introduced to engineer a glycine-binding site in GBP have however reduced the affinity for strychnine significantly away from that of the wild-type template as well as from that observed with recombinant homomeric GlyR samples. The thermodynamic parameters [Supplementary Fig. S1(*e*)] indicate that whilst the enthalpic contribution to strychnine binding is similar for *Ac*AChBP and GBP, there is a significant penalty in the entropic contribution that explains the reduced affinity of GBP for this ligand. We speculate that this may be linked to a reduction in thermal stability of GBP relative to the wild-type protein and/or be influenced by the increased flexibility introduced into loop C. The observation does have an important implication for our use of the GBP surrogate, suggesting that care should be applied when using this system with larger ligands that might engage with residues distant from those involved directly in neurotransmitter binding.

The WT *Ac*AChBP, GlyR-a3 homomer and GBP complexes display different orientations of strychnine in the binding site that are directly linked to the substitutions that have been incorporated to produce GBP (Fig. 5). An overlay of the protein structures (data not shown) places the tertiary amines within 2 Å of each other, but the ligands adopt different poses,

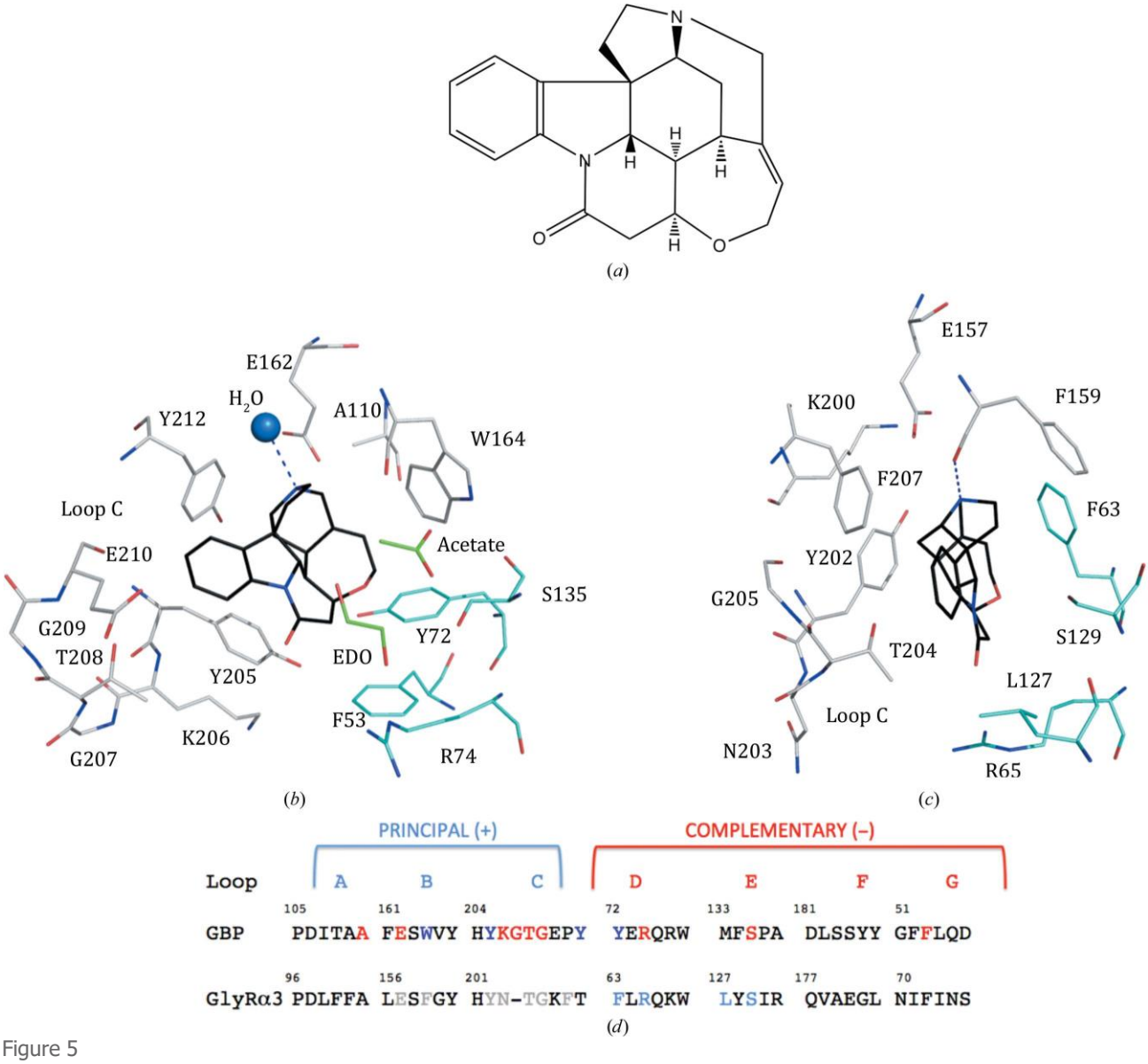


Figure 5
Strychnine bound to GBP. (*a*) The chemical structure of the natural product. (*b*) The key residues and orientation of strychnine bound to GBP. A similar color scheme as shown in Fig. 3 is used, with C positions of strychnine in black and C positions of acetate and ethanediol (EDO) in green. (*c*) The binding of strychnine to the human GlyR-a3 homomer from PDB entry 5cfb (Huang *et al.*, 2015); the residue numbers in the PDB entry are retained. (*d*) For comparative purposes the alignment of GBP [see Fig. 1(*b*)] with human GlyR-a3 is shown using the numbering scheme of the PDB entry. Residues shown in (*c*) are shown in gray for the principal side and in cyan for the complementary side.

essentially pivoting around this N atom. In the GBP–strychnine complex the tertiary amine is protonated and donates a hydrogen bond to a highly ordered water that forms hydrogen bonds to the carbonyl groups of Ser163 and Tyr166 and contributes to a solvent network that burrows into the protein fold (data not shown). The Y110A substitution removes the possibility of van der Waals interactions between the aromatic side chain and any ligand. The G162E change directs a polar side chain into the bindng site and this forces the strychnine to adopt a different orientation. On the other side of the binding site, the Q74R substitution serves to place a polar group further into the cleft. This would clash with the ligand orientation noted in the wild-type *Ac*AChBP complex and thus works in concert with the presence of Ala110 and Glu162 to position the ligand. The I135S and T53F changes allow Tyr72 to adopt a different rotamer conformation: in the former by providing space and in the latter with stabilizing van der Waals interactions. The position of Phe53 also serves to stabilize the side-chain position of Arg74. The change in orientation of strychnine results in the amide carbonyl O atom rotating by about 90° and relocating by almost 8 Å directed towards Met133 and with the indole system placed to interact with loop C. Here also there are significant changes to the protein structure. The side chain of Tyr205 adopts a different rotamer, participates in van der Waals interactions with strychnine and now occupies the space that is filled by the Cys207–Cys208 disulfide bond in WT *Ac*AChBP, in essence forming a lid over the binding site. Tyr212 also displays a different rotamer, partially filling the space vacated by Tyr205, and this allows the strychnine indole moiety to bind under loop C.

4. Conclusions

We sought to investigate the orthosteric binding site of the heteromeric α1(−)/β3(+) form of human GlyR but circumventing the experimental difficulties of working with a multi-subunit membrane-bound protein. Based on existing sequence, structural and functional data, we considered which amino acids of the homolog *Ac*AChBP might be substituted, allowing us to create a convenient surrogate system. In stages, we modified *Ac*AChBP and characterized variant proteins to interrogate the binding site. Ultimately, the highly stable *Ac*AChBP framework has allowed us to introduce nine amino-acid substitutions that have resulted in a stable surrogate for a heteromeric neurotransmitter site that binds glycine. This system could be exploited in early-stage drug discovery by use as a target for the screening of chemical libraries, for structural elucidation of receptor–ligand interactions, and for biophysical characterization of kinetic and thermodynamic parameters relating to ligand binding.

Acknowledgements

Author contributions are as follows. AD, TGH and WNH participated in research design. AD, PT, JOS, MJJ and HP conducted experiments. AD, PT, JOS, MJJ, HP, TGH and WNH performed data analysis. AD, TGH and WNH wrote or contributed to the writing of the manuscript. We thank the ESRF and Diamond Light Source synchrotrons, especially local contacts, and Dr Paul Fyfe for support.

Funding information

This work was supported by The Wellcome Trust (grant 094090).

References

Albuquerque, E. X., Pereira, E. F. R., Braga, M. F. M., Matsubayashi, H. & Alkondon, M. (1998). *Toxicol. Lett.* 102 – 103, 211–218.
Bergmann, R., Kongsbak, K., Sørensen, P. L., Sander, T. & Balle, T. (2013). *PLoS One*, 8, e52323.
Brams, M., Pandya, A., Kuzmin, D., van Elk, R., Krijnen, L., Yakel, J. L., Tsetlin, V., Smit, A. B. & Ulens, C. (2011). *PLoS Biol.* 9, e1001034.
Burgos, C. F., Yévenes, G. E. & Aguayo, L. G. (2016). *Mol. Pharmacol.* 90, 318–325.
Celie, P. H., van Rossum-Fikkert, S. E., van Dijk, W. J., Brejc, K., Smit, A. B. & Sixma, T. K. (2004). *Neuron*, 41, 907–914.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D66, 12–21.
Corringer, P.-J., Poitevin, F., Prevost, M. S., Sauguet, L., Delarue, M. & Changeux, J.-P. (2012). *Structure*, 20, 941–956.
Du, J., Lü, W., Wu, S., Cheng, Y. & Gouaux, E. (2015). *Nature (London)*, 526, 224–229.
Durisic, N., Godin, A. G., Wever, C. M., Heyes, C. D., Lakadamyali, M. & Dent, J. A. (2012). *J. Neurosci.* 32, 12915–12920.
Dutertre, S., Becker, C. M. & Betz, H. (2012). *J. Biol. Chem.* 287, 40216–40223.
Eadsforth, T. C., Gardiner, M., Maluf, F. V., McElroy, S., James, D., Frearson, J., Gray, D. & Hunter, W. N. (2012). *PLoS One*, 7, e35973.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D60, 2126–2132.
Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D69, 1204–1214.
García-Colunga, J. & Miledi, R. (1999). *Proc. Natl Acad. Sci. USA*, 96, 4113–4118.
Grudzinska, J., Schemm, R., Haeger, S., Nicke, A., Schmalzing, G., Betz, H. & Laube, B. (2005). *Neuron*, 45, 727–739.
Hansen, S. B., Sulzenbacher, G., Huxford, T., Marchot, P., Taylor, P. & Bourne, Y. (2005). *EMBO J.* 24, 3635–3646.
Hansen, S. B., Talley, T. T., Radić, Z. & Taylor, P. (2004). *J. Biol. Chem.* 279, 24197–24202.
Haynes, W. M. (2015). Editor. *CRC Handbook of Chemistry and Physics*, 95th ed., pp. 5–103. Boca Raton: CRC Press.
Helbig, I., Scheffer, I. E., Mulley, J. C. & Berkovic, S. F. (2008). *Lancet Neurol.* 7, 231–245.
Hibbs, R. E., Sulzenbacher, G., Shi, J., Talley, T. T., Conrod, S., Kem, W. R., Taylor, P., Marchot, P. & Bourne, Y. (2009). *EMBO J.* 28, 3040–3051.
Huang, X., Chen, H., Michelsen, K., Schneider, S. & Shaffer, P. L. (2015). *Nature (London)*, 526, 277–280.
Huang, X., Chen, H. & Shaffer, P. L. (2017). *Structure*, 25, 945–950.
Huang, X., Shaffer, P. L., Ayube, S., Bregman, H., Chen, H., Lehto, S. G., Luther, J. A., Matson, D. J., McDonough, S. I., Michelsen, K., Plant, M. H., Schneider, S., Simard, J. R., Teffera, Y., Yi, S., Zhang, M., DiMauro, E. F. & Gingras, J. (2017). *Nat. Struct. Mol. Biol.* 24, 108–113.
Imlach, W. L. (2017). *Pharmacol. Res.* 116, 93–99.
Jensen, A. A., Gharagozloo, P., Birdsall, N. J. & Zlotos, D. P. (2006). *Eur. J. Pharmacol.* 539, 27–33.
Kabsch, W. (2010). *Acta Cryst.* D66, 125–132.
Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. (2015). *Nat. Protoc.* 10, 845–858.

Kesters, D., Thompson, A. J., Brams, M., van Elk, R., Spurny, R., Geitmann, M., Villalgordo, J. M., Guskov, A., Danielson, U. H., Lummis, S. C., Smit, A. B. & Ulens, C. (2013). *EMBO Rep.* 14, 49–56.
Khalaf, A. I., Huggan, J. K., Suckling, C. J., Gibson, C. L., Stewart, K., Giordani, F., Barrett, M. P., Wong, P. E., Barrack, K. L. & Hunter, W. N. (2014). *J. Med. Chem.* 57, 6479–6494.
Laverty, D., Desai, R., Uchański, T., Masiulis, S., Stec, W. J., Malinauskas, T., Zivanov, J., Pardon, E., Steyaert, J., Miller, K. W. & Aricescu, A. R. (2019). *Nature (London)*, 565, 516–520.
Lemoine, D., Jiang, R., Taly, A., Chataigneau, T., Specht, A. & Grutter, T. (2012). *Chem. Rev.* 112, 6285–6318.
Lynch, J. W. (2009). *Neuropharmacology*, 56, 303–309.
Lynch, J. W., Zhang, Y., Talwar, S. & Estrada-Mondragon, A. (2017). *Adv. Pharmacol.* 79, 225–253.
Masiulis, S., Desai, R., Uchański, T., Serna Martin, I., Laverty, D., Karia, D., Malinauskas, T., Zivanov, J., Pardon, E., Kotecha, A., Steyaert, J., Miller, K. W. & Aricescu, A. R. (2019). *Nature (London)*, 565, 454–459.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* 40, 658–674.
Morales-Perez, C. L., Noviello, C. M. & Hibbs, R. E. (2016). *Nature (London)*, 538, 411–415.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D67, 355–367.
Olsen, R. W. (2018). *Neuropharmacology*, 136, 10–22.
Phulera, S., Zhu, H., Yu, J., Claxton, D. P., Yoder, N., Yoshioka, C. & Gouaux, E. (2018). *Elife*, 7, e39383.
Pless, S. A., Hanek, A. P., Price, K. L., Lynch, J. W., Lester, H. A., Dougherty, D. A. & Lummis, S. C. (2011). *Mol. Pharmacol.* 79, 742–748.
Pless, S. A., Leung, A. W. Y., Galpin, J. D. & Ahern, C. A. (2011). *J. Biol. Chem.* 286, 35129–35136.
Price, K. L., Lillestol, R. K., Ulens, C. & Lummis, S. C. (2016). *ACS Chem. Neurosci.* 7, 1641–1646.
Reinscheid, F., Schmidt, M., Abromeit, H., Liening, S., Scriba, G. K. E. & Reinscheid, U. M. (2016). *J. Mol. Struct.* 1106, 200–209.
Richter, L., de Graaf, C., Sieghart, W., Varagic, Z., Mo¨rzinger, M., de Esch, I. J., Ecker, G. F. & Ernst, M. (2012). *Nat. Chem. Biol.* 8, 455–464.
Rucktooa, P., Haseler, C. A., van Elk, R., Smit, A. B., Gallagher, T. & Sixma, T. K. (2012). *J. Biol. Chem.* 287, 23283–23293.
Ruepp, M. D., Wei, H., Leuenberger, M., Lochner, M. & Thompson, A. J. (2017). *Neuropharmacology*, 119, 48–61.
Sauguet, L., Shahsavar, A. & Delarue, M. (2015). *Biochim. Biophys. Acta*, 1850, 511–523.
Shahsavar, A., Gajhede, M., Kastrup, J. S. & Balle, T. (2016). *Basic Clin. Pharmacol. Toxicol.* 118, 399–407.
Shan, Q., Han, L. & Lynch, J. W. (2012). *J. Biol. Chem.* 287, 21244–21252.
Shiang, R., Ryan, S. G., Zhu, Y.-Z., Hahn, A. F., O'Connell, P. & Wasmuth, J. J. (1993). *Nat. Genet.* 5, 351–358.
Sievers, F. & Higgins, D. G. (2014). *Curr. Protoc. Bioinformatics*, 30, 3.13.1–3.13.16.
Sixma, T. K. & Smit, A. B. (2003). *Annu. Rev. Biophys. Biomol. Struct.* 32, 311–334.
Smart, O. S., Womack, T. O., Sharff, A., Flensburg, C., Keller, P., Paciorek, W., Vonrhein, C. & Bricogne, G. (2014). *grade* v.1.2.9. Global Phasing Ltd, Cambridge, England.
Thompson, A. J., Lester, H. A. & Lummis, S. C. (2010). *Q. Rev. Biophys.* 43, 449–499.
Walsh, R. M. Jr, Roh, S. H., Gharpure, A., Morales-Perez, C. L., Teng, J. & Hibbs, R. E. (2018). *Nature (London)*, 557, 261–265.
Webb, T. I. & Lynch, J. W. (2007). *Curr. Pharm. Des.* 13, 2350–2367.
Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta Cryst.* D67, 235–242.
Wo¨hri, A. B., Hillertz, P., Eriksson, P. O., Meuller, J., Dekker, N. & Snijder, A. (2013). *Mol. Membr. Biol.* 30, 169–183.
Yang, Z., Ney, A., Cromer, B. A., Ng, H.-L., Parker, M. W. & Lynch, J. W. (2007). *J. Neurochem.* 100, 758–769.
Yang, Z., Taran, E., Webb, T. I. & Lynch, J. W. (2012). *Biochemistry*, 51, 5229–5231.
Yu, R., Hurdiss, E., Greiner, T., Lape, R., Sivilotti, L. & Biggin, P. C. (2014). *Biochemistry*, 53, 6041–6051.
Zhu, S., Noviello, C. M., Teng, J., Walsh, R. M. Jr, Kim, J. J. & Hibbs, R. E. (2018). *Nature (London)*, 559, 67–72.

A Structural Rationale for *N*-Methylbicuculline Acting as a Promiscuous Competitive Antagonist of Inhibitory Pentameric Ligand-Gated Ion Channels

Mathew J. Jones,[a] Alice Dawson,[a] Tim G. Hales,[b] and William N. Hunter*[a]

Bicuculline, a valued chemical tool in neurosciences research, is a competitive antagonist of specific $GABA_A$ receptors and affects other pentameric ligand-gated ion channels including glycine, nicotinic acetylcholine and 5-hydroxytryptamine type 3 and the $GABA_A$ receptor–bicuculline interactions. *N*-methylbicuculline displays a similar conformation in different structures, but adopts distinct orientations enforced by interactions and steric blocks with key residues and plasticity in the binding sites. These features explain the promiscuous activity of bicuculline-binding protein (GBP), which provides a surrogate for the culline against the principal inhibitory pentameric ligand-gated ion channels in the CNS.

receptor (GlyR). The 2.4 a resolution crystal structure of the GBP:*N*-methylbicuculline complex, sequence and structural alignments reveal similarities and differences between GlyR receptor–bicuculline interactions. *N*-methylbicuculline displays a similar conformation in different structures, but adopts distinct orientations enforced by interactions and steric blocks with key residues and plasticity in the binding sites. These features explain the promiscuous activity of bicuculline against the principal inhibitory pentameric ligand-gated ion channels in the CNS.

Introduction

The alkaloid bicuculline competitively antagonises activation of the inhibitory $GABA_A$ receptor ($GABA_AR$) by the native agonist *g*-aminobutyric acid.[1–3] This natural product played a seminal role in early studies of synaptic transmission, in particular helping to characterise the role of *g*-aminobutyric acid as a neurotransmitter.[3,4] Subsequently, concerns linked to compound instability, poor solubility and activities on other receptors were raised.[5,6] Bicuculline has poor aqueous solubility and is highly susceptible to hydrolysis of the lactone moiety, with a half-life of 45 minutes at physiological pH, to produce inactive bicucine.[7] These complicating factors can be overcome by using the *N*-methylbicuculline salts (Figure 1 A), which are more stable and soluble.[2,8,9] However, an issue previously described is the haphazard reporting of which variant has been used in experiments;[2] a point we shall mention again. The alkaloid is also active against related pentameric ligand-gated ion channels (pLGICs) including the other inhibitory system, the glycine receptor (GlyR),[10] and the excitatory serotonin and nicotinic acetylcholine receptors (nAChR). This complication is compounded by activity against the unrelated calcium-activated potassium channels.[11,12] A range of IC_{50} values for activity against different receptors is presented in Table S1 in the Supporting Information.

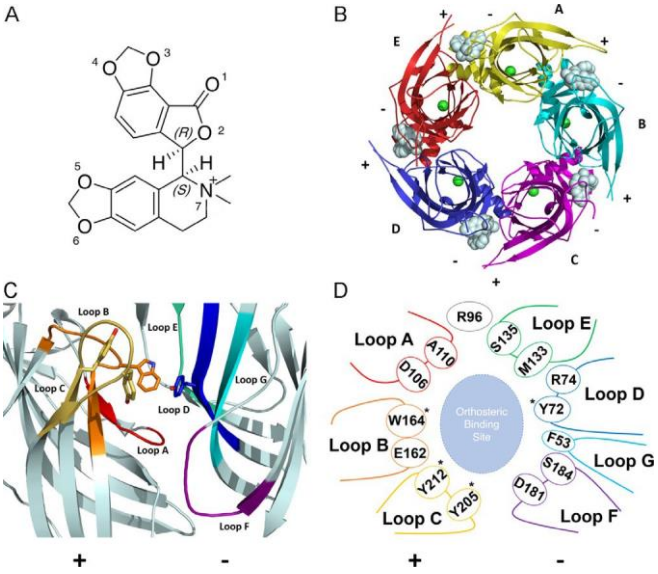


Figure 1. A) The structure of (@)-*N*-methylbicuculline. N and O atoms are numbered. B) Ribbon diagram of GBP showing the five subunits and the principal [+] and complementary [@] sides at the orthosteric binding sites. *N*-methylbicuculline (grey van der Waals spheres) occupies each site. Chloride (green spheres) bound in each subunit are shown. C) Molecular image of an orthosteric binding pocket of GBP showing, the principal subunit-donating loops A (red), B (orange), C (yellow) and the complementary subunit-contributing loops D (blue), E (blue), F (purple), G (cyan). The aromatic cage side chains (Tyr72, Trp164, Tyr205 and Tyr212) are shown as sticks with C positions coloured according to which loop they belong with. D) Schematic representation of the GBP-binding site, starred residues are responsible for the aromatic cage. Arg96 is not associated with any of the loops.

Bicuculline is GABA competitive and interacts with the orthosteric site of pLGICs.[3] The compound also acts as an inverse agonist, inhibiting $GABA_AR$-subtype activity independent

[a] *M. J. Jones, Dr. A. Dawson, Prof. W. N. Hunter*
Division of Biological Chemistry and Drug Discovery School of Life Sciences
University of Dundee
Dow St, Dundee, DD1 5EH (UK)
E-mail: w.n.hunter@dundee.ac.uk

[b] *Prof. T. G. Hales*
Division of Systems Medicine, School of Medicine
Ninewells Hospital, University of Dundee
Dundee DD1 9SY (UK)

Supporting information and the ORCID identification numbers for the authors of this article can be found under https://doi.org/10.1002/cbic.201900680.

of agonist binding to the orthosteric site.[13,14] As receptor-specific residues at key positions influence pLGIC ligand specificity and affinity,[15–17] we sought to understand how this semirigid compound was able to antagonise the transmission of inhibitory potentials in the distinctive GlyR and $GABA_AR$ systems. These receptors consist of intracellular, transmembrane and extracellular domains.[15,16] The acetylcholine-binding protein from *Aplysia californica* (*Ac*AChBP) shares sequence identity and structural homology with the extracellular domain (ECD) of pLGICs, and has been exploited as a surrogate system.[18–20] We engineered an AChBP derivative called glycine-binding protein (GBP) as a surrogate for the physiologically relevant *b*[+]/*a*[@] heteromeric interface of GlyR.[21] The orthosteric binding site is formed at the interface of two subunits with three loops (termed A–C) contributed by the principal [+]-subunit; the [@]-complementary subunit donates another four loops, D–G (Figure 1 B–D). We have characterised the interactions between *N*-methylbicuculline, also termed (@)-bicuculline methiodide, with *Ac*AChBP and GBP. Affinity and thermodynamic data are derived from a tryptophan fluorescence-quenching assay and isothermal titration calorimetry (ITC). We attempted to co-crystallise *N*-methylbicuculline with wild-type *Ac*AChBP and variants generated by site-directed mutagenesis to investigate aspects of GlyR and $GABA_AR$ activity, but to no avail. We did, however, obtain a crystal structure of the *N*-methyl derivative with GBP at 2.4 a resolution that provides a description of interactions in the neurotransmitter or orthosteric binding site and implies how the ligand modulates GlyR activity. Comparisons with $GABA_AR$ amino acid sequences and a structure derived by cryogenic-electron microscopy (cryo-EM)[22] inform a discussion about the promiscuous activity of the alkaloid against this subgroup of the pLGIC family.

Results and Discussion

Bicuculline, and the more stable *N*-methyl derivative are primarily hydrophobic, and semirigid with a single C@C bond about which rotation defines the alignment of the dioxolophthalide and dioxoloisoquinolinium moieties with respect to each other (Figure 1 A). The solvent-accessible surface area (SASA) of *N*-methylbicuculline is around 520 a^2, with a polar surface area of only about 70 a^2. There are six oxygen atoms capable of accepting hydrogen bonds and with an estimated pK_a of 14.7; under physiological conditions, bicuculline is protonated, and a hydrogen bond can be donated by the quaternary ammonium. The *N*-methyl derivative lacks this hydrogen-bond-donating capacity.

There is a requirement to use DMSO to dissolve the ligand at a concentration required to investigate binding. Given the potential complications of using this solvent,[23] we first ascertained how it might affect GBP. We characterised the thermal stability of GBP by recording the inflection point of a melting curve in different levels (0 to 10 %) of DMSO. The results indicated that 2% DMSO could be appropriate. Next, a time-course experiment indicated that, for binding assays, the 2% DMSO level was not deleterious to the sample. Control experiments were carried for ITC and fluorescence measurements allowing us to take into account the presence of DMSO in the binding assays.

The binding of *N*-methylbicuculline to GBP and *Ac*AChBP was investigated by using a fluorescence-quenching assay exploiting the presence of Trp164 in the binding site; it gave dissociation constant (K_d) values of 8.7 : 0.5 and 1.2 : 0.1 *m*m, respectively (Figure S1). An orthogonal ITC assay resulted in thermodynamic dissociation constant (K_D) values of 29.6 : 10.0 and 4.7 : 1.6 *m*m (Figure S2). We note, however, that the molar ratios for *N*-methylbicuculline binding (mol ligand/mol pentamer) derived from the ITC data are not 5:1, as would be expected for a single-site binding event, but rather 4:1 for the association with *Ac*AChBP and nearly 12:1 with GBP. In comparison, use of a [^{3}H]strychnine competition binding assay gave IC_{50} values of 5–6 *m*m for bicuculline acting on the GlyR,[24,25] whereas electrophysiology experiments recorded values in the range 169–300 *m*m for *N*-methylbicuculline.[26,27] When tested against mammalian nAChRs, again in an electrophysiology assay, bicuculline retained antagonist properties, with IC_{50} values in the range 12–34 *m*m; moreover, with Hill coefficients demonstrated to be close to unity.[28] It has been noted, from using ITC,[28] that certain ligands, for example, carbamylcholine, bind the pentameric *Ac*AChBP with a molar ratio of 2.5:1, yet there is no evidence for cooperativity or allosteric transitions in this protein. Carbamylcholine displays a similar affinity for *Ac*AChBP, 7.6 : 0.4 *m*m,[29] as *N*-methylbicuculline, and the reason for the low molar ratio is unclear. Other ligands (e.g., acetylcholine) fit the 5:1 ratio. Our ITC data were derived from curves with low *c* values of between 3 and 7. Although not optimal, such values are considered acceptable[30] with the proviso that some caution should be exercised when drawing conclusions. In our case, we note a consistency in the data derived from two distinct biophysical assays. We note also that the binding of *N*-methylbicuculline and bicuculline occurs with comparable affinity to *Ac*AChBP, GBP and to members of the pLGIC superfamily. Contrary to observations with *Ac*AChBP, the molar ratio with GBP might indicate the presence of multiple binding sites.

The crystal structure of the GBP:*N*-methylbicuculline complex has a pentamer in the asymmetric unit, with each orthosteric site occupied by a single molecule of ligand (Figures 1B and S3, Table S2) and no evidence for any other binding site. Of course, the conditions under which crystals are obtained are different from those employed for binding studies. The subunits are labelled A–E, and ligands are assigned to the subunit that forms the principal side of the binding site. We imposed noncrystallographic symmetry (NCS) in the initial refinement of the complex, but on observing deviations, released these restraints in the binding site. This in particular applies to key residues Tyr205 and Tyr212 in different orthosteric sites and will be discussed below. The electron density is well defined for most of the polypeptide chains, indeed in places sufficiently so that dual rotamers were modelled. However, at the periphery of the binding sites involving subunits B–E, the electron density is diffuse for part of loop C, and several residues could not be modelled reliably. Gly207 and Thr208 were omitted in subunits B–E, Lys206 also from B, C and E. Our observa-

tion is consistent with conformational flexibility previously noted for this part of the binding site and which is relevant to function.[15,20] The crystallographic order of the ligands differs as indicated by variability in definition of the electron density and the average *B*-factors. Our interpretation of the crystallographic data is a model with two similar poses of the ligand in a 3:2 ratio, which identify the orientation of the ligand bound to a site that possesses a degree of conformational flexibility. Pose I is observed for ligands B, D, E and pose II for A and C. The *B*-factors, or displacement parameters, indicate that the most ordered subunits are A and B, with average *B*-factors of 43.5 and 40.3 a^2, respectively. The most ordered ligand displaying pose I is at the interface formed between subunits B[+]/C[@] with an average *B*-factor of 65.2 a^2 and real-space correlation coefficient (RSCC) value of 0.90. The most ordered pose II is at the A[+]/B[@] interface with an average *B*-factor of 95.7 a^2 and RSCC = 0.88. The SASA values are similar for each pose, varying between 519 and 532 a^2, and when *N*-methylbicuculline is bound to GBP, approximately 60 % of the SASA is lost. We describe the details of the most ordered pose I because this correlates with the highest level of crystallographic order (Figure 2) and comment briefly on pose II (Figure S4).

A feature, deep in the orthosteric binding site, is the well-ordered arrangement of aromatic residues Trp164, Tyr212 and Tyr72, which creates a *p*-electron-rich environment to interact with the positively charged methylated amine. Such an interaction is an important feature of pLGIC ligand complexes.[17,18,31] In addition, there are van der Waals interactions between ligand and protein involving these residues. The side chain of Glu162, juxtaposed between Trp164 and Tyr212, is directed to the quaternary amine, thereby making an electrostatic contribution to ligand binding. On the complementary side, the phthalide part of the ligand makes van der Waals contacts to Phe53, Arg74, Met133 and Ser135. The last two residues act like a wedge between the phthalide and isoquinolium components and, together with Tyr212 from the other subunit, help to orient the ligand. The dioxoloisoquinolium occupies a depression between loops B and C of the principal subunit. Here, there are van der Waals interactions with main-chain atoms of the tripeptide segment Val165, Tyr166 and Ser167 and side-chain atoms of Tyr205, Glu210 and Tyr212. The side chain of Val125 also has van der Waals interactions with this end of the ligand; here the position of Arg96 is noted.

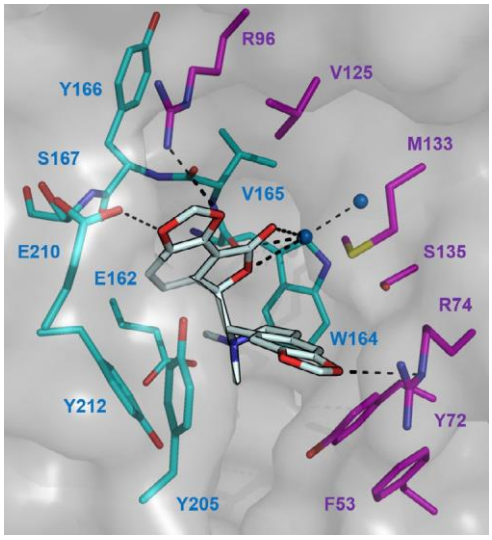


Figure 2. *N*-methylbicuculline binding to GBP. The ligand, pose I, is shown as sticks with C positions silver. The protein surface is presented as a semitransparent van der Waals surface (grey). Residues from the [+]-principal (C positions cyan) and [@]-complementary (C positions magenta) subunits are shown. N, O and S positions are blue, red and yellow, respectively. Two water molecules are depicted as blue spheres. Putative hydrogen bonds are represented as black dashed lines.

The dioxolo group associated with the phthalide can accept two hydrogen bonds donated by Arg96 on the complementary side and a protonated Glu210 on the principal side. Although the geometry is not ideal, the distance of 3.3 a suggests that the dioxolo group attached to the isoquinolium may accept a hydrogen bond donated by the side chain of Arg74. The other oxygen is solvent accessible, 4.1 a from the hydroxy group of Tyr205 on the flexible loop C. The remaining two hydrogen-bond-acceptor groups are from the lactone. These oxygen atoms are 3.1 and 3.2 a from a well-ordered water molecule (*B*-factor of 31 a^2), with an environment and geometry suggestive of a bifurcated hydrogen bond. The water in turn donates a hydrogen bond to the carbonyl of Trp164 then accepts a hydrogen bond from another highly ordered water, (*B*-factor of 31 a^2). An overlay of the *Ac*AChBP nicotine complex (PDB ID: 5O87)[18] with the structure reported here indicates that the first water maps to the position of the pyridine N atom of nicotine. The second water represents a highly conserved hydration point that forms hydrogen bonds to main-chain residues Ile123 and Ile135 in *Ac*AChBP, Ser135 in GBP, and to another water that continues a solvent network through to the surface of the protein (data not shown). This ordered solvent structure is consistently noted in high-resolution structures of AChBP ligand complexes.[15,20]

In pose II, the conformation of the ligand is preserved, with a similar orientation of the dioxolophthalide and dioxoloisoquinolinium entities with respect to each other, and the quaternary amine occupies the same relative position in the orthosteric site (Figures S4 and S5 A). Pose II participates in similar interactions to those described for pose I, for example, the key cation–*p* interaction of the quaternary amine with the protein. However, the orientation differs slightly in that, for pose II, the ligand pivots as a rigid body about the amine, and the dioxoloisoquinolinium is placed closer to Phe53 and Tyr72, further from Met133. Although an ordered water molecule occupies the same position as that discussed above, the lactone is now too far for a hydrogen bond to form. The dioxolophthalide is positioned further from the tripeptide segment Val165–Tyr166–Ser167 on the principal side. The side chains of Tyr205 and Tyr212 adopt different rotamer conformations compared to those involved in binding pose I, and accommodate this slight rigid-body adjustment whilst maintaining van der Waals interactions with the ligand. The hydrogen bonds with Arg96 and Glu210, however, are lost.

We considered first whether the GBP:*N*-methylbicuculline complex was representative of how a GlyR *b*[+]/*a*1[@] ortho-

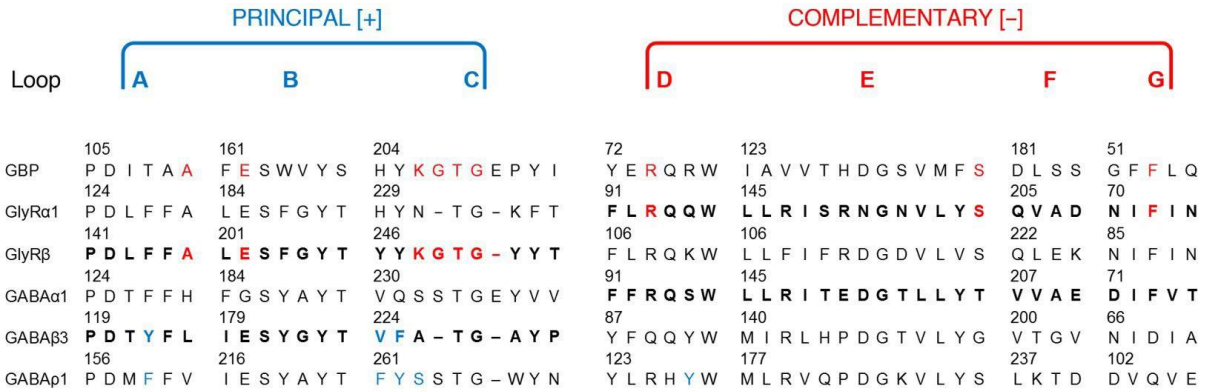


Figure 3. Sequence alignment of segments involved in creating the orthosteric binding sites of GBP, two human GlyR forms (*a*1 and *b*) and three $GABA_A$-R forms (*a*1, *b*3, *1*1). The loops are identified and split into principal and complementary sides. Residues in red were engineered into AChBP to create GBP, a surrogate for a heteromeric GlyR orthosteric site and are key to ligand interactions. Residues that contribute to heteromeric binding sites are shown in bold.

steric site might interact with the ligand. The alignment of sequences and the crystal structure of the human GlyR-*a*3 homomer in complex with strychnine (PDB ID: 5CFB)[32] were used to inform on similarity between the orthosteric sites of GBP and the GlyR *b*[+]/*a*1[@] combination (Figure 3). Arg96, a contribution from the complementary side of the orthosteric site, is not usually considered as part of the binding site and is not included in Figure 3. This residue corresponds to Asp112 in GlyR-*a*1. The contributions of 15 residues in GBP that interact with *N*-methylbicuculline have been described earlier. Seven of the residues are strictly conserved in the relevant human GlyR subunits (Phe53, Arg74, Ser135, Glu162, Tyr166, Tyr205 and Tyr212). A further four involve conservative substitutions (Met133Leu, Trp164Phe, Tyr72Phe, Ser167Thr). The differences involve the substitutions Arg96Asp, Val125Arg and Val165Gly, as well as Glu210. The Val165Gly substitution is unlikely to be significant as the residue forms van der Waals interactions with the ligand by using the main chain. Glu210 and Tyr212 are on the flexible loop C, some uncertainty about which residues on the GlyR *b*-subunit they would align with. If GBP Tyr212 aligns with GlyR-*b* Tyr253, as we think most likely (Figure 3), then the interactions with the ligand and loop C would be conserved. Tyr252 could also provide stabilising associations including a hydrogen-bond-donor group to interact with the ligand. That leaves Val125 to consider together with Arg96. These residues are close together on adjacent *b*-strands. In human GlyR-*a*1, they correspond to Arg147 and Asp112, respectively, and both are conserved in GlyR-*a*3. A superposition with the GlyR-*a*3 crystal structure indicates that the Arg96Asp difference directs the acidic group away from the binding site, but critically, the Val125Arg substitution places the guanidinium groups at the same position to interact with the ligand (Figure S6).

Alkaloids carrying a tetrahydroisoquinoline core can induce convulsions.[10] We were particularly interested in features relevant to recognition of that part of *N*-methylbicuculline. There are nine residues that interact with the dioxoloisoquinolinium entity (Phe53, Tyr72, Arg74, Met133, Ser135, Glu162, Trp164, Tyr205 and Try212). Six of these are strictly conserved, and three represent conservative substitutions in GlyR (Tyr72Phe, Met133Leu and Trp164Phe). This suggests that the *N*-methylbicuculline complex is indeed representative of how isoquinolium antagonists bind a heteromeric GlyR orthosteric site to effect competitive antagonism.

Figure 3 also presents alignments of GBP with three $GABA_AR$ subunits, this being the other inhibitory pLGIC. Again, Arg96, which corresponds to Arg112 in human $GABA_A$-*a*1, is not shown. The recent cryo-EM structure of the complex between bicuculline and the human $GABA_A$-$[a1]_2[b3]_2[g]$ receptor[22] offered the opportunity for comparison and further analysis. First a comment; there is an inconsistency in the description of what was used to derive the structure in PDB ID: 6HUK.[19] The publication and PDB entry describe bicuculline methochloride being used, but the coordinates and chemical diagram indicate bicuculline in both *b*3[+]/*a*1[@] sites, and that is used in our comparisons. We calculate RSCC values of 0.84 and 0.81 for the ligands, comparable to values in our crystal structure.

The conformation of the semirigid molecule is essentially identical between the two complex structures (Figure S5 B). An overlay of two subunits that create an orthosteric site in each protein positions the quaternary ammonium groups within 0.8 a. The environment of the cationic group is similar in the two structures. The acidic Glu162 of GBP is strictly conserved as Glu180, and the cluster of aromatic residues also maintained. These are Phe53, Tyr72, Trp164, Tyr205 and Tyr212 in GBP, which correspond to Phe73, Phe92, Tyr182, Phe225 and Tyr230, respectively, in $GABA_AR$ (Figure 4). A significant difference involves Thr108 in GBP, which is Tyr122 in the $GABA_AR$-*b*3 sequence. The increase in size of the side chain places the tyrosine hydroxy group into the cation binding pocket, thereby helping to orient the ligand.

However, although the interactions between the receptors and the cationic group are conserved, the orientation of the ligands in the orthosteric sites is different (Figure 4). The isoquinolium entities occupy a similar position in the binding site but are orthogonal to each other, and their distinct orientations are accompanied by concerted adjustments of side-chain rotamers for several of the aromatic residues deep in the binding site. This alteration of the isoquinolium, together with a tilt about the ammonium, places the dioxolophthalide components of the ligands in completely different positions. In the

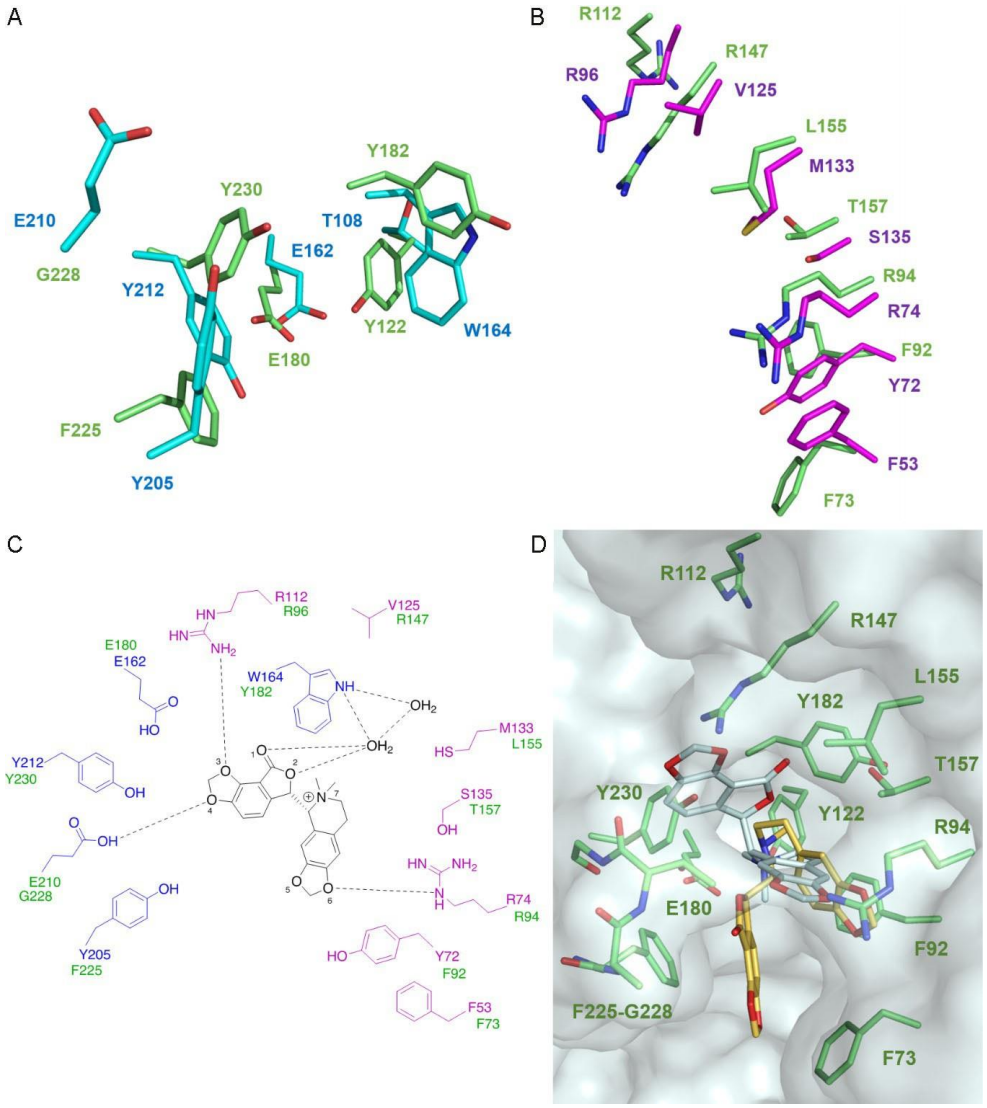


Figure 4. Comparison of the key residues of GBP and GABA$_A$R. The alignment of the structures is based on a least-squares overlay with the cryo-EM structure (PDB ID: 6HUK). A) The principal subunits with GBP residue side chains shown as sticks with C positions coloured cyan and O red, the corresponding GABA$_A$R residues are shown with C positions green. B) The complementary subunits with GBP C positions coloured magenta. C) A schematic diagram of pose I in the GBP:*N*-methylbicuculline complex. Blue and magenta distinguish residues of GBP (+) and (@) sides, respectively. The corresponding residues in GABA$_A$R are labelled in green. D) An overlay to show the distinct orientations of *N*-methylbicuculline bound to GBP (silver C) and bicuculline bound to GABA$_A$R (yellow C) in the GABA$_A$R orthosteric site. GABA$_A$R is shown as a semitransparent van der Waals surface (grey) with key residues in the binding pocket shown as in (A) and (B). The numbering of the GABA$_A$R residues is taken from the Uniprot entries (*a*1 P14867 and *b*3 P28472).

case of GBP, this part of the ligand is placed over towards the principal loop C on one side and the complementary side loop E on the other. In the GABA$_A$R structure, the dioxolophthalide component is directed towards the other end of the orthosteric site, near loop G (Figure 4 D). The nature of the amino acid at two positions on the complementary side and one on loop C of the principal side appear to contribute to the observation of two distinct orientations. In GBP, the positions of Arg96 and Val125, discussed above, align to Arg112 and Arg147 in the GABA$_A$R orthosteric site. The juxtaposition of the two arginine residues in the GABA$_A$R orthosteric site places the side chain of Arg147 to hydrogen bond with Tyr230 from loop C on the principal side to, in effect, provide a steric block preventing the dioxolophthalide from adopting the orientation observed when in complex with GBP. The crystal structure of bicuculline itself has been determined.[33] Superposition of this structure on that of the models derived from the protein complex structures indicates that, in isolation, a different conformation is observed with the two ring systems rotated about 90° relative to each other (Figure S5 C). Our comparisons therefore suggest that, although the semirigid alkaloid can adopt different conformations, only one is observed when the ligand binds to inhibitory pLGICs. However, this conformation can be present in two distinct orientations in distinct neurotransmitter binding sites, so explaining the promiscuous activity of bicuculline on these inhibitory ion channels.

The *1*-class GABA$_A$R, also known as the GABA$_C$ receptor, is insensitive to bicuculline,[34] so we sought to test our under-

standing by exploring this aspect of ion-channel pharmacology. The alignment of the human $GABA_AR$-*1*1 sequence with the sequences and structures described is presented in Figure 3. A previous attempt to promote a gain of function in $GABA_AR$-*1*1with respect to bicuculline antagonism was based on molecular modelling, site-directed mutagenesis and electrophysiology.[34] The combination of Tyr127Ser (loop D) and Phe159Tyr (loop A) substitutions, and the tripeptide Phe261–Tyr262–Ser263 (loop C) changed to a dipeptide Val–Phe gave the largest increase in bicuculline sensitivity, though still at least an order of magnitude reduced from that of the $GABA_A$-*b*3[+]/*a*1[@] system. The introduction of the serine would be predicted to open up space behind a conserved arginine (Arg125 in $GABA_AR$-*1*1, Arg74 in GBP) and allow it to act to interact with bicuculline. The Phe159Tyr change would place the hydroxy group to form stabilising interactions with the quaternary ammonium and create an environment to interact with the cationic group similar to position of Tyr122 on the *b*3 subunit of the $GABA_A$ *b*3[+]/*a*1[@] structure. The changes to a tripeptide on loop C might simply provide space to accommodate the ligand or to place the phenylalanine to mimic Phe226 of the *b*3 structure. That these changes only produce a limited gain of function indicates that other features must be important.

Our comparisons drew attention to the role of loop C and Phe53 in GBP, conserved as Phe73 in the $GABA_AR$-*a*1 subunit. In the $GABA_A$-*b*3[+]/*a*1[@] orthosteric binding site, Phe73 aligns to Gln104 in the *1*1 sequence. The presence of a polar glutamine would remove van der Waals interactions between the phenyl group and both dioxolo groups of the ligand, and would be likely to prevent bicuculline binding in the orientation noted in the cryo-EM structure. Loop C is more highly conserved with the *a*1 sequence, which is on the principal side not the complementary one. It is difficult without structural data on the *1*1 receptor to be certain about what happens with loop C, but the presence of a bulky Trp267 on the *1*1 subunit could potentially block access to the ligand. Further data would however be required to address the selectivity issues of the $GABA_A$-*1*1 receptor more completely.

Conclusions

In summary, *N*-methylbicuculline binds to GBP with low-micromolar affinity, comparable to that displayed against members of the pLGIC family (Table S1). The crystal structure of the complex with GBP has been elucidated with the ligand modelled in two similar poses that represent an orientation of the compound in the binding site. Comparisons of sequences and structures identify significant similarities between GBP and GlyR orthosteric sites, and we conclude that the orientation of *N*-methylbicuculline is representative of how the alkaloid acts as a competitive antagonist against GlyR. When bicuculline binds the human $GABA_AR$, the natural product displays the same molecular conformation as when bound to GBP, with conservation of interactions involving the quaternary ammonium group. However, the molecule adopts a different orientation in the binding site. This observation is likely a consequence of only a few specific amino acid differences between the two proteins. The binding of this promiscuous competitive antagonist appears to be driven by the charge–*p* interaction of the quaternary ammonium with the protein and van der Waals interactions that can be accommodated by flexibility inherent in the pLGIC orthosteric sites.

Experimental Section

Protein production: A recombinant source of *Ac*AChBP (Uniprot ID: Q8WSF8) and GBP with a C terminus tobacco etch virus cleavage site and His_6 tag were produced in baculovirus-infected *Sf*9 insect cells by using the Bac-to-Bac expression system (ThermoFisher). Suspension High Five insect cells, cultured in Express Five medium plus 100 U $mL^{@1}$ penicillin/streptomycin and 2 mm L-glutamine (Thermo-Fisher), were used for protein production. Typically, 15 V 10^5 cells $mL^{@1}$ were infected with 5% of baculovirus carrying the appropriate gene and incubated at 278C in shaking flasks for 48 h before being harvested by centrifugation (1500 *g*, 10 min, 12 8C followed by 4000 *g*, 10 min, 128C). The proteins are secreted out to the medium, and by using the Sartojet system with a 10 kDa cut-off Sartocon Slice filter (Sartorius), the medium was exchanged for buffer A (50 mm Tris·HCl, 250 mm NaCl, pH 7.5), and the sample was concentrated. The protein solution was applied to a 5 mL Ni^{2+} HisTrap column (GE Life Sciences) equilibrated in buffer A for immobilised metal anion chromatography. The column was washed with 15 column volumes of buffer A + 7.5 % buffer B (50 mm Tris·HCl, 250 mm NaCl, 800 mm imidazole pH 7.5), then the product was eluted with 30 column volumes by using a combination of a stepped and linear gradient of buffer B. A native-page gel (Figure S7 A) identified the presence of monomer, the desired pentamer and a higher-order multimer, possibly a dimer of pentamers. For size-exclusion chromatography, a Superdex 200 10/300 GL column was equilibrated overnight in buffer A, then samples were loaded and run over 1.5 column volumes (Figure S7 B). The retention time of each peak was recorded, and the molecular weight was deduced from a previously determined calibration curve. The use of stain-free SDS-PAGE gels (Bio-Rad, Figure S7 C) allowed us to confirm the presence of the protein, and fractions corresponding to the desired pentameric assembly were pooled, and samples were concentrated by using 10 kDa centrifugal concentrators (Pall).

Crystallographic analysis: GBP at a concentration of 4 mg $mL^{@1}$ in buffer A was incubated with 2 mm bicuculline methiodide (Sigma–Aldrich; 20 mm stock in buffer A, 20 % DMSO), diluted in buffer A for 1 h before crystallisation trials with commercially available PEGS (Qiagen) and JCSG (Molecular Dimensions) screens. Sitting-drops (final volume 0.2 *m*L) containing a 1:1 mix of protein and reservoir (a volume of 50 *m*L was used in the plates) were prepared with a Rigaku Phoenix automated dispenser and incubated at 188C for one month. A suitable crystal (rectangular prism, 0.7 V 0.25 V 0.25 mm) appeared from a condition with a reservoir of 0.2 m ammonium formate and 20 % (*w*/*v*) PEG 3350. The crystal was immersed in liquid nitrogen, then placed under a stream of nitrogen gas at around @170 8C, and diffraction data were collected with a Rigaku M007HF copper-anode generator, Varimax Cu-VHF optics, Saturn 944HG + CCD detector and AFC-11 4-axis partial *c* goniometer. The data were integrated with XDS[35] and scaled with AIMLESS;[36] the structure was solved by molecular replacement with PHASER,[37] exploiting the already refined structure of GBP with glycine (PDB ID: 5OAN).[19] Multiple rounds of automated restrained refinement were completed by using REFMAC5,[38] interspersed with model adjustment based on inspection of electron and difference density maps in COOT.[39] NCS restraints were employed at

the onset of the refinement, but when real differences became apparent, these were released. Ligand models and restraints were generated with the GRADE server [Global Phasing: http://grade.globalphasing.org/cgi-bin/grade/server.cgi]. Ligands, water molecules and chloride ions, together with several dual rotamers were incorporated into the model. Asn91 was glycosylated and *N*-acetyl-d-glucosamine was modelled onto several subunits at this position. Several residues that were not well defined by the electron density were omitted from the model. All the above software was available through the CCP4 suite.[40] All structural figures were prepared by using PyMOL (www.pymol.org) and annotated in Microsoft PowerPoint. Crystallographic statistics are given in Table S2.

Assessing the effect of DMSO on the thermal stability of GBP: Solutions of GBP (0.5 mg mL$^{@1}$) were prepared with varying concentrations of DMSO (0–10 %) and incubated at room temperature for 5 min. The inflection point (Ti) of the unfolding transition as the samples were heated was determined based on the ratio of fluorescence at 350:330 nm on a Nanotemper Tycho NT.6 instrument. Experiments were carried out in triplicate, then the mean and standard deviation of each measurement were determined and plotted (Figure S8). A further assessment was carried out in which GPB was incubated with 2% DMSO at room temperature, and the Ti was determined at 15 min intervals over 45 mins. Experiments were conducted in triplicate, then the mean and standard deviation of each were calculated. These values were then compared to the 0% DMSO control. The data indicated that over 45 mins, the time period required for the biophysical assays, there is no determinantal effect on protein stability from a 2% level of DMSO.

Fluorescence quenching assay: Stock solutions of 10 *m*g mL$^{@1}$ GBP and *Ac*AChBP were prepared, along with a 4 mm stock of bicuculline methiodide (100 mm stock in DMSO) in buffer A. Data were collected on a LS-55 PerkinElmer spectrometer with the detector sensitivity set to 750 V. Protein samples (volume 2 mL) were excited at a wavelength of 280 nm, and emissions between 300 and 400 nm were monitored. Additions of 1–2 *m*L of the *N*-methylbicuculline stock were dispensed, each followed by mixing, for a total of 20 *m*L. Experiments were carried out in triplicate, and the percentage change in fluorescence was calculated. Control measurements were conducted by adding buffer to the protein solutions whilst matching the concentration of DMSO. These were then subtracted from the percentage change in fluorescence to provide a correction for the presence of DMSO (Figure S9).

Isothermal titration calorimetry: Experiments were carried out by using a PEAQ-ITC (MicroCal, Malvern Panalytical) at 258C. GBP and *Ac*AChBP solutions (40 *m*m) were prepared by dialysis against buffer A at 4 8C overnight, then DMSO was added to match that in the titrant. Bicuculline methiodide (2 mm, 2 % DMSO) was prepared in the same buffer. The injection needle acted as a paddle to stir the cell contents at 750 rpm. An initial injection of 0.4 *m*L was followed by twelve 3 *m*L injections at 3 min intervals. Data were analysed by using the software supplied by the manufacturer assuming a one-binding-site model with mean composite controls: buffer–buffer, buffer–protein, ligand–buffer considered. Titrations were conducted in triplicate. The controls, original traces plus derived titration curves and parameters with average values are given in Figure S10 and Table S3. Representative data are presented in Figure S2E.

Sequence and structure alignments: Sequences for GlyR (Uniprot codes: a_1 P23415 and *b* P48167) and $GABA_A$ (Uniprot codes: *a*1 P14867, *b*3 P28472, *1*1 P24046) were extracted from Uniprot and aligned by using Clustal Omega on the Jalview platform.[41] GBP (PDB ID: 5OBH) and the $GABA_A$ receptor (PDB ID: 6HUK) files were retrieved from the Protein Data Bank and superimposed in COOT by SSM superimpose. Images were then produced in PyMOL, after using the align function to obtain an optimal overlay of the two structures, with ten cycles of refinement.

Acknowledgements

Funded with support from a Wellcome Trust PhD studentship and equipment grant [094090]. We thank Drs. Paul Fyfe and Paul Davis for many useful discussions.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: acetylcholine-binding proteins · alkaloids · competitive antagonists · crystal structures · $GABA_A$ receptors · glycine receptor

[1] D. R. Curtis, A. W. Duggan, D. Felix, G. A. R. Johnston, *Nature* 1970, *226*, 1222–1224.
[2] G. A. R. Johnston, P. M. Beart, D. R. Curtis, C. J. A. Game, R. M. McCulloch, R. M. Maclachlan, *Nature New Biol.* 1972, *240*, 219–220.
[3] G. A. R. Johnston, *Br. J. Pharmacol.* 2013, *169*, 328–336.
[4] D. R. Curtis, A. W. Duggan, D. Felix, G. A. R. Johnston, *Brain Res.* 1971, *32*, 69–96.
[5] D. W. Straughan, M. J. Neal, M. A. Simmonds, G. G. S. Collins, R. G. Hill, *Nature* 1971, *233*, 352–354.
[6] V. Seutin, S. W. Johnson, *Trends Pharmacol. Sci.* 1999, *20*, 268–270.
[7] R. W. Olsen, M. Ban, T. Miller, *Brain Res.* 1976, *102*, 283–299.
[8] S. F. Pong, L. T. Graham, Jr., *Brain Res.* 1972, *42*, 486–490.
[9] J. Krall, T. Balle, N. Krogsgaard-Larsen, T. E. Sørensen, P. Krogsgaard-Larsen, U. Kristiansen, B. Frølund, *Adv. Pharmacol.* 2015, *72*, 201–227
[10] P. Li, M. Slaughter, *Vis. Neurosci.* 2007, *24*, 513–521.
[11] R. Khawaled, A. Bruening-Wright, J. P. Adelman, J. Maylie, *Pflegers Arch.* 1999, *438*, 314–321.
[12] D. Strøbaek, T. D. Jørgensen, P. Christophersen, P. K. Ahring, S. P. Olesen, *Br. J. Pharmacol.* 2000, *129*, 991–999.
[13] S. Ueno, J. Bracamontes, C. Zorumski, D. S. Weiss, J. H. Steinbach, *J. Neurosci.* 1997, *17*, 625–634.
[14] M. R. McCartney, T. Z. Deeb, T. N. Henderson, T. G. Hales, *Mol. Pharmacol.* 2007, *71*, 539–548.
[15] M. Nys, D. Kesters, C. Ulens, *Biochem. Pharmacol.* 2013, *86*, 1042–1053.
[16] D. Lemoine, R. Jiang, A. Taly, T. Chataigneau, A. Specht, T. Grutter, *Chem. Rev.* 2012, *112*, 6285–6318.
[17] P. J. Corringer, F. Poitevin, M. S. Prevost, L. Sauguet, M. Delarue, J. P. Changeux, *Structure* 2012, *20*, 941–956.
[18] T. K. Sixma, A. B. Smit, *Annu. Rev. Biophys. Biomol. Struct.* 2003, *32*, 311–334.
[19] A. Shahsavar, M. Gajhede, J. S. Kastrup, T. Balle, *Basic Clin. Pharmacol. Toxicol.* 2016, *118*, 399–407.
[20] L. Sauguet, A. Shahsavar, M. Delarue, *Biochim. Biophys. Acta* 2015, *1850*, 511–523.
[21] A. Dawson, P. Trumper, J. Oliveira de Souza, H. Parker, M. J. Jones, T. G. Hales, W. N. Hunter, *IUCrJ* 2019, *6*, 1014–1023.
[22] S. Masiulis, R. Desai, T. Uchański, I. S. Martin, D. Laverty, D. Karia, T. Malinauskas, J. Zivanov, E. Pardon, A. Kotecha, J. Steyaert, K. W. Miller, A. R. Aricescu, *Nature* 2019, *565*, 454–459.
[23] A. Tjernberg, N. Markova, W. J. Griffiths, D. Hall8n, *J. Biomolec. Screen.* 2006, *11*, 131–137.
[24] A. Goldinger, W. E. Meller, *Neurosci. Lett.* 1980, *16*, 91–95.
[25] J. C. Marvizjn, J. V#zquez, M. G. Calvo, F. Mayor, A. R. Gjmez, F. Valdiveso, J. Benavides, *Mol. Pharmacol.* 1986, *30*, 590–597.
[26] N. Tokutomi, M. Kaneda, N. Akaike, *Br. J. Pharmacol.* 1989, *97*, 353–360.
[27] H. Sun, T. K. Machu, *Eur. J. Pharmacol.* 2000, *391*, 243–249.

[28] A. Demurom, E. Palma, F. Eusebi, R. Miledi, *Neuropharmacology* 2001, *41*, 854–861.
[29] P. H. Celie, S. E. van Rossum-Fikkert, W. J. van Dijk, K. Brejc, A. B. Smit, T. K. Sixma, *Neuron* 2004, *41*, 907–914.
[30] W. B. Turnbull, A. H. Daranas, *J. Am. Chem. Soc.* 2003, *125*, 14859–14866.
[31] W. Zhong, J. P. Gallivan, Y. Zhang, L. Li, H. A. Lester, D. A. Dougherty, *Proc. Natl. Acad. Sci. USA* 1998, *95*, 12088–12093.
[32] X. Huang, H. Chen, K. Michelsen, S. Schneider, P. L. Shaffer, *Nature* 2015, *526*, 277–280.
[33] C. Gorinsky, D. S. Moss, *J Cryst. Mol. Struct.* 1973, *3*, 299–307.
[34] J. Zhang, F. Xue, Y. Chang, *Mol. Pharmacol.* 2008, *74*, 941–951.
[35] W. Kabsch, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 2010, *66*, 125–132.
[36] P. R. Evans, G. N. Murshudov, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 2013, *69*, 1204–1214.
[37] A. J. McCoy, R. W. Grosse-Kuntsleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* 2007, *40*, 658–674.
[38] G. N. Murshudov, P. Skub#k, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 2011, *67*, 355–367.
[39] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 2004, *60*, 2126–2132.
[40] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 2011, *67*, 235–242.
[41] A. M. Waterhouse, J. B. Procter, D. M. Martin, M. Clamp, G. J. Barton, *Bioinformatics* 2009, *25*, 1189–1191.

Manuscript received: November 6, 2019
Accepted manuscript online: December 20, 2019
Version of record online: February 12, 2020