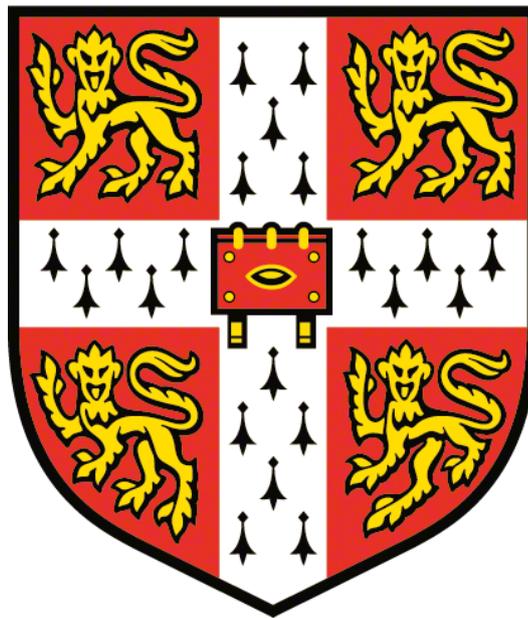


**ENGINEERING THE
ANGIOTENSIN II TYPE 1 RECEPTOR
FOR STRUCTURAL STUDIES**



JENNIFER ANN THOMAS

**THIS DISSERTATION IS SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY**

TABLE OF CONTENTS

TABLE OF CONTENTS.....	i
LIST OF FIGURES.....	iii
LIST OF TABLES.....	v
DECLARATION.....	vi
ACKNOWLEDGEMENTS.....	vii
ABSTRACT.....	viii
JOURNAL ARTICLES.....	ix
ABBREVIATIONS.....	x
CHAPTER 1 INTRODUCTION.....	1
1.1 Introduction to GPCRs.....	1
1.2 Crystallisation and structure determination of GPCRs.....	3
1.3 Requirements for GPCR crystallisation and strategies developed to overcome impediments.....	10
1.4 The angiotensin II type 1 receptor.....	14
1.5 Aim of the research presented in this dissertation.....	18
CHAPTER 2 THE SUITABILITY OF AT₁R AS A CANDIDATE FOR STRUCTURAL STUDIES.....	19
2.1 Introduction.....	19
2.2 Materials and Methods.....	24
2.3 Results.....	30
2.4 Discussion.....	46
CHAPTER 3 IMPROVING THE EXPRESSION OF AT₁R IN MAMMALIAN CELLS..	51
3.1 Introduction.....	51
3.2 Methods.....	55
3.3 Results.....	57
3.4 Discussion.....	71
CHAPTER 4 ASSESSMENT OF AT₁R LIGANDS FOR STABILISATION.....	73
4.1 Introduction.....	73
4.2 Methods.....	75
4.3 Results.....	76
4.4 Discussion.....	89
CHAPTER 5 ENGINEERING AT₁R FOR USE IN STRUCTURAL STUDIES	91
5.1 Introduction.....	91
5.2 Materials and Methods.....	93
5.3 Results.....	95
5.4 Discussion.....	114

CHAPTER 6	DETECTION OF UNFOLDED RECOMBINANT MEMBRANE PROTEINS IN EUKARYOTIC CELLS	117
6.1	Introduction.....	117
6.2	Materials and Methods.....	119
6.3	Results.....	122
6.4	Discussion.....	133
CHAPTER 7	OVERALL DISCUSSION AND CONCLUSIONS.....	135
	LIST OF REFERENCES	140
APPENDIX 1	AT₁R AMINO ACID SEQUENCE ALIGNMENTS.....	152
APPENDIX 2	INSECT CELL EXPRESSION TEST DATA	156
APPENDIX 3	FLUORESCENCE ACTIVATED CELL SORTING DATA	158
APPENDIX 4	SUBMITTED JOURNAL ARTICLE	160

LIST OF FIGURES

Figure 1.1 GPCR ligand efficiency	2
Figure 1.2 Time line and milestones of GPCR structure determination	4
Figure 1.3 Common structural characteristics of class A GPCRs	6
Figure 1.4 Signal transduction in GPCRs.....	7
Figure 1.5 Comparison of active and inactive β_2 AR structures.....	9
Figure 1.6 Activation and G protein coupling for β_2 AR	10
Figure 1.7 Comparison of three strategies used for the crystallisation of A_2A R	13
Figure 1.8 Agonist activation and desensitisation of AT_1 R.....	15
Figure 2.1 Schematic of the three methods for the thermostability assay.	20
Figure 2.2 FACS gating based on PI staining, cell size and optical homogeneity.....	27
Figure 2.3 Thermostability of AT_1 R in detergent.....	32
Figure 2.4 In-gel fluorescence analysis of N-linked glycosylation site mutations in AT_1 R	34
Figure 2.5 Confocal microscopy analysis of N-linked glycosylation site mutations in AT_1 R.....	35
Figure 2.6 Ligand binding analysis of N-linked glycosylation site mutations in AT_1 R	36
Figure 2.7 Optimisation of bv AT_1 R-H ₁₀ and bv AT_1 R-LS-H ₁₀ expressed in Sf9 cells	37
Figure 2.8 Glycosylation of AT_1 R across four insect cell lines.....	38
Figure 2.9 Expression of AT_1 R in three insect cell lines.....	38
Figure 2.10 Increased expression of AT_1 R-GFP-H ₁₀ in a stable cell line compared to transient transfection.....	40
Figure 2.11 N-linked glycosylation of AT_1 R in a stable mammalian cell line	41
Figure 2.12 Optimisation of AT_1 R expression from the stable polyclonal cell line iHEK(AT_1 R-GFP-H ₁₀).....	42
Figure 2.13 Stability of DDM-solubilised AT_1 R expressed in three different systems bound to the antagonist [¹²⁵ I]-Sar ¹ ..	43
Figure 2.14 Five-fold higher functional expression of AT_1 R in mammalian cells compared to insect cells	44
Figure 2.15 DDM solubilises considerable amounts of inactive AT_1 R produced in the baculovirus expression system	45
Figure 2.16 Size exclusion chromatography of AT_1 R expressed in mammalian cells compared to insect cells	45
Figure 3.1 FACS based methods for increasing expression in mammalian cells	53
Figure 3.2 Strategy for the selection of highly-expressing clonal cell lines	55
Figure 3.3 Optimisation of the number of cells per well deposited for FACS	57
Figure 3.4 Isolation of clonal cell lines from a population with high basal expression leads to an increase in AT_1 R expression	58
Figure 3.5 Correlation of median FACS fluorescence intensity	59
Figure 3.6 Clonal cell lines show a high level of inducibility	60
Figure 3.7 Increase in expression of AT_1 R-GFP-H ₁₀ in stable cell lines over time.....	61
Figure 3.8 In-gel fluorescence shows an increase in expression of AT_1 R-GFP-H ₁₀ through the creation of clonal cell lines.....	63
Figure 3.9 FSEC analysis of AT_1 R-GFP-H ₁₀ produced from clonal cell lines	64
Figure 3.10 AT_1 R produced in the clonal cell line iHEK415 is predominantly localised to the cell surface	64
Figure 3.11 Assessment of clonal cell lines derived from iHEK(AT_1 R-GFP-H ₁₀)	65
Figure 3.12 Clonal cell lines show high levels of expression long term	66
Figure 3.13 Significant increase in AT_1 R expression through a combination of FACS selection and the use of sodium butyrate	67
Figure 3.14 Sodium butyrate does not significantly increase the amount of unfolded AT_1 R expressed	68
Figure 3.15 Increasing levels of AT_1 R expression through FACS selection and induction with sodium butyrate.....	68
Figure 3.16 FSEC analysis of AT_1 R-GFP-H ₁₀ produced from iHEK415 grown in suspension	70
Figure 4.1 Structure of AT_1 R agonists and antagonists.....	74
Figure 4.2 Variation in expression of AT_1 R-GFP-H ₁₀ in the clonal cell line iHEK415	76
Figure 4.3 Optimisation of AT_1 R storage conditions	77
Figure 4.4 Optimisation of sample loop size for analytical FSEC	77
Figure 4.5 Analytical FSEC-based ligand screen.....	81
Figure 4.6 Optimisation of azilsartan incubation time	82

Figure 4.7 Analytical FSEC-based ligand screen with experimental ligands	82
Figure 4.8 Analysis of AT ₁ R with different concentrations of Sar ¹	83
Figure 4.9 Analysis of AT ₁ R for different periods of time with Sar ¹	84
Figure 4.10 Effects of azilsartan on AT ₁ R is mitigated with the use of LMNG	85
Figure 4.11 Calibration of Superdex 200 10/300 GL	86
Figure 5.1 Removal of positive charges on the N-terminus of AT ₁ R increases expression.....	95
Figure 5.2 In-gel fluorescence analysis of N-terminus mutants.....	96
Figure 5.3 Inclusion of a leader sequence increases AT ₁ R expression in mammalian cells	97
Figure 5.4 Assessment of cell line created to express AT ₁ R-LS-GFP-H ₁₀	97
Figure 5.5 Human AT ₁ R probability of disorder	98
Figure 5.6 Assessment of cell lines created to express C-terminal truncations of AT ₁ R.....	100
Figure 5.7 Truncation of AT ₁ R at proline 321 decreases expression.....	101
Figure 5.8 FSEC analysis of AT ₁ R-GFP expression from the iGNTI25 and iHEK26 stable cell line.....	101
Figure 5.9 Assessment of cell lines created to express AT ₁ R-T4L and C-terminal truncations of AT ₁ R.....	104
Figure 5.10 T _m assay of AT ₁ R-T4L fusion.....	105
Figure 5.11 FSEC analysis of AT ₁ R expression from the iGNTI42 stable cell line	106
Figure 5.12 The addition of antagonist and the use of CHS did not stabilise AT ₁ R expressed from the stable cell line iGNTI42	106
Figure 5.13 The use of LMNG did not stabilise AT ₁ R-CL3-T4L	107
Figure 5.14 T _m assay of AT ₁ R produced from the stable cell line iGNTI25	108
Figure 5.15 Optimisation of AT ₁ R(Del. K323-E359)-GFP expression in the stable cell line iGNTI25	109
Figure 5.16 FSEC analysis of AT ₁ R(Del. K323-E359)-GFP expression in the iGNTI25 stable cell line.....	110
Figure 5.17 Degradation of AT ₁ R-GFP produced from the polyclonal cell line iHEK(AT ₁ R-GFP-H ₁₀) when grown in suspension culture	111
Figure 5.18 FACS analysis of iGNTI25 grown in large scale	111
Figure 6.1 Misfolded AT ₁ R produced by the baculovirus expression system is poorly solubilised by either DDM or digitonin	123
Figure 6.2 Expression of A ₁ R in mammalian cells compared to insect cells.....	125
Figure 6.3 Misfolded A ₁ R produced by the baculovirus expression system is poorly solubilised by either DDM or digitonin	126
Figure 6.4 A ₁ R-GL26 is misfolded when expressed in mammalian cells.....	128
Figure 6.5 Misfolded A ₁ R-GL26 is poorly solubilised by either DDM or digitonin	129
Figure 6.6 Misfolded β ₁ AR is poorly solubilised by either DDM or digitonin	131
Figure 6.7 Misfolded SERT produced by the baculovirus expression system is poorly solubilised by either DDM or digitonin	132
Appendix Figure 1 Alignment of AT ₁ R with GPCRs of known structure.....	152
Appendix Figure 2 Alignment of vertebrate AT ₁ Rs shows a high level of conservation	153
Appendix Figure 3 Alignment of AT ₁ R with GPCR-T4L fusions.....	155
Appendix Figure 4 Alignment of human and rat AT ₁ R.....	155
Appendix Figure 5 Optimisation of bvAT ₁ R-H10 and bvAT ₁ R-LS-H ₁₀ expressed in Sf21 or Hi5 cells.....	156
Appendix Figure 6 Size exclusion chromatography of AT ₁ R expressed in insect cells	157

LIST OF TABLES

Table 2.1 X-ray diffraction structures of mammalian membrane proteins determined from protein overexpressed in mammalian cells	23
Table 2.2 Comparison of baculovirus and mammalian expression systems for the production of recombinant membrane proteins	23
Table 2.3 AT ₁ R N-linked glycosylation mutants for mammalian expression	25
Table 2.4 Apparent thermostability of AT ₁ R under different assay conditions.....	33
Table 4.1 Binding affinities of selected AT ₁ R agonists and antagonists	74
Table 4.2 Estimated molecular weight of AT ₁ R-GFP-H ₁₀	87
Table 4.3 Molecular weights of proteins expressed from three cell lines.....	88
Table 5.1 AT ₁ R alterations for mammalian expression.....	94
Table 5.2 Thermostability of AT ₁ R-T4L fusion.....	105
Table 5.3 AT ₁ R purifications attempted.....	113
Table 6.1 Percentage identity of transmembrane regions of human adenosine receptors	127
Appendix Table 1 FACS analysis of monoclonal cell lines from the polyclonal cell line iHEK(AT ₁ R-GFP-H ₁₀)	158

DECLARATION

This dissertation is the result of research carried out between June 2011 and May 2014 at the Medical Research Council Laboratory of Molecular Biology in Cambridge, UK. This dissertation is the result of my own work. Experiments were designed, performed and analysed by the author, except for the handling of [¹²⁵I] labelled ligands, which was carried out by Chris Tate and Simone Weyand. Nothing in this dissertation is the outcome of work done in collaboration, except where specifically indicated in the text.

The work described in this dissertation is not substantially the same as any that I have submitted, or is being currently submitted, for a degree, diploma or other qualification at the University of Cambridge, or any other University or similar institution. I further state that no substantial part of this dissertation has already been submitted, or is being currently submitted, for any such degree, diploma or other qualification at the University of Cambridge, or any other University or similar institution.

This dissertation does not exceed 60,000 words.

Jennifer A. Thomas
St Edmund's College
Cambridge
30 September 2014

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor Chris Tate for offering me a position in his Group and for providing help, insights, advice and guidance over the course of my research. In particular, his personal involvement in elements of my experimental work and his unstinting and generous support during the composition of my thesis were both invaluable.

I would also like to thank my second supervisor, Jan Löwe, both for the perspectives that he was able to offer me scientifically and for his ongoing support.

I am of course also very grateful to have worked with all of the members of the Tate Group, each of whom has helped me in many different ways. I wanted especially to thank Simone Weyand for both her professional assistance and friendship; it was a pleasure to collaborate with her in a number of areas. I am also very much indebted both to Tony Warne for his advice and assistance with the baculovirus expression system and to Juni Andréll for her help and guidance with the mammalian expression system.

I have also benefited from the support of family and friends in both the US and UK and would like to single out Tammy Pacyna for her interest in my work and her continuous encouragement. Finally, I would like to thank my husband, Peter Thomas, without whom none of this would have been possible. He has been with me every step of the way and often served as a much needed voice of reason. Thank you for your patience, kindness, love and support through this period and always.

Assuming the Riemann hypothesis...

ABSTRACT

G protein-coupled receptors (GPCRs) are eukaryotic integral membrane proteins that perform transmembrane signal transduction. Due to their pivotal role in a wide range of essential physiological functions GPCRs represent a high proportion of all drug targets. High resolution X-ray structures of GPCRs are however underrepresented in the Protein Data Bank. This is due to their instability in detergent, low expression levels and the presence of misfolded receptors in many heterologous expression systems.

The objective of this project was to engineer the angiotensin II type 1 receptor (AT₁R), a human GPCR, to make it suitable for structural studies. It was determined that detergent-solubilised AT₁R was thermostable with antagonist bound with an apparent T_m of ~45°C, which was sufficiently stable for purification without further thermostabilisation by rational mutagenesis. Two expression systems were then evaluated for large-scale production of AT₁R, namely baculovirus-mediated expression in insect cells and mammalian expression in HEK293 cells. Radioligand binding assays showed that only the mammalian system produced sufficient quantities of active AT₁R for structural studies. Expression in the mammalian system was further optimised to approximately 6 mg/L. An AT₁R-GFP fusion was created to examine membrane localisation using confocal laser scanning microscopy, to assay expression levels, to select highly expressing monoclonal cell lines using fluorescence activated flow cytometry and to develop a fluorescence size-exclusion chromatography-based assay to examine the suitability of 12 different ligands for co-crystallization. AT₁R was also engineered to facilitate crystallisation, including C-terminal truncations to remove predicted disordered regions and bacteriophage T4-lysozyme being added to the third intracellular loop to provide additional points of contact for crystallisation, which increased the apparent T_m by approximately 10°C. All modified versions of AT₁R were assessed for expression, stability and monodispersity. Additionally a rapid western blotting based assay was developed for the detection of unfolded membrane proteins, which will have wide applicability in the field.

JOURNAL ARTICLES

Part of the work presented in this dissertation was submitted to the *Journal of Molecular Biology* on 15 August 2014 and accepted for publication on 28 October 2014. The manuscript is included in Appendix 4:

Quality control in eukaryotic membrane protein overproduction.

Jennifer A. Thomas & Christopher G. Tate

A second manuscript is in preparation:

Generation of stable tetracycline-inducible HEK293 cell lines overexpressing mammalian G protein coupled receptors and transporters.

Juni M. Andréll, Jennifer A. Thomas & Christopher G. Tate

ABBREVIATIONS

A ₁ R	adenosine A ₁ receptor	HRP	horseradish peroxidase
A _{2A} R	adenosine A _{2A} receptor	iGnTI ⁻	tetracycline-inducible N-acetylglucosaminyl transferase I deficient HEK293S cells
ACE	angiotensin converting enzyme	iHEK	tetracycline-inducible human embryonic kidney cells
AT ₁ R	angiotensin II type 1 receptor	IRES	internal ribosome entry site
β ₁ AR	β ₁ adrenergic receptor	LCP	lipidic cubic phase
β ₂ AR	β ₂ adrenergic receptor	LMNG	lauryl maltose neopentyl glycol
BRIL	apocytochrome b ₅₆₂ RIL	LS	leader sequence
BSA	bovine serum albumin	Nb	nanobody
Cas9	CRISPR-associated protein-9 nuclease	NG	n-nonyl β-D-glucopyranoside
CHS	cholesteryl hemisuccinate	NOP	nociception opioid peptide receptor
CL	cellular loop	NTS1	neurotensin receptor
CMV	human cytomegalovirus	OG	n-octyl β-D-glucopyranoside
ConA	concanavalin A	OTG	n-octyl β-D-thioglucopyranoside
CRISPR	clustered regularly interspaced short palindromic repeats	PBS	phosphate buffered saline
D ₃ R	dopamine D ₃ receptor	PDB	protein data bank
DDM	n-dodecyl β-D-maltopyranoside	PI	propidium iodide
DEER	double electron-electron resonance	PNGase F	peptide-N-glycosidase F
DHA	dihydroalprenolol	RONN	regional order neural network
DM	n-decyl β-D-maltopyranoside	RTI-55	2β-carbomethoxy-3β-(4-iodophenyl)tropane
DPCPX	dipropylcyclopentylxanthine	Sar ¹	Sar ¹ -Ile ⁸ -angiotensin II
dpm	disintegrations per minute	Sar ¹ Val ⁵	Sar ¹ Val ⁵ L-Br ₅ Phe ⁸ angiotensin II
ECL	extracellular loop	SDS	sodium dodecyl sulphate
<i>E. coli</i>	<i>Escherichia coli</i>	SEC	size exclusion chromatography
Endo H	endoglycosidase H	SEM	standard error of the mean
ER	endoplasmic reticulum	SERT	serotonin transporter
F _{ab}	fragment antigen-binding	T4L	T4 lysozyme
FACS	fluorescence activated cell sorting	TEV	tobacco etch virus
FC12	fos-choline-12	T _m	apparent melting temperature
FSEC	fluorescence-detection size-exclusion chromatography	TMD	transmembrane domain
GFP	enhanced green fluorescent protein	tsβ ₁ AR	thermo stable β ₁ adrenergic receptor
GPCR	G protein-coupled receptor	ZZ	tandem IgG binding domain (ZZ) based on <i>S. aureus</i> protein A
G _t	transducin		
H ₁₀	decahistidine		
H	helix		

CHAPTER 1 INTRODUCTION

1.1 Introduction to GPCRs

1.1.1 GPCR function and diversity

G protein-coupled receptors (GPCRs) are a class of integral membrane proteins which act as signal transducers, responding to stimuli as diverse as hormones and light. Encoded by more than 800 genes ¹, GPCRs encompass the largest and most diverse family of proteins in the human body ². Multiple complex intracellular signalling pathways are modulated by GPCRs, including responses to the majority of neurotransmitters and hormones; they are also key components of vision, olfaction and taste ³. Characteristically, GPCRs span the plasma membrane with seven α -helices and bind heterotrimeric G proteins to elicit their signal transduction pathway, although they are also known to signal independently from their eponymous pathway, mainly through coupling with the scaffold protein arrestin ^{3, 4}.

Within the human genome, GPCR sequences have been categorised into four major families based on sequence identity within the transmembrane domain (TMD) regions ^{5, 6}. The four families are comprised of the rhodopsin family (class A), the secretin and adhesion family (class B), the glutamate family (class C) ⁵ and the smoothed and frizzled family (class F) ⁶. In the human genome, class A is the largest group with 714 members ⁷ and binds the most varied set of ligands, ranging in size from small molecules to glycoproteins ⁸. Class A receptors are involved in diverse physiological processes such as vision, olfaction and immune response regulation, whereas some of the functions of class B and C receptors include blood glucose maintenance and synaptic transmission ⁹. In humans, class F consists of 10 frizzled receptors which control Wnt signalling and one smoothed receptor which controls the hedgehog pathway ⁶. Most class A receptors differ typically from others classes by having a relatively short N-terminus as compared to the long, heavily glycosylated N-terminus generally seen in class B, C and F receptors ⁸. Class B receptors consist primarily of orphan receptors and characteristically have two pairs of β -sheets and three disulphide bridges on their N-terminus, whereas class C receptors contain a 'Venus fly trap' domain on their N-terminus and class F receptors contain a cysteine rich domain on their N-terminus ^{6, 8}. Orthosteric ligand binding in class A occurs in the TMD bundle, whereas for class B receptors it extends from the TMD to the N-terminus, whilst for family C receptors the

CHAPTER 1 Introduction

evidence suggests that the agonist binds the Venus fly trap domain, which indirectly activates the TMD site⁸. For class F GPCRs, it is thought that the cysteine rich domain on the N-terminus of the receptors is involved in ligand binding⁶.

1.1.2 GPCRs as pharmaceutical drug targets

Given the role of GPCRs in a diverse array of cellular processes, it is not surprising that they comprise one of the largest families in the druggable genome and thus form a major class of targets for experimental drugs¹⁰. GPCRs exist in an equilibrium between an inactive conformation (R) and active conformation (R*), and the binding of a ligand can shift the balance towards one or other of these states¹¹. Ligand efficiency defines the biological response of a receptor to the binding of a ligand^{11, 12} (Figure 1.1). Ligands that act on GPCRs fall into two main categories, agonists and antagonists, which respectively either activate the receptor or block binding of the agonist^{11, 13}. Agonists can be further classified as full, partial or weak depending on the level of biological response observed¹³. Additionally, basal activity, or activity in the absence of agonist, can be suppressed by an inverse agonist^{12, 13} whereas neutral agonists neither activate nor suppress basal activity of the receptors¹³.

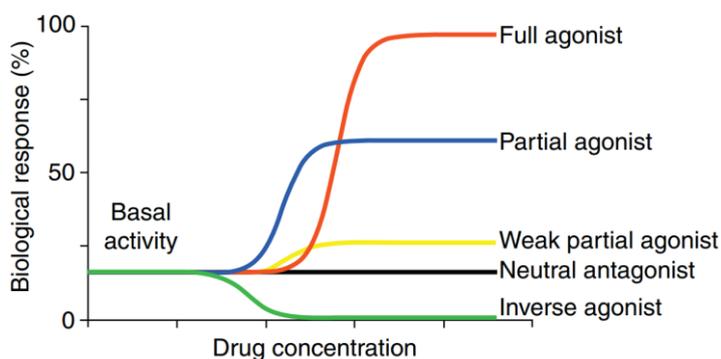


Figure 1.1 GPCR ligand efficiency

The five main classes of GPCR ligands based on ligand efficiency. Agonists activate the receptor to varying degrees (full, partial or weak). Antagonists block the binding of agonists and fall into two major groups; neutral antagonists which neither activate nor suppress basal activity of the receptors and inverse agonists which reduce basal activity. Figure adapted from¹³.

GPCR signalling can be further influenced through allosteric modulators which bind to sites on the receptor separate from the natural binding pocket^{4, 11}. Allosteric modulators can act by either changing the signalling pathway or altering the binding affinity of the endogenous ligand¹¹. For example, it was observed both pharmacologically and structurally that intramembrane Na⁺ ions act as an allosteric modulator of the adenosine A_{2A} receptor (A_{2A}R) by decreasing the affinity for the agonist with increasing Na⁺ concentration^{14, 15, 16, 17}.

CHAPTER 1 Introduction

Agonist binding can activate receptors in two main ways. Firstly, it is possible for an agonist to physically displace stabilising interactions, which in turn causes a conformational change that is stabilised by a new set of interactions. Secondly, agonists can act as a bridge, which creates new opportunities for interactions between TMDs that could, for example, stabilise the active conformation. For larger ligands such as peptides, there is also the possibility that a combination of mechanisms might apply¹². Coupling of cognate G proteins to a receptor can increase its binding affinity through allosteric effects. For example binding of G_s to β_2 AR increases its affinity for the agonist isoproterenol a hundred-fold¹⁸.

1.2 Crystallisation and structure determination of GPCRs

1.2.1 Milestones in GPCR structure determination

In an effort to understand GPCR function it is important to have a number of high-resolution structures of the receptors. The first three-dimensional GPCR crystal structure obtained was bovine rhodopsin at 2.8 Å¹⁹. Rhodopsin was the first GPCR to yield crystals due to its abundance from natural sources, stability in harsh detergents and the ability to control its conformational state²⁰. The first high-resolution rhodopsin structure had a P4₁ space group and the crystal contacts were formed between the hydrophilic domains of the receptor¹⁹. Subsequent rhodopsin structures have formed crystal contacts between extracellular loops (ECLs) and cytoplasmic loops (CLs), which had a P3₁ space group and possibly showed a more native configuration of CL3, since it closely matched that seen in electron diffraction structures of 2-dimensional crystals (Gebhard Schertler, personal communication)²¹. Other rhodopsin structures have been further refined to 2.2 Å²². The next high-resolution GPCR structures were of the human β_2 adrenergic receptor (β_2 AR)^{23, 24}, the avian β_1 adrenergic receptor (β_1 AR)²⁵ and the human A_{2A}AR¹⁶. These three structures represented a leap forward in GPCR structure determination because they illustrated the utility of newly developed protein engineering and crystallisation techniques (Section 1.3). With the advent of generic methodologies to both stabilise GPCRs and to make them more amenable to crystallography, the number and diversity of GPCR structures has dramatically increased (Figure 1.2). To date there are over 25 unique GPCR structures in the protein data bank (PDB) representing both active and inactive conformations and bound to ligands as diverse as small molecules and peptides⁹. A major milestone in GPCR structure determination was when the active state of β_2 AR in complex with its cognate G protein, G_s, was solved at 3.5 Å resolution²⁶. Stabilisation of this complex required fusion of β_2 AR with bacteriophage T4

CHAPTER 1 Introduction

lysozyme (T4L) as well as the binding of a single-chain camelid antibody fragment (nanobody) to the heterotrimeric G protein²⁷ (Section 1.2.4). More recently, structures of GPCRs from classes other than class A have been determined (Figure 1.2). These include structures of class B receptors for corticotropin-releasing factor¹²⁸ and glucagon²⁹; class C receptors, metabotropic glutamate receptor 1³⁰ and metabotropic glutamate receptor 5³¹; and a class F receptor for smoothed³². However all of these structures lack their N-terminal domains, which are essential for binding the native agonists.

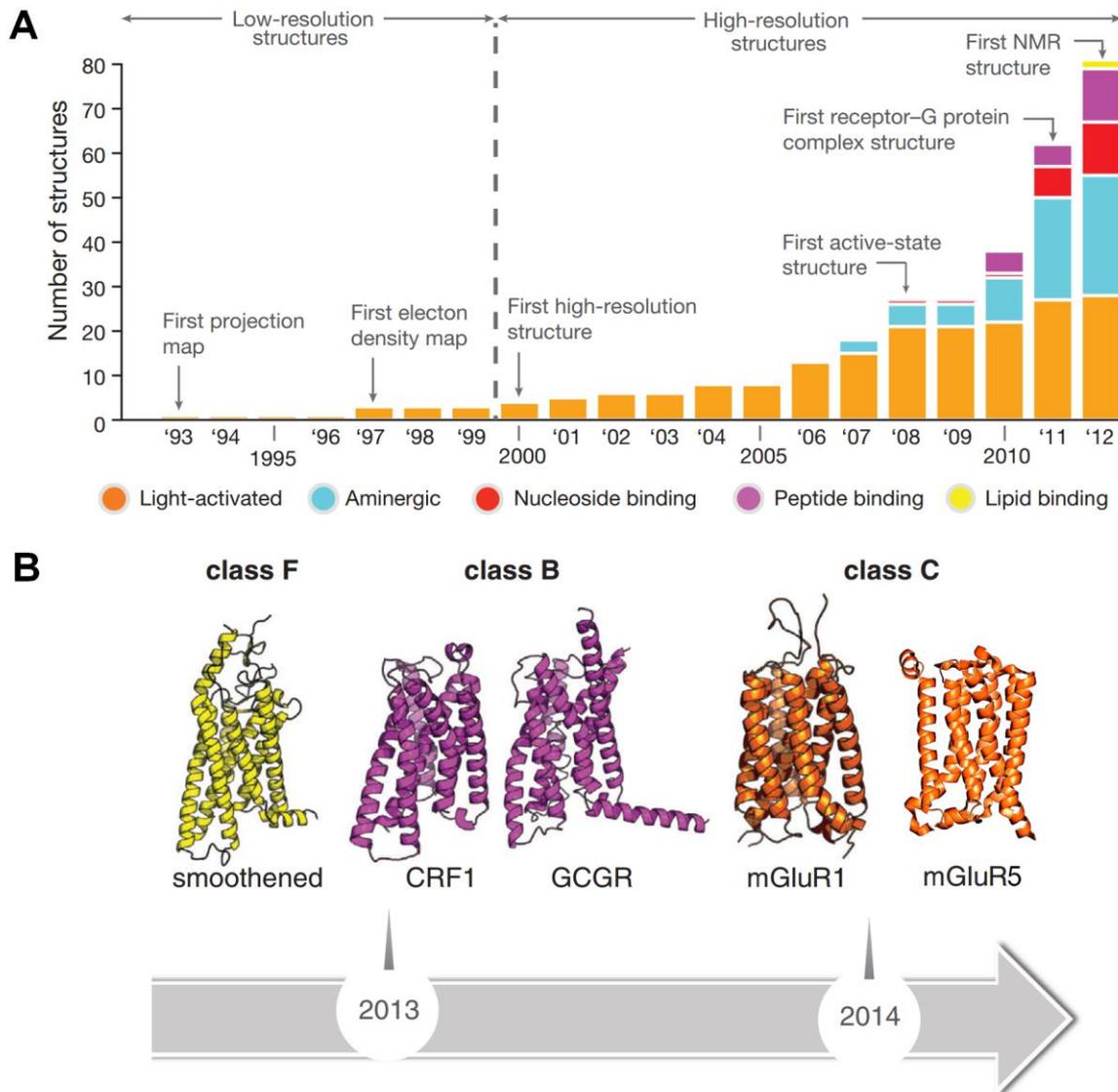


Figure 1.2 Time line and milestones of GPCR structure determination

(A) Bar graph indicating the increase in the number of GPCR structures from 1993 to 2012. Figure adapted from³³. **(B)** Milestones in GPCR structure determination 2013 to present. Shown are structures representing the smoothed receptor, corticotropin-releasing factor (CRF1), glucagon receptor (GCGR), metabotropic glutamate receptor 1 (mGluR1) and metabotropic glutamate receptor 5 (mGluR5). Figure adapted from⁹.

CHAPTER 1 Introduction

1.2.2 Common structural characteristics of GPCRs

To compare common structural characteristics of class A GPCRs, an alignment of the high resolution structures of β_1 AR and the dopamine D₃ receptor (D₃R) is shown in Figure 1.3. All GPCRs contain seven α -helical TMDs which are oriented approximately perpendicular to the membrane. The TMD helix bundle serves to transmit extracellular signals in the ligand binding pocket approximately 30 Å to the G protein coupling region³⁴. GPCRs contain an extracellular N-terminus and intracellular C-terminus and their TMDs are connected by a series of three cytoplasmic loops (CL1-3) and three extracellular loops (ECL1-3)², which form the most structurally variable regions of GPCRs³³. In class A receptors, the cavity of the ligand binding pocket is primarily located in the TMD bundle and usually only the side chains of the residues are involved in ligand binding³⁵. In the β_1 AR and D₃R structures, the ligands interact primarily with helix 3 (H3), H5 and H7³⁵. Located towards the centre of the TMD bundle, it is thought that H3 functions as the structural and functional hub since virtually every residue either forms contacts with another TMD or the cognate G protein³³. In inactive structures, Arg^{3.50} (Ballesteros Weinstein numbering³⁶ used in superscript throughout the document) in the highly conserved D[E]RY sequence, frequently forms a salt bridge with Asp[Glu]^{3.49}³⁷. Another key feature of H3 is a second salt bridge formed between Arg^{3.50} and Asp^{6.30} which creates an ionic lock that potentially holds together the intracellular ends of H3 and H6³⁷. The CWXP motif located near the bottom of the ligand binding pocket is thought to act as a toggle switch in some receptors, thus controlling receptor signalling through Trp^{6.48}. Activation of the receptor through agonist binding is thought to cause a rotational change in Trp^{6.48} which sets in motion a series of alterations in residues which in turn extend to the bottom of the TMD bundle³⁵. A further highly conserved motif is NPXXY which is located on the intracellular end of H7 and it serves as another switch controlling activation of the receptor. In inactive GPCR structures it is commonly found that the side chain of Tyr^{7.53} is oriented towards H1, H2 or H7, whereas in active structures it orientates toward H7 and interacts with H3 and H6³⁷. ECL2 often contains one or more disulphide bonds and it is sometimes involved in ligand binding, as observed for example in peptide receptors and A_{2A}R. In rhodopsin structures, this region contains β -sheets, whereas in many GPCRs, such as β_1 AR and β_2 AR, this region contains an α -helix³⁵. CL2 along with H5 and H6 are involved in G protein coupling²⁶.

CHAPTER 1 Introduction

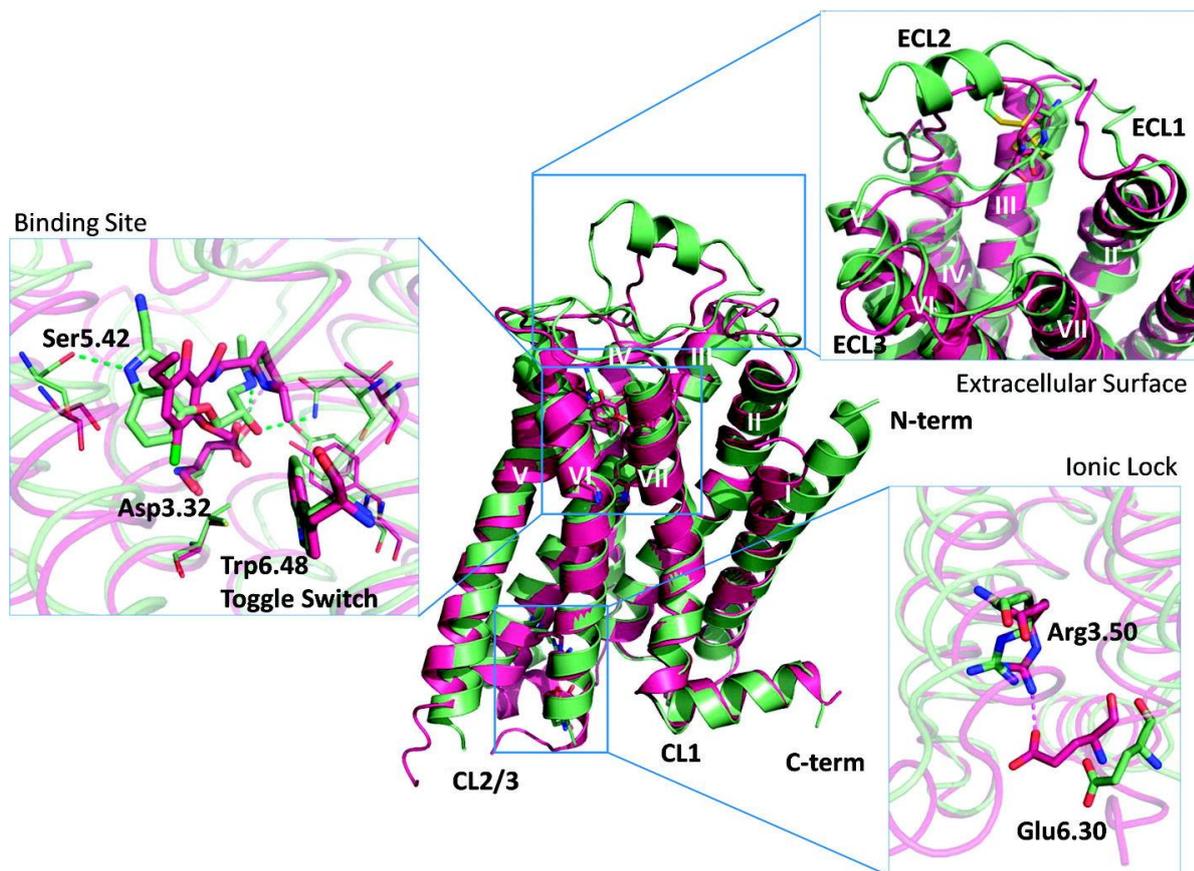


Figure 1.3 Common structural characteristics of class A GPCRs

The structures of β_1 AR (green, 2VT4) and D₃R (red, 3PBL) are shown as examples of class A high-resolution structures determined by X-ray crystallography. The seven helices are labelled using Roman numerals and are connected by a series of three CLs and three ECLs. Ligand binding involves the binding pocket in the TMD bundle, in particular to residues in H3, H5 and H7. The ionic lock is a salt bridge which holds together H3 and H6. In some receptors activation of the receptor causes a rotational change in Trp^{6.48} (toggle switch) which allows the receptor to transmit its signal to the bottom of the TMD bundle. For a full description of the key features of the structures see Section 1.2.2. Ballesteros Weinstein numbering is shown throughout³⁶. Adapted from⁸.

1.2.3 Signal transduction in GPCRs

Originally in the two-state model, GPCRs were thought of as simple bimodal switches that were either in an active or inactive state and, indeed, this model is accurate for describing rhodopsin activation¹². In the inactive state, rhodopsin is covalently bound to 11-*cis*-retinal¹⁹. Exposure to light isomerises 11-*cis*-retinal to all-*trans*-retinal which activates the receptor in milliseconds^{38, 39}. Basal activity varies greatly among receptors. Rhodopsin has virtually no basal activity, which is essential for its light-sensing function, whereas other receptors, such as the histamine H₃ receptor, have much higher levels of basal activity⁴⁰. Although rhodopsin has several characterised intermediate conformations, none of them have direct roles in interacting with signalling proteins²⁰. Therefore the two-state model holds true since absorption of a single photon of light causes maximal activation and coupling to transducin (G_t)¹². However, this is not the case for β_2 AR which can bind and signal through

CHAPTER 1 Introduction

both G_s and G_i as well as β -arrestin³. Even when saturated with a high affinity agonist, only a small portion of receptors manage to achieve the fully active state at any given time²⁰. The concept of β_2 AR being able to signal through several pathways is illustrated by the binding of the drug carvedilol, which acts as an inverse agonist for the G_s pathway, but as a partial agonist for the arrestin pathway¹⁸. Figure 1.4 is a simplified diagram displaying the wide array of signalling pathways activated by β_2 AR. GPCRs are now thought to exist in a variety of discrete conformations which are influenced by a number of agents including ligands, regulatory proteins, ions, pH and lipids²⁰.

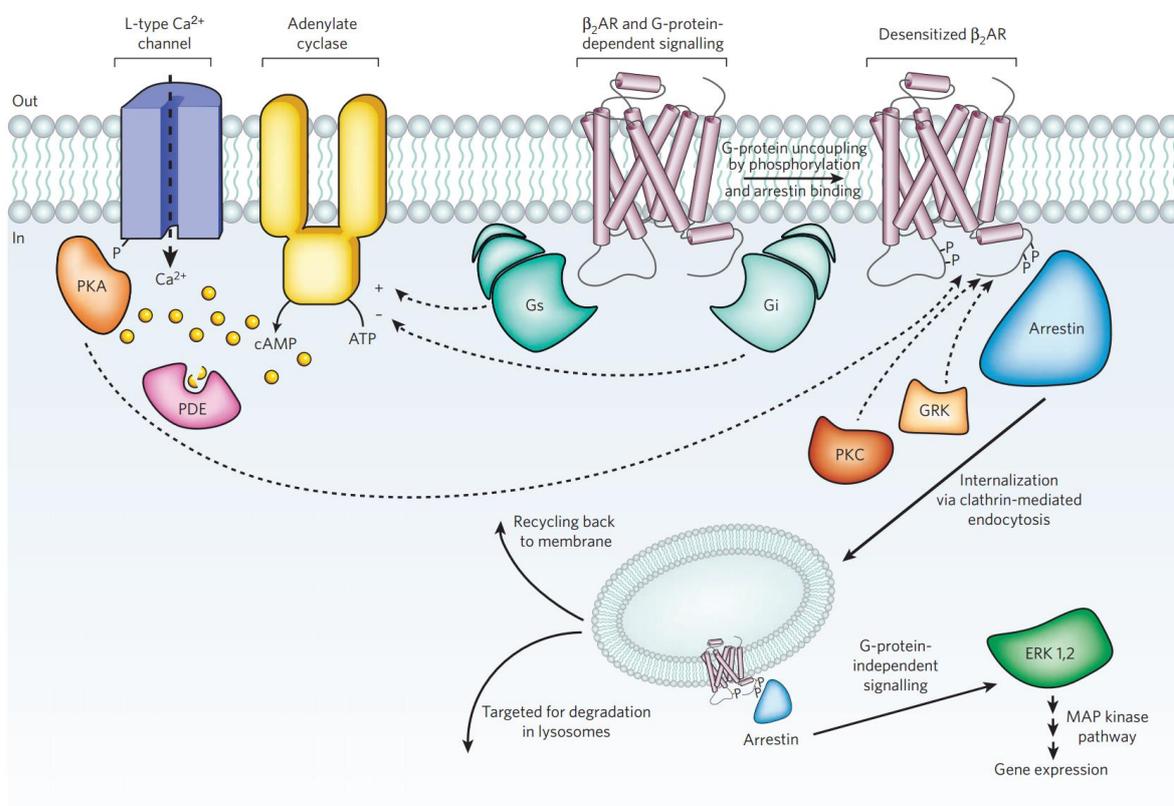


Figure 1.4 Signal transduction in GPCRs

A wide array of signalling pathways can be activated by β_2 AR. β_2 AR can bind to two different heterotrimeric G proteins, G_s and G_i , which differentially regulate adenylate cyclase and can in turn generate cyclic AMP (cAMP) leading to the activation of protein kinase A (PKA). PKA controls the activity of a number of proteins including L-type Ca^{2+} channels and β_2 AR itself forming a feedback loop. cAMP is downregulated by phosphodiesterase (PDE). G-protein-coupled receptor kinases (GRKs) and protein kinase C (PKC) phosphorylate β_2 AR after agonist-induced activation, which in turn allows β -arrestin coupling. β -arrestin can then signal through extracellular signal-regulated kinases (ERKs) and promote the internalisation of β_2 AR through clathrin-coated pits. Adapted from³.

1.2.4 Conformational changes in the activation of GPCRs

Initially, all GPCR structures were of receptors in the inactive conformation, because this state was the most stable and therefore more amenable to crystallisation. Full activation of a GPCR usually requires binding of an agonist and often also of the cognate G protein. To

CHAPTER 1 Introduction

date only three receptors have been crystallised in both inactive and fully active conformations, these include rhodopsin^{19, 38, 39}, β_2 AR^{23, 26, 41, 42}, and the M₂ muscarinic receptor^{43, 44}. However only β_2 AR has been crystallised in complex with the heterotrimeric G protein, G_s,²⁶ as well as a G protein mimetic, nanobody (Nb) 80⁴² as agonist alone was not enough to fully stabilise β_2 AR in the active conformation⁴⁵. The β_2 AR structures allow for a direct comparison of inactive and active conformations with a G protein bound (Figure 1.5). The most striking structural changes observed upon receptor activation are a 14 Å outward shift in the cytoplasmic end of H6; followed by the extension by two helical turns and an outward movement of the cytoplasmic end of H5; and finally more minor rearrangements in H7²⁶. Primarily, in this structure, it is the outward shift in H5 and H6 that forms a cavity which allows for G protein coupling. Small changes in the ligand binding pocket, largely through the formation of additional hydrogen bonds, are thought to set off a chain of rearrangements down the TMDs which ultimately lead to the large outward shift seen in H5 and H6⁴⁶. NMR studies of activated β_2 AR bound to Nb80 remain consistent with the movements shown upon activation⁴⁷. Additionally, in the structure of rhodopsin bound to a C-terminal G_t peptide, H6 shows a 6 Å outward displacement^{38, 39}. Further validating this model of activation, double electron-electron resonance (DEER) spectroscopy of rhodopsin has shown a 5 Å outward displacement of H6 upon activation of the receptor, along with smaller movements in H7⁴⁸. By contrast, crystal structures of A_{2A}AR in the intermediate-active state reveal only a 3 Å outward displacement of H6 upon agonist binding, which is smaller than that observed in either β_2 AR or rhodopsin, presumably due to the lack of a bound G protein^{49, 50}. The β_2 AR-Nb80 complex shows striking similarity to the β_2 AR-G_s complex with an overall root mean square deviation between the two structures of 0.6 Å. The only major differences are that Arg131 interacts with the nanobody in the β_2 AR-Nb80 structure whereas it packs against G_{αs} in the β_2 AR-G_s structure and H6 does not move as much in the β_2 AR-Nb80 structure^{26, 42} (Figure 1.5).

Until recently the mechanism for how GPCRs activate G proteins and release GDP from the G_α subunit was unknown⁵¹. With the structure of the β_2 AR-G_s complex, it is now possible to understand this mechanism (Figure 1.6). G proteins are able to couple to the GPCR after it is activated by an agonist and the cellular domains of H5 and H6 form a large pocket to enable this binding to take place.

CHAPTER 1 Introduction

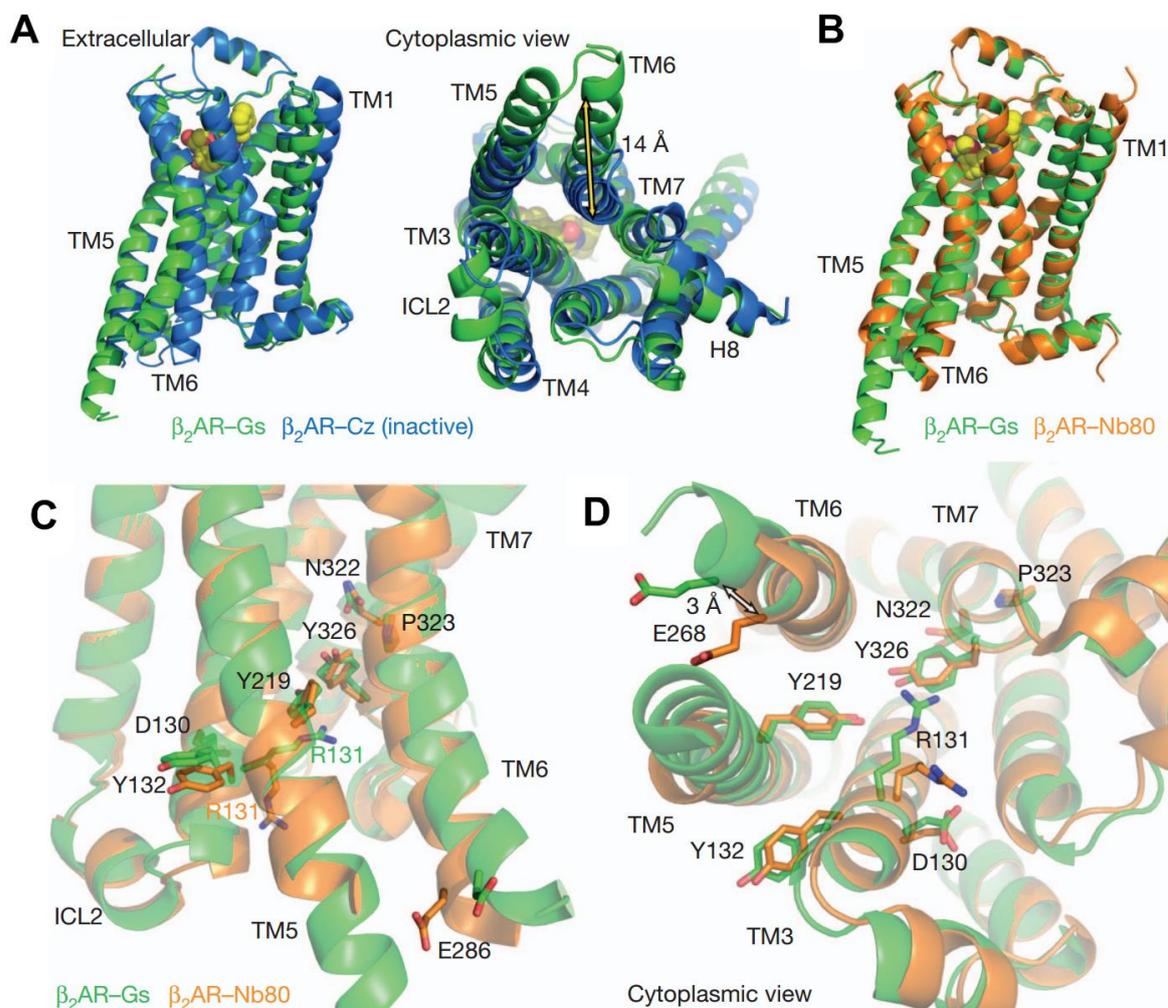


Figure 1.5 Comparison of active and inactive β_2 AR structures

(A) Side and cytoplasmic views of the β_2 AR-G_s structure (green) compared to the inactive carazolol (CZ)-bound β_2 AR structure (blue). The largest changes in conformation are located in the intracellular end of H5 and H6 (note Hs are annotated by TMs in the diagram), which allows the G protein to bind. H5 is extended by two helical turns whereas H6 is moved outward by 14 Å as measured at the α -carbons of Glu268 (yellow arrow) in the two structures. **(B)** Comparison of β_2 AR-G_s (green) with nanobody-stabilised active state (β_2 AR-Nb80) (orange). **(C)** The positions of residues in the E/DRY and NPXXY motifs and other key residues of the β_2 AR-G_s and β_2 AR-Nb80 structures. All residues occupy very similar positions except Arg131 which in the β_2 AR-Nb80 structure interacts with the nanobody. **(D)** Cytoplasmic view of residues shown in illustrating a 3 Å shift in H6 (C). Adapted from ²⁶.

From the crystal structure of the β_2 AR-G_s complex, CL2, H5 and H6 of β_2 AR form a large interface with G _{α s} ²⁶. The formation of the complex allows the exchange of GDP for GTP and subsequent activation of the α , β and γ subunits of the G protein. Eventually they are deactivated and recycled. One of the major challenges in crystallising the β_2 AR-G_s complex was the variable position of the α -helical domain of the G _{α s} subunit as revealed by electron microscopy (Figure 1.6). This was resolved by binding Nb35 to the G _{α s} subunit, which locked it into one position ²⁶.

CHAPTER 1 Introduction

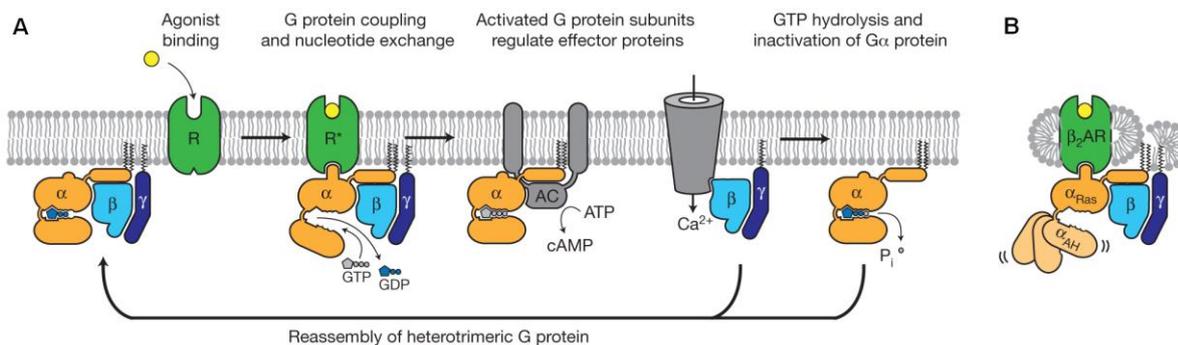


Figure 1.6 Activation and G protein coupling for β_2 AR

(A) Agonist binding of β_2 AR causes a rearrangement in the cytoplasmic ends of the TMDs which allows coupling of the receptor to the G α_s heterotrimer (α , β and γ). Formation of the β_2 AR-G α_s complex leads to the release of GDP which in turn allows for the binding of GTP which causes the α , β and γ subunits to disassociate from β_2 AR. The subunits are then free to activate their associated effector proteins (adenylate cyclase (AC) and calcium channels in this case). The G α_s heterotrimer is then reassembled from the α , β and γ subunits after GTP is hydrolysed to GDP. (B) Detergent solubilised, nucleotide-free β_2 AR-G α_s complex, illustrating the complexity of working with a nucleotide free G α_s . Two nucleotide binding subunits comprise G α_s , the α -helical domain (α AH) and Ras domain (α Ras). When nucleotides are removed, the α AH domain position varies relative to the α Ras. Adapted from ²⁶.

1.2.5 Conformational changes in GPCR- β -arrestin coupling

Significantly less is understood about the structural changes that occur in GPCRs when β -arrestin couples, compared to when G proteins bind. Structures of β_1 AR bound to biased agonists have not identified a clear mechanism for why some ligands are biased and why some are not ⁵². However, there is some evidence to suggest that the cytoplasmic end of H7 alters conformation upon arrestin binding ⁵³, which correlates with the importance of the phosphorylated state of the C-terminus in dictating the model of arrestin binding ⁵⁴. Single particle electron microscopy has revealed that the C-terminus and H5 in β_2 AR are important for coupling with arrestin-1 ⁵⁵. However the resolution of the structure was only about 30 Å, thus insufficient detail was resolved to define the model of arrestin binding. A high resolution structure of an arrestin bound GPCR would help to explain the mechanism of arrestin coupling.

1.3 Requirements for GPCR crystallisation and strategies developed to overcome impediments

1.3.1 Difficulties in crystallising GPCRs

Despite the increase in the number of GPCR structures in the past five years, they remain challenging targets for crystallography. This is due to several reasons. GPCRs exist in equilibrium between two main conformations, a low energy inactive R state which is

CHAPTER 1 Introduction

favoured in the absence of ligand and a higher energy active R* state that binds G proteins⁵⁶. This conformational heterogeneity is not conducive to crystallisation and therefore it is essential to lock the receptors into one state. GPCRs have a relatively low hydrophilic surface area which can be used to form crystals⁴⁶. Compounding this problem, GPCRs are inherently flexible and unstable⁴⁶ and the harsh short chain detergents, which are required to expose hydrophilic faces, act to further destabilise them. The majority of GPCRs are not present from natural sources in quantities sufficient for structure determination^{57, 58}.

Therefore, structural projects depend upon using an overexpression system capable of producing milligram amounts of active receptor, but unfortunately, low expression levels are typically obtained for mammalian membrane proteins produced from recombinant sources⁵⁹. In addition, post-translational modifications, such as N-glycosylation, are often required for efficient mammalian membrane protein expression, folding and targeting to the plasma membrane. As shown for the N-glycosylated membrane protein, the serotonin transporter (SERT), the expression of fully active eukaryotic proteins cannot always be accomplished in heterologous expression systems⁶⁰. Expression of GPCRs will be further discussed in Chapter 2. Finally, the stability of the GPCR needs to be sufficient for purification in the detergent-solubilised state and also for the days or weeks required for crystallisation. Thus, for each potential structural biology project on GPCRs, three problems need to be overcome to optimise the probability of success: (1) milligram amounts of functional receptor expressed in a heterologous system; (2) sufficient stability of either the receptor alone or in complex with a ligand to allow purification and crystallisation; (3) a method to lock the receptor in a single conformation.

1.3.2 Strategies for the crystallisation of GPCRs

Different strategies have been developed to overcome issues with GPCR crystallisation. Two separate approaches are used to increase the hydrophilic surface area of membrane proteins, which in turn allows the use of milder detergents with larger micelles. $\beta_{2A}R$ was co-crystallised with a fragment antigen-binding (F_{ab}) antibody which provided sufficient hydrophilic surface contacts to allow for crystallisation in bicelles, which are lipid rich and consequently much larger than the detergent micelles usually employed²³. However the resultant structure was at a resolution of 3.4 Å and thus insufficient to model side chains in the extracellular portion of the receptor or to delineate the ligand. Another example of a F_{ab} being used to stabilise a GPCR was the co-crystallisation of a conformationally-specific F_{ab} with $A_{2A}R$ ⁶¹ (Figure 1.7), which yielded a structure at a resolution of 2.7 Å. The use of

CHAPTER 1 Introduction

antibodies to stabilise receptors took a leap forward with the use of nanobodies (Nbs). These are derived from antibodies found in camelids, which are comprised of a single domain fragment that maintains the antigen-binding capacity of traditional antibodies, but which are only 15 kDa in size²⁷. Nbs are thought to be more stable than traditional antibodies, including F_{abs}, and can be easily expressed in bacteria and yeast, however their production requires milligram quantities of the purified protein that they are to be raised against²⁷. The resulting Nbs then need to be screened to isolate ones which recognise a single conformation of the protein of interest²⁷. A structure of active-state β_2 AR in complex with Nb80 was solved at 3.5 Å, however this structure also required the use of a fusion protein, bacteriophage T4 lysozyme (T4L; discussed in more detail below), to further increase crystal contacts⁴² (Figure 1.5). Interestingly, Nb80 increased the affinity of the agonist isoproterenol by a hundredfold, which was similar to the increase observed when G_s binds to β_2 AR²⁷, indicating that Nb80 mimics some of the effects of the G protein. A further nanobody, Nb35, was raised against β_2 AR crosslinked to the G_s heterotrimer²⁶ and was essential during crystallisation to prevent the β_2 AR-G_s complex from dissociating²⁶. The resulting structure also employed T4L to increase crystal contacts and was solved at 3.2 Å²⁶ (Figure 1.5).

As mentioned above, another strategy for increasing the likelihood of crystal contacts forming is to create a fusion with a small soluble protein such as T4L. This approach was also used to crystallise β_2 AR bound to a high-affinity inverse agonist and resulted in a 2.4 Å resolution structure²⁴. The crystals were grown in lipidic cubic phase, because vapour diffusion crystal trials did not yield diffraction-quality crystals. While the addition of a fusion protein can introduce flexibility, the careful placement of T4L in CL3 was successful in reducing flexibility in this region, however this particular placement can also prevent G protein coupling⁴¹. Instead T4L can be fused to the N-terminus to maintain the ability of the receptor to bind G proteins²⁶. Another example of a GPCR-T4L fusion is A_{2A}AR-T4L that was solved at 2.7 Å resolution⁵⁰ (Figure 1.7). Thusfar the use of T4L fusion has been employed to crystallise over eleven unique GPCRs⁶². This technique has also been extended to include other small soluble proteins such as apocytochrome b₅₆₂RIL (BRIL). A fusion of A_{2A}AR with BRIL inserted into CL3 also resulted in a 1.8 Å structure¹⁴. The technique of creating a fusion with a small soluble protein however relies on the availability of high-affinity ligands with slow dissociation rates, which are essential to both make the receptor sufficiently stable to allow purification and crystallisation and to stabilise the receptor in a single conformation¹³.

CHAPTER 1 Introduction

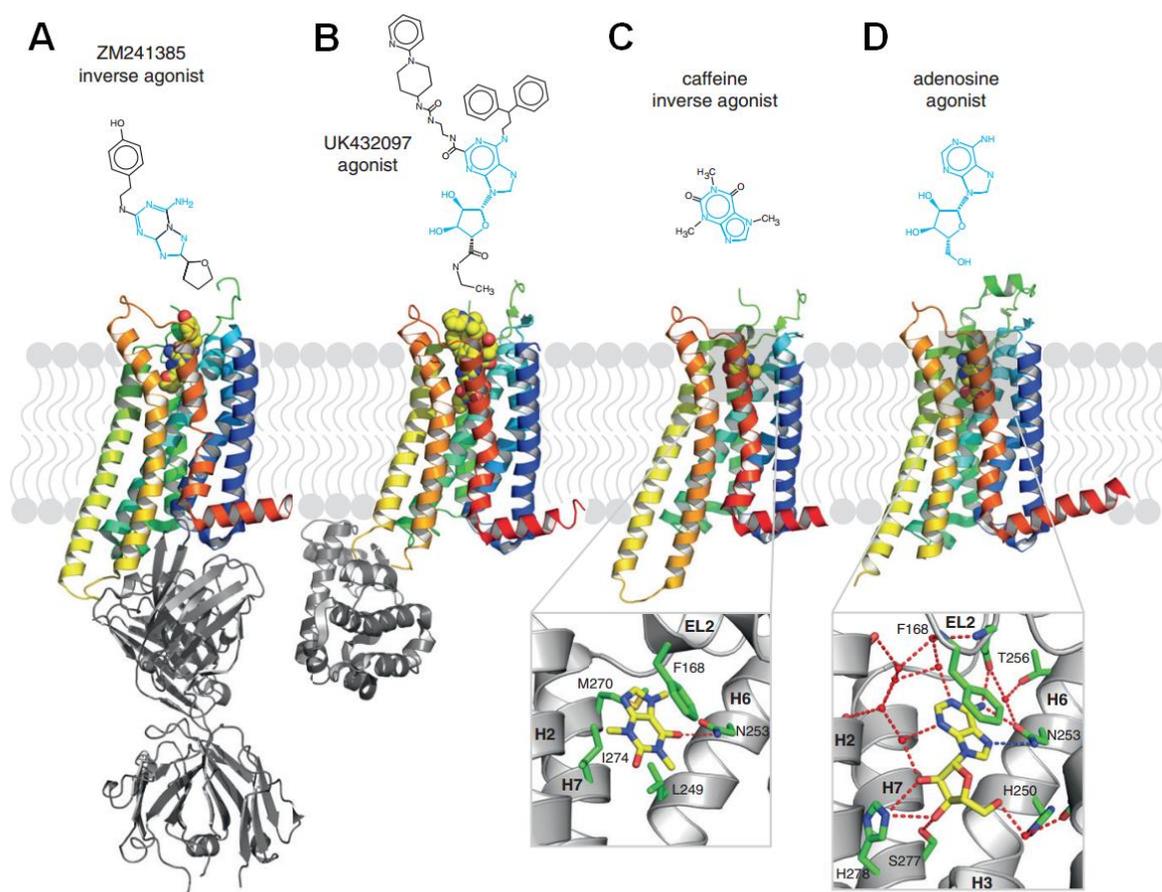


Figure 1.7 Comparison of three strategies used for the crystallisation of A_{2A}R

Ribbon diagram in rainbow colouration (N-terminus blue, C-terminus red) illustrating three different techniques used for structure determination of A_{2A}R bound to endogenous and synthetic ligands. Strategies employed include increasing the hydrophilic area through the use of a F_{ab} (A), creation of a fusion with T4 lysozyme in CL3 (B) or conformational thermostabilisation (C, D). A cartoon version of the lipid bilayer is shown for perspective. Ligands are shown in space-filling representation (C, yellow; N, blue; O, red). Above the receptors, the portions of the ligands that are similar to adenosine are shown in blue. Detailed inserts show the binding pocket for caffeine and adenosine: part of H3 has been removed for clarity; red broken lines, hydrogen bonds; blue broken lines, polar interactions; red spheres, water molecules; interacting amino acid side chains (C, green; N, blue; O, red); ligands are shown in stick representation (C, yellow; N, blue; O, red). PDB codes: (A), A_{2A}R bound to F_{ab} (grey) 3VG9⁶¹; (B), A_{2A}R-T4L (T4L, grey) 3QAK⁵⁰; (C), thermostabilised A_{2A}R bound to caffeine⁶³ 3RFM; (D), thermostabilised A_{2A}R bound to adenosine⁴⁹ 2YDO. Adapted from¹³.

Another mechanism for enhancing the likelihood of crystal formation stems from the observation that point mutations can increase the stability of a detergent-solubilised membrane protein⁶⁴. This offers the opportunity to engineer proteins that are stable enough to allow the use of short chain detergents for crystallisation. While these detergents expose the greatest surface area of the membrane protein, they are by nature very destabilising and therefore denature many proteins. Thus the crystallisation of membrane proteins in detergents such as nonylglucoside or octylglucoside only works for thermally stable proteins. Conformational thermostabilisation consists of evolving proteins through mutagenesis to be thermally stable whilst simultaneously selecting for proteins locked preferentially in one conformation^{17, 65, 66, 67, 68, 69}. The choice of assay to measure an apparent T_m is discussed in Chapter 2. For β₁AR, conformational thermostability was

CHAPTER 1 Introduction

achieved through selecting stable mutants in complex with the antagonist dihydroalprenolol in the inactive conformation⁶⁹. Thermostabilisation allowed the protein to be crystallised in the harsh detergent n-octyl β -D-thioglucoopyranoside (OTG) with cyanopindolol bound²⁵. Conformational thermostabilisation of GPCRs has allowed several further crystal structures to be obtained, most recently a class C GPCR, the metabotropic glutamate receptor 5³¹ and a class B GPCR, the corticotropin-releasing factor receptor²⁸. The same approach has also been used to determine the structures of several class A receptors including: the adenosine A_{2A} in both an R*-like state⁴⁹ (Figure 1.7) and an R state⁶³ (Figure 1.7); the neurotensin receptor in an R*-like state⁷⁰; and β_1 AR in the R state bound to agonists, partial agonists, weak partial agonists and biased agonists^{25, 52, 71, 72, 73}. The advantage of conformational thermostabilisation over the other techniques mentioned here is that it allows for the crystallisation of GPCRs with low-affinity ligands¹³.

1.4 The angiotensin II type 1 receptor

1.4.1 Control of the cardiovascular system by AT₁R

A highly potent hormone, recognised as a key regulator of blood pressure, was identified by two separate groups in the 1940s. The substance was separately named hypertensin and angiotonin and as a compromise, the portmanteau angiotensin was adopted as the standard nomenclature⁷⁴. The precursor to angiotensin, angiotensinogen, is a 56 kDa glycoprotein that is released into the blood primarily from the liver. It is cleaved by the protease renin to ten amino acids, which are then further cleaved by angiotensin converting enzyme (ACE) to create the biologically active octapeptide angiotensin II (Asp-Arg-Val-Tyr-Ile-His-Pro-Phe)⁷⁵ (Figure 1.8). Angiotensin II is the endogenous ligand for the angiotensin II type 1 receptor (AT₁R) which is a member of the class A (rhodopsin) family of GPCRs⁷⁶. AT₁R is involved in homeostasis of the cardiovascular system. It is expressed in a variety of tissues including vascular smooth muscle, adrenal glands, the kidneys and the brain⁷⁶. Binding of the endogenous agonist angiotensin II causes activation of G_q-mediated phospholipase C, which in turn stimulates inositol phosphate responses, Ca²⁺ signal generation and protein kinase C activation^{77, 78} (Figure 1.8). Agonist activation of AT₁R also elicits several intercellular signalling pathways including Ras and JAK/STAT⁷⁵. Cellular responses to angiotensin II binding are tissue specific and include vasoconstriction, sodium re-absorption, increased thirst and ventricular hypertrophy⁷⁶.

CHAPTER 1 Introduction

Independent of the eponymous G protein pathway, AT₁R can signal through binding of the scaffold protein β -arrestin. As part of GPCR desensitization, Ser and Thr residues on the C-terminus of AT₁R are phosphorylated by a variety of kinases, which then allows β -arrestin to bind^{78, 79} (Figure 1.8). Even after AT₁R is targeted to endocytic vesicles, it can signal via β -arrestin, since unlike most GPCRs such as β_2 AR, the complex is stable and not readily dephosphorylated⁸⁰. From the perspective of β -arrestin binding, AT₁R is categorised as a class B receptor, which means that AT₁R is internalised along with β -arrestin and can still activate numerous intercellular signalling pathways including, Src, ERK and JNK3^{81, 82} (Figure 1.8). These pathways in turn regulate cellular processes such as proliferation and apoptosis⁸².

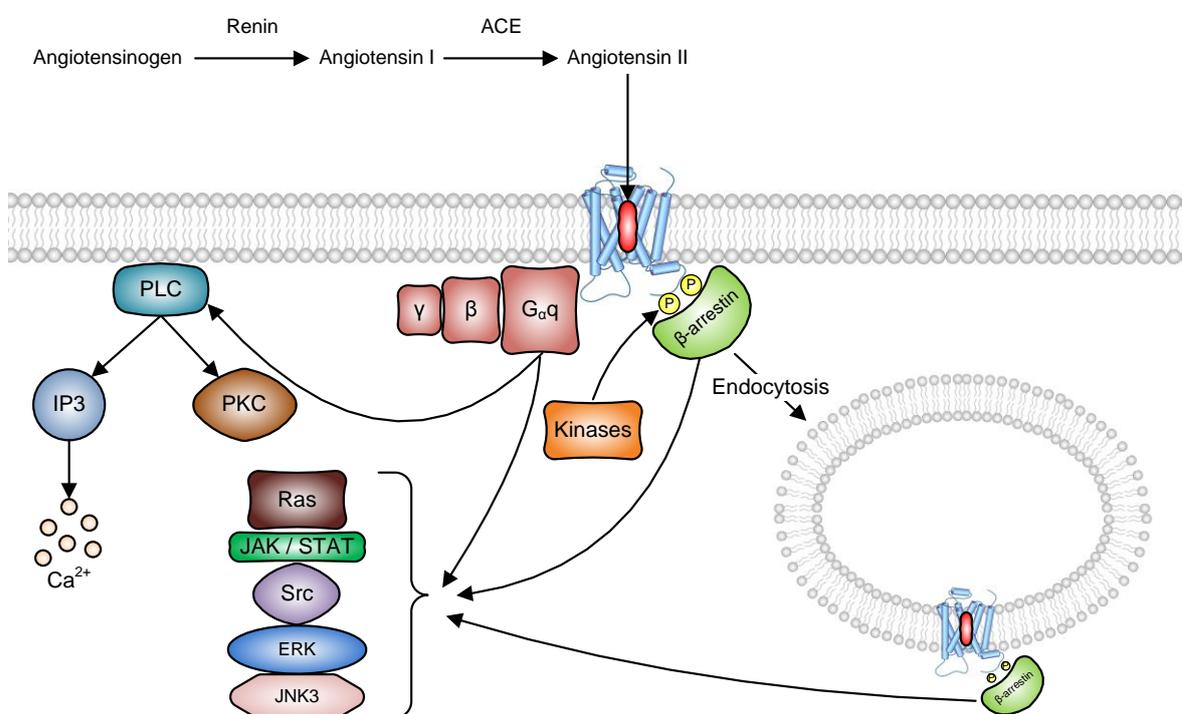


Figure 1.8 Agonist activation and desensitisation of AT₁R

A wide array of signalling pathways can be activated by AT₁R and a simplified version is shown here. Upon agonist binding, AT₁R binds primarily to the heterotrimeric G protein, G_q which in turn activates PLC (phospholipase C), which then initiates IP (inositol phosphate) responses including IP₃, Ca²⁺ signal generation and PKC (protein kinase C) activation. PKC controls the activity of a number of proteins including AT₁R itself, forming a feedback loop. Activation of AT₁R also mediates several intracellular signalling pathways including Ras and JAK/STAT. G-protein-coupled receptor kinases and PKC phosphorylate AT₁R after agonist-induced activation, which in turn allows β -arrestin coupling. β -arrestin can then signal through Ras (rat sarcoma GTPase), JAK/STAT (janus kinase/signal transducer and activator of transcription), Src (sarcoma tyrosine kinase), ERK (extracellular signal-regulated kinase) and JNK3 (c-Jun-N-terminal kinase 3) and it promotes the internalisation of AT₁R through clathrin-coated pits, where it can still signal. Adapted from³.

1.4.2 Analysis of the primary structure of AT₁R

Expression cloning was first used to isolate AT₁R cDNA from bovine adrenal cells and rat aortic vascular smooth muscle cells which were both abundant sources of the receptor and in

CHAPTER 1 Introduction

subsequent years AT₁R cDNA was obtained from several other vertebrates⁷⁵. An amino acid sequence alignment of AT₁R between the human receptor and that of other vertebrate species shows an extremely high level of conservation (Appendix 1). An additional alignment of human AT₁R with GPCRs of known structure (Appendix 1) allows for the postulation of the location of the transmembrane domains. Sequence homology between AT₁ receptors is greatest in the transmembrane domains and three intracellular loops and is least within the 4 extracellular regions and the C-terminus. There are potentially two highly conserved disulphide bonds present in AT₁R. As seen in many GPCR structures, there is the possibility of a disulphide bond between ECL1 and ECL2. AT₁R also contains an additional pair of cysteine residues putatively located in the N-terminal region and ECL3, whereas the other cysteine residues are most likely located in transmembrane domains. These disulphide bonds could impart a level of stability to the receptor⁸³. As an example, a disulphide bond was engineered to connect the N-terminus and ECL3 of rhodopsin, increasing its thermostability by 10°C⁸⁴. AT₁R possesses three highly conserved N-linked glycosylation consensus sequences (Asn-X-Ser/Thr, where X is not Pro) in extracellular segments. Asn4 is at the N-terminus of the receptor while Asn176 and Asn188 are in ECL2. This source of heterogeneity could pose a problem for crystallisation. N-linked glycosylation of AT₁R will be further discussed in Chapter 2.

AT₁R has been shown to be held in an inactive conformation and agonist binding is thought to destabilise these interactions. Specifically, Asn111^{3,35} interacts with Asn295^{7,46} to produce a locking mechanism⁸⁵. Mutation of Asn111^{3,35} to alanine produces a constitutively active receptor⁸⁶ and Asn111^{3,35} has been shown to interact with Tyr4 of angiotensin II⁸⁷. Therefore it is logical to postulate that agonist binding might destabilise the interactions between Asn111^{3,35} and Asn295^{7,46}¹². AT₁R possesses the highly conserved DRY motif and mutation of this sequence prevents G protein coupling after agonist binding, however it does not diminish receptor phosphorylation, internalisation or recruitment of β-arrestin⁸⁸.

1.4.3 Pharmacology of AT₁R

As the gatekeeper of the cardiovascular system, AT₁R has generated considerable pharmacological interest over many years⁷⁶ and a number of ligands have been designed which block it. Initial attempts at designing AT₁R antagonists were complicated by the need to work with peptide analogues of angiotensin II. As such, the first AT₁R antagonist, saralasin (Sar¹-Val⁵-Ala⁸ angiotensin II), had to be administered intravenously and even

CHAPTER 1 Introduction

presented partial agonist action in some patients^{76, 89}. The first non-peptide AT₁R antagonist to reach clinical trials was losartan, followed shortly by a large number of derivative compounds, such as valsartan and candesartan⁷⁶. Unlike the peptide analogues, these compounds showed oral bioavailability and were effective at reducing blood pressure⁸⁹. Further development of non-peptide AT₁R antagonists was aided by cloning of the *Xenopus laevis* AT₁R DNA sequence. The amphibian receptor bound angiotensin II with a similar affinity to its mammalian counterpart, but it did not recognise losartan⁹⁰. Mutation of non-conserved amino acids in the rat AT₁R to those of the amphibian version obliterated losartan binding, whereas only small changes in the affinities for angiotensin II and the peptide antagonist Sar¹-Ile⁸ angiotensin II (Sar¹) were observed. Residues Val108^{3,32}, Ala163^{4,60}, Thr198^{5,41}, Ser252^{6,47}, Leu300^{7,51} and Phe301^{7,52} were shown to be of particular importance for non-peptide antagonist binding to AT₁R and all were postulated to be located in transmembrane domains⁹⁰. The non-peptide agonist, L-162,313 has a physiological effect similar to the endogenous agonist⁹¹. The most recent AT₁R blocker to gain FDA approval, azilsartan, had a similar binding affinity to other non-peptide antagonists, but showed stronger inverse agonism, as determined by inositol phosphate production, and stayed bound to the receptor for a longer period of time⁹². The structures of eleven AT₁R ligands along with their binding affinities are shown in Chapter 4.

AT₁R ligands also include a number of biased agonists such as Sar¹-Ile⁴-Ile⁸ angiotensin II, which was the first ligand to be discovered that favoured signalling through β -arrestin rather than G_q^{93, 94}. Further compounds have since been developed that showed an even stronger bias towards β -arrestin signalling, such as TRV120027, which exhibited a 30-fold increase towards β -arrestin 2 recruitment over Sar¹-Ile⁴-Ile⁸ angiotensin II^{95, 96}. This discovery of functional selectivity illustrated that the classical model of GPCR activation from the R to R* state was incomplete and multiple ligand-specific conformations (R^{*n}) could exist, each of which could be capable of eliciting different downstream signalling⁹⁷. Although several AT₁R biased agonists have been developed, there is currently no inverse agonist for both the G protein and β -arrestin pathways. For example, although Sar¹-Ile⁴-Ile⁸ angiotensin II was originally shown to be biased towards β -arrestin signalling, additional assays which directly monitored the interactions between AT₁R and the transducers (as opposed to assessing downstream signals such as cAMP or ERK1/2 phosphorylation) showed that it partially activated G proteins⁹⁸.

CHAPTER 1 Introduction

1.5 Aim of the research presented in this dissertation

Although over 25 unique high-resolution GPCRs structures now exist, relatively little is known about the conformational dynamics of GPCRs, in particular the structural changes that occur after G proteins and arrestin bind. To date, there is only one structure of a GPCR bound to a G protein, β_2 AR in complex with G_s ²⁶. Since AT_1R couples to G_q it presents an opportunity to understand G protein coupling with G proteins other than G_s . Additionally, since AT_1R forms a stable complex with arrestin and stays bound to arrestin longer than most GPCRs⁸⁰, it presents an opportunity to further understand GPCR-arrestin interactions.

The aim of the work presented here was to develop an effective heterologous expression system for the production of milligram quantities of AT_1R for structure determination. In addition, the effect of N-linked glycosylation on protein expression and stability had to be determined, as well as the best combination of ligand and detergent to produce an optimally stable receptor suitable for purification.

CHAPTER 2 THE SUITABILITY OF AT₁R AS A CANDIDATE FOR STRUCTURAL STUDIES

2.1 Introduction

Before structure determination of AT₁R could proceed several potential obstacles needed to be assessed and, where necessary, overcome. The first aim was to examine the stability of detergent-solubilised AT₁R. The second was to determine the expression levels of functional AT₁R in heterologous systems and to determine the most efficient system for producing milligram quantities.

Prior to the work presented here, the stability of detergent-solubilised AT₁R was unknown. Several analytical methods exist for measuring protein stability with respect to temperature. These include sedimentation velocity^{99, 100}, circular dichroism spectroscopy^{101, 102, 103}, dynamic light scattering^{99, 104} and the CPM unfolding assay¹⁰⁵. However, all these methods require purified protein and, in the case of the biophysical techniques, usually milligram quantities are necessary. In contrast, the CPM assay was developed to work on microgram quantities of membrane proteins. The CPM assay uses a maleimide fluorochrome which is nonfluorescent in its unbound state and fluoresces upon covalent attachment to accessible cysteine residues. A further benefit is that it is unaffected by the presence of detergent¹⁰⁵. Another method for determining protein stability is the use of green fluorescent protein (GFP) fusions coupled with fluorescence-detection size exclusion chromatography (FSEC)^{106, 107}. The advantage of this method is that it uses unpurified material, however minor changes in stability are not easily measured by this technique. A further method to assess protein stability is the radioactive ligand-bound apparent thermostability (T_m) assay^{17, 65, 66, 67, 68, 69}. The T_m assay has the advantages of not requiring purified material, being rapid and easy to use on multiple samples, and requiring only a few nanograms of the target membrane protein. The apparent T_m of a protein is defined as the temperature at which 50% of the ligand remains bound after a 30 minute incubation. The T_m assay has been successfully used to evaluate thermostabilising mutations in GPCRs, which in turn has allowed several crystal structures to be obtained. These include: metabotropic glutamate receptor 5³¹, corticotropin-releasing factor receptor²⁸ adenosine A_{2A} in both an R*-like state⁴⁹ and an R state⁶³, neurotensin receptor⁷⁰ and β₁-adrenergic^{25, 52, 71, 72, 73}. Given these advantages, the T_m assay was selected to examine the stability of AT₁R.

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

The turkey (*Meleagris gallopavo*) β_1 -adrenergic receptor was shown to have an apparent T_m that was 22°C higher in the detergent DDM than the orthologous human protein, this was despite having 76% sequence identity¹⁰⁸. This indicates that the amino acid sequence of a protein governs its thermostability, which therefore can be improved by alteration of the protein through making point mutations or by creating a fusion protein. The addition of a ligand can also increase the apparent T_m of a protein^{67, 68}.

Three methods for performing the thermostability assay were developed. In the [-] format, the detergent-solubilised receptor is heated before ligand is added and in the [+] format ligand is added to the solubilised receptor before heating. However, not all proteins are sufficiently stable to be measured in these conditions. Therefore the [Super +] format was developed, where ligand is added to the membrane preparation and allowed to bind to the receptor, the receptor is then solubilised in detergent and the mixture is finally heated (Figure 2.1).

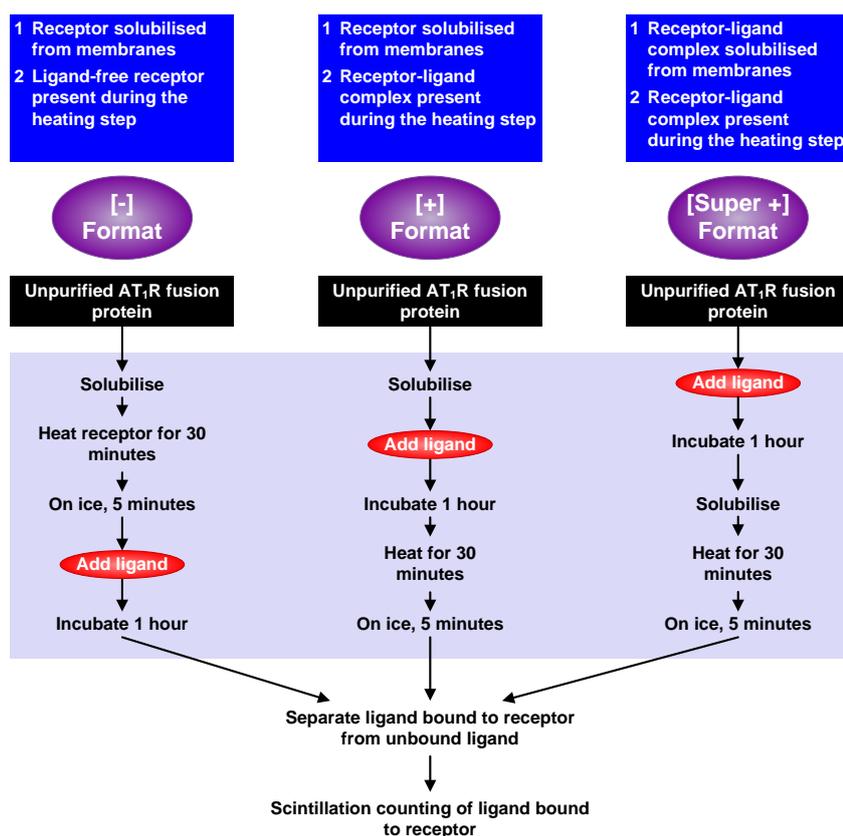


Figure 2.1 Schematic of the three methods for the thermostability assay.

Figure adapted from⁶⁸.

The apparent T_m derived from the [-] format is indicative of the overall stability of the ligand-free receptor, whereas the apparent T_m derived from the [+] and [Super +] formats is

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

descriptive of the stability of the ligand-bound receptor⁶⁸. Assessment of the apparent T_m of AT₁R in all formats would give a good indication of the likelihood of crystal formation and suggest whether more work was needed to stabilise the receptor by mutagenesis. The next step would then be to produce sufficient quantities of AT₁R for structural studies.

If the apparent T_m of AT₁R was deemed to be sufficiently high, focus would then shift to obtaining enough correctly folded AT₁R to proceed with crystallisation trials. The majority of GPCRs are not present from natural sources in quantities sufficient for structure determination^{57, 58}. Therefore, structural projects depend on using an overexpression system capable of producing increased amounts of active receptor. In theory AT₁R, could be overexpressed in *Escherichia coli* (*E. coli*) and yeasts, but these have had limited success in producing sufficient amounts of mammalian membrane proteins for structural biology, most likely because these systems have not evolved to fold mammalian membrane proteins^{59, 109, 110}. To date only one GPCR, the neurotensin receptor, expressed recombinantly in *E. coli* has resulted in a crystal structure¹¹¹, whereas GPCRs expressed in yeast have produced two structures, the adenosine A_{2A} receptor⁶¹ and H₁ histamine receptor¹¹². The first GPCR structure from recombinant sources utilised mammalian cells⁸⁴ (Table 2.1). However the most widely used system for the overexpression of GPCRs is recombinant baculovirus-infected insect cells. This system has produced recombinant GPCRs that have resulted in over 100 X-ray crystal structures. In a survey of 17 unique GPCR high resolution structures, 14 used the baculovirus expression system¹¹³. One of the highest expression levels reported for the baculovirus system was the production of 360 pmol of the β_1 AR receptor per mg of solubilised membrane protein¹¹⁴ which lead to a 2.7 Å resolution structure²⁵. Expression of AT₁R in insect cells from recombinant baculovirus was reported to produce approximately 30 pmol of receptor per mg of total membrane protein with an apparent K_D for angiotensin II similar to that of the native receptor¹¹⁵. These promising findings were a good starting point for the work described here, but the published results showed insufficient characterisation to fully understand the behaviour of AT₁R. For example, the authors did not establish whether unfolded protein was present. Shukla *et al.*¹¹⁵ also expressed AT₁R in BHK cells using the Semliki Forest virus expression system to produce 32 pmol of receptor per mg of total membrane protein. These results however should be treated with caution since this system often produces protein which is misfolded and retained in the ER¹¹⁰. Further attempts at high level expression of AT₁R using mammalian cells have been limited to amounts sufficient for pharmacological study. AT₁R has been expressed by transient transfection in COS-7 cells¹¹⁶ and in CHO cells by both transient transfection and the construction of a

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

stable cell line which expressed AT₁R constitutively from a human cytomegalovirus (CMV) promoter¹¹⁷. The CHO cell line that stably expressed AT₁R produced 12,000 molecules of AT₁R per cell or 0.8 µg per litre, assuming 1 million cells per ml. Strachan *et al.*⁹⁷ created a stable HEK293 cell line which also expressed AT₁R constitutively from a CMV promoter. It produced 7.1 pmol of receptor per mg of membranes. While the receptor maintained its pharmacological characteristics in these expression systems, none produced amounts sufficient for structural studies.

AT₁R possesses three highly conserved N-linked glycosylation consensus sequences (Asn-X-Ser/Thr, where X is not Pro) in its extracellular segments. Asn4 is at the N-terminus of the receptor while Asn176 and Asn188 are in extracellular loop two (ECL2; Appendix 1). A consideration for structural studies is whether to produce protein without post translational modification or to remove this source of heterogeneity during the purification process. It is therefore important to understand the role of N-linked glycosylation in terms of protein expression and integrity. Lanctôt *et al.*¹¹⁶ reported that an AT₁R mutant lacking all N-linked glycosylation sites showed five times lower expression than the wild type receptor while retaining the same affinity for the peptide antagonist Sar¹. The glycosylation-deficient mutant was also present at higher concentrations in the endoplasmic reticulum-golgi apparatus complex than in the plasma membrane. Jayadev *et al.*¹¹⁸ reported similar findings with a decrease in functional expression for the glycosylation-deficient AT₁R mutant. In addition they found that ligand-binding affinities were unaffected. This suggests that glycosylation of AT₁R is required for efficient cell surface expression. Baculovirus-driven expression in insect cells is known to be inefficient at producing protein with full N-linked glycosylation^{57,58}, whereas mammalian cell overexpression systems have been shown to be effective at producing complex glycosylated membrane proteins for structural studies⁶⁰. Recombinant membrane protein produced using the mammalian system has resulted in nine membrane protein structures: bovine rhodopsin^{38, 84, 119}, the human ammonia transporter RhCG¹²⁰, the human GABA_A receptor β3⁷¹, the *Drosophila melanogaster* dopamine transporter¹²¹, the human serotonin 5-HT₃ receptor¹²², the *Xenopus laevis* NMDA receptor¹²³ and the rat AMPA receptor¹²⁴. Most of these used the combination of a strong CMV promoter with a tetracycline-inducible system¹²⁵ (Table 2.1). This system has an advantage over constitutive expression systems when expressing proteins which might be toxic to the cell. After optimisation the inducible mammalian expression system produced 10 mg of opsin, the apo-form of rhodopsin, per litre of culture, assuming 10 million cells per ml of culture¹²⁶. The inducible mammalian system also produced 1.0 mg/L (assuming 1.2 million

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

cells per ml of culture) of the rat (*Rattus norvegicus*) neurotensin receptor (NTS1) and cell surface expression of NTS1 was nearly threefold higher in mammalian than insect cells¹²⁷. Together these results indicate that the tetracycline-inducible mammalian system under the control of a strong promoter is capable of producing sufficient authentically folded recombinant membrane protein for structural studies.

Both the mammalian and baculovirus systems were worth investigating for the overexpression of AT₁R. The baculovirus system is more established for the overexpression of GPCRs; however the near-native mammalian system appears to be better at expressing functional proteins with complete post translational modifications (Table 2.2).

Table 2.1 X-ray diffraction structures of mammalian membrane proteins determined from protein overexpressed in mammalian cells

Membrane protein	Source	Cells used for protein production †	Resolution of structure	Reference
Rhodopsin N2C/D282C	Bovine	COS-7	3.4 Å	84
Ammonia transporter RhCG	Human	HEK293S(TetR)-GnTI ⁻	2.1 Å	120
Rhodopsin N2C/D282C/D113Q	Bovine	HEK293S(TetR)-GnTI ⁻	3.0 Å	38
Rhodopsin N2C/D282C/M257Y	Bovine	HEK293S(TetR)-GnTI ⁻	3.3 Å	119
Dopamine transporter	<i>Drosophila</i>	HEK293S-GnTI ⁻	3.0 Å	121
GABA _A receptor β3	Human	HEK293S-GnTI ⁻	3.0 Å	71
Serotonin 5-HT ₃ receptor	Mouse	HEK293S(TetR)	3.5 Å	122
NMDA receptor	<i>Xenopus laevis</i>	HEK293S-GnTI ⁻	3.7 Å	123
AMPA receptor	Rat	HEK293S-GnTI ⁻	3.5 Å	124

† TetR indicates the use of the tetracycline inducible system. Adapted from¹¹⁰.

Table 2.2 Comparison of baculovirus and mammalian expression systems for the production of recombinant membrane proteins

	Expression in recombinant baculovirus-infected insect cells	Inducible mammalian expression
System maturity	Established	Emerging
Expression levels	Generally high	Variable
Quick investigation	None	4 days (via transient transfections)
Initiation period	1 month (via virus generation)	2-3 months (via stable cell lines)
Large-scale production	A few days	A few weeks
Advantages	<ul style="list-style-type: none"> • Membranes contain cholesterol • Some post-translational modification <ul style="list-style-type: none"> ○ Including limited glycosylation 	<ul style="list-style-type: none"> • Membranes contain cholesterol • Full post-translational modification <ul style="list-style-type: none"> ○ Including full glycosylation

2.2 Materials and Methods

2.2.1 Materials

The radiolabeled agonist [¹²⁵I]-angiotensin II and antagonist [¹²⁵I]-Sar¹-Ile⁸-angiotensin II (Sar¹) were purchased from Perkin Elmer. Unlabelled Sar¹ was purchased from Source Biosciences. Unlabelled angiotensin II was purchased from Tocris Biosciences. The detergents n-dodecyl β-D-maltopyranoside (DDM), n-decyl β-D-maltopyranoside (DM), n-nonyl β-D-glucopyranoside (NG), n-octyl β-D-glucopyranoside (OG) and lauryl maltose neopentyl glycol (LMNG) were purchased from Anatrace. Digitonin was purchased from Calbiochem. Cholesteryl hemisuccinate (CHS), sodium dodecyl sulphate (SDS) and sodium butyrate were purchased from Sigma. A tetracycline-inducible HEK293 cell line, T-REx™-293, (iHEK) was purchased from Invitrogen. A tetracycline-inducible HEK293S cell line lacking N-acetylglucosaminyltransferase I (iGnTI⁻) was kindly provided by Philip J. Reeves (Massachusetts Institute of Technology) ^{126, 128}. Anti-penta-histidine antibody conjugated to horseradish peroxidase (anti-pentaHis-HRP) was purchased from Qiagen. Peptide-N-glycosidase F (PNGase F), endoglycosidase H (Endo H), DpnI and all other restriction enzymes were purchased from New England Biolabs.

2.2.2 Mammalian and baculovirus expression construct design

Expression of AT₁R in mammalian cells was performed using derivatives of pcDNA4/TO (Invitrogen). The serotonin transporter (SERT) cDNA was inserted using EcoRV/NotI, followed by a cassette in the NotI/ApaI sites encoding enhanced green fluorescent protein (GFP) ^{129, 130} and a decahistidine (H₁₀) tag (plasmid pJMA111, kindly provided by Juni Andréll, MRC Laboratory of Molecular Biology). The cDNA clone for human AT₁R was obtained from the Missouri S&T cDNA Resource Center (www.cdna.org), amplified by polymerase chain reaction, flanked with EcoRV and NotI sites and cloned into the corresponding sites of pJMA111 to create plasmid pJAP2, which expressed AT₁R-GFP-H₁₀. For generating baculoviruses, AT₁R was cloned into the BamHI/EcoRI sites of the transfer vector pBacPAK8 (Cloneteck), creating plasmid pJAP15. Additionally, AT₁R was cloned into the BamHI/EcoRI sites in plasmid pAcGP67-B (BD Biosciences) in order to utilise the acidic glycoprotein gp67 signal sequence (LS) preceding the N-terminus of AT₁R, creating plasmid pJAP16, which expressed AT₁R-LS-H₁₀. All baculovirus sequences were engineered to contain a C-terminal tobacco etch virus (TEV) cleavage site and H₁₀ tag. All constructs were verified by DNA sequencing at Source Biosciences, UK.

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

2.2.3 Mutagenesis of AT₁R

Mutants were generated by PCR using pJAP2 as a template and the QuikChange II methodology (Stratagene); however KOD hot start polymerase (Novagen) was substituted for PfuUltra High-Fidelity DNA Polymerase. DpnI-digested PCR mixes were transformed into competent *E. coli* strain DH5 α , colonies were picked and grown in 5 ml Luria Bertani media containing 100 μ g/ml ampicillin with shaking overnight at 37°C. DNA was extracted using a Qiagen Miniprep kit. All constructs were verified by DNA sequencing at Source Biosciences, UK (Table 2.3).

Table 2.3 AT₁R N-linked glycosylation mutants for mammalian expression

Putative N-linked glycosylation sites (Asn-X-Ser/Thr, where X is not Pro) were identified by aligning AT₁R with other GPCRs of known structure to determine the position of the extracellular regions (Appendix 1). Alanine was substituted for asparagine in order to remove the glycosylation sites.

Plasmid name	Inserts and mutation(s)	Vector backbone
pJAP2	AT ₁ R-GFP-H ₁₀	pJMA111 (pcDNA4/TO)
pJAP6	AT ₁ R(N4A)-GFP-H ₁₀	pJAP2
pJAP7	AT ₁ R(N176A)-GFP-H ₁₀	pJAP2
pJAP8	AT ₁ R(N188A)-GFP-H ₁₀	pJAP2
pJAP11	AT ₁ R(N4A+N176A+N188A)-GFP-H ₁₀	pJAP2
pJAP12	AT ₁ R(N176A+N188A)-GFP-H ₁₀	pJAP2
pJAP13	AT ₁ R(N4A+N176A)-GFP-H ₁₀	pJAP2
pJAP14	AT ₁ R(N4A+N188A)-GFP-H ₁₀	pJAP2

2.2.4 Transient transfection, generation of stable cell lines and protein expression

Mammalian expression plasmids for the expression of AT₁R and mutants were amplified in *E. coli* strain DH5 α , purified using a Maxi-prep kit (Qiagen) and transiently transfected (GeneJuice, Novagen) into adherent mammalian iHEK cells or iGnTI⁻ cells following the manufacturer's protocol. Cells were grown in Dulbecco's modified Eagle's media supplemented with 10% tetracycline-free foetal bovine serum (Invitrogen) and 5 μ g/ml Blastidicin (Invitrogen) and incubated at 37°C in an atmosphere containing 5% CO₂. Cells were passaged every 3-4 days using standard tissue culture techniques. Expression of plasmids was induced by addition of 1 μ g/ml tetracycline and incubation at 37°C for 24 h. Stable cell lines were generated by selection with media containing 200 μ g/ml Zeocin (Invitrogen). After expression cells were washed twice in phosphate buffered saline (PBS), counted using the Countess Automated Cell Counter (Invitrogen), pelleted (1,200 \times g for 5 minutes) and resuspended at 10 million cells per ml in ice cold cell buffer (50 mM Tris pH

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

7.4, 150 mM NaCl supplemented with Complete EDTA-Free Protease Inhibitor Cocktail (Roche)). Cell suspensions were flash frozen in liquid nitrogen and stored at -80°C.

2.2.5 Baculovirus generation, protein expression and cell harvest

Recombinant baculoviruses that expressed AT₁R-H₁₀ and AT₁R-leader sequence (LS)-H₁₀ were made using the BaculoGold Baculovirus Expression System according to manufacturer's protocol (BD Bioscience). Viruses were isolated by plaque purification¹³¹ and screened for expression by western blotting using an anti-pentaHis-HRP antibody. Recombinant baculoviruses were passaged two times in Sf9 cells to obtain high titre stocks. Viruses were used to infect either Sf9, Sf21 or Hi5 cells as indicated on the figure for either 48 hours, or the time indicated for expression optimisation assays. After protein expression, cells were counted using the Countess Automated Cell Counter (Invitrogen), pelleted (1,200 × g for 5 minutes), washed twice in PBS, and the cell pellet was resuspended at 10 million cells per ml in ice cold cell buffer (50 mM Tris pH 7.4, 150 mM NaCl supplemented with Complete EDTA-Free Protease Inhibitor Cocktail (Roche)). Cell suspensions were flash frozen in liquid nitrogen and stored at -80°C.

2.2.6 Western blotting and deglycosylation

Cell suspensions were sonicated briefly and the total protein concentration determined using the Bradford method¹³². Samples were solubilised in the detergent DDM at 1% (w/v) final concentration at 4°C for 1 hour. Sodium dodecyl sulphate (SDS)-loading buffer was added to the supernatant (corresponding to approximately 150,000 cells), and samples were separated on a 4-20% tris glycine gel (Invitrogen) and transferred to nitrocellulose using standard techniques. Membranes were probed with anti-pentaHis-HRP at a dilution of 1:1,000 and developed using enhanced chemiluminescence (GE Healthcare). Where indicated, 2 µl of PNGase F or Endo H was added to 15 µl of the supernatant and incubated at 37°C for 1 hour prior to SDS-PAGE to remove N-linked glycosylation.

2.2.7 In-gel fluorescence

The protocol for western blotting was followed except the gel was visualised with the Typhoon TRIO variable mode imager (GE Healthcare) using standard fluorescein amidite settings to detect GFP fluorescence.

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

2.2.8 Laser scanning confocal microscopy

For live cell imaging, cells were grown on 35 mm glass bottom culture dishes under standard conditions and visualised on an inverted LSM 710 Laser Scanning Microscope (Carl Zeiss Ltd, UK) with 63× oil-immersion objective and a 1.4 numerical aperture. The laser was set to an excitation wavelength of 488 nm for GFP visualisation.

2.2.9 Flow cytometry

To perform fluorescence-activated cell sorting (FACS) analysis, cells were first washed with PBS and then harvested in PBS and analysed on a FACSCalibur II (Becton Dickinson). To examine GFP fluorescence the laser was set to an excitation wavelength of 488 nm and 10,000 events were counted for each measurement using the FL-1 detector (530 nm). Results were evaluated with the programme FlowJo. To exclude dead cells from the population, cells were stained with propidium iodide (PI) (1 µg/ml final concentration) and analysed as for GFP fluorescence, however the FL-2 detector (585 nm) was used. Cells were gated based on PI uptake, size (based on forward scatter) and optical homogeneity (based on side scatter) to only select whole live cells (G1) (Figure 2.2). Analysis was performed on the G1 population. G1 cell counts were normalised with the largest cell count set as 100%; Y-axes of normalised graphs are accordingly labelled “% of Max”. GFP expression above background was established by measuring the maximum fluorescence intensity of non-transfected parental cells and calculating the percent of the G1 population that was above this point (cells expressing GFP (%)).

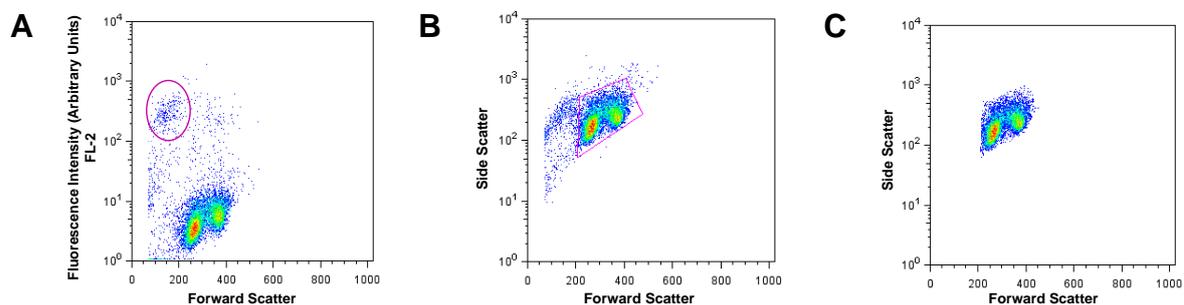


Figure 2.2 FACS gating based on PI staining, cell size and optical homogeneity

Colour density plots of iHEK cells stained with PI. **(A)** In order to exclude dead cells from the analysis, cells were stained with PI and fluorescence was detected using the FL-2 detector. The most intensely fluorescent cells which took up PI (pink ellipse) were predominantly below 200 on the forward scatter scale. Based on this, cell counts of below 200 on the forward scatter scale as well as the highest counts on the forward and side scatter scales were excluded from the population and only live, single cells were included in the analysis (pink quadrilateral) **(B)** resulting in population G1 **(C)**.

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

2.2.10 Thermostability assay of detergent-solubilised AT₁R

The cell suspension containing unpurified AT₁R was sonicated briefly and diluted into buffer (50 mM Tris pH 7.4, 5 mM MgCl₂, 1 mM EDTA, 0.1% (w/v) bovine serum albumin (BSA), 150 mM NaCl, 40 µg/ml bacitracin). For agonist binding [¹²⁵I]-angiotensin II and unlabelled angiotensin II were added to give final concentrations of 0.5 nM and 100 nM, respectively. For antagonist binding [¹²⁵I]-Sar¹ and unlabelled Sar¹ were added to give final concentrations of 0.5 nM and 100 nM, respectively. AT₁R present in the membrane was incubated for 1 hour at room temperature with either agonist or antagonist before chilling on ice and solubilising in 1% (w/v, final concentration) DDM, DM, NG, OG or 1% DDM with 0.1% CHS as indicated for the [+] assay format. The samples were then heated at varying temperatures for 30 minutes and the [¹²⁵I]-Sar¹-bound receptor was separated from the free radioligand by gel filtration spin columns as described previously^{67, 68, 69}. For the [-] assay format, the cell suspension was solubilised in 1% (w/v, final concentration) DDM, LMNG, digitonin or 1% DDM with 0.1% CHS as indicated for 1 hour on ice before heating and adding ligand for 1 hour on ice. Background was determined by adding radioligand to non-transfected parental mammalian cells or uninfected insect cells. Results were evaluated by nonlinear regression using GraphPad Prism.

2.2.11 Detergent-solubilised radioligand binding assays

Cell suspensions were sonicated briefly and the total protein concentration determined using the Bradford method¹³². Cells were diluted into buffer (150 mM NaCl, 50 mM Tris pH 7.4) and incubated with [¹²⁵I]-Sar¹ at the same concentration as for the T_m assay (1 h, 4°C), and then solubilised in a final concentration of 1% (w/v) DDM for 1 hour at 4°C. Bound and free radioligand were separated on gel filtration spin columns as per section 2.2.10.

2.2.12 Fluorescence-detection size exclusion chromatography and size exclusion chromatography detected by western blotting

The void volume (8.16 ml) of the Superdex 200 10/300 (24 ml) (GE healthcare) was determined by running blue dextran through the column and observing where it eluted using A₂₈₀. For fluorescence-detection size exclusion chromatography (FSEC), approximately 5 million iHEK(AT₁R-GFP-H₁₀) cells were thawed on ice and sonicated briefly. Cells were incubated at room temperature for 1 hour with 40 nM Sar¹ before chilling on ice and solubilising in 1% DDM (w/v, final concentration) followed by centrifugation at 280,000 ×

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

g for 30 minutes at 4°C. The supernatant was then passed through a 0.22 µm filter and injected onto a Superdex 200 10/300 column pre-equilibrated with running buffer (0.03% (w/v) DDM, 50 mM Tris pH 7.4, 150 mM NaCl and 1µM Sar¹). The fluorescence of eluent was detected by a Hitachi fluorometer (mV) set to an excitation of 488 nm and emission of 525 nm. To detect AT₁R-H₁₀ produced in Sf9 cells, approximately 5 million cells were sonicated, incubated with ligand, solubilised and centrifuged as above. The eluent was detected by western blotting (Section 2.2.6) and bands corresponding to AT₁R-H₁₀ were quantified using ImageJ.

2.3 Results

2.3.1 Apparent thermostability of AT₁R in detergent

The apparent T_m of a protein is good indicator of the likelihood of the protein to form crystals in short chain detergents by vapour diffusion. For this reason, the apparent T_m was the first parameter considered in assessing the suitability of AT₁R as a candidate for structural studies. To measure the apparent T_m of AT₁R, it was expressed from the plasmid pJAP2 (expressing AT₁R-GFP-H₁₀) in the tetracycline-inducible cell line HEK293 (iHEK). Detergent-solubilised AT₁R-GFP-H₁₀ was assayed for its thermostability with the peptide agonist [¹²⁵I]-angiotensin II and with the peptide antagonist [¹²⁵I]-Sar¹. Additionally, two different assay methods were used to test thermostability. In the [Super +] condition, the agonist angiotensin II or antagonist Sar¹ was added to AT₁R present in the membrane and allowed to bind for 1 hour before the receptor-ligand complex was solubilised in detergent; after heating for 30 minutes, the amount of ligand-bound receptor was determined. In the [-] condition, AT₁R was first solubilised in detergent, then heated in the absence of ligand for 30 minutes; radioligand was then added, allowed to bind, and the amount of ligand-bound receptor was determined.

The thermostability of AT₁R was first tested in the [-] format using [¹²⁵I]-Sar¹, but no binding was detected after solubilising with either the detergent DDM alone or DDM with CHS (Figure 2.3 and Table 2.4). Consequently a sigmoidal dose response curve could not be fitted for these conditions. Although the detergent digitonin appeared to give the highest apparent T_m for the ligand-free condition, only a small proportion of AT₁R was solubilised in a functional form and therefore digitonin was not a good choice of detergent. LMNG was the only detergent condition that gave a reliable apparent T_m of $25.3^\circ\text{C} \pm 0.3$ (n=3) and with a disintegrations per minute (dpm) count that was well above background in the [-] format. This suggests that AT₁R is unstable in the absence of ligand and therefore subsequent assays were performed in the [Super +] format. Angiotensin II-bound AT₁R had a low apparent T_m of $23.4^\circ\text{C} \pm 1.5$ (n=3) in the mild detergent DDM and also the amount of functional receptor was very low. In contrast Sar¹-bound AT₁R was shown to be more thermostable with an apparent T_m of $46.5^\circ\text{C} \pm 0.3$ (n=4) in DDM and $31.8^\circ\text{C} \pm 0.3$ (n=4) in the harsher detergent OG. In both detergents there was a similar amount of functional receptor with the antagonist bound. Unlike many receptors, CHS in combination with DDM gave a lower apparent T_m of $26.3^\circ\text{C} \pm 0.5$ (n=3) than DDM alone, suggesting that CHS destabilised the receptor. Two

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

conclusions were drawn from the thermostability data. First, the large discrepancy in apparent T_m between the [Super +] and [-] format assays suggested that binding of the peptide antagonist Sar¹ significantly enhances the stability of AT₁R. Secondly, the apparent T_m of OG solubilised AT₁R bound to the peptide antagonist Sar¹ was remarkably high for a GPCR and it was plausible that it could be purified and crystallised under these conditions, or crystallised in lipidic cubic phase (LCP). Therefore the next step was to determine the best expression system and the first step in this process was to determine the impact of N-linked glycosylation on AT₁R expression.

2.3.2 The role of N-linked glycosylation on AT₁R expression and stability

Since antagonist-bound AT₁R was found to be thermostable with an apparent T_m of 46.5°C in DDM the next step was to find a suitable expression system to produce enough recombinant protein for structural studies. AT₁R has three putative N-linked glycosylation sites at Asn4, Asn176 and Asn188 (Appendix 1). N-linked glycosylation can be required for the expression and stability of a protein and therefore this can potentially limit the choice of expression system. Additionally, N-linked glycosylation causes heterogeneity in a protein, which reduces the likelihood of crystallisation, and therefore the ideal construct for the overexpression of AT₁R would eliminate this. Before removal of N-linked glycosylation on AT₁R could be achieved, the impact of N-linked glycosylation on trafficking to the cell surface, levels of expression and stability needed to be determined. To facilitate this investigation, three mutations were introduced to the putative N-linked glycosylation sites of AT₁R-GFP-H₁₀ (Table 2.3). The effects of these mutations were assessed by confocal microscopy, in-gel fluorescence, digestion with glycosidases and radioactive ligand binding assays. To ascertain whether AT₁R was a glycoprotein when expressed in mammalian cells, two different enzymes, Endo H and PNGase F were tested for their ability to remove N-glycans. AT₁R-GFP-H₁₀ was expressed by transient transfection in two mammalian cell lines, iHEK and a derivative cell line, which lacks N-acetylglucosaminyltransferase I activity (GnTI⁻) and expresses homogeneous N-glycans comprised of only seven sugars (Man₅GlcNAc₂)^{126, 128}. As shown in the gel in Figure 2.4, two forms of N-linked glycosylated AT₁R exist in the HEK cells, AT₁R modified with high mannose N-glycans (Endo H sensitive) and AT₁R modified with complex N-glycans (Endo H resistant and PNGase F sensitive). The iGnTI⁻ cells produced only the high mannose form of AT₁R. Both cell lines produced similar amounts of AT₁R polypeptide and unglycosylated AT₁R was not observed.

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

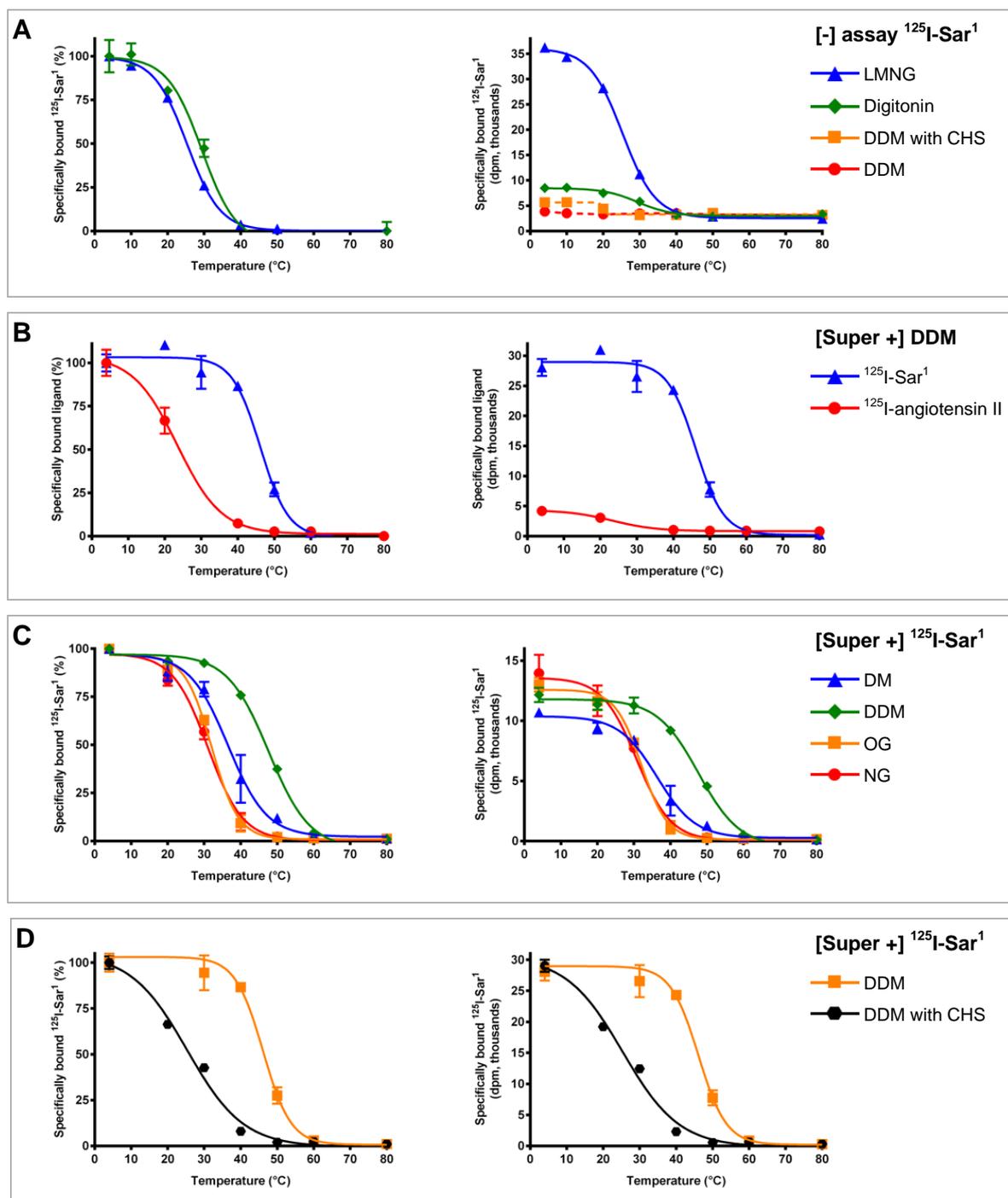


Figure 2.3 Thermostability of AT₁R in detergent

AT₁R was expressed from plasmid pJAP2 in the mammalian cell line iHEK (expression of AT₁R-GFP-H₁₀) and induced for 24 hours with 1 µg/ml tetracycline prior to being assayed. Graphs on the left side had the amount of specifically bound radioligand normalised to 100% at 4°C and 0% at 80°C to allow for comparison of the slopes of the curves. Graphs on the right are in dpm to emphasise differences in the total amount of functional detergent-solubilised AT₁R. **(A)** Stability curve of Sar¹-bound AT₁R in the [-] condition with several detergents; LMNG (blue triangles), digitonin (green diamonds), DDM with CHS (yellow squares) and DDM (red circles). Dotted lines indicate that a sigmoidal dose response curve could not be accurately fitted. The results are from a single experiment performed in triplicate and plotted as a mean value ± SEM. **(B)** The stability of AT₁R was measured in the [Super +] condition with the antagonist Sar¹ (blue triangles) or agonist angiotensin II (red circles) bound. The results are from a single experiment performed in triplicate and plotted as a mean value ± SEM. **(C)** Stability curve of Sar¹-bound AT₁R in the [Super +] condition with several detergents; DM (blue triangles), DDM (green diamonds), OG (yellow squares) and NG (red circles). Results are from two independent experiments performed in duplicate and plotted as a mean value ± SEM. **(D)** Stability curve of Sar¹-bound AT₁R in the [Super +] condition; DDM (yellow squares) and DDM with CHS (black hexagons). The results are from a single experiment performed in triplicate and plotted as a mean value ± SEM. Results for all figures are summarised in Table 2.4.

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

Table 2.4 Apparent thermostability of AT₁R under different assay conditions

Ligand	Assay format	Detergent (w/v)	T_m (°C)	SEM
Sar ¹	[-]	1% LMNG	25.3	0.3
Sar ¹	[-]	1% digitonin	29.5 [†]	1.3
Sar ¹	[-]	1% DDM, 0.1% CHS	Unable to fit curve	—
Sar ¹	[-]	1% DDM	Unable to fit curve	—
Angiotensin II	[Super +]	1% DDM	23.4	1.5
Sar ¹	[Super +]	1% DDM	46.5	0.3
Sar ¹	[Super +]	1% DM	36.4	1.0
Sar ¹	[Super +]	1% NG	30.6	0.5
Sar ¹	[Super +]	1% OG	31.8	0.3
Sar ¹	[Super +]	1% DDM, 0.1% CHS	26.3	0.5

[†] Digitonin has the highest apparent T_m for the [-] condition, however it had significantly lower amounts of functional detergent-solubilised AT₁R in comparison to LMNG.

These data indicate that AT₁R is indeed heavily glycosylated and therefore attempts were made to reduce this by mutating the consensus N-glycosylation sequence from Asn-X-Ser/Thr to Ala-X-Ser/Thr. In-gel fluorescence of an SDS-PAGE gel loaded with an equal amount of cells in each lane showed that the single mutations N176A and N188A appeared to give reasonable expression of AT₁R but the N4A mutations significantly decreased expression levels. No AT₁R was detected when all three N-linked glycosylation sites were mutated. Further experiments were therefore performed where combinations of two N-glycosylation sites were mutated and analysed by confocal microscopy and ligand binding assays. The confocal microscopy data showed that any single deletion of an N-linked glycosylation site did not significantly impair trafficking to the cell surface (Figure 2.5). In contrast, of all the possible combinations of double mutants tested, only N4A+N188A and N4A+N176A gave significant amounts of cell surface expression, whereas the N176+N188A AT₁R mutant was expressed predominantly within intracellular membranes. The triple mutant was also expressed mainly intracellularly. Transfected cells expressing either wild-type AT₁R or one of the N-glycosylation mutants grew well, with the exceptions of the N176+N188A mutant and the triple mutant, both of which produced round detached cells in the confocal images. The double mutants N4A+N188A and N176A+N188A as well as the triple mutant had fewer cells positive for GFP expression in comparison to the wild type receptor and the other glycosylation mutants. It is difficult to ascertain expression levels from confocal images, so therefore the amount of DDM-solubilised AT₁R mutants was determined using the [Super +] format assay described in section 2.2.11.

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

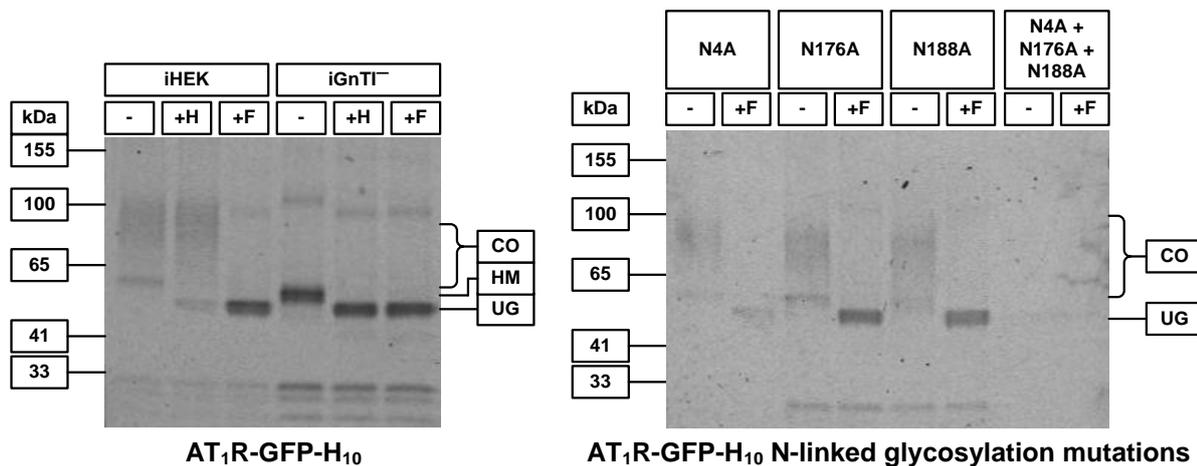


Figure 2.4 In-gel fluorescence analysis of N-linked glycosylation site mutations in AT₁R

Plasmid pJAP2 (expressing AT₁R-GFP-H₁₀) and N-glycosylation mutants of AT₁R (Table 2.3) were transiently transfected into iHEK or iGnTI⁻ cells and induced for 24 hours with 1 µg/ml tetracycline for 24 hours prior to being assayed by SDS-PAGE and visualised by in-gel fluorescence. Removal of N-linked glycosylation was achieved by treatment of cell samples with either Endo H (+H), PNGase F (+F) or no enzyme (-). The different glycosylated forms of AT₁R are indicated; complex N-glycans (CO), high mannose core N-glycans (HM) and unglycosylated (UG). An equal amount of cells was loaded in each lane.

The radioligand binding assays showed that only three combinations of mutations gave expression levels similar to AT₁R with all three N-glycosylation sites intact (Figure 2.6). Mutations of either Asn176 or Asn188 did not significantly decrease expression levels. In contrast N4A gave a seven-fold reduction in expression, but surprisingly, combinations of the mutations N4A and N188A gave a 70% increase in expression levels. All other combinations of mutations reduced expression levels relative to the fully N-glycosylated receptor with the lowest expression levels observed for the double mutant N4A+N176A and the triple mutant. Another important consideration is the stability of each of these mutants. The stability was tested for all of the mutants in a two-point thermostability assay (Figure 2.6) and it appeared that removal of the N-glycosylation sites did not greatly affect stability for the N188A mutant or the double mutant N4A+N188A. The mutant N176A+N188A showed a slight decrease in stability, whereas the measurement for the mutant N176A had a large error and should be treated with caution. The lack of results for the other mutants containing the N4A mutation suggests that they did not express well in a functional form.

In conclusion, these data suggest that the only mutant that may be worth considering for further structural studies is the N4A+N188A mutant, but this still indicates that expression of an N-glycosylated version of AT₁R needed to be explored. This meant pursuing expression in either the baculovirus expression system or in mammalian cells.

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

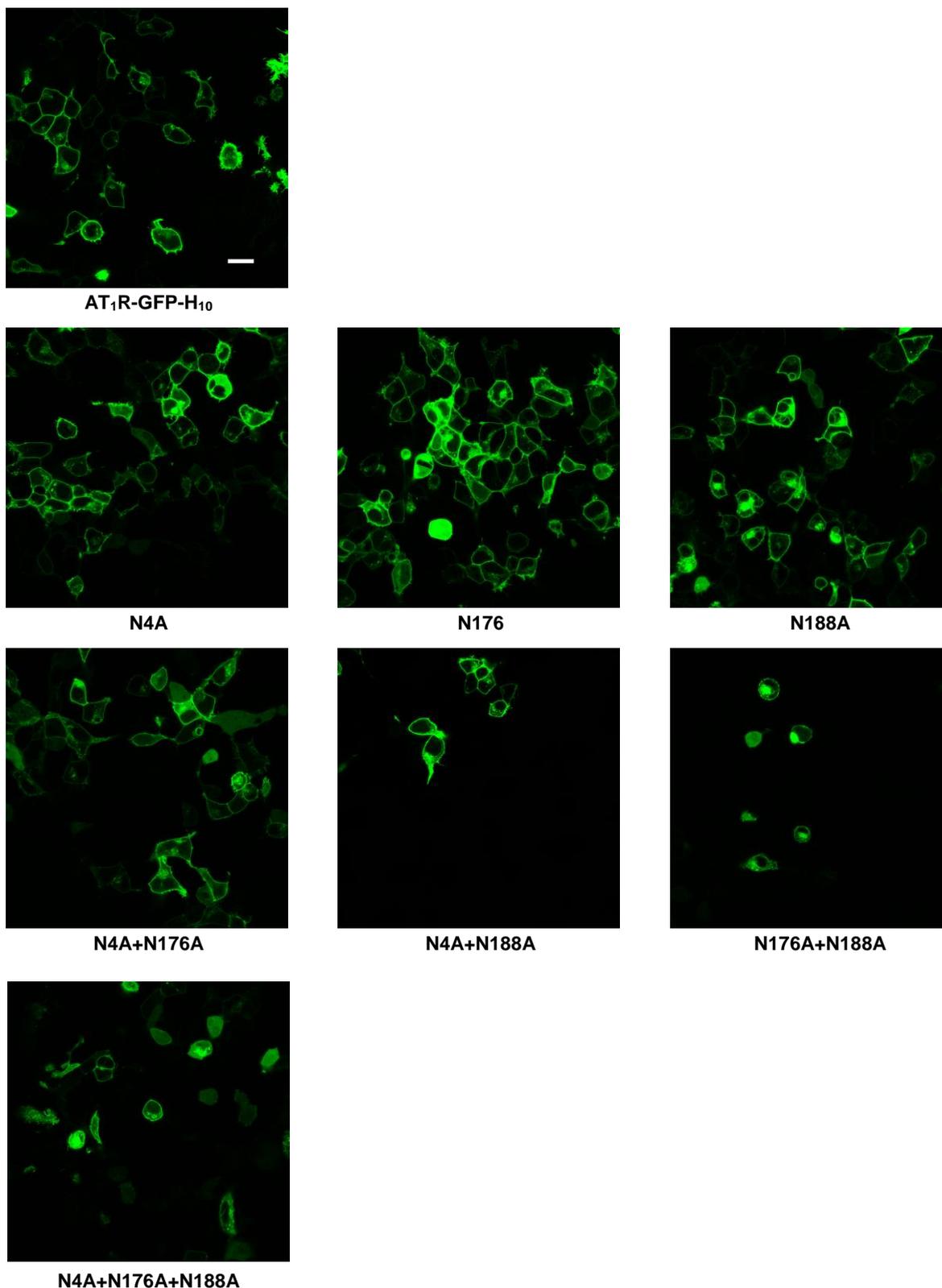


Figure 2.5 Confocal microscopy analysis of N-linked glycosylation site mutations in AT₁R

Plasmid pJAP2 (expressing AT₁R-GFP-H₁₀) and N-glycosylation mutants of AT₁R (Table 2.3) were transiently transfected into iHEK cells and induced for 24 hours with 1 μ g/ml tetracycline for 24 hours prior to being assayed. Confocal micrographs of iHEK cells expressing mutated forms of AT₁R. Mutations of AT₁R are indicated below each image. The gain and offset of the laser were set to be equal for all images. Scale bar represents 20 μ m. Non-transfected parental cells showed no fluorescence (data not shown).

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

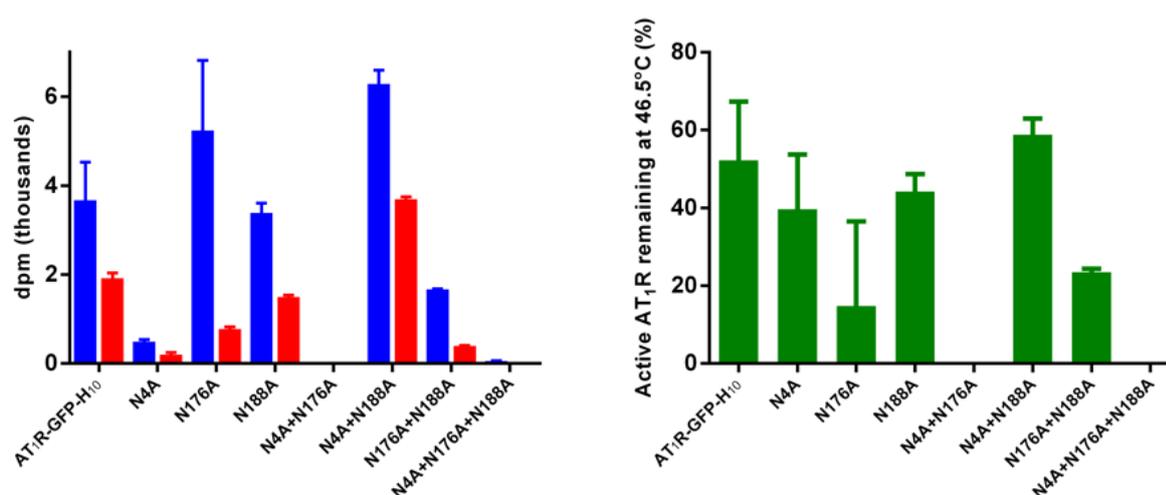


Figure 2.6 Ligand binding analysis of N-linked glycosylation site mutations in AT₁R

Plasmid pJAP2 (expressing AT₁R-GFP-H₁₀) and N-glycosylation mutants of AT₁R (Table 2.3) were transiently transfected into iHEK cells and induced for 24 hours with 1 μ g/ml tetracycline for 24 hours prior to being assayed. Antagonist [¹²⁵I]-Sar¹ binding was carried out on AT₁R N-glycosylation mutants at 4°C (blue bars) and 46.5°C (red bars). From these data, the percent of active AT₁R remaining at 46.5°C was calculated (green bars). The results are from a single experiment performed in triplicate and plotted as a mean value \pm SEM.

2.3.3 Baculovirus-mediated expression of AT₁R in insect cells

From the data presented here the extent of the role of N-linked glycosylation in determining the thermostability of AT₁R is not clear, however it does affect expression levels. Given this, the choice of expression systems for the production of recombinant AT₁R focused on two systems, baculovirus mediated expression in insect cells and mammalian expression using the iHEK cell line. Baculovirus mediated expression of recombinant GPCRs in insect cells has been utilised many times for the production of protein for crystallography, which has resulted in over a hundred GPCR structures. Given the success of this system, recombinant baculovirus expressing AT₁R from the polyhedrin promoter were constructed.

Two baculoviruses that expressed AT₁R were created using the Baculo Gold method, referred to as bvAT₁R-H₁₀ and bvAT₁R-LS-H₁₀. The difference between the two AT₁R constructs was that the leader sequence for acidic glycoprotein 67 from *Autographa californica* was inserted at the N-terminus of AT₁R to make AT₁R-LS-H₁₀, since leader sequences have been shown to improve expression levels⁵⁸. To optimise expression three cell lines were tested (Sf9, Sf21 and Hi5) and cell samples were analysed at 24 hour intervals over 96 hours post infection. The presence of bvAT₁R-H₁₀ or bvAT₁R-LS-H₁₀ was detected by western blotting (Figure 2.7). Samples were deglycosylated using the enzyme Endo H and the amounts of glycosylated and unglycosylated AT₁R were quantified using ImageJ (Figure 2.8). All conditions tested from 48 hours onwards showed the presence of a

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

large amount of unglycosylated AT₁R, but only small amounts were observed at 24 hours post infection for bvAT₁R-LS-H₁₀ expressed in Sf9 cells. The expression levels of AT₁R produced from either bvAT₁R-H₁₀ or bv-AT₁R-LS-H₁₀ in the three insect cell lines (Sf9, Sf21 and Hi5) were quantified by using the [Super +] assay format and measuring binding of the antagonist [¹²⁵I]-Sar¹ to DDM-solubilised AT₁R (Figure 2.9). The best condition for expressing AT₁R in insect cells was Hi5 cells infected with bvAT₁R-LS-H₁₀ for 48 hours. This produced approximately 1.8 million molecules of AT₁R per cell (n=6), which, assuming 1 million cells per ml of culture, equates to 0.1 mg of AT₁R per litre. Further analyses are presented in Section 2.3.5 where baculovirus expression is compared to mammalian expression. Given that glycosylation was shown to be important for cell surface expression, mammalian expression was also investigated since this system is able to achieve near native post translational modification.

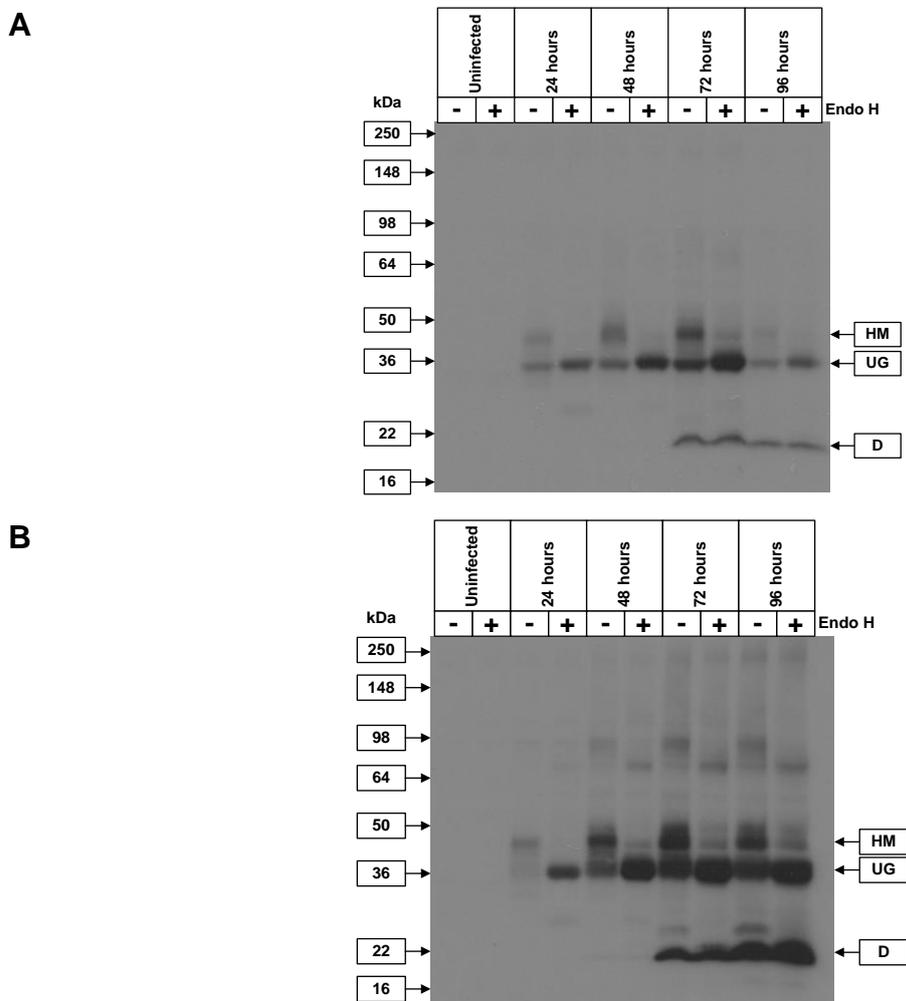


Figure 2.7 Optimisation of bvAT₁R-H₁₀ and bvAT₁R-LS-H₁₀ expressed in Sf9 cells

Western blot of proteins from whole Sf9 cells expressing **(A)** bvAT₁R-H₁₀ or **(B)** bvAT₁R-LS-H₁₀. The blot was probed with an anti-pentaHis-HRP conjugated antibody. N-linked glycosylation was removed using Endo H where indicated (+). The high mannose form of AT₁R (HM), unglycosylated (UG) and putative degradation products (D) are indicated. Bands corresponding to HM and UG forms of AT₁R were quantified using Image J (Figure 2.8). This assay was repeated for each cell line (Sf9, Sf21 and Hi5) and for each virus (bvAT₁R-H₁₀ and bvAT₁R-LS-H₁₀); Appendix 2.

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

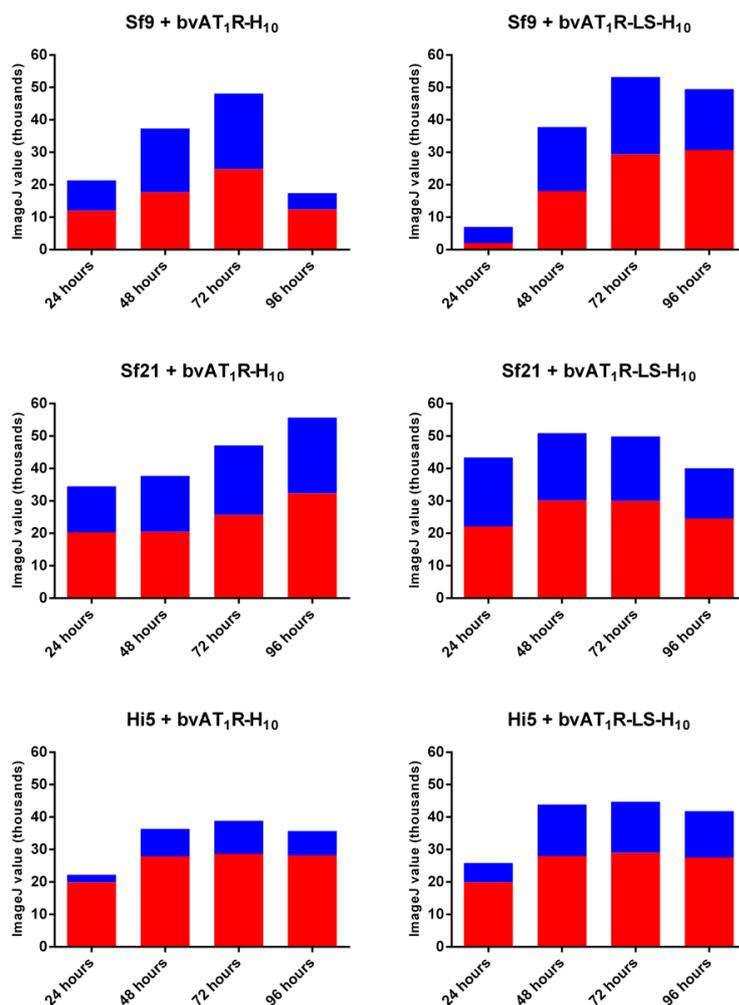


Figure 2.8 Glycosylation of AT₁R across four insect cell lines

AT₁R was expressed from recombinant baculoviruses bvAT₁R-H₁₀ or bvAT₁R-LS-H₁₀ in three insect cell lines (Sf9, Sf21 and Hi5) over 96 hours and analysed by western blotting. Bands on the western blot were quantified using ImageJ and correspond to either the high manose, glycosylated form of AT₁R (blue bars) or unglycosylated AT₁R (red bars). Western blot data are shown in Figure 2.7 and Appendix 2.

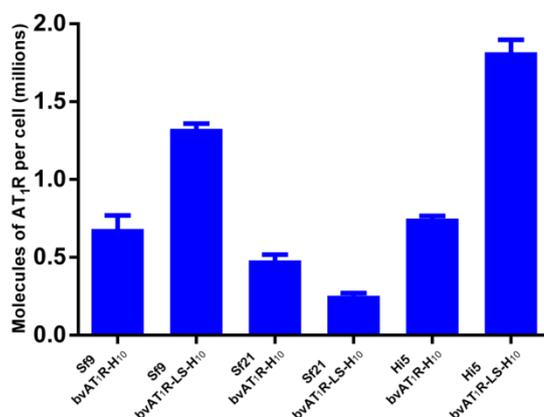


Figure 2.9 Expression of AT₁R in three insect cell lines

AT₁R was expressed from either the virus bvAT₁R-H₁₀ or bvAT₁R-LS-H₁₀ and used to infect the insect cell line indicated (Sf9, Sf21 or Hi5). Cells were harvested 48 hours post infection. The amount of functional AT₁R in each insect cell line was determined by measuring specific binding of the antagonist [¹²⁵I]-Sar¹. After the addition of ligand, membranes were solubilised in DDM and non-bound ligand was separated from receptor-ligand complex on gel filtration spin columns and measured by liquid scintillation counting. Each data point was determined in triplicate from two independent experiments and was plotted as mean ± SEM.

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

2.3.4 Mammalian expression of AT₁R

Given the low levels of N-linked glycosylation of AT₁R observed in insect cells, mammalian cell expression systems were investigated further, because they produce mostly proteins with full post translational modification. The mammalian expression plasmid pJAP2, based on pcDNA4/TO, was created to express AT₁R-GFP-H₁₀ under the control of a tetracycline-inducible CMV promoter. Transient transfection of plasmid pJAP2 into iHEK cells and induction with tetracycline gave low levels of expression of AT₁R-GFP-H₁₀, therefore a polyclonal cell line stably expressing AT₁R was created. 24 hours post transfection, the cells were split into six populations containing 50 thousand to 500 thousand cells per well in a 6-well plate. Cells with a stable incorporation of plasmid pJAP2 were selected by adding Zeocin to the culture media, grown for 4-6 weeks until confluent in the 6-well plate, then passaged twice before a subset was induced with tetracycline and screened for GFP expression by FACS analysis (Section 2.2.4). The most highly expressing polyclonal cell line, iHEK(AT₁R-GFP-H₁₀), showed a dramatic increase in expression of AT₁R-GFP-H₁₀ in comparison to transient transfection of plasmid pJAP2 and it produced high levels of AT₁R-GFP-H₁₀ even after the cells were passaged for 64 days (Figure 2.10). Confocal microscopy indicated that AT₁R-GFP-H₁₀ in the polyclonal cell line iHEK(AT₁R-GFP-H₁₀) was expressed homogenously on the cell surface in comparison to transient transfection of plasmid pJAP2 (Figure 2.10). A 96-hour expression test of iHEK(AT₁R-GFP-H₁₀) showed that the majority of AT₁R was glycosylated (Figure 2.11) unlike the results seen for insect cells (Figure 2.8). Attempts to increase expression of AT₁R in the stable cell line iHEK(AT₁R-GFP-H₁₀) by lengthening the induction time and the addition of agonist and antagonist ligands were not successful (Figure 2.12). However adding 5 mM sodium butyrate to the culture at the time of induction had a dramatic impact on expression levels as measured by FACS (Figure 2.12). This led to an increase in the elution of AT₁R-GFP-H₁₀ from a size exclusion column, indicating that the protein remained correctly folded. However further increasing the amount of sodium butyrate in culture caused cell death, therefore 5 mM was selected to achieve a balance between increased expression and the health of cells.

The majority of AT₁R produced in the mammalian system was glycosylated. Expression in the mammalian system could be increased through creation of a stable cell line as well as including sodium butyrate in the media at the time of induction. A direct comparison of the amounts of functional AT₁R between insect and mammalian systems was therefore required.

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

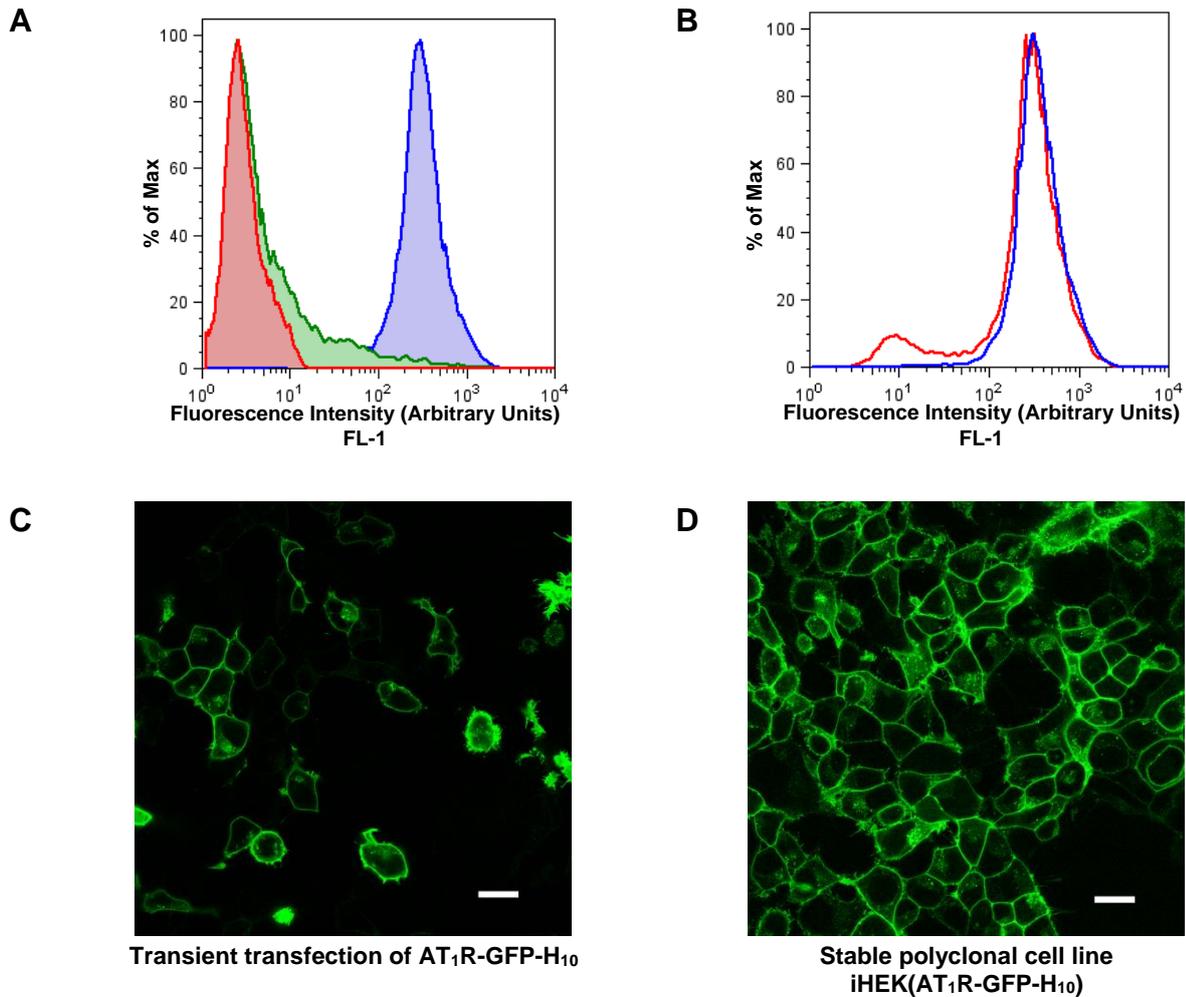


Figure 2.10 Increased expression of AT₁R-GFP-H₁₀ in a stable cell line compared to transient transfection

A stable cell line that expressed AT₁R-GFP-H₁₀ was created by selection for the integration of plasmid pJAP2 into the genome of iHEK cells with Zeocin. For FACS analysis cells were harvested in PBS and analysed on the FACSCalibur II for GFP fluorescence using the FL-1 detector. Cell counts have been normalised (% of Max).

(A) FACS analysis of iHEK cells. Non-transfected iHEK parental cells (red, median fluorescence 1.73), transient transfection of plasmid pJAP2 (expressing AT₁R-GFP-H₁₀) into iHEK cells and induced for 24 hours (green, median fluorescence 2.53) and stable polyclonal cell line iHEK(AT₁R-GFP-H₁₀) induced for 24 hours (blue, median fluorescence 267). **(B)** The expression levels of AT₁R-GFP-H₁₀ were compared in cells grown continuously for 64 days (blue line, median fluorescence 319) and freshly thawed cells (red line, median fluorescence 250). Both cell lines were induced for 24 hours with tetracycline before FACS analysis. **(C)** Confocal micrograph of AT₁R-GFP-H₁₀ expressed by transient transfection of plasmid pJAP2 into HEK cells and **(D)** the stable cell line iHEK(AT₁R-GFP-H₁₀).

The laser was not set to be equal for both images. The scale bar represents 20 μm. Cells were induced for 24 hours with tetracycline prior to imaging.

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

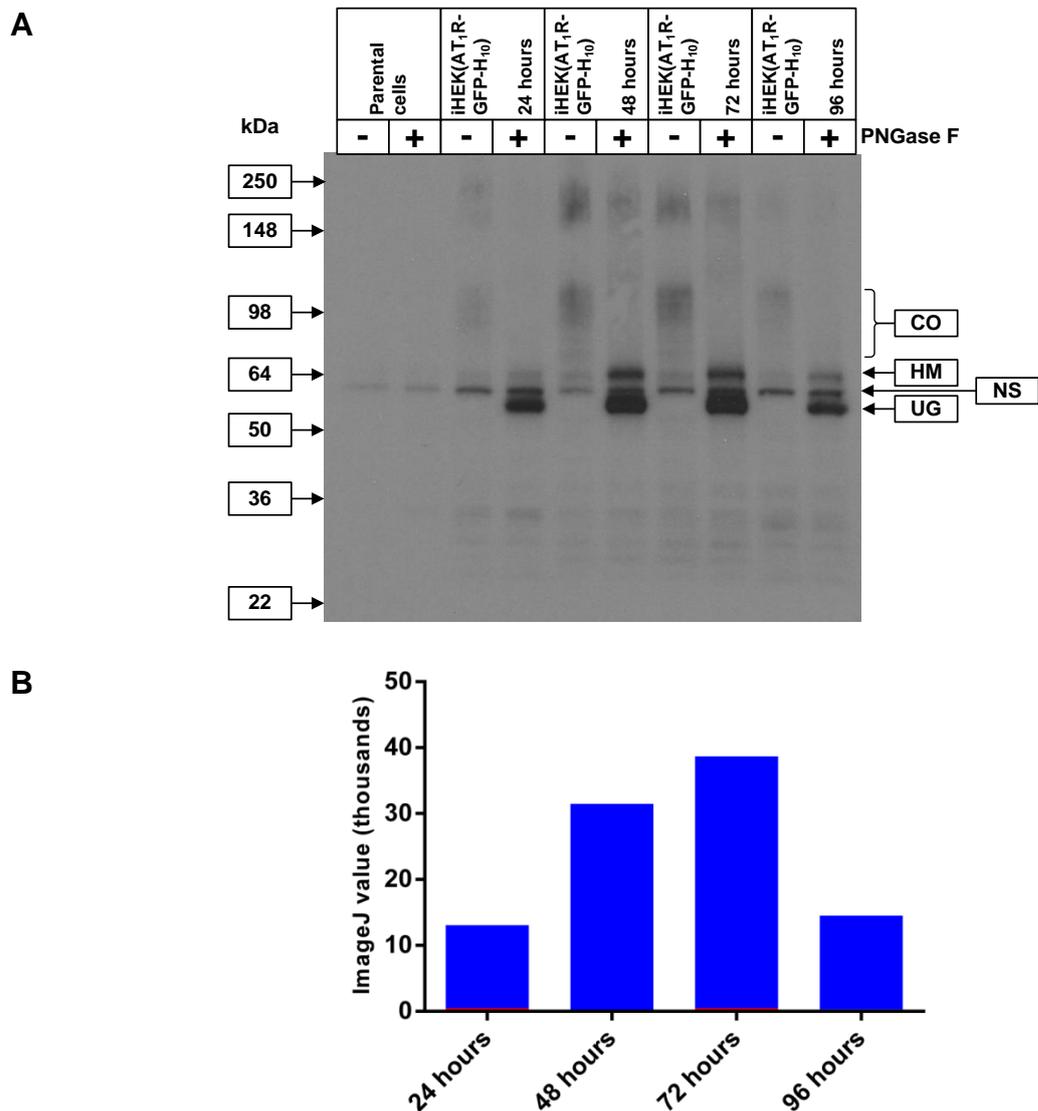


Figure 2.11 N-linked glycosylation of AT₁R in a stable mammalian cell line

(A) AT₁R was expressed from the stable cell line iHEK(AT₁R-GFP-H₁₀) over 96 hours and analysed by western blotting. The blot was probed with an anti-pentaHis-HRP conjugated antibody. The different glycosylated forms of AT₁R are indicated; complex N-glycans (CO), high mannose core N-glycans (HM) and unglycosylated (UG). Removal of N-linked glycosylation was achieved by treatment of cell samples with PNGase F (+F). A nonspecific band (NS) is indicated. An equal amount of cells was loaded in each lane. **(B)** Bands on the western blot from the expression assay corresponding to the complex glycosylated form of AT₁R (blue bars) and unglycosylated AT₁R (red bars, not visible for 48 and 96 hours) were quantified using ImageJ.

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

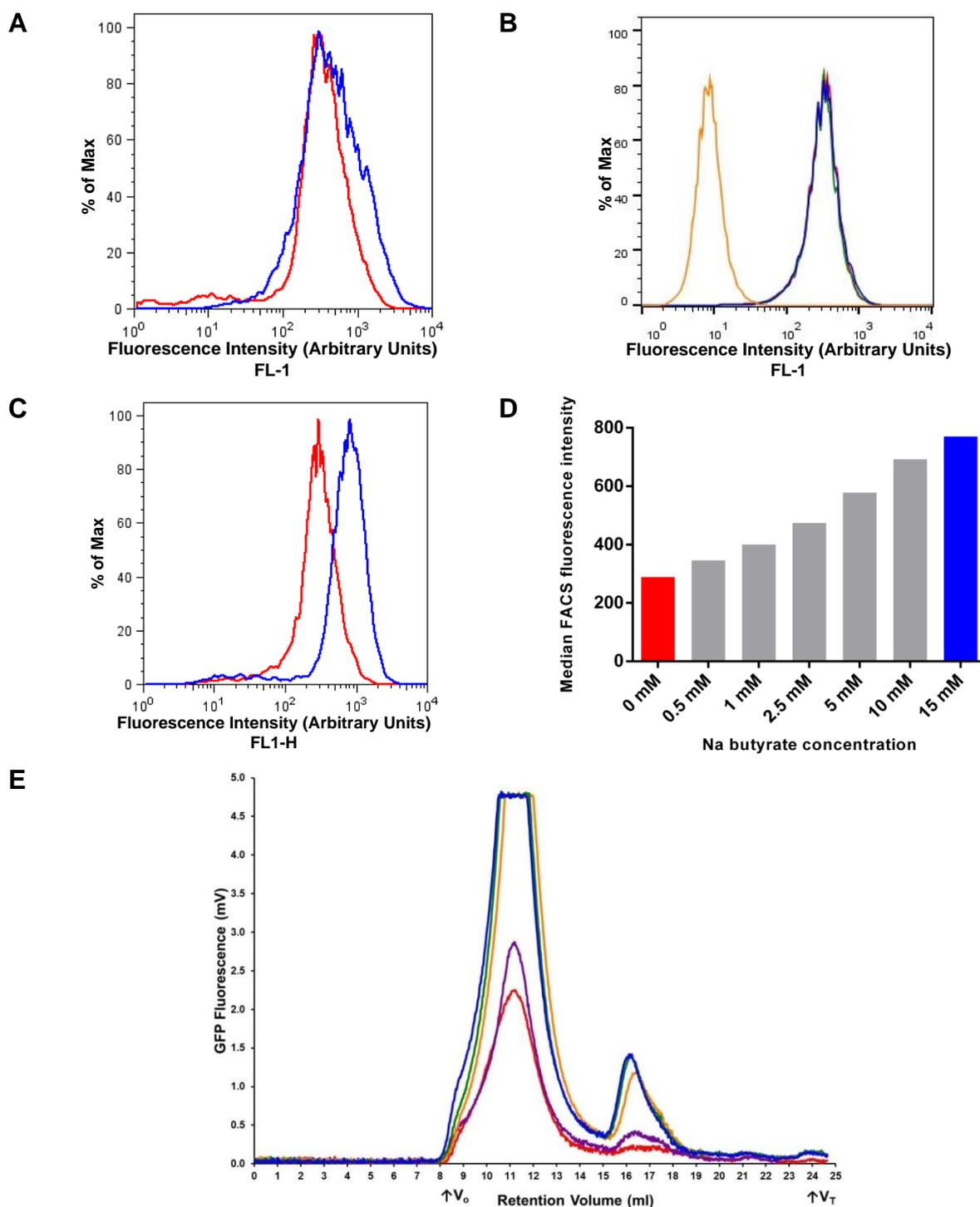


Figure 2.12 Optimisation of AT₁R expression from the stable polyclonal cell line iHEK(AT₁R-GFP-H₁₀)

(A) FACS analysis of GFP expression in the stable cell line iHEK(AT₁R-GFP-H₁₀) induced for 24 hours (red line, median fluorescence 250) and 48 hours (blue line, median fluorescence 319). **(B)** Induction of expression with ligands was tested; uninduced iHEK(AT₁R-GFP-H₁₀), (yellow line; median fluorescence 8); induction of iHEK(AT₁R-GFP-H₁₀) with tetracycline only, (blue line; median fluorescence 274); induction of iHEK(AT₁R-GFP-H₁₀) with tetracycline and 1 μ M Sar¹, (green line; median fluorescence 269); induction of iHEK(AT₁R-GFP-H₁₀) with tetracycline and 1 μ M angiotensin II, (red line; median fluorescence 271). **(C)** FACS analysis of GFP expression in the stable cell line iHEK(AT₁R-GFP-H₁₀) induced for 24 hours with either tetracycline (red line) or with tetracycline and 15 mM Na butyrate (blue line). **(D)** Median FACS fluorescence intensities are shown in the bar graph. **(E)** FSEC analysis of AT₁R-GFP-H₁₀. The cell line iHEK(AT₁R-GFP-H₁₀) was induced with tetracycline and the amount of Na butyrate indicated for 24 hours; tetracycline only (red), 1 mM Na butyrate (purple), 5 mM Na butyrate (yellow), 10 mM Na butyrate (green) and 15 mM Na butyrate (blue). 40 nM of Sar¹ was added to 5 million cells and allowed to bind for 1 hour at 23°C before being solubilised in 1% DDM. The elution of iHEK(AT₁R-GFP-H₁₀) was detected using GFP fluorescence (mV). The void (V_0) and total column volumes (V_T) are indicated.

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

2.3.5 Advantages of mammalian cell expression of AT₁R compared baculovirus-mediated expression of AT₁R in insect cells

In order to determine the best overexpression system for AT₁R, it was expressed in three different systems: baculovirus-mediated expression in insect cells (bvAT₁R-H₁₀) and expression in mammalian cells by either transient transfection of plasmid pJAP2 into iHEK cells or by the creation of the stable cell line iHEK(AT₁R-GFP-H₁₀). The apparent T_m of Sar¹-bound AT₁R solubilised in DDM produced from the three different systems was determined (Figure 2.13). The apparent T_m found in each case was similar within experimental error and the mean across the three systems was 46.6°C ± 0.6 (n=3). This suggested that the biophysical properties of AT₁R expressed in all three systems were the same.

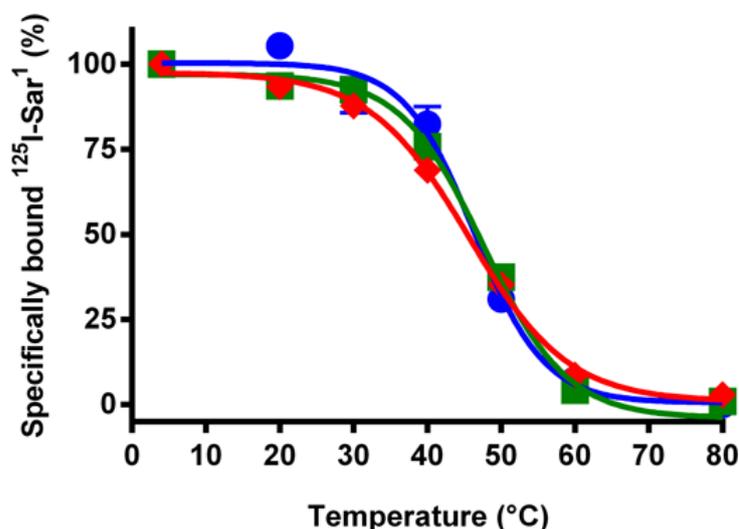


Figure 2.13 Stability of DDM-solubilised AT₁R expressed in three different systems bound to the antagonist [¹²⁵I]-Sar¹

AT₁R was expressed using three different expression systems: baculovirus bvAT₁R-H₁₀ in Sf9 cells (blue circles), transient transfection of plasmid pJAP2 into iHEK cells (expression of AT₁R-GFP-H₁₀) (green squares), stable clonal cell line iHEK(AT₁R-GFP-H₁₀) (red diamonds). The apparent T_m values of AT₁R expressed in each system are: Sf9 cells, 46.4°C ± 0.8 (n=3); iHEK transient transfection, 47.7°C ± 0.5 (n=3); iHEK(AT₁R-GFP-H₁₀), 45.7°C ± 0.7 (n=3). Each data point was determined in triplicate and was plotted as a mean value ± SEM.

Quantification of the functional expression levels of AT₁R produced in either mammalian or insect cells was performed using the [Super +] assay format and measuring binding of the antagonist [¹²⁵I]-Sar¹ to DDM-solubilised AT₁R. There was a fivefold increase in functional expression of AT₁R in the iHEK(AT₁R-GFP-H₁₀) cell line compared to the best expression in insect cells (Figure 2.14). The stable mammalian cell line produced 9.1 million molecules of AT₁R per cell (n=6), which assuming 1 million cells per ml of culture equates to 0.6 mg of AT₁R per litre. The best condition for expression in insect cells, Hi5 cells infected with bvAT₁R-H₁₀ for 48 hours, produced 1.8 million molecules of AT₁R per cell (n=6), which equates to 0.1 mg of AT₁R per litre.

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

Another important factor in choosing which expression system to use was the production of misfolded receptor which needs to be minimised. A western blot containing equal amounts of active material showed that there was a large amount of DDM-solubilised misfolded AT₁R produced by the baculovirus system (Figure 2.15). However, when AT₁R was solubilised and analysed by size exclusion chromatography (SEC), both the insect cell and mammalian cell systems produced AT₁R of a similar size and homogeneity indicating that most of the unfolded AT₁R in the insect system is removed during solubilisation, centrifugation and passage through the column pre-filter and the SEC resin (Figure 2.16).

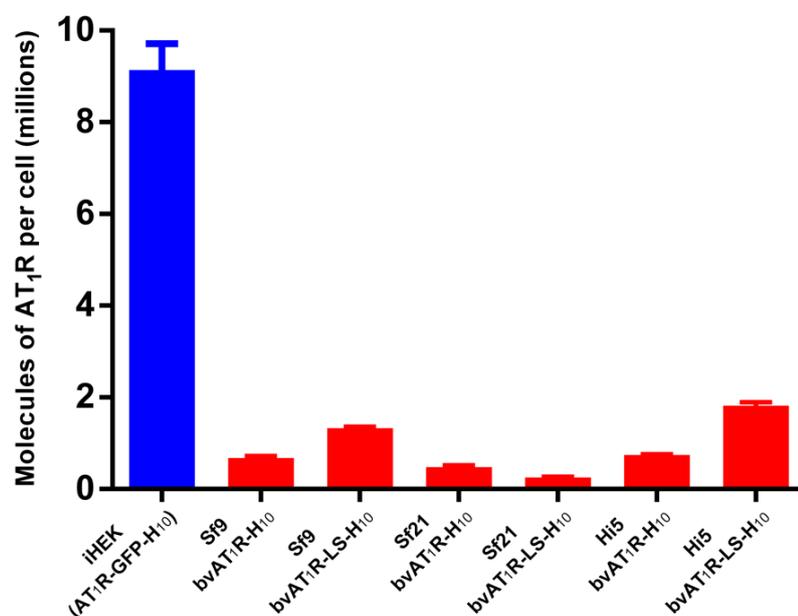


Figure 2.14 Five-fold higher functional expression of AT₁R in mammalian cells compared to insect cells

The amount of functional AT₁R in each expression system was determined by measuring specific binding of the antagonist [¹²⁵I]-Sar¹. Mammalian expression was from the polyclonal cell line iHEK(AT₁R-GFP-H₁₀) (blue bar). Baculoviral expression was performed either in Sf9, Sf21 or Hi5 cells (red bars). After the addition of ligand, membranes were solubilised in DDM and non-bound ligand was separated from receptor-ligand complex on gel filtration spin columns and measured by liquid scintillation counting. Each data point was determined in triplicate from two independent experiments and was plotted as mean ± SEM.

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

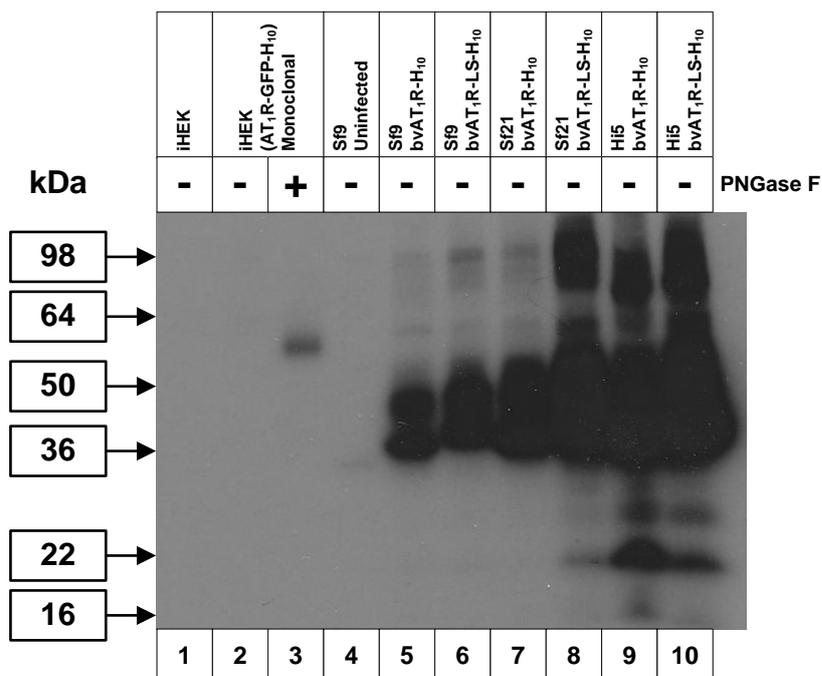


Figure 2.15 DDM solubilises considerable amounts of inactive AT₁R produced in the baculovirus expression system

Western blot of DDM-solubilised AT₁R, with equal amounts of active receptor per sample (lanes 2, 3, 5-10). The blot was probed with an anti-pentaHis-HRP conjugated antibody. Lane 1, iHEK parental cells; lanes 2 and 3, iHEK(AT₁R-GFP-H₁₀) stable clonal cell line (Chapter 3); lane 4, uninfected Sf9 cells; lanes 5-10, bvAT₁R-H₁₀ infected insect cells. N-linked glycosylation was removed using PNGase F where indicated (+). AT₁R was expressed either in the stable mammalian cell line iHEK(AT₁R-GFP-H₁₀) or by using the recombinant baculoviruses bvAT₁R-H₁₀ and bvAT₁R-LS-H₁₀ to infect Sf9, Sf21 and Hi5 cells as indicated. iHEK cell lines were induced with 1 µg/ml tetracycline for 24 hours and insect cells were infected with recombinant baculovirus for 48 hours. The amount of functional AT₁R was determined by measuring specific binding of the antagonist [¹²⁵I]-Sar¹ (Figure 2.14).

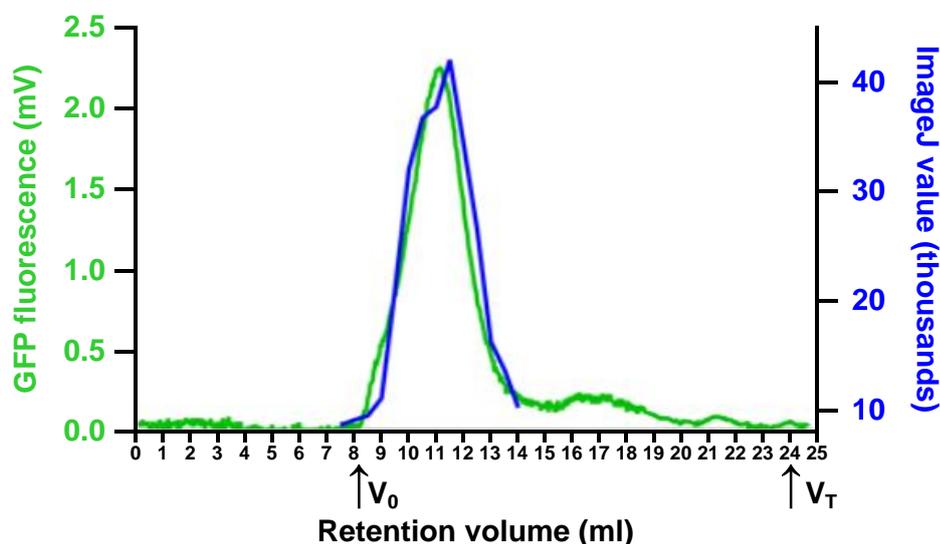


Figure 2.16 Size exclusion chromatography of AT₁R expressed in mammalian cells compared to insect cells

Size exclusion chromatography was carried out using a Superdex 200 10/300 (24 ml) column. The elution of iHEK(AT₁R-GFP-H₁₀) was detected using GFP fluorescence (mV). The elution of bvAT₁R-H₁₀ was detected by western blotting and band quantification (ImageJ value). See Appendix 2 for western blot. iHEK(AT₁R-GFP-H₁₀) shows a symmetrical peak whereas bvAT₁R-H₁₀ shows two peaks however both systems show elution of a protein of a similar size and homogeneity. The void (V_0) and total column volumes (V_T) are indicated.

2.4 Discussion

Stability is an important factor for indicating the ability of membrane proteins to form crystals^{133, 134, 135, 136}, however there is no specific level of stability above which crystallisation is guaranteed. Several methods have been devised to measure stability and it is not always easy to compare the data since the techniques measure stability in different ways. A prevalent method for measuring stability is the CPM assay which utilises a thiol specific dye¹⁰⁵. When measured using the CPM assay, the nociception opioid peptide receptor (NOP) had an apparent T_m of 57°C in DDM and 69°C in DDM with CHS¹⁰⁴. NOP was successfully crystallised in DDM with CHS resulting in a 3.0 Å structure, however apocytochrome b₅₆₂RIL (BRIL) was fused to the N-terminus¹³⁷, which might have raised the apparent T_m , but this was not recorded. Using the CPM assay on an A_{2A}-T4 lysozyme fusion protein, which was bound to each of five different ligands, demonstrated that only the ligand with the highest apparent T_m resulted in a crystal structure (ZM 241385; 62°C)¹⁶. Another method for measuring thermostability is the radioactive ligand-bound apparent T_m assay. Using this assay, partly purified A_{2A}R-GL26 solubilised in DM had an apparent T_m of 44.5°C⁶⁷. Removal of an N-linked glycosylation site produced A_{2A}R-GL31 which was crystallised in octylthioglucoside (T_m of 31°C) and resulted in a 3.0 Å resolution structure bound to adenosine and a 2.6 Å resolution structure bound to NECA⁴⁹. It is difficult to compare the stability of A_{2A} between both assays because it was measured in different detergents and utilised different techniques. However both methods have been successfully employed to evaluate conditions which produced high resolution structures.

For structure determination of AT₁R to proceed the receptor needs to be sufficiently thermostable. The radioactive ligand-bound apparent T_m assay was chosen to measure the stability of AT₁R since it has become a key parameter in GPCR structure determination and can be used on unpurified material. The apparent T_m of AT₁R was first measured in the [-] format assay, where the detergent-solubilised receptor is heated before ligand is added. The only detergent condition that gave a reliable apparent T_m in the [-] format was LMNG (25.3°C). The instability of AT₁R in the [-] format assay was indicated by the fact that no binding in DDM was observed coupled with the low apparent T_m in LMNG. Two other membrane proteins also shown to be unstable without ligand bound are a GPCR, A_{2A}R-GL26⁶⁷ and a transporter, rat (*Rattus norvegicus*) serotonin transporter (SERT)⁶⁵. Both of these membrane proteins had to be assayed using the [Super +] format apparent T_m assay, where ligand is added prior to solubilisation. Given that AT₁R was unstable in the [-] format

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

assay all further assays were performed using the [Super +] format. Using this format the apparent thermostability of antagonist-bound AT₁R was shown to be remarkably high in all detergents tested. In the mild detergent DDM the apparent T_m of AT₁R was 0.5°C higher than a thermostabilised version of DDM-solubilised SERT (SAH6)⁶⁵, which has resulted in crystals (Andréll, Edwards and Tate; unpublished). Additionally, the apparent T_m for A_{2A}R-GL26 solubilised in OG was 1.3°C less than the one obtained for AT₁R in the same detergent. What this high apparent T_m showed is that antagonist-bound AT₁R was a good candidate for structure determination by X-ray crystallography in its native form. Additionally, several GPCRs, which are presumed to have low stability, have been crystallised using alternative techniques. For example, β₂AR was crystallised with bacteriophage T4 lysozyme (T4L) inserted into cellular loop 3 (CL3) and produced a 2.4 Å resolution structure^{24, 41}. The β₂AR-T4L fusion was more resistant to proteases than the native receptor but its thermostability was not directly measured. If AT₁R does not crystallise in its native form then alternatively a fusion of T4L into CL3 of AT₁R might provide a path to crystallisation.

Glycosylation can impart flexibility and heterogeneity to a protein¹³⁸; therefore it was important to examine the role of N-linked glycosylation in the expression and stability of AT₁R and to establish whether the receptor could be expressed devoid of any glycosylation. Enzymatic deglycosylation of AT₁R directly demonstrated the attachment of N-linked glycans to the receptor produced from both insect and mammalian systems. The glycosylation deficient iGnTI⁻ cell line produced AT₁R with only the core Man₅GlcNAc₂ in comparison to the heterogeneous glycosylation seen in the iHEK cell line. The iGnTI⁻ cell line not only produced more homogeneous glycosylation, but it allowed different options during crystallisation. Complex glycosylation in iHEK cells must be removed with PNGase F which causes the Asn linked to the GlcNAc to be converted to Asp¹³⁸. In contrast the glycosylation produced in iGnTI⁻ cells can be removed with either PNGase F or Endo H, the latter leaving one GlcNAc residue covalently bound to the protein, which can remarkably improve stability. Deletion of two N-linked glycosylation sites did not dramatically change expression levels but deletion of all three sites significantly reduced expression. These results however should be treated with caution since the experiments were carried out using transient transfection and therefore can give less reproducible results than those obtained from stable cell lines. Also it is worth examining amino acid substitutions other than Ala in the N-linked glycosylation sequence, for example Asn to Asp or Ser/Thr to hydrophilic amino acids, to explore whether this maintains the expression levels seen for the

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

wild type receptor. Additionally, fusion of BRIL to the N-terminus of NOP negated the need for three N-linked glycosylation sites¹³⁷ and might be a worthwhile approach for AT₁R. What the results here suggested is that AT₁R needs to have at least one N-linked glycosylation site intact for expression of the correctly folded receptor. Although this is unusual for the majority of GPCR structures obtained thus far, it is similar to what was seen for bovine rhodopsin. Rhodopsin has two N-linked glycosylation sites at Asn2 and Asn15. Removal of the N-glycosylation site at Asn2 had no effect, but removal of Asn15 reduced signal transduction and cell surface expression. Deletion of both of the N-linked glycosylation sites also reduced cell surface expression¹³⁹. All recombinant sources of bovine rhodopsin used for crystal structures thus far are confined to mammalian cells¹¹³. These results are also similar to what has been shown for the mammalian transporter SERT. N-linked glycosylation has been established as a requirement for expression of correctly folded SERT or necessary for its stability since a SERT mutant lacking N-linked glycosylation showed 20-fold lower expression compared to the native transporter¹⁴⁰. Another example of a protein dependent on N-linked glycosylation for expression is the extracellular domain of the insulin receptor which was successfully expressed in CHO cells modified to produce limited N-linked glycosylation¹⁴¹. An engineered monomer of the ectodomain of the insulin receptor, also produced in the modified CHO cells, contained 15 N-linked glycans but could still be crystallised using two different F_{ab} fragments, allowing structure determination at 3.8 Å resolution¹⁴². The homopentameric GABA_A receptor β3 also had 15 N-linked glycosylation sites necessary for expression however these were cleaved prior to crystallisation with endoglycosidase F1 resulting in crystals that were used to determine a 3.0 Å resolution structure⁷¹. A two-point thermostability assay showed that removal of the N-glycosylation sites did not greatly affect stability of AT₁R, but even if removal does prove necessary, it has been shown that it is possible to crystallise GPCRs with glycosylation intact. Although this approach might be more difficult, the first rhodopsin crystal structure was purified from native sources and N-linked glycosylation was not removed¹⁹. Additionally, a rhodopsin mutant lacking one glycosylation site (Asn2) was expressed in the iGnTI⁻ system, which has restricted homogenous N-linked glycosylation, and this resulted in a structure where the core GlcNAc₂-Man₁ was ordered and formed crystal contacts¹¹⁹. These results open up the possibility of producing and crystallising AT₁R with N-linked glycosylation intact.

Since a minimum of one N-linked glycosylation site is necessary for high levels of expression of AT₁R, two overexpression systems were investigated thoroughly: expression

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

in insect cells by recombinant baculovirus and expression in mammalian cells. Also, the expression of functional eukaryotic membrane proteins has had significantly more success in eukaryotic rather than prokaryotic systems, possibly due to the slower rates of translation and folding seen in the latter system¹⁴³. Expression of GPCRs in insect cells has proven to be successful, resulting in over a hundred crystal structures, however in a study of 16 related human GPCRs expressed in insect cells, yields of correctly folded material varied dramatically⁵⁷. Attempts at expressing each of the five muscarinic receptor subtypes under the same conditions in insect cells resulted in yields of between 0.6 and 16 pmol/mg however the three different opioid receptor subtypes all expressed at 1-2 pmol/mg¹⁴⁴. These data indicated that there is no reliable way to predict the yields of a particular expression system. Given that trial and error is the current best approach to determining optimal expression conditions, expression of AT₁R in insect cells was optimised by examining different insect cell lines, different infection times and the inclusion of acidic glycoprotein 67 leader sequence. However the only condition explored that had a measurable effect on expression levels was the inclusion of the leader sequence. The best insect cell expression condition was Hi5 cells infected with bvAT₁R-LS-H₁₀ for 48 hours (1.8 million molecules of AT₁R per cell). Assuming 1 million cells per ml of culture this condition produced approximately 0.1 mg of AT₁R per litre. The best condition in the absence of a leader sequence was also Hi5 cells infected with for 48 hours. This produced 0.05 mg of AT₁R per litre. In addition to the low expression levels, the insect cell system also produced a large amount of unfolded AT₁R peptide which would likely hinder crystallisation. This result was similar to what had been previously shown for rat SERT when expressed in insect cells⁶⁰.

In contrast to the baculovirus system, nearly all of the AT₁R produced in the mammalian system was glycosylated and localised at the cell surface. A tetracycline inducible system was chosen for expression of AT₁R in mammalian cells since it has previously been shown that overexpression of some GPCRs is toxic to cells. For example, a constitutively active mutant of rhodopsin (E113Q+E134Q+M257Y) could not be successfully expressed in a constitutive mammalian system¹²⁶, but the tetracycline-inducible system¹²⁵ was successful for production of this constitutively active mutant of rhodopsin¹²⁸ as well as for the large scale production of NTS1¹²⁷ and SERT¹⁴⁵. Expression levels of AT₁R from transient transfection were highly variable and too low to measure accurately. Based on FACS analysis of 10 thousand cells, generation of a cell line stably expressing AT₁R-GFP-H₁₀ dramatically increased expression in comparison to transient transfection. The stable cell line, iHEK(AT₁R-GFP-H₁₀), also showed less variability in expression levels and was more

CHAPTER 2 The Suitability of AT₁R as a Candidate for Structural Studies

uniformly expressed on the cell surface in comparison to transient transfection. The cell line iHEK(AT₁R-GFP-H₁₀) produced 0.6 mg per litre of receptor when induced for 24 hours with tetracycline. Several methods were investigated to increase AT₁R expression in the stable mammalian cell line including: lengthening the expression time; adding agonists and antagonists at the time of induction; and induction with sodium butyrate and tetracycline. The only condition that increased expression levels was the inclusion of sodium butyrate at the time of induction. Sodium butyrate has also been shown to increase GPCR expression in iHEK cell lines stably expressing rhodopsin^{126, 128} and the neurotensin receptor¹²⁷.

In conclusion, the T_m assay data suggested that AT₁R bound to the high affinity peptide antagonist Sar¹ is thermostable enough to consider purification and crystallisation in the wild type form. The mammalian system proved to be superior to the baculovirus system because of the higher levels of active material present in the former and the presence of unfolded material in the latter.

CHAPTER 3 IMPROVING THE EXPRESSION OF AT₁R IN MAMMALIAN CELLS

3.1 Introduction

AT₁R presents a challenge to overexpress and crystallise because it is an integral membrane protein that contains three N-linked glycosylation sites in regions predicted to be on the extracellular surface of the cell. More than 50% of proteins encoded by the human genome are predicted to be glycosylated¹⁴⁶, however glycosylated proteins only represent around 10% of all structures in the Protein Data Bank (PDB)¹⁴⁷. A common hindrance to the determination of glycoprotein structures is their requirements for expression of correctly folded material. In particular, they often need complex folding machinery and post-translational modification, which are only found in mammalian cells¹³⁸. Human membrane proteins are frequently glycosylated and therefore they benefit from being expressed in mammalian systems. In addition, the near native lipid composition of mammalian cells is also advantageous for their overexpression^{144, 148}. It has become routine to use mammalian cells for functional studies, but protein obtained from mammalian hosts accounts for only ~3% of unique structures in the PDB¹⁴⁹, although this figure has grown by 40% in the last two years¹³⁸. Despite their advantages, there has been a reluctance to use mammalian cells to overexpress membrane proteins, which is probably due to the expense of the system, the length of time involved in generating large volumes of cells and the low yields frequently obtained^{144, 148, 150}.

If low expression levels of the target protein are obtained in mammalian cells, there are only a few published methods for increasing expression levels further. It might be reasonable to assume that increasing the stability of a protein could increase expression levels, however this was not the case for the expression of a thermostabilised version of SERT in HEK293 cells⁶⁵, despite this expression system providing the highest amounts of functional protein out of seven expression systems examined⁶⁰. Chaudhary *et al.*¹⁴⁸ described a method for transiently transfecting multiple constructs tagged with GFP into HEK293 GnTI⁻ cells, screening the solubilised protein for expression levels with western blotting and further analysing the recombinant protein with FSEC. While this method allows for a rapid assessment of the amounts of protein produced, expression levels are highly variable in transient transfection and often too low to be accurately measured by FSEC. Chaudhary *et*

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

al. further expanded upon this method as applied to deriving cell lines stably expressing the protein of interest. Since integration of the transfected plasmid into the genome of the host was a random event, the authors found that screening up to 24 polyclonal cell lines resulted in increased expression, presumably by finding a cell line which had the plasmid integrated into an area of high expression or which had been integrated more than once. While the use of stable cell lines might address the variability seen in transient transfection, this method is time consuming and relatively low-throughput. Another method developed to increase expression levels in mammalian cells used fluorescence activated cell sorting (FACS)¹⁵¹ to select highly expressing polyclonal cell lines¹⁵⁰. This method used a constitutive CMV promoter to co-express the protein of interest and GFP, which served as a marker for expression, with the GFP downstream from an internal ribosome entry site (IRES) to allow independent translation. After incorporation of the bicistronic plasmid into the host's genome, the resulting stable cell line was subjected to FACS and those cells that formed the top 0.1% of GFP fluorescence were selected (Figure 3.1). This process was repeated until sufficient expression levels were achieved and was used successfully on the rat serotonin receptor subtype 2c (5HT_{2c}) to produce approximately 3 million molecules per cell (2.5 mg of receptor from 1-5 litres of culture)¹⁵⁰. The main disadvantage of this technique was that FACS was required every time cells were grown on a large scale in order for protein production to maintain high expression levels. A further way to increase expression of membrane proteins in mammalian cells is the use of sodium butyrate in the culture media at the time of induction. Sodium butyrate is thought to inhibit histone deacetylation, thus increasing acetylation levels of histones, which in turn leads to less compact DNA packing that allows better access to the transcription machinery. Examples of the success of this approach include increased expression of rhodopsin^{126, 128}, the neurotensin receptor¹²⁷ and the human Rh C glycoprotein¹²⁰. Finally, protein engineering can be used to increase expression of the protein of interest. This will be discussed in more detail in Chapter 4.

Recently, the FACS-based method of selecting highly expressing clonal cell lines¹⁵⁰ was refined in our lab for use with the T-RExTM tetracycline-inducible expression system (Andréll & Tate, unpublished). A polyclonal cell line was produced that stably expressed the protein of interest as a GFP-fusion under the control of the inducible CMV promoter. GFP served as a marker to select and assess highly expressing clonal cell lines by FACS (Figure 3.1).

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

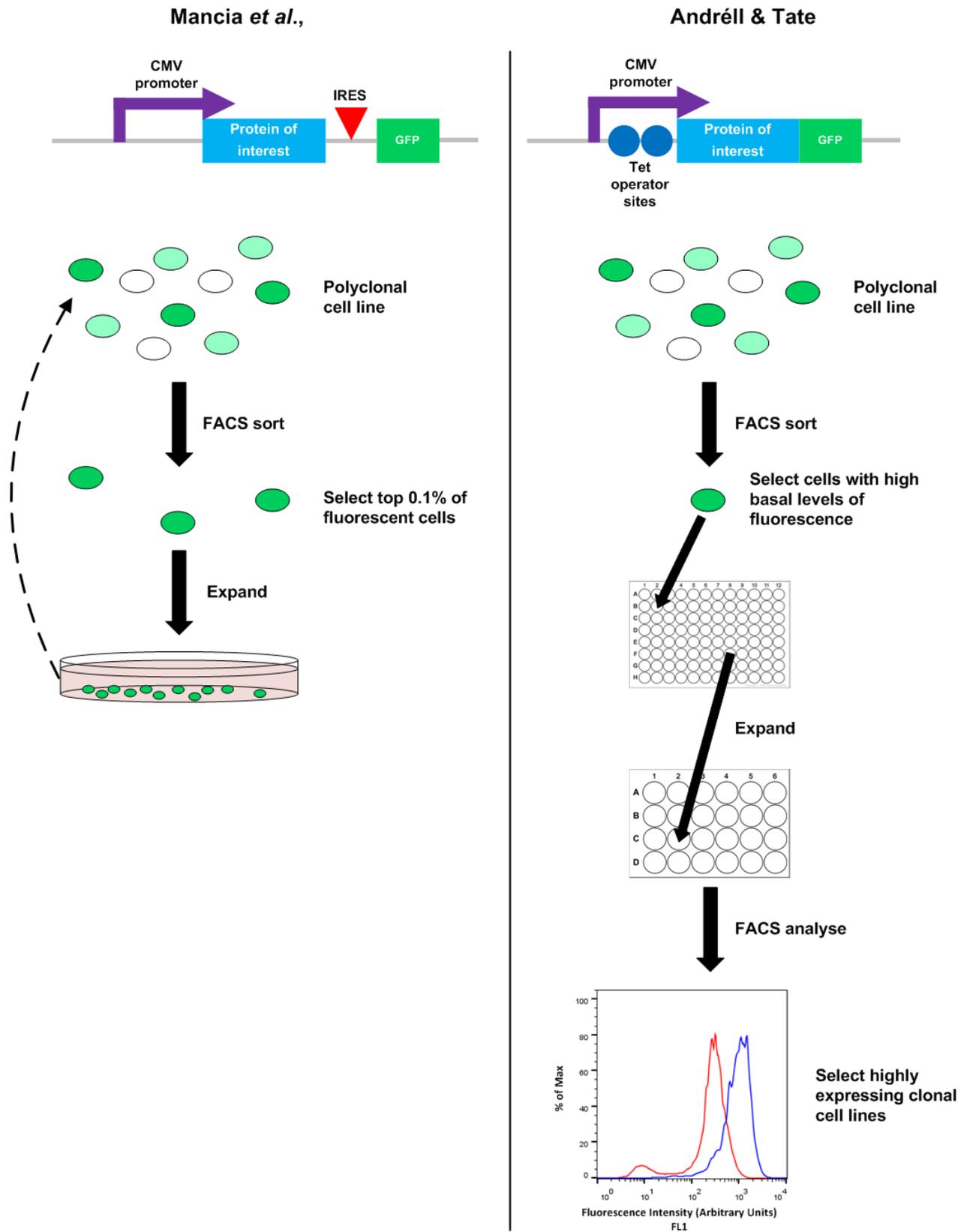


Figure 3.1 FACS based methods for increasing expression in mammalian cells

Schematic of two separate methods for increasing expression in mammalian cells using GFP as a marker for expression. On the left is the method from ¹⁵⁰ which utilises a constitutive expression system and an IRES sequence upstream of GFP. This method generates highly expressing polyclonal cell lines. On the right is the method from Andréll & Tate (unpublished). It utilises a tetracycline-inducible system and selects cells that have a high level of expression in the uninduced state. This method generates highly expressing clonal cell lines. Figure adapted from ¹⁵⁰.

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

The cells which showed the highest levels of expression were selected in either the tetracycline-induced or uninduced state. The most successful approach so far was to select cells that showed a high level of expression in the uninduced state (Andréll & Tate, unpublished). Although this method also takes about two to three months to complete, it has the advantage of being high-throughput and not requiring multiple rounds of FACS.

Limited options are currently available for increasing yields of recombinant protein from mammalian hosts. The work presented in this chapter further supports the FACS-based method for creating clonal stable cell lines which greatly increased expression levels in comparison to a traditional polyclonal cell line.

3.2 Methods

3.2.1 Generation of highly-expressing clonal cell lines using fluorescence activated cell sorting

Highly-expressing clonal AT₁R-GFP-H₁₀ cell lines were selected from the uninduced polyclonal cell line iHEK-AT₁R-GFP-H₁₀ (Section 2.3.4) using a MoFlo High Speed Cell Sorter (Beckman Coulter) with the coherent sapphire laser set to 488 nm to excite GFP (Figure 3.2). Ten cells were deposited per well in a 96-well plate and allowed to grow until confluent under standard conditions. The cells were then transferred into 24-well plates. When the cells reached ~50% confluence they were induced for 24 hours with fresh media with 1 µg/ml tetracycline. Both induced and uninduced samples were analysed on a FACSCalibur II (Becton Dickinson) (Section 2.2.9). The cell lines which showed the greatest increase in GFP expression were retained.

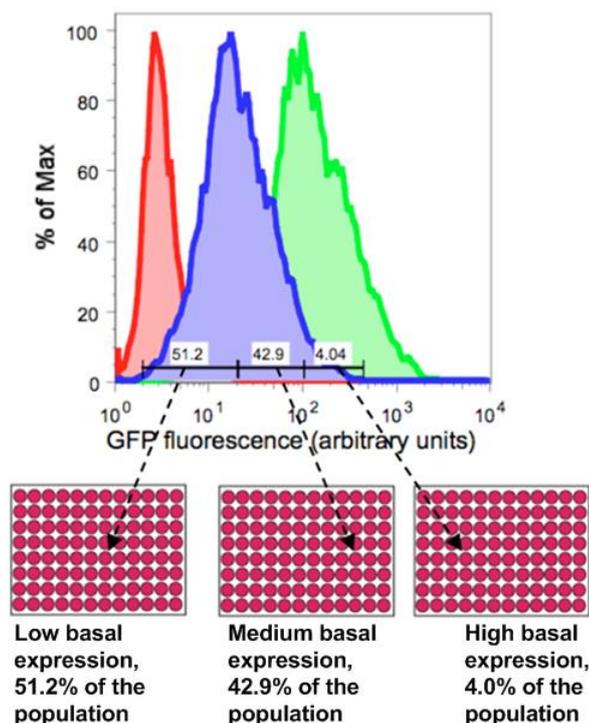


Figure 3.2 Strategy for the selection of highly-expressing clonal cell lines

FACS analysis of iHEK cells. Cells were harvested in PBS and analysed on the FACSCalibur II for GFP fluorescence using the FL-1 detector. Cell counts have been normalised (% of Max). iHEK parental cells (red), uninduced cell line stably expressing SERT-GFP (blue), induced cell line stably expressing SERT-GFP (green). Approximately 4% of the uninduced stable cell line shows high levels of basal expression. Isolating and screening individual cells from this population leads to the creation of highly-expressing clonal cell lines. Figure adapted from Juni Andréll, unpublished.

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

3.2.2 Fixing and staining cells for analysis by confocal laser scanning microscopy

To facilitate membrane staining, cells were grown on 35 mm glass bottom culture dishes, induced for 24 hours under standard conditions and fixed using 2% paraformaldehyde⁶⁰. After washing with PBS, membranes were selectively stained using a 10 µg/ml solution of concanavalin A (conA)–Alexa Fluor 647 conjugate (Invitrogen) in PBS for 10 minutes at room temperature. After further washing with PBS, cells were stored in fresh PBS with 0.02% Na azide at 4°C protected from light. Cells were visualised on a Leica TCS SP8 STED inverted laser scanning microscope with 63× oil-immersion objective and a 1.4 numerical aperture. The white light laser was set to a wavelength of 488 nm to excite GFP and to 633 nm for Alexa Fluor 647; the pinhole emission wavelength was set to 580 nm.

3.2.3 Large scale culture of mammalian cells

For suspension cultures, cells were grown in FreeStyle™ 293 Expression Medium (Gibco) supplemented with 5% tetracycline-free foetal bovine serum at a density of one million cells per ml in 2-litre roller bottles (Corning) and incubated at 37°C and 200 r.p.m. Cells were induced at a density of 1-2 million cells per ml by adding 1 µg/ml tetracycline and incubating at 37°C for 24 h. For growth in HYPERFlask™ (Corning) the manufacturer's protocol was followed.

3.3 Results

3.3.1 Development of a novel strategy to increase AT₁R expression in the mammalian system

Although the stable polyclonal mammalian cell line, iHEK(AT₁R-GFP-H₁₀) was shown to have much higher expression of AT₁R than any of the insect cell conditions investigated (0.6 mg/L compared to 0.1 mg/L, assuming 1 million cells per ml of culture), this was below what was reported for the expression of other GPCRs in mammalian hosts, such as 10 mg of opsin, per litre of culture, assuming 10 million cells per ml of culture ¹²⁶ and 1.0 mg of the NTS1 per litre of culture, assuming 1.2 million cells per ml of culture ¹²⁷. Therefore another strategy was used to increase expression of AT₁R in mammalian cells. This technique relied on creating an AT₁R-GFP-fusion where GFP served as a marker for expression under the control of a tetracycline-inducible CMV promoter. A polyclonal cell line expressing AT₁R-GFP-H₁₀ was created in iHEK cells through antibiotic selection (Section 2.3.4) and this served as the base from which clonal cell lines were created. Fluorescence activated cell sorting was used to select cells which showed high levels of GFP expression in the uninduced state and a separate selection was also performed for those cells which showed moderate levels of expression (Figure 3.2). From each of these two populations one cell was deposited per well in a 96-well plate, however only two cell lines survived from the first population and one from the second. To determine the minimum number of FACS-selected cells per well that are necessary to yield over 50% of the clonal cell lines surviving seven days after sorting, different numbers of cells per well were deposited into a 96-well plate and assessed for viability after seven days in culture (Figure 3.3). As a result of this optimisation ten cells were deposited per well for all further plates.

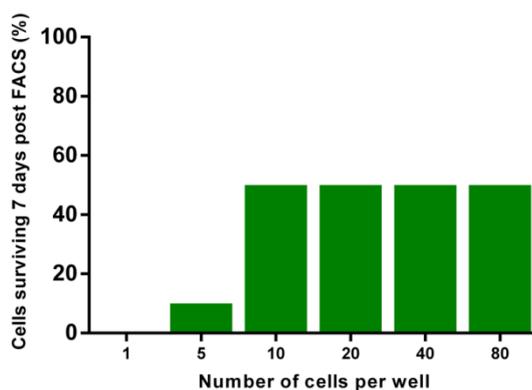


Figure 3.3 Optimisation of the number of cells per well deposited for FACS

In order to determine the minimum number of cells per well which led to at least 50% of wells containing viable cells, varying numbers of uninduced iHEK(AT₁R-GFP-H₁₀) cells were deposited per well in duplicate and the approximate number of cells surviving 7 days post sorting was assessed under a light microscope.

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

The clonal cell lines were allowed to grow until confluent in the 96-well plates and then transferred into 24-well plates in triplicate. When the cell lines reached ~50% confluence in the 24-well plates, one plate was induced for 24 hours with fresh media containing 1 µg/ml tetracycline while the second plate was kept as an uninduced sample and a third plate was retained as a stock. In total 72 clonal cell lines in the induced and uninduced states were analysed for GFP expression using FACS analysis (Appendix 3). Of the total 72 clonal cell lines analysed, 24 were derived from the population with moderate basal expression and 48 were derived from the population with high basal expression. The clonal cell lines which showed the most dramatic increase in expression in comparison to the polyclonal iHEK(AT₁R-GFP-H₁₀) cell line came from the population with a high basal level of expression (Figure 3.4).

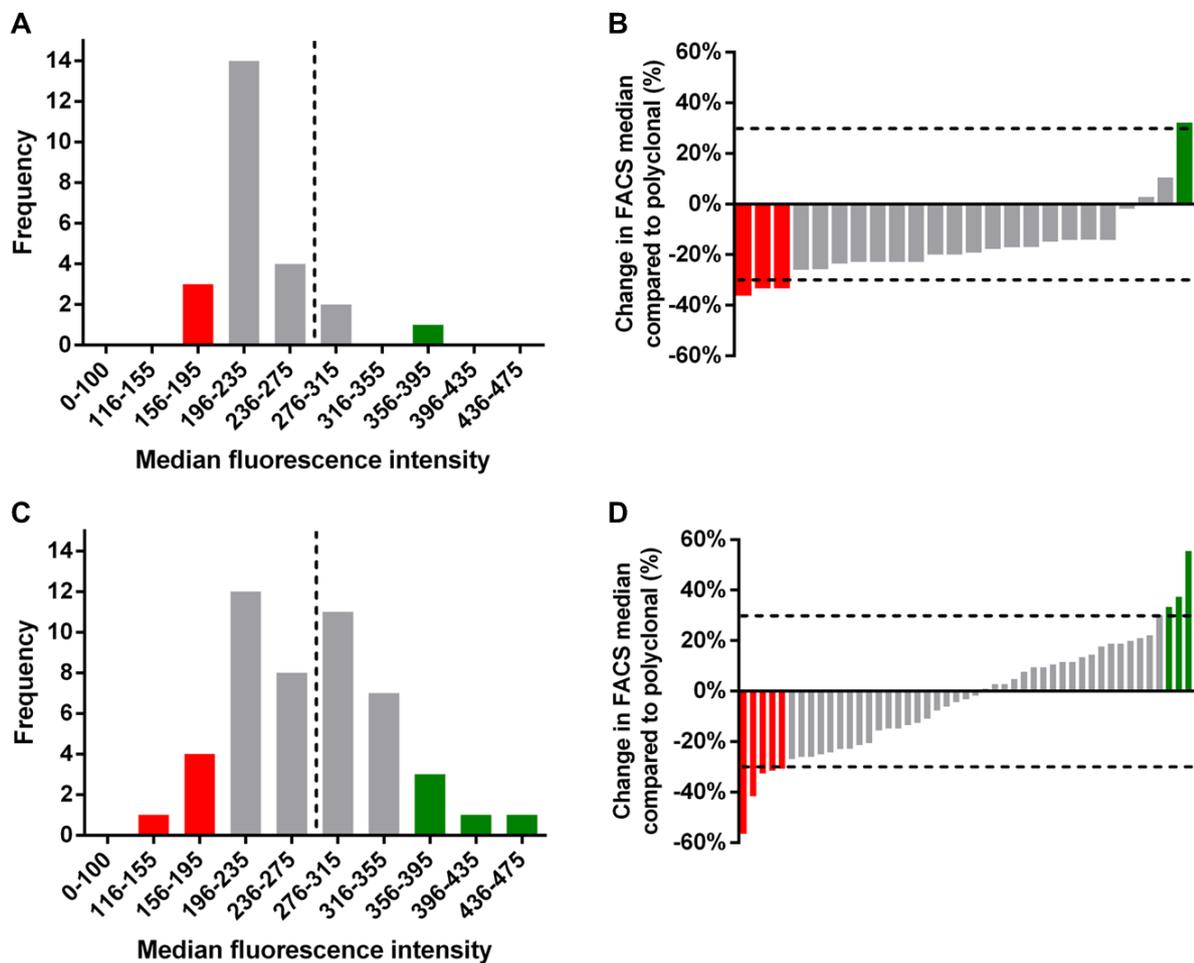


Figure 3.4 Isolation of clonal cell lines from a population with high basal expression leads to an increase in AT₁R expression

Frequency distribution histogram of FACS analysis data (Appendix 3). Clonal cell lines were isolated from two different populations of the polyclonal cell line iHEK(AT₁R-GFP-H₁₀): those with moderate basal expression (**A** and **B**) and those with high basal expression (**C** and **D**). The polyclonal cell line had a median fluorescence intensity of 276; represented by the dashed black line in the graphs (**A**) and (**C**) and 0% in the graphs (**B**) and (**D**). For the graphs (**B**) and (**D**) each bar represents an individual cell line and dashed lines are at $\pm 30\%$ of the relative expression in the polyclonal cell line, which equated to ± 2 bins on the histograms (**A** and **C**).

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

13% of the clonal cell lines derived from the population with moderate basal expression showed a > 30% reduction in FACS median in comparison to the polyclonal cell line and only 4% showed a > 30% increase in FACS median (Figure 3.4). In contrast, 10% of the clonal cell lines derived from the population with high basal expression showed a > 30% reduction in FACS median in comparison to the polyclonal cell line and 10% showed a > 30% increase in FACS median (Figure 3.4). It was not possible to predict expression levels from the FACS median of uninduced cells because all of the cell lines tested had medians between 8-14 (iHEK parental cells, 6), whereas after induction the FACS median was 120-437 (Figure 3.5). This shift in FACS median between uninduced and induced states is clearly shown in (Figure 3.6). The clonal cell lines which showed the greatest increase in expression of AT₁R-GFP-H₁₀ were grown in culture for a further nine days and assessed for expression using FACS (Figure 3.7).

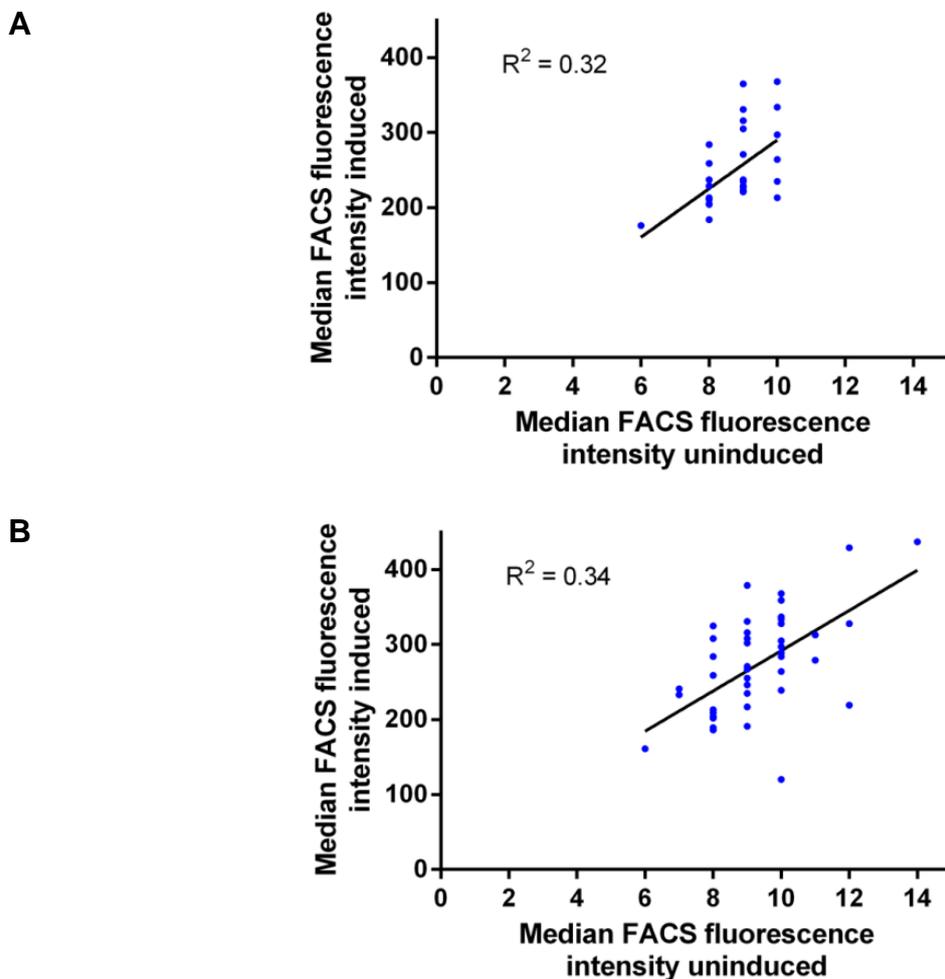
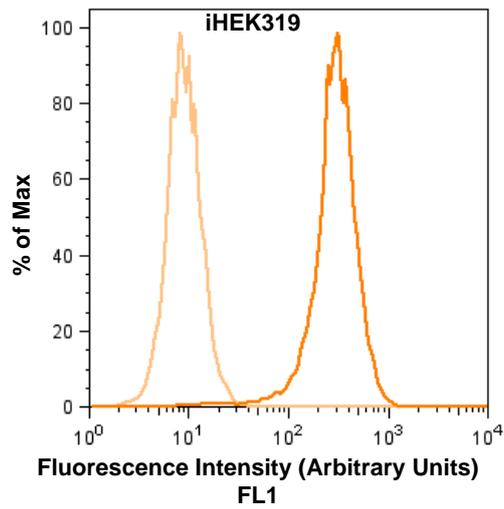


Figure 3.5 Correlation of median FACS fluorescence intensity

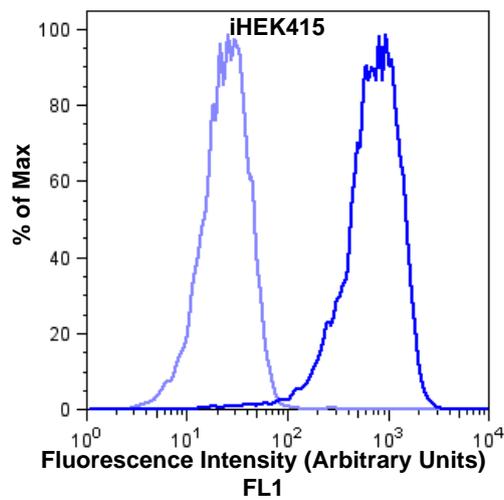
The median FACS fluorescence intensity data (Appendix 3) of the clonal cell lines in the induced state was correlated with that in the uninduced state. Linear regression was used to establish a line of best fit for cells derived from the population with moderate basal expression (**A**) and those derived from the population with high basal expression (**B**) from the original FACS sort (Figure 3.2).

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

A



B



C

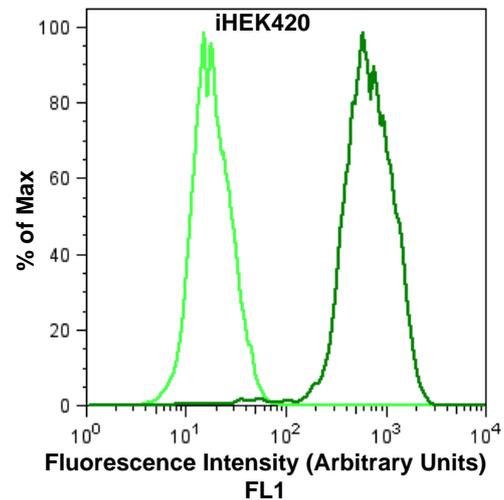
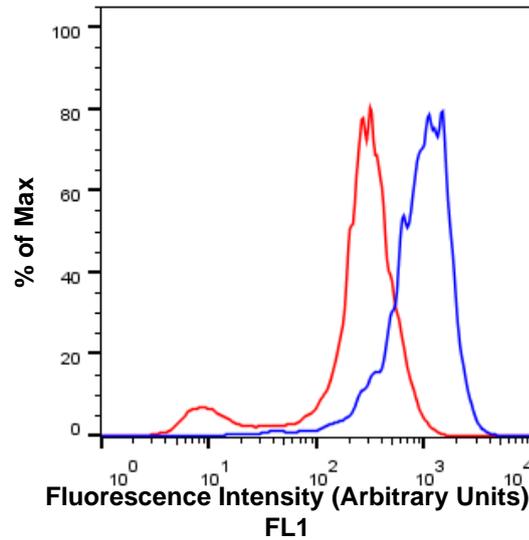


Figure 3.6 Clonal cell lines show a high level of inducibility

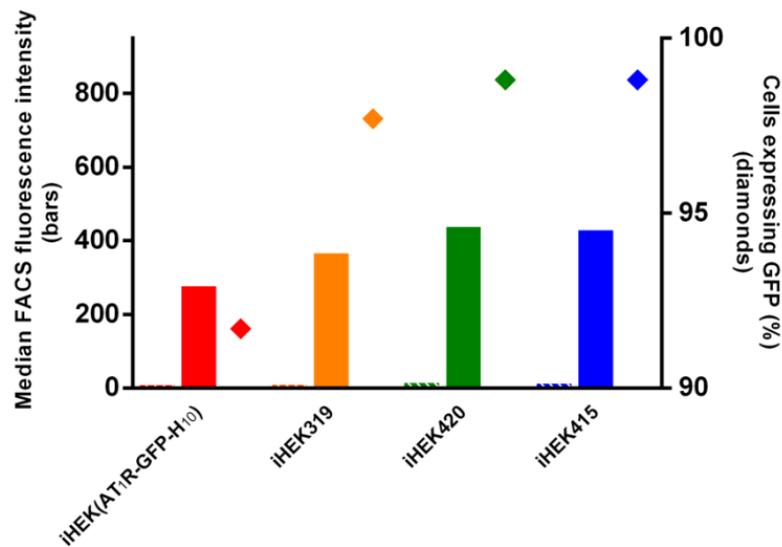
FACS analysis of iHEK cells. Cells were harvested in PBS and analysed on the FACSCalibur II for GFP fluorescence using the FL-1 detector. Cell counts have been normalised (% of Max). Light colours represent uninduced whereas dark colours represent cells induced with tetracycline for 24 hours. **(A)** Clonal iHEK319 cell line, derived from the population with moderate basal expression (orange). **(B)** Clonal iHEK415 cell line, derived from the population with high basal expression (blue). **(C)** Clonal iHEK420 cell line, derived from the population with high basal expression (green).

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

A



B



C

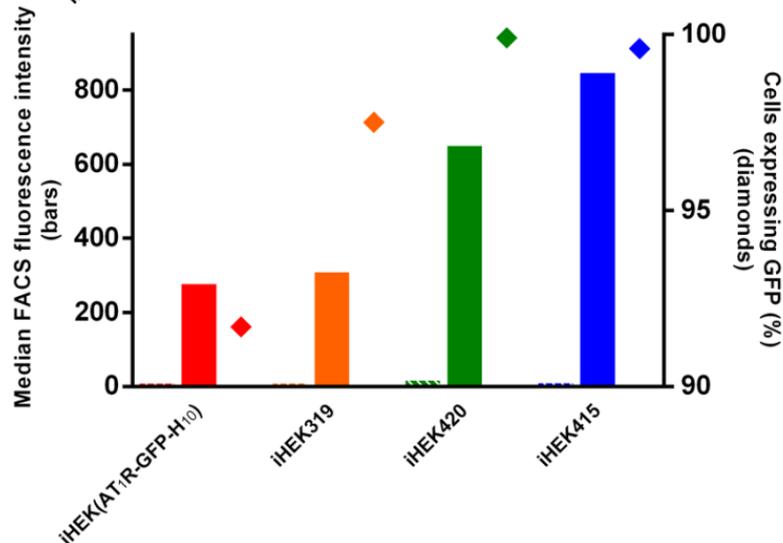


Figure 3.7 Increase in expression of AT₁R-GFP-H₁₀ in stable cell lines over time

FACS analysis of iHEK cell lines was performed on cells harvested in PBS and analysed on the FACSCalibur II for GFP fluorescence using the FL-1 detector. Cell counts have been normalised (% of Max). **(A)** Polyclonal parental cell line, iHEK(AT₁R-GFP-H₁₀) after induction with tetracycline (red), clonal iHEK415 cell line after induction with tetracycline (blue). **(B and C)** Median FACS fluorescence intensity; uninduced (hatched bars), induced (solid bars), percentage of cells showing GFP fluorescence (diamonds). The expression levels of AT₁R-GFP-H₁₀ from the initial analysis **(B)** were compared with cells grown continuously for further 9 days and then induced with tetracycline **(C)**.

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

3.3.2 Characterisation of AT₁R expression in clonal mammalian cell lines

As a verification of expression of AT₁R, cell suspensions derived from the most highly expressing cell lines were equalised for total protein, separated by SDS-PAGE and visualised by in-gel fluorescence (Figure 3.8). The bands corresponding in size to AT₁R-GFP were quantified with ImageJ and this data was correlated to the median fluorescence intensity from FACS analysis of the samples prior to harvest (Figure 3.8). The strong association ($R^2 = 0.9$, Figure 3.8) between the two sets of data suggests that GFP fluorescence as measured by FACS is a good indicator of AT₁R expression due to the absence of GFP un-associated with AT₁R (*e.g.* through proteolysis). The cell lines which showed the greatest increase in AT₁R expression upon tetracycline-induction were also analysed by FSEC with the same number of cells in each sample (Section 2.2.12). All four cell lines produced a major peak at 11 ml retention volume that corresponded to AT₁R-GFP-H₁₀. However the size of the peak varied markedly between the cell lines, with the largest peak produced from iHEK415 and the smallest peak from the polyclonal parental cell line iHEK(AT₁R-GFP-H₁₀). Intermediate-sized peaks were observed for the cell lines iHEK420 and iHEK425. In addition, a second fluorescent product was observed at a retention volume of 17.5 ml in the iHEK425 cell line, which probably represented free GFP (Figure 3.9). The cell line iHEK415 was also examined by laser scanning confocal microscopy (Figure 3.10). The plasma membranes were stained with Alexa Fluor 647 conjugated to conA and co-localisation of AT₁R-GFP was shown on the cell surface indicating that AT₁R-GFP was expressed predominantly on the plasma membrane in the iHEK415 cell line. As a final check for the presence of correctly folded material, the clonal cell lines which showed the greatest increase in AT₁R expression were analysed by radioligand binding using the antagonist [¹²⁵I]Sar¹ (Section 2.2.11). In a one-point binding assay, the approximate number of functional AT₁R molecules per cell was determined with 15 million copies per cell observed in cell lines iHEK415 and iHEK420. This compared to 9 million copies per cell in the polyclonal parental cell line and 2 million copies per cell in iHEK425 (Figure 3.11). The cell lines iHEK415 and iHEK420 showed only a modest reduction in expression levels of AT₁R-GFP after induction when the cell lines had been cultured for over 20 days (Figure 3.12). The combination of the FSEC data and the ligand binding data was used to determine which cell line was the best at producing AT₁R for further studies. The low number of copies per cell and the presence of free GFP in FSEC suggested that AT₁R-GFP-H₁₀ expressed in the cell line iHEK425 was being degraded. Cell lines iHEK420 and iHEK425 apparently produced the most functional AT₁R as assessed by ligand binding (15 million

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

copies per cell), but the FSEC peak for AT₁R-GFP-H₁₀ appeared to be considerably smaller for iHEK420 than that observed for iHEK415. This could be because AT₁R-GFP-H₁₀ produced in iHEK420 was either poorly solubilised or more prone to aggregation than the AT₁R-GFP-H₁₀ produced in iHEK415, although the reason for this is unclear. Therefore, based on this analysis, iHEK415 was considered the best cell line for the production of AT₁R-GFP-H₁₀, which, based on the ligand binding data, equates to 1 mg per litre of cells assuming 1 million cells per ml.

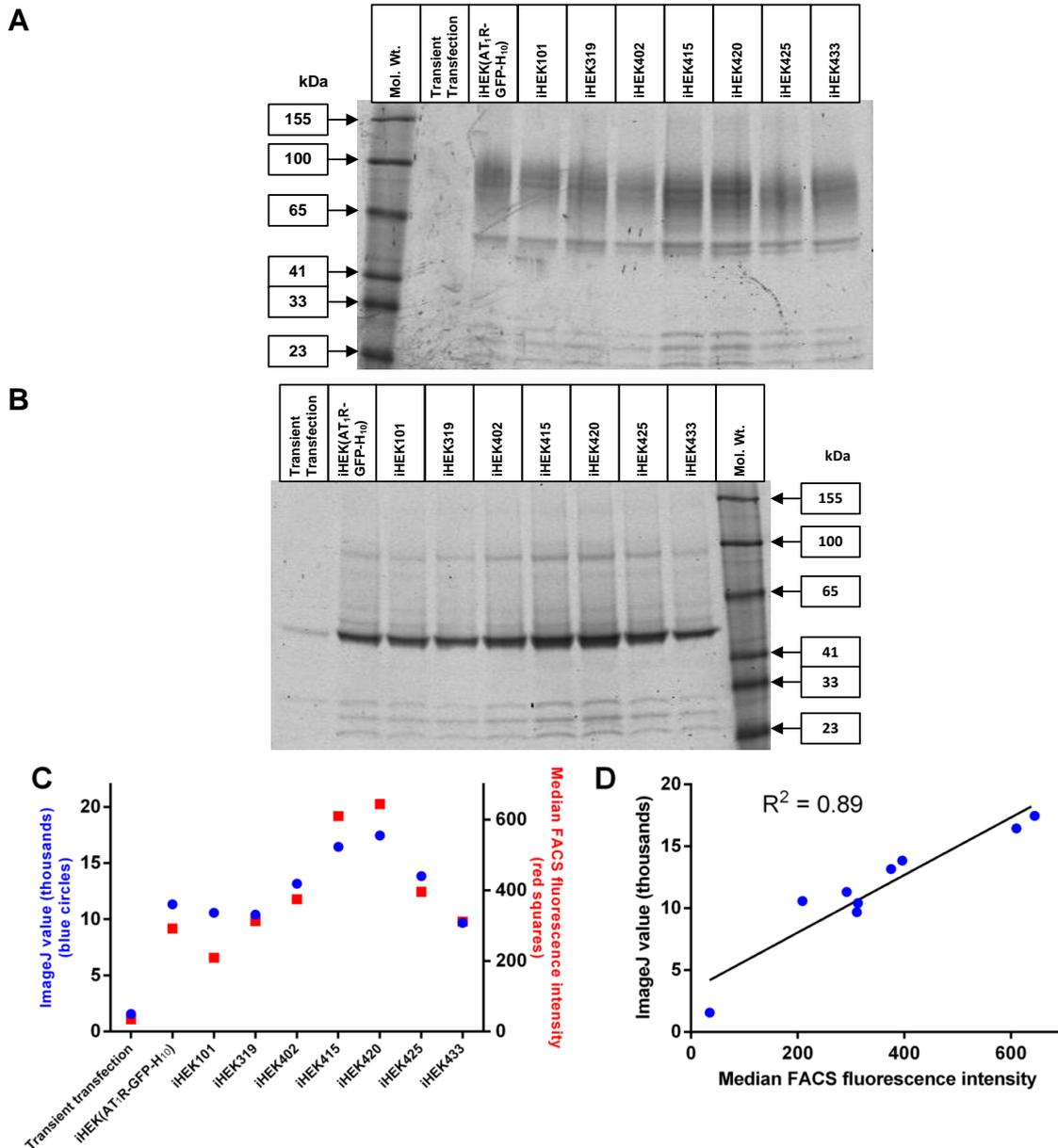


Figure 3.8 In-gel fluorescence shows an increase in expression of AT₁R-GFP-H₁₀ through the creation of clonal cell lines

iHEK cells were transiently transfected with pJAP2 (expression of AT₁R-GFP-H₁₀) and polyclonal and clonal stable cell lines were generated as described before (Section 2.3.4 and Figure 3.2). All cells were induced for 24 hours with 1 μ g/ml tetracycline prior to being assayed. An equal amount of cells was loaded in each lane. **(A)** In-gel fluorescence of AT₁R-GFP-H₁₀ expression from transient transfection, polyclonal and clonal stable cell lines. **(B)** In-gel fluorescence of AT₁R-GFP-H₁₀ after removal of N-linked glycosylation by PNGase F. **(C)** Bands corresponding to deglycosylated AT₁R-GFP-H₁₀ were quantified with ImageJ (blue circles) and compared to their respective median FACS fluorescence intensity values (red squares). **(D)** Linear regression was used to establish a line of best fit.

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

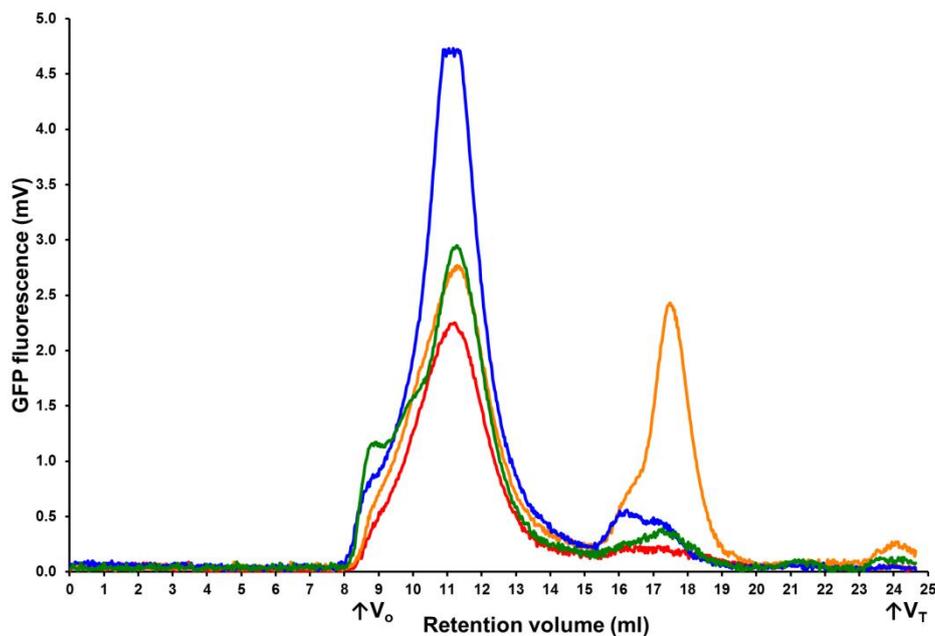


Figure 3.9 FSEC analysis of AT₁R-GFP-H₁₀ produced from clonal cell lines

The clonal cell lines derived from iHEK(AT₁R-GFP-H₁₀) were induced with tetracycline for 24 hours. 40 nM of Sar¹ was added to 5 million cells and allowed to bind for 1 hour at 23°C before being solubilised in 1% DDM. Samples were then evaluated on a Superdex 200 10/300 GL column. Polyclonal cell line iHEK(AT₁R-GFP-H₁₀) (red); clonal cell lines: iHEK420 (green), iHEK425 (orange) and iHEK415 (blue). The elution of AT₁R-GFP-H₁₀ was detected using GFP fluorescence (mV). The void (V_0) and total column volumes (V_T) are indicated.

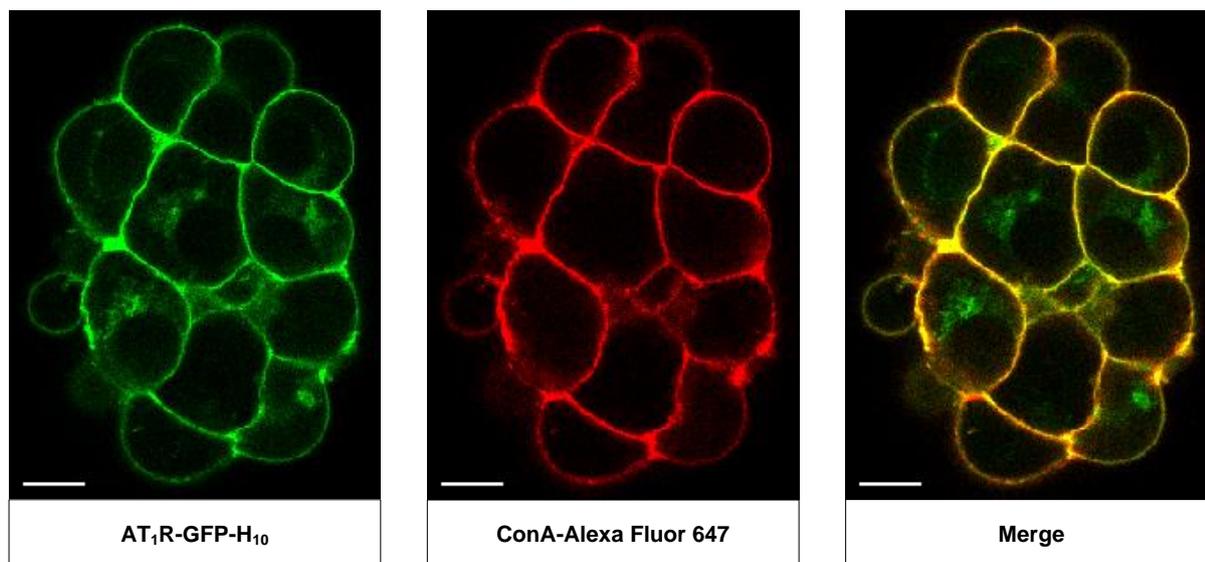


Figure 3.10 AT₁R produced in the clonal cell line iHEK415 is predominantly localised to the cell surface

Confocal micrographs of the iHEK415 clonal cell line (expression AT₁R-GFP-H₁₀) after 24 hours induction with tetracycline. Cells were fixed using paraformaldehyde and the plasma membrane was defined by staining with Alexa Fluor 647-conjugated conA prior to visualisation. Unlabelled iHEK parental cells showed no fluorescence (not shown). The scale bar represents 10 μ m.

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

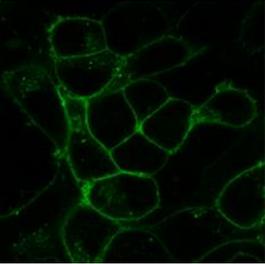
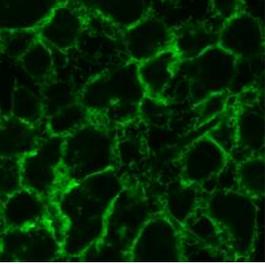
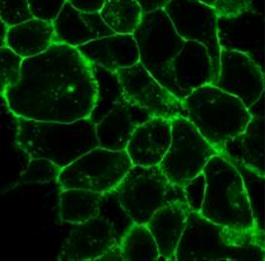
Cell line name	iHEK(AT ₁ R-GFP-H ₁₀)		Polyclonal		
Open reading frame schematic	AT ₁ R		GFP	Strep II	H ₁₀
Parental cell line	iHEK		Master Gain	1107	
Median FACS fluorescence intensity	265		Digital Gain	1.24	
Functional expression levels (copies/cell)	9 million ± 0.5 (n=6)		Digital Offset	0.00	
Cell line name	iHEK415		Clonal		
Open reading frame schematic	AT ₁ R		GFP	Strep II	H ₁₀
Parental cell line	iHEK(AT ₁ R-GFP-H ₁₀)		Master Gain	696	
Median FACS fluorescence intensity	846		Digital Gain	0.30	
Functional expression levels (copies/cell)	15 million ± 0.4 (n=6)		Digital Offset	2457.60	
Cell line name	iHEK420		Clonal		
Open reading frame schematic	AT ₁ R		GFP	Strep II	H ₁₀
Parental cell line	iHEK(AT ₁ R-GFP-H ₁₀)		Master Gain	952	
Median FACS fluorescence intensity	649		Digital Gain	1.00	
Functional expression levels (copies/cell)	15 million ± 1.5 (n=6)		Digital Offset	113.75	
Cell line name	iHEK425		Clonal		
Open reading frame schematic	AT ₁ R		GFP	Strep II	H ₁₀
Parental cell line	iHEK(AT ₁ R-GFP-H ₁₀)				
Median FACS fluorescence intensity	418				
Functional expression levels (copies/cell)	2 million ± 0.2 (n=6)				

Figure 3.11 Assessment of clonal cell lines derived from iHEK(AT₁R-GFP-H₁₀)

FACS was used on the uninduced polyclonal cell line iHEK(AT₁R-GFP-H₁₀) to select clonal cell lines which showed the highest amount of GFP expression in the uninduced state. Assessment of the resulting clonal cell lines was based on confocal microscopy and the average counts for FACS analysis and [¹²⁵I]-Sar¹ binding.

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

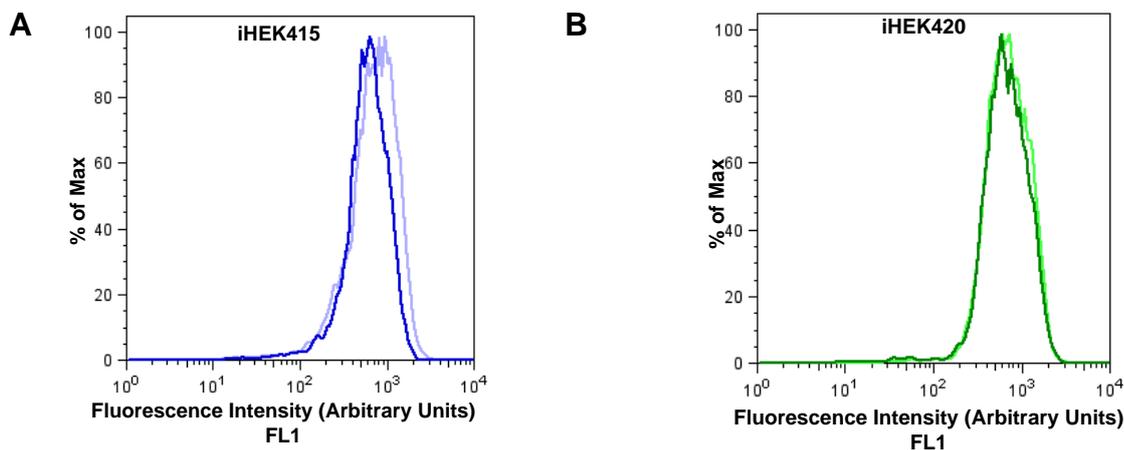


Figure 3.12 Clonal cell lines show high levels of expression long term

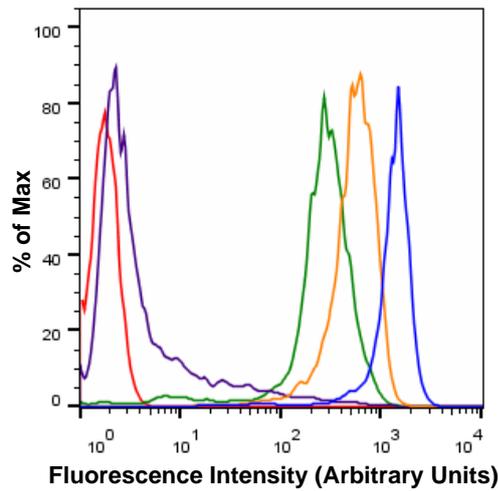
FACS analysis of iHEK cells. Cells were induced for 24 hours with tetracycline and harvested in PBS and analysed on the FACSCalibur II for GFP fluorescence using the FL-1 detector. Cell counts have been normalised (% of Max). **(A)** The expression levels of AT₁R-GFP-H₁₀ in the clonal cell line iHEK415 were compared in cells grown in culture for 26 days (dark blue line, FACS median 629) and freshly thawed cells (light blue line, FACS median 737). **(B)** The expression levels of AT₁R-GFP-H₁₀ in the clonal cell line iHEK420 were compared in cells grown in culture for 21 days (dark green line, FACS median 547) and freshly thawed cells (light green line, FACS median 649).

3.3.3 Dramatic increase in expression of AT₁R in mammalian cells through a combination of FACS selection and use of sodium butyrate

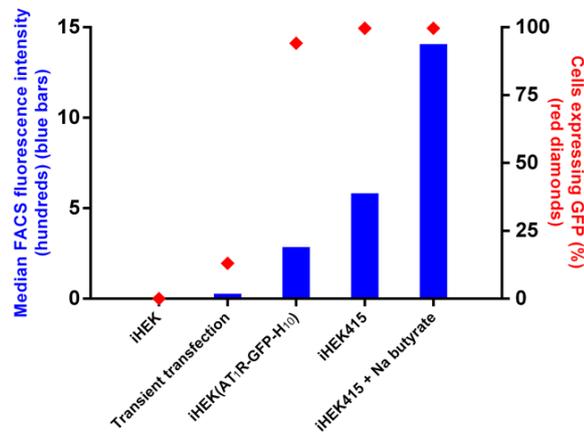
The process of selecting clonal cell lines on the basis of high basal expression of AT₁R-GFP led to a 66% increase in expression in comparison to the polyclonal cell line based on the radioligand binding data. As it was previously noted that adding sodium butyrate at the time of induction increased active expression of AT₁R in the polyclonal cell line (Section 2.3.4), the effect of 5 mM sodium butyrate added to the iHEK415 cell line was tested. FACS analysis showed a dramatic increase in expression and radioligand binding experiments demonstrated that functional levels of AT₁R in the iHEK415 cell line increased to approximately 26 million receptors per cell or 1.8 mg of AT₁R per litre of culture assuming 1 million cells per ml (Figure 3.13). Sodium butyrate in combination with the generation of a clonal cell line produced a nearly a threefold increase in AT₁R expression over the polyclonal cell line iHEK(AT₁R-GFP-H₁₀). A western blot loaded with equal amounts of active AT₁R per lane indicated that the increase in expression was not at the expense of correct folding of the receptor since equal signal intensity was seen for all conditions (Figure 3.14). As an additional check on the quality of the AT₁R produced, an equal amount of cells transiently transfected with the plasmid pJAP2 (expressing AT₁R-GFP-H₁₀), the polyclonal cell line iHEK(AT₁R-GFP-H₁₀) or the clonal cell line iHEK415 were separated by SDS-PAGE and visualised by in-gel fluorescence (Figure 3.15). Digestion of these samples with the glycosidase PNGase F indicated that AT₁R was glycosylated in all of the conditions tested.

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

A



B



C

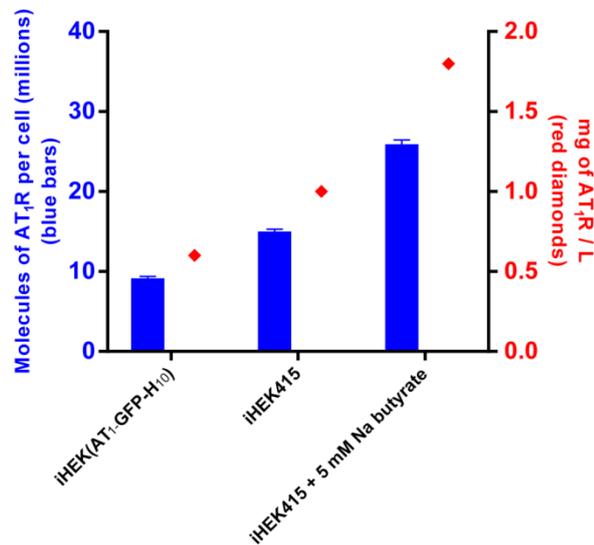


Figure 3.13 Significant increase in AT₁R expression through a combination of FACS selection and the use of sodium butyrate

FACS analysis of iHEK cells was performed on cells harvested in PBS and analysed on the FACSCalibur II for GFP fluorescence using the FL-1 detector. Cell counts have been normalised (% of Max). **(A)** Histogram of GFP fluorescence intensity. Non-transfected parental iHEK cells (red), transient transfection of pJAP2 (expressing AT₁R-GFP-H₁₀) (purple), polyclonal stable cell line iHEK(AT₁R-GFP-H₁₀) (green), clonal cell line iHEK415 (orange), iHEK415 induced with 5 mM Na butyrate (blue). **(B)** Median FACS fluorescence intensity (blue bars) are compared with the number of cells positive for GFP expression (%) (red diamonds). **(C)** The amount of functional DDM-solubilised AT₁R in each cell line was determined by measuring specific binding of the antagonist [¹²⁵I]-Sar¹ (blue bars). The results are from two independent experiments performed in triplicate and plotted as a mean value ± SEM. From these data, milligrams of AT₁R per litre of cells, assuming 1 million cells per ml, were calculated (red diamonds).

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

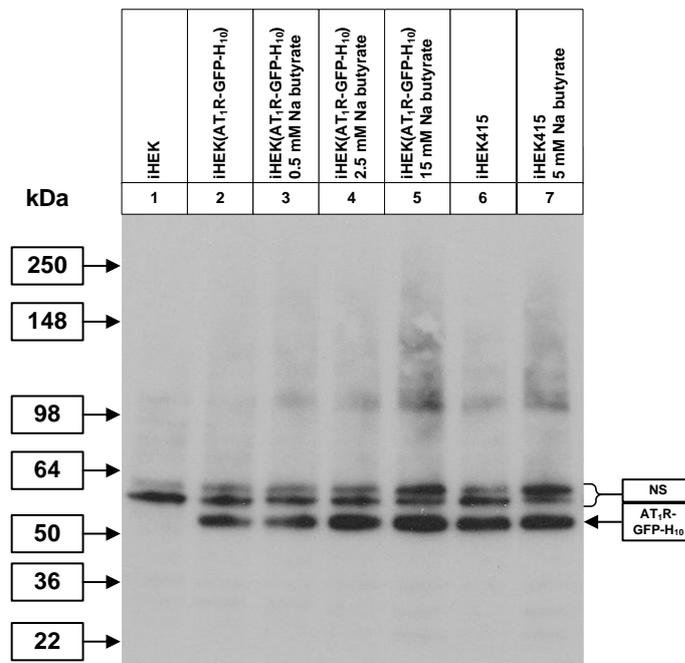


Figure 3.14 Sodium butyrate does not significantly increase the amount of unfolded AT₁R expressed

Western blot of DDM-solubilised AT₁R, with equal amounts of active receptor per sample (lanes 2-7). The blot was probed with an anti-pentaHis-HRP conjugated antibody. Lane 1, iHEK parental cells; lanes 2-5, iHEK(AT₁R-GFP-H₁₀) stable polyclonal cell line; lanes 6-7 clonal cell line, iHEK415. N-linked glycosylation was removed using PNGase F on all samples. Cells were induced with 1 μ g/ml tetracycline for 24 hours. The amount of functional AT₁R was determined by measuring specific binding of the antagonist [¹²⁵I]-Sar¹. Two non-specific bands are indicated (NS).

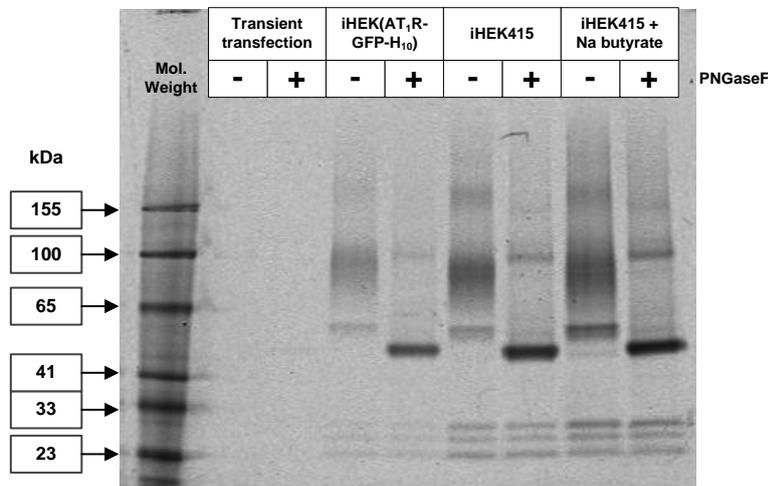


Figure 3.15 Increasing levels of AT₁R expression through FACS selection and induction with sodium butyrate

In-gel fluorescence of AT₁R-GFP-H₁₀ expression from transient transfection, polyclonal and clonal stable cell lines. iHEK cells were transiently transfected with pJAP2 (expression of AT₁R-GFP-H₁₀). Polyclonal and clonal stable cell lines were generated as described before (Section 2.3.4 and Figure 3.2). All cells were induced for 24 hours with 1 μ g/ml tetracycline or 1 μ g/ml tetracycline with 5 mM Na butyrate as indicated for 24 hours prior to being assayed. An equal amount of cells was loaded in each lane.

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

3.3.4 Large scale growth of iHEK415

The best clonal cell line for expressing AT₁R was determined to be iHEK415. In order to obtain the milligram quantities required for purification and crystallisation of the receptor, the next step was to explore large scale growth. iHEK415 was grown adherently in T-175 tissue culture flasks (172 cm² surface area) and induced for 24 hours with tetracycline. Each flask yielded approximately 2.5 x 10⁷ million cells. Assuming 1 million cells per ml, one litre of cells would yield 1 x 10⁹ cells, therefore approximately 40 T-175 flasks would be required to obtain approximately 2 mg of AT₁R. Growing the iHEK415 cell line in this fashion would require a large amount of time and resources, therefore suspension culture of iHEK415 was investigated. iHEK415 cells initially grown adherently were harvested and resuspended in Free Style media supplemented with 5% tetracycline-free foetal bovine serum at 1 million cells per ml in 2 litre roller bottles and placed into a shaking incubator at 37°C. The suspension cultures of iHEK415 quickly aggregated onto the sides of the flasks, a result which was also observed when the cell concentration was diluted to 0.5 million cells per ml. The iHEK415 suspension cultures were checked every 24 hours and diluted as necessary. The cells were induced at 1.5 million cells per ml for 24 hours and analysed by FSEC (Figure 3.16). The cell line iHEK415 grown in suspension culture produced a peak at 11 ml retention volume that corresponded to AT₁R-GFP-H₁₀. However a second larger fluorescent product was observed at a retention volume of 17.5 ml, which probably represented free GFP. Since suspension culture did not provide a favourable environment to grow iHEK415 in large scale, adherent cultures using roller bottles was next examined. This however also failed since the cells would not adhere to the sides of the bottles. The next method that was investigated was growth in Hyperflasks (1720 cm² surface area per flask). These flasks proved to be technically challenging to use when changing the media for induction since iHEK cells are only loosely adherent and therefore large losses of cells occurred. From four Hyperflasks tested the average yield was 1.5 x 10⁸ million cells per flask.

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

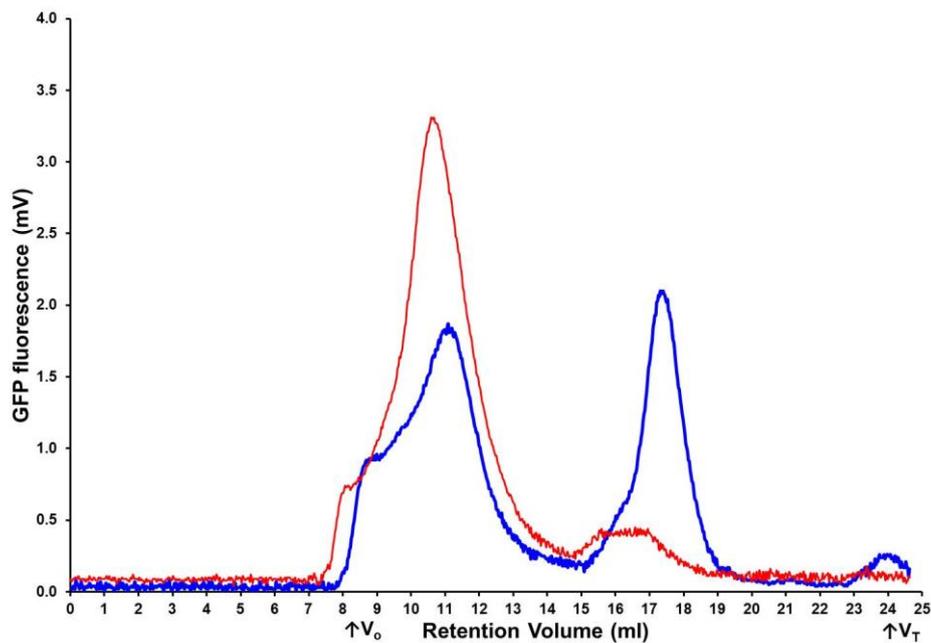


Figure 3.16 FSEC analysis of AT₁R-GFP-H₁₀ produced from iHEK415 grown in suspension

The clonal cell line iHEK415 was grown in either suspension culture (blue line) or adherent culture (red line) and induced with tetracycline for 24 hours. 40 nM of Sar¹ was added to 5 million cells and allowed to bind for 1 hour at 23°C before being solubilised in 1% DDM. Samples were then evaluated on a Superdex 200 10/300 GL column. The elution of AT₁R-GFP-H₁₀ was detected using GFP fluorescence (mV). The void (V₀) and total column volumes (V_T) are indicated.

3.4 Discussion

The main reason for creating a clonal cell line was to increase expression levels. The polyclonal cell line iHEK(AT₁R-GFP-H₁₀) had nine million molecules of AT₁R per cell, whereas creation of a clonal cell line increased this number to 15 million molecules of AT₁R per cells, representing nearly a two-fold increase. Functional yields of AT₁R from the clonal cell line iHEK415 were further increased to 26 million molecules of AT₁R per cells or approximately two mg of AT₁R per litre of culture assuming one million cells per ml. Mancina *et al.*¹⁵⁰ described a FACS-based method for increasing expression levels of membrane proteins in human cells that expressed proteins constitutively from the strong CMV promoter. Optimisation of this system produced 3 million molecules per cell of the rat serotonin receptor subtype 2c (5HT_{2C}) or 2.5 mg of receptor from 1-5 litres of culture, assuming 10⁷ cells per ml. However this method required iterative rounds of FACS each time large volumes of cells were required for receptor purification¹⁵⁰. Nevertheless, the technique described by Mancina *et al.*¹⁵⁰, illustrated the value of FACS as a method for increasing expression levels. The method presented here used an inducible promoter to express the receptor, rather than a constitutive promoter; this approach had the advantage that cells grown in the uninduced state expressed very low levels of AT₁R. Therefore the cells maintained a high level of expression after induction despite being in culture for over twenty days in the uninduced state. Thus a FACS sort was not required each time a large batch of cells was required, as observed for the constitutive expression system¹⁵⁰.

Another example of a mammalian membrane protein highly over expressed in mammalian cells was the human β_2 adrenergic receptor, which was expressed at 200 pmoles/mg in a stable CHO cell line¹⁵². However, these levels decreased dramatically when the cells were left in culture for an extended period of time, possibly due to toxicity effects¹⁰⁹. A decrease in expression levels in mammalian stable cell lines upon extended culture under a constitutive promoter has been a common observation. Examples include, the β_2 adrenergic receptor¹⁵³, the bradykinin B₂ receptor¹⁵⁴, SERT⁶⁰ and rhodopsin^{126, 128, 155}. When comparative studies have been performed between inducible and constitutive mammalian expression systems, there has been a 4-12-fold increase in expression with the use of the tetracycline-inducible system¹¹⁰. The decline in expression seen with a constitutive promoter could have been due to the excessive metabolic demands on the cell for protein biosynthesis, the intrinsic activity of the protein or there being insufficient amounts of specific molecular chaperones required for folding¹¹⁰. Thus the main advantage of the

CHAPTER 3 Improving the Expression of AT₁R in Mammalian Cells

inducible mammalian system is the cells grow at a normal rate, unencumbered by the overproduction of recombinant proteins, until they reach the required density, at which point they can be induced to produce protein. That inducible expression in mammalian cells is far superior to the use of constitutive promoters is perhaps unsurprising given that this has been known in bacterial and yeast expression systems for several decades.

The clonal cell line iHEK415 overexpressed AT₁R to give a final yield of 2 mg of receptor per litre of culture, assuming 1 million cells per ml. However, whilst the clonal cell line iHEK415 showed little loss in expression upon prolonged growth on a small scale, it was not possible to adapt the cell line to suspension culture. Growing the cell line adherently for the production of AT₁R for structural studies would have been very resource and labour intensive. Therefore a new stable cell line was required that could be grown on a large scale. As this would take 2-3 months to do it was decided to also engineer AT₁R to improve the probability of crystallography, by removing N-linked glycosylation sites and removing flexible regions at the C-terminus, and also other potential improvements in expression levels were explored (Chapter 5).

CHAPTER 4 ASSESSMENT OF AT₁R LIGANDS FOR STABILISATION

4.1 Introduction

As a key component of the cardiovascular system, AT₁R has generated considerable pharmacological interest⁷⁶ and numerous ligands have been designed which block the receptor. The structures of several AT₁R ligands are illustrated in Figure 4.1 and their respective binding affinities are shown in Table 4.1. For more background on AT₁R pharmacology see Section 1.4.3.

The choice of a ligand for stabilisation and co-crystallisation with a GPCR is crucial for obtaining well-diffracting crystals and there are several desirable characteristics of a ligand that are thought to correlate with its usefulness in GPCR crystallography. For example, for wild type A_{2A}R-T4L, out of eight ligands examined the only structures obtained were with ligands that gave the highest apparent T_ms; UK-432097⁵⁰ and ZM241385¹⁶. Further structures of A_{2A}R bound to ligands with a lower apparent T_m required the use of techniques to stabilise the receptor such as thermostabilisation^{63, 49} or the use of antibodies⁶¹. An ideal ligand should have a high binding affinity (pM to low nM), a slow off rate and it should stabilise the receptor in a single conformation¹³. Since there are multiple ligands with tight binding affinities for AT₁R, FSEC was used to screen for the best candidates for stabilisation and co-crystallisation with AT₁R.

CHAPTER 4 Assessment of AT₁R Ligands for Stabilisation

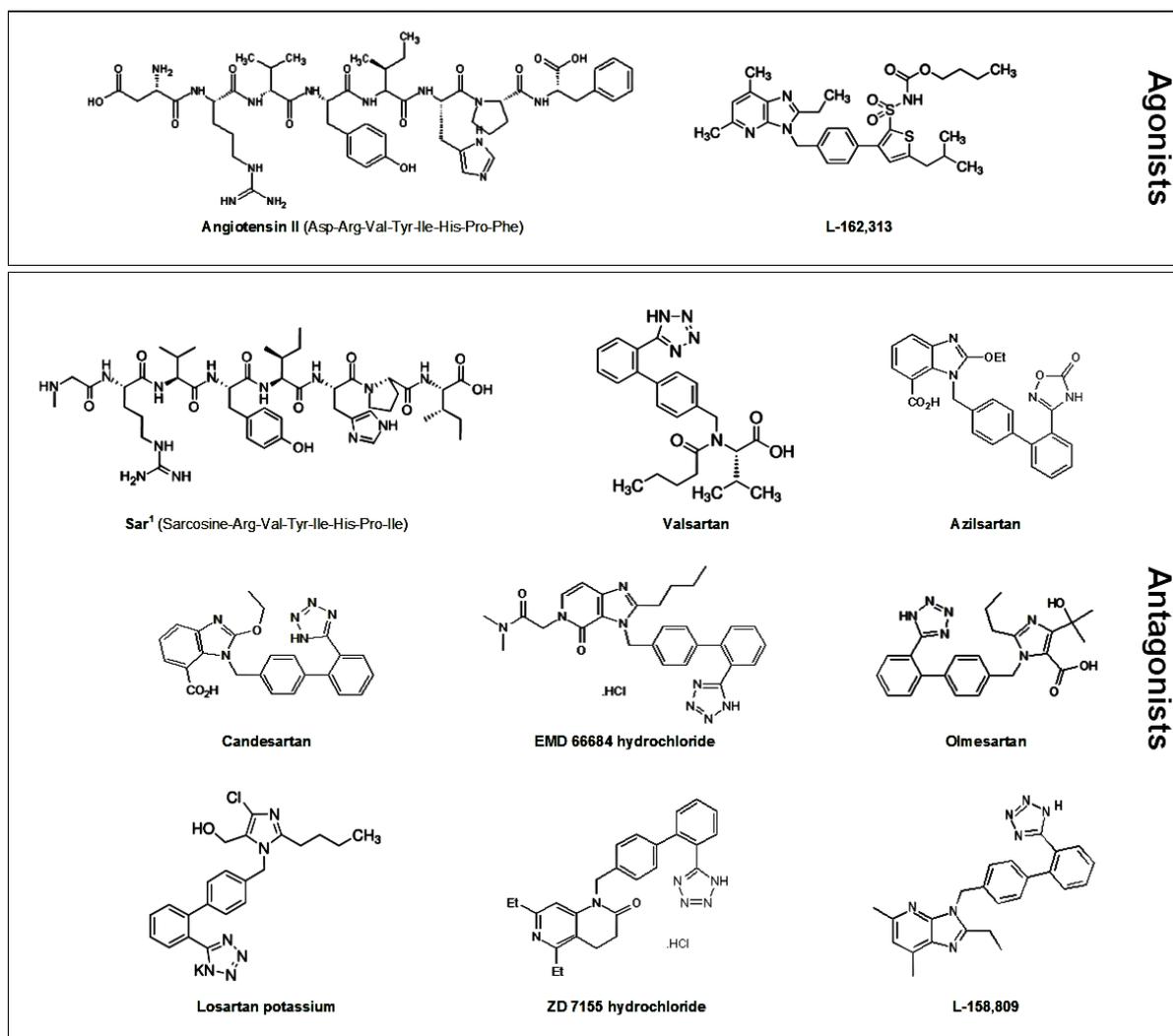


Figure 4.1 Structure of AT₁R agonists and antagonists

Structures of selected AT₁R agonists and antagonists, which are either peptides (the amino acid sequence is shown adjacent to the name) or non-peptide ligands. Figure adapted from the websites of the relevant manufacturers, see section 4.2.1.

Table 4.1 Binding affinities of selected AT₁R agonists and antagonists

Ligand	Action	Binding affinity (nM) † (species)	Reference
Angiotensin II	Agonist	0.9 (human)	156
L-162,313	Agonist	23.5 (rat)	157
Sar ¹	Antagonist	0.2 (human)	158
Valsartan	Antagonist	1.4 (rat)	159
Azilsartan	Antagonist	3.1 (rat)	160
Candesartan	Antagonist	2.8 (human)	117
EMD 66684 HCl	Antagonist	0.7 (rat); IC ₅₀	161
Olmesartan	Antagonist	2.3 (rat)	162
Losartan K	Antagonist	6.7 (human)	117
ZD 7155 HCL	Antagonist	3.8 (guinea pig); IC ₅₀	163
L-158,809	Antagonist	0.7 (rat)	164

† Binding affinities are apparent K_Ds unless otherwise stated

4.2 Methods

4.2.1 Materials

Sar¹Ile⁸ angiotensin II (Sar¹) was purchased from Source Biosciences and L-162,313 was purchased from Sigma. Angiotensin II, valsartan, azilsartan, candesartan, EMD 66684 hydrochloride, olmesartan, losartan potassium and ZD 7155 hydrochloride were purchased from Tocris Biosciences. L-158,809 and Sar¹Val⁵L-Br₅Phe⁸ were kindly provided by Emanuel Escher (Université, Sherbrooke). Molecular weight standards for SEC were purchased from Sigma and GE Healthcare.

4.2.2 Analytical FSEC

The FSEC method described in section 2.2.12 was adapted for analytical use with the following alterations. To control for expression of AT₁R-GFP, cells of the same passage number were induced, harvested, aliquoted and analysed together. Cells were incubated at room temperature with the ligand indicated for 1 hour prior to solubilisation on ice with the detergent indicated. In the condition where no ligand was used, cells were incubated at room temperature for 1 hour before being solubilised on ice with the detergent indicated.

Approximately 5 million iHEK-AT₁R-GFP cells at 10 million cells per ml (i.e. 500 µl of cell suspension) were solubilised in the detergent indicated and loaded into a 200 µl sample loop.

4.2.3 Gel filtration column calibration

Molecular weight standards were separated on the Superdex 200 10/300 GL using a 200 µl sample loop according to the manufacturer's instructions. The partition coefficient (K_{av}) was calculated as per the equation below for each standard.

$$K_{av} = (V_e - V_o) / (V_t - V_o)$$

V_e = elution volume of the protein standard, V_o = void volume and V_t = total column volume. V_o was determined experimentally by measuring the V_e of blue dextran. V_t was obtained from the manufacturer. Linear regression was then used to establish a line of best fit relating the K_{av} and \log_{10} of the molecular weights.

4.3 Results

4.3.1 Development of an analytical FSEC-based ligand screen

Since AT₁R is a medically relevant receptor, over twenty ligands that bind with nanomolar affinity have been developed. To determine which of these ligands was the best candidate for co-crystallisation with AT₁R, an analytical FSEC-based ligand screen was developed. Points of optimisation for the assay included controlling the expression levels of AT₁R in the iHEK415 clonal cell line, establishing the best storage condition for the harvested cells and determining the optimum size of the sample loop to use. Expression levels of AT₁R in the clonal cell line iHEK415 varied by passage number and the frozen stock used (Figure 4.2). This variation in expression level could have easily resulted in a change in peak height on FSEC. To prevent this, cells of the same stock and passage number were induced, harvested, stored in aliquots and analysed in parallel.

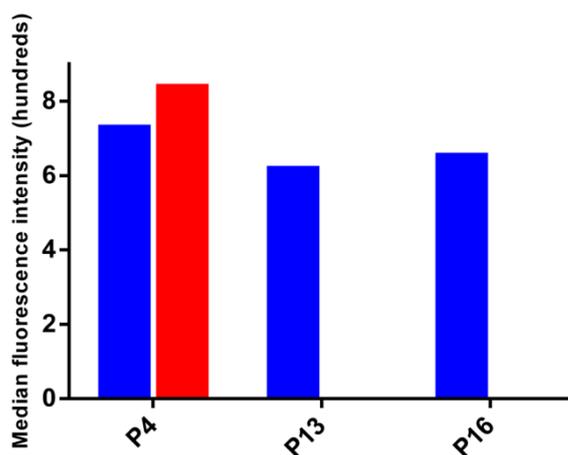


Figure 4.2 Variation in expression of AT₁R-GFP-H₁₀ in the clonal cell line iHEK415

FACS analysis of GFP expression in the stable cell line iHEK415 induced for 24 hours with tetracycline. Median fluorescence intensities are shown for cells grown continuously for either 4 passages (P4), 13 passages (P13) or 16 passages (P16). One set of cells was passaged 16 times (blue bars) and a separate set of cells was passaged 4 times (red bar). One passage is defined as when the cells have become confluent, are then resuspended and placed at a low dilution into new media, which normally occurs every 2-3 days.

To establish the optimum storage temperature for cells expressing AT₁R-GFP-H₁₀, after induction and harvest cells were stored at either -20°C or -80°C for approximately 130 days before being analysed by FSEC. AT₁R-GFP-H₁₀ was found to be more stable when stored at -80°C in comparison to storage at -20°C, as indicated by both the reduction in fluorescence signal and a shift towards the void volume as measured by FSEC (Figure 4.3). Therefore all further cell suspensions were flash frozen in liquid nitrogen and stored at -80°C. Another point of optimisation was the size of the sample loop. In order to perform reproducible analytical gel filtration it was important that both the sample loop was overloaded (leaving no dead volume) and that the fluorescence intensity was not saturated so that the peak could

CHAPTER 4 Assessment of AT₁R Ligands for Stabilisation

be evaluated in its entirety. It was found that a 200 μ l sample loop fulfilled both of these criteria better than the 500 μ l sample loop and therefore this was used for all further FSEC experiments (Figure 4.4).

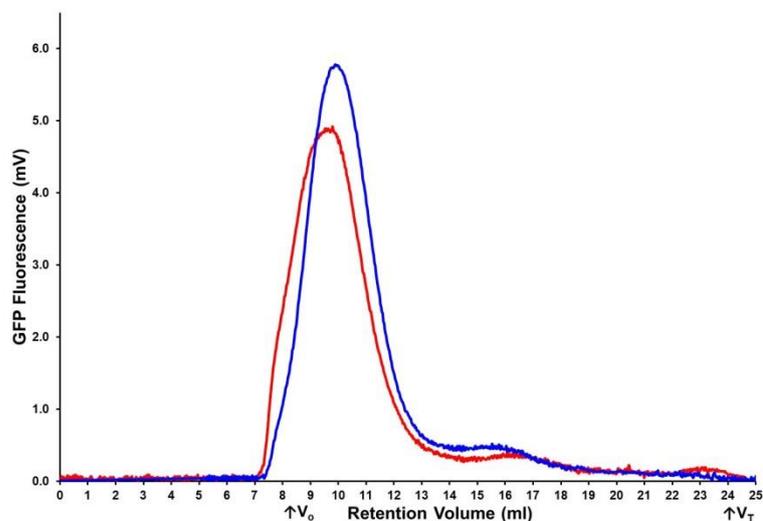


Figure 4.3 Optimisation of AT₁R storage conditions

The clonal cell line iHEK415 (expressing AT₁R-GFP-H₁₀) was induced with tetracycline for 24 hours and stored at either -20°C (red line) or -80°C (blue line) for approximately 130 days before analysis by FSEC. The elution of iHEK(AT₁R-GFP-H₁₀) was detected using GFP fluorescence (mV). The void (V_0) and total column volumes (V_T) are indicated.

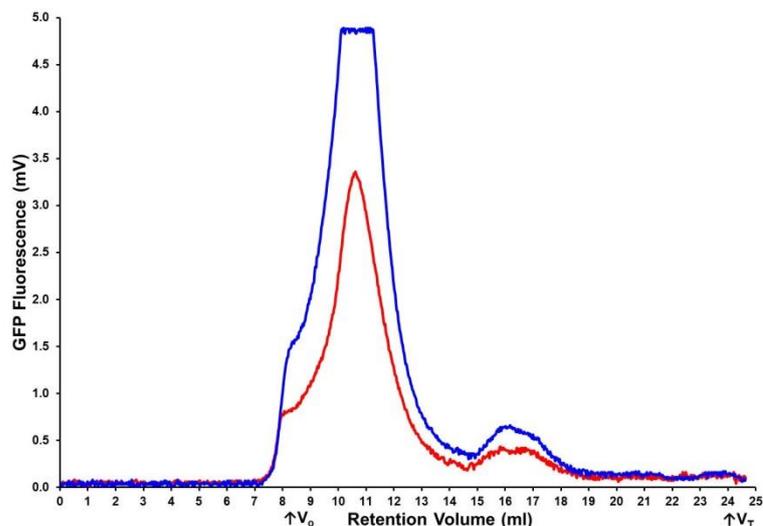


Figure 4.4 Optimisation of sample loop size for analytical FSEC

The clonal cell line iHEK415 (expressing AT₁R-GFP-H₁₀) was induced with tetracycline for 24 hours and analysed by FSEC using either a 200 μ l sample loop (red line) or a 500 μ l sample loop (blue line). The elution of iHEK(AT₁R-GFP-H₁₀) was detected using GFP fluorescence (mV). The void (V_0) and total column volumes (V_T) are indicated.

CHAPTER 4 Assessment of AT₁R Ligands for Stabilisation

4.3.2 Results of FSEC-based ligand screen

The analytical FSEC-based ligand screen was used to determine which ligands, when bound to AT₁R, produced a monodispersed FSEC peak. To achieve this, ten ligands were evaluated and compared to the receptor in the absence of ligand (Figure 3.1). All ligands were added to the AT₁R cell suspension at a concentration higher than their respective binding affinities (Table 4.1) and incubated for one hour at room temperature, then solubilised in DDM for 1 hour. Insoluble material was removed by centrifugation and the samples were analysed by FSEC. AT₁R in the absence of ligand gave a symmetrical peak on FSEC, but it was noticeable that none of the receptor-ligand complexes produced a peak of similar size or symmetry (Figure 4.5). The antagonist Sar¹ was first tested since it is the same ligand that was used for the radioligand binding assays. This was used at a concentration of 40 nM, which was approximately 200 times apparent K_D. The peak produced by this concentration maintained its symmetry although it showed a reduction in intensity and an increase in the void area, in comparison to the no ligand condition. It was previously thought that the more concentrated the ligand was in solution the more stable the receptor would be; therefore Sar¹ was tested at a concentration of 100 µM. At this higher concentration, the peak shifted to the left, decreased in intensity and became less symmetrical in comparison to 40 nM Sar¹, which are all indications that the receptor was less stable in 100 µM Sar¹ compared to 40 nM Sar¹. Sar¹ is known to exhibit partial agonist activity⁷⁶, therefore a ligand which exhibits only inverse agonism, azilsartan^{92, 160}, was tested at two concentrations, 40 nM and 100 µM. Neither concentration of azilsartan gave a symmetrical peak as observed in the absence of ligand; however 100 µM azilsartan gave an increase in peak height in comparison to 40 nM. The endogenous agonist angiotensin II, which was tested at a concentration of 40 nM, showed an increase in the void area in comparison to the no ligand condition as did the synthetic agonist L-162,313, which also showed a reduction in intensity in comparison to the no ligand condition. Since none of the ligands tested produced as symmetrical and intense a peak as the no ligand condition, six other antagonists were tested at the higher concentration, 100 µM. None of these additional ligands gave a peak as symmetrical as observed in the absence of ligand and showed varying amounts of aggregated receptor in the void volume. However, out of these ligands the peak produced from losartan-bound AT₁R was the most symmetrical and had the greatest intensity. The shift in retention volume at the greatest peak height from the no ligand condition is further discussed in terms of the change in molecular weight in section 4.3.3. In order to determine whether the length of time that azilsartan had to bind to AT₁R was affecting the results, AT₁R was incubated at room temperature with

CHAPTER 4 Assessment of AT₁R Ligands for Stabilisation

azilsartan for either 1 or 2 hours prior to solubilisation. Increasing the amount of time that azilsartan had to bind to AT₁R did not seem to dramatically increase the signal intensity or result in a more symmetrical peak and incubation for 1.5 or 2 hours seemed to increase the amount of free GFP, possibly indicating that the AT₁R-GFP fusion protein was being degraded (Figure 4.6).

Since none of the ligands tested showed a peak as symmetrical and as intense as the no ligand condition, two experimental ligands, L-158,809 and Sar¹Val⁵L-Br₅Phe⁸ angiotensin II, were also screened by analytical FSEC. L-158,809 was chosen because it was an unsurmountable antagonist *i.e.* it showed slow dissociation from AT₁R. The peptide antagonist Sar¹Val⁵L-Br₅Phe⁸ angiotensin II was chosen because it has neither G_q nor β-arrestin stimulation properties (E. Escher, personal communication). Both experimental ligands were insoluble at a final concentration of 100 μM, therefore they were tested at 40 nM and at 1 μM. Both L-158,809 and Sar¹Val⁵L-Br₅Phe⁸ angiotensin II gave a more symmetrical peak than azilsartan, however there was a reduction in signal in comparison to the no ligand condition which was more pronounced at a concentration of 1 μM (Figure 4.7).

The data from the ligand screen were unusual from two perspectives. First, despite the high affinity of the ligands, adding a ligand did not improve the characteristic of the FSEC peak compared to when no ligand was bound, and in fact the quality usually got worse. Second, the high concentrations of ligands apparently led to a decrease in the ‘quality’ of the detergent-solubilised AT₁R, as defined by peak height and peak symmetry. Subsequently, several concentrations of Sar¹ from 2 nM to 100 μM were analysed by FSEC (Figure 4.8) to identify if there was an ideal concentration of ligand to use. The Sar¹ concentrations of 2 nM, 10 nM and 100 nM gave a peak height similar to the no ligand condition, although there was a slight shift in the retention volume for the greatest peak height for ligand concentrations of 10 nM and 100 nM in comparison to the no ligand condition (Figure 4.8). The significance of these changes in peak height was assessed by a radioligand binding assay and it was found that DDM-solubilised AT₁R in the presence of either 40 nM or 100 nM Sar¹ gave similar levels of specific binding (Figure 4.8). This suggested that either the differences in peak height in the FSEC experiments were not significant or that another parameter uncontrolled in the FSEC experiments was an important factor which needed to be considered.

CHAPTER 4 Assessment of AT₁R Ligands for Stabilisation

Given the previous FSEC results, another parameter that could affect the FSEC result was tested *i.e.*, the time during which the receptor was exposed to the ligand. AT₁R in membranes was incubated with Sar¹ for 1 hour at room temperature, then solubilised in DDM for approximately 16 hours, insoluble material was removed by centrifugation and the samples were analysed by FSEC. It was found that there was a dramatic reduction in signal for the condition which was solubilised with Sar¹ overnight in comparison to the no ligand condition (Figure 4.9). However, what was striking was that once more the binding data did not change when measured over a 45 hour period (Figure 4.9).

All of the above assays were performed using DDM to solubilise AT₁R. The effect of using a different detergent, LMNG, was also examined. The combination of AT₁R with LMNG seemed to produce aggregation (Figure 4.10). It appeared that the effects of azilsartan binding to AT₁R were mitigated when the detergent LMNG was used, both with different concentrations and overnight incubations, however, without further study, it was impossible to tell whether the effects seen were because of the use of LMNG or azilsartan.

CHAPTER 4 Assessment of AT₁R Ligands for Stabilisation

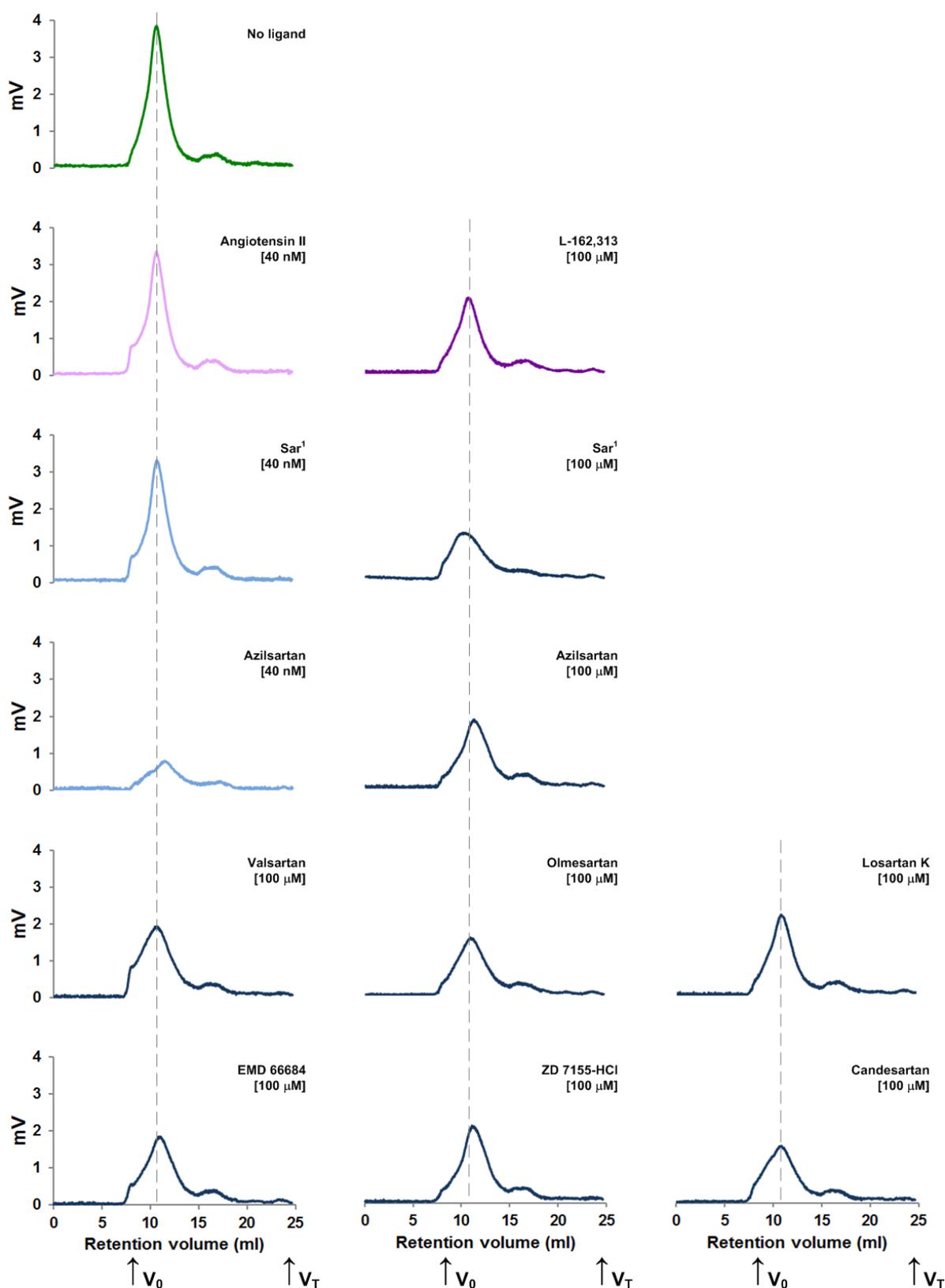


Figure 4.5 Analytical FSEC-based ligand screen

The clonal cell line iHEK415 (expressing AT₁R-GFP-H₁₀) was induced with tetracycline for 24 hours, incubated with the ligand at the concentration indicated (or no ligand was added), solubilised in DDM and evaluated by analytical FSEC using a 200 μ l sample loop. The elution of iHEK(AT₁R-GFP-H₁₀) was detected using GFP fluorescence (mV). Dashed lines indicate the middle of the 'no ligand' condition peak (green line). Agonists (purple lines) and antagonists (blue lines) are shown. Dark lines indicate that a higher concentration of 100 μ M was used. The void (V_0) and total column volumes (V_T) are indicated.

CHAPTER 4 Assessment of AT₁R Ligands for Stabilisation

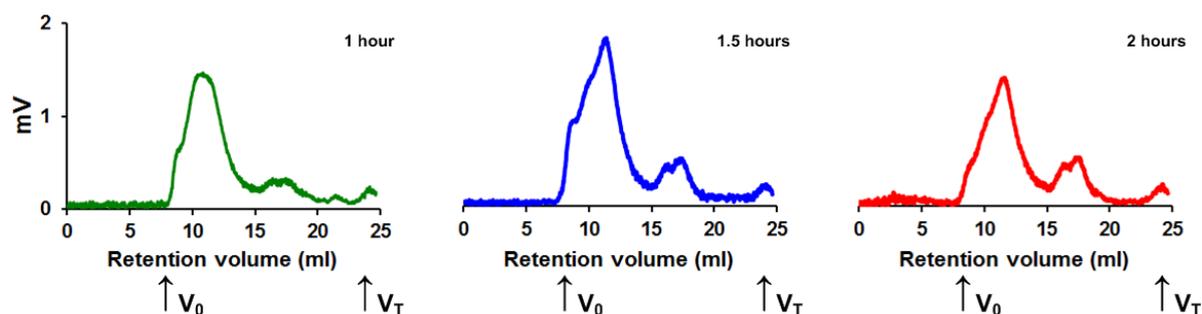


Figure 4.6 Optimisation of azilsartan incubation time

The clonal cell line iHEK415 (expressing AT₁R-GFP-H₁₀) was induced with tetracycline for 24 hours and prior to analysis by FSEC, incubated with 1 mM azilsartan for 1 hour (green line), 1.5 hours (blue line) or 2 hours (red line) at room temperature. The elution of iHEK(AT₁R-GFP-H₁₀) was detected using GFP fluorescence (mV). The void (V_0) and total column volumes (V_T) are indicated.

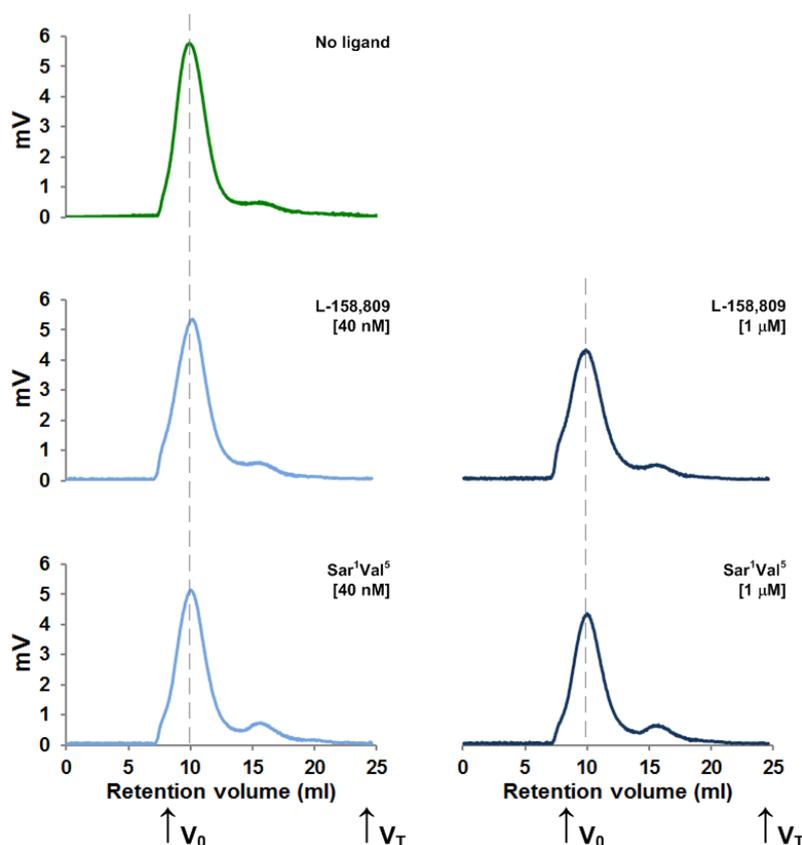


Figure 4.7 Analytical FSEC-based ligand screen with experimental ligands

The clonal cell line iHEK415 (expressing AT₁R-GFP-H₁₀) was induced with tetracycline for 24 hours, incubated with the ligand at the concentration indicated (or no ligand was added), solubilised in DDM and evaluated by analytical FSEC using a 200 μl sample loop. The elution of iHEK(AT₁R-GFP-H₁₀) was detected using GFP fluorescence (mV). Dashed lines indicate the middle of the 'no ligand' condition peak (green line). Antagonists are indicated (blue lines). Dark lines indicate that a higher concentration of 1 μM was used. The void (V_0) and total column volumes (V_T) are indicated. Sar¹Val⁵L-Br₅Phe⁸ Angiotensin II is abbreviated (Sar¹Val⁵).

CHAPTER 4 Assessment of AT₁R Ligands for Stabilisation

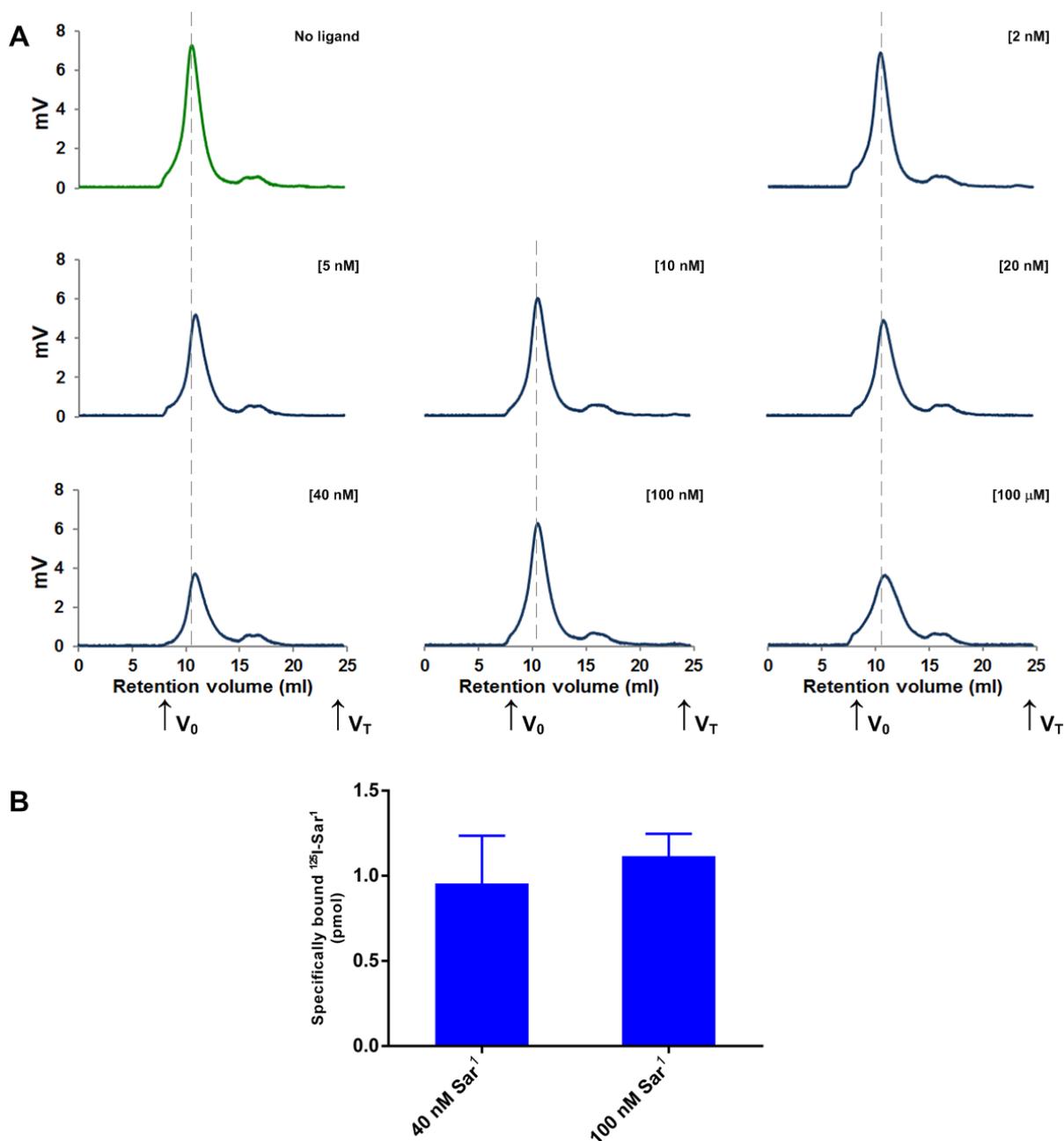


Figure 4.8 Analysis of AT₁R with different concentrations of Sar¹

(A) The clonal cell line iHEK415 (expressing AT₁R-GFP-H₁₀) was induced with tetracycline for 24 hours, incubated with Sar¹ at the concentration indicated (or no ligand was added), solubilised in DDM and evaluated by analytical FSEC using a 200 μl sample loop. The elution of iHEK(AT₁R-GFP-H₁₀) was detected using GFP fluorescence (mV). Dashed lines indicate the middle of the 'no ligand' condition peak (green line). Different concentrations of Sar¹ (blue lines) are shown. The void (V₀) and total column volumes (V_T) are indicated. **(B)** The amount of functional AT₁R in each condition was determined by measuring specific binding of the antagonist [¹²⁵I]-Sar¹ (Section 2.2.11). After the addition of ligand, membranes were solubilised in DDM and non-bound ligand was separated from receptor-ligand complex on gel filtration spin columns and measured by liquid scintillation counting. An identical number of cells (55 thousand) was used for each assay. Each data point was determined in triplicate and was plotted as mean ± SEM.

CHAPTER 4 Assessment of AT₁R Ligands for Stabilisation

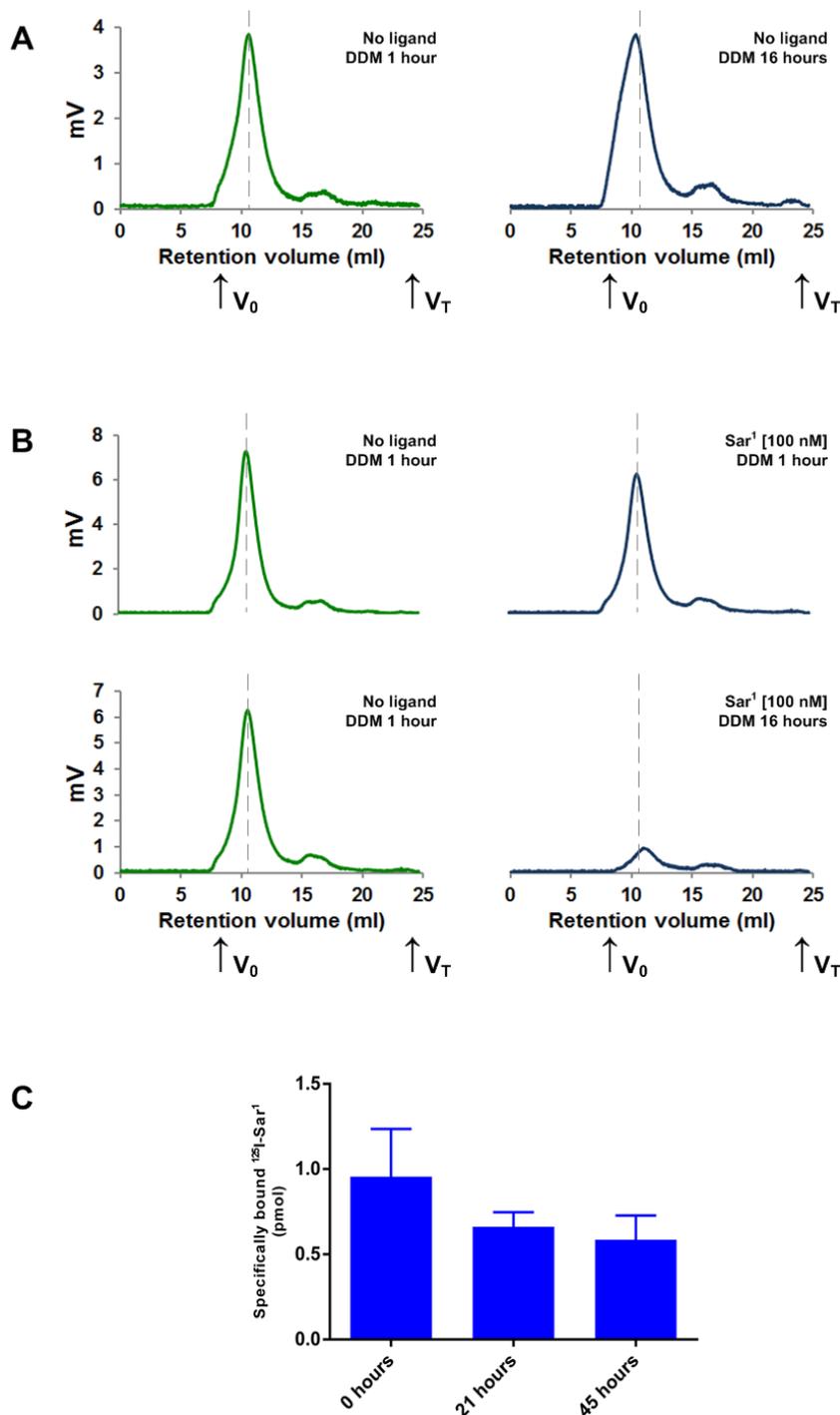


Figure 4.9 Analysis of AT₁R for different periods of time with Sar¹

The clonal cell line iHEK415 (expressing AT₁R-GFP-H₁₀) was induced with tetracycline for 24 hours, incubated with the ligand at the concentration indicated (or no ligand was added), solubilised in DDM and evaluated by analytical FSEC using a 200 μ l sample loop. The elution of iHEK(AT₁R-GFP-H₁₀) was detected using GFP fluorescence (mV). Dashed lines indicate the middle of the 'no ligand' condition peak (green line). The void (V_0) and total column volumes (V_T) are indicated. **(A)** Different solubilisation times in DDM are shown: 1 hour (green line) and 16 hours (blue line). **(B)** Different solubilisation times are shown with DDM in the presence of Sar¹: 1 hour with no ligand (green line) and 16 hours with Sar¹ (blue line). **(C)** The amount of functional AT₁R in each condition was determined by measuring specific binding of the antagonist [¹²⁵I]-Sar¹. After the addition of ligand, membranes were solubilised in DDM and non-bound ligand was separated from receptor-ligand complex on gel filtration spin columns and measured by liquid scintillation counting. An identical number of cells (55 thousand) were used per assay. Each data point was determined in triplicate and was plotted as mean \pm SEM.

CHAPTER 4 Assessment of AT₁R Ligands for Stabilisation

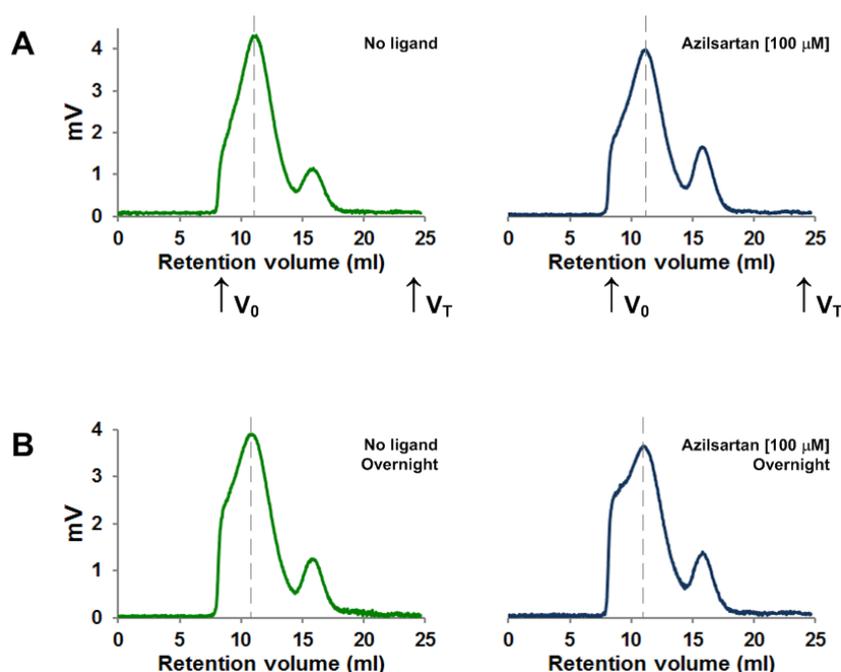


Figure 4.10 Effects of azilsartan on AT₁R is mitigated with the use of LMNG

The clonal cell line iHEK415 (expressing AT₁R-GFP-H₁₀) was induced with tetracycline for 24 hours, **(A)** incubated with no ligand (green line) or 100 μ M azilsartan (blue line), solubilised in LMNG and evaluated by analytical FSEC using a 200 μ l sample loop. **(B)** Overnight solubilisation in LMNG is shown without ligand (green line) and with 100 μ M azilsartan (blue line).

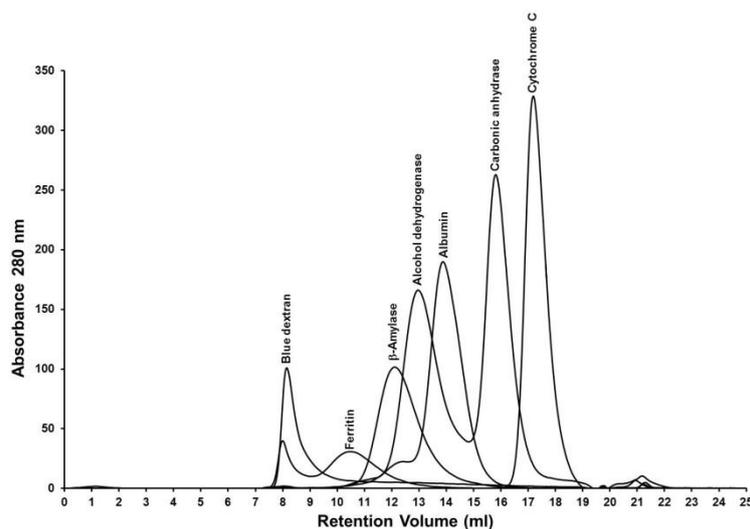
4.3.3 Change in apparent molecular weight of AT₁R as determined by analytical FSEC

In order to understand the change in retention volume at the highest point of the peak in terms of molecular weight, the column used for FSEC was calibrated with molecular weight standards (Figure 4.11). This enabled the estimation of the apparent molecular weight of AT₁R under the various conditions tested (Table 4.2). Although the column used was the most appropriate one available for the size range that AT₁R-GFP eluted, it did not separate proteins with good resolution in this range (Figure 4.11). From these data, the size of wild type AT₁R-GFP, solubilised in DDM with no ligand bound (Table 4.2; numbers 1, 14, 19, 20 and 33) varied from 490 kDa to 660 kDa with an average weight of 510 kDa. The molecular weight of AT₁R-GFP solubilised in DDM with ligand bound (Table 4.2; numbers 2-13, 15-18, 21 and 22) varied from 290 kDa to 650 kDa with an average weight of 460 kDa, which within experimental error is similar to without ligand bound. From the amino acid sequence, the estimated molecular weight for wild type AT₁R-GFP is 72 kDa (Table 4.3). The DDM detergent micelle can add approximately 100 kDa of mass^{100, 165}. The mass of the three complex N-glycans was calculated by plotting the apparent mobility of glycosylated AT₁R-GFP against the log molecular weight of the molecular weight standards on the gel from a western blot (Figure 2.11) and these were estimated to add 15-60 kDa.

CHAPTER 4 Assessment of AT₁R Ligands for Stabilisation

Therefore the macromolecular weight of the protein plus the detergent micelle and the three complex N-glycans should be about 230 kDa, which is much less than the apparent molecular mass obtained from FSEC. There are a number of possibilities for explaining these observed discrepancies in apparent molecular weight and they are discussed in Section 4.4. In contrast, AT₁R-T4 lysozyme (T4L) fusion protein (produced from the cell line iGNTI42; Chapter 5) solubilised in DDM without ligand (Table 4.2; numbers 34-36) varied in apparent weight from 230 kDa to 280 kDa with an average molecular weight of 250 kDa without ligand, whereas the condition with ligand had an apparent molecular weight of 260 (Table 4.2; number 38). Another observation from the data was that the apparent weight of AT₁R-T4L-GFP in LMNG (Table 4.2; number 31) was lower than in DDM (190 kDa compared to 250 kDa) which was possibly due to LMNG having a smaller detergent micelle. LMNG also produced a smaller apparent molecular weight for wild type AT₁R-GFP (Table 4.2; numbers 27 and 29; average 373 kDa).

A



B

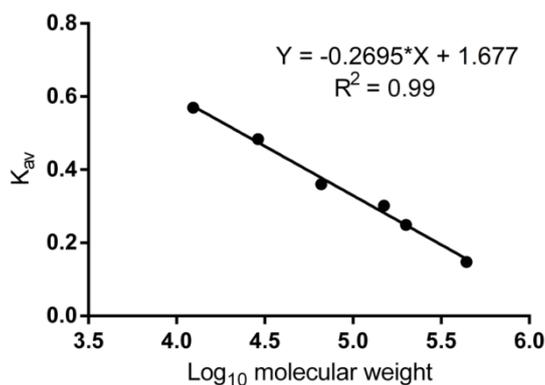


Figure 4.11 Calibration of Superdex 200 10/300 GL

(A) The Superdex 200 10/300 GL was calibrated using molecular weight standards according to the manufacturer's protocol. The retention volumes of each of the protein standards is as follows (molecular weights in parentheses (kDa)): blue dextran, 8.2 ml (2000); ferritin, 10.5 ml (440); β -amylase, 12.1 ml (200); alcohol dehydrogenase, 13.0 ml (150); albumin, 13.9 ml (66); carbonic anhydrase, 15.8 ml (29); cytochrome C, 17.2 ml (12.4). **(B)** Linear regression was used to establish a line of best fit relating the partition coefficient (K_{av}) and \log_{10} of the molecular weights (Section 4.2.3).

CHAPTER 4 Assessment of AT₁R Ligands for Stabilisation

Table 4.2 Estimated molecular weight of AT₁R-GFP-H₁₀

Cells expressing AT₁R-GFP-H₁₀ were solubilised in the detergent indicated and analysed by FSEC. The Superdex 200 10/300 GL was calibrated by running samples of known molecular weight (Figure 4.11). The ligand Sar¹Val⁵L-Br₅Phe⁸ angiotensin II is abbreviated (Sar¹Val⁵). The cell line iHEK415 expressed wild type AT₁R-GFP-H₁₀, whereas iGNTI25 expressed AT₁R Del.K323-E359-GFP-ZZ (Chapter 5) and iGNTI42 expressed LS-FLAG-AT₁R(1-228)-T4L-AT₁R(229-323)-TEV-AT₁R(323-359)-GFP-H₁₀ (Chapter 5); see Table 4.3 for more information. Shading is used to group samples that were measured at the same time and are therefore directly comparable and fall into the following groups: Group 1, samples 1-13; Group 2, samples 14-18; Group 3, samples 19-20; Group 4, samples 21-22; Group 5, samples 23-25; Group 6, sample 26; Group 7, samples 27-28; Group 8, 29-32; Group 9, samples 33-38.

Index	Cell line (Na butyrate)	Ligand	Detergent (solubilisation time)	Peak height (mV)	Peak height retention volume (ml)	Estimated molecular weight (kDa)
1	iHEK415	No ligand	DDM (1 hour)	3.85	10.32	520
2	iHEK415	40 nM Angiotensin II	DDM (1 hour)	3.36	10.62	440
3	iHEK415	100 µM L-162,313	DDM (1 hour)	2.05	10.62	440
4	iHEK415	40 nM Sar ¹	DDM (1 hour)	3.31	10.59	450
5	iHEK415	100 µM Sar ¹	DDM (1 hour)	1.29	10.02	610
6	iHEK415	40 nM Azilsartan	DDM (1 hour)	0.80	11.44	290
7	iHEK415	100 µM Azilsartan	DDM (1 hour)	1.86	11.23	320
8	iHEK415	100 µM Candesartan	DDM (1 hour)	1.54	10.77	410
9	iHEK415	100 µM Valsartan	DDM (1 hour)	1.94	10.64	440
10	iHEK415	100 µM EMD 66684	DDM (1 hour)	1.85	10.88	390
11	iHEK415	100 µM ZD 7155 HCl	DDM (1 hour)	2.09	11.09	340
12	iHEK415	100 µM Olmesartan	DDM (1 hour)	1.56	10.85	390
13	iHEK415	100 µM Losartan K	DDM (1 hour)	2.18	10.75	410
14	iHEK415	No ligand	DDM (1 hour)	5.78	9.89	660
15	iHEK415	40 nM Sar ¹ Val ⁵	DDM (1 hour)	5.15	10.03	610
16	iHEK415	1 µM Sar ¹ Val ⁵	DDM (1 hour)	4.37	10.00	620
17	iHEK415	40 nM L-158,809	DDM (1 hour)	5.36	10.14	570
18	iHEK415	1 µM L-158,809	DDM (1 hour)	4.35	9.90	650
19	iHEK415	No ligand	DDM (1 hour)	7.28	10.45	490
20	iHEK415	No ligand	DDM (overnight)	3.85	10.35	510
21	iHEK415	100 nM Sar ¹	DDM (1 hour)	6.28	10.51	470
22	iHEK415	100 nM Sar ¹	DDM (overnight)	0.95	10.93	380
23	iGNTI25	40 nM Sar ¹	DDM (1 hour)	1.83	10.93	380
24	iGNTI25 (1 mM)	40 nM Sar ¹	DDM (1 hour)	3.77	11.12	340
25	iGNTI25 (5 mM)	40 nM Sar ¹	DDM (1 hour)	2.90	10.99	360
26	iHEK(AT ₁ R-GFP-H ₁₀)	10 µM Angiotensin II	Digitonin (1 hour)	0.99	11.14	340
27	iHEK415	No ligand	LMNG (overnight)	3.90	10.69	430
28	iHEK415	100 µM Azilsartan	LMNG (overnight)	3.65	10.83	400
29	iHEK415	No ligand	LMNG (1 hour)	4.33	11.23	320
30	iHEK415	100 µM Azilsartan	LMNG (1 hour)	3.99	11.17	330
31	iGNTI42	No ligand	LMNG (1 hour)	0.74	12.16	190
32	iGNTI42	100 µM Azilsartan	LMNG (1 hour)	0.85	12.16	190
33	iHEK415	No ligand	DDM (1 hour)	5.66	10.83	400
34	iGNTI42	No ligand	DDM (1 hour)	0.38	11.49	280
35	iGNTI42 (2.5 mM)	No ligand	DDM (1 hour)	0.62	11.83	230
36	iGNTI42 (5 mM)	No ligand	DDM (1 hour)	0.92	11.68	250
37	iGNTI42	No ligand	DDM + CHS (1 hour)	1.06	11.63	260
38	iGNTI42	1 µM Sar ¹	DDM (1 hour)	0.41	11.60	260

CHAPTER 4 Assessment of AT₁R Ligands for Stabilisation

Table 4.3 Molecular weights of proteins expressed from three cell lines

Cell line name	Recombinant protein expressed	Estimated molecular weight of the protein (kDa)	N-linked glycosylation sites (type of sugars)
iHEK415 *	AT ₁ R-GFP-H ₁₀	72	3 (complex glycosylation)
iGNTI25 †	AT ₁ R(Del.K323-E359)-GFP-ZZ	86	3 (high mannose)
iGNTI42 †	LS-FLAG-AT ₁ R(1-228)-T4L-AT ₁ R(229-323)-TEV-AT ₁ R(324-359)-GFP-H ₁₀	93	3 (high mannose)

* The clonal cell line iHEK415 is discussed in detail in Chapter 3. † Cell lines iGNTI25 and iGNTI42 are discussed in detail in Chapter 5.

4.4 Discussion

Given the availability of several ligands which tightly bind to AT₁R, an FSEC-based ligand screen was performed to determine which of these would best stabilise the receptor. It was found that none of the receptor-ligand complexes tested had an FSEC peak as symmetrical and as intense as the condition with no ligand. One possibility was that while the ligands might be an inverse agonist for one pathway (such as G protein coupling) they may act as an agonist for another pathway (such as β -arrestin coupling). In the active conformation that couples to either a G protein or β -arrestin, GPCRs tend to be much less stable than when in the inactive conformation. Although none of the ligands tested here have been shown to be biased agonists in the literature, the effect of concentration might cause these ligands to act in such a fashion by increasing the amount of time a receptor is bound to a ligand and hence increasing the probability of a conformational change. Another possibility was that the ligands may be behaving in a manner analogous to protean agonism. This theory describes the potential for a ligand to change its activity from agonism to inverse agonism if the agonist produces a conformation that is of lower efficacy than the constitutively active conformation^{166, 167, 168}. This means that a ligand can activate receptors in the inactive state and act as an inverse agonist to receptors showing a constitutive level of activity. Protean agonism has thus far only been observed in systems with a high level of basal activity¹⁶⁹ and AT₁R had been thought to be held in an inactive conformation⁸⁸. However it is difficult to determine what happens to receptors in detergent since both G protein and β -arrestin coupling assays cannot be performed on detergent-solubilised receptors. Nevertheless, the theory of protean agonism suggests that the pharmacological characteristics of a receptor are dependent on its environment. This, combined with concentration effects, might explain the behaviour of AT₁R shown here. Thermostabilisation of the receptor might enable AT₁R to be locked in one conformation diminishing or entirely abolishing these effects. Another approach would be to use T4L to stabilise the receptor, however the only AT₁R-T4L construct to be expressed at suitable levels proved to be unstable as assayed by FSEC (Chapter 5).

From the amino acid sequence, the molecular weight of AT₁R-GFP-H₁₀ was predicted to be 72 kDa, however, as determined by FSEC, the average observed molecular weight of AT₁R-GFP-H₁₀ without ligand bound was 510 kDa. This discrepancy could potentially be explained by the increase in the Stokes radius of AT₁R-GFP due to the presence of three complex N-glycans, which could appreciably affect the mobility of the protein on SEC and

CHAPTER 4 Assessment of AT₁R Ligands for Stabilisation

therefore cause it to elute at a lower volume than expected. Interestingly, the cell line iGNTI25 (Chapter 5), which has all three glycosylation sites intact but only produced the mannose core, had an average molecular weight of 350 kDa with ligand bound. Since this weight was also much larger than what was predicted from the amino acid sequence, glycosylation alone does not explain the increase in weight. Another possibility for the increase in molecular weight was the detergent micelle. However, the detergent micelle from DDM only contributes approximately 100 kDa of mass^{100, 165}. Using a detergent assay¹⁷⁰ on the purified protein would determine precisely how much weight the detergent micelle added. Similarly, repeating FSEC on the deglycosylated receptor could establish the contribution of N-glycans to the weight of AT₁R, but using PNGase F on unpurified material is prohibitively expensive. If these assays confirm that AT₁R is still much larger than predicted, two theories might explain this. First, AT₁R might have either G proteins or β -arrestin still bound which could add over 100 kDa or 50 kDa respectively. A high salt wash (1 M or above) of the membranes prior to solubilisation would disrupt any protein-protein interactions and allow for this theory to be tested. Second, AT₁R-T4L in DDM had an apparent molecular weight much closer to the estimated weight of AT₁R-GFP in DDM with three complex N-glycans. Since the apparent molecular weight of AT₁R-T4L was half that of AT₁R without the T4L fusion, it is possible that AT₁R exists in detergent solution as a dimer and T4L prohibits the formation of AT₁R dimers. If AT₁R is normally a dimer, this might also explain why the FSEC signal was very low for all of the AT₁R-T4L-GFP conditions. This could be because detergent-solubilised AT₁R might be significantly more stable in the dimeric form or it might prevent the binding of a G protein.

In conclusion, the most symmetrical FSEC peak observed was with no ligand bound. Further assays needed to be performed on the purified material, therefore the next steps were to express AT₁R in large scale and to purify the receptor.

CHAPTER 5 ENGINEERING AT₁R FOR USE IN STRUCTURAL STUDIES

5.1 Introduction

The use of FACS to select highly expressing clonal stable cell lines increased the expression of AT₁R two-fold in comparison to the stable polyclonal cell line. Although this increase in expression was substantial, it was not without its problems. For example, the FACS sort method takes 2-3 months to complete on top of the 2 months required to create the parental polyclonal stable cell line, therefore other methods of increasing expressing were investigated. Additionally, it was important to establish whether AT₁R would benefit from further engineering in order to make it more amenable for crystallography. Specifically, it was unknown whether AT₁R contained flexible regions which might hinder crystal formation and whether the strategy of adding a small soluble protein, such as bacteriophage T4 lysozyme (T4L), would improve the stability of AT₁R and increase the likelihood of crystal contacts forming.

It has been observed that there is a charge-bias in membrane proteins, with more positively charged residues in the cytoplasmic loops compared to the extracellular loops^{171, 172}. The positive-inside rule indicates that short protein segments (less than 70-80 residues long) containing Arg and Lys residues are translocated across the plasma membrane two to four times less frequently than segments not containing those residues¹⁷². Therefore, insertion of GPCRs into the plasma membrane might be negatively impacted if positively charged residues were present on the N-terminus. Since AT₁R contained four positively charged residues on its N-terminus (Lys12, Arg13, Lys20 and Arg23) this might explain why expression of AT₁R was low. Another strategy to increase expression of membrane proteins with an extracellular N-terminus is the inclusion of a leader sequence (LS). LSs are short amino acid sequences which are generally non-conserved and are comprised of a hydrophobic core flanked by polar amino acids¹⁷³. During biosynthesis of a membrane protein, the ribosome is first targeted to the membrane where it docks to the heterotrimeric Sec61 complex which efficiently exports the C-terminal portion of a LS to the endoplasmic reticulum lumen, hence resulting in the translocation of the N-terminus of the GPCR¹⁷⁴. The use of a LS for the expression of AT₁R in the baculovirus system doubled the amounts of

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

active AT₁R (Section 2.4), therefore it was worth investigating its utility in the mammalian system.

Flexible regions in a protein often present potential obstacles for crystallography. A common approach to removing flexible regions of a protein is the use of limited proteolysis on the purified protein, however this method is not always 100% reproducible and can lead to protein heterogeneity. Another method to remove flexible regions from a protein is the use of site-directed mutagenesis to produce suitable truncations. For GPCRs, non-conserved regions are potentially flexible and an alignment of the receptor of interest can be performed with GPCRs of known structure to determine these areas. In particular, regions in the N-terminus, the C-terminus and CL3 are often flexible. For AT₁R, both the N-terminus and CL3 are short, however it does contain a long non-conserved C-terminus (Appendix 1). Additionally, algorithms such as the regional order neural network (RONN) software¹⁷⁵ can be used to predict disordered regions. For the β_1 -adrenergic receptor truncations at the C-terminus and CL3 were necessary in order to obtain well-diffracting crystals¹⁶⁵. Using CHO cell lines stably expressing C-terminal deletions of rat AT₁R cells Conchon *et al.*¹⁷⁶ showed that deletions of the C-terminus of AT₁R up to residue 314 exhibited a similar affinity for angiotensin II as the wild type receptor and did not change expression levels. However, Gaèborik *et al.*¹⁷⁷ also examined the effect of C-terminal deletions on the rat AT₁R up to residue 309 and found that, while binding affinities were unchanged, expression of the receptor decreased with increasing the length of the truncation and truncation at residue 309 reduced expression to 4% of the wild type receptor. If AT₁R is to be truncated to remove flexible regions, the impacts on expression and ligand binding need to be carefully monitored.

A common strategy for increasing the hydrophilic area of GPCRs, as well as potentially stabilising them, is to create a fusion with a small soluble protein such as bacteriophage T4 lysozyme (T4L). This method was successfully employed to crystallise the β_2 AR and resulted in a 2.4 Å resolution structure^{24, 41}. The crystals were grown in lipidic cubic phase (LCP), because vapour diffusion crystal trials did not result in the formation of well-diffracting crystals. While the addition of a fusion protein can introduce flexibility, the specific placement of lysozyme in CL3 was successful in reducing flexibility in this region. T4L has since been used to successfully produce numerous high-resolution GPCR structures and it was therefore worthwhile investigating this strategy for the crystallisation of AT₁R.

5.2 Materials and Methods

5.2.1 Materials

IgG Sepharose 6 FastFlow and Q Sepharose were purchased from GE Healthcare. Anti-FLAG M2 affinity gel and Concanavalin A resin were purchased from Sigma.

5.2.2 Further alterations of AT₁R

Site directed mutagenesis (Section 2.2.3) was used to remove positive residues from the N-terminus of AT₁R expressed from the plasmid pJAP2 (Section 2.2.3). A leader sequence (LS) based on the 5HT_{3A} serotonin receptor (MRLCIPQVLLALFLSMLTGPGECS) was inserted before the sequence of AT₁R. In order to label AT₁R the Snap tag sequence (New England Biolabs) was added to its N-terminus. A sequence encoding a tandem IgG binding domain (ZZ) based on *S. aureus* protein A¹⁷⁸ was added to the C-terminus of AT₁R-GFP to enable affinity purification. The sortase sequence (LPETGGGRR)^{179, 180, 181, 182} was added to the truncated C-terminus of AT₁R to facilitate the attachment of peptides to AT₁R. The tobacco etch virus (TEV) recognition sequence (ENLYFQG) was added to promote site specific proteolysis^{183, 184}. The nucleotide sequences for the inserts in plasmids pJAP24, 25 and 32 were synthesised by IDT and cloned into the EcoRI/XbaI sites of pcDNA4/TO. Cysteine-free bacteriophage T4 lysozyme (N2-Y161 with the mutations C54T and C97A)⁷⁰ was codon optimised for expression in mammalian cells and inserted between residues Ile228 and Gln229 of AT₁R. See Table 5.1 for a full description of the plasmids used.

5.2.3 Large scale suspension culture of mammalian cells

For large scale suspension culture of mammalian cells, 2 litres of cells at a density of 0.5-1 million cells per ml were added to a 10-litre Cellbag (GE Healthcare), connected to a WAVE Bioreactor (GE Healthcare), incubated at 37°C with 5% CO₂ and rocked at 20 r.p.m. with an angle of 6°. The cell density was checked daily and the culture was expanded by diluting it to 1 million cells per ml with fresh media when the cells reached a density over 1.5 million cells per ml until the desired total volume was reached (i.e. 10 litres). The r.p.m. were gradually increased to 30 and the angle to 7° as the culture was expanded. Cells were induced at a density of 1-2 million cells per ml by adding tetracycline to a final concentration of 1 µg/ml and incubating at 37°C for 24 h.

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

Table 5.1 AT₁R alterations for mammalian expression

Plasmid name	Inserts and mutation(s) †	Vector backbone
pJAP4	LS-Snap tag- AT ₁ R-GFP-H ₁₀	pcDNA4/TO
pJAP17	AT ₁ R(K12Q)-GFP-H ₁₀	pcDNA4/TO
pJAP18	AT ₁ R(R13Q)-GFP-H ₁₀	pcDNA4/TO
pJAP19	AT ₁ R(K20Q)-GFP-H ₁₀	pcDNA4/TO
pJAP20	AT ₁ R(R23Q)-GFP-H ₁₀	pcDNA4/TO
pJAP22	LS-AT ₁ R-GFP-H ₁₀	pcDNA4/TO
pJAP24	LS-FLAG tag-TEV-AT ₁ R(Del. P321-E359)-TEV-GFP-H ₁₀	pcDNA4/TO
pJAP25	LS-FLAG tag-TEV-AT ₁ R(Del. K323-E359)-Sortase-GFP-TEV-ZZ	pcDNA4/TO
pJAP26	LS- FLAG tag-TEV- AT ₁ R(N4A+N188A+Del.P321-E359)-TEV-GFP-H ₁₀	pcDNA4/TO
pJAP30	LS-FLAG tag-TEV- AT ₁ R(Del. K323-E359)-TEV-GFP-H ₁₀	pcDNA4/TO
pJAP32	LS-FLAG tag-TEV- AT ₁ R(I228-Q229 CL3-T4L-Del. K323-E359)-TEV-GFP-H ₁₀	pcDNA4/TO
pJAP40	LS-FLAG tag-TEV- AT ₁ R(I228-Q229 CL3-T4L-Del. K323-E359)-TEV-GFP-ZZ-H ₁₀	pcDNA4/TO
pJAP42	LS-FLAG tag-TEV- AT ₁ R(I-228)-T4L-AT ₁ R(229-323)-TEV-AT ₁ R(324-359)-GFP-H ₁₀	pcDNA4/TO
pJAP43	LS-FLAG tag-TEV- AT ₁ R(Del. K323-E359)-TEV-GFP-ZZ-H ₁₀	pcDNA4/TO

† The inclusion of a leader sequence (LS), the ZZ domain of protein A (ZZ) and bacteriophage T4 lysozyme (T4L) into cellular loop 3 (CL3) of AT₁R are indicated.

5.3 Results

5.3.1 Effects of N-terminal alterations and inclusion of a leader sequence on AT₁R expression in mammalian cells

The positive-inside rule indicates that short protein segments containing Arg and Lys residues are translocated across the plasma membrane less efficiently than segments not containing those residues^{171, 172}. In order to test whether this had any impact on the expression of AT₁R, the N-terminus of AT₁R was mutated. AT₁R contained four positively charged residues in its N-terminal region: Lys12, Arg13, Lys20 and Arg23. Therefore, expression of AT₁R could potentially be reduced by the presence of these positively charged amino acids. To test this theory, each of the positively charged residues was mutated to the polar uncharged residue Gln (Table 5.1). The mutated plasmids were transiently transfected into iHEK cells and analysed by both FACS (Figure 5.1) and in-gel fluorescence (Figure 5.2). There was a noticeable increase in expression of AT₁R when the positive residues on the N-terminus were mutated to Gln. In particular mutations K12Q and R23Q appeared to have the greatest impact on increasing expression (Figure 5.1).

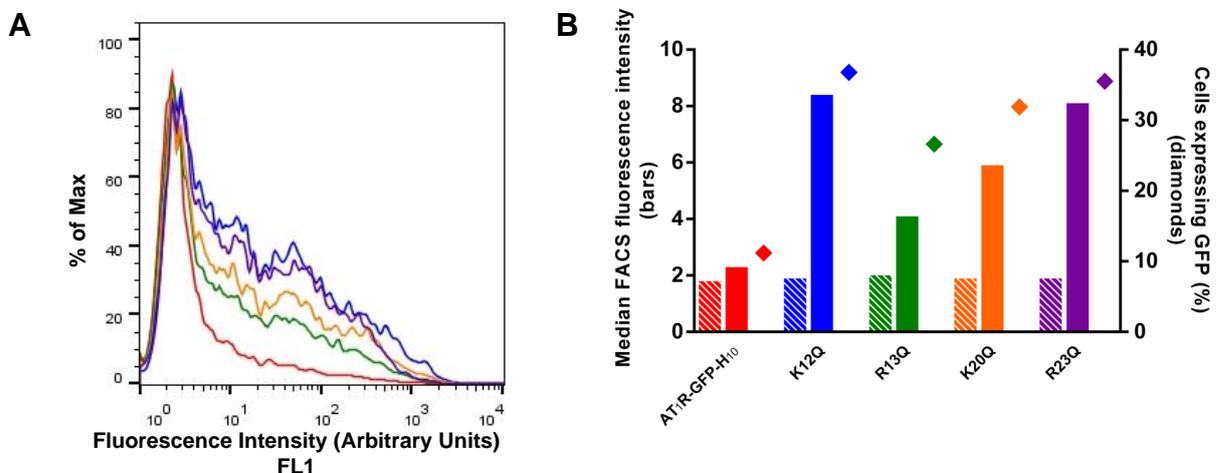


Figure 5.1 Removal of positive charges on the N-terminus of AT₁R increases expression

For FACS analysis, iHEK cells were transiently transfected with plasmids expressing either AT₁R-GFP-H₁₀ or the mutations indicated and induced for 24 hours, harvested in PBS and analysed on the FACSCalibur II for GFP fluorescence using the FL-1 detector. Cell counts have been normalised (% of Max). **(A)** Histogram of GFP fluorescence intensity. AT₁R-GFP-H₁₀ (red), AT₁R(K12Q)-GFP-H₁₀ (blue), AT₁R(R13Q)-GFP-H₁₀ (green), AT₁R(K20Q)-GFP-H₁₀ (orange) and AT₁R(R23Q)-GFP-H₁₀ (purple). **(B)** Median FACS fluorescence intensity; uninduced (hatched bars), induced (solid bars) and cells positive for GFP expression (%) (diamonds).

In-gel fluorescence of N-terminal positively charged AT₁R mutants showed that despite the alterations, they were still glycosylated and the gel band intensity roughly correlated with the percentage of positive cells (Figure 5.2). Despite these positive findings, N-terminal residues had previously been found to be important for ligand binding in peptide receptor

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

structures, such as the neurotensin receptor⁷⁰. Disruption of the ligand binding pocket would not be a favourable outcome, therefore, other means for increasing AT₁R expression were investigated.

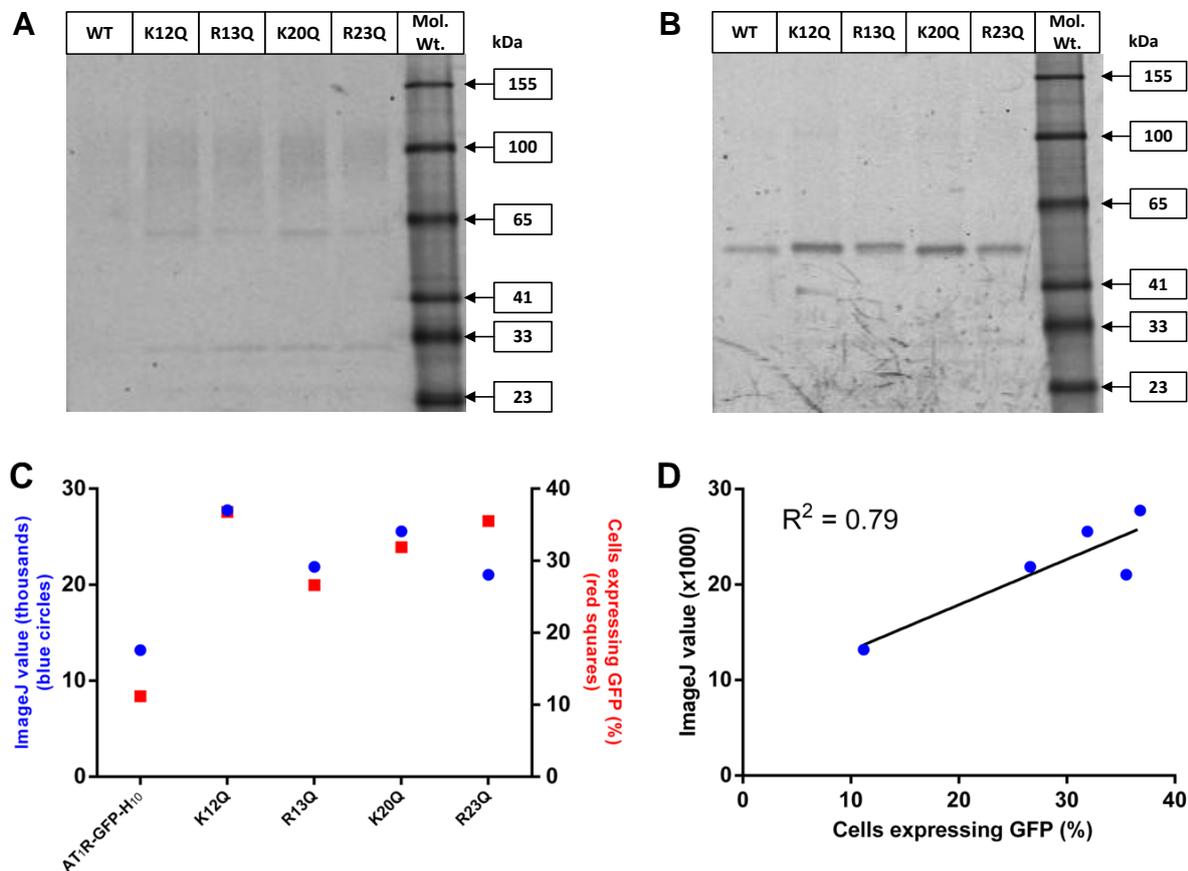


Figure 5.2 In-gel fluorescence analysis of N-terminus mutants

N-terminal mutants of AT₁R were transiently transfected into iHEK cells, induced for 24 hours with tetracycline and analysed by in-gel fluorescence. An equal amount of cells was loaded in each lane (A). N-linked glycosylation was removed by treatment with PNGase F (B). Bands corresponding to deglycosylated AT₁R were quantified by using ImageJ (blue circles) and plotted against cells expressing GFP (red squares) (C). Linear regression was used to establish a line of best fit for the percentage of positive cells and the Image J values (D).

The addition of a leader sequence (LS) in front of the AT₁R sequence doubled its expression in the baculovirus system (Section 2.4) therefore the same approach was investigated for mammalian cells. A LS based on the 5HT₃A serotonin receptor was added in front of the AT₁R sequence (Table 5.1), the plasmid was transiently transfected into iHEK cells and these were analysed by FACS (Figure 5.3). Inclusion of a LS gave an increase in FACS median of 3 to 4. Additionally, an iHEK cell line stably expressing AT₁R with a LS (iHEK4; Figure 5.4) showed a 50% increase in expression of AT₁R-GFP as measured by FACS analysis (Figure 5.3). The inclusion of a LS was an ideal way to increase expression since it did not require the 2-3 months that FACS selection took (Chapter 3), nor was it likely to interfere with ligand binding, which N-terminal mutations could have done.

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

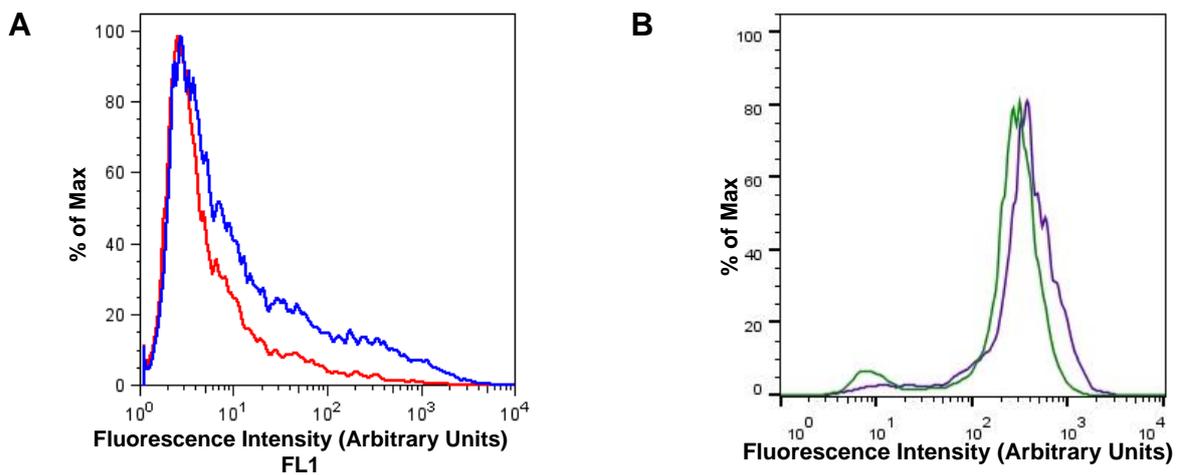


Figure 5.3 Inclusion of a leader sequence increases AT₁R expression in mammalian cells

Histogram of GFP fluorescence intensity. For FACS analysis cells were induced for 24 hours, harvested in PBS and analysed on the FACSCalibur II for GFP fluorescence using the FL-1 detector. Cell counts have been normalised (% of Max). **(A)** iHEK cells were transiently transfected with plasmids expressing either AT₁R-GFP-H₁₀, (red line; median fluorescence intensity of 3) or AT₁R-LS-GFP-H₁₀, (blue line; median fluorescence intensity of 4). **(B)** FACS analysis of the stable polyclonal cell line iHEK(AT₁R-GFP-H₁₀) (green line; median fluorescence intensity of 293) and iHEK4 (expressing LS-Snap Tag-AT₁R-GFP-H₁₀) (see Figure 5.4 for construct information) (purple line; median fluorescence intensity of 425).

Cell line name	iHEK4						
Open reading frame schematic	Leader Sequence	Snap Tag	AT₁R	Factor Xa	GFP	Strep II	H₁₀
Parental cell line	iHEK				Master Gain	1107	
Median FACS fluorescence intensity	425				Digital Gain	1.24	
					Digital Offset	0.00	

Figure 5.4 Assessment of cell line created to express AT₁R-LS-GFP-H₁₀

Assessment of the iHEK4 polyclonal cell line was based on confocal microscopy and the average counts for FACS analysis.

5.3.2 The impact of AT₁R C-terminal truncations on expression

To determine whether AT₁R had any disordered regions, which might hinder crystallisation, the amino acid sequence of human AT₁R was analysed by regional order neural network (RONN) software¹⁷⁵ (Figure 5.5). This software used an algorithm to detect potentially disordered regions in AT₁R. The N-terminus of AT₁R was predicted to be ordered, but the C-terminus, starting from Ala324, was predicted to be disordered (Figure 5.5).

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

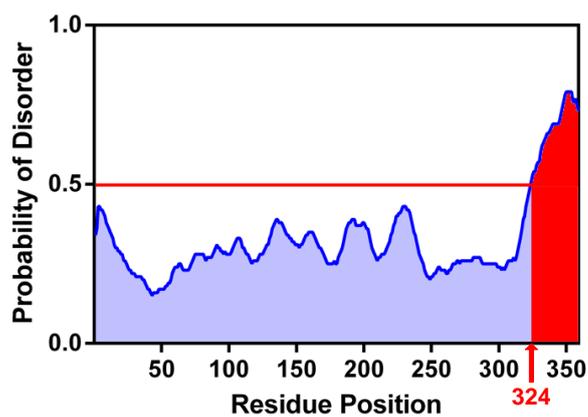


Figure 5.5 Human AT₁R probability of disorder

The amino acid sequence of human AT₁R was entered into the RONN software algorithm¹⁷⁵ to detect natively disordered segments. A probability of 0.5 or greater suggests that the residue was likely to be disordered. Human AT₁R was predicted to have a disordered C-terminus starting at alanine 324.

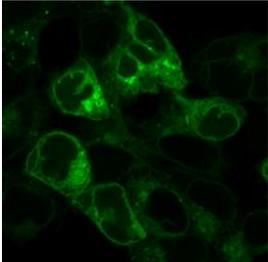
To remove the potentially disordered C-terminus, AT₁R was truncated in two different ways, either P321-E359 or K323-E359. Six new cell lines that stably expressed truncated versions of AT₁R were created (Figure 5.6). From the FACS analysis and radioligand binding data only cell lines that contained P321-P322 showed favourable expression (Figure 5.7). Cell lines iHEK25 and iGNTI25, which contained P321-P322, showed a 19-fold increase in molecules of AT₁R per cell in comparison to cell lines iHEK24 and iHEK26, both of which did not contain the two proline residues. Additionally, when analysed by FSEC, the iGNTI25 cell line showed a peak around the volume expected for AT₁R-GFP (approximately 11 ml) (Chapter 4), whereas the cell line iHEK26 only showed a peak at about 16 ml, which is the volume where GFP alone elutes, possibly indicating that AT₁R in the iHEK26 cell had been degraded (Figure 5.8). The effects of C-terminal truncations on AT₁R were unexpected since previously it had been shown that C-terminal truncation of the rat version of AT₁R up to amino acid 314 did not alter its expression in CHO cells¹⁷⁶. However the opposite was found to be true in a further study¹⁷⁷. One possible explanation for the decrease in expression was that the truncations were near to or located in helix 8. An alignment of AT₁R with GPCRs of known structure (Appendix 1) indicated that H8 of AT₁R might be located in this region. If the C-terminal truncations removed a portion of H8, then AT₁R would be severely destabilised and rapidly degraded. This might explain the drop in expression. For further analysis of the iGNTI25 cell line see Section 5.3.4.

After examination of the effects of N-linked glycosylation upon AT₁R expression and stability, it was deemed that the only N-linked glycosylation mutant worth further exploring was AT₁R(N4A+N188A) (Section 2.3.2). Therefore, two cell lines stably expressing AT₁R N-linked glycosylation mutations (N4A and N188A) were created, one in iHEK cells

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

(iHEK26) and one in iGnTI⁻ cells (iGNTI26) (Figure 5.6). Neither of these stable cell lines showed favourable levels of expression and therefore were not further investigated.

However, unfortunately both contained C-terminal deletions that removed P321-P322, therefore it is unknown whether the reduction in AT₁R expression was due to the lack of N-glycosylation sites or the C-terminal truncation.

Cell line name		iHEK24					
Open reading frame schematic							
Parental cell line	iHEK					Master Gain	811
Median FACS fluorescence intensity	97					Digital Gain	1.25
Functional expression levels (copies/cell)	0.3 million ± 0.01 (n=3)					Digital Offset	28.00

Cell line name		iGNTI24					
Open reading frame schematic							
Parental cell line	iGnTI ⁻						
Median FACS fluorescence intensity	75						

Cell line name		iHEK25						
Open reading frame schematic								
Parental cell line	iHEK							
Median FACS fluorescence intensity	965							
Functional expression levels (copies/cell)	6 million ± 0.2 (n=6)							

Continued...

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

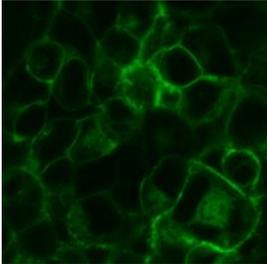
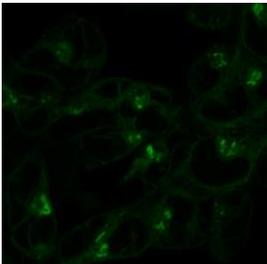
Cell line name		iGNTI25								
Open reading frame schematic		Leader Sequence	FLAG	★ TEV	AT ₁ R(Del. K323-E359)	Sortase	GFP	★ TEV	ZZ	
Parental cell line	iGnTl ⁻								Master Gain	811
Median FACS fluorescence intensity	655								Digital Gain	1.25
Functional expression levels (copies/cell)	6 million ± 0.2 (n=3)								Digital Offset	2818.69
Cell line name		iHEK26								
Open reading frame schematic		Leader Sequence	FLAG	★ TEV	AT ₁ R(N4A+N188A+ Del. P321-E359)	★ TEV	GFP	H ₁₀		
Parental cell line	iHEK								Master Gain	757
Median FACS fluorescence intensity	78								Digital Gain	1.25
Functional expression levels (copies/cell)	0.4 million ± 0.1 (n=3)								Digital Offset	28.00
Cell line name		iGNTI26								
Open reading frame schematic		Leader Sequence	FLAG	★ TEV	AT ₁ R(N4A+N188A+ Del. P321-E359)	★ TEV	GFP	H ₁₀		
Parental cell line	iGnTl ⁻									
Median FACS fluorescence intensity	66									

Figure 5.6 Assessment of cell lines created to express C-terminal truncations of AT₁R

Assessment of polyclonal cell lines was based on confocal microscopy and the average counts for FACS analysis and [¹²⁵I]-Sar¹ binding.

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

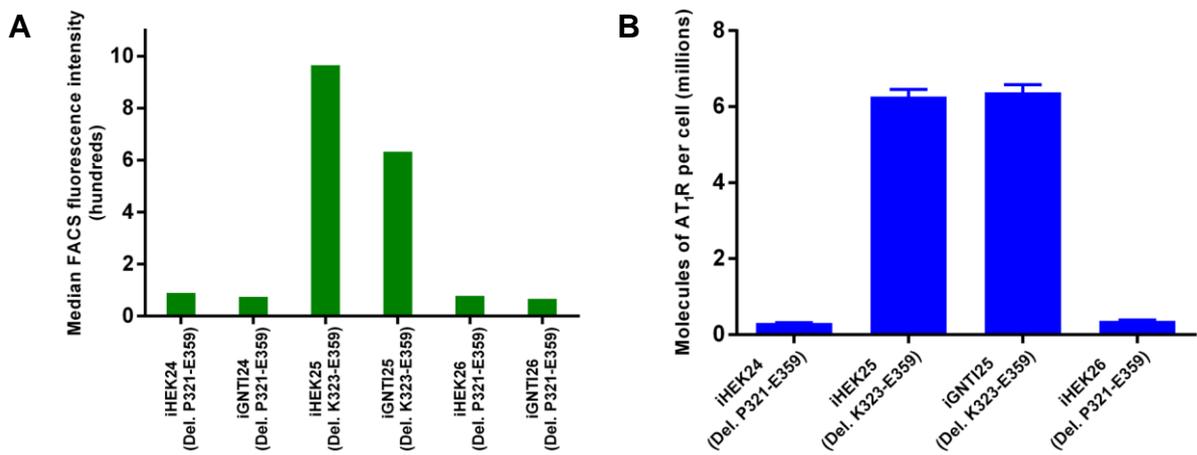


Figure 5.7 Truncation of AT₁R at proline 321 decreases expression

Six stable cell lines were created from three separate constructs and two different cell lines: three in iHEK cells and three in iGnT1⁻ cells. See Figure 5.6 for construct information. For each of these cell lines, the process was repeated six times and the highest expressing population was retained. **(A)** Cells were harvested in PBS and analysed on the FACSCalibur II for GFP fluorescence using the FL-1 detector. **(B)** The amount of functional DDM-solubilised AT₁R in each cell line was determined by measuring specific binding of the antagonist [¹²⁵I]-Sar¹. The assay was performed in at least triplicate and plotted as a mean value ± SEM.

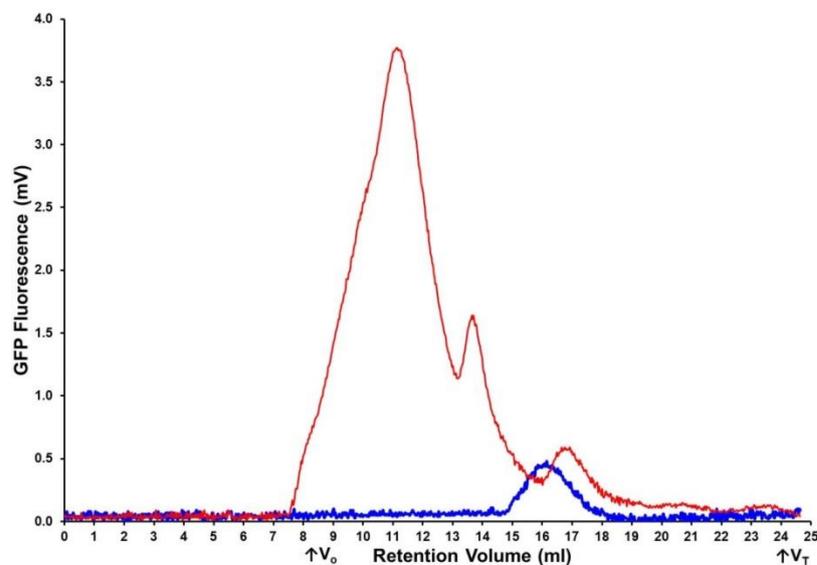


Figure 5.8 FSEC analysis of AT₁R-GFP expression from the iGNTI25 and iHEK26 stable cell line

The cell lines iGNTI25 (expressing AT₁R(Del. K323-E359)-GFP-H₁₀) (red line) and iHEK26 (expressing AT₁R(Del. P321-E359)-GFP-H₁₀) (blue line) were induced with tetracycline for 24 hours and analysed by FSEC. The elution of AT₁R-GFP-H₁₀ was detected using GFP fluorescence (mV). The void (V_0) and total column volumes (V_T) are indicated.

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

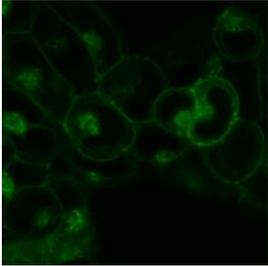
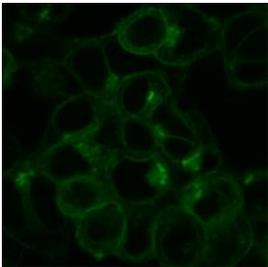
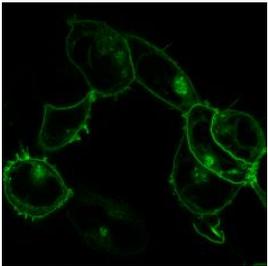
5.3.3 Stability gained by the insertion of T4 lysozyme into intra-cellular loop three of AT₁R

In order to increase the likelihood of crystal contacts forming and to potentially increase the stability of AT₁R, T4L was inserted into intracellular loop three (CL3) of AT₁R. In order to determine where to place T4L, AT₁R was aligned with other GPCR-T4L fusions (Appendix 1). From this alignment the ideal area to place T4L in CL3 was determined to be between residues I228-Q229. Four separate cell lines stably expressing the AT₁R-CL3-T4L fusion were created as well as two cell lines without the T4L fusion; the last two cell lines acted as a control to see whether the inclusion of T4L affected expression (Figure 5.9). The only cell line of this group that showed favourable expression was iGNTI42 which expressed approximately 3 million copies of AT₁R-CL3-T4L per cell (Figure 5.9).

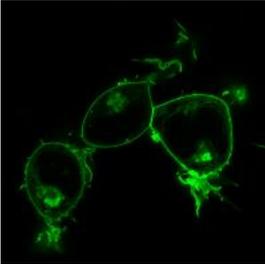
As a way to measure the stability of the AT₁R-CL3-T4L fusion, the apparent T_m of AT₁R-CL3-T4L expressed in the stable cell line iGNTI42 (expressing AT₁R-T4L; Table 5.1) was measured (Section 2.2.10). From this assay, antagonist-bound AT₁R-CL3-T4L fusion showed an 11°C increase in the apparent T_m in comparison to wild type receptor in the [Super +] format assay (Figure 5.10 and Table 5.2). Interestingly, the [-] format assay showed that antagonist-bound AT₁R-CL3-T4L was 6°C less stable than the wild type receptor (Figure 5.10 and Table 5.2). Additionally, when examined by FSEC, the AT₁R-CL3-T4L fusion had a significantly lower signal than the wild type receptor, where an equal number of cells were examined, as well as a peak that was shifted to the right (Figure 5.11). The addition of the antagonist Sar¹ and the use of CHS did not help to stabilise the AT₁R-CL3-T4L fusion (Figure 5.12), neither did the use of the mild detergent LMNG, nor addition of the antagonist azilsartan (Figure 5.13).

Since the [Super +] format assay measures the stability of the ligand binding pocket and the [-] format assay is an indicator of the overall stability of a protein, a possible explanation for the reduction in signal seen in the FSEC analysis could be that while T4L stabilised the ligand binding pocket of AT₁R, it destabilised the other regions of the receptor. Given this, it was decided to progress with the iGNTI25 cell line which did not contain T4L but contained a C-terminal truncation of AT₁R, thus making it more suitable for crystallography than the wild type protein and which showed the highest level of AT₁R expression of all of the C-terminal truncations.

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

Cell line name		iGNTI30					
Open reading frame schematic							
Parental cell line	iGnTI ⁻					Master Gain	969
Median FACS fluorescence intensity	122					Digital Gain	1.25
						Digital Offset	1458.85
Cell line name		iHEK32					
Open reading frame schematic							
Parental cell line	iHEK					Master Gain	1044
Median FACS fluorescence intensity	248					Digital Gain	1.25
						Digital Offset	1458.85
Cell line name		iGNTI32					
Open reading frame schematic							
Parental cell line	iGnTI ⁻						
Median FACS fluorescence intensity	121						
Cell line name		iGNTI40					
Open reading frame schematic							
Parental cell line	iGnTI ⁻					Master Gain	880
Median FACS fluorescence intensity	86					Digital Gain	1.0
						Digital Offset	0
		<i>Continued...</i>					

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

Cell line name		iGNTI42									
Open reading frame schematic		Leader Sequence	FLAG	★ TEV	AT ₁ R (1-228)	T4L	AT ₁ R (229-323)	★ TEV	AT ₁ R (324-359)	GFP	H ₁₀
Parental cell line	iGnTl ⁻									Master Gain	748
Median FACS fluorescence intensity	346									Digital Gain	1.25
Functional expression levels (copies/cell)	3 million ± 0.6 (n=3)									Digital Offset	0

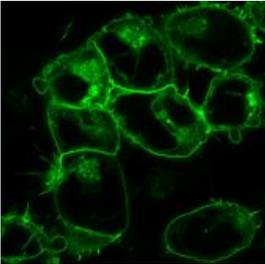
Cell line name		iGNTI43									
Open reading frame schematic		Leader Sequence	FLAG	★ TEV	AT ₁ R(Del. K323-E359)	★ TEV	GFP	ZZ	H ₁₀		
Parental cell line	iGnTl ⁻									Master Gain	789
Median FACS fluorescence intensity	175									Digital Gain	1.24
										Digital Offset	0

Figure 5.9 Assessment of cell lines created to express AT₁R-T4L and C-terminal truncations of AT₁R

Assessment of polyclonal cell lines was based on confocal microscopy and the average counts for FACS analysis and [¹²⁵I]-Sar¹ binding.

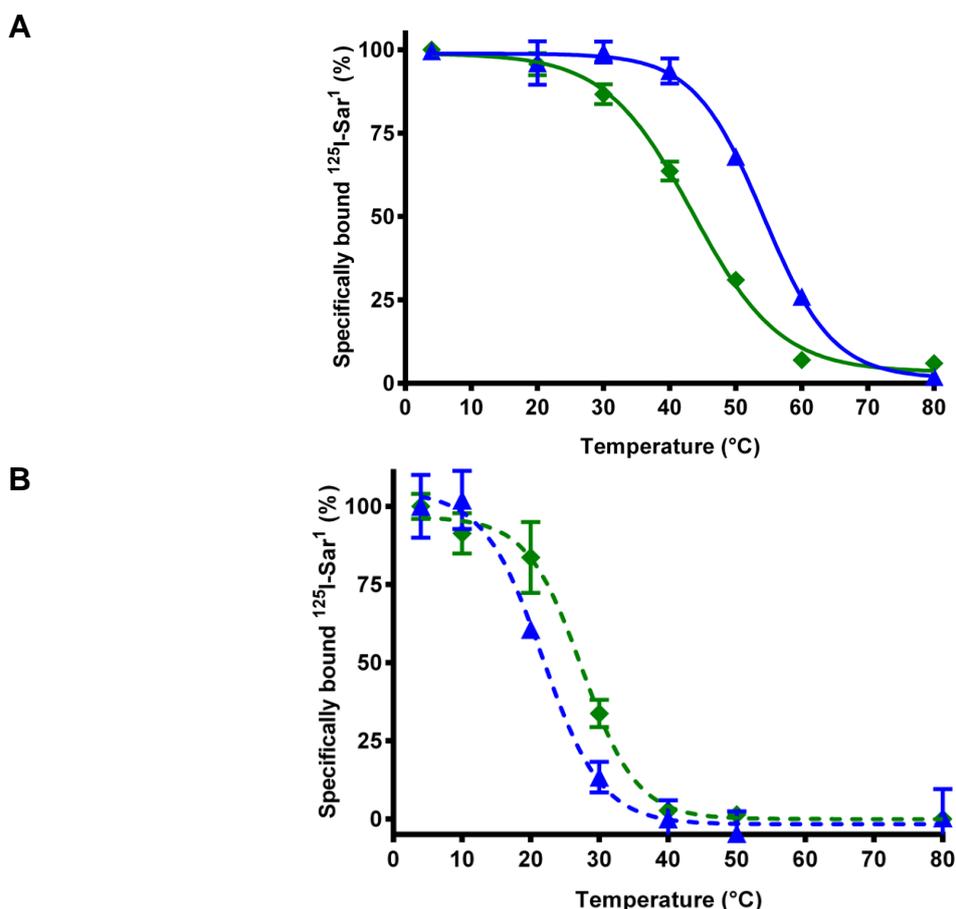


Figure 5.10 T_m assay of AT₁R-T4L fusion

(A) Stability of DDM-solubilised AT₁R bound to the antagonist [¹²⁵I]-Sar¹ performed in the [Super +] format. AT₁R was expressed from the stable cell line iHEK415 (expressing AT₁R-GFP-H₁₀), green diamonds, and iGNT142 (expressing AT₁R-T4L-GFP-H₁₀), blue triangles. **(B)** Stability of LMNG-solubilised AT₁R bound to the antagonist [¹²⁵I]-Sar¹ performed in the [-] format. AT₁R was expressed from the stable cell line iHEK415 (expressing AT₁R-GFP-H₁₀), green diamonds, and iGNT142 (expressing AT₁R-T4L-GFP-H₁₀), blue triangles. Each data point was determined in triplicate and was plotted as a mean value ± SEM. Results are summarised in Table 5.2.

Table 5.2 Thermostability of AT₁R-T4L fusion

CL3 Fusion	Assay Format	Detergent (w/v)	T _m (°C)	SEM
None	[Super +]	1% DDM	43.5	0.7
T4L	[Super +]	1% DDM	54.2	0.8
None	[-]	1% LMNG	27.4	1.1
T4L	[-]	1% LMNG	21.8	1.4

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

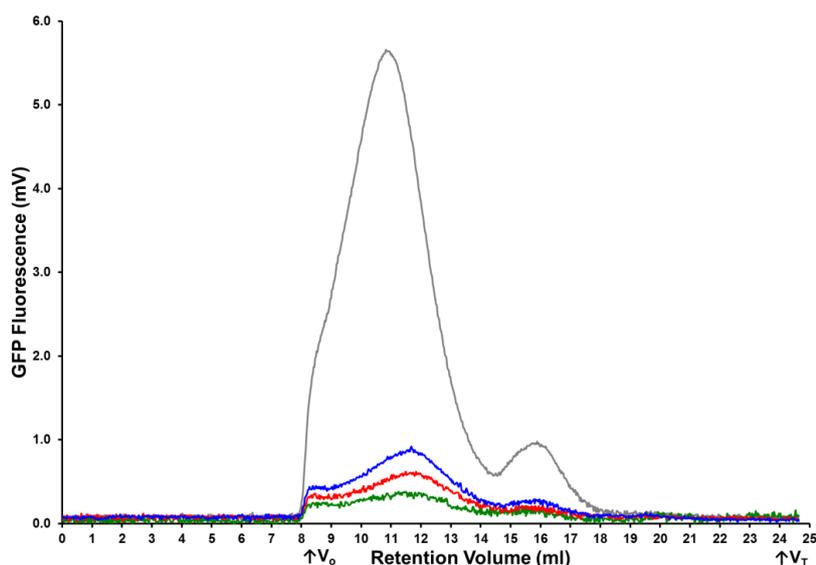


Figure 5.11 FSEC analysis of AT₁R expression from the iGNTI42 stable cell line

The cell line iGNTI42 (expressing AT₁R-T4L-GFP-H₁₀) was induced with tetracycline only (green line), tetracycline with 2.5 mM Na butyrate (red line) or tetracycline with 5 mM Na butyrate (blue line) for 24 hours and analysed by FSEC. The clonal cell line iHEK415 (expressing AT₁R-GFP-H₁₀) was induced with tetracycline only for 24 hours (grey line). An equal amount of cells was analysed for each run. The elution of AT₁R-GFP was detected using GFP fluorescence (mV). The void (V_0) and total column volumes (V_T) are indicated.

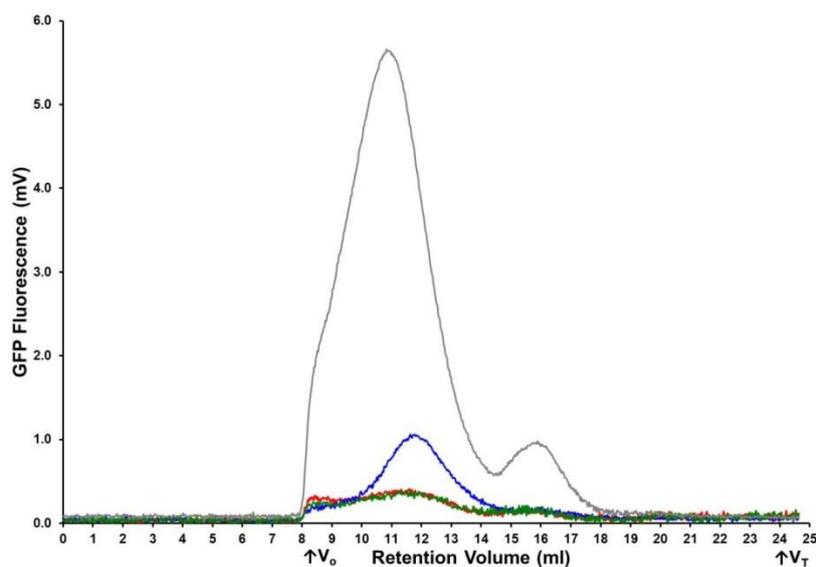


Figure 5.12 The addition of antagonist and the use of CHS did not stabilise AT₁R expressed from the stable cell line iGNTI42

The cell line iGNTI42 (expressing AT₁R-T4L-GFP-H₁₀) was induced with tetracycline only for 24 hours and analysed by FSEC. 1 μ M of Sar¹ was added to 5 million iGNTI42 cells and allowed to bind for 1 hour at 23°C before being solubilised in 1% DDM (red line) or no ligand was added prior to solubilisation (green line). Additionally, 5 million iGNTI42 cells were solubilised in 1% DDM with CHS without ligand (blue line). The clonal cell line iHEK415 (expressing AT₁R-GFP-H₁₀) was induced with tetracycline for 24 hours and analysed by FSEC. 5 million iHEK415 cells were solubilised in 1% DDM without ligand (grey line). An equal amount of cells was analysed for each run. The elution of AT₁R-GFP was detected using GFP fluorescence (mV). The void (V_0) and total column volumes (V_T) are indicated.

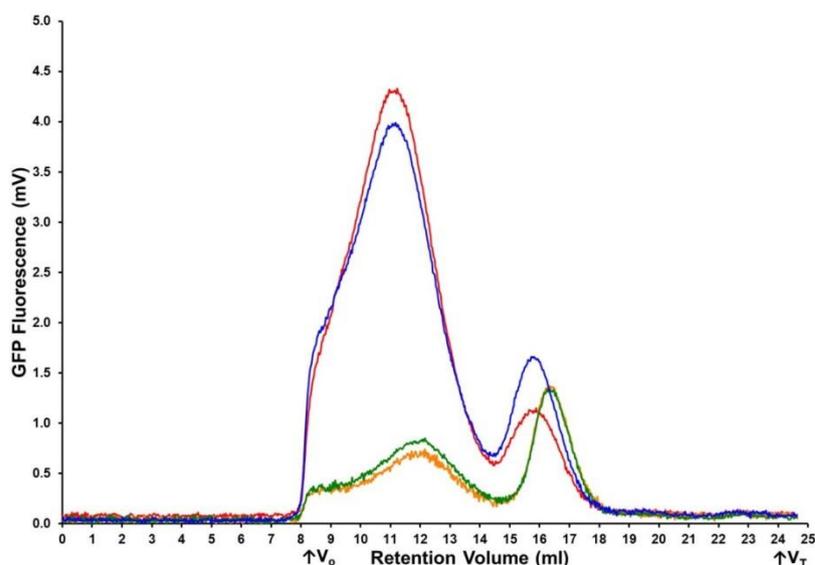


Figure 5.13 The use of LMNG did not stabilise AT₁R-CL3-T4L

The cell line iGNTI42 (expressing AT₁R-T4L-GFP-H₁₀) was induced with tetracycline for 24 hours and analysed by FSEC. 100 μ M of azilsartan was added to 5 million iGNTI42 cells and allowed to bind for 1 hour at 23°C before being solubilised in 1% LMNG (orange line) or no ligand was added prior to solubilisation (green line). The clonal cell line iHEK415 (expressing AT₁R-GFP-H₁₀) was induced with tetracycline for 24 hours and analysed by FSEC. 100 μ M of azilsartan was added to 5 million iHEK415 cells and allowed to bind for 1 hour at 23°C before being solubilised in 1% LMNG (blue line) or no ligand was added prior to solubilisation (red line). An equal amount of cells was analysed for each run. The elution of AT₁R-GFP was detected using GFP fluorescence (mV). The void (V_0) and total column volumes (V_T) are indicated.

5.3.4 Analysis of the stable cell line iGNTI25

The stable cell line iGNTI25 (truncation of K323 to E359) showed the highest level expression of AT₁R in comparison to all C-terminal truncations of AT₁R (6 million copies of AT₁R per cell compared to less than 0.4 million copies for all other AT₁R truncations; Figure 5.6) therefore this cell line was further investigated. It was unknown whether C-terminal truncations of AT₁R affected the stability of the receptor, therefore the stability of AT₁R(Del. K323-E359) produced from the iGNTI25 cell line was measured using the apparent T_m assay (Section 2.2.10). The apparent T_m of antagonist-bound AT₁R expressed in the iGNTI25 cell line ($45.2^\circ\text{C} \pm 0.8$; Figure 5.14) was similar to the apparent T_m of the wild type protein ($46.5^\circ\text{C} \pm 0.3$; Section 2.3.1) in DDM.

The next step was to optimise expression of AT₁R in the iGNTI25 cell line. As shown previously (Sections 2.3.4 and 3.3.3), the addition of sodium butyrate to the culture media at the time of induction can increase the amount of AT₁R produced, and therefore this was tested in the stable cell line iGNTI25. As measured by FACS analysis, the addition of 1 mM sodium butyrate to the culture media increased expression of AT₁R by twofold in comparison to no sodium butyrate, whereas the addition of 5 mM sodium butyrate increased expression of AT₁R by threefold (Figure 5.15).

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

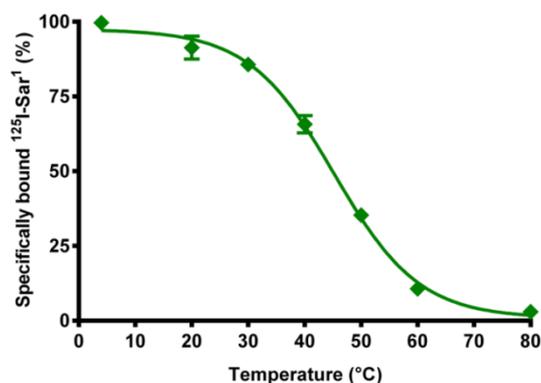


Figure 5.14 T_m assay of AT₁R produced from the stable cell line iGNTI25

Stability of DDM-solubilised AT₁R expressed from the stable cell line iGNTI25 bound to the antagonist [¹²⁵I]-Sar¹. The apparent T_m value of AT₁R expressed from iGNTI25 is 45.2°C ± 0.8. Each data point was determined in triplicate and was plotted as a mean value ± SEM.

However, the addition of 5 mM sodium butyrate did not increase the in-gel fluorescence signal (Figure 5.15). Also, FSEC analysis showed a reduction in signal with the use of 5 mM sodium butyrate in comparison to the addition of 1 mM sodium butyrate (Figure 5.16). A possible explanation for this was AT₁R reached a limit of expression with the inclusion of a LS and 1 mM sodium butyrate. Therefore 1 mM sodium butyrate was used for all further iGNTI25 cultures.

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

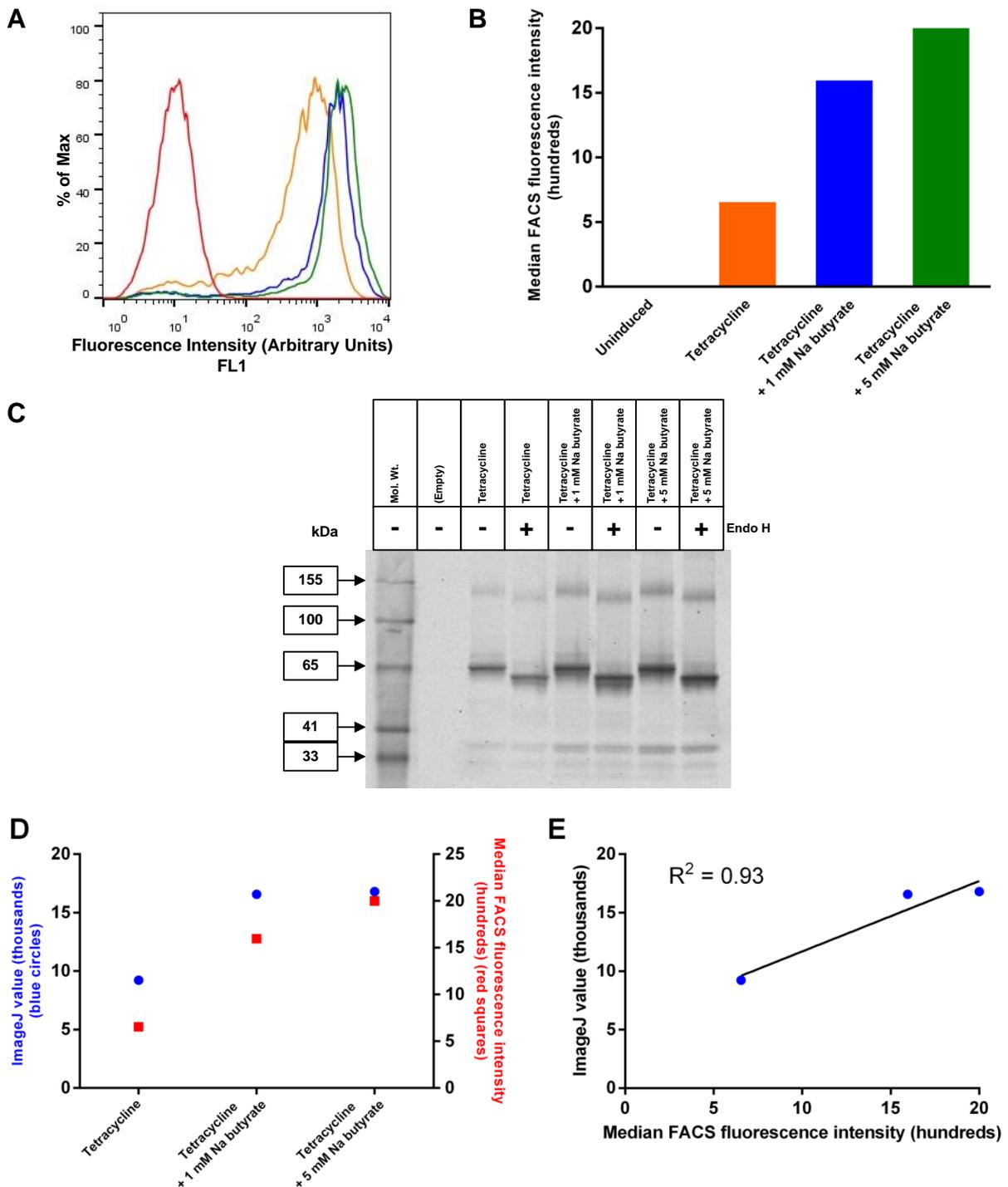


Figure 5.15 Optimisation of AT₁R(Del. K323-E359)-GFP expression in the stable cell line iGNTi25

FACS analysis of iGNTi25 cells. Cells were harvested in PBS and analysed on the FACSCalibur II for GFP fluorescence using the FL-1 detector. Cell counts have been normalised (% of Max). **(A)** Histogram of GFP fluorescence intensity. Uninduced (red), induced with 1 µg/ml tetracycline (orange), induced with 1 µg/ml tetracycline and 1 mM Na butyrate (blue) and induced with 1 µg/ml tetracycline and 5 mM Na butyrate (green). **(B)** Median FACS fluorescence intensity. Colours as in (A). **(C)** In-gel fluorescence of AT₁R-GFP expression from the iGNTi25 stable cell line. An equal amount of cells was loaded in each lane. N-linked glycans were removed using Endo H where indicated (+). **(D)** Bands corresponding to deglycosylated AT₁R were quantified using ImageJ (blue circles) and plotted against the median FACS fluorescence intensity (red squares). **(E)** Linear regression was used to establish a line of best fit for the median FACS fluorescence intensity and the Image J values.

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

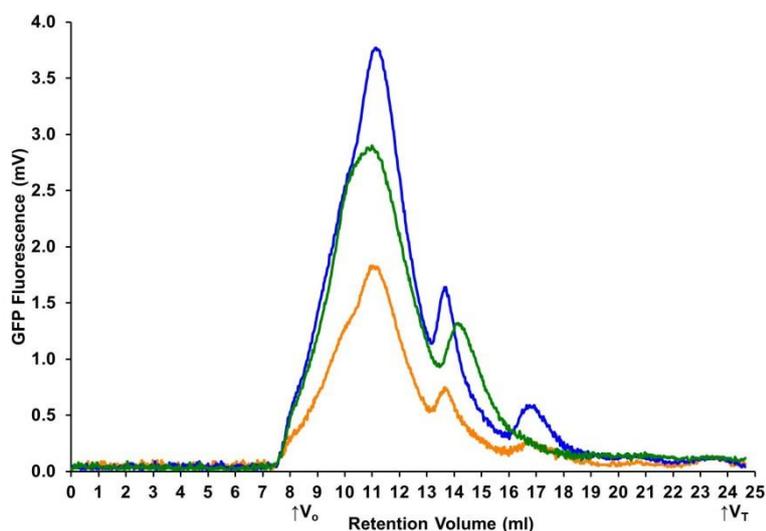


Figure 5.16 FSEC analysis of AT₁R(Del. K323-E359)-GFP expression in the iGNTI25 stable cell line

The iGNTI25 cell line was induced with tetracycline (orange), tetracycline and 1 mM Na butyrate (blue) and tetracycline and 5 mM Na butyrate (green) for 24 hours. 40 nM of Sar¹ was added to 5 million cells and allowed to bind for 1 hour at 23°C before being solubilised in 1% DDM. An equal number of cells was analysed for each run. The elution of AT₁R-GFP-H₁₀ was detected using GFP fluorescence (mV). The void (V₀) and total column volumes (V_T) are indicated.

5.3.5 Large scale growth of mammalian cells

After optimising expression of AT₁R in the iGNTI25 cell line, the next step was to grow iGNTI25 in large scale in order to obtain the amounts of AT₁R required for crystallisation trials. Previous attempts to express wild type AT₁R in the polyclonal cell line iHEK(AT₁R-GFP-H₁₀) using a Wave Bioreactor system, led to the degradation of AT₁R, which was demonstrated by the presence of free GFP in in-gel fluorescence (Figure 5.17). The cell line iHEK(AT₁R-GFP-H₁₀) when adapted to suspension culture tended to form clumps and could not be grown above a density of 1.5×10^6 cells per ml without a large increase in the number of dead cells. The cell line iHEK(AT₁R-GFP-H₁₀) when grown in a Wave Bioreactor system produced ten litres of culture with 1.3×10^6 cells per ml at time of harvest and had a median FACS fluorescence intensity of 227.

In contrast, the polyclonal cell line iGNTI25 did not form clumps and was able to be grown at a density of 3.4×10^6 cells per ml in the Wave Bioreactor system. Propidium iodide (PI) staining of this population indicated that the cell line iGNTI25 showed minimal cell death (2.3% of the total population) (Figure 5.18). The cell line iGNTI25 was induced with tetracycline and 1 mM Na butyrate when the density reached approximately 1.5×10^6 cells per ml. Ten litres of iGNTI25 cells at 3.4×10^6 cells per ml were harvested and they had a median FACS fluorescence intensity of 1407. The clonal cell line iHEK415 induced with tetracycline and 5 mM sodium butyrate has a similar median FACS fluorescence intensity of

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

1400 and, using a radioligand binding assay, this cell line was determined to contain approximately 26 million molecules of AT₁R per cell (Section 3.3.3). Based on this, it was estimated that there was approximately 6 mg of AT₁R per litre of culture, assuming 3.4×10^6 cells per ml.

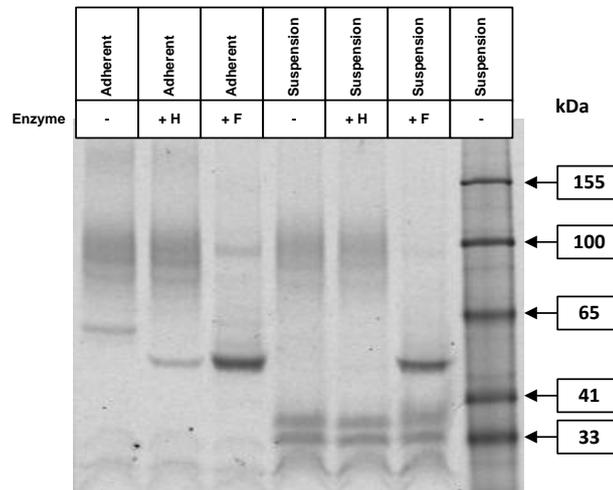


Figure 5.17 Degradation of AT₁R-GFP produced from the polyclonal cell line iHEK(AT₁R-GFP-H₁₀) when grown in suspension culture

In-gel fluorescence of AT₁R-GFP expression from the iHEK(AT₁R-GFP-H₁₀) stable cell line grown adherently or in suspension as indicated. An equal amount of cells was loaded in each lane. N-linked glycans were removed using Endo H (+ H) or PNGase F (+ F) where indicated.

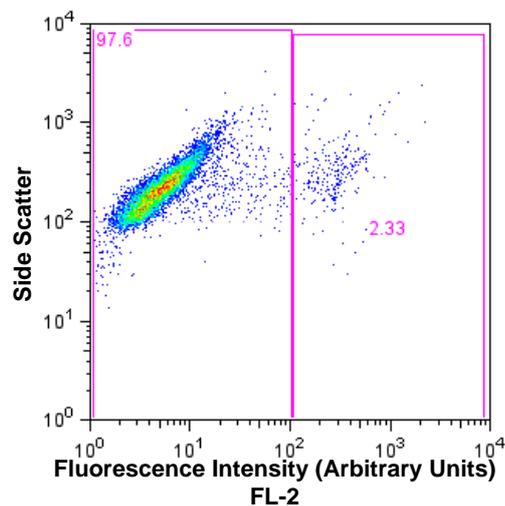


Figure 5.18 FACS analysis of iGNTI25 grown in large scale

Colour density plots of iHEK cells stained with propidium iodide (PI). The cell line iGNTI25 was grown in a Wave Bioreactor. For FACS analysis, a sample of cells was stained with PI (Section 2.2.9) and analysed using the FL-2 detector to detect fluorescence. Numbers inside the pink rectangles represent the percentage of cells falling within the relevant rectangle.

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

5.3.6 Purification of AT₁R

Two stable mammalian cell lines produced sufficient material for the purification of AT₁R, the monoclonal cell line iHEK415 and the polyclonal cell line iGNTI25. Only iGNTI25 could be grown in large scale suspension culture (Section 5.3.5), therefore efforts to purify AT₁R focused on this cell line. Recently a mammalian dynein complex was purified from the baculovirus system using a two-step purification consisting of the recombinant protein binding to IgG Sepharose and then being directly cleaved from the resin with the use of TEV protease, resulting in a homogenous product¹⁸⁵. Since the mammalian dynein complex was expressed at a level similar to AT₁R (2 mg of dynein complex from 1 litre of culture, assuming 1-2 x 10⁶ cells per ml)¹⁸⁵ it was reasonable to assume that this purification process would work well with AT₁R. However this was not the case. AT₁R did not adhere to the IgG Sepharose and ended up mostly in the flowthrough (Table 5.3). One possible explanation for this was that the binding of AT₁R to the resin was hindered due to the presence of 1% detergent (final concentration), which was not required for the purification of dynein (a soluble protein). However, reducing the amount of detergent used during the solubilisation step to 0.5% (final concentration) did not reduce the loss in the flow through. Therefore, an Anti-FLAG M2 affinity gel was investigated. Despite optimising the salt concentration and examining the use of other detergents, this resin also resulted in major losses of AT₁R in the flow through (Table 5.3).

One potential explanation for this is that negatively charged DNA present in the cell lysate was adhering to the resin and preventing AT₁R from binding. As a method to remove DNA, a negative purification of AT₁R was attempted with Q Sepharose (Table 5.3). Some of the DNA was retained on the resin, however the resulting AT₁R in the flow through still did not bind to the FLAG resin, indicating that either insufficient amounts of DNA were removed or some other factor was influencing the binding of AT₁R such as the presence of detergent. Another possibility was that the resin had a very low binding capacity. A potential way to alleviate this problem would be to produce anti-FLAG antibodies and attach them to a resin. Finally, Concanavalin A resin, which binds glycoproteins, was used. This had the opposite effect to that observed with the other resins. Once AT₁R was bound it could not be removed from the resin (Table 5.3) however, a glucose gradient was used for elution whereas elution with methyl α -D-mannopyranoside might have been more effective.

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

Table 5.3 AT₁R purifications attempted

Tag / binding recognition	Resin	Optimisations	Result
Protein A	IgG Sepharose	<ul style="list-style-type: none">• 1 and 0.5% final detergent concentration for solubilisation	<ul style="list-style-type: none">• Major loss of protein in the flow through
FLAG tag	Anti-FLAG M2 affinity gel	<ul style="list-style-type: none">• High and low salt• DDM, DDM + CHS and LMNG	<ul style="list-style-type: none">• ~80% loss of protein in the flow through
Negatively charged substances	Q Sepharose	<ul style="list-style-type: none">• Gradient of NaCl for binding• Negative purification to remove contaminating DNA	<ul style="list-style-type: none">• Major loss of protein in the flow through• Some DNA bound
Glycoproteins	Concanavalin A	<ul style="list-style-type: none">• Elution with a glucose gradient	<ul style="list-style-type: none">• AT₁R would not elute from column

5.4 Discussion

Protein engineering is often a requirement for obtaining well-diffracting crystals. It can be used to increase expression of the target protein, to remove N-linked glycosylation sites and flexible regions and to stabilise the protein. The FACS sort method used to create clonal cell lines (Chapter 3) was effective at increasing AT₁R expression, however it was a time consuming process. It has been observed that there is a charge-bias among proteins inserted into the plasma membrane, in particular residues with positively charged amino acids are more likely to be located in the cytoplasm^{171, 172}. Since the N-terminus of AT₁R contained four positively charged residues, removal of one or more of them appeared to be a promising means for increasing expression. This proved to be the case for the expression of N-terminal mutants of AT₁R transiently transfected into iHEK cells, however each positively charged amino acid was examined in isolation. It would be interesting to see whether removal of more than one positively charged residue further increased expression and also whether this increase was seen in stable cell lines created from N-terminal mutants. Additionally, examining the effects of using amino acids other than Gln might be valuable. Ultimately this technique was not further investigated since it was unknown whether N-terminal alterations of AT₁R might negatively impact ligand binding. Instead the use of a leader sequence (LS) was examined as a means to increase expression of AT₁R in mammalian cells. This approach had previously been successfully employed in the baculovirus system (Chapter 2) and therefore using it in the mammalian system was a logical extension. The addition of a LS to AT₁R expressed in the mammalian system, in combination with the use of sodium butyrate at the time of induction, removed the need to use the FACS sort method because expression levels were comparable (both induced cell lines iHEK415 and iGNTI25 gave a median FACS fluorescence intensity of approximately 1400).

The use of RONN software¹⁷⁵ in combination with an alignment of AT₁R with other GPCRs of known structure, predicted that the C-terminus was flexible and might prohibit well-diffracting crystals from forming. There was conflicting information available about the impact of C-terminal truncations on the expression of AT₁R. Both rat and human AT₁R consist of 359 residues and share 95% sequence identity (Appendix 1). CHO cell lines stably expressing C-terminal deletions of rat AT₁R up to residue 314 showed that neither the binding affinity nor the expression levels of AT₁R were impacted¹⁷⁶. However, C-terminal deletions of the rat AT₁R up to residue 309, expressed by transient transfection in COS-7 cells, showed that, despite binding affinities remaining unchanged, expression of AT₁R

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

decreased with increasing the length of the truncation¹⁷⁷. It was therefore necessary to establish whether it was possible to remove the putatively disordered C-terminus of AT₁R without negatively impacting expression. To facilitate this, mammalian cell lines stably expressing two different truncations of AT₁R were examined; P321-E359 and K323-E359. Only cell lines containing P321 and P322 showed high levels of expression of AT₁R possibly indicating that truncation of AT₁R beyond K323 destabilised the receptor. The apparent T_m of antagonist-bound AT₁R(Del. K323-E359) solubilised in DDM was similar to that of the wild type receptor, demonstrating that the stability of the altered protein was not negatively impacted by the truncation.

The addition of T4L to increase the hydrophilic area of a GPCR, particularly in combination with the use of LCP, has become a routine method for the crystallization of GPCRs. A β_2 AR-CL3-T4L fusion has previously been carefully engineered with the intention of stabilising the receptor rather than introducing further flexibility and resulted in a high-resolution structure^{24, 41}. Since then numerous high-resolution GPCR structures have been obtained using this method, therefore the creation of an AT₁R-T4L fusion appeared to be worthwhile. Using the [Super +] format apparent T_m assay, antagonist-bound AT₁R-CL3-T4L showed an 11°C increase in apparent T_m in comparison to the wild type receptor. However using the [-] format assay, AT₁R-CL3-T4L showed a 6°C reduction in apparent T_m in comparison to the wild type receptor. Additionally, FSEC analysis of AT₁R-CL3-T4L demonstrated a dramatic reduction in signal intensity in comparison to the wild type receptor, which persisted despite optimisation. A possible explanation for this effect was that while T4L stabilised the binding pocket, the other domains of the receptor were destabilised. Another possibility was that the introduction of T4L caused an increase in the flexibility of CL3, thus destabilising the receptor. A method for determining whether this was happening would be to introduce rigid linkers at the AT₁R-T4L junction, or other placements of T4L. Finally AT₁R might exist as a dimer and the addition of T4L might inhibit the formation of dimers, which could also destabilise the receptor (Chapter 4).

Given the inconsistent behaviour of AT₁R-CL3-T4L, it was decided to proceed with a version of AT₁R that did not contain T4L but had a C-terminal truncation which did not affect expression. The cell line iGNTI25, which expressed AT₁R(Del. K323-E359) grew well in the Wave Bioreactor system and it was estimated that there was approximately 6 mg of AT₁R per litre of culture, assuming 3.4×10^6 cells per ml. Unfortunately, it was not possible to purify the receptor derived from this cell line. It is likely that the antibody based

CHAPTER 5 Engineering AT₁R for Use in Structural Studies

purification of AT₁R was unsuccessful because of the presence of detergent in the solubilisate. One potential way to alleviate this problem would be to use a tag such as H₁₀. The binding of H₁₀ to IMAC resin is not known to be affected by detergent and it should be sufficiently long that it would not be occluded by the detergent micelle. However the only cell line which was successfully grown in suspension, iGNTI25, did not express AT₁R with a H₁₀ tag. Therefore, a new stable cell line would need to be generated. Alternatively, the genome of iGNTI25 could be edited to contain the H₁₀ tag using the clustered regularly interspaced short palindromic repeats (CRISPR) / CRISPR-associated protein-9 nuclease (Cas9) system from *Streptococcus pyogenes*¹⁸⁶.

CHAPTER 6 DETECTION OF UNFOLDED RECOMBINANT MEMBRANE PROTEINS IN EUKARYOTIC CELLS

6.1 Introduction

One of the most important aspects of membrane protein overexpression is knowing both how much of the target protein is folded correctly and how much is misfolded. However, detecting the presence of misfolded recombinant membrane proteins is often very difficult. For *E. coli* a method was devised that relies upon the creation of a C-terminal GFP-fusion protein; if the target membrane protein is correctly folded, GFP is fluorescent, but if the membrane protein misfolds, it forms aggregates within the cell and the GFP does not fluoresce^{183, 187}. On SDS-PAGE, the fluorescent GFP fusion protein is observed by in-gel fluorescence. A western blot of the same gel probed with anti-GFP antibody identifies both folded and misfolded fusion protein. It is usually the case that the functional GFP fusion protein migrates further than the misfolded GFP fusion protein. This system has been adapted to work in yeast expression systems¹⁸⁸. However, for eukaryotic expression systems, GFP often remains fluorescent despite being fused to misfolded proteins^{189, 190}. Prior to the work described in this thesis detection of misfolded eukaryotic membrane proteins relied upon either radioligand binding experiments⁶⁰ or examination of the void area produced from an FSEC trace¹⁴⁸.

It has been observed previously that SERT expressed using the baculovirus system showed differential amounts of peptide on a western blot depending on the ‘harshness’ of the detergent used (Tate, unpublished data). This was consistent with the observation that most of SERT expressed in insect cells was misfolded, whereas subsequent work showed that the majority of SERT expressed in a mammalian cell line was correctly folded⁶⁰. However, in most baculovirus expression trials it is not possible to compare western blots because either a radioligand is not available for the target or a stable mammalian cell line is not available for comparison. This means that a simple assay would be of great utility in detecting whether misfolded protein was expressed in any expression system. To address this, the observation of differential solubility of misfolded and correctly folded membrane protein in detergent of different ‘harshness’ was used to develop an assay. In order to assess the efficacy of this assay targets were chosen that had been expressed in both mammalian cells and insect cells, and which also had a radioligand binding assay. The assay was used to

CHAPTER 6 Detection Of Unfolded Membrane Proteins

determine how much functional membrane protein was solubilised in two mild detergents (digitonin or DDM). The amount of detergent-solubilised target protein was then determined on a western blot and compared between different types of detergents *i.e.* the harsh detergents SDS and FC12 and the mild detergents digitonin and DDM. It was anticipated that if all the target protein was correctly folded then the western blot signal should be identical in all four samples. If there was misfolded protein present, then this would be observed by an increase in the amount of target protein seen in the samples solubilised in either SDS or FC12.

6.2 Materials and Methods

6.2.1 Materials

All radiolabelled ligands were purchased from PerkinElmer: [³H]-dihydroalprenolol ([³H]-DHA), [³H]-dipropylcyclopentylxanthine ([³H]-DPCPX) and [¹²⁵I]-2β-carbomethoxy-3β-(4-iodophenyl)tropane ([¹²⁵I]RTI-55). The detergent fos-choline-12 (FC12) was purchased from Anatrace.

6.2.2 Constructs

Expression in mammalian cells was performed using derivatives of pcDNA4/TO (Invitrogen) (Section 2.2.2 for a description of mammalian expression vectors for SERT and AT₁R). The cDNA for the human adenosine A₁ receptor (A₁R) (Missouri S&T cDNA Resource Center) was cloned into the EcoRV/NotI sites of pJMA111 (Section 2.2.2) to create plasmid pJAP34, which expressed A₁R-GFP-H₁₀. In an effort to create a thermostable A₁R receptor, 4 mutations that stabilised the adenosine A_{2A} receptor (A_{2A}R) in the active state (L48A, A54L, T65A, Q89A⁶⁷) were transferred to A₁R (mutations L51A, A57L, L68A, Q92A). In addition, the mutations N148G and N159G were included to remove the putative N-linked glycosylation sites. To remove flexible regions, the N-terminus was truncated between Pro2 and Ile5, the C-terminus was truncated at Phe307 and the sequence VLRQQEPEPKAA was added to the C-terminus, thus generating A₁R-GL26. A synthetic cDNA encoding A₁R-GL26 (Life Technologies) was cloned into the EcoRV/NotI sites in pJMA111 creating pJAP37, which expressed A₁R-GL26-GFP-H₁₀. For generating baculoviruses, A₁R was cloned into the XhoI/EcoRI sites of the transfer vector pBacPAK8 (Clontech), and A₁R-GL26 was cloned into the EcoRI/EagI sites of the same vector, creating plasmids pJAP44 and pJAP33 respectively (see Section 2.2.2 for AT₁R baculovirus transfer vector creation). All baculovirus sequences were engineered to contain a C-terminal tobacco etch virus (TEV) cleavage site and H₁₀ tag. All constructs were verified by DNA sequencing (Source Biosciences, UK).

6.2.3 Transient transfection, generation of stable cell lines and protein expression

Mammalian expression plasmids for the expression of A₁R (pJAP34) and A₁R-GL26 (pJAP37) were amplified, transiently transfected into adherent mammalian iHEK cells or

CHAPTER 6 Detection Of Unfolded Membrane Proteins

iGnTI⁻ cells, grown in culture and induced with tetracycline as described previously (Section 2.2.4). Stable cell lines were generated by selection with media containing Zeocin (Section 2.2.4). An iGnTI⁻ stable cell line expressing a thermostable mutant of SERT, SERT-SAH9⁶⁵, iGnTI⁻ (SERT-SAH9-GFP-H₁₀), was kindly provided by J. Andréll. Cells were induced and harvested as described previously (Section 2.2.4).

6.2.4 Recombinant baculovirus generation and protein expression

Recombinant baculoviruses that expressed either A₁R or A₁R-GL26 were generated, isolated and screened for expression as described previously (Section 2.2.5). Recombinant baculovirus that expressed SERT with a H₁₀ tag at its C-terminus was created as previously reported^{131, 140}. Recombinant baculovirus that expressed β₁AR with a H₁₀ tag at its C-terminus was kindly provided by R. Nehme (MRC Laboratory of Molecular Biology) and a thermostable β₁AR fused to thioredoxin (tsβ₁AR) was kindly provided by T. Warne (MRC Laboratory of Molecular Biology). Recombinant baculoviruses were passaged twice in Sf9 cells to obtain high titre stocks. Viruses were used to infect either Sf9, Sf21 or Hi5 cells for 48 or 72 hours as indicated. Cells were harvested as described previously (Section 2.2.5).

6.2.5 Differential solubility western blotting

Cell suspensions were sonicated briefly and the total protein concentration determined using the Bradford assay¹³². Samples were then solubilised in the detergent indicated (SDS, FC12, DDM or digitonin; all at 1% (w/v) final concentration) at either 4°C (FC12, DDM, digitonin) or 20°C (SDS) for 1 hour. The solubilisate was centrifuged at 280,000 × g for 30 minutes at 4°C to remove the insoluble fraction. SDS-loading buffer was added to the supernatant (corresponding to approximately 150,000 cells) and samples were separated on a 4-20% tris glycine gel and transferred to nitrocellulose using standard techniques. Membranes were probed with anti-pentaHis-HRP (Section 2.2.6). Where indicated, PNGase F was added to the supernatant prior to SDS-PAGE (Section 2.2.6).

6.2.6 Detergent-solubilised and membrane bound radioligand binding assays

Detergent solubilised radioligand assays were performed as described previously (Section 2.2.11) with the following additions: [³H]-DHA was used at a final concentration of 200 nM and [³H]-DPCPX at 39 nM in 150 mM NaCl, 50 mM Tris pH 7.4. [¹²⁵I]-RTI-55 was used at a concentration of 1 nM in PBS. Binding of [³H]-DHA, [³H]-DPCPX and [¹²⁵I]-RTI-55 was

CHAPTER 6 Detection Of Unfolded Membrane Proteins

on ice for 1 hour. To determine the amount of SERT, AT₁R, β₁AR or A₁R present in cell membranes, binding assays were performed without the samples being solubilised.

Separation of receptor-bound and free radioligand was achieved by filtration through 96-well glass fibre filter plates (Millipore) pre-treated with 0.1% polyethyleneimine⁶⁵ except for [¹²⁵I]-Sar¹ where no polyethyleneimine was used. Background for both assays was determined by adding radioligand to non-transfected parental mammalian cells or uninfected insect cells.

6.3 Results

6.3.1 Development of an assay to detect unfolded recombinant membrane proteins in eukaryotic cells

The first step in purifying any membrane protein is its extraction from the lipid bilayer using a detergent, therefore the ability of four detergents to solubilise AT₁R was examined. It was observed that AT₁R produced in the baculovirus system gave a considerably stronger signal on a western blot when harsher detergents (SDS and FC12) were used in comparison to when milder detergents (DDM and digitonin) were used, indicating that SDS and FC12 were solubilising more AT₁R peptide than DDM and digitonin (Figure 6.1). This finding contrasted with the equal signal seen across all detergent conditions for AT₁R produced in the mammalian system (Figure 6.1). Combined with the knowledge that the baculovirus system produced large quantities of unfolded AT₁R, whereas the mammalian system did not (Section 2.3.5), these findings were further investigated using radioligand binding experiments.

For AT₁R produced by both the mammalian and baculovirus system [¹²⁵I]Sar¹-bound AT₁R was only detected using the mild detergents DDM and digitonin (Figure 6.1). Across both systems, only a small amount of binding was observed when FC12 was used and no binding was detected when SDS was used. In order to compare the radioligand binding data across both systems, the data in Figure 6.1 were normalised. However it should be noted that the stable mammalian cell line expressed 20 times more functional AT₁R than was observed in Sf9 cells. In addition, there was twice as much [¹²⁵I]Sar¹-bound AT₁R detected when the receptor was solubilised with DDM than when the receptor was bound to membranes (Figure 6.1). Since freeze-thawed membranes were used for this assay, the lower signal seen in membranes could be due to the membrane-impermeant peptide [¹²⁵I]-Sar¹ not being able to access inside-out vesicles which were present along with rightside-out vesicles. Given that no [¹²⁵I]Sar¹-bound AT₁R was detected for the SDS condition, it was conceivable that the signal seen on the western blot for this detergent condition represented misfolded AT₁R protein. Conversely, given the high levels of radioligand binding seen for the DDM condition, an assumption can be made that DDM solubilised only active AT₁R. Therefore a discrepancy in signal on the western blots between harsh and mild detergents could indicate the presence of misfolded material. It was next important to establish whether this remained the case for other membrane proteins.

CHAPTER 6 Detection Of Unfolded Membrane Proteins

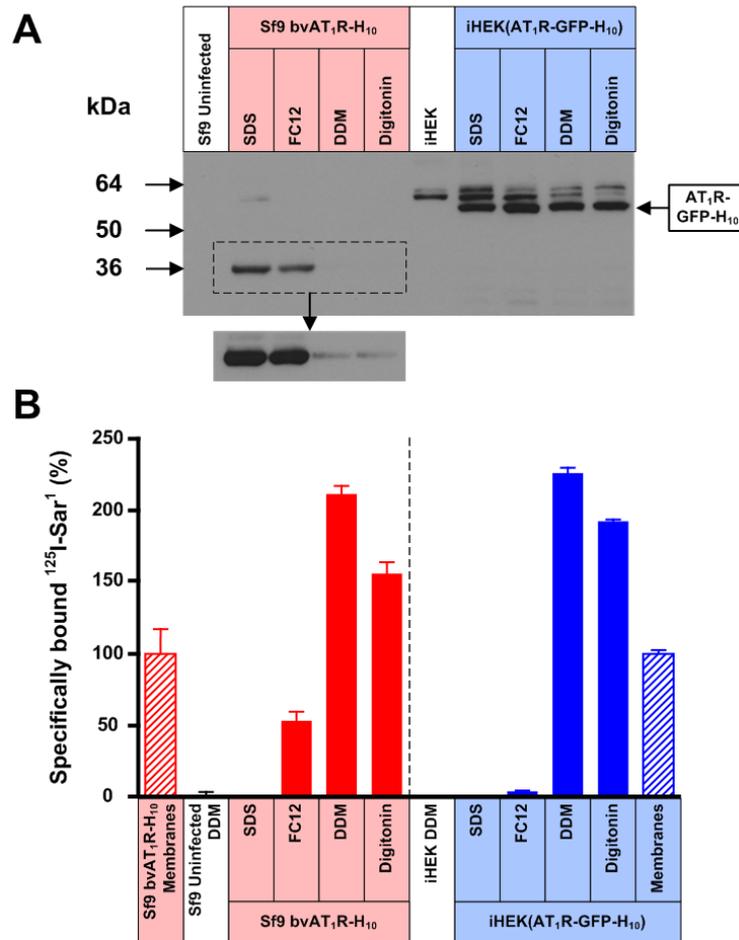


Figure 6.1 Misfolded AT₁R produced by the baculovirus expression system is poorly solubilised by either DDM or digitonin

(A) Western blot of AT₁R solubilised from whole cells using four different detergents (SDS, FC12, DDM or digitonin) and probed with an anti-pentaHis-HRP conjugated antibody. Each lane contains an equal amount of total protein and N-linked glycosylation was removed from all samples using PNGaseF prior to SDS-PAGE. AT₁R was expressed either in the stable mammalian cell line iHEK(AT₁R-GFP-H₁₀) or by using the recombinant baculovirus bvAT₁R-H₁₀ to infect Sf9 cells. The iHEK cell line was induced with 1 µg/ml tetracycline for 24 hours and Sf9 cells were infected for 48 hours. The western blot insert is a 7 times longer exposure. **(B)** The amount of functional detergent-solubilised AT₁R was determined by measuring specific binding of the antagonist [¹²⁵I]-Sar¹. After the addition of ligand, membranes were solubilised in the detergent indicated and non-bound ligand was separated from receptor-ligand complex on gel filtration spin columns and measured by liquid scintillation counting: filled red bars, AT₁R expressed in Sf9 cells; filled blue bars, AT₁R expressed in iHEK cells. The amount of AT₁R in membranes (non-solubilised) was determined by separation of receptor-bound and free radioligand by filtration through glass fibre plates: hatched red bars, AT₁R expressed in Sf9 cells; hatched blue bars, AT₁R expressed in iHEK cells. For ease of comparison, binding data have been normalised with respect to AT₁R in membranes (100%), which is equivalent to 1,400 ± 240 dpm (n=2; 380 fmoles/million cells) for baculovirus-infected Sf9 cells and 12,000 ± 300 dpm (n=2; 8.8 pmoles/million cells) for iHEK(AT₁R-GFP-H₁₀) cells. Absolute levels of AT₁R therefore cannot be compared meaningfully between the two expression systems using this bar graph. Binding assays for AT₁R contained either 150,000 Sf9 cells or 55,000 iHEK cells. Each data point was determined in duplicate or triplicate from a single experiment and was plotted as mean ± SEM.

CHAPTER 6 Detection Of Unfolded Membrane Proteins

6.3.2 Expression of A₁R in insect and mammalian cells

The next membrane protein to be investigated was the adenosine A₁ receptor (A₁R). Although structures of A_{2A}R in inactive^{16, 61, 63} and active-like^{49, 50} conformations have been determined, no structures of A₁R had been published to date, therefore this presented an opportunity for further study. Human A₁R contains two N-linked glycosylated sites in EL2 (Asn148 and Asn159). For the same reasons previously described in this thesis for AT₁R (Chapter 2) two expression systems for the production of recombinant A₁R were chosen, baculovirus mediated expression in insect cells and mammalian expression using the iGNTI⁻ cell line. A baculovirus that expressed A₁R from the polyhedrin promoter (bvA₁R-H₁₀) was created using the Baculo Gold method and used to infect Sf9 cells for 72 hours. Additionally a cell line stably expressing A₁R-GFP-H₁₀ was created using iGNTI⁻ cells. Cells were induced for 24 hours with tetracycline prior to harvest. Interestingly, in the uninduced state the cell line A₁R-GFP-H₁₀ grew significantly slower (over two weeks to reach confluency in a T-75 flask) than the parental cells or cell lines stably expressing AT₁R-GFP-H₁₀, presumably due to high basal activity of the receptor. In an attempt to diminish these effects, the inverse agonist DPCPX was added to the culture medium at final concentrations of 1 μM, 10 μM and 100 μM, but none of the conditions increased the growth rate of the cells.

When A₁R was expressed in the cell line iGNTI⁻ (A₁R-GFP-H₁₀) there was a reduction in apparent molecular weight when the enzyme PNGase F was used (Figure 6.2), presumably due to the removal of N-linked glycosylation. This was not seen for A₁R produced in insect cells (Figure 6.2). As measured by [³H]-DPCPX binding, similar amounts of A₁R was produced in mammalian cells compared to insect cells (Figure 6.2). However, a western blot containing equal amounts of active material showed that the baculovirus produced a significant amount of unfolded receptor whereas the mammalian system produced none (Figure 6.2).

A western blot of A₁R using four different detergent conditions produced a similar pattern to that seen for AT₁R. Specifically, A₁R expressed in insect cells showed a much stronger signal in harsh detergents (SDS and FC12) in comparison to mild detergents (DDM and digitonin) whereas an equal signal was seen across all conditions when A₁R was expressed in mammalian cells (Figure 6.3). A₁R-[³H]-DPCPX binding was noticeably less for all detergent solubilised conditions in comparison to binding in membranes (Figure 6.3), presumably due to the instability of the detergent-solubilised receptor.

CHAPTER 6 Detection Of Unfolded Membrane Proteins

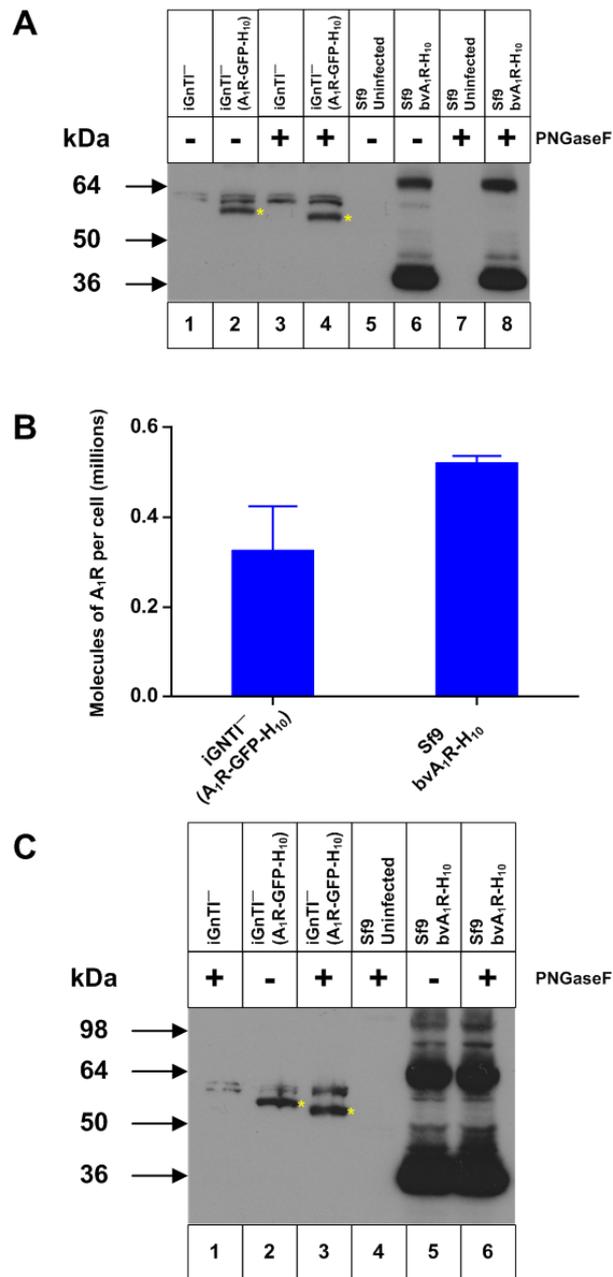


Figure 6.2 Expression of A₁R in mammalian cells compared to insect cells

(A) Western blot of whole cells expressing A₁R solubilised in SDS. Lanes 1 and 3, iGnTI⁻ parental cells; lanes 2 and 4, iGnTI⁻(A₁R-GFP-H₁₀) stable cell line; lanes 5 and 7, uninfected Sf9 cells; lanes 6 and 8, bvA₁R-H₁₀ infected Sf9 cells. N-linked glycosylation was removed using PNGaseF where indicated (+). Bands corresponding to A₁R-GFP-H₁₀ in mammalian cells are indicated with a yellow asterisk (*). The iGnTI⁻ cell line was induced with 1 μg/ml tetracycline for 24 hours and insect cells were infected with recombinant baculovirus for 72 hours. The blot was probed with an anti-pentaHis-HRP conjugated antibody. **(B)** The amount of functional A₁R in each expression system was determined by measuring specific binding of the antagonist [³H]-DPCPX. After the addition of ligand, membranes were solubilised in DDM and non-bound ligand was separated from receptor-ligand complex on gel filtration spin columns and measured by liquid scintillation counting. Each data point was determined in duplicate and was plotted as mean ± SEM. **(C)** Western blot of DDM-solubilised A₁R, with equal amounts of active receptor per sample (lanes 2, 3, 5, 6). The blot was probed with an anti-pentaHis-HRP conjugated antibody. Lane 1, iGnTI⁻ parental cells; lanes 2 and 3, iGnTI⁻(A₁R-GFP-H₁₀) stable cell line; lane 4, uninfected Sf9 cells; lanes 5-6, bvA₁R-H₁₀ infected Sf9 cells. N-linked glycosylation was removed using PNGase F where indicated (+). A₁R was expressed either in the stable mammalian cell line iGnTI⁻(A₁R-GFP-H₁₀) or by using the recombinant baculovirus bvA₁R-H₁₀ to infect Sf9 cells. Bands corresponding to A₁R-GFP-H₁₀ in mammalian cells are indicated with a yellow asterisk (*). The iGnTI⁻ cell line was induced with 1 μg/ml tetracycline for 24 hours and insect cells were infected with recombinant baculovirus for 72 hours. The amount of functional A₁R was determined by measuring specific binding of the antagonist [³H]-DPCPX.

CHAPTER 6 Detection Of Unfolded Membrane Proteins

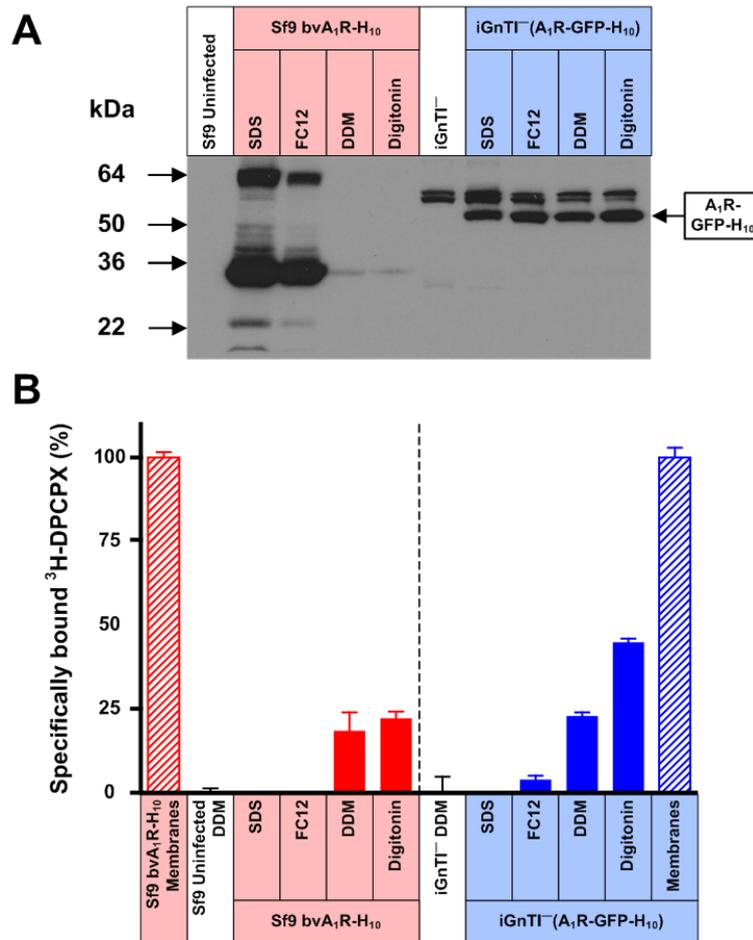


Figure 6.3 Misfolded A₁R produced by the baculovirus expression system is poorly solubilised by either DDM or digitonin

(A) Western blot of A₁R solubilised from whole cells using four different detergents (SDS, FC12, DDM or digitonin) and probed with an anti-pentaHis-HRP conjugated antibody. Each lane contains an equal amount of total protein and N-linked glycosylation was removed from all samples using PNGase F prior to SDS-PAGE. A₁R was expressed either in the stable mammalian cell line iGnTI⁻(A₁R-GFP-H₁₀) or by using the recombinant baculovirus bvA₁R-H₁₀ to infect Sf9 cells. The iGnTI⁻ cell line was induced with 1 μg/ml tetracycline for 24 hours and Sf9 cells were infected for 72 hours. **(B)** The amount of functional detergent-solubilised A₁R was determined by measuring specific binding of the antagonist [³H]-DPCPX. After the addition of ligand, membranes were solubilised in the detergent indicated and non-bound ligand was separated from receptor-ligand complex on gel filtration spin columns and measured by liquid scintillation counting: filled red bars, A₁R expressed in Sf9 cells; filled blue bars, A₁R expressed in iGnTI⁻ cells. The amount of A₁R in membranes (non-solubilised) was determined by separation of receptor-bound and free radioligand by filtration through glass fibre plates: hatched red bars, A₁R expressed in Sf9 cells; hatched blue bars, A₁R expressed in iGnTI⁻ cells. For ease of comparison, binding data have been normalised with respect to A₁R in membranes (100%), which is equivalent to 120,000 ± 2000 dpm (n=3; 2.9 pmoles/million cells) for baculovirus-infected Sf9 cells and 7,500 ± 250 dpm (n=3; 3.8 pmoles/million cells) for iGnTI⁻(A₁R-GFP-H₁₀) cells. Absolute levels of A₁R therefore cannot be compared meaningfully between the two expression systems using this bar graph. Binding assays for A₁R contained either 150,000 Sf9 cells or 7,500 iGnTI⁻ cells. Each data point was determined in duplicate or triplicate from a single experiment and was plotted as mean ± SEM.

CHAPTER 6 Detection Of Unfolded Membrane Proteins

6.3.3 Generation of a thermostable version of A₁R

Thermostabilising mutations in turkey β_1 AR have been successfully transferred to both human β_1 AR and human β_2 AR¹⁰⁸. An alignment of transmembrane regions of adenosine receptors showed that A₁R had a 51% identity with A_{2A}R (Table 6.1). Therefore, in an effort to thermostabilise A₁R, 4 mutations that stabilised the A_{2A}R in the active state (L48A, A54L, T65A, Q89A⁶⁷) were transferred to A₁R (mutations L51A, A57L, L68A, Q92A) thus generating A₁R-GL26. A baculovirus that expressed A₁R-GL26 from the polyhedrin promoter (bvA₁R-GL26-H₁₀) was created using the Baculo Gold method and used to infect Sf9 cells for 72 hours. Additionally, a cell line stably expressing A₁R-GL26-GFP-H₁₀ was created using iGNTI⁻ cells. Cells were induced for 24 hours with tetracycline prior to harvest. Confocal microscopy of the cell line iGNTI⁻(A₁R-GL26-GFP-H₁₀) showed that the vast majority of A₁R-GL26 was internalised and presumably unfolded (Figure 6.4). This was in contrast to wild type A₁R produced in mammalian cells which was predominantly localised to the cell surface (Figure 6.4). A western blot using different detergents showed that there was a significantly stronger signal observed for the harsh detergents SDS and FC12 when A₁R-GL26 was expressed in either mammalian or insect cells (Figure 6.4). [³H]-DPCPX binding showed that, similar to A₁R binding, the signal was highest for membrane-bound A₁R-GL26 and only a weak signal was detected for receptor solubilised with the mild detergents DDM and digitonin (Figure 6.5). When combined with the radioligand binding assays, the differential signal intensity seen in the western blot when harsh detergents were used indicated that the majority of the receptor was unfolded in both expression systems.

Table 6.1 Percentage identity of transmembrane regions of human adenosine receptors

	A ₁ R	A _{2A} R	A _{2B} R	A ₃ R
A ₁ R		51	48	50
A _{2A} R	51		63	42
A _{2B} R	48	63		40
A ₃ R	50	42	40	

Table generated using ClustalW2¹⁹¹.

CHAPTER 6 Detection Of Unfolded Membrane Proteins

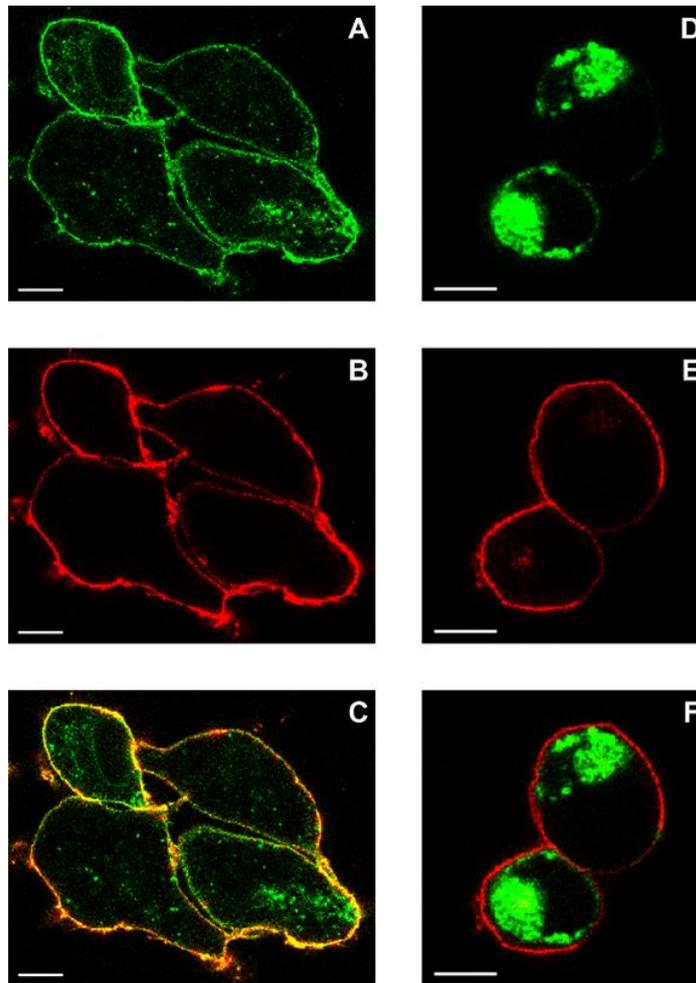


Figure 6.4 A₁R-GL26 is misfolded when expressed in mammalian cells

(A-C) Confocal micrographs of the iGnTI⁻(A₁R-GFP-H₁₀) cell line after 24 hours induction with tetracycline. Cells were fixed using paraformaldehyde and the plasma membrane was defined by staining with Alexa Fluor 647-conjugated con A prior to visualisation. Unlabelled iGnTI⁻ parental cells showed no fluorescence (not shown). The scale bar represents 10 μm. **(D-F)** Confocal micrographs of the iGnTI⁻(A₁R-GL26-GFP-H₁₀) cell line after 24 hours induction with tetracycline. Cells were fixed using paraformaldehyde and the plasma membrane was defined by staining with Alexa Fluor 647-conjugated con A prior to visualisation. Unlabelled iGnTI⁻ parental cells showed no fluorescence (not shown). The scale bar represents 10 μm.

CHAPTER 6 Detection Of Unfolded Membrane Proteins

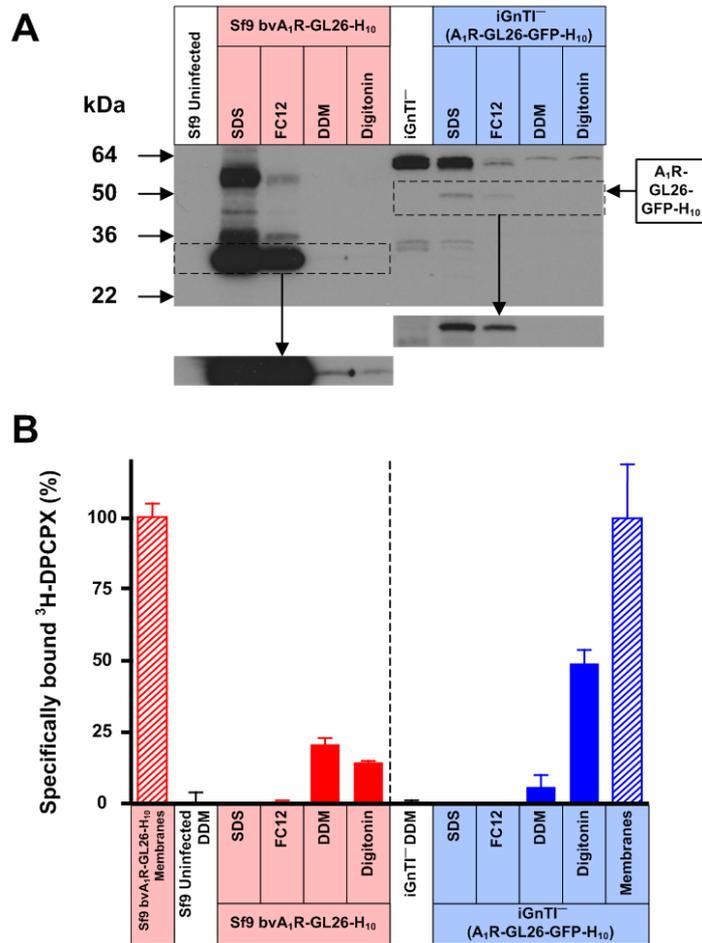


Figure 6.5 Misfolded A₁R-GL26 is poorly solubilised by either DDM or digitonin

(A) Western blot of A₁R-GL26 solubilised from whole cells using four different detergents (SDS, FC12, DDM or digitonin) and probed with an anti-pentaHis-HRP conjugated antibody. Each lane contains an equal amount of total protein. A₁R was expressed either in the stable mammalian cell line, iGnT⁻(A₁R-GL26-GFP-H₁₀), or by using the recombinant baculovirus bvA₁R-GL26-H₁₀ to infect Sf9 cells. The iGnT⁻ cell line was induced with 1 μg/ml tetracycline for 24 hours and Sf9 cells were infected for 72 hours. The western blot inserts are a 4 times longer exposure. **(B)** The amount of functional detergent-solubilised A₁R-GL26 was determined by measuring specific binding of the antagonist [³H]-DPCPX. After the addition of ligand, membranes were solubilised in the detergent indicated and non-bound ligand was separated from receptor-ligand complex on gel filtration spin columns and measured by liquid scintillation counting: filled red bars, A₁R-GL26 expressed in Sf9 cells; filled blue bars, A₁R-GL26 expressed in iGnT⁻ cells. The amount of A₁R-GL26 in membranes (non-solubilised) was determined by separation of receptor-bound and free radioligand by filtration through glass fibre plates: hatched red bars, A₁R-GL26 expressed in Sf9 cells; hatched blue bars, A₁R-GL26 expressed in iGnT⁻ cells. For ease of comparison, binding data have been normalised with respect to A₁R-GL26 in membranes (100%), which is equivalent to 17,400 ± 800 dpm (n=3; 435 fmoles/million cells) for baculovirus-infected Sf9 cells and 2,000 ± 350 (n=2; 48 fmoles/million cells) for iGnT⁻(A₁R-GL26-GFP-H₁₀) cells. Absolute levels of A₁R-GL26 therefore cannot be compared meaningfully between the two expression systems using this bar graph. Binding assays for A₁R-GL26 contained 150,000 iGnT⁻ or Sf9 cells. Each data point was determined in duplicate or triplicate from a single experiment and was plotted as mean ± SEM.

CHAPTER 6 Detection Of Unfolded Membrane Proteins

6.3.4 Differential solubility of β_1 AR and SERT

To further verify the accuracy of the differential solubility assay, two additional membrane proteins were examined, the turkey β_1 adrenergic receptor (β_1 AR) and the rat serotonin transporter (SERT). β_1 AR was examined because its structure has been determined numerous times bound to several different ligands^{25, 52, 71, 72, 73} and, in all cases, the recombinant protein was produced using the baculovirus system^{114, 165}. Therefore it was interesting to determine whether the baculovirus expression system produced misfolded β_1 AR as it did with AT_1R and A_1R . A western blot of wild type β_1 AR with N-terminal and C-terminal deletions solubilised in different detergents showed that there was more peptide solubilised by the harsh detergents SDS and DDM in comparison to the more mild detergents DDM and digitonin (Figure 6.6). The [³H]-DHA binding assay showed that β_1 AR was functional when solubilised in the detergents digitonin, DDM and FC12 (Figure 6.6). When SDS was used no [³H]-DHA binding was observed (Figure 6.6). Combining the results of the binding assay with the strong signal seen on the western blot for the SDS condition indicated that there was misfolded β_1 AR present, but considerably less than observed for either AT_1R or A_1R . A thermostable version of β_1 AR fused to thioredoxin (ts β_1 AR) also showed evidence of misfolded receptor with considerably more receptor solubilised by SDS and FC12 than either digitonin or DDM (Figure 6.6). Interestingly, receptor that contained an uncleaved leader sequence was solubilised only by SDS or FC12, suggesting that it was predominantly misfolded.

The final membrane protein to be examined was SERT. It had previously been shown that the majority of SERT produced in the baculovirus system was misfolded and the most efficient system for producing it was inducible HEK293 cells⁶⁰. Therefore it was interesting to determine whether this result was replicated with the detergent assay. Indeed, the pattern of a strong western blot signal in SDS and FC12 with no [¹²⁵I]-RTI-55 binding observed for the FC12 condition (SDS binding was not measured) was repeated for SERT produced in the baculovirus system (Figure 6.7). This was not the case for SERT produced in the mammalian system where an equal signal was seen on the western blot for all detergent conditions tested (Figure 6.7). Once more, the detergent assay described here detected the presence of unfolded SERT when produced using the baculovirus system but not the mammalian system.

CHAPTER 6 Detection Of Unfolded Membrane Proteins

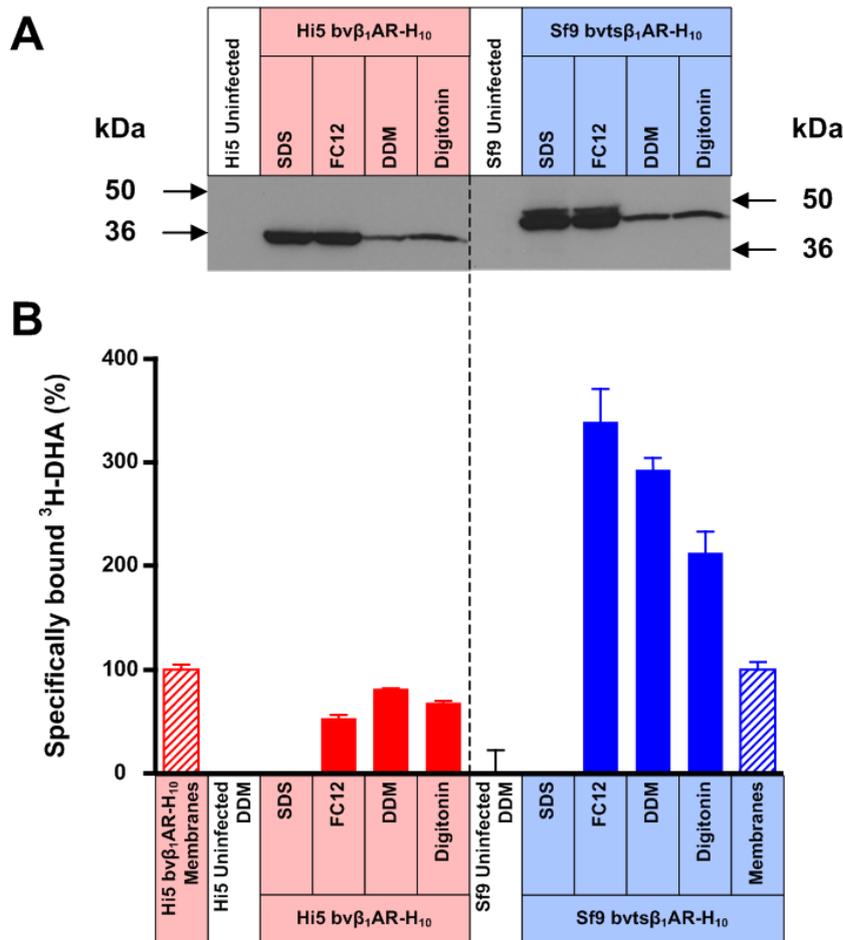


Figure 6.6 Misfolded β_1 AR is poorly solubilised by either DDM or digitonin

(A) Western blot of β_1 AR solubilised from whole cells using four different detergents (SDS, FC12, DDM or digitonin) and probed with an anti-pentaHis-HRP conjugated antibody. Each lane contains an equal amount of total protein. β_1 AR was expressed by using the recombinant baculovirus bv β_1 AR-H₁₀ to infect Hi5 cells. ts β_1 AR was expressed by using the recombinant baculovirus bvts β_1 AR-H₁₀ to infect Sf9 cells. Hi5 and Sf9 cells were infected for 48 hours. The dashed line indicates two separate blots. **(B)** The amount of functional detergent-solubilised β_1 AR and ts β_1 AR was determined by measuring specific binding of the antagonist [³H]-DHA. After the addition of ligand, membranes were solubilised in the detergent indicated and non-bound ligand was separated from receptor-ligand complex on gel filtration spin columns and measured by liquid scintillation counting: filled red bars, β_1 AR expressed in Hi5 cells; filled blue bars, ts β_1 AR expressed in Sf9 cells. The amount of β_1 AR in membranes (non-solubilised) was determined by separation of receptor-bound and free radioligand by filtration through glass fibre plates: hatched red bars, β_1 AR expressed in Hi5 cells; hatched blue bars, ts β_1 AR expressed in Sf9 cells. For ease of comparison, binding data have been normalised with respect to β_1 AR in membranes (100%), which is equivalent to $11,000 \pm 550$ dpm ($n=3$; 6.1 pmoles/million cells) for baculovirus-infected Hi5 cells and $2,600 \pm 190$ ($n=3$; 1.4 pmoles/L) for bvts β_1 AR-H₁₀ infected Sf9 cells. Absolute levels of β_1 AR therefor cannot be compared meaningfully between the two constructs using this bar graph. All binding assays for β_1 AR and ts β_1 AR contained 8,300 cells. Each data point was determined in duplicate or triplicate from a single experiment and was plotted as mean \pm SEM.

CHAPTER 6 Detection Of Unfolded Membrane Proteins

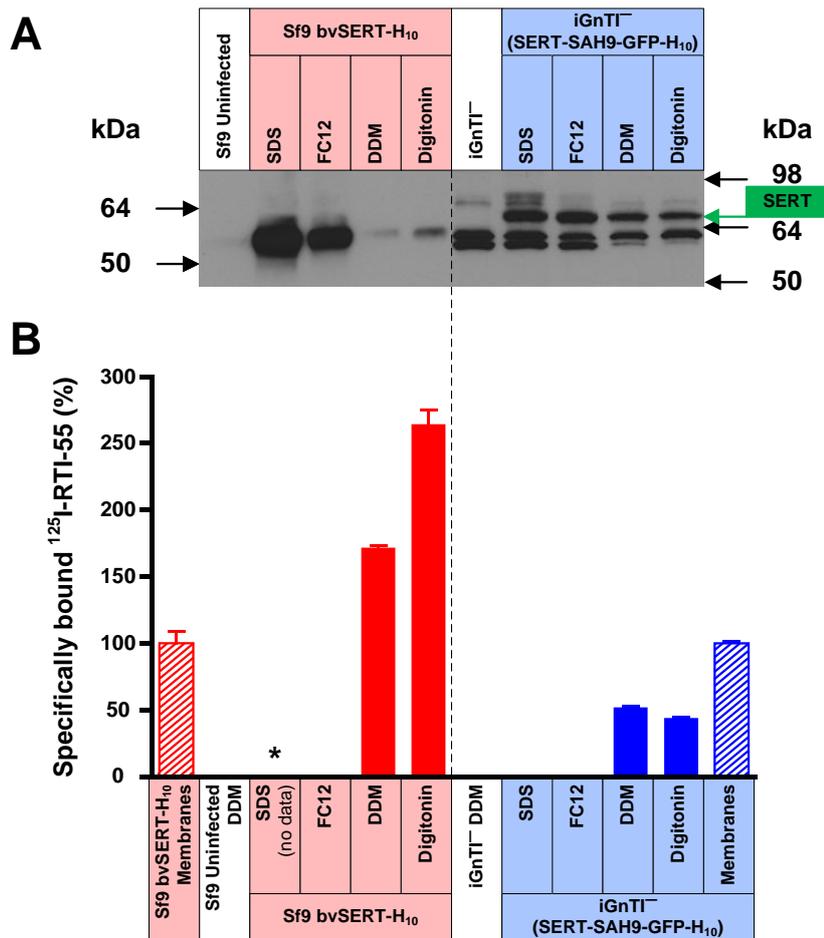


Figure 6.7 Misfolded SERT produced by the baculovirus expression system is poorly solubilised by either DDM or digitonin

(A) Western blot of SERT solubilised from whole cells using four different detergents (SDS, FC12, DDM or digitonin) and probed with an anti-pentaHis-HRP conjugated antibody. Each lane contains an equal amount of total protein. SERT was expressed either in the stable mammalian cell line iGnTI⁻ (SERT-SAH9-GFP-H₁₀) or by using the recombinant baculovirus bvSERT-H₁₀ to infect Sf9 cells. The iGnTI⁻ cell line was induced with 1 µg/ml tetracycline for 24 hours and Sf9 cells were infected for 48 hours. The dashed line indicates separate blots. **(B)** The amount of functional detergent-solubilised SERT was determined by measuring specific binding of the ligand [¹²⁵I]-RTI-55. After the addition of ligand, membranes were solubilised in the detergent indicated and non-bound ligand was separated from receptor-ligand complex on gel filtration spin columns and measured by liquid scintillation counting: filled red bars, SERT expressed in Sf9 cells; filled blue bars, SERT-SAH9 expressed in iGnTI⁻ cells; *, not determined. The amount of SERT in membranes (non-solubilised) was determined by separation of receptor-bound and free radioligand by filtration through glass fibre plates: hatched red bars, SERT expressed in Sf9 cells; hatched blue bars, SERT-SAH9 expressed in iGnTI⁻ cells. For ease of comparison, binding data have been normalised with respect to SERT in membranes (100%), which is equivalent to 10,200 ± 950 dpm (n=2; 75.7 fmoles/million cells) for baculovirus-infected Sf9 cells and 35,400 ± 420 dpm (n=2; 730 fmoles/million cells) for iGnTI⁻ (Sert-SAH9-GFP-H₁₀) cells. Absolute levels of SERT therefore cannot be compared meaningfully between the two expression systems using this bar graph. Binding assays for SERT contained either 28,000 Sf9 cells or 10,000 iGnTI⁻ cells. Each data point was determined in duplicate or triplicate from a single experiment and was plotted as mean ± SEM.

6.4 Discussion

When considering the results of the differential solubility experiments, it is interesting to note that all of the membrane proteins produced using the baculovirus system suffered from a varying degree of misfolding. This was particularly notable for β_1 AR produced using the baculovirus system, a system which has been used to determine several structures of this protein with different ligands bound^{25, 52, 71, 72, 73}. However, the amount of misfolded material present was much less than for the other baculovirus-generated membranes proteins examined, which might account for the ability of β_1 AR to be crystallised. It is possible that aggregated material is removed during the purification process through the use of filters or size exclusion chromatography, or that the misfolded material might precipitate during crystallisation. In contrast, the mammalian system produced no detectable misfolded material for any of the membrane proteins examined. However, mammalian expression using stable cell lines is not suitable in all circumstances. For A_1 R, large scale growth using stable mammalian cell lines would potentially be hindered by the slow growth rate of the cells, presumably due to high basal activity of the receptor. One potential way around this would be to create an A_1 R-T4L fusion in intracellular loop three, which would prevent G protein coupling⁴¹. Another possibility is to express A_1 R in a mammalian system but, instead of creating stable cell lines, a virus could be used to infect cells once they reach a high enough density. Recent success using the BacMam expression system¹⁹² suggests that this could be useful for the expression of A_1 R.

The technique of using detergents of different levels of harshness to solubilise a membrane protein prior to western blotting and examining the resulting signal intensity creates a rapid assay for detecting the presence of unfolded membrane proteins. Given that consistent results were obtained for four different membrane proteins there is no longer a need to couple the western blot with radioligand binding assays as it is indicated that the western blot assay alone would reliably predict the presence of misfolded protein in other cases. Therefore this assay could be of great benefit when dealing with proteins for which a ligand does not exist or where radioligand binding experiments are not possible. The western blot assay described here can be simplified further by using just two detergents, one harsh (SDS) and one mild (digitonin) which should provide a sufficient indication of the presence of misfolded material. There has been an emphasis on screening membrane proteins in a high-throughput system to find a candidate suitable for structural studies¹⁹³. The method described here can be easily used with high-throughput screening of eukaryotic membrane

CHAPTER 6 Detection Of Unfolded Membrane Proteins

proteins. It also can be used to rapidly determine the best choice of expression system. Therefore, this simple and rapid assay should be of great value to the field. This assay would also be useful for people who are relatively new to the field of membrane proteins since it is a simple diagnostic tool to use. Finally, if the yield of a purification is significantly lower than what was expected then this could be due to the presence of unfolded material which precipitated over the course of the purification. This assay will help to avoid misleading conclusions being drawn in these circumstances.

CHAPTER 7 OVERALL DISCUSSION AND CONCLUSIONS

For most GPCRs, the amount of protein that can be isolated from native sources is significantly less than would be necessary for structure determination^{57, 58}. For this reason, a necessary precursor to structural projects is the overexpression of the GPCR in a relevant system. To date, the most frequently used overexpression system for GPCRs has been the baculovirus-infected insect cell system¹¹³. However, this system is not generally applicable to all GPCRs, which vary enormously in both their complexity and the number of their N-linked glycosylation sites. The technique adopted in this project of creating mammalian cell lines that stably express AT₁R was very successful, resulting in 26 million molecules of AT₁R per cell or approximately 2 mg of AT₁R per litre, assuming 1 million cells per ml. After further optimisation of the construct and through the use of a Wave Bioreactor system, the stable mammalian cell line iGNTI25 produced 6 mg of AT₁R per litre of culture, at a cell density of 3.4×10^6 cells per ml. Inducible expression of AT₁R in stable mammalian cell lines remained consistent despite being in culture for over twenty days and the expressed receptor was of high quality with virtually all of the AT₁R produced being N-glycosylated and virtually no misfolded receptor present. This was in stark contrast to AT₁R produced in the baculovirus system, which showed limited amounts of N-glycosylation, large amounts of misfolded material and only 1.8 million molecules of functional AT₁R per cell (0.1 mg of AT₁R per litre of culture, at a cell density of 1 million cells per ml). Given this, the mammalian system was shown to be superior to the baculovirus system for the overexpression of AT₁R.

Despite their advantages, there are a number of drawbacks to creating tetracycline-inducible stable mammalian cell lines to overexpress GPCRs for structural studies. There is a good probability that at least a few of the stable mammalian cell lines will result in favourable expression levels. However, out of fifteen cell lines created that stably expressed AT₁R, only five showed levels of expression consistent with obtaining the milligram quantities of AT₁R required for structural studies. Also, not all of the stable cell lines could be grown in large scale, for example the clonal cell line iHEK415 could not be grown efficiently in suspension. In addition, there is no reliable way to alter the protein once it has integrated into the genome, which makes it difficult to optimise a construct for crystallography. One possibility for overcoming this problem is to use the CRISPR / Cas9 system¹⁸⁶ to edit the genome of the stable cell line, although there are many potential problems in getting this system to work. Another major drawback is that inducible stable mammalian cell lines do

CHAPTER 7 Overall Discussion and Conclusions

not work for some GPCRs such as A₁R. Despite using an inducible system and an inverse agonist in the culture media, a cell line which stably expressed A₁R did not grow rapidly enough to enable large-scale growth, *i.e.* it took over two weeks for the cells to reach confluency in a T-75 flask. The reason for the poor growth of the A₁R cell line is unclear, but could be due to the basal activity of the receptor. A more general problem with using mammalian cells as an expression system is that the amount of time it takes to generate a stable cell line, approximately two months, presents a significant obstacle. One potential way to shorten the time frames is to use transient transfection. However, expression levels in transient transfection are often too low for purification purposes. Also, expression levels obtained using this method did not necessarily correlate with expression levels observed in stable cell lines. Testing multiple constructs in parallel is usually necessary before the optimum final construct for crystallisation is determined and this is particularly onerous if stable cell lines are used. For example, when inserting a small soluble protein such as T4L to increase the likelihood of crystal contacts forming, it would often be necessary to screen about a hundred constructs to obtain a highly expressed chimera with no flexible regions^{41, 62} and this would not be feasible using stable cell lines. Transient transfection can be used in a high-throughput mode, however the difficulties described above, such as low expression, can be problematic. Another possibility for screening hundreds of constructs is to use a viral delivery system that expresses the gene of interest. Unlike transient transfection, viral systems efficiently infect all of the cells in a population and thereby normally achieve high levels of expression. One such system is the baculovirus-mediated gene transduction of mammalian cells (BacMam) system¹⁹². This system was used to overproduce an engineered dopamine transporter¹²¹ and both the NMDA¹²³ and AMPA¹²⁴ receptors. This approach allowed the structures of these three proteins to be determined, thus proving the utility of the BacMam system for the expression of integral membrane proteins. Although the BacMam system is compatible with high-throughput methods, it can still take up to one month to generate a baculovirus. Another virus system that can be used to create high level recombinant protein expression in mammalian cells and which can also be used in a high-throughput capacity is the lentivirus system^{194, 195, 196}. The drawbacks of the lentivirus system are that it has not yet been tested for high level expression of mammalian membrane proteins and that its use often requires working under biosafety level 2 conditions, which would be difficult with large-scale cultures. An issue common to several viral expression systems is that they may express predominantly unfolded material in the endoplasmic reticulum of the host¹¹⁰, which would not be suitable for purification and crystallisation.

CHAPTER 7 Overall Discussion and Conclusions

Nevertheless, viral delivery systems seem worth investigating given that the alternative of creating stable cell lines is far from ideal.

From the data presented in this thesis, it is unclear whether N-linked glycosylation will prove to be an obstacle for crystallising AT₁R. It is possible to obtain well diffracting crystals of some proteins with glycosylation intact. For example, the first crystal structure of rhodopsin was obtained from native sources and with its complex N-glycans intact¹⁹. In addition, structure determination of constitutively active metarhodopsin-II relied on the presence of N-glycans as they formed crystal contacts¹¹⁹. Glycosylation can be a source of heterogeneity for crystallisation, thus potentially preventing the formation of well-diffracting crystals, but the N-glycans can be removed. Structures of the P2Y receptor¹⁹⁷,¹⁹⁸, dopamine D₃ receptor¹⁹⁹ and CXCR4 chemokine receptor²⁰⁰ were all obtained after the recombinant protein was fully deglycosylated using PNGase F. However AT₁R might be unstable with all of its sugars removed. If this proved to be the case, an alternative approach would be to express the receptor in an iGNTI cell line, which only produces a high mannose core. These core N-glycans can then be removed using Endo H, which leaves one GlcNAc residue covalently bound, which could potentially improve the stability of the protein. It is not possible to predict which strategy would be the most beneficial and this would have to be determined empirically. N-glycosylation could be removed by mutating one or more of the N-linked glycosylation sites of AT₁R to alanine, but it was not possible to remove all three N-glycosylation sites while maintaining expression levels. Another option would be to mutate the sites to an hydrophilic amino acid, such as Ser, Lys or Glu which might give very different results from using Ala. It might then be feasible to mutate all of the N-glycosylation sites without significantly affecting both stability and expression.

To assess the stability of AT₁R, this project used two criteria, the apparent T_m assay and FSEC. Using the [Super +] format apparent T_m assay, detergent-solubilised AT₁R was found to have a T_m of 47°C with the antagonist Sar¹ bound. This was deemed to be sufficiently stable to attempt crystallisation without further stabilisation. However, when Sar¹ was incubated with AT₁R overnight, the FSEC peak was dramatically reduced in signal in comparison to the receptor being incubated with no ligand overnight. This indicated that incubation of AT₁R with Sar¹ for an extended period of time was somehow destabilising the receptor and, on account of this, crystallisation with antagonist present might be problematic. There are several approaches to dealing with this problem. Conformational thermostabilisation^{17, 65, 66, 67, 68, 69}, particularly in combination with FSEC, might prove beneficial at stabilising AT₁R with ligand bound. It is interesting to note that the addition of

CHAPTER 7 Overall Discussion and Conclusions

T4L into CL3 of AT₁R increased the apparent T_m by 11°C when the [Super +] format assay was used in comparison to the wild type receptor. However when the apparent T_m was measured using the [-] format assay, T4L reduced the apparent T_m by 6°C in comparison to the wild type receptor. Therefore, in any future mutagenesis projects of AT₁R, it would be advisable to measure thermostability using both [Super +] and [-] format assays. Another possibility for increasing the stability of AT₁R with ligand bound is to perform a charge scan on the intracellular side of AT₁R which might remove unfavourable charge-charge interactions and allow additional hydrogen bonds to form. An alanine scan of the turkey β₁AR revealed that R68A was stabilising and R68S was even more stabilising⁶⁹. The reason for this was subsequently revealed in the β₁AR structure which showed that R68 was adjacent in three dimensional space to R355 on H8 and the mutant R68S enabled the formation of two hydrogen bonds between the intracellular domains of H1 and H8⁵². In theory, E68 might have been even more stabilising, however this was not tested. Salt bridges could also be introduced to increase thermostability. β₂AR was determined to be more thermostable than β₁AR¹⁰⁸ possibly due to the presence of a salt bridge between Asp192 and Lys305²³ which β₁AR lacks²⁵. This was engineered into β₁AR and shown to improve thermostability by 5°C⁶⁶. Another approach to stabilising AT₁R is to bind either a G protein or arrestin to the receptor. However, this would be extremely difficult since no applicable generic methodology exists and AT₁R couples most frequently to G_q, which is both difficult to express and unstable in detergent (Nehme and Tate, unpublished). Additionally, an antibody could be used to stabilise the intracellular face of AT₁R (Section 1.3.2). β₂AR was stabilised in the agonist conformation by a camelid antibody fragment (nanobody 80)⁴² and the A_{2A}R was locked into an inverse agonist position by the binding of a F_{ab} fragment (Fab2838) to its cytoplasmic face⁶¹. To produce an antibody that recognises AT₁R, purified protein would need to be obtained and a ligand that locks the receptor into one conformation would need to be identified.

The discrepancy between the apparent thermostability of AT₁R, as measured using a radioligand, and the FSEC data raises the question of why AT₁R is behaving so differently from receptors such as β₁AR and A_{2A}R. One possibility is that there is a time dependent dislocation between the ligand binding pocket and the intracellular side of the receptor. This behaviour might not be observed in membranes where transmembrane helices of the receptor are subjected to lateral pressure forces generated by the lipid bilayer. However, once the lateral pressure is removed by solubilisation with detergent, the receptor may

CHAPTER 7 Overall Discussion and Conclusions

become more dynamic. One method for examining the dynamics of a receptor is double electron-electron resonance (DEER) and this could be used to study the differences in AT₁R dynamics in detergent micelles compared to membranes. DEER has been used to map conformational changes in rhodopsin after binding visual arrestin²⁰¹. However, this method requires purified protein that contains a pair of spin labels. Another possibility for explaining the unusual behaviour of AT₁R is that, in terms of β -arrestin recruitment, it belongs to class B, which means that the receptor-arrestin binding interaction is strong and enduring, so that AT₁R internalises with β -arrestin still bound^{82, 202}. Whether class B receptors exhibit a different range or magnitude of conformational dynamics compared to class A receptors is unknown, but if they do then this could help to explain the unusual features of AT₁R stability. Another possibility is that the high concentration of Arg/Lys residues on the intracellular face of AT₁R makes this region less stable than the intracellular face of other receptors.

In order for structure determination of AT₁R to proceed, it is clear that the receptor needs to be stabilised either by mutagenesis or the inclusion of a binding protein as mentioned above. Once this is achieved, the work presented in this thesis demonstrates that the inducible mammalian system is the best for producing the milligram quantities of AT₁R needed for structural studies.

LIST OF REFERENCES

- 1 Fredriksson, R., Lagerstrom, M. C., Lundin, L. G. & Schioth, H. B. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Molecular pharmacology* **63**, 1256-1272, doi:10.1124/mol.63.6.1256 (2003).
- 2 Gether, U. Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. *Endocrine reviews* **21**, 90-113, doi:10.1210/edrv.21.1.0390 (2000).
- 3 Rosenbaum, D. M., Rasmussen, S. G. & Kobilka, B. K. The structure and function of G-protein-coupled receptors. *Nature* **459**, 356-363, doi:10.1038/nature08144 (2009).
- 4 Heng, B. C., Aubel, D. & Fussenegger, M. An overview of the diverse roles of G-protein coupled receptors (GPCRs) in the pathophysiology of various human diseases. *Biotechnology advances* **31**, 1676-1694, doi:10.1016/j.biotechadv.2013.08.017 (2013).
- 5 Lagerstrom, M. C. & Schioth, H. B. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nature reviews. Drug discovery* **7**, 339-357, doi:10.1038/nrd2518 (2008).
- 6 Schulte, G. International Union of Basic and Clinical Pharmacology. LXXX. The class Frizzled receptors. *Pharmacological reviews* **62**, 632-667, doi:10.1124/pr.110.002931 (2010).
- 7 Isberg, V. *et al.* GPCRDB: an information system for G protein-coupled receptors. *Nucleic acids research* **42**, D422-425, doi:10.1093/nar/gkt1255 (2014).
- 8 Congreve, M., Langmead, C. J., Mason, J. S. & Marshall, F. H. Progress in structure based drug design for G protein-coupled receptors. *Journal of medicinal chemistry* **54**, 4283-4311, doi:10.1021/jm200371q (2011).
- 9 Venkatakrisnan, A. *et al.* Structured and disordered facets of the GPCR fold. *Current opinion in structural biology* **27C**, 129-137, doi:10.1016/j.sbi.2014.08.002 (2014).
- 10 Hopkins, A. L. & Groom, C. R. The druggable genome. *Nature reviews. Drug discovery* **1**, 727-730, doi:10.1038/nrd892 (2002).
- 11 Galandrin, S., Oligny-Longpre, G. & Bouvier, M. The evasive nature of drug efficacy: implications for drug discovery. *Trends in pharmacological sciences* **28**, 423-430, doi:10.1016/j.tips.2007.06.005 (2007).
- 12 Kobilka, B. K. & Deupi, X. Conformational complexity of G-protein-coupled receptors. *Trends in pharmacological sciences* **28**, 397-406, doi:10.1016/j.tips.2007.06.003 (2007).
- 13 Tate, C. G. A crystal clear solution for determining G-protein-coupled receptor structures. *Trends in biochemical sciences* **37**, 343-352, doi:10.1016/j.tibs.2012.06.003 (2012).
- 14 Liu, W. *et al.* Structural basis for allosteric regulation of GPCRs by sodium ions. *Science* **337**, 232-236, doi:10.1126/science.1219218 (2012).
- 15 Gao, Z. G., Kim, S. K., Ijzerman, A. P. & Jacobson, K. A. Allosteric modulation of the adenosine family of receptors. *Mini reviews in medicinal chemistry* **5**, 545-553 (2005).
- 16 Jaakola, V. P. *et al.* The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **322**, 1211-1217, doi:10.1126/science.1164772 (2008).
- 17 Magnani, F., Shibata, Y., Serrano-Vega, M. J. & Tate, C. G. Co-evolving stability and conformational homogeneity of the human adenosine A2a receptor. *Proceedings*

REFERENCES

- of the National Academy of Sciences of the United States of America **105**, 10744-10749, doi:10.1073/pnas.0804396105 (2008).
- 18 Shoichet, B. K. & Kobilka, B. K. Structure-based drug screening for G-protein-coupled receptors. *Trends in pharmacological sciences* **33**, 268-272, doi:10.1016/j.tips.2012.03.007 (2012).
- 19 Palczewski, K. *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**, 739-745 (2000).
- 20 Manglik, A. & Kobilka, B. The role of protein dynamics in GPCR function: insights from the beta2AR and rhodopsin. *Current opinion in cell biology* **27**, 136-143, doi:10.1016/j.ceb.2014.01.008 (2014).
- 21 Kobilka, B. K. G protein coupled receptor structure and activation. *Biochimica et biophysica acta* **1768**, 794-807, doi:10.1016/j.bbamem.2006.10.021 (2007).
- 22 Okada, T. *et al.* The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. *Journal of molecular biology* **342**, 571-583, doi:10.1016/j.jmb.2004.07.044 (2004).
- 23 Rasmussen, S. G. *et al.* Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature* **450**, 383-387, doi:10.1038/nature06325 (2007).
- 24 Cherezov, V. *et al.* High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **318**, 1258-1265, doi:10.1126/science.1150577 (2007).
- 25 Warne, T. *et al.* Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* **454**, 486-491, doi:10.1038/nature07101 (2008).
- 26 Rasmussen, S. G. *et al.* Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature* **477**, 549-555, doi:10.1038/nature10361 (2011).
- 27 Steyaert, J. & Kobilka, B. K. Nanobody stabilization of G protein-coupled receptor conformational states. *Current opinion in structural biology* **21**, 567-572, doi:10.1016/j.sbi.2011.06.011 (2011).
- 28 Hollenstein, K. *et al.* Structure of class B GPCR corticotropin-releasing factor receptor 1. *Nature* **499**, 438-443, doi:10.1038/nature12357 (2013).
- 29 Siu, F. Y. *et al.* Structure of the human glucagon class B G-protein-coupled receptor. *Nature* **499**, 444-449, doi:10.1038/nature12393 (2013).
- 30 Wu, H. *et al.* Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science* **344**, 58-64, doi:10.1126/science.1249489 (2014).
- 31 Dore, A. S. *et al.* Structure of class C GPCR metabotropic glutamate receptor 5 transmembrane domain. *Nature* **511**, 557-562, doi:10.1038/nature13396 (2014).
- 32 Wang, C. *et al.* Structure of the human smoothed receptor bound to an antitumour agent. *Nature* **497**, 338-343, doi:10.1038/nature12167 (2013).
- 33 Venkatakrisnan, A. J. *et al.* Molecular signatures of G-protein-coupled receptors. *Nature* **494**, 185-194, doi:10.1038/nature11896 (2013).
- 34 Stevens, R. C. *et al.* The GPCR Network: a large-scale collaboration to determine human GPCR structure and function. *Nature reviews. Drug discovery* **12**, 25-34, doi:10.1038/nrd3859 (2013).
- 35 Congreve, M., Dias, J. M. & Marshall, F. H. Structure-based drug design for G protein-coupled receptors. *Progress in medicinal chemistry* **53**, 1-63, doi:10.1016/b978-0-444-63380-4.00001-9 (2014).
- 36 Ballesteros, J. A. & Weinstein, H. in *Methods in Neurosciences* Vol. Volume 25 (ed C. Sealfon Stuart) 366-428 (Academic Press, 1995).
- 37 Katritch, V., Cherezov, V. & Stevens, R. C. Structure-function of the G protein-coupled receptor superfamily. *Annual review of pharmacology and toxicology* **53**, 531-556, doi:10.1146/annurev-pharmtox-032112-135923 (2013).
- 38 Standfuss, J. *et al.* The structural basis of agonist-induced activation in constitutively active rhodopsin. *Nature* **471**, 656-660, doi:10.1038/nature09795 (2011).

REFERENCES

- 39 Choe, H. W. *et al.* Crystal structure of metarhodopsin II. *Nature* **471**, 651-655, doi:10.1038/nature09789 (2011).
- 40 Seifert, R. & Wenzel-Seifert, K. Constitutive activity of G-protein-coupled receptors: cause of disease and common property of wild-type receptors. *Naunyn-Schmiedeberg's archives of pharmacology* **366**, 381-416, doi:10.1007/s00210-002-0588-0 (2002).
- 41 Rosenbaum, D. M. *et al.* GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science* **318**, 1266-1273, doi:10.1126/science.1150609 (2007).
- 42 Rasmussen, S. G. *et al.* Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor. *Nature* **469**, 175-180, doi:10.1038/nature09648 (2011).
- 43 Haga, K. *et al.* Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature* **482**, 547-551, doi:10.1038/nature10753 (2012).
- 44 Kruse, A. C. *et al.* Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature* **504**, 101-106, doi:10.1038/nature12735 (2013).
- 45 Yao, X. J. *et al.* The effect of ligand efficacy on the formation and stability of a GPCR-G protein complex. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 9501-9506, doi:10.1073/pnas.0811437106 (2009).
- 46 Kobilka, B. K. Structural insights into adrenergic receptor function and pharmacology. *Trends in pharmacological sciences* **32**, 213-218, doi:10.1016/j.tips.2011.02.005 (2011).
- 47 Nygaard, R. *et al.* The dynamic process of beta(2)-adrenergic receptor activation. *Cell* **152**, 532-542, doi:10.1016/j.cell.2013.01.008 (2013).
- 48 Altenbach, C., Kusnetzow, A. K., Ernst, O. P., Hofmann, K. P. & Hubbell, W. L. High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 7439-7444, doi:10.1073/pnas.0802515105 (2008).
- 49 Lebon, G. *et al.* Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature* **474**, 521-525, doi:10.1038/nature10136 (2011).
- 50 Xu, F. *et al.* Structure of an agonist-bound human A2A adenosine receptor. *Science* **332**, 322-327, doi:10.1126/science.1202793 (2011).
- 51 Chung, K. Y. Structural Aspects of GPCR-G Protein Coupling. *Toxicological research* **29**, 149-155, doi:10.5487/TR.2013.29.3.149 (2013).
- 52 Warne, T., Edwards, P. C., Leslie, A. G. & Tate, C. G. Crystal structures of a stabilized beta1-adrenoceptor bound to the biased agonists bucindolol and carvedilol. *Structure* **20**, 841-849, doi:10.1016/j.str.2012.03.014 (2012).
- 53 Nobles, K. N. *et al.* Distinct phosphorylation sites on the beta(2)-adrenergic receptor establish a barcode that encodes differential functions of beta-arrestin. *Science signaling* **4**, ra51, doi:10.1126/scisignal.2001707 (2011).
- 54 Liu, J. J., Horst, R., Katritch, V., Stevens, R. C. & Wuthrich, K. Biased signaling pathways in beta2-adrenergic receptor characterized by 19F-NMR. *Science* **335**, 1106-1110, doi:10.1126/science.1215802 (2012).
- 55 Shukla, A. K. *et al.* Visualization of arrestin recruitment by a G-protein-coupled receptor. *Nature* **512**, 218-222, doi:10.1038/nature13430 (2014).
- 56 Deupi, X. & Kobilka, B. K. Energy landscapes as a tool to integrate GPCR structure, dynamics, and function. *Physiology* **25**, 293-303, doi:10.1152/physiol.00002.2010 (2010).
- 57 Akermoun, M. *et al.* Characterization of 16 human G protein-coupled receptors expressed in baculovirus-infected insect cells. *Protein expression and purification* **44**, 65-74, doi:10.1016/j.pep.2005.04.016 (2005).

REFERENCES

- 58 Massotte, D. G protein-coupled receptor overexpression with the baculovirus-insect cell system: a tool for structural and functional studies. *Biochimica et biophysica acta* **1610**, 77-89 (2003).
- 59 Tate, C. G. Overexpression of mammalian integral membrane proteins for structural studies. *FEBS letters* **504**, 94-98 (2001).
- 60 Tate, C. G. *et al.* Comparison of seven different heterologous protein expression systems for the production of the serotonin transporter. *Biochimica et biophysica acta* **1610**, 141-153 (2003).
- 61 Hino, T. *et al.* G-protein-coupled receptor inactivation by an allosteric inverse-agonist antibody. *Nature* **482**, 237-240, doi:10.1038/nature10750 (2012).
- 62 Mathew, E., Ding, F. X., Naider, F. & Dumont, M. E. Functional fusions of T4 lysozyme in the third intracellular loop of a G protein-coupled receptor identified by a random screening approach in yeast. *Protein engineering, design & selection : PEDS* **26**, 59-71, doi:10.1093/protein/gzs070 (2013).
- 63 Dore, A. S. *et al.* Structure of the adenosine A(2A) receptor in complex with ZM241385 and the xanthenes XAC and caffeine. *Structure* **19**, 1283-1293, doi:10.1016/j.str.2011.06.014 (2011).
- 64 Faham, S. & Bowie, J. U. Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *Journal of molecular biology* **316**, 1-6, doi:10.1006/jmbi.2001.5295 (2002).
- 65 Abdul-Hussein, S., Andrell, J. & Tate, C. G. Thermostabilisation of the serotonin transporter in a cocaine-bound conformation. *Journal of molecular biology* **425**, 2198-2207, doi:10.1016/j.jmb.2013.03.025 (2013).
- 66 Miller, J. L. & Tate, C. G. Engineering an ultra-thermostable beta(1)-adrenoceptor. *Journal of molecular biology* **413**, 628-638, doi:10.1016/j.jmb.2011.08.057 (2011).
- 67 Lebon, G., Bennett, K., Jazayeri, A. & Tate, C. G. Thermostabilisation of an agonist-bound conformation of the human adenosine A(2A) receptor. *Journal of molecular biology* **409**, 298-310, doi:10.1016/j.jmb.2011.03.075 (2011).
- 68 Shibata, Y. *et al.* Thermostabilization of the neurotensin receptor NTS1. *Journal of molecular biology* **390**, 262-277, doi:10.1016/j.jmb.2009.04.068 (2009).
- 69 Serrano-Vega, M. J., Magnani, F., Shibata, Y. & Tate, C. G. Conformational thermostabilization of the beta1-adrenergic receptor in a detergent-resistant form. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 877-882, doi:10.1073/pnas.0711253105 (2008).
- 70 White, J. F. *et al.* Structure of the agonist-bound neurotensin receptor. *Nature* **490**, 508-513, doi:10.1038/nature11558 (2012).
- 71 Miller, P. S. & Aricescu, A. R. Crystal structure of a human GABAA receptor. *Nature* **512**, 270-275, doi:10.1038/nature13293 (2014).
- 72 Warne, T. *et al.* The structural basis for agonist and partial agonist action on a beta(1)-adrenergic receptor. *Nature* **469**, 241-244, doi:10.1038/nature09746 (2011).
- 73 Moukhametzianov, R. *et al.* Two distinct conformations of helix 6 observed in antagonist-bound structures of a beta1-adrenergic receptor. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 8228-8232, doi:10.1073/pnas.1100185108 (2011).
- 74 Braun-Menendez, E. & Page, I. H. Suggested Revision of Nomenclature--Angiotensin. *Science* **127**, 242, doi:10.1126/science.127.3292.242-a (1958).
- 75 Guo, D. F., Sun, Y. L., Hamet, P. & Inagami, T. The angiotensin II type 1 receptor and receptor-associated proteins. *Cell research* **11**, 165-180, doi:10.1038/sj.cr.7290083 (2001).
- 76 de Gasparo, M., Catt, K. J., Inagami, T., Wright, J. W. & Unger, T. International union of pharmacology. XXIII. The angiotensin II receptors. *Pharmacological reviews* **52**, 415-472 (2000).

REFERENCES

- 77 Hunyady, L. & Catt, K. J. Pleiotropic AT1 receptor signaling pathways mediating physiological and pathogenic actions of angiotensin II. *Molecular endocrinology* **20**, 953-970, doi:10.1210/me.2004-0536 (2006).
- 78 Qian, H., Pipolo, L. & Thomas, W. G. Association of beta-Arrestin 1 with the type 1A angiotensin II receptor involves phosphorylation of the receptor carboxyl terminus and correlates with receptor internalization. *Molecular endocrinology* **15**, 1706-1719, doi:10.1210/mend.15.10.0714 (2001).
- 79 Oakley, R. H., Laporte, S. A., Holt, J. A., Barak, L. S. & Caron, M. G. Molecular determinants underlying the formation of stable intracellular G protein-coupled receptor-beta-arrestin complexes after receptor endocytosis*. *The Journal of biological chemistry* **276**, 19452-19460, doi:10.1074/jbc.M101450200 (2001).
- 80 Tohgo, A. *et al.* The stability of the G protein-coupled receptor-beta-arrestin interaction determines the mechanism and functional consequence of ERK activation. *The Journal of biological chemistry* **278**, 6258-6267, doi:10.1074/jbc.M212231200 (2003).
- 81 Thomas, W. G., Qian, H. & Smith, N. J. When 6 is 9: 'uncoupled' AT1 receptors turn signalling on its head. *Cellular and molecular life sciences : CMLS* **61**, 2687-2694, doi:10.1007/s00018-004-4245-2 (2004).
- 82 Pierce, K. L. & Lefkowitz, R. J. Classical and new roles of beta-arrestins in the regulation of G-protein-coupled receptors. *Nature reviews. Neuroscience* **2**, 727-733, doi:10.1038/35094577 (2001).
- 83 Betz, S. F. Disulfide bonds and the stability of globular proteins. *Protein science : a publication of the Protein Society* **2**, 1551-1558, doi:10.1002/pro.5560021002 (1993).
- 84 Standfuss, J. *et al.* Crystal structure of a thermally stable rhodopsin mutant. *Journal of molecular biology* **372**, 1179-1188, doi:10.1016/j.jmb.2007.03.007 (2007).
- 85 Balmforth, A. J., Lee, A. J., Warburton, P., Donnelly, D. & Ball, S. G. The conformational change responsible for AT1 receptor activation is dependent upon two juxtaposed asparagine residues on transmembrane helices III and VII. *The Journal of biological chemistry* **272**, 4245-4251 (1997).
- 86 Groblewski, T. *et al.* Mutation of Asn111 in the third transmembrane domain of the AT1A angiotensin II receptor induces its constitutive activation. *The Journal of biological chemistry* **272**, 1822-1826 (1997).
- 87 Noda, K. *et al.* The active state of the AT1 angiotensin receptor is generated by angiotensin II induction. *Biochemistry* **35**, 16435-16442, doi:10.1021/bi961593m (1996).
- 88 Gaborik, Z. *et al.* The role of a conserved region of the second intracellular loop in AT1 angiotensin receptor activation and signaling. *Endocrinology* **144**, 2220-2228, doi:10.1210/en.2002-0135 (2003).
- 89 Timmermans, P. B. *et al.* Angiotensin II receptors and angiotensin II receptor antagonists. *Pharmacological reviews* **45**, 205-251 (1993).
- 90 Ji, H., Leung, M., Zhang, Y., Catt, K. J. & Sandberg, K. Differential structural requirements for specific binding of nonpeptide and peptide antagonists to the AT1 angiotensin receptor. Identification of amino acid residues that determine binding of the antihypertensive drug losartan. *The Journal of biological chemistry* **269**, 16533-16536 (1994).
- 91 Kivlighn, S. D. *et al.* Discovery of L-162,313: a nonpeptide that mimics the biological actions of angiotensin II. *The American journal of physiology* **268**, R820-823 (1995).
- 92 Ojima, M. *et al.* In vitro antagonistic properties of a new angiotensin type 1 receptor blocker, azilsartan, in receptor binding and function studies. *The Journal of*

REFERENCES

- pharmacology and experimental therapeutics* **336**, 801-808, doi:10.1124/jpet.110.176636 (2011).
- 93 Wei, H. *et al.* Independent beta-arrestin 2 and G protein-mediated pathways for angiotensin II activation of extracellular signal-regulated kinases 1 and 2. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 10782-10787, doi:10.1073/pnas.1834556100 (2003).
- 94 Holloway, A. C. *et al.* Side-chain substitutions within angiotensin II reveal different requirements for signaling, internalization, and phosphorylation of type 1A angiotensin receptors. *Molecular pharmacology* **61**, 768-777 (2002).
- 95 Violin, J. D., Soergel, D. G., Boerrigter, G., Burnett, J. C., Jr. & Lark, M. W. GPCR biased ligands as novel heart failure therapeutics. *Trends in cardiovascular medicine* **23**, 242-249, doi:10.1016/j.tcm.2013.01.002 (2013).
- 96 Violin, J. D. *et al.* Selectively engaging beta-arrestins at the angiotensin II type 1 receptor reduces blood pressure and increases cardiac performance. *The Journal of pharmacology and experimental therapeutics* **335**, 572-579, doi:10.1124/jpet.110.173005 (2010).
- 97 Strachan, R. T. *et al.* Divergent transducer-specific molecular efficacies generate biased agonism at a G protein-coupled receptor (GPCR). *The Journal of biological chemistry* **289**, 14211-14224, doi:10.1074/jbc.M114.548131 (2014).
- 98 Wisler, J. W., Xiao, K., Thomsen, A. R. & Lefkowitz, R. J. Recent developments in biased agonism. *Current opinion in cell biology* **27**, 18-24, doi:10.1016/j.ceb.2013.10.008 (2014).
- 99 Inagaki, S. *et al.* Modulation of the interaction between neurotensin receptor NTS1 and Gq protein by lipid. *Journal of molecular biology* **417**, 95-111, doi:10.1016/j.jmb.2012.01.023 (2012).
- 100 Butler, P. J., Ubarretxena-Belandia, I., Warne, T. & Tate, C. G. The Escherichia coli multidrug transporter EmrE is a dimer in the detergent-solubilised state. *Journal of molecular biology* **340**, 797-808, doi:10.1016/j.jmb.2004.05.014 (2004).
- 101 Lau, F. W. & Bowie, J. U. A method for assessing the stability of a membrane protein. *Biochemistry* **36**, 5884-5892, doi:10.1021/bi963095j (1997).
- 102 Corin, K. *et al.* Structure and function analyses of the purified GPCR human vomeronasal type 1 receptor 1. *Scientific reports* **1**, 172, doi:10.1038/srep00172 (2011).
- 103 O'Malley, M. A., Naranjo, A. N., Lazarova, T. & Robinson, A. S. Analysis of adenosine A(2)a receptor stability: effects of ligands and disulfide bonds. *Biochemistry* **49**, 9181-9189, doi:10.1021/bi101155r (2010).
- 104 Thompson, A. A. *et al.* GPCR stabilization using the bicelle-like architecture of mixed sterol-detergent micelles. *Methods* **55**, 310-317, doi:10.1016/j.ymeth.2011.10.011 (2011).
- 105 Alexandrov, A. I., Mileni, M., Chien, E. Y., Hanson, M. A. & Stevens, R. C. Microscale fluorescent thermal stability assay for membrane proteins. *Structure* **16**, 351-359, doi:10.1016/j.str.2008.02.004 (2008).
- 106 Mancusso, R., Karpowich, N. K., Czyzewski, B. K. & Wang, D. N. Simple screening method for improving membrane protein thermostability. *Methods* **55**, 324-329, doi:10.1016/j.ymeth.2011.07.008 (2011).
- 107 Kawate, T. & Gouaux, E. Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. *Structure* **14**, 673-681, doi:10.1016/j.str.2006.01.013 (2006).
- 108 Serrano-Vega, M. J. & Tate, C. G. Transferability of thermostabilizing mutations between beta-adrenergic receptors. *Molecular membrane biology* **26**, 385-396, doi:10.3109/09687680903208239 (2009).

REFERENCES

- 109 Grisshammer, R. & Tate, C. G. Overexpression of integral membrane proteins for structural studies. *Quarterly reviews of biophysics* **28**, 315-422 (1995).
- 110 Andrell, J. & Tate, C. G. Overexpression of membrane proteins in mammalian cells for structural studies. *Molecular membrane biology* **30**, 52-63, doi:10.1009/09687688.2012.703703 (2013).
- 111 Egloff, P. *et al.* Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E655-662, doi:10.1073/pnas.1317903111 (2014).
- 112 Shimamura, T. *et al.* Structure of the human histamine H1 receptor complex with doxepin. *Nature* **475**, 65-70, doi:10.1038/nature10236 (2011).
- 113 Maeda, S. & Schertler, G. F. Production of GPCR and GPCR complexes for structure determination. *Current opinion in structural biology* **23**, 381-392, doi:10.1016/j.sbi.2013.04.006 (2013).
- 114 Warne, T., Chirside, J. & Schertler, G. F. Expression and purification of truncated, non-glycosylated turkey beta-adrenergic receptors for crystallization. *Biochimica et biophysica acta* **1610**, 133-140 (2003).
- 115 Shukla, A. K., Reinhart, C. & Michel, H. Comparative analysis of the human angiotensin II type 1a receptor heterologously produced in insect cells and mammalian cells. *Biochemical and biophysical research communications* **349**, 6-14, doi:10.1016/j.bbrc.2006.07.210 (2006).
- 116 Lanctot, P. M., Leclerc, P. C., Escher, E., Leduc, R. & Guillemette, G. Role of N-glycosylation in the expression and functional properties of human AT1 receptor. *Biochemistry* **38**, 8621-8627, doi:10.1021/bi9830516 (1999).
- 117 Vanderheyden, P. M., Fierens, F. L., De Backer, J. P., Fraeyman, N. & Vauquelin, G. Distinction between surmountable and insurmountable selective AT1 receptor antagonists by use of CHO-K1 cells expressing human angiotensin II AT1 receptors. *British journal of pharmacology* **126**, 1057-1065, doi:10.1038/sj.bjp.0702398 (1999).
- 118 Jayadev, S. *et al.* N-linked glycosylation is required for optimal AT1a angiotensin receptor expression in COS-7 cells. *Endocrinology* **140**, 2010-2017, doi:10.1210/endo.140.5.6689 (1999).
- 119 Deupi, X. *et al.* Stabilized G protein binding site in the structure of constitutively active metarhodopsin-II. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 119-124, doi:10.1073/pnas.1114089108 (2012).
- 120 Gruswitz, F. *et al.* Function of human Rh based on structure of RhCG at 2.1 Å. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 9638-9643, doi:10.1073/pnas.1003587107 (2010).
- 121 Penmatsa, A., Wang, K. H. & Gouaux, E. X-ray structure of dopamine transporter elucidates antidepressant mechanism. *Nature* **503**, 85-90, doi:10.1038/nature12533 (2013).
- 122 Hassaine, G. *et al.* X-ray structure of the mouse serotonin 5-HT₃ receptor. *Nature* **512**, 276-281, doi:10.1038/nature13552 (2014).
- 123 Lee, C. H. *et al.* NMDA receptor structures reveal subunit arrangement and pore architecture. *Nature* **511**, 191-197, doi:10.1038/nature13548 (2014).
- 124 Durr, K. L. *et al.* Structure and dynamics of AMPA receptor GluA2 in resting, pre-open, and desensitized states. *Cell* **158**, 778-792, doi:10.1016/j.cell.2014.07.023 (2014).
- 125 Yao, F. *et al.* Tetracycline repressor, tetR, rather than the tetR-mammalian cell transcription factor fusion derivatives, regulates inducible gene expression in mammalian cells. *Human gene therapy* **9**, 1939-1950, doi:10.1089/hum.1998.9.13-1939 (1998).

REFERENCES

- 126 Reeves, P. J., Kim, J. M. & Khorana, H. G. Structure and function in rhodopsin: a tetracycline-inducible system in stable mammalian cell lines for high-level expression of opsin mutants. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 13413-13418, doi:10.1073/pnas.212519199 (2002).
- 127 Xiao, S., White, J. F., Betenbaugh, M. J., Grisshammer, R. & Shiloach, J. Transient and stable expression of the neurotensin receptor NTS1: a comparison of the baculovirus-insect cell and the T-REx-293 expression systems. *PloS one* **8**, e63679, doi:10.1371/journal.pone.0063679 (2013).
- 128 Reeves, P. J., Callewaert, N., Contreras, R. & Khorana, H. G. Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 13419-13424, doi:10.1073/pnas.212519299 (2002).
- 129 Chalfie, M. Green fluorescent protein. *Photochemistry and photobiology* **62**, 651-656 (1995).
- 130 Tsien, R. Y. The green fluorescent protein. *Annual review of biochemistry* **67**, 509-544, doi:10.1146/annurev.biochem.67.1.509 (1998).
- 131 Tate, C. G. Baculovirus-mediated expression of neurotransmitter transporters. *Methods in enzymology* **296**, 443-455 (1998).
- 132 Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry* **72**, 248-254 (1976).
- 133 Scott, D. J., Kummer, L., Tremmel, D. & Pluckthun, A. Stabilizing membrane proteins through protein engineering. *Current opinion in chemical biology* **17**, 427-435, doi:10.1016/j.cbpa.2013.04.002 (2013).
- 134 Robertson, N. *et al.* The properties of thermostabilised G protein-coupled receptors (StaRs) and their use in drug discovery. *Neuropharmacology* **60**, 36-44, doi:10.1016/j.neuropharm.2010.07.001 (2011).
- 135 Tate, C. G. Practical considerations of membrane protein instability during purification and crystallisation. *Methods in molecular biology* **601**, 187-203, doi:10.1007/978-1-60761-344-2_12 (2010).
- 136 Sonoda, Y. *et al.* Tricks of the trade used to accelerate high-resolution structure determination of membrane proteins. *FEBS letters* **584**, 2539-2547, doi:10.1016/j.febslet.2010.04.015 (2010).
- 137 Thompson, A. A. *et al.* Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic. *Nature* **485**, 395-399, doi:10.1038/nature11085 (2012).
- 138 Aricescu, A. R. & Owens, R. J. Expression of recombinant glycoproteins in mammalian cells: towards an integrative approach to structural biology. *Current opinion in structural biology* **23**, 345-356, doi:10.1016/j.sbi.2013.04.003 (2013).
- 139 Kaushal, S., Ridge, K. D. & Khorana, H. G. Structure and function in rhodopsin: the role of asparagine-linked glycosylation. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 4024-4028 (1994).
- 140 Tate, C. G. & Blakely, R. D. The effect of N-linked glycosylation on activity of the Na(+)- and Cl(-)-dependent serotonin transporter expressed using recombinant baculovirus in insect cells. *The Journal of biological chemistry* **269**, 26303-26310 (1994).
- 141 Lou, M. *et al.* The first three domains of the insulin receptor differ structurally from the insulin-like growth factor 1 receptor in the regions governing ligand specificity. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 12429-12434, doi:10.1073/pnas.0605395103 (2006).

REFERENCES

- 142 McKern, N. M. *et al.* Structure of the insulin receptor ectodomain reveals a folded-
over conformation. *Nature* **443**, 218-221, doi:10.1038/nature05106 (2006).
- 143 Grisshammer, R. Understanding recombinant expression of membrane proteins.
Current opinion in biotechnology **17**, 337-340, doi:10.1016/j.copbio.2006.06.001
(2006).
- 144 Sarramegna, V., Talmont, F., Demange, P. & Milon, A. Heterologous expression of
G-protein-coupled receptors: comparison of expression systems from the standpoint
of large-scale production and purification. *Cellular and molecular life sciences* :
CMLS **60**, 1529-1546 (2003).
- 145 Tate, C. G. & Schertler, G. F. Engineering G protein-coupled receptors to facilitate
their structure determination. *Current opinion in structural biology* **19**, 386-395,
doi:10.1016/j.sbi.2009.07.004 (2009).
- 146 Apweiler, R., Hermjakob, H. & Sharon, N. On the frequency of protein
glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochimica
et biophysica acta* **1473**, 4-8 (1999).
- 147 Chang, V. T. *et al.* Glycoprotein structural genomics: solving the glycosylation
problem. *Structure* **15**, 267-273, doi:10.1016/j.str.2007.01.011 (2007).
- 148 Chaudhary, S. *et al.* Efficient expression screening of human membrane proteins in
transiently transfected Human Embryonic Kidney 293S cells. *Methods* **55**, 273-280,
doi:10.1016/j.ymeth.2011.08.018 (2011).
- 149 Nettleship, J. E., Assenberg, R., Diprose, J. M., Rahman-Huq, N. & Owens, R. J.
Recent advances in the production of proteins in insect and mammalian cells for
structural biology. *Journal of structural biology* **172**, 55-65,
doi:10.1016/j.jsb.2010.02.006 (2010).
- 150 Mancia, F. *et al.* Optimization of protein production in mammalian cells with a
coexpressed fluorescent marker. *Structure* **12**, 1355-1360,
doi:10.1016/j.str.2004.06.012 (2004).
- 151 Ibrahim, S. F. & van den Engh, G. High-speed cell sorting: fundamentals and recent
advances. *Current opinion in biotechnology* **14**, 5-12 (2003).
- 152 Lohse, M. J. Stable overexpression of human beta 2-adrenergic receptors in
mammalian cells. *Naunyn-Schmiedeberg's archives of pharmacology* **345**, 444-451
(1992).
- 153 Chelikani, P., Reeves, P. J., Rajbhandary, U. L. & Khorana, H. G. The synthesis and
high-level expression of a beta2-adrenergic receptor gene in a tetracycline-inducible
stable mammalian cell line. *Protein science : a publication of the Protein Society* **15**,
1433-1440, doi:10.1110/ps.062080006 (2006).
- 154 Camponova, P. *et al.* High-level expression and purification of the human bradykinin
B(2) receptor in a tetracycline-inducible stable HEK293S cell line. *Protein
expression and purification* **55**, 300-311, doi:10.1016/j.pep.2007.04.020 (2007).
- 155 Reeves, P. J., Thurmond, R. L. & Khorana, H. G. Structure and function in
rhodopsin: high level expression of a synthetic bovine opsin gene and its mutants in
stable mammalian cell lines. *Proceedings of the National Academy of Sciences of the
United States of America* **93**, 11487-11492 (1996).
- 156 Grone, H. J., Simon, M. & Fuchs, E. Autoradiographic characterization of
angiotensin receptor subtypes in fetal and adult human kidney. *The American journal
of physiology* **262**, F326-331 (1992).
- 157 Hunyady, L. *et al.* Dependence of AT1 angiotensin receptor function on adjacent
asparagine residues in the seventh transmembrane helix. *Molecular pharmacology*
54, 427-434 (1998).
- 158 Van Der Hee, R. M., Deurholt, T., Gerhard, C. C. & De Groene, E. M. Comparison
of 3 AT1 receptor binding assays: filtration assay, ScreenReady Target, and WGA

REFERENCES

- Flashplate. *Journal of biomolecular screening* **10**, 118-126, doi:10.1177/1087057104271330 (2005).
- 159 de Gasparo, M. & Whitebread, S. Binding of valsartan to mammalian angiotensin AT1 receptors. *Regul Pept* **59**, 303-311 (1995).
- 160 Miura, S., Okabe, A., Matsuo, Y., Karnik, S. S. & Saku, K. Unique binding behavior of the recently approved angiotensin II receptor blocker azilsartan compared with that of candesartan. *Hypertension research : official journal of the Japanese Society of Hypertension* **36**, 134-139, doi:10.1038/hr.2012.147 (2013).
- 161 Mederski, W. W. *et al.* Non-peptide angiotensin II receptor antagonists: synthesis and biological activity of a series of novel 4,5-dihydro-4-oxo-3H-imidazo[4,5-c]pyridine derivatives. *Journal of medicinal chemistry* **37**, 1632-1645 (1994).
- 162 Miura, S. *et al.* Small molecules with similar structures exhibit agonist, neutral antagonist or inverse agonist activity toward angiotensin II type 1 receptor. *PLoS one* **7**, e37974, doi:10.1371/journal.pone.0037974 (2012).
- 163 Yoshida, K., Perich, R., Casley, D. J. & Johnston, C. I. Hypotensive effect of ZD7155, an angiotensin II receptor antagonist, parallels receptor occupancy in two-kidney, one-clip Goldblatt hypertensive rats. *Journal of hypertension* **16**, 645-655 (1998).
- 164 Chen, T. B., Lotti, V. J. & Chang, R. S. Characterization of the binding of [3H]L-158,809: a new potent and selective nonpeptide angiotensin II receptor (AT1) antagonist radioligand. *Molecular pharmacology* **42**, 1077-1082 (1992).
- 165 Warne, T., Serrano-Vega, M. J., Tate, C. G. & Schertler, G. F. Development and crystallization of a minimal thermostabilised G protein-coupled receptor. *Protein expression and purification* **65**, 204-213 (2009).
- 166 Kenakin, T. Functional selectivity through protean and biased agonism: who steers the ship? *Molecular pharmacology* **72**, 1393-1401, doi:10.1124/mol.107.040352 (2007).
- 167 Kenakin, T. Inverse, protean, and ligand-selective agonism: matters of receptor conformation. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **15**, 598-611, doi:10.1096/fj.00-0438rev (2001).
- 168 Kenakin, T. Pharmacological proteus? *Trends in pharmacological sciences* **16**, 256-258 (1995).
- 169 Gbahou, F. *et al.* Protean agonism at histamine H3 receptors in vitro and in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 11086-11091, doi:10.1073/pnas.1932276100 (2003).
- 170 Urbani, A. & Warne, T. A colorimetric determination for glycosidic and bile salt-based detergents: applications in membrane protein research. *Analytical biochemistry* **336**, 117-124, doi:10.1016/j.ab.2004.09.040 (2005).
- 171 von Heijne, G. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *Journal of molecular biology* **225**, 487-494 (1992).
- 172 von Heijne, G. & Gavel, Y. Topogenic signals in integral membrane proteins. *European journal of biochemistry / FEBS* **174**, 671-678 (1988).
- 173 von Heijne, G. Patterns of amino acids near signal-sequence cleavage sites. *European journal of biochemistry / FEBS* **133**, 17-21 (1983).
- 174 Zimmermann, R., Eyrich, S., Ahmad, M. & Helms, V. Protein translocation across the ER membrane. *Biochimica et biophysica acta* **1808**, 912-924, doi:10.1016/j.bbamem.2010.06.015 (2011).
- 175 Yang, Z. R., Thomson, R., McNeil, P. & Esnouf, R. M. RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. *Bioinformatics* **21**, 3369-3376, doi:10.1093/bioinformatics/bti534 (2005).

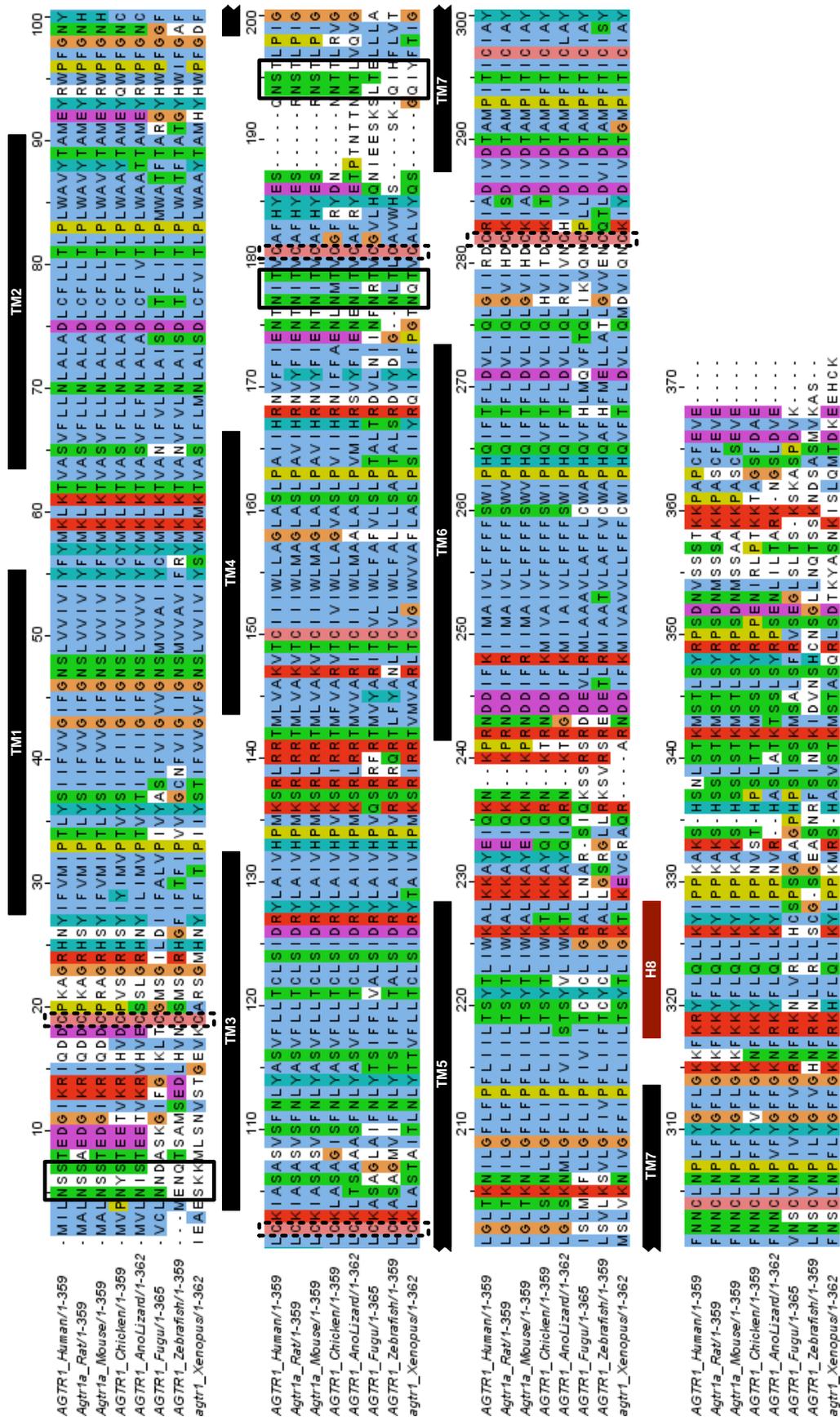
REFERENCES

- 176 Conchon, S., Peltier, N., Corvol, P. & Clauser, E. A noninternalized nondesensitized truncated AT1A receptor transduces an amplified ANG II signal. *The American journal of physiology* **274**, E336-345 (1998).
- 177 Gaborik, Z. *et al.* Requirement of membrane-proximal amino acids in the carboxyl-terminal tail for expression of the rat AT1a angiotensin receptor. *FEBS letters* **428**, 147-151 (1998).
- 178 Nilsson, B. *et al.* A synthetic IgG-binding domain based on staphylococcal protein A. *Protein engineering* **1**, 107-113 (1987).
- 179 Ton-That, H., Mazmanian, S. K., Faull, K. F. & Schneewind, O. Anchoring of Surface Proteins to the Cell Wall of Staphylococcus aureus : SORTASE CATALYZED IN VITRO TRANSPEPTIDATION REACTION USING LPXTG PEPTIDE AND NH₂-GLY₃SUBSTRATES. *Journal of Biological Chemistry* **275**, 9876-9881, doi:10.1074/jbc.275.13.9876 (2000).
- 180 Mao, H. A self-cleavable sortase fusion for one-step purification of free recombinant proteins. *Protein expression and purification* **37**, 253-263, doi:10.1016/j.pep.2004.06.013 (2004).
- 181 Popp, M. W. & Ploegh, H. L. Making and breaking peptide bonds: protein engineering using sortase. *Angewandte Chemie* **50**, 5024-5032, doi:10.1002/anie.201008267 (2011).
- 182 Levary, D. A., Parthasarathy, R., Boder, E. T. & Ackerman, M. E. Protein-protein fusion catalyzed by sortase A. *PloS one* **6**, e18342, doi:10.1371/journal.pone.0018342 (2011).
- 183 Drew, D., Lerch, M., Kunji, E., Slotboom, D. J. & de Gier, J. W. Optimization of membrane protein overexpression and purification using GFP fusions. *Nature methods* **3**, 303-313, doi:10.1038/nmeth0406-303 (2006).
- 184 Drew, D. *et al.* GFP-based optimization scheme for the overexpression and purification of eukaryotic membrane proteins in *Saccharomyces cerevisiae*. *Nature protocols* **3**, 784-798, doi:10.1038/nprot.2008.44 (2008).
- 185 Schlager, M. A., Hoang, H. T., Urnavicius, L., Bullock, S. L. & Carter, A. P. In vitro reconstitution of a highly processive recombinant human dynein complex. *The EMBO journal* **33**, 1855-1868, doi:10.15252/emj.201488792 (2014).
- 186 Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823, doi:10.1126/science.1231143 (2013).
- 187 Drew, D. E., von Heijne, G., Nordlund, P. & de Gier, J. W. Green fluorescent protein as an indicator to monitor membrane protein overexpression in *Escherichia coli*. *FEBS letters* **507**, 220-224 (2001).
- 188 Newstead, S., Kim, H., von Heijne, G., Iwata, S. & Drew, D. High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 13936-13941, doi:10.1073/pnas.0704546104 (2007).
- 189 Illing, M. E., Rajan, R. S., Bence, N. F. & Kopito, R. R. A rhodopsin mutant linked to autosomal dominant retinitis pigmentosa is prone to aggregate and interacts with the ubiquitin proteasome system. *The Journal of biological chemistry* **277**, 34150-34160, doi:10.1074/jbc.M204955200 (2002).
- 190 Saliba, R. S., Munro, P. M., Luthert, P. J. & Cheetham, M. E. The cellular fate of mutant rhodopsin: quality control, degradation and aggresome formation. *Journal of cell science* **115**, 2907-2918 (2002).
- 191 Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948, doi:10.1093/bioinformatics/btm404 (2007).
- 192 Dukkipati, A., Park, H. H., Waghray, D., Fischer, S. & Garcia, K. C. BacMam system for high-level expression of recombinant soluble and membrane

REFERENCES

- glycoproteins for structural studies. *Protein expression and purification* **62**, 160-170, doi:10.1016/j.pep.2008.08.004 (2008).
- 193 Mancia, F. & Love, J. High throughput platforms for structural genomics of integral membrane proteins. *Current opinion in structural biology* **21**, 517-522, doi:10.1016/j.sbi.2011.07.001 (2011).
- 194 Gaillet, B. *et al.* High-level recombinant protein production in CHO cells using lentiviral vectors and the cumate gene-switch. *Biotechnology and bioengineering* **106**, 203-215, doi:10.1002/bit.22698 (2010).
- 195 Matrai, J., Chuah, M. K. & VandenDriessche, T. Recent advances in lentiviral vector development and applications. *Molecular therapy : the journal of the American Society of Gene Therapy* **18**, 477-490, doi:10.1038/mt.2009.319 (2010).
- 196 Cockrell, A. S. & Kafri, T. Gene delivery by lentivirus vectors. *Molecular biotechnology* **36**, 184-204 (2007).
- 197 Zhang, K. *et al.* Structure of the human P2Y12 receptor in complex with an antithrombotic drug. *Nature* **509**, 115-118, doi:10.1038/nature13083 (2014).
- 198 Zhang, J. *et al.* Agonist-bound structure of the human P2Y12 receptor. *Nature* **509**, 119-122, doi:10.1038/nature13288 (2014).
- 199 Chien, E. Y. *et al.* Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* **330**, 1091-1095, doi:10.1126/science.1197410 (2010).
- 200 Wu, B. *et al.* Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **330**, 1066-1071, doi:10.1126/science.1194396 (2010).
- 201 Kim, M. *et al.* Conformation of receptor-bound visual arrestin. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 18407-18412, doi:10.1073/pnas.1216304109 (2012).
- 202 DeWire, S. M., Ahn, S., Lefkowitz, R. J. & Shenoy, S. K. Beta-arrestins and cell signaling. *Annu Rev Physiol* **69**, 483-510, doi:10.1146/annurev.ph.69.013107.100021 (2007).

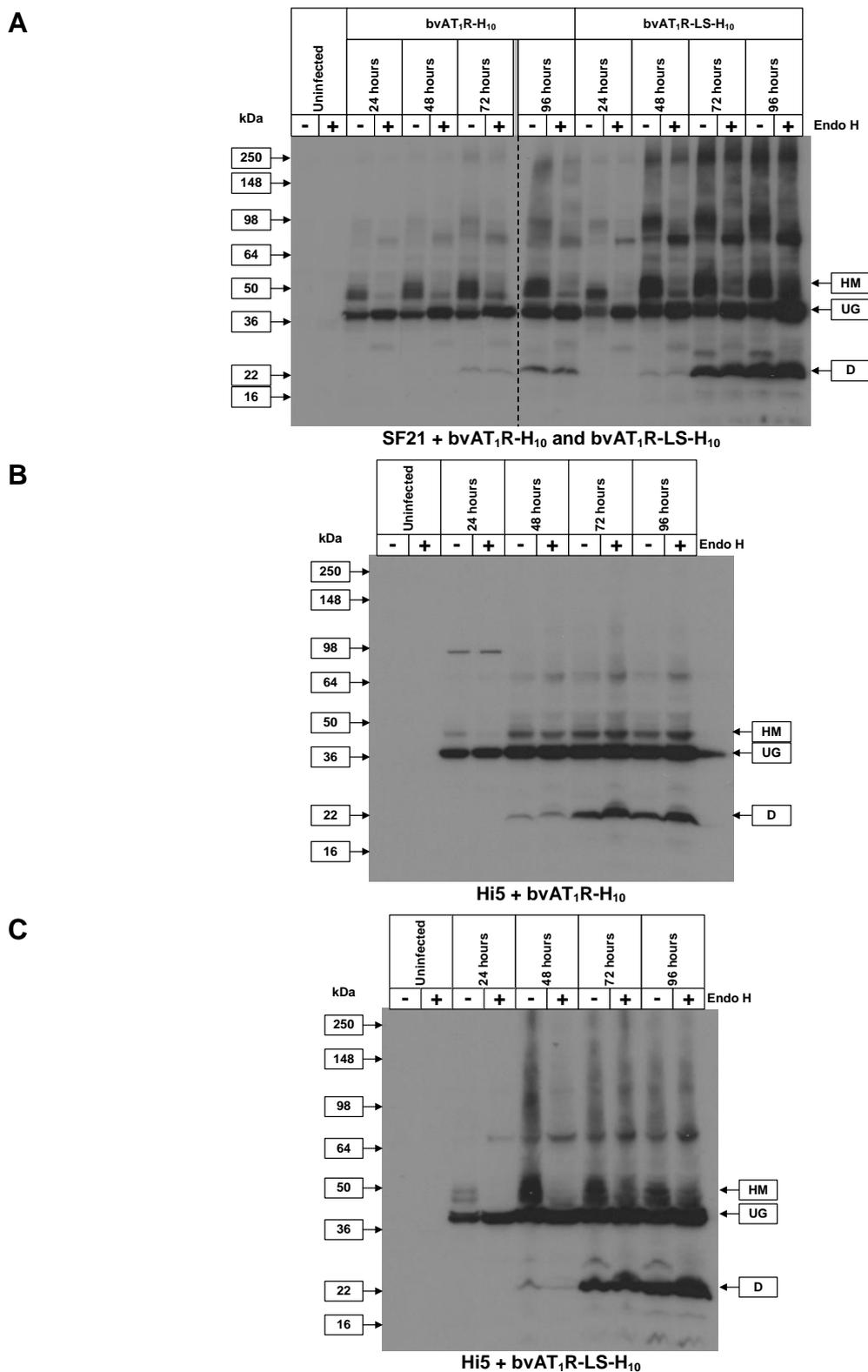
APPENDIX 1 AT₁R Amino Acid Sequence Alignments



Appendix Figure 2 Alignment of vertebrate AT₁Rs shows a high level of conservation

ClustalW alignment¹⁹¹ of human, rat, mouse, chicken, anole lizard, fugu, zebrafish and *Xenopus laevis* AT₁R amino acid sequences. Putative N-linked glycosylation sites are shown in solid black boxes, cysteine residues which may form disulphide bridges are shown in dashed black boxes. Putative TMs are shown in solid black boxes above the sequence and were determined by aligning AT₁R with GPCRs of known structure (Appendix Figure 1). Conserved residue properties are grouped by colour, e.g. blue represents a hydrophobic residue.

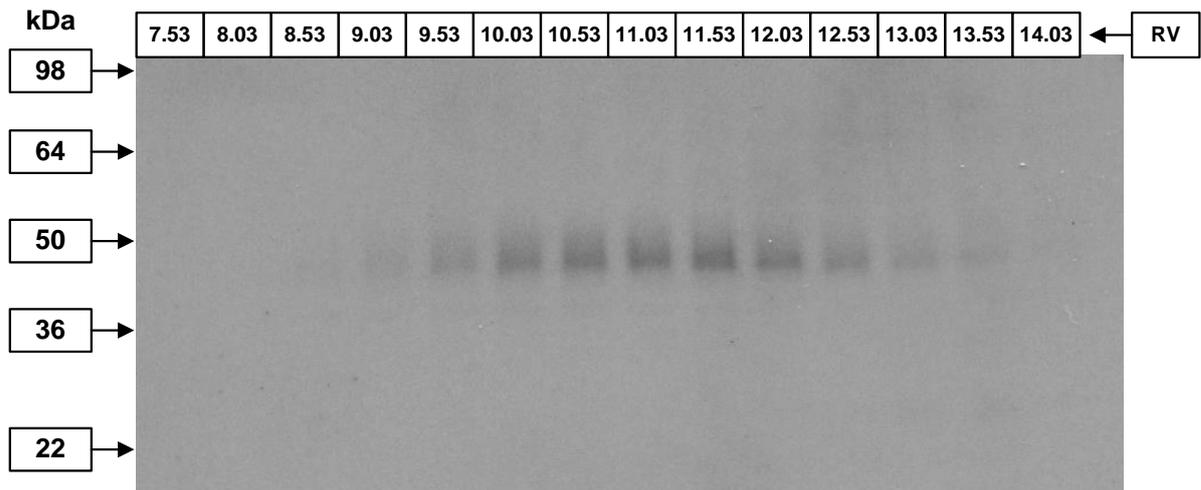
APPENDIX 2 INSECT CELL EXPRESSION TEST DATA



Appendix Figure 5 Optimisation of bvAT₁R-H₁₀ and bvAT₁R-LS-H₁₀ expressed in Sf21 or Hi5 cells

Western blot of proteins from whole cells. The blot was probed with an anti-pentaHis-HRP conjugated antibody. **(A)** Sf21 cells infected with either bvAT₁R-H₁₀ or bvAT₁R-LS-H₁₀ as indicated. The dashed line indicates two separate blots. **(B)** Hi5 cells infected with bvAT₁R-H₁₀. **(C)** Hi5 cells infected with bvAT₁R-LS-H₁₀. N-linked glycosylation was removed using Endo H where indicated (+). The high manose form of AT₁R (HM), unglycosylated (UG) and putative degradation products (D) are indicated. Bands corresponding to HM and UG forms of AT₁R were quantified using Image J (Figure 2.8).

APPENDIX 1 Alignments



Appendix Figure 6 Size exclusion chromatography of AT₁R expressed in insect cells

Western blot of detergent-solubilised AT₁R-H₁₀ from Sf9 cells analysed by SEC. The blot was probed with an anti-pentaHis-HRP conjugated antibody. The retention volume (RV) of the sample is indicated in ml. Bands corresponding to AT₁R were quantified using Image J (Figure 2.16).

APPENDIX 1 Alignments

APPENDIX 3 FLUORESCENCE ACTIVATED CELL SORTING DATA

Appendix Table 1 FACS analysis of monoclonal cell lines from the polyclonal cell line iHEK(AT₁R-GFP-H₁₀)

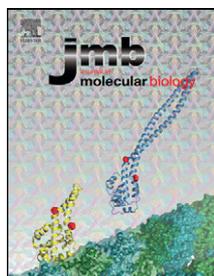
FACS analysis of iHEK cells. Cells were harvested in PBS and analysed on the FACSCalibur II for GFP fluorescence using the FL-1 detector. Monoclonal cell lines derived from the population with moderate basal expression were named iHEK3XX. Monoclonal cell lines derived from the population with high basal expression were named iHEK4XX. Green indicates cell lines that are significantly better than the polyclonal cell line iHEK(AT₁R-GFP-H₁₀) for the respective population from which it was derived (*i.e.* high or moderate basal expression).

Cell line	Induced			Uninduced		Cells expressing GFP (%)
	FACS Fluorescence intensity		Cells expressing GFP (%)	FACS Fluorescence intensity		
	Mean	Median		Mean	Median	
iHEK301	376	205	96.1	14	8	15.4
iHEK302	446	237	97.4	19	8	17.8
iHEK303	440	235	96.9	23	10	22.2
iHEK304	422	284	97.3	16	8	16.6
iHEK305	458	271	97.6	15	9	17.6
iHEK306	445	237	96.4	15	8	16.6
iHEK307	428	229	97.4	15	8	16.0
iHEK308	495	227	97.5	18	9	20.4
iHEK309	523	305	98.0	17	9	16.5
iHEK310	392	176	96.4	13	6	12.6
iHEK311	410	204	96.6	17	8	17.7
iHEK312	423	221	97.5	22	9	21.7
iHEK313	407	237	97.3	17	9	17.4
iHEK314	436	213	97.5	16	8	16.5
iHEK315	415	213	97.5	20	8	19.0
iHEK316	418	211	97.2	14	8	15.2
iHEK317	401	184	96.8	22	8	21.6
iHEK318	439	221	97.3	17	9	19.7
iHEK319	619	365	97.7	14	9	14.4
iHEK320	360	184	95.3	18	8	19.4
iHEK321	414	223	97.5	17	9	21.2
iHEK322	414	229	96.5	17	9	20.8
iHEK323	375	213	96.4	15	8	16.4
iHEK324	420	213	96.1	18	10	23.2
<hr/>						
iHEK401	334	235	93.7	14	9	14.2
iHEK402	581	368	98.5	17	10	19.4
iHEK403	457	264	98.0	17	10	19.2
iHEK404	345	204	96.9	14	8	15.2
iHEK405	487	316	97.9	14	9	14.2
iHEK406	504	331	98.2	14	9	13.6
iHEK407	591	334	98.4	18	10	20.4
iHEK408	409	259	98.0	13	8	12.5
iHEK409	452	297	97.8	22	10	20.6
iHEK410	373	255	97.5	18	9	16.9
iHEK411	365	204	97.6	13	8	13.3

APPENDIX 1 Alignments

Cell line	Induced			Uninduced		
	FACS Fluorescence intensity		Cells expressing GFP (%)	FACS Fluorescence intensity		Cells expressing GFP (%)
	Mean	Median		Mean	Median	
iHEK412	336	186	97.0	16	8	15.4
iHEK413	358	191	97.5	15	9	15.9
iHEK414	362	213	97.6	17	8	15.9
iHEK415	748	429	98.8	20	12	26.3
iHEK416	520	308	98.0	13	8	12.7
iHEK417	352	241	96.3	12	7	10.1
iHEK418	354	233	97.8	12	7	10.6
iHEK419	507	289	98.4	26	10	20.2
iHEK420	671	437	98.8	21	14	28.1
iHEK421	363	209	97.3	14	8	15.0
iHEK422	400	235	97.9	17	9	18.3
iHEK423	486	302	97.0	—	—	—
iHEK424	390	189	96.3	15	8	15.2
iHEK425	625	379	98.7	26	9	19.3
iHEK426	527	359	98.2	18	10	18.1
iHEK427	381	219	97.0	18	12	24.9
iHEK428	430	271	96.9	16	9	17.1
iHEK429	489	325	98.3	12	8	10.4
iHEK430	377	213	96.5	14	8	14.0
iHEK431	528	239	97.2	21	10	21.6
iHEK432	445	284	97.7	12	8	11.4
iHEK433	311	161	93.1	11	6	10.6
iHEK434	584	302	97.8	25	9	24.2
iHEK435	497	305	97.7	15	10	19.4
iHEK436	423	246	97.1	14	9	17.4
iHEK437	477	267	97.9	16	9	18.7
iHEK438	409	207	96.5	14	8	15.5
iHEK439	449	279	96.9	22	11	23.8
iHEK440	481	308	97.4	15	9	16.7
iHEK441	535	313	96.9	25	11	26.0
iHEK442	454	120	73.0	23	10	24.3
iHEK443	510	337	97.3	16	10	17.0
iHEK444	367	202	96.8	15	8	16.7
iHEK445	500	328	97.5	14	10	16.5
iHEK446	452	284	95.4	14	10	17.4
iHEK447	513	328	96.9	28	12	26.8
iHEK448	438	217	97.4	23	9	23.8
iHEK(AT ₁ R-GFP-H ₁₀)	310	276	91.7	10	8	5.7
iHEK parental	—	—	—	8	6	5.6

APPENDIX 4 SUBMITTED JOURNAL ARTICLE



Quality Control in Eukaryotic Membrane Protein Overproduction

Jennifer A. Thomas and Christopher G. Tate

MRC Laboratory of Molecular Biology, Cambridge Biomedical Campus, Francis Crick Avenue, Cambridge CB2 0QH, UK

Correspondence to Christopher G. Tate: cgt@mrc-lmb.cam.ac.uk.

<http://dx.doi.org/10.1016/j.jmb.2014.10.012>

Edited by B. Poolman

Abstract

The overexpression of authentically folded eukaryotic membrane proteins in milligramme quantities is a fundamental prerequisite for structural studies. One of the most commonly used expression systems for the production of mammalian membrane proteins is the baculovirus expression system in insect cells. However, a detailed analysis by radioligand binding and comparative Western blotting of G protein-coupled receptors and a transporter produced in insect cells showed that a considerable proportion of the expressed protein was misfolded and incapable of ligand binding. In contrast, production of the same membrane proteins in stable inducible mammalian cell lines suggested that the majority was folded correctly. It was noted that detergent solubilisation of the misfolded membrane proteins using either digitonin or dodecylmaltooside was considerably less efficient than using sodium dodecyl sulfate or foscholine-12, whilst these detergents were equally efficient at solubilising correctly folded membrane proteins. This provides a simple and rapid test to suggest whether heterologously expressed mammalian membrane proteins are indeed correctly folded, without requiring radioligand binding assays. This will greatly facilitate the high-throughput production of fully functional membrane proteins for structural studies.

© 2014 MRC Laboratory of Molecular Biology. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/3.0/>).

Introduction

Structure determination of integral membrane proteins requires the production of milligrammes of pure, authentically folded protein for crystallisation [1]. As a natural prerequisite, the protein needs to be expressed in one of a number of heterologous expression systems, such as *Escherichia coli*, yeasts, insect cells or mammalian cells [2]. A number of expression strategies have been developed for each host system and many are now efficient for expression trials of hundreds of proteins in parallel [3]. A popular strategy for the expression of membrane proteins in *E. coli* is to generate fusion proteins with green fluorescent protein (GFP), which can be used as an indicator for both the quantity of protein expressed [4] and its relative stability upon detergent solubilisation by fluorescence-detection size-exclusion chromatography (FSEC) [5]. The utility of this strategy is that fluorescence of the fusion protein expressed in bacteria discriminates

between correctly folded membrane protein (the GFP tag is fluorescent) and misfolded, aggregated membrane protein (the GFP tag is not fluorescent) [6,7]. In *E. coli*, it appears that the misfolded membrane protein promotes the formation of inclusion bodies and, once in an aggregate, the GFP is unable to fold and attain fluorescence. However, in eukaryotic cells, such as yeasts, insect cells used in the baculovirus expression system and in mammalian cells, GFP tagged to a membrane protein remains fluorescent regardless of whether the membrane protein is misfolded in the endoplasmic reticulum or correctly folded in the plasma membrane [8–11]. Higher eukaryotes have an efficient quality control system in the endoplasmic reticulum so that only folded proteins exit the endoplasmic reticulum, whilst misfolded proteins are retained for degradation [12]. Thus, GFP is not an appropriate marker for the folding status of membrane proteins produced using either mammalian cells or the baculovirus expression system, although it is still

useful in analysing the stability of a membrane protein in different detergents by FSEC.

The baculovirus expression system has proven efficient for the production of many eukaryotic membrane proteins, such as G protein-coupled receptors (GPCRs) [13], some of which have been crystallised and their structures determined [14]. However, recombinant baculovirus is not a panacea and there are many proteins that are poorly expressed and there have also been reports that some membrane proteins are expressed predominantly in a misfolded state [2,15]. This is a serious problem for structural biology, as it is not obvious from current methodology whether an overexpressed membrane protein is predominantly folded or misfolded. If misfolded material is inadvertently purified, then this will likely have a detrimental effect on the ability of the sample to crystallise and may also adversely affect the quality of any crystals obtained. The best way to determine whether misfolded material is present is to perform quantitative Western blotting to assess the total amount of membrane protein expressed in conjunction with radioligand binding assays to determine how much is functional [16,17]. However, this is expensive, difficult to perform and is also impossible for the majority of membrane proteins that do not possess high-affinity radioligands. It is also unclear whether the presence of misfolded overexpressed membrane protein is a rare event or whether it is commonly observed. We have therefore studied a number of membrane proteins produced both in stable mammalian cell lines and using the baculovirus expression system. The data show that all the four membrane proteins analysed are expressed in the baculovirus system as a mixture of folded and misfolded proteins, whereas mammalian cell lines are much more efficient at producing only correctly folded membrane proteins. A simple comparative detergent solubilisation assay is described, which is an excellent indicator for the presence of misfolded membrane proteins.

Results

Comparative expression of the angiotensin II type 1 receptor in insect cells and stable mammalian cell lines

The human angiotensin II type 1 receptor (AT₁R) is a GPCR with the typical predicted architecture of seven transmembrane regions with the N-terminus on the extracellular surface of the cell. The receptor contains three N-linked glycosylation sites with one in the N terminal region (Asn4) and two in extracellular loop 2 (Asn176 and Asn188). Two expression systems were used for the production of AT₁R, the

baculovirus expression system and stable tetracycline-inducible mammalian cell lines (the T-Rex system). AT₁R was expressed with a C-terminal decahistidine tag (H₁₀) from the polyhedrin promoter in the recombinant baculovirus bvAT₁R-H₁₀. In inducible mammalian HEK293 cells (iHEK), AT₁R was expressed with a C-terminal GFP-H₁₀ tag from the CMV promoter after induction with tetracycline; the stable cell line iHEK(AT₁R-GFP-H₁₀) was generated through random integration of the plasmid in the genome followed by fluorescence-activated cell sorting to isolate a high-expressing clonal cell line. Initial analysis of expression was performed by Western blotting using an anti-penta-His horseradish peroxidase (HRP) conjugated antibody for detection (Fig. 1). AT₁R was extensively N-glycosylated in mammalian cells, which could be removed by treatment with PNGase F to yield a major product AT₁R-GFP-H₁₀ of apparent molecular mass of 60 kDa; no unglycosylated AT₁R-GFP-H₁₀ was visible in untreated cells. In Sf9 cells, AT₁R-H₁₀ was expressed as a mixture of glycosylated and unglycosylated receptor, which yielded a single major product (apparent molecular mass of 36 kDa) after treatment with PNGase F. The blot in Fig. 1 contained the same number of cells per lane; thus, an assessment of band intensities by eye suggested that there were similar levels of AT₁R expressed from the baculovirus expression system and from the stable mammalian cell line. However, despite the apparently similar levels of AT₁R polypeptide expressed in Sf9 and iHEK cells, radioligand binding assays showed that there was 20× more functional AT₁R expressed per cell in mammalian cells compared to the best baculoviral expression observed (Fig. 1). This implied that a large proportion of AT₁R expressed in insect cells was misfolded and incapable of binding antagonist.

Detergent solubilisation is the first step in the purification of a membrane protein; thus, the ability of AT₁R expressed stably in the iHEK cell line iHEK(AT₁R-GFP-H₁₀) to be solubilised by four different detergents was tested. The four detergents used in order of decreasing “harshness” [18] were sodium dodecyl sulfate (SDS), foscholine-12 (FC12), dodecylmaltoside (DDM) and digitonin. Digitonin was the mildest detergent used and it is very effective in maintaining membrane proteins in a functional state. DDM is one of the most popular mild detergents used for membrane protein purification, but it is a little harsher than digitonin. Only very few bacterial membrane proteins are sufficiently stable to maintain their integrity in either FC12 or SDS; thus, no ligand binding would be expected for AT₁R solubilised in either SDS or FC12. All four detergents were equally effective at solubilising AT₁R polypeptide expressed in iHEK cells (Fig. 2). However, as expected from the differing “harshness” of the detergents, only DDM and digitonin maintained the

integrity of ^{125}I -Sar¹-bound AT₁R so that receptor-bound radioligand could be detected (Fig. 2). In contrast, bound radioligand was not detected when SDS was used to solubilise ^{125}I -Sar¹-bound AT₁R, and only a small amount was detected when FC12 was used. Assays on DDM-solubilised AT₁R measured nearly twice as much receptor as detected in membranes (Fig. 2), which could be due to freeze-thawed membranes being a mixture of both right-side-out vesicles and inside-out vesicles, and the membrane-impermeant peptide ^{125}I -Sar¹ could bind only to AT₁R in the rightside-out vesicles.

Detergent solubilisation of ^{125}I -Sar¹-bound AT₁R from Sf9 cell membranes after expression from the recombinant baculovirus bvAT₁R-H₁₀ followed a similar pattern to that observed from the stable mammalian iHEK cell line; that is, double the amount of radioligand was observed in DDM-solubilised receptor compared to membranes and no binding was detected when SDS was used. Note that the binding data in Fig. 2 are normalised for ease of comparison, whereas in actuality, there is 20-fold less functional AT₁R per cell in Sf9 cells compared to the stable mammalian cell line. However, the Western blotting data of AT₁R produced in Sf9

cells are different from the analogous data from the iHEK cell line. Orders of magnitude more AT₁R polypeptide was solubilised from Sf9 cell membranes by SDS or FC12 compared to either DDM or digitonin (Fig. 2). It is reasonable to assume from the ^{125}I -Sar¹ binding data that DDM solubilised all the functional AT₁R and therefore the difference between the signal on the Western blot for DDM-solubilised AT₁R and SDS-solubilised AT₁R represents misfolded AT₁R.

There is a significant difference between the Western blotting signal for SDS-solubilised AT₁R and DDM-solubilised AT₁R from Sf9 cells and that difference represents an amount of misfolded AT₁R that can be solubilised by SDS but not by DDM. However, there may be more misfolded AT₁R present in Sf9 cell membranes than suggested from the differential solubility in SDS *versus* DDM because it is plausible that DDM can also solubilise some AT₁R that cannot bind antagonist. To assess this possibility, we diluted membranes from the stable mammalian cell line iHEK(AT₁R-GFP-H₁₀) and insect cell membranes expressing AT₁R-H₁₀ to give the same amount of functional AT₁R per millilitre, solubilised in DDM and then analysed by Western blotting. The data (Fig. 3) showed clearly that there was considerably more AT₁R polypeptide solubilised from Sf9 cells than from the mammalian cell line and that this difference is due to misfolded receptor given that there was the same amount of functional receptor per lane. Efforts to decrease the amount of misfolded AT₁R in the baculovirus

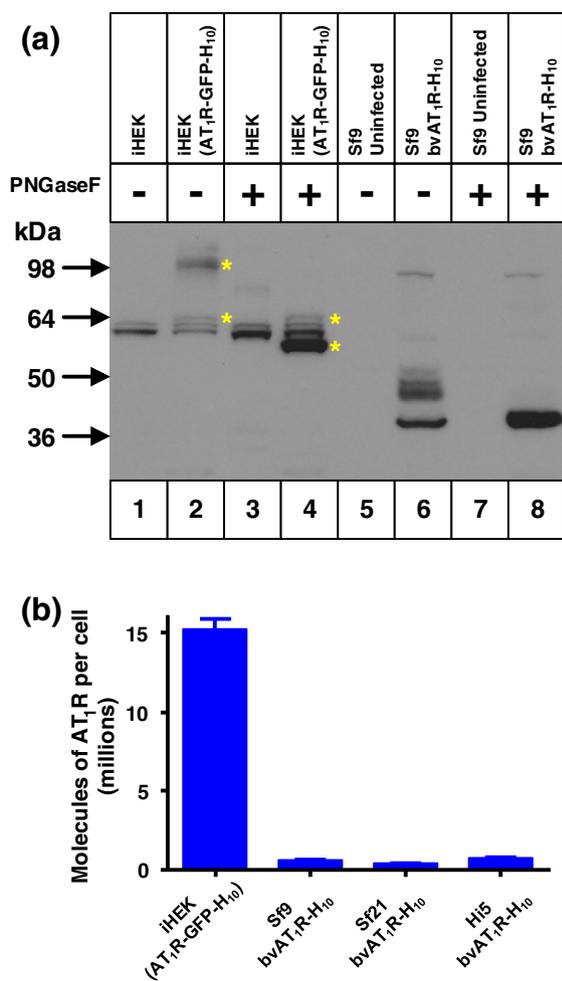


Fig. 1. Functional expression of AT₁R in mammalian cells that is 5-fold higher compared to insect cells. (a) Western blot of whole cells expressing AT₁R solubilised in SDS. Lanes 1 and 3, iHEK parental cells; lanes 2 and 4, iHEK(AT₁R-GFP-H₁₀) stable clonal cell line; lanes 5 and 7, uninfected Sf9 cells; lanes 6 and 8, bvAT₁R-H₁₀ infected Sf9 cells. N-Linked glycosylation was removed using PNGase F where indicated (+). Bands corresponding to AT₁R-GFP-H₁₀ in mammalian cells are indicated with a yellow asterisk (*). iHEK cell lines were induced with 1 $\mu\text{g}/\text{ml}$ tetracycline for 24 h and insect cells were infected with recombinant baculovirus for 48 h. The blot was probed with an anti-pentaHis-HRP conjugated antibody. (b) The amount of functional AT₁R in each expression system was determined by measuring specific binding of the antagonist [^{125}I]Sar¹. Baculoviral expression was performed in Sf9, Sf21 or Hi5 cells. After the addition of ligand, membranes were solubilised in DDM and non-bound ligand was separated from receptor–ligand complex on gel-filtration spin columns and measured by liquid scintillation counting. [^{125}I]Sar¹-bound AT₁R is stable in DDM, but not in SDS. The amount of functional AT₁R was most accurately determined after solubilisation with DDM to ensure that all the receptor was accessible to ligand (see Fig. 2, where twice as much receptor could be measured upon solubilisation in DDM compared to in membranes). Each data point was determined in triplicate from two independent experiments and was plotted as mean \pm SEM (standard error of the mean).

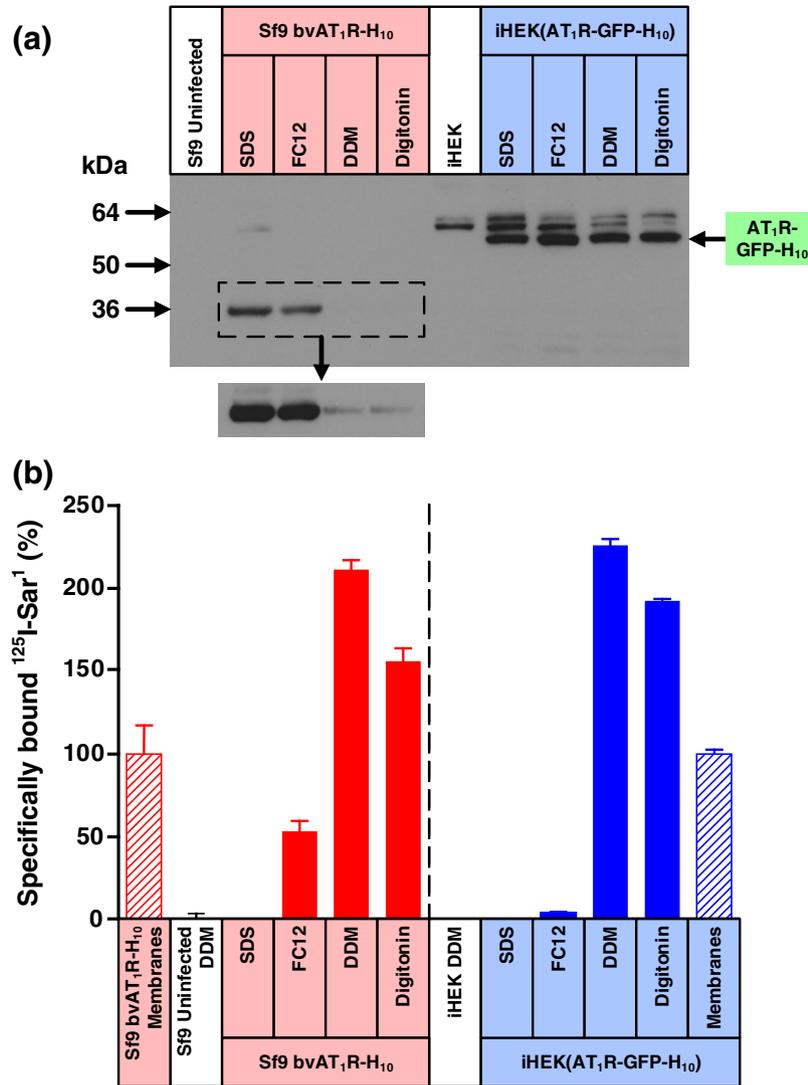


Fig. 2. Misfolded AT₁R produced by the baculovirus expression system is poorly solubilised either by DDM or digitonin. (a) Western blot of AT₁R solubilised from whole cells using four different detergents (SDS, FC12, DDM or digitonin) and probed with an anti-pentaHis-HRP conjugated antibody. Each lane contains an equal amount of total protein and N-linked glycosylation was removed from all samples using PNGase F prior to SDS-PAGE. AT₁R was expressed either in the stable mammalian cell line iHEK(AT₁R-GFP-H₁₀) or by using the recombinant baculovirus bvAT₁R-H₁₀ to infect Sf9 cells. The iHEK cell line was induced with 1 µg/ml tetracycline for 24 h and Sf9 cells were infected for 48 h. The Western blot insert is a 7× longer exposure. (b) The amount of functional detergent-solubilised AT₁R was determined by measuring specific binding of the antagonist [¹²⁵I]Sar¹. After the addition of ligand, membranes were solubilised in the detergent indicated and non-bound ligand was separated from receptor–ligand complex on gel-filtration spin columns and measured by liquid scintillation counting: red-filled bars, AT₁R expressed in Sf9 cells; blue-filled bars, AT₁R expressed in iHEK cells. The amount of AT₁R in membranes (non-solubilised) was determined by separation of receptor-bound and free radioligand by filtration through glass fibre plates: red hatched bars, AT₁R expressed in Sf9 cells; blue hatched bars, AT₁R expressed in iHEK cells. For ease of comparison, binding data have been normalised with respect to AT₁R in membranes (100%), which is equivalent to 1400 ± 240 dpm (*n* = 2; 380 fmol per million cells) for baculovirus-infected Sf9 cells and 12,000 ± 300 dpm (*n* = 2; 8.8 pmol per million cells) for iHEK(AT₁R-GFP-H₁₀) cells. Absolute levels of AT₁R therefore cannot be compared meaningfully between the two expression systems using this bar graph (see Fig. 1). Binding assays for AT₁R contained either 150,000 Sf9 cells or 55,000 iHEK cells. Each data point was determined in duplicate or triplicate from a single experiment and was plotted as mean ± SEM.

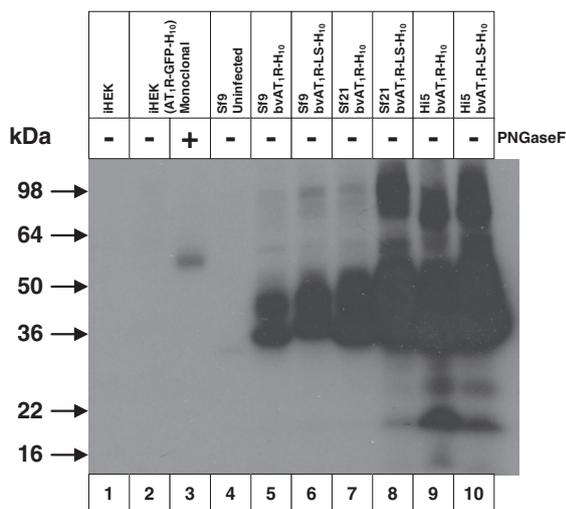


Fig. 3. DDM solubilises considerable amounts of inactive AT₁R produced in the baculovirus expression system. (a) Western blot of DDM-solubilised AT₁R, with equal amounts of active receptor per sample (lanes 2, 3 and 5–10). The blot was probed with an anti-penta-His-HRP conjugated antibody. Lane 1, iHEK parental cells; lanes 2 and 3, iHEK(AT₁R-GFP-H₁₀) stable clonal cell line; lane 4, uninfected Sf9 cells; lanes 5–10, bvAT₁R-H₁₀ infected insect cells. N-Linked glycosylation was removed using PNGase F where indicated (+). AT₁R was expressed either in the stable mammalian cell line iHEK(AT₁R-GFP-H₁₀) or by using the recombinant baculoviruses bvAT₁R-H₁₀ and bvAT₁R-LS-H₁₀ to infect Sf9, Sf21 and Hi5 cells as indicated. iHEK cell lines were induced with 1 µg/ml tetracycline for 24 h and insect cells were infected with recombinant baculovirus for 48 h. The amount of functional AT₁R was determined by measuring specific binding of the antagonist [¹²⁵I]Sar¹.

expression system either by using different cell lines (Sf21, Hi5) or by including an N-terminal signal sequence on AT₁R were ineffective (Fig. 3).

In order to ascertain the quality of AT₁R expressed in either mammalian cells or insect cells, we analysed two biophysical parameters of the detergent-solubilised receptor. Firstly, the thermostability of DDM-solubilised AT₁R was determined and the apparent T_m values of [¹²⁵I]-Sar¹-bound AT₁R expressed in either Sf9 cells or mammalian cells were found to be identical (Sf9 cells, 46 ± 0.8 °C; iHEK(AT₁R-GFP-H₁₀), 46 ± 0.7 °C). Secondly, the mobility on size-exclusion chromatography (SEC) of AT₁R expressed using the baculovirus expression system in Sf9 cells or from the stable mammalian cell line was compared and also found to be very similar (Fig. 4). These data, coupled with the similarity in pharmacology between AT₁R expressed in the two cell types [19], suggest that there is no significant difference between correctly folded AT₁R produced in the baculovirus expression system and AT₁R produced in the stable mammalian cell line.

The presence of misfolded protein upon overexpression from recombinant baculovirus is not uncommon

The presence of substantial amounts of misfolded AT₁R upon production in the baculovirus expression system raised the question of whether this is specific for AT₁R or whether other membrane proteins also exhibited this property. As it is not possible to test rigorously all membrane proteins, a careful selection was made of interesting test cases. The avian β₁-adrenergic receptor (β₁AR) is a well-characterised GPCR and its structure has been determined bound to many different ligands of different efficacy [20–23]. All of the β₁AR crystals were grown from protein expressed in either Sf9 or Hi5 cells using recombinant

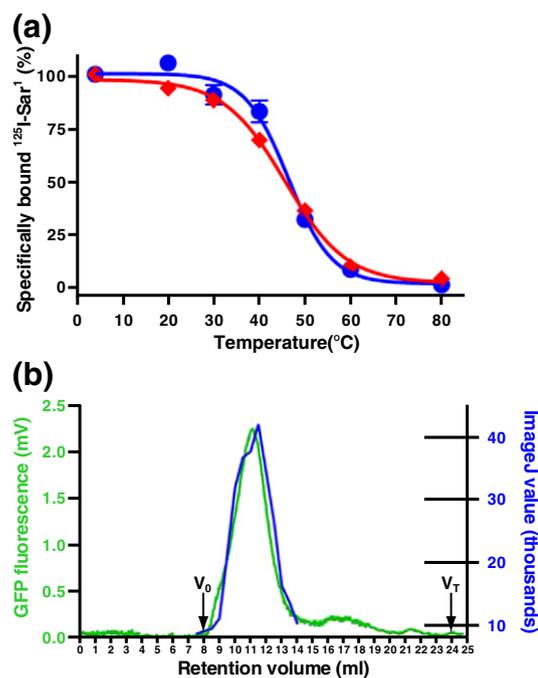


Fig. 4. Comparison of AT₁R expressed in mammalian cells and insect cells. (a) Stability of DDM-solubilised AT₁R bound to the antagonist [¹²⁵I]Sar¹. AT₁R was expressed using three different expression systems: blue circles, baculovirus bvAT₁R in Sf9 cells; red diamonds, stable clonal cell line iHEK(AT₁R-GFP-H₁₀). The apparent T_m values of AT₁R expressed in each system are as follows: Sf9 cells, 46 ± 0.8 °C; iHEK(AT₁R-GFP-H₁₀), 46 ± 0.7 °C. Each data point was determined in triplicate and was plotted as a mean value ± SEM. (b) SEC was carried out using a Superdex 200 10/300 (24 ml) column. The elution of iHEK(AT₁R-GFP-H₁₀) was detected using GFP fluorescence (mV). The elution of bvAT₁R-H₁₀ was detected by Western blotting and band quantification (ImageJ value). iHEK(AT₁R-GFP-H₁₀) shows a symmetrical peak whereas bvAT₁R-H₁₀ shows two peaks; however, both systems show elution of a protein of a similar size. The void (V_0) and total (V_T) column volumes are indicated.

baculoviruses [24,25]. The assays described for AT₁R were therefore repeated using wild-type β_1 AR with truncations at the N-terminus and at the C-terminus (bv β_1 AR-H₁₀), which facilitates expression of a homogeneous protein. Comparison of the amount of β_1 AR-H₁₀ solubilised by the different detergents clearly indicates that a large proportion of the receptor is indeed misfolded, as suggested by the higher proportion of receptor solubilised by either SDS or FC12 compared

to either DDM or digitonin (Fig. 5). Efforts to improve the proportion of folded protein by using a thermostable β_1 AR mutant fused to the N-terminus to a leader sequence and a well-folded soluble protein (thioredoxin) did not increase the proportion of correctly folded β_1 AR (Fig. 5). However, it is interesting to note that receptor containing an uncleaved leader sequence was only extracted by SDS or FC12, suggesting that this sub-population of receptor was probably mainly

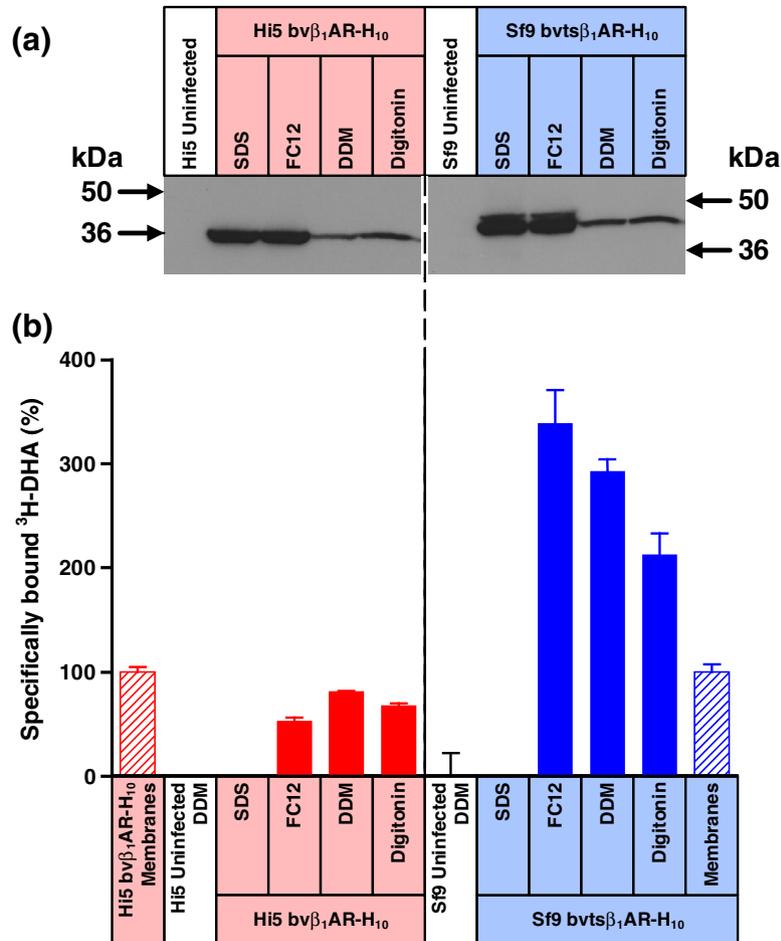


Fig. 5. Misfolded β_1 AR is poorly solubilised either by DDM or digitonin. (a) Western blot of β_1 AR solubilised from whole cells using four different detergents (SDS, FC12, DDM or digitonin) and probed with an anti-pentaHis-HRP conjugated antibody. Each lane contains an equal amount of total protein. β_1 AR was expressed by using the recombinant baculovirus bv β_1 AR-H₁₀ to infect Hi5 cells. ts β_1 AR was expressed by using the recombinant baculovirus bvts β_1 AR-H₁₀ to infect Sf9 cells. Hi5 and Sf9 cells were infected for 48 h. The broken line indicates two separate blots. (b) The amount of functional detergent-solubilised β_1 AR and ts β_1 AR was determined by measuring specific binding of the antagonist [³H]DHA. After the addition of ligand, membranes were solubilised in the detergent indicated and non-bound ligand was separated from receptor–ligand complex on gel-filtration spin columns and measured by liquid scintillation counting: red-filled bars, β_1 AR expressed in Hi5 cells; blue-filled bars, ts β_1 AR expressed in Sf9 cells. The amount of β_1 AR in membranes (non-solubilised) was determined by separation of receptor-bound and free radioligand by filtration through glass fibre plates: red hatched bars, β_1 AR expressed in Hi5 cells; blue hatched bars, ts β_1 AR expressed in Sf9 cells. For ease of comparison, binding data have been normalised with respect to β_1 AR in membranes (100%), which is equivalent to $11,000 \pm 550$ dpm ($n = 3$; 6.1 pmol per million cells) for baculovirus-infected Hi5 cells and 2600 ± 190 ($n = 3$; 1.4 pmol/l) for bvts β_1 AR-H₁₀ infected Sf9 cells. All binding assays for β_1 AR and ts β_1 AR contained 8300 cells, and therefore, comparison on absolute levels of receptor can be directly compared. Each data point was determined in duplicate or triplicate from a single experiment and was plotted as mean \pm SEM.

misfolded. In addition, it is unlikely that the fusion protein was efficiently trafficked to the cell surface given that FC12 extraction resulted in a 3-fold increase in the amount of receptor binding obtained compared to when membranes were used.

In a second example, we compared the expression of the adenosine A₁ receptor (A₁R) in both the stable mammalian cell line iGnTI⁻(A₁R-GFP-H₁₀)

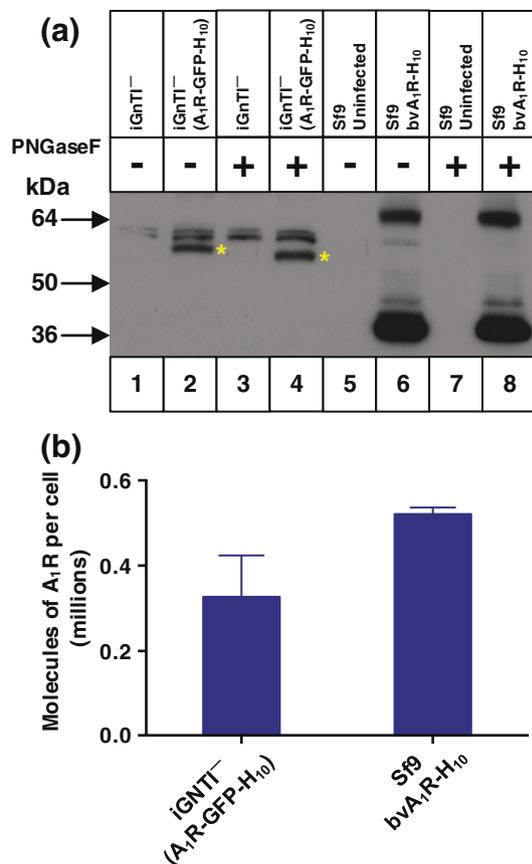


Fig. 6. Expression of A₁R in mammalian cells compared to insect cells. (a) Western blot of whole cells expressing A₁R solubilised in SDS. Lanes 1 and 3, iGnTI⁻ parental cells; lanes 2 and 4, iGnTI⁻(A₁R-GFP-H₁₀) stable cell line; lanes 5 and 7, uninfected Sf9 cells; lanes 6 and 8, bvA₁R-H₁₀ infected Sf9 cells. N-Linked glycosylation was removed using PNGase F where indicated (+). Bands corresponding to A₁R-GFP-H₁₀ in mammalian cells are indicated with a red asterisk (*). The iGnTI⁻ cell line was induced with 1 µg/ml tetracycline for 24 h and insect cells were infected with recombinant baculovirus for 72 h. The blot was probed with an anti-pentaHis-HRP conjugated antibody. (b) The amount of functional A₁R in each expression system was determined by measuring specific binding of the antagonist [³H]DPCPX. After the addition of ligand, membranes were solubilised in DDM and non-bound ligand was separated from receptor–ligand complex on gel-filtration spin columns and measured by liquid scintillation counting. Each data point was determined in duplicate and was plotted as mean ± SEM.

and the insect cells using the baculovirus expression system (bvA₁R-H₁₀). N-Linked glycosylated sites are found in extracellular regions of A₁R (extracellular loop 2; Asn148 and Asn159), which produces a glycosylated form in mammalian cells that can be reduced in molecular weight by treatment with PNGase F, whereas the majority of the receptor is unglycosylated in Sf9 cells (Fig. 6). Expression of A₁R-H₁₀ in insect cells gave comparative Western blots analogous to those observed for AT₁R-H₁₀, with SDS and FC12 extracting orders of magnitude more polypeptide from insect cell membranes compared to either DDM or digitonin, consistent with a large excess of misfolded receptor produced in the baculovirus expression system (Fig. 7). In contrast, all the detergents used to solubilise A₁R-GFP-H₁₀ from a stable mammalian cell line were equally efficacious, indicating that there is minimal misfolded receptor in these cells (Fig. 7). The low levels of antagonist binding activity observed for A₁R is a consequence of the poor stability of this receptor in detergent solutions. A₁R also provided a nice example of the usefulness of confocal microscopy in defining whether a receptor is likely to be correctly folded. A₁R-GFP-H₁₀ is expressed in the stable cell line predominantly at the cell surface whereas a mutant of A₁R that contained multiple changes introduced to try and facilitate crystallisation (A₁R-GL26-GFP-H₁₀; see [Methods](#)) was expressed predominantly in intracellular membranes (Fig. 8). The confocal data correlated well with the Western blotting data. The misfolded mutant A₁R-GL26-GFP-H₁₀ was only efficiently extracted from mammalian cells with SDS (Fig. 9), whereas the wild-type receptor was extracted equally efficiently using either digitonin or SDS (Fig. 7).

The final example we tested was the serotonin transporter (SERT). The expression of SERT has been studied intensively [16,17,26] and was the first example along with rhodopsin that showed the utility of mammalian cells for the overexpression of functional membrane protein using the tetracycline-inducible HEK293 cell line [27,28]. Here we demonstrated that the Western blot data show an identical pattern of results with constructs based on wild-type A₁R and AT₁R, namely, similar amounts of extractable SERT-SAH9-GFP-H₁₀ from the mammalian cell line, regardless of the detergent used, whereas there are orders of magnitude more SDS-extractable SERT-H₁₀ in Sf9 cells compared to the amount solubilised by digitonin or DDM (Fig. 10).

A simplified assay for the detection of misfolded membrane proteins

Analysis of the data in [Figs. 1–10](#) suggests that the salient conclusions of this paper, that is, that the majority of AT₁R, A₁R and SERT constructs

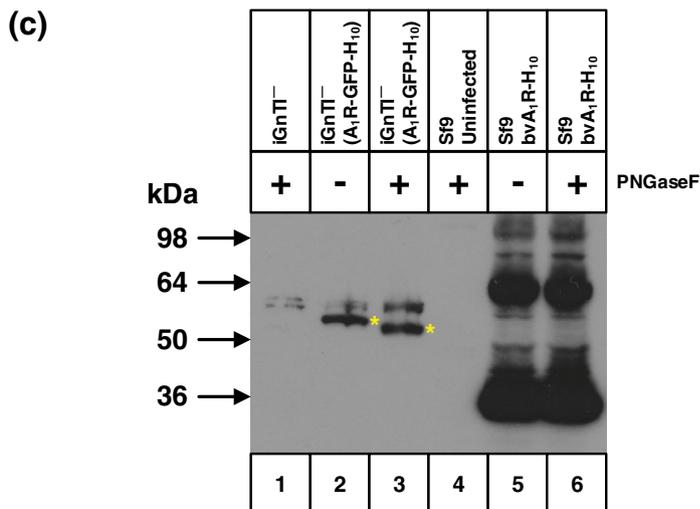
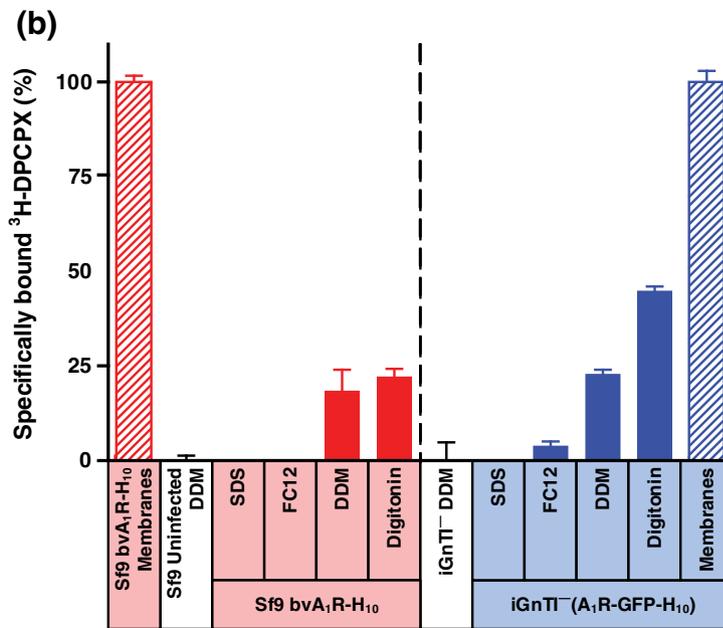
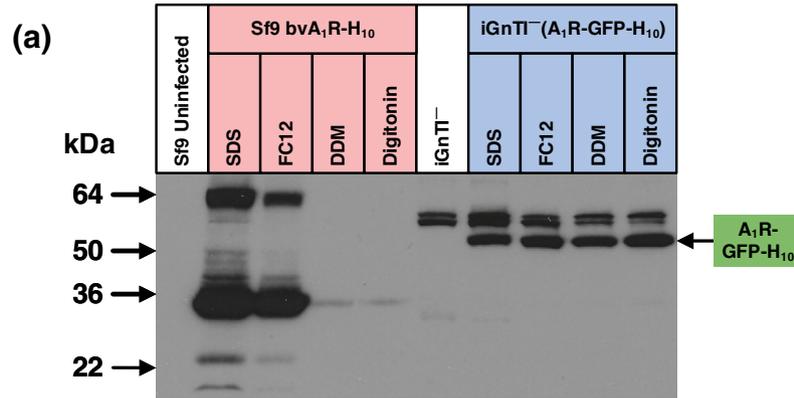


Fig. 7 (legend on next page)

expressed in insect cells were misfolded, whereas expression in mammalian cells produced correctly folded protein, could be deduced with a fraction of the work. Comparison of two lanes in each Western blot, namely, SDS-extracted protein and digitonin-extracted protein, is sufficient to draw the relevant conclusions. Importantly, this obviates the need for radioligand binding assays and a stable mammalian cell line for each membrane protein to be studied. Radioligands have been developed for only a small fraction of membrane proteins and not all radioligands are of sufficiently high affinity (100 nM or better) to make them suitable for assays on detergent-solubilised membrane proteins. In addition, construction of stable mammalian cell lines can take many months and sometimes the cell lines grow very poorly due to basal activity of the membrane protein. This was noticeable for the stable A₁R cell line iGnTI⁻(A₁R-GFP-H₁₀) developed here, which grew very poorly compared to the stable cell line iGnTI⁻(A₁R-GL26-GFP-H₁₀) expressing the inactive A₁R mutant, despite the use of an inducible promoter.

Using the methodology described in this paper, it would be relatively simple to test 50 or so different membrane protein expression trials in a day. However, if hundreds of samples are to be tested in 96-well plates, then the ultracentrifugation step will become limiting and will need to be replaced using filtration through low-protein-binding 0.2- μ m filters. The use of a dot-blot apparatus and semi-quantification of the resulting signals in relation to a known standard would be sufficient to define how much functional membrane protein could be extracted and whether or not extra

precautions may be required to remove potential misfolded protein during purification. It would also be possible to measure the fluorescence of a GFP-tagged membrane protein, rather than performing a Western blot, to improve further the high-throughput capabilities of this assay.

Discussion

A commonly held misconception, particularly amongst investigators new to the membrane protein field, is that if a membrane protein can be expressed into a membrane within a cell and can be extracted with detergent, then that membrane protein is folded authentically. Over the last 30 years, there have been sporadic reports of overexpressed membrane proteins in *E. coli*, yeast or the baculovirus expression system being predominantly misfolded and inactive [2]. The work presented here demonstrates that the baculovirus expression system is particularly prone to producing misfolded membrane proteins, even of apparently uncomplex GPCRs that were expressed over 20 years ago. However, the simple assay proposed here will rapidly demonstrate whether misfolded membrane protein is indeed present. A few words of caution are warranted with regard to the differential solubility assay. Firstly, we have tested the assay on membrane proteins expected to be expressed in the plasma membrane of mammalian cells, which is efficiently solubilised by DDM. This is evident from the similar levels of solubilisation between DDM and SDS of correctly folded membrane proteins in the plasma membrane.

Fig. 7. Misfolded A₁R produced by the baculovirus expression system is poorly solubilised either by DDM or digitonin. (a) Western blot of A₁R solubilised from whole cells using four different detergents (SDS, FC12, DDM or digitonin) and probed with an anti-pentaHis-HRP conjugated antibody. Each lane contains an equal amount of total protein and N-linked glycosylation was removed from all samples using PNGase F prior to SDS-PAGE. A₁R was expressed either in the stable mammalian cell line iGnTI⁻(A₁R-GFP-H₁₀) or by using the recombinant baculovirus bvA₁R-H₁₀ to infect Sf9 cells. The iGnTI⁻ cell line was induced with 1 μ g/ml tetracycline for 24 h and Sf9 cells were infected for 72 h. (b) The amount of functional detergent-solubilised A₁R was determined by measuring specific binding of the antagonist [³H]DPCPX. After the addition of ligand, membranes were solubilised in the detergent indicated and non-bound ligand was separated from receptor-ligand complex on gel-filtration spin columns and measured by liquid scintillation counting: red-filled bars, A₁R expressed in Sf9 cells; blue-filled bars, A₁R expressed in iGnTI⁻ cells. The amount of A₁R in membranes (non-solubilised) was determined by separation of receptor-bound and free radioligand by filtration through glass fibre plates: red hatched bars, A₁R expressed in Sf9 cells; blue hatched bars, A₁R expressed in iGnTI⁻ cells. For ease of comparison, binding data have been normalised with respect to A₁R in membranes (100%), which is equivalent to 120,000 \pm 2000 dpm (n = 3; 2.9 pmol per million cells) for baculovirus-infected Sf9 cells and 7500 \pm 250 dpm (n = 3; 3.8 pmol per million cells) for iGnTI⁻(A₁R-GFP-H₁₀) cells. Absolute levels of A₁R therefore cannot be compared meaningfully between the two expression systems using this bar graph. Binding assays for A₁R contained either 150,000 Sf9 cells or 7500 iGnTI⁻ cells. Each data point was determined in duplicate or triplicate from a single experiment and was plotted as mean \pm SEM. (c) Western blot of DDM-solubilised A₁R, with equal amounts of active receptor per sample (lanes 2, 3, 5 and 6). The blot was probed with an anti-pentaHis-HRP conjugated antibody. Lane 1, iGnTI⁻ parental cells; lanes 2 and 3, iGnTI⁻(A₁R-GFP-H₁₀) stable cell line; lane 4, uninfected Sf9 cells; lanes 5 and 6, bvA₁R-H₁₀ infected Sf9 cells. N-Linked glycosylation was removed using PNGase F where indicated (+). A₁R was expressed either in the stable mammalian cell line iGnTI⁻(A₁R-GFP-H₁₀) or by using the recombinant baculovirus bvA₁R-H₁₀ to infect Sf9 cells. Bands corresponding to A₁R-GFP-H₁₀ in mammalian cells are indicated with a yellow asterisk (*). The iGnTI⁻ cell line was induced with 1 μ g/ml tetracycline for 24 h and insect cells were infected with recombinant baculovirus for 72 h. The amount of functional A₁R was determined by measuring specific binding of the antagonist [³H]DPCPX binding.

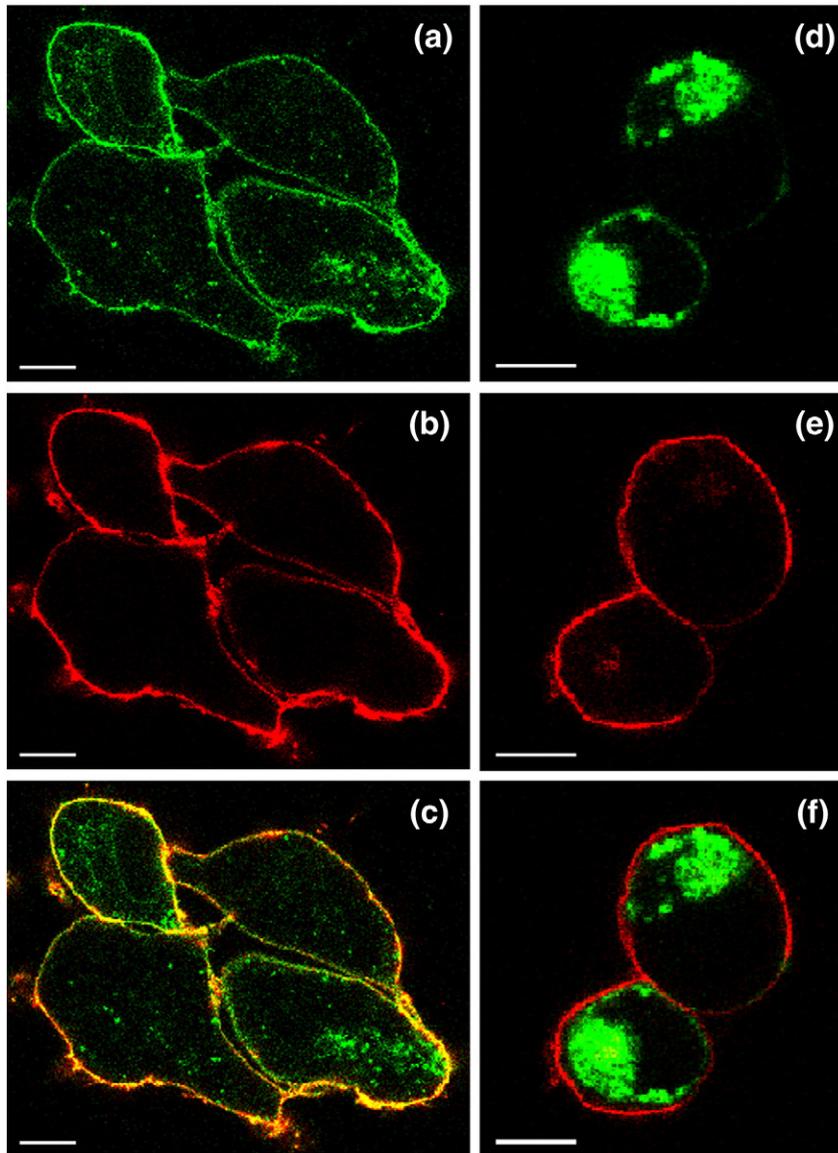


Fig. 8. A₁R-GL26 is misfolded when expressed in mammalian cells. (a–c) Confocal micrographs of the iGnTI⁻ (A₁R-GFP-H₁₀) cell line after 24 h of induction with tetracycline. Cells were fixed using paraformaldehyde and the plasma membrane was defined by staining with Alexa Fluor 647-conjugated conA prior to visualisation. Unlabelled iGnTI⁻ parental cells showed no fluorescence (data not shown). The scale bar represents 10 μm. (d–f) Confocal micrographs of the iGnTI⁻ (A₁R-GL26-GFP-H₁₀) cell line after 24 h of induction with tetracycline. Cells were fixed using paraformaldehyde and the plasma membrane was defined by staining with Alexa Fluor 647-conjugated conA prior to visualisation. Unlabelled iGnTI⁻ parental cells showed no fluorescence (data not shown). The scale bar represents 10 μm.

Secondly, we are using the assay as a guide rather than as an exact measure for determining the number of molecules of the target membrane protein that are correctly folded compared to the number of molecules that are misfolded.

Knowing that a proportion of an expressed membrane protein is misfolded is important. Many efforts have been made to parallelise expression of membrane proteins to facilitate high-throughput post-genomic approaches to determine rapidly membrane protein structures [3]. Although it has proven possible

to do this for bacterial membrane proteins, it has proven harder to replicate these strategies for mammalian membrane proteins, partly because yields of membrane protein suggested from the quantification of polypeptide expressed have not reflected the yield of purified membrane protein. There are two factors that could explain this. Firstly, as described here, most of the membrane protein could be expressed in a misfolded state and therefore cannot be purified in mild detergents. Secondly, membrane proteins are often unstable in detergent and therefore they become

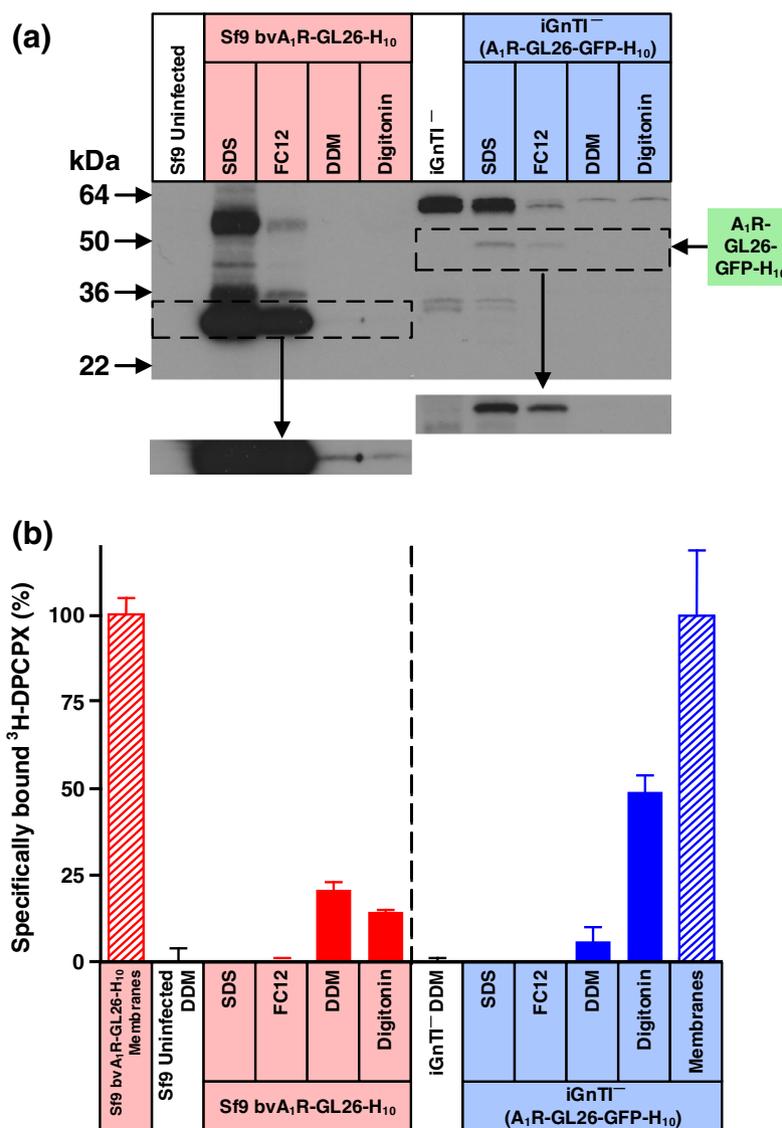


Fig. 9. Misfolded A₁R-GL26 is poorly solubilised either by DDM or digitonin. (a) Western blot of A₁R solubilised from whole cells using four different detergents (SDS, FC12, DDM or digitonin) and probed with an anti-pentaHis-HRP conjugated antibody. Each lane contains an equal amount of total protein. A₁R was expressed either in the stable mammalian cell line [iGnTI⁻(A₁R-GL26-GFP-H₁₀)] or by using the recombinant baculovirus bvA₁R-GL26-H₁₀ to infect Sf9 cells. The iGnTI⁻ cell line was induced with 1 μg/ml tetracycline for 24 h and Sf9 cells were infected for 72 h. The Western blot inserts are a 4× longer exposure. (b) The amount of functional detergent-solubilised A₁R-GL26 was determined by measuring specific binding of the antagonist [³H]DPCPX. After the addition of ligand, membranes were solubilised in the detergent indicated and non-bound ligand was separated from receptor–ligand complex on gel-filtration spin columns and measured by liquid scintillation counting: red-filled bars, A₁R-GL26 expressed in Sf9 cells; blue-filled bars, A₁R-GL26 expressed in iGnTI⁻ cells. The amount of A₁R-GL26 in membranes (non-solubilised) was determined by separation of receptor-bound and free radioligand by filtration through glass fibre plates: red hatched bars, A₁R-GL26 expressed in Sf9 cells; blue hatched bars, A₁R-GL26 expressed in iGnTI⁻ cells. For ease of comparison, binding data have been normalised with respect to A₁R-GL26 in membranes (100%), which is equivalent to 17,400 ± 800 dpm (*n* = 3; 435 fmol per million cells) for baculovirus-infected Sf9 cells and 2000 ± 350 (*n* = 2; 48 fmol per million cells) for iGnTI⁻(A₁R-GL26-GFP-H₁₀) cells. Absolute levels of A₁R therefore cannot be compared meaningfully between the two expression systems using this bar graph. Binding assays for A₁R-GL26 contained 150,000 cells. Each data point was determined in duplicate or triplicate from a single experiment and was plotted as mean ± SEM.

inactive and aggregate during solubilisation and purification. The assay described here will define which is the problematic step, thus directing resources

to solving the relevant problem. For example, knowing that the majority of membrane protein is misfolded in the baculovirus expression system suggests that using

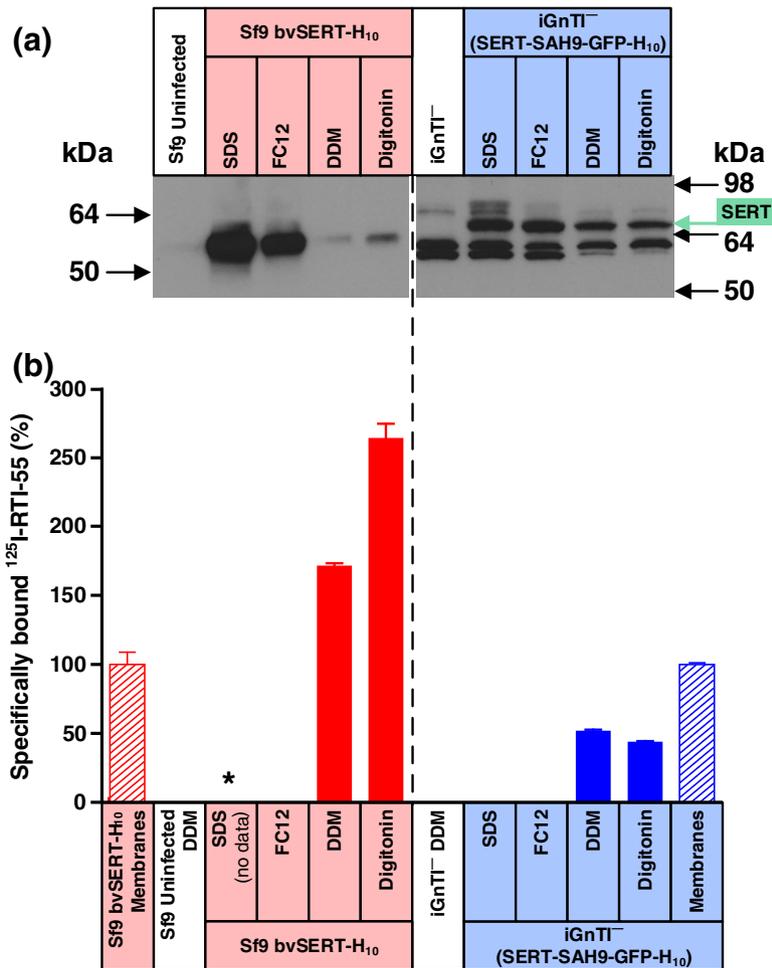


Fig. 10. Misfolded SERT produced by the baculovirus expression system is poorly solubilised by either DDM or digitonin. (a) Western blot of SERT solubilised from whole cells using four different detergents (SDS, FC12, DDM or digitonin) and probed with an anti-pentaHis-HRP conjugated antibody. Each lane contains an equal amount of total protein. SERT was expressed either in the stable mammalian cell line iGnTI⁻ (SERT-SAH9-GFP-H₁₀) or by using the recombinant baculovirus bvSERT-H₁₀ to infect Sf9 cells. The iGnTI⁻ cell line was induced with 1 µg/ml tetracycline for 24 h and Sf9 cells were infected for 48 h. The broken line indicates separate blots. (b) The amount of functional detergent-solubilised SERT was determined by measuring specific binding of the ligand [¹²⁵I]RTI-55. After the addition of ligand, membranes were solubilised in the detergent indicated and non-bound ligand was separated from receptor–ligand complex on gel-filtration spin columns and measured by liquid scintillation counting: red-filled bars, SERT expressed in Sf9 cells; blue-filled bars, SERT-SAH9 expressed in iGnTI⁻ cells; *, not determined. The amount of SERT in membranes (non-solubilised) was determined by separation of receptor-bound and free radioligand by filtration through glass fibre plates: red hatched bars, SERT expressed in Sf9 cells; blue hatched bars, SERT-SAH9 expressed in iGnTI⁻ cells. For ease of comparison, binding data have been normalised with respect to SERT in membranes (100%), which is equivalent to 10,200 ± 950 dpm (*n* = 2; 75.7 fmol per million cells) for baculovirus-infected Sf9 cells and 35,400 ± 420 dpm (*n* = 2; 730 fmol per million cells) for iGnTI⁻ (Sert-SAH9-GFP-H₁₀) cells. Therefore, absolute levels of SERT cannot be compared meaningfully between the two expression systems using this bar graph. Binding assays for SERT contained either 28,000 Sf9 cells or 10,000 iGnTI⁻ cells. Each data point was determined in duplicate or triplicate from a single experiment and was plotted as mean ± SEM.

stable inducible mammalian cell lines could improve yields [29]. In the work described here, the AT₁R expressed in the baculovirus expression system would yield only 0.1 mg/l of functional receptor, whereas the stable clonal cell line iHEK(AT₁R-GFP-H₁₀) would yield 0.5 mg/l. However, the major advantage of using

the mammalian expression system is that there is little or no misfolded AT₁R expressed.

Is the misfolded membrane protein expressed in insect cells a potential problem for downstream purification and crystallisation? Even though DDM is a mild detergent and cannot solubilise misfolded

protein as well as SDS or FC12, misfolded AT₁R is the major component of DDM-solubilised insect cell membranes. In the initial stages of a project, this could be highly misleading, as it would appear that major losses were being incurred on, for example, the first Ni²⁺-affinity column, when in actual fact, it may be the case that the only protein lost was the misfolded material and that the yields of the correctly folded protein were around 80–90%. In the worst instance, researchers may note that FC12 extracts more of the target protein than DDM and then waste many years trying to purify and crystallise this material, not knowing that the target protein was likely to be totally inactive. Interestingly, the work here shows that β_1 AR is expressed as a mixture of both folded and misfolded receptors, but β_1 AR was purified and crystallised and its structure was determined without knowing this. Two effects may help in reducing the impact of misfolded membrane proteins on crystallisation trials. Firstly, SEC is a frequently used step in protein purification and will effectively remove any misfolded protein. Secondly, misfolded membrane proteins have a tendency to aggregate; thus, this portion may just “disappear”, either through retention on columns by non-specific effects or by being unable to pass through pre-filters that are normally present upstream of columns run on automated protein purification equipment. Thirdly, during crystallisation, any remaining inactive protein will precipitate more readily than the folded protein, hopefully allowing crystals to form later on.

Why are misfolded membrane proteins produced in the baculovirus expression system? Although there are many potential differences between insect cells and mammalian cells that may reduce the efficiency of membrane protein folding (e.g., potential specificity and amounts of molecular chaperones, different lipid composition, etc.), there are two overriding factors that have to be considered. Baculoviruses are lytic viruses, and one of the first effects of the virus is to impair the cells' secretory pathway, which is precisely where membrane proteins are folded. Thus, during the infection cycle, the rate of secretion decreases and it is also observed that post-translational modifications such as N-glycosylation also decrease [30]. In addition, the polyhedrin promoter is one of the strongest known eukaryotic promoters, resulting in the polyhedrin mRNA transcript representing over 20% of the cellular polyadenylated RNA [31,32] and polyhedrin representing over 50% of the total cellular protein upon infection of a wild-type baculovirus [33]. Thus, it is highly likely that production of too much mRNA of a target membrane protein, which could well overwhelm the secretory pathway due to insufficient folding factors, in combination with an impairment in the secretory pathway caused by the baculovirus, combines to facilitate the production of misfolded membrane proteins. It is interesting to note that where careful comparisons have been made with mamma-

lian expression systems that utilise viruses with strong promoters, such as the semiliki forest virus expression system, misfolded and inactive membrane protein has also been observed [19,34,35]. Thus, the current successes with the production of authentically folded membrane proteins in mammalian cells for structural studies are all about ensuring that there is a balance between the amount of mRNA produced and the ability of the membrane protein to fold [29]. This will be different for each membrane protein and will have to be optimised empirically on a case-by-case basis. However, the differential solubility assay described here will ensure that expression of only the correctly folded membrane protein will be optimised.

Materials and Methods

Materials

All radiolabelled ligands were purchased from PerkinElmer: [¹²⁵I]sar¹-Ile⁸-angiotensin II ([¹²⁵I]Sar¹), [³H]dihydroalprenolol ([³H]DHA), [³H]dipropylcyclopentylxanthine ([³H]DPCPX) and [¹²⁵I]2 β -carbomethoxy-3 β -(4-iodophenyl)tropane ([¹²⁵I]RTI-55). The detergents *n*-dodecyl β -D-maltopyranoside (DDM) and fos-choline-12 (FC12) were purchased from Anatrace; SDS was purchased from Sigma and digitonin was purchased from Calbiochem. Anti-penta-histidine antibody conjugated to HRP (anti-pentaHis-HRP) was purchased from Qiagen. A tetracycline-inducible HEK293 cell line, T-Rex™-293 (iHEK), was purchased from Invitrogen. A tetracycline-inducible HEK293S cell line lacking *N*-acetylglucosaminyltransferase I (iGnTI⁻) was kindly provided by Philip J. Reeves (Massachusetts Institute of Technology) [36].

Methods

Constructs

Expression in mammalian cells was performed using derivatives of pcDNA4/TO (Invitrogen). The serotonin transporter cDNA was inserted into the EcoRV/NotI restriction sites in pcDNA4/TO, for expression from the tetracycline-inducible CMV promoter, and then a cassette encoding enhanced GFP, the StrepII tag and a decahistidine (H₁₀) tag was inserted after SERT in the NotI/ApaI sites (plasmid pJMA111, kindly provided by J. Andréll, MRC Laboratory of Molecular Biology). The cDNA clone for human angiotensin II type 1 receptor (AT₁R) was obtained from the Missouri S&T cDNA Resource Center[†], amplified by polymerase chain reaction, flanked with EcoRV and NotI sites and cloned into the corresponding sites of pJMA111 to create plasmid pJAP2, which expressed AT₁R-GFP-H₁₀. Additionally, the cDNA for the human adenosine A₁R (Missouri S&T cDNA Resource Center) was cloned similarly into the EcoRV/NotI sites to create plasmid pJAP34, which expressed A₁R-GFP-H₁₀. In an effort to create a thermostable A₁R receptor, four mutations that stabilised the adenosine A_{2A} receptor in the active state (L48A, A54L, T65A, Q89A) [37] were transferred to A₁R (mutations L51A, A57L, L68A, Q92A). In addition, the mutations N148G and N159G were included to remove the putative N-linked

glycosylation sites. To remove flexible regions, we truncated the N-terminus between Pro2 and Ile5, truncated the C-terminus at Phe307 and also added the sequence VLRQQEPPFKAA to the C-terminus, thus generating A₁R-GL26. A synthetic cDNA encoding A₁R-GL26 (Life Technologies) was cloned into the EcoRV/NotI sites in pJMA111 creating pJAP37, which expressed A₁R-GL26-GFP-H₁₀. For generating baculoviruses, AT₁R was cloned into the BamHI/EcoRI sites of the transfer vector pBacPAK8 (Clontech), A₁R was cloned into the XhoI/EcoRI sites and A₁R-GL26 was cloned into the EcoRI/EagI sites, creating plasmids pJAP15, pJAP44 and pJAP33, respectively. Additionally, AT₁R was cloned into the BamHI/EcoRI sites in plasmid pAcGP67-B (BD Biosciences) in order to utilise the acidic glycoprotein gp67 signal sequence (LS) preceding the N-terminus of AT₁R, creating plasmid pJAP16, which expressed AT₁R-LS-H₁₀. All baculovirus sequences were engineered to contain a C-terminal tobacco etch virus cleavage site and H₁₀ tag. All constructs were verified by DNA sequencing (Source Biosciences, UK).

Transient transfection, generation of stable cell lines and protein expression

Mammalian expression plasmids for the expression of AT₁R (pJAP2), A₁R (pJAP34) and A₁R-GL26 (pJAP37) were amplified in *E. coli* strain DH5 α , purified using a Maxi-prep kit (Qiagen) and transiently transfected (Gene-Juice, Novagen) into adherent mammalian iHEK cells or iGnTI⁻ cells following the manufacturer's protocol. Cells were grown in Dulbecco's modified Eagle's media supplemented with 10% tetracycline-free foetal bovine serum (Invitrogen) and 5 μ g/ml blasticidin (Invitrogen) and incubated at 37 °C in an atmosphere containing 5% CO₂. Expression of plasmids was induced by addition of 1 μ g/ml tetracycline and incubated at 37 °C for 24 h. Stable cell lines were generated by selection with media containing 200 μ g/ml Zeocin (Invitrogen). An iGnTI⁻ stable cell line expressing a thermostable mutant of SERT, SERT-SAH9 (J. Andr  ll and C. Tate, unpublished results; Ref. [38]) and (iGnTI⁻ SERT-SAH9-GFP-H₁₀) was kindly provided by J. Andr  ll. A highly expressing clonal AT₁R-GFP-H₁₀ cell line was selected from a polyclonal cell line using fluorescence-activated cell sorting. After expression, cells were washed twice in phosphate-buffered saline (PBS), counted using the Countess Automated Cell Counter (Invitrogen), pelleted (1200g for 5 min) and resuspended at 10 million cells per millilitre in ice-cold cell buffer [50 mM Tris (pH 7.4) and 150 mM NaCl supplemented with Complete EDTA (ethylenediaminetetraacetic acid)-Free Protease Inhibitor Cocktail (Roche)]. Cell suspensions were flash frozen in liquid nitrogen and stored at -80 °C.

Recombinant baculovirus generation and protein expression

Recombinant baculoviruses that expressed AT₁R, A₁R or A₁R-GL26 were generated using the BaculoGold Baculovirus Expression System according to manufacturer's protocol (BD Bioscience). Viruses were isolated by plaque purification and screened for expression by Western blotting using an anti-pentaHis-HRP antibody. Recombinant baculovirus that expressed SERT with a H₁₀ tag at its C-terminus was previously described [16,26]. Recombinant baculovirus that expressed β_1 AR with a H₁₀

tag at its C-terminus [24] was kindly provided by R. Nehme (MRC Laboratory of Molecular Biology) and a thermostable β_1 AR fused to thioredoxin (ts β_1 AR) was kindly provided by T. Warne (MRC Laboratory of Molecular Biology). Recombinant baculoviruses were passaged twice in Sf9 cells to obtain high titre stocks. Viruses were used to infect Sf9, Sf21 or Hi5 cells for 48 or 72 h as indicated. After protein expression, cells were counted using the Countess Automated Cell Counter (Invitrogen), pelleted (1200g for 5 min) and washed twice in PBS, and the cell pellet was resuspended at 10 million cells per millilitre in ice-cold cell buffer [50 mM Tris (pH 7.4) and 150 mM NaCl supplemented with Complete EDTA-Free Protease Inhibitor Cocktail (Roche)]. Cell suspensions were flash frozen in liquid nitrogen and stored at -80 °C.

Western blotting

Cell suspensions were sonicated briefly and the total protein concentration was determined using the Bradford assay [39]. Samples were then solubilised in the detergent indicated [SDS, FC12, DDM or digitonin; all at 1% (w/v) final concentration] at either 4 °C (FC12, DDM, digitonin) or 20 °C (SDS) for 1 h. For blots corresponding to the differential solubility assay, the solubilise was centrifuged at 280,000g for 30 min at 4 °C to remove the insoluble fraction. SDS-loading buffer was added to the supernatant (corresponding to approximately 150,000 cells), and samples were separated on a 4–20% Tris glycine gel and transferred to nitrocellulose using standard techniques. Membranes were probed with anti-pentaHis-HRP at a dilution of 1:1000 and developed using enhanced chemiluminescence (GE Healthcare). Where indicated, 2 μ l of PNGase F (New England Biolabs) was added to 15 μ l of the supernatant and incubated at 37 °C for 1 h prior to SDS-PAGE to remove N-linked glycosylation.

Thermostability assay of detergent-solubilised AT₁R

The cell suspension containing unpurified AT₁R was sonicated briefly and diluted into buffer [50 mM Tris (pH 7.4), 5 mM MgCl₂, 1 mM EDTA, 0.1% (w/v) bovine serum albumin, 150 mM NaCl and 40 μ g/ml bacitracin]. [¹²⁵I]Sar¹ and unlabelled Sar¹ were added to give final concentrations of 0.5 nM and 100 nM, respectively, and incubated for 1 h at room temperature before chilling on ice and solubilising in 1% DDM (w/v, final concentration). The samples were then heated at varying temperatures for 30 min and the [¹²⁵I]Sar¹-bound receptor was separated from the free radioligand by gel-filtration spin columns as described previously [40–43]. Background was determined by adding radioligand to non-transfected parental mammalian cells or uninfected insect cells. Each reaction was performed in triplicate. Results were evaluated by nonlinear regression using GraphPad Prism.

Detergent-solubilised and membrane-bound radioligand binding assays

Cell suspensions were sonicated briefly and the total protein concentration was determined using the Bradford assay [39]. Cells were then diluted into buffer [150 mM NaCl and 50 mM Tris (pH 7.4)], incubated with the

respective radioligand (1 h, 4 °C) and solubilised in a final concentration of 1% detergent (DDM, FC12, digitonin, SDS) for 1 h at 4 °C. [³H]DHA was used at a final concentration of 200 nM and [³H]DPCPX was used at a final concentration of 39 nM in 150 mM NaCl and 50 mM Tris (pH 7.4). [¹²⁵I]RTI-55 was used at a concentration of 1 nM in PBS. [¹²⁵I]Sar¹ was used as per the thermostability assay mentioned above. Bound and free radioligands were separated on gel-filtration spin columns as above.

To determine the amount of SERT, AT₁R, β₁AR or A₁R present in cell membranes, we performed binding assays as mentioned above but without the samples being solubilised with detergent. Separation of receptor-bound and free radioligands was achieved by filtration through a 96-well glass fibre filter plates (Millipore) pre-treated with 0.1% polyethyleneimine [38] except for [¹²⁵I]Sar¹ where no polyethyleneimine was used. Background for both assays was determined by adding radioligand to non-transfected parental mammalian cells or uninfected insect cells.

FSEC and SEC analysed by Western blotting

The void volume (8.16 ml) of the Superdex 200 10/300 (24 ml) (GE healthcare) was determined by running blue dextran through the column and observing where it eluted using A₂₈₀. For FSEC, approximately 5 million iHEK(AT₁R-GFP-H₁₀) cells were thawed on ice and sonicated briefly. Cells were incubated at room temperature for 1 h with 40 nM Sar¹ before chilling on ice and solubilising in 1% DDM (w/v, final concentration). Followed by centrifugation at 280,000g for 30 min at 4 °C. The supernatant was then passed through a 0.22-μm filter and injected onto a Superdex 200 10/300 column pre-equilibrated with running buffer [0.03% (w/v) DDM, 50 mM Tris (pH 7.4), 150 mM NaCl and 1 μM Sar¹]. The fluorescence of eluent was detected by a Hitachi fluorometer (mV) set to an excitation of 488 nm and an emission of 525 nm. Approximately 5 million cells were sonicated, incubated with ligand, solubilised and centrifuged as described above, in order to detect bvAT₁R-H₁₀ produced in Sf9 cells. The eluent was detected by Western blotting as described above and bands corresponding to bvAT₁R-H₁₀ were quantified by densitometry using ImageJ.

Fixing and staining cells for analysis by confocal laser-scanning microscopy

Cells were grown on 35-mm glass bottom culture dishes, induced for 24 h under standard conditions and fixed using 2% paraformaldehyde [28]. After washing with PBS, we selectively stained membranes using a solution (10 μg/ml) of concanavalin A (ConA)-Alexa Fluor 647 conjugate (Invitrogen) in PBS for 10 min at room temperature. After washing with PBS, we stored icells n fresh PBS with 0.02% Na azide at 4 °C protected from light. Cells were visualised on a Leica TCS SP8 STED inverted laser-scanning microscope with 63× oil-immersion objective and a 1.4 numerical aperture. The white light laser was set to a wavelength of 488 nm to excite GFP and to 633 nm for Alexa Fluor 647 with the pinhole emission wavelength set to 580 nm.

Acknowledgements

J.A.T. was the recipient of an MRC-funded studentship and research in the laboratory of C.G.T. is funded by a core grant from the Medical Research Council (MRC U105197215).

Received 15 August 2014;

Received in revised form 12 October 2014;

Accepted 13 October 2014

Available online 28 October 2014

Keywords:

eukaryotic membrane proteins;
overexpression;
G protein-coupled receptors;
transporters

www.cdna.org

Abbreviations used:

FSEC, fluorescence-detection size-exclusion chromatography; GPCR, G protein-coupled receptor; A₁R, A₁ receptor; GFP, green fluorescent protein; PBS, phosphate-buffered saline.

References

- [1] Bill RM, Henderson PJ, Iwata S, Kunji ER, Michel H, Neutze R, et al. Overcoming barriers to membrane protein structure determination. *Nat Biotechnol* 2011;29:335–40.
- [2] Grisshammer R, Tate CG. Overexpression of integral membrane proteins for structural studies. *Q Rev Biophys* 1995;28:315–422.
- [3] Mancía F, Love J. High throughput platforms for structural genomics of integral membrane proteins. *Curr Opin Struct Biol* 2011;21:517–22.
- [4] Drew D, Lerch M, Kunji E, Slotboom DJ, de Gier JW. Optimization of membrane protein overexpression and purification using GFP fusions. *Nat Methods* 2006;3:303–13.
- [5] Kawate T, Gouaux E. Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. *Structure* 2006;14:673–81.
- [6] Drew DE, von Heijne G, Nordlund P, de Gier JW. Green fluorescent protein as an indicator to monitor membrane protein overexpression in *Escherichia coli*. *FEBS Lett* 2001;507:220–4.
- [7] Geertsma ER, Groeneveld M, Slotboom DJ, Poolman B. Quality control of overexpressed membrane proteins. *Proc Natl Acad Sci USA* 2008;105:5722–7.
- [8] Newstead S, Kim H, von Heijne G, Iwata S, Drew D. High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* 2007;104:13936–41.
- [9] Haggie PM, Stanton BA, Verkman AS. Diffusional mobility of the cystic fibrosis transmembrane conductance regulator mutant, delta F508-CFTR, in the endoplasmic reticulum

- measured by photobleaching of GFP-CFTR chimeras. *J Biol Chem* 2002;277:16419–25.
- [10] Illing ME, Rajan RS, Bence NF, Kopito RR. A rhodopsin mutant linked to autosomal dominant retinitis pigmentosa is prone to aggregate and interacts with the ubiquitin proteasome system. *J Biol Chem* 2002;277:34150–60.
- [11] Saliba RS, Munro PM, Luthert PJ, Cheetham ME. The cellular fate of mutant rhodopsin: quality control, degradation and aggresome formation. *J Cell Sci* 2002;115:2907–18.
- [12] Ellgaard L, Helenius A. Quality control in the endoplasmic reticulum. *Nat Rev Mol Cell Biol* 2003;4:181–91.
- [13] Akermoun M, Koglin M, Zvalova-looss D, Folschweiller N, Dowell SJ, Gearing KL. Characterization of 16 human G protein-coupled receptors expressed in baculovirus-infected insect cells. *Protein Expression Purif* 2005;44:65–74.
- [14] Venkatakrishnan AJ, Deupi X, Lebon G, Tate CG, Schertler GF, Babu MM. Molecular signatures of G-protein-coupled receptors. *Nature* 2013;494:185–94.
- [15] Tate CG. Overexpression of mammalian integral membrane proteins for structural studies. *FEBS Lett* 2001;504:94–8.
- [16] Tate CG, Blakely RD. The effect of N-linked glycosylation on activity of the Na(+)- and Cl(-)-dependent serotonin transporter expressed using recombinant baculovirus in insect cells. *J Biol Chem* 1994;269:26303–10.
- [17] Tate CG, Whiteley E, Betenbaugh MJ. Molecular chaperones stimulate the functional expression of the cocaine-sensitive serotonin transporter. *J Biol Chem* 1999;274:17551–8.
- [18] Tate CG. Practical considerations of membrane protein instability during purification and crystallisation. *Methods Mol Biol* 2010;601:187–203.
- [19] Shukla AK, Reinhart C, Michel H. Comparative analysis of the human angiotensin II type 1a receptor heterologously produced in insect cells and mammalian cells. *Biochem Biophys Res Commun* 2006;349:6–14.
- [20] Moukhametzianov R, Warne T, Edwards PC, Serrano-Vega MJ, Leslie AG, Tate CG, et al. Two distinct conformations of helix 6 observed in antagonist-bound structures of a beta1-adrenergic receptor. *Proc Natl Acad Sci USA* 2011;108:8228–32.
- [21] Warne T, Moukhametzianov R, Baker JG, Nehme R, Edwards PC, Leslie AG, et al. The structural basis for agonist and partial agonist action on a beta(1)-adrenergic receptor. *Nature* 2011;469:241–4.
- [22] Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, Henderson R, et al. Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* 2008;454:486–91.
- [23] Miller-Gallacher JL, Nehme R, Warne T, Edwards PC, Schertler GF, Leslie AG, et al. The 2.1 Å resolution structure of cyanopindolol-bound beta1-adrenoceptor identifies an intramembrane Na⁺ ion that stabilises the ligand-free receptor. *PLoS One* 2014;9:e92727.
- [24] Warne T, Serrano-Vega MJ, Tate CG, Schertler GF. Development and crystallization of a minimal thermostabilised G protein-coupled receptor. *Protein Expression Purif* 2009;65:204–13.
- [25] Warne T, Chirnside J, Schertler GF. Expression and purification of truncated, non-glycosylated turkey beta-adrenergic receptors for crystallization. *Biochim Biophys Acta* 2003;1610:133–40.
- [26] Tate CG. Baculovirus-mediated expression of neurotransmitter transporters. *Methods Enzymol* 1998;296:443–55.
- [27] Reeves PJ, Kim JM, Khorana HG. Structure and function in rhodopsin: a tetracycline-inducible system in stable mammalian cell lines for high-level expression of opsin mutants. *Proc Natl Acad Sci USA* 2002;99:13413–8.
- [28] Tate CG, Haase J, Baker C, Boorsma M, Magnani F, Vallis Y, et al. Comparison of seven different heterologous protein expression systems for the production of the serotonin transporter. *Biochim Biophys Acta* 2003;1610:141–53.
- [29] Andrell J, Tate CG. Overexpression of membrane proteins in mammalian cells for structural studies. *Mol Membr Biol* 2013;30:52–63.
- [30] Jarvis DL, Summers MD. Glycosylation and secretion of human tissue plasminogen activator in recombinant baculovirus-infected insect cells. *Mol Cell Biol* 1989;9:214–23.
- [31] Rohel DZ, Cochran MA, Faulkner P. Characterization of two abundant mRNAs of *Autographa californica* nuclear polyhedrosis virus present late in infection. *Virology* 1983;124:357–65.
- [32] Adang MJ, Miller LK. Molecular cloning of DNA complementary to mRNA of the baculovirus *Autographa californica* nuclear polyhedrosis virus: location and gene products of RNA transcripts found late in infection. *J Virol* 1982;44:782–93.
- [33] Miyamoto C, Smith GE, Farrell-Towt J, Chizzonite R, Summers MD, Ju G. Production of human c-myc protein in insect cells infected with a baculovirus expression vector. *Mol Cell Biol* 1985;5:2860–5.
- [34] Sen S, Jaakola VP, Heimo H, Engstrom M, Larjoomaa P, Scheinin M, et al. Functional expression and direct visualization of the human alpha 2B-adrenergic receptor and alpha 2B-AR-green fluorescent fusion protein in mammalian cell using Semliki Forest virus vectors. *Protein Expression Purif* 2003;32:265–75.
- [35] Shukla AK, Haase W, Reinhart C, Michel H. Biochemical and pharmacological characterization of the human bradykinin subtype 2 receptor produced in mammalian cells using the Semliki Forest virus system. *Biol Chem* 2006;387:569–76.
- [36] Reeves PJ, Callewaert N, Contreras R, Khorana HG. Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible *N*-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc Natl Acad Sci USA* 2002;99:13419–24.
- [37] Lebon G, Bennett K, Jazayeri A, Tate CG. Thermostabilisation of an agonist-bound conformation of the human adenosine A(2A) receptor. *J Mol Biol* 2011;409:298–310.
- [38] Abdul-Hussein S, Andrell J, Tate CG. Thermostabilisation of the serotonin transporter in a cocaine-bound conformation. *J Mol Biol* 2013;425:2198–207.
- [39] Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 1976;72:248–54.
- [40] Serrano-Vega MJ, Magnani F, Shibata Y, Tate CG. Conformational thermostabilization of the beta1-adrenergic receptor in a detergent-resistant form. *Proc Natl Acad Sci USA* 2008;105:877–82.
- [41] Shibata Y, Gvozdenovic-Jeremic J, Love J, Kloss B, White JF, Grisshammer R, et al. Optimising the combination of thermostabilising mutations in the neurotensin receptor for structure determination. *Biochim Biophys Acta* 2013;1828:1293–301.
- [42] Magnani F, Shibata Y, Serrano-Vega MJ, Tate CG. Co-evolving stability and conformational homogeneity of the human adenosine A2a receptor. *Proc Natl Acad Sci USA* 2008;105:10744–9.
- [43] Shibata Y, White JF, Serrano-Vega MJ, Magnani F, Aloia AL, Grisshammer R, et al. Thermostabilization of the neurotensin receptor NTS1. *J Mol Biol* 2009;390:262–77.