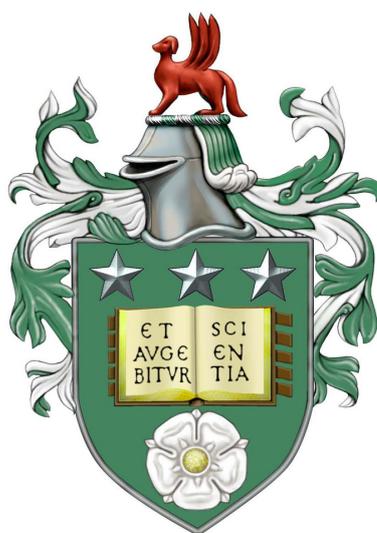


**Structural and Pharmacological Analysis of Novel
Adenosine A_{2A} Receptor Ligands to Guide Rational Drug
Design**

by
Claudia Tania Zilian

Submitted in accordance with the requirements
for the degree of Doctor of Philosophy



The University of Leeds
School of Biomedical Sciences & Astbury Centre for Structural
Molecular Biology
February 2022

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors of this work has been explicitly indicated below. The candidate confirms that appropriate credit had been given within the thesis where reference has been made to the work of others.

The work presented in Appendix A has been published in the following articles:

- [1] **Stohrer, C.**, Horrell, S., Meier, S., Sans, M., Von Stetten, D., Hough, M., Goldman, A., Monteiro, D.C. and Pearson, A.R. Homogeneous batch micro-crystallization of proteins from ammonium sulfate. *Acta Crystallographica Section D: Structural Biology*, **77**(2), pp. 194–204 (2021).

The candidate (CS) was involved in conceiving the study, planning and performing experiments regarding ADC crystallisation and wrote the manuscript. CS, AG, DCFM, and ARP conceived the study and planned the experiments. CS, SH, SM, and MS produced the proteins and carried out the crystallisation experiments. DCFM, DVS and ARP collected and analysed the X-ray diffraction data. CS, AG, DCFM, and ARP wrote the manuscript. All authors assisted in manuscript revision and editing.

Figures 1, 2, and S1 of [1] were used in Appendix A.

Permission of redistribution: Article [1] is published under an open-access licence, which permits unrestricted use, distribution and reproduction in any medium if appropriate credit is given to the original author(s), indication is given when changes were made and a link to the license is provided. Link to licence: <https://creativecommons.org/licenses/by/4.0/>

- [2] Monteiro, D.C., Von Stetten, D., **Stohrer, C.**, Sans, M., Pearson, A.R., Santoni, G., Van Der Linden, P. and Trebbin, M. 3D-MiXD: 3D-printed X-ray-compatible microfluidic devices for rapid, low-consumption serial synchrotron crystallography data collection in flow. *IUCrJ*, **7**(2) pp. 207–219 (2020).

The candidate (CS) produced the protein and crystals for the study and assisted in setting up the device for data collection and collecting the diffraction data. DCFM, ARP, and MT conceived the study and planned the experiments. CS, SM, and DCFM produced the proteins and crystals for the experiment. DCFM developed the original design of the device and manufactured it with assistance of PVDL. DCFM, DVS, CS, and GS collected the X-ray diffraction data and DCFM, DVS, ARP and GS analysed and processed it. DCFM, ARP, and MT wrote the manuscript. All authors assisted in manuscript revision and editing.

Additional publications not included in this thesis:

- [3] Cecchetti, C., Strauss, J., **Stohrer, C.**, Naylor, C., Pryor, E., Hobbs, J., Tanley, S., Goldman, A. and Byrne, B. A novel high-throughput screen for identifying lipids that stabilise membrane proteins in detergent based solution. *PLoS ONE*, **16**(7) pp. 1–20 (2021).

The candidate (CS) expressed and purified the A_{2A}R for stability testing and was involved in interpreting the data and writing the manuscript. CC, JS, JH, BB, and AG conceived the study and planned the experiments. CC, JS, CN, EP, and BB carried out literature research on lipids and designed the lipid screen. CC and JS produced the first batch of the RAMP Lipid Screen at Imperial College London. CC, JS, and CS expressed and purified test proteins for stability testing carried out by JS. CC, JS, CS, AG, and BB wrote the manuscript. All authors assisted in manuscript revision and editing.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis maybe published without proper acknowledgement. The right of Claudia Tania Zilian to be identified as Author of this work has been asserted by Claudia Tania Zilian in accordance with the Copyright, Designs and Patents act 1988.

© 2022 The University of Leeds and Claudia Tania Zilian

Acknowledgements

Firstly, I would like to thank my supervisors Dan Donnelly, Adrian Goldman, and Christos Pliotas, for their continued support and guidance through the many challenges and difficult decisions that had to be taken during this PhD. You not only helped me to develop this project but also helped me to progress as a scientist.

I also want to thank the RAMP consortium and especially Monika Budayova-Spano for giving me the opportunity to be a part of such a fantastic network. I truly enjoyed all the great workshops we had and learned so much about membrane protein crystallography and being a good scientist. I also want to thank my industrial collaborators from Novartis, who gave me the opportunity to dip my toes into drug discovery. Special thanks to Veli-Pekka Jakkola for sharing all his knowledge about the A_{2A}R. I further want to thank his successor Matthias Haffke and Trixi Brandl for dealing with legal affairs and supplying the compounds. A true highlight during my PhD was the secondment at the Universität Hamburg. The main reasons for this positive sentiment are the great supervision and support from Arwen Pearson and Diana Monteiro. I learned so much about micro-crystallography during this time and even got the opportunity to write a first author paper. Thank you so much!

I also need to thank the Janggen-Pöhn foundation for the financial support so I could extend my PhD for an additional year. Especially in times of Covid-19, getting some additional time was a great relief.

I was incredibly lucky to get allocated an amazing Lab group. Many thanks to the amazing postdocs Maren Tompson, Steve Harborne, Julie Heggelund and Francis Totanes who welcomed me into the lab and took the time to answer all my questions. Special thanks also to James Hillier who helped me to get more self-confidence in my crystal-harvesting skills and held my hand during each beamtime. People from this group were not only great colleagues but also true friends! Thank you, Jessica and Maria, for being my personal motivators, Jack and Jannik for taking me out climbing and Andreas for well-deserved coffee breaks and everyone else in the lab: Ana, Anchal, Brendan, Craig, Emily, and Katie for taking out my plates from the incubator, splitting my cells and just making the lab such a great place. I also want to thank the lab members of the Pliotas lab for welcoming me into their group, especially Bolin Wang for all his help with my last-minute attempt at cryo-EM. A big thank you to Alex Holmes and Alex Flynn who were such a big help with the docking studies, as well as Amelia Lesiuk, Chi Trinh and Pirjo Johnson who were always around to solve our daily problems.

A huge thank you to Jamie Henderson. I actually don't know where to start the list, but being my personal chauffeur, the best lab buddy you can wish for, and proofreading every single word of my thesis are surely at the top. You were always there for me, and I can't thank you enough!

Finally, I would like to thank my family for their support throughout the years. I couldn't have done this without you! Diego, I really enjoy having someone in the family who actually understands what I'm complaining about. Last, but not least, I want to thank my now husband Moritz. I'm eternally grateful for all the support you have given me through these past four years, starting from moving countries, to listening to all my stories and complaints about lab life, and helping me with any computer related issue, including formatting this thesis. By now you're probably an expert in structural biology yourself! Therefore, I think it's more than appropriate that our name is printed on the front of this thesis.

Abstract

Extracellular adenosine signalling is mainly conferred through adenosine receptors, including the adenosine A_{2A} receptor (A_{2A}R), modulating a variety of different physiological responses through the human body. Because of its versatile role, A_{2A}R represents a valuable drug target for neurodegenerative diseases and in immunotherapy for cancer.

This work aims to characterise the receptor-ligand interactions for a selection of novel A_{2A}R screening hits identified by Novartis, to uncover structure-activity relationships that can be exploited for lead identification and further ligand optimisation. To this end, a combination of pharmacological, *in silico*, and structural methods were employed.

Four of the provided hit ligands were characterised as having nanomolar affinity and were selected for structural characterisation. The best hit, DEL-1, displayed a pIC_{50} of 7.3 ± 0.19 in competition binding assays and inhibited agonist-mediated receptor function with a pIC_{50} of 7.7 ± 0.18 . A series of co-crystallisation trials were undertaken using differently stabilised receptor constructs, and adding the ligand at different points during the purification protocol. Additionally, *in meso* soaking experiments were explored. While these strategies were successful with the established antagonist ZM241385, no diffracting crystals were obtained with the novel compounds. *In silico* ligand docking studies predicted that DEL-1 and DEL-2 extend into a subordinate, potentially allosteric binding pocket, likely inducing different receptor conformations that might have hampered the success of the crystallographic studies.

The performed pharmacological and docking studies with FOC-6 suggested ligand binding to a completely different receptor site, possibly a lipid pocket, that prevented the ZM241385–A_{2A}R complex from forming well ordered crystals. Cryo-EM studies are ventured to characterise this interaction.

Overall, this study forms a basis for rational chemical optimisation of the selected hits, which might lead to the discovery of novel, subtype selective or allosteric A_{2A}R ligands and suggests novel strategies for structural characterisation of receptor ligand complexes.

Contents

1	Introduction	1
1.1	Target-based and structure-based drug discovery	1
1.1.1	Experimental methods for lead discovery	3
1.1.2	Computer-aided methods for structure-based drug design	5
1.1.3	The use of protein structures in SBDD	7
1.2	The family of G protein-coupled receptors as drug targets	10
1.2.1	Physiological GPCR functions and signalling	12
1.2.2	Regulating GPCR signalling with ligands and drugs	14
1.3	Strategies to determine GPCR structures	16
1.3.1	Receptor stabilisation for structural studies	17
1.3.2	Crystallisation and structure determination of GPCRs	20
1.3.3	Cryo-EM structural approaches for GPCR drug discovery	21
1.4	The Adenosine A _{2A} Receptor	23
1.4.1	Adenosine signalling	23
1.4.2	Efforts targeting A _{2A} R signalling as therapy in human diseases	24
1.4.3	Characterisation of A _{2A} R structure	26
1.4.4	A _{2A} R ligands and structure-based drug discovery	32
1.5	Screening for novel A _{2A} R ligands	34
1.6	Aims of this project	35
2	Materials and Methods	37
2.1	Standard protocols	37
2.1.1	Buffers and chemicals	37
2.1.2	Media and agar plates	37
2.1.3	SDS-PAGE	38
2.1.4	Western blotting	38
2.1.5	Native-Page	38
2.2	Molecular biology	39
2.2.1	Primers and genes	39
2.2.2	Plasmid backbones	39
2.2.3	Heat-shock transformation	39
2.2.4	Polymerase Chain Reaction	40
2.2.5	Colony PCR	40
2.2.6	Agarose gel electrophoresis	40

2.2.7	DNA purification	41
2.2.8	Infusion cloning	41
2.3	Cell culture	42
2.3.1	Insect cell culture	42
2.3.2	Bac-to-Bac baculovirus expression	42
2.3.3	Protein expression in mammalian cells	44
2.4	Protein purification	47
2.4.1	Protein Quantification	47
2.4.2	A _{2A} R membrane preparation	47
2.4.3	A _{2A} R membrane preparation for binding assays	48
2.4.4	A _{2A} R purification for crystallisation	48
2.5	Biochemical and pharmacological assays	51
2.5.1	Radioligand binding	51
2.5.2	LANCE cAMP assay	52
2.5.3	Gel shift assay	53
2.5.4	nanoDSF	54
2.6	Crystallography	54
2.6.1	Crystallisation screens	54
2.6.2	Ligand exchange	55
2.6.3	<i>In-meso</i> crystallisation	55
2.6.4	<i>In-meso</i> soaking	56
2.7	Data collection and Structure solving	56
2.7.1	Semi-automated serial data collection	56
2.7.2	Data processing	56
2.8	Electron microscopy	57
2.8.1	Receptor-Fab complex preparation	57
2.8.2	Negative stain	57
2.8.3	Cryo-grid preparation	58
2.8.4	Cryo-EM data collection	58
2.9	<i>In silico</i> structural analysis	58
2.9.1	Generating a homology model for the A _{2B} R	58
2.9.2	Docking with Glide	59
2.9.3	Docking using constraints	59
2.9.4	Docking with Autodock Vina	60
3	Pharmacological characterisation of novel A_{2A}R drug-screening hits	61
3.1	A novel selection of A _{2A} R ligand hits	61
3.2	Measuring concentration-response curves and affinity	64
3.2.1	Setting up a competition assay for A _{2A} R	64
3.2.2	Determining the binding constants for the novel ligands	66
3.2.3	Comparison to insect cell expression and effect of different constructs	69
3.3	Determining ligand antagonism	71
3.4	Thermostabilising effect of the compounds	73
3.4.1	Melting temperatures determined using a gel shift assay	74

3.4.2	Determining the melting temperature by nanoDSF	75
3.5	Discussion	76
4	<i>In silico</i> structural studies of A_{2A} receptor-ligand interactions	79
4.1	Validation of docking approach	79
4.2	Docking of novel ligands	82
4.2.1	Comparison of predicted binding for different bioisosteres	84
4.2.2	Docking poses of ligands containing the same chemical scaffold	86
4.2.3	Docking novel compounds using knowledge-based binding constraints	89
4.2.4	Screening for alternative binding sites for FOC-6	91
4.3	Predicting receptor subtype selectivity	92
4.3.1	Docking the novel ligands into an inactive state A ₁ R crystal structure	93
4.3.2	Docking the novel ligands into an inactive state A _{2B} R homology model	97
4.4	Discussion	102
4.4.1	Validity of the predictions	102
4.4.2	Possibilities for chemotype optimisation	102
4.4.3	Possible allosteric ligands	103
4.4.4	Possible subtype selective ligands	104
5	Towards experimental structures of novel A_{2A} receptor ligand complexes	107
5.1	Crystallisation strategies for different receptor-ligand complexes	107
5.2	Receptor construct choice	108
5.3	Expression of A _{2A} R in <i>Sf9</i> cells for structural studies	109
5.4	Effects of C-terminal HRV-cleavage site	110
5.5	Crystallisation trials with A _{2A} -StaR2-bRIL construct	114
5.5.1	Crystal soaking	114
5.5.2	Ligand exchange	115
5.5.3	Co-purification	117
5.5.4	Improving protein quality	118
5.6	Crystallisation trials with the A _{2A} -bRIL construct	120
5.6.1	Crystallisation and data collection with ZM241385	120
5.6.2	Attempts to exchange ZM241385 with novel ligands	122
5.6.3	Ligand exchange LUF5834 and theophylline	125
5.6.4	ZM241385–FOC-6 complex	126
5.7	Alternatives to crystallisation: Cryo-EM	127
5.8	Discussion	129
6	Overall conclusions and future directions	133
6.1	Key findings of the project	133
6.1.1	Identification of novel A _{2A} R binders	133
6.1.2	Strategies for the determination of early stage ligand co-structures	134
6.1.3	Identification of novel potential subtype-selective and allosteric ligands	135
6.2	Future directions	136
6.2.1	Strategies to further advance structural work	136

6.2.2	Biophysical characterisation of receptor-ligand interactions	137
6.2.3	Functional aspects of ligand binding	138
6.3	Significance of the work	139
A	Exploring novel ways in which crystallography might assist drug discovery	141
A.1	Introduction	143
A.1.1	Strategies for serial time-resolved crystallography	143
A.1.2	Crystal requirements for time-resolved studies	144
A.1.3	Micro-crystallisation strategies	145
A.1.4	Aim	146
A.2	Methods	147
A.3	Micro-crystallisation of ADC	147
A.3.1	Ammonium sulphate concentration to control crystal size	147
A.3.2	Reproducibility of the crystallisation protocol	149
A.4	Micro-crystallisation of other model enzymes	150
A.5	Discussion	151
A.6	Application of microcrystals: testing 3D printed microfluidic chips	152
B	Supplementary material	155
	Bibliography	163

List of Figures

1.1	Target-based drug design process	2
1.2	Principle of DEL screening	4
1.3	Different classes of human GPCR receptors and their ligand-binding sites	11
1.4	G protein-coupled receptor signalling	13
1.5	Effects of different ligands on the biological response of receptors	15
1.6	Stabilising receptor modifications for structural studies of GPCRs	18
1.7	Strategies for GPCR structure determination	21
1.8	Adenosine signalling	24
1.9	Milestone structures of A _{2A} R solved along the receptor's conformational trajectory	27
1.10	Orthosteric binding site of A _{2A} R	29
1.11	Structural changes upon receptor activation	30
1.12	Exemplary A _{2A} R ligands relevant for drug discovery	33
2.1	Plasmid backbone of the most commonly used vectors in this study	39
3.1	Determination of optimal amount of membrane protein for binding assays	65
3.2	Effect of increased DMSO concentration on binding assays	66
3.3	Competition binding curves recorded with different ligands	68
3.4	Binding assays on stabilised constructs from <i>Sf9</i> membranes	70
3.5	Receptor activation upon NECA stimulation	71
3.6	Measuring cellular cAMP production upon NECA activation of A _{2A} R transfected CHO cells after treatment with different inhibitors	72
3.7	Attempt of trying to determine thermostabilising effects using the gel shift assay	74
3.8	Thermostabilising effect of ligands recorded by nanoDSF	76
4.1	Comparison of docking poses for reference compounds to crystal structures	81
4.2	Correlation of docking scores with the previously determined affinities	82
4.3	Docking poses for DEL-1 and FOC-1	84
4.4	Comparison of the docked pose for FOC-2 to the binding pose of the 1,2,4-triazines 4g in the co-crystal structure 3UZA	84
4.5	Comparison between the docking of FOC-3 and the co-crystal structure of triazole compound Cmpd-1	86
4.6	Comparison of docking poses of DEL-2 and DEL-5	87
4.7	Comparison of docking poses of DEL-3 and DEL-6	88

4.8	Docking of DEL-3 and DEL-6 with and without positional π -constraints	90
4.9	Blind docking of FOC-6 into the whole receptor	91
4.10	Alignment of different adenosine receptors	93
4.11	Docking of DEL-4 into the A ₁ R	95
4.12	Docking of DEL-5 and DEL-2 into the wider binding pocket of A ₁ R compared to the docked pose of DEL-5 in A _{2A} R	96
4.13	Homology models generated for A _{2B} R	98
4.14	Docking of ZM241385 and FOC-3 for the A _{2B} R compared to A _{2A} R	99
5.1	Schematic of purification and crystallisation process used for structure determination of receptor-ligand complexes	108
5.2	Example of expression test	109
5.3	Purification of A _{2A} -StaR2-bRIL ^{HRV-HIS}	110
5.4	Crystals obtained for the A _{2A} -StaR2-bRIL ^{HRV-HIS} -ZM241385 complex	111
5.5	Improvement of SEC profile for A _{2A} -StaR2-bRIL ^{HRV-HIS} -theophylline complex	112
5.6	Results for A _{2A} -StaR2-bRIL purification without the HRV cleavage site	113
5.7	Crystal soaking of A _{2A} -StaR2-bRIL-theophylline complex	115
5.8	Crystals obtained after ligand exchange	117
5.9	Co-purification of A _{2A} -StaR2-bRIL in complex with FOC-2	118
5.10	Attempt to optimise A _{2A} -StaR2-bRIL-theophylline protein quality	119
5.11	Purification and crystallisation of A _{2A} -bRIL construct with ZM241385	121
5.12	Ligand exchange displacing ZM241385 with three novel compounds	123
5.13	Electron density maps for the ligand binding site of A _{2A} -bRIL-ZM241385 complex after the attempted ligand exchange	124
5.14	Ligand exchange experiments using A _{2A} -bRIL purified in complex with theophylline or LUF5834	125
5.15	Significantly smaller crystals were obtained after adding FOC-6 to A _{2A} -bRIL-ZM241385 complex	126
5.16	Cryo-EM studies on the A _{2A} R-bRIL-Fab complex	128
A.1	Simplified protein crystallisation diagram	148
A.2	Effect of increasing ammonium sulphate concentration	149
A.3	Effects of buffer and protein age	150
A.4	Schematic of a generic crystallisation phase diagram	152

List of Tables

1.1	Comparison of the different adenosine receptor subtypes	25
2.1	Working concentration for commonly used antibiotics	37
2.2	Standard PCR protocol	40
2.3	Volume of media to be added before transfection	45
2.4	DNA to PEI ratios for HEK293 transfection	46
2.5	DNA to TrasIT-X2 ratios for CHO cell transfection	46
2.6	Ligand concentration added to each purification step	50
2.7	Parameters of reference measurement to convert cpm in dpm	52
3.1	List of novel compounds containing a monocyclic core	62
3.2	List of novel compounds containing a bicyclic core	63
3.3	Affinities of novel ligands determined through competition binding	68
3.4	Antagonist potency of novel ligands determined by recording concentration response curves using the LANCE cAMP assay	73
4.1	Docking scores and residues predicted to form polar interactions for A _{2A} R	83
4.2	Docking scores obtained through blind docking into the whole extracellular portion of A _{2A} R using Autodock vina	92
4.3	Docking scores and residues predicted to form polar interactions for A ₁ R	94
4.4	Docking scores and residues predicted to form polar interactions for A _{2B} R homology model	100
4.5	Comparison of ligand Glide scores for docking against different AR subtypes	101
5.1	Summary of collected data for A _{2A} -StaR2-bRIL ligand exchanged from theophylline to ZM241385	116
B.1	Composition of the DoE screen	155
B.2	List of used primers	158
B.3	List of used A _{2A} R constructs	159
B.4	Docking scores and residues predicted to form polar interactions for ligand docking into the A _{2A} R (4EIY) determined using a positional π -constraint	160
B.5	Data collection and refinement statistics	161

Abbreviations

β_2 AR	β_2 Adrenergic receptor
3D	Three dimensional
A ₁ R	Adenosine A ₁ receptor
A _{2A} R	Adenosine A _{2A} receptor
A _{2B} R	Adenosine A _{2B} receptor
A ₃ R	Adenosine A ₃ receptor
AC	Adenylyl cyclase
ADA	Adenosine deaminase
ADP	Adenosine diphosphate
AM ₁ R	Adrenomedullin 1 receptor
AMP	Adenosine monophosphate
AR	Adenosine receptor
ATP	5'-Adenosine triphosphate
B _{max}	Maximum specific binding
BAM	Biased allosteric modulators
BCA	Bicinchoninic acid
BRET	Bioluminescence resonance energy transfer
bRIL	Thermostabilised apocytochrome b562RIL
BSA	Bovine serum albumin
cAMP	Cyclic adenosine monophosphate
CASR	Calcium sensing receptor
CB ₁	Cannabinoid receptor type 1
CHO	Chinese hamster ovary
CHS	Cholesteryl hemisuccinate
Cl-Adenosine	2-Chloroadenosine
CMV	Cytomegalovirus
CNS	Central nervous system
cpm	Counts per minute
cryo-EM	Cryogenic electron microscopy
CV	Column volumes

D ₂ R	Dopamine D ₂ receptor
DAG	Diacylglycerol
DDM	N-Dodecyl- β -D-Maltoside
DECLs	DNA-encoded chemical libraries
DEER	Double-electron-electron resonance
DM	Decylmaltoside
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DoE	Design of experiment
dpm	Disintegrations per minute
DSF	Differential scanning fluorimetry
ECD	Extracellular domain
ECL	Extracellular loop
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
ENT	Equilibrative nucleoside transporter
EPR	Electron paramagnetic resonance
Fab	Antigen-binding fragment
FBDD	Fragment based drug design
FBS	Fetal bovine serum
FDA	U.S. Food and Drug Administration
g	Gravitational constant
GCGR	Glucagon receptor
GDP	Guanosine diphosphate
GEF	RhoGTPase nucleotide exchange factor
GIRK	G protein-coupled inwardly rectifying potassium
GLP1R	Glucagon-like peptide 1 receptor
GPCR	G protein-coupled receptor
GRK	G protein-coupled receptor kinase
GTP	Guanosine triphosphate
HEK293	Human embryonic kidney 293
HIS	Histidine
HRP	Horseradish peroxidase
HT	High throughput
IC ₅₀	Half maximal inhibitory concentration

ICL	Intracellular loop
IP ₃	Inositol triphosphate
K_D	Dissociation constant
K_i	Inhibitory dissociation constant
LB	Lysogeny broth
LCP	Lipidic cubic phase
LMNG	Lauryl maltose neopentyl glycol
MAG	Monoacylglycerol
MAO	Mechanism of action
MAPK	Mitogen-activated protein kinase
MD	Molecular dynamics
mGlu ₅	Metabotropic glutamate receptor 5
microED	Microcrystal electron diffraction
MOI	Multiplicity of infection
MS	Mass spectrometry
NAM	Negative allosteric modulator
nb35	Nanobody 35
NECA	5'-N-Ethylcarboxamidoadenosine
Ni-NTA	Nickel-nitrilotriacetic acid
NMR	Nuclear magnetic resonance
NR2B	Ionotropic receptor NMDA type 2B
PAGE	Polyacrylamide gel electrophoresis
PAM	Positive allosteric modulator
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein data bank
PEI	Polyethylamine
PH	Polyhedron
PIP ₂	Phosphatidylinositol 4,5-bisphosphate
PKA	Protein kinase A
PKC	Protein kinase C
PLC	Phospholipase C
PVDF	Polyvinylidene difluoride
QMEAN	Qualitative model energy analysis
QSAR	Quantitative structure-activity relationship

RAMPs	Receptor activity modifying proteins
RhoGEFs	RhoGTPase nucleotide exchange factors
RMSD	Root-mean-square deviation
RNA	Ribonucleic acid
RT	Room temperature
SAR	Structure-activity relationships
SBDD	Structure-based drug design
SD	Standard deviation
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
SEM	Standard error of the mean
<i>Sf9</i>	<i>Spodoptera frugiperda 9</i>
SFX	Serial femtosecond crystallography
SHG	Second harmonic generation
SMO	Smoothened receptor
SPR	Surface plasmon resonance
SSX	Serial synchrotron crystallography
T4L	T4-lysozyme
T _m	Melting temperature
TBST	Tris-buffered saline with Tween-20
TM	Transmembrane helix
TR-FRET	Time-resolved fluorescence resonance energy transfer
UV	Ultraviolet
vdW	Van der Waals
WT	Wild type
XFEL	X-ray free electron lasers

Chapter 1

Introduction

1.1 Target-based and structure-based drug discovery

One of the major challenges of today's society is the steadily increasing health care costs. Part of this problem is caused by the tremendously high costs and long timelines associated with research and development of novel drugs. The average cost of bringing a drug to market (including drug failures) amounts to about \$2.6 billion [4] and takes approximately 15 years [5]. Thus, there is a big interest from the private, as well as the public sector, to design drug discovery processes that are more rational and economically efficient.

Traditional drug discovery strategies (classical forward pharmacology) are based on phenotypical screening, in which cells or whole organisms are hit with large libraries of compounds, looking for a desired phenotypical effect. With the advances in molecular biology and genomics, the idea arose to use a defined molecular target, that is hypothesised to have an important role in disease, as starting point for a more rational drug discovery approach, promising to be a more specific, efficient and rapid process for lead discovery and optimisation. Thus, in contrast to the classical approach, target-based drug discovery (also known as reverse pharmacology) first requires the identification of macromolecular target (Figure 1.1) [6]. If the target is known and assayable, it can then be investigated by ligand-based approaches. In many ways this is similar to phenotypic screening approaches. Additionally, this method employs statistical approaches to correlate ligand activity to structural information [7]. Furthermore, if the structure of the target protein is known, structure-based drug design (SBDD) strategies can be employed to determine structure activity relationships (SAR), and with these predict biological activity from molecular structure [8]. Ligand- and structure-based approaches are highly synergetic and are therefore combined in most drug-discovery programs.

Most SBDD efforts target proteins; nonetheless, successful projects targeting structured ribonucleic acids (RNAs) have been reported as well [9, 10]. Drugs that are developed against a disease caused by a human pathogen are usually devised to completely inhibit an essential pathogenic function, leading to its death. In contrast, if a disease is caused by the malfunction of physiological processes in the human body, drugs are designed to modulate the function of

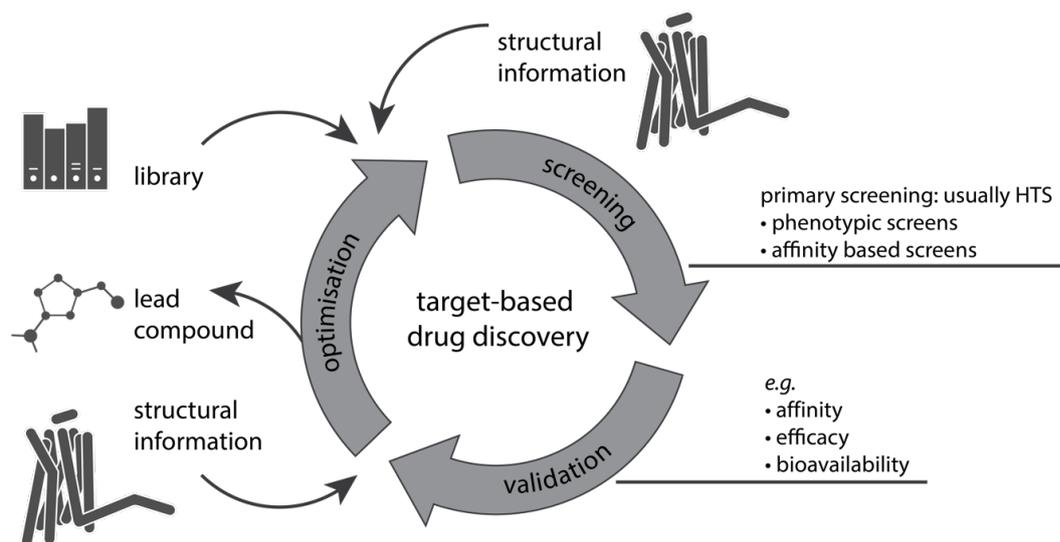


Figure 1.1: Target-based drug design process. Target-based drug design is a highly iterative process including many rounds of screening, validation, and optimisation. Structural information is an extremely valuable input in several stages of the cycle, from target and hit identification to compound optimisation.

the target, rather than total inhibition. A specially interesting family of human protein drug targets are G protein-coupled receptors (GPCRs), which are targeted by about 35% of all marketed drugs [11, 12].

A SBDD project requires a solved target structure. For a target where there is no structure available yet, and a ligand-based approach is not feasible or desirable, a SBDD project starts with its cloning, purification, and experimental structure determination, either by nuclear magnetic resonance (NMR), X-ray crystallography, or more recently cryogenic electron microscopy (cryo-EM). Alternatively, structure prediction methods such as homology modelling can be used [13, 13, 14]. In a second step, the ligand binding site needs to be determined. Ideally, this is a defined pocket with a variety of potential hydrogen bond donors and acceptors, such as found in the active site of enzymes. Allosteric binding sites and target sites for protein-protein interactions can be more difficult to locate, since these might not represent actual pockets, but might be rather flat, large and hydrophobic. Ligand-target complex structures, such as obtained from co-crystallisation studies, can be of great value to identify such a site.

Once the structure and target site are identified, there are different experimental, as well as computer aided strategies, to identify a first selection of binding molecules. These compounds are then further characterised with biochemical assays to describe their biological and pharmacological properties, such as efficacy, affinity, and potency to identify first lead compounds. Usually, a promising lead displays at least micromolar inhibition *in vitro* [15]. This is then followed by a second cycle of structure determination, now of the target in complex with the lead from first selection round, to reveal potential interaction sites that can be optimised to increase potency. A structure-based drug discovery process is highly iterative. It usually includes many

cycles of synthesis, biochemical and structural characterisation, and ligand optimisation, until binding and the specificity of the compound for the target are at a point at which animal and then clinical trials can be realised (Figure 1.1).

Besides improving affinity and specificity of the compound, there are several other factors which need to be considered when optimising a lead compound to ensure its drug-likeness. These are physicochemical properties *i.e.* the compounds molecular mass, the number of atoms, its partition coefficient, solubility, and the number of present hydrogen bond donors and acceptors that would make a lead compound a likely orally active drug in humans. Lipinski’s “rule of five” is one of the guidelines which is used to evaluate these properties during the drug discovery process [16].

1.1.1 Experimental methods for lead discovery

After having identified a protein target, interaction or affinity-based screens are applied to screen large numbers of diverse molecules in order to identify novel lead compounds for drug development. For the discovery of biomolecules (*e.g.* antibodies) display technologies such as phage- or yeast-display are a rapid and inexpensive strategy, providing protein-binding peptides or antibodies within weeks. These approaches are however, limited to the development of biotherapeutics and antibodies, as they rely on the natural transcription and translation machinery.

For the discovery of novel small molecule drugs, large numbers of chemically diverse compounds are screened, in an attempt to cover as much of the chemical space as possible. This requires huge, expensive compound libraries, which are tested one at a time in biochemical or cell-based signalling assays. To do this with high throughput (HT-screening), sophisticated robotic equipment is required, which is often not available in academia or smaller companies. Hits must be validated and can then be used as starting points in structure-activity cycles which aim to generate a lead compound.

An alternative approach to efficiently sample the chemical space, is to use smaller libraries with low molecular mass compounds (<250 Da, therefore also called fragments) of low complexity [17]. This results in significantly lower costs, broadening the user availability of fragment-based drug design (FBDD) compared to conventional HT-screening. FBDD is especially interesting for targets that lack a deep or well-defined ligand binding pocket and are also a great way to circumvent patents by discovering ligands of novel chemical classes [18]. However, fragments usually bind with a low affinity to the protein target, requiring highly sensitive biophysical detection approaches such as NMR and surface plasmon resonance (SPR). Using structural knowledge about the target-fragment interaction, these fragments can then be grown or linked into larger molecules that form starting points for classical SAR cycles.

1.1.1.1 DNA-encoded libraries

Deoxyribonucleic acid (DNA) encoded chemical libraries (DECLs or DELs) are collections of combinatorial fragment compounds, in which each chemical structure is tagged with a DNA identification barcode. These libraries were designed with the aim to combine the benefits from display technologies with synthetic methods for compound preparation for the discovery of small molecules. There are several approaches for DEL screening proposed [19, 20, 21, 22, 23], but

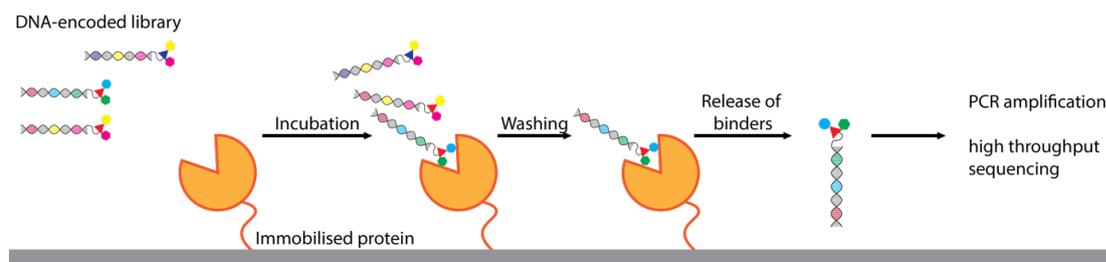


Figure 1.2: Principle of DEL screening. DELs are combinatorial libraries in which each compound is tagged with a unique DNA barcode. The coloured shapes represent different functional compound fragments. For DEL screening the library is incubated with a purified and immobilised protein target (orange packman shape). Non-binding compounds are washed off and binders can be released, *e.g.* by heating up the sample. The eluted library compounds are enriched by PCR amplification and identified by reading out their barcode ID with high throughput sequencing. Subsequently, they can be resynthesized and tested.

usually the target protein is isolated and immobilised and then incubated in the presence of the entire library. Subsequently, DNA conjugates with low target affinity are removed by several washing steps. The binders are identified by analysing the sequence enrichment, either by high-throughput DNA sequencing or using DNA microarray chips (Figure 1.2). Thus, compared to conventional HT-screening, the DNA barcodes of the DEL compounds allow screening of an entire library in one step, rather than one compound at a time, significantly speeding up the screening process and eliminating the need for HT-robotics [24].

DELs are combinatorial libraries, meaning they are prepared by assembling sets of building blocks in predefined geometries. Consequently, besides the actual chemistry of the different building blocks, the number of different sets of building blocks, synthetic cycles, and the assembly geometry are crucial parameters for the design of such a library. However, the selection of chemical reaction for library assembly needs to be considered. Ideally, reactions for DEL synthesis should be high-yielding, generate nearly no side products and, most importantly, need to be DNA-compatible, which limits the choice to only a fraction of organic reactions [24]. The big advantage of adopting a combinatorial approach is the exponential relationship between library size and the number of assembled sets of building blocks. However, adding sets of building blocks inevitably increases the molecular mass of the compounds, which according to Lipinski’s “rule of five” should be less than 500 Da [16]. Moreover, compounds with larger molecular masses tend to display higher polar surface areas or lipophilicity, which are also unfavourable characteristics for a drug’s pharmacokinetics. Therefore, all published libraries use no more than four different sets of building blocks [24]. The structural parameters of these libraries vary markedly, because the ideal compound geometry may vary depending on the protein target. A DEL designed to fit a deep narrow binding pocket might yield no hits for a target with a shallow interaction surface and *vice versa*. Structural information about the protein target is therefore essential for successful hit discovery.

1.1.2 Computer-aided methods for structure-based drug design

Apart from experimental approaches there are several computer-aided methods which are used to identify, but also to further optimise lead compounds.

The most common approach is virtual screening or docking of compound libraries. For this, each compound in the library is virtually docked into the binding site and is scored based on its predicted interaction with the target [25]. To this end a variety of different docking programs have been developed, *e.g.* AutoDock [26], DOCK [27], and Glide [28] to name only a few. To determine the predicted binding pose, they perform multiple iterations, sampling the conformational space until the minimum energy state is attained. After each step, scoring functions evaluate the quality of the current pose, guiding the search algorithm towards relevant ligand conformations. Besides pose prediction, these scoring functions also help to distinguish between active and non-active compounds and to predict binding affinities. Depending on the task and the program used, one or a combination of different scoring functions are utilised; thus, making it impossible to compare the scores between programs. However, most of the used scoring functions fall into one of three general categories [29].

(1) Force field-based scores are calculated based on the physical atomic interactions between the ligand and the target, such as electrostatic and van der Waals (vdW) interactions, as well as bond stretching, bending and torsion forces. To prevent a bias on Coulombic electrostatic interactions and therefore a selection of only highly charged compounds, solvent effects have to be considered as well. This can be done using techniques such as free energy perturbation and thermodynamic integration [30].

(2) Knowledge-based scoring functions are based on the statistical analysis of preferred protein-ligand atom pair interactions. These can directly be obtained from the occurrence frequency of atom pairs in databases such as the protein data bank (PDB). The derived potentials are used as predictors for the binding free energy of the complex. Even though this correlation is rather arbitrary, compared to other scoring functions, knowledge-based approaches circumvent the task of balancing many opposing contributions to binding and are therefore relatively fast [31].

(3) The last class are empirical scoring functions, which aim to reproduce experimental affinity data [32]. To develop this type of scoring function a training data set, consisting of lots of experimental protein-ligand complex structures, associated with the corresponding affinities, is required. First, essential interaction characteristics are numerically captured by defining a set of descriptors. These descriptors are often meant to represent interaction terms or energy contributions to the free binding energy, such as hydrogen bonds, hydrophobic contacts, or the loss of torsional degrees of freedom, and are derived from structural data with particular model assumptions [33]. However, it is also possible to use simpler and purely numerical measures, such as element dependent or atom-type dependent distance counts [34, 35]. Then the weight of the different descriptors and a relationship between the descriptors and the empirical affinity is established. This is classically done using multiple linear regression, but more recently machine-learning methods based on non-linear regression have become popular. Machine-learning approaches are particularly advantageous with large training sets and descriptor spaces and have significantly contributed to the recent advances in affinity-prediction performance [36].

Despite the broad variety of different scoring functions that have been developed over the last couple of years, the desired general reliability of the predictions has not yet been reached and the accuracy of each score is usually tied to a specific application. To take the advantages and balance the error rates of different scoring functions, consensus or composite scoring techniques are now widely used, employing collective scores, rather than to rely on a single scoring function, to increase the possibility of true positive selection [29, 37, 38].

Each docking experiment usually starts with the preparation of the protein structure and the ligands. This includes the addition of hydrogen atoms and assigning the protonation and tautomeric states of residues. Additionally, small molecules, such as ions and water molecules can be added to the structure, if it is known that they play structural roles that are crucial for the conformation of the target. As part of this protein pre-processing step, some docking programs such as Glide also represent the shape and properties of the target as different sets of fields on a grid, providing progressively more accurate scoring of the ligand pose [28]. The next step in the docking process is the generation of initial ligand conformations. To this end, Glide uses a series of hierarchical filters to search for possible locations of the ligand in the binding-site. Starting from these poses, the ligand is minimised in the field of the receptor using a molecular mechanics energy function. For the lowest-energy poses nearby torsion minima are examined to properly orient peripheral groups and to further optimise internal torsion angles. Thus, Glide is an example of a docking program that uses composite scoring functions [28].

Similar to the experimental fragment-based screening, compound fragments, such as benzene rings, carbonyl groups, amino groups etc., can also be positioned virtually into the binding site, in a process called *de novo* generation of compounds. In contrast to try and dock entire compounds, fragments have the advantage that the binding site can be explored more fully. The fragments are then scored and linked *in silico* to create druglike compounds. This allows the discovery of novel scaffolds for inhibitors [39]. However, the big disadvantage of this strategy is that these compounds need to be newly synthesised before they can be tested in the laboratory, which is not necessarily straight forward. Further, even though this process is automated along most steps of the route, some work still needs to be done manually, making it quite laborious.

The big advantage of virtual screening is that huge numbers of compounds can be screened without having to purchase or synthesise them all, this reduces costs and speeds up the hit discovery processes. However, when libraries of hundreds of thousands (or even millions) of compounds are screened, this also results in a considerable number of hits. Therefore, post-docking analysis is usually performed to prioritise the compounds. This generally involves the visualisation of the docked compound and looking for the presence of specific intermolecular interactions and other predetermined requirements [40]. Finally, each selected hit is brought into the wet lab for biochemical evaluation. Most of the time this includes affinity assays as well as a measure of compound efficacy and bioavailability. Promising leads re-enter the structural determination process. If a co-structure is obtained, this ligand can be used as a starting point for further modification, which aim to maximise complementary interactions in the target site and thereby develop the compound into an inhibitor. This process, also called inspection, does not need to be started with a virtual screening hit, but is also applicable if initial co-structures bound to *e.g.* the natural ligand, or if hits from other screening methods are available [41]. The newly designed compounds are again scored for binding using evaluative scoring algorithms.

When comparing the obtained co-structures of novel leads with their predicted binding modes there are only a few examples which show significant differences [42], but in most of the cases the docked and experimental conformations are highly similar, displaying a root-mean-square deviation (RMSD) of less than 2 Å [43]. Some of these observed differences might arise if protein and ligand flexibility are not taken into consideration sufficiently. Some enzymes and receptors can undergo large conformational changes upon ligand binding and in many cases a ligand stabilises a specific subset of conformations. However, often protein structures are only available in one specific conformation, limiting docking to this specific state. If the target is known to display a high degree of flexibility, molecular dynamics (MD) simulations can be used to produce alternative conformational states [44]. MD can also be useful for post-docking analysis to estimate the stability of a proposed ligand-receptor complex [45].

One of the major drawbacks of virtual screening is that many of the used libraries ignore protonation and tautomerism effects, as well as the ionisation state of the compounds. Therefore, significant hits can be overlooked [39]. Nonetheless, the identification of several inhibitors and antagonists demonstrates the potential of computer aided drug discovery strategies. The discovery of drugs targeting the human immunodeficiency virus [46], tuberculosis [47], diabetic neuropathy [48], and cancer [49] are some examples of success cases.

1.1.3 The use of protein structures in SBDD

SBDD has become an integral part of each drug discovery project and the three-dimensional (3D) structures of protein-ligand complexes are now routinely used in all parts of the drug discovery process. They provide detailed information about the intermolecular features of recognition and binding of a ligand, the analysis of different binding conformations, ligand-protein interactions, and the identification of unknown binding pockets. Further, in combination with computational methods, such as molecular dynamics, the elucidation of conformational changes resulting from ligand binding also give mechanistic insights. In combination with combinatorial chemistry, structural information also allows the design of focused compound libraries. As these libraries are designed with the specific target in mind, fewer compounds need to be screened and higher hit rates and discernible structure-activity relationships can be obtained, which facilitate the follow up of these hits.

To date, the most important methods for harnessing information about protein structure in SBDD are X-ray crystallography, NMR, homology modelling and recently also cryo-EM; X-ray crystallography still being the most dominant technique.

1.1.3.1 Structure determination of receptor-ligand complexes by X-ray crystallography for SBDD

Technical advances and the automation of many of the steps during the crystallisation and data collection process allow structure determination at high throughput along the entire drug discovery process, from the initial selection of the protein target, the identification of binding molecules, hit to lead optimisation, to the multi-threaded process by which the final properties of the clinical candidate are developed. Ideally, for each round of optimisation, a novel structure of the protein-ligand is solved. To this end, it is important to have a highly robust crystallisation system, to which a variety of different compounds can be added, without perturbing protein

crystallisation. This is especially challenging with proteins that require the stabilising effect of a ligand for crystallisation, *e.g.* many membrane proteins. Very often, exchanging this ligand for another compound with different properties, changes the crystallisation conditions or even leads to the protein not crystallising.

The most straight forward approach to get a structure of a receptor-ligand complex is co-purification of the receptor bound to the ligand of interest; however, this involves the addition of excess compound throughout the entire process, which requires increased materials and time resources. Especially in early drug discovery stages, most compounds are only synthesised on a milligram scale and thus, the required amounts of compound for co-purification might not be available.

A more resource-saving strategy is to perform ligand exchange on purified protein, at a point where volumes are already highly reduced. In contrast to when a low affinity ligand is exchanged for a higher affinity ligand during the purification process, where the low affinity ligand is diluted away, in this case it stays in solution, resulting in a ligand mixture. Thus, to maximise the success of the ligand exchange, the concentration of the novel ligand should be as high as possible. Using the law of mass action and the definition of occupancy, it can be estimated that near full occupancy ($\sim 91\%$) requires a ligand concentration of about ten times the dissociation constant (K_D) value. This might be more difficult to reach with low affinity compounds, especially if solubility is limited. Furthermore, when working with lower volumes, dimethyl sulfoxide (DMSO) concentration, which is commonly used to dissolve most compounds, might become problematic as well, as it gets less diluted out while trying to reach high ligand concentrations. However, high DMSO concentrations might hamper crystallisation.

Soaking of small molecule ligands into protein crystals has become one of the standard techniques for structure-based drug design, allowing high-throughput structure determination of protein-ligand complexes with reduced ligand consumption. However, the transfer of these techniques to membrane protein targets still remains a major challenge [50, 51, 52]. The primary caveat of soaking experiments is the need for a reliable crystallisation system, generating crystals with an accessible ligand-binding site, so added ligand can readily diffuse in. Classically this would use the *apo*-protein; however, often it is also possible to displace co-crystallised ligands with ligands of a similar or equal binding affinity [53]. Indeed, for targets such as GPCRs, where the *apo*- and ligand-bound state are distinct, co-crystal structures with similar ligands may be a better starting point, as the ligand induced conformational changes might otherwise disrupt crystal contacts. But even with a robust crystallisation system in hand, technical barriers in the experimental setup can be encountered when attempting soaking experiments with membrane proteins. One reason for this is that high resolution structure determination often relies on *in meso* crystallisation (further discussed in chapter 1.3.2) [54]. *In meso* crystallisation trials are typically set up using sandwich plates, squeezing the mesophase between two glass plates, which makes crystal harvesting difficult and time-consuming. Post crystallisation treatments such as soaking are basically impossible, as the top glass plate needs to be cut open to access the crystals. One method to circumvent this problem is to use plastic covers, which can be more easily cut open and resealed to apply the ligand of interest [50]. Another method employs the

use of *in-situ* plates, in which crystallisation occurs between two plastic films. Individual wells can then be cut out for post-crystallisation treatment [55, 56]. However, the manual handling of the crystals in both approaches still remains a major challenge [57].

SBDD typically requires structures at a resolution below 2.5 Å; however, more importantly, the structure should have a high data to parameter ratio and the placement of residues in the electron density map needs to be unambiguous [58]. As for any other interpretation of protein structures, the quality of a crystal structure should be carefully assessed before using it for SBDD. To this end, the standard quality indicators in crystallography are considered. The quality of the model at the ligand binding site can further be assessed by comparing the temperature factors (B-factor) of atoms in the region of interest to the average temperature factor of the molecule. A high B-factor can reflect disorder due to motion of the residue or ligand or be a general indication of error.

Most crystal structures contain additional information about the position of small molecules, such as ions or waters, which often play crucial roles in protein function. About 65% of the protein-ligand complexes solved by crystallography contain at least one water molecule that is involved in ligand-receptor recognition [59]. These molecules are often located deep inside the binding pocket mediating multiple hydrogen bonds between the ligand and the target and are therefore important to take into consideration for docking experiments. There are two main strategies on how to deal with these waters: they can either be removed from the structure and be replaced by the ligand mimicking the interaction network, or they can be considered as a part of the target and be included in the docking process, mitigating the formation of favourable hydrogen bonding networks [40]. Only strongly bound water molecules should be retained in the structure in SBDD and docking strategies. To discriminate them from displaceable molecules, the positions of structural waters from multiple crystal structures should be compared, whenever possible [60]. Additionally, there exist a variety of different algorithms which help to guide the decision of either keeping or removing a water molecule, *e.g.* free energy perturbation calculation using Monte Carlo statistical mechanics simulations [61] or the analysis of geometric parameters of the protein environment surrounding the water molecule [62].

1.1.3.2 Other methods for protein structure determination for SBDD

The second empirical strategy for structure determination in SBDD is NMR. Besides providing information about protein structure and dynamics, NMR spectroscopy is also a very valuable tool for FBDD [63], as it can reliably detect weak ligand binding (single digit millimolar affinity). In contrast to other biophysical techniques such as SPR, NMR has the advantage that not only the affinity but also the binding site and mode of a fragment can be determined as well. This can be especially interesting for the identification of novel allosteric binding sites.

Homology models

If no structure can be obtained experimentally, but the structure of a homologous protein is known, a homology model can be employed for drug design. There are multiple online recourses available, which generate and evaluate these models, one example being SWISS-MODEL¹. To evaluate a model, SWISS-MODEL calculates a confidence factor per residue,

¹The Centre for Molecular Life Sciences, University of Basel: <https://swissmodel.expasy.org>

based on the amount of structural information used, to create that part of the model. A lower confidence number reflects a higher number of templates and, therefore, an increased accuracy [64]. However, besides checking confidence scores, a homology model should always also be evaluated by looking at the stereochemical soundness of the structure, *i.e.* the Ramachandran plot, bond lengths and angles, planarity, and packing.

Cryo-electron microscopy

Cryo-EM emerged to one of the most popular techniques to solve protein structures, especially if working with larger complexes and more challenging targets such as membrane proteins. Even though there are still no reports about drug discovery campaigns guided by cryo-EM, the method has gained interest in SBDD, as several structures with small molecules bound have now been published (*e.g.* β -galactosidase in complex with an inhibitor or the glucagon-like peptide 1 receptor (GLP1R) bound to a peptide and a non-peptide agonist [65, 66]). Some of the big advantages of cryo-EM, over X-ray crystallography, are the reduced sample consumption, possibly allowing the study of proteins isolated from their native source, as well as the flexibility to change sample conditions, *e.g.* allowing a structure to be solved at its native pH, rather than the one at which it happens to crystallise. However, despite all the technical advances, grid preparation is still very laborious, making it difficult to implement HT-processes. This focuses the application of cryo-EM studies in SBDD on the target and hit identification phases and the understanding of the molecular mode of action of drug molecules, while at the moment they still are less suited for SAR guided hit optimisation purposes [67].

1.2 The family of G protein-coupled receptors as drug targets

G protein-coupled receptors (GPCRs) are the largest family of membrane proteins. They recognise a large variety of extracellular stimuli and mediate cellular responses to most hormones, metabolites, cytokines and neurotransmitters. As such, they are involved in basically every physiological process in the human body and are closely linked to human diseases, including Alzheimer's disease, cancer, diabetes, obesity, cardio vascular diseases as well as psychiatric disorders. Since GPCRs are placed at the cell surface, and are often cell or tissue-specific, they promise high druggability, as drugs do not have to be transported across the cell membrane. Therefore, they are the most successful class of drug targets [11]. In fact, mining different drug data bases revealed that about 34% of all drugs approved by the U.S. Food and Drug Administration (FDA) target GPCRs [12] and that GPCRs represent 19% of all targets for drugs approved in the five years between 2015–2020 [68].

Even though different GPCRs may vary a lot in function and structure, they share a general topology consisting of seven transmembrane helices (TMI–TMVII) connected by three extracellular loops (ECL1–ECL3) and three intracellular loops (ICL1–ICL3), an extracellular N-terminus and an intracellular C-terminus. While the structural architecture of the transmembrane domains is relatively well conserved, the other parts are highly diverse and account for the large variety in ligand recognition and interaction partners. The over 800 human GPCR genes can be grouped into five different classes according phylogenetic characteristics, called

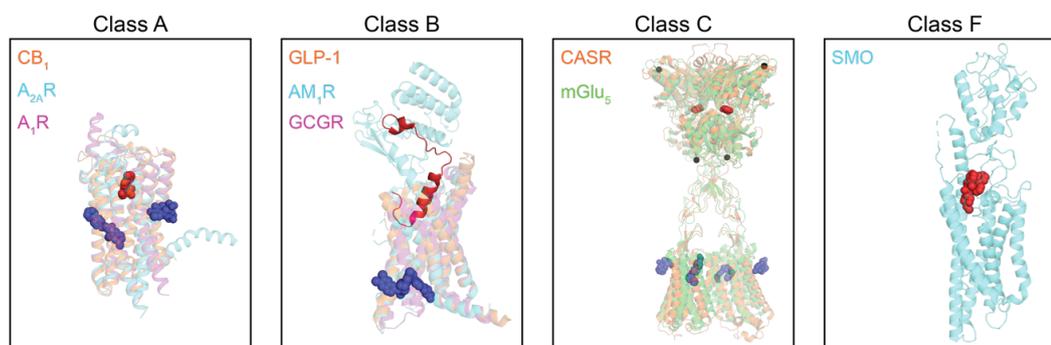


Figure 1.3: Different classes of human GPCR receptors and their ligand-binding sites.

Crystal or cryo-EM structures of representative GPCR-ligand complexes from classes A, B, C, and F, presenting diverse ligand-binding sites. For the class of adhesion GPCRs no experimental structural data is available so far. For class A the non-rhodopsin GPCRs adenosine A_{2A} receptor ($A_{2A}R$, PDB: 2YDO) is shown in complex with its endogenous ligand and cannabinoid receptor type 1 (CB_1 , PDB: 6KQI) and adenosine A_1 receptor (A_1R , PDB: 7LD3) are shown in complex with allosteric modulators. Class B is represented by the adrenomedullin 1 receptor (AM_1R PDB: 6UUN) in complex with the endogenous peptide hormone, the glucagon receptor ($GCGR$, PDB: 5EE7) in complex with a small molecule antagonist and the glucagon-like peptide-1 ($GLP-1$, PDB: 7EVM) with a bound allosteric modulator. Class C GPCRs form constitutive dimers. The cryo-EM structures of the calcium sensing receptor ($CASR$, PDB: 7M3G) bound to an amino acid, four calcium ions (black spheres) and an allosteric modulator as well as the metabotropic glutamate receptor 5 ($mGlu_5$, PDB: 7FD8) in complex with an allosteric modulator are shown. Finally, class F is represented by smoothed receptor (SMO , PDB: 6XBL) bound to a sterol. Receptors are shown in cartoon representation and the ligands are shown as spheres. Orthosteric ligands are highlighted in red while allosteric binders are coloured in blue.

Glutamate (also known as class C receptors), Rhodopsin-like (family A receptors), Adhesion, Frizzled/Taste2 (class F receptors) and Secretin (family B receptors) class (GRAFS system) [69]. Further subdivision is often made according to the native ligand recognised by the receptor. The class A receptor family forms the largest and most studied group of receptors, including chemokine and aminergic receptors, some peptide receptors such as the opioid and oxytocin receptors and nucleotide receptors such as adenosine receptors.

A feature that makes GPCRs great drug targets is that they often contain accessible, well-defined ligand binding pockets. However, the domains in recognising the endogenous ligand can vary a lot depending on the receptor (Figure 1.3). Glutamate receptors for example, are constitutive dimeric receptors with a large extracellular domain (ECD) containing a Venus flytrap module that is utilised for ligand recognition [70]. Similarly, in frizzled receptors the ligand is recognised by a cysteine-rich domain in the ECD [71]. In contrast, class A receptors usually have very short N-termini and ligands are recognised by a ligand-binding pocket in the transmembrane bundle. Finally, for class B GPCRs the ligand-binding site is formed by both extracellular and transmembrane domains [72].

1.2.1 Physiological GPCR functions and signalling

1.2.1.1 Canonical signalling pathways through G protein and arrestin coupling

Upon ligand binding, GPCRs activate at least one of several canonical and noncanonical signalling partners. Ligand binding induces a conformational change of the receptor that allows coupling of heterotrimeric G proteins on the intracellular side of the receptor. These G proteins consist of a G_α , a G_β and a G_γ subunit. Depending on the nucleotide bound, they are either in an active or an inactive state. The nucleotide binding pocket is located at the G_α subunit in between its two subdomains, the Ras-like domain and the α -helical domain. Interaction of the extreme C-terminus of the G_α Ras-like domain with a GPCR leads to nucleotide exchange of bound guanosine diphosphate (GDP) for guanosine triphosphate (GTP) and the activation of the G protein [73]. There are 16 genes in humans that encode α -subunits, which are classified into four functional families G_s , $G_{i/o}$, $G_{q/11}$ and $G_{12/13}$, activating different downstream signalling pathways. The C-terminus of the α -subunit is the main contributor for coupling specificity [74]. However, the molecular basis for this remains poorly understood and things have become even more complex, as it is now recognised that most GPCRs can couple, selectively, to members of multiple G protein families [75].

After nucleotide exchange the G protein is released from the receptor and the G_α and $G_{\beta\gamma}$ subunits dissociate, allowing the modulation of downstream signalling partners such as enzymes and channels that alter cellular levels of second messengers (Figure 1.4) [76, 77]. The activated G_s subunit stimulates the activity of adenylyl cyclase (AC), which is responsible for the formation of the second messenger cyclic adenosine monophosphate (cAMP). Conversely, $G_{i/o}$ activation leads to the inhibition of this same enzyme, reducing levels of the second messenger as well as the simulation of G protein-coupled inwardly rectifying potassium (GIRK) channels. Besides other effects, cAMP directly activates a number of proteins including protein kinase A (PKA) and ion channels, which successively regulate functional activities of cells via a variety of different pathways [78]. cAMP dependant pathways are a very common feature across many different GPCRs and cAMP levels are therefore often used as a readout for receptor activation.

$G_{q/11}$ is coupled to the activation of phospholipase C (PLC), a membrane-associated enzyme that hydrolyses phosphatidylinositol 4,5-bisphosphate (PIP_2) into inositol triphosphate (IP_3) and diacylglycerol (DAG). The produced IP_3 binds to receptors on the endoplasmic reticulum (ER), leading to the release of calcium ions from internal stores. Further, DAG facilitates the activation of protein kinase C (PKC). Both responses participate in diverse signalling to evoke different cellular events [79].

$G_{\alpha12}$ and $G_{\alpha13}$ directly activate RhoGTPase nucleotide exchange factors (RhoGEFs) to regulate the activity of the GTPase RhoA, which is involved in the moderation of a variety of cytoskeleton related processes, such as the formation of stress fibres [80].

Not only the G_α subunit, but also the $G_{\beta\gamma}$ heterodimer acts as signal transducer and activates a wide range of effector targets. Some of the first cryo-EM structures of G protein-coupled receptor complexes revealed that besides the G_α , the G_β subunit also forms some contacts with the receptor, but their importance is still not clear [81, 82, 83, 84]. Since the $G_{\beta\gamma}$ heterodimers do not undergo major conformational changes upon G protein activation, their activity is regulated by the G_α subunit. Dissociation of the trimeric G protein upon

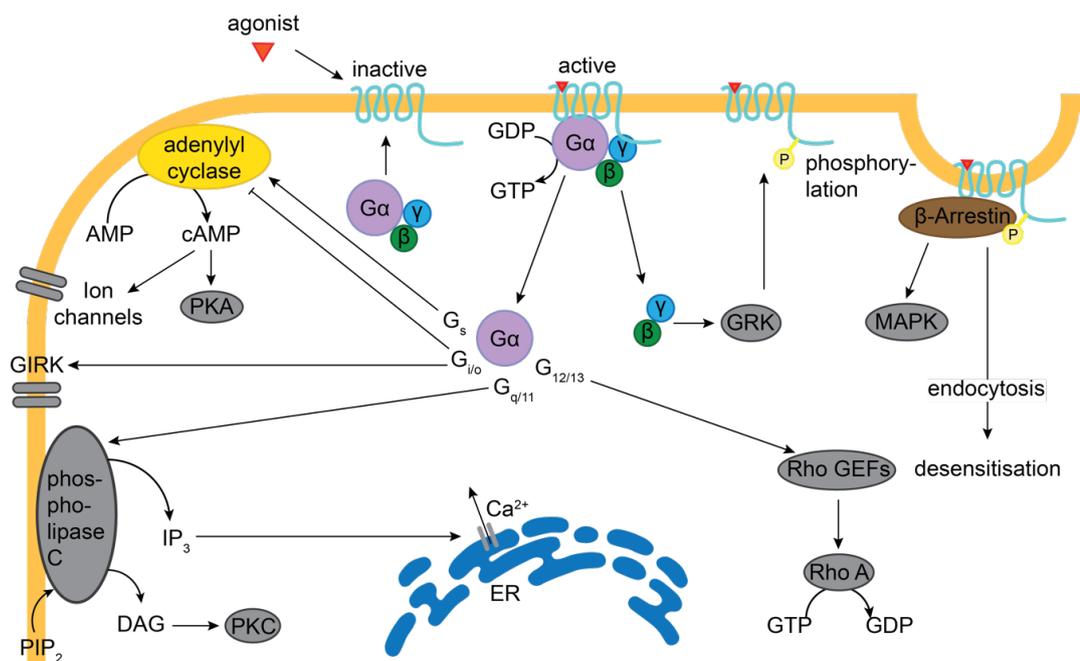


Figure 1.4: G protein-coupled receptor signalling. Canonical GPCR signalling occurs via coupling to heterotrimeric G proteins (G_{α} , G_{β} , and G_{γ}). Upon activation the G_{α} and $G_{\beta\gamma}$ subunits dissociate and initiate downstream signalling cascades. G_{α} proteins can be divided in four main families with different signalling profiles. G_s and $G_{i/o}$ activate and inhibit cAMP production, respectively. G_q signals through PLC, resulting in IP_3 and DAG production, and $G_{12/13}$ through Rho A mediated pathways. Activated receptor is phosphorylated, leading to the recruitment of β -arrestins and receptor internalisation.

activation exposes the $G_{\beta\gamma}$ protein-protein interface that can bind to downstream interaction partners. One of these targets is G protein-coupled receptor kinase 2 (GRK2), which functions in a negative feedback loop [85].

The agonist activated conformation of the receptor is phosphorylated by GRKs [86]. This allows the recruitment of β -arrestins to the receptor, which leads to receptor desensitization through two separate mechanisms. On the one hand, the β -arrestins compete with the G proteins for binding to the receptor and therefore prevent binding of subsequent G proteins, and on the other hand, they mediate clathrin dependent endocytosis, and thus remove the receptor from the plasma membrane and ligand exposure. However, further investigation of these signalling processes revealed several additional layers of complexity: β -arrestins can also mediate G protein independent signalling pathways, such as the mitogen-activated protein kinase (MAPKs) pathways [87, 88]. Certain ligands can bias the receptor and specifically activate one signalling pathway without the other, also referred to as biased signalling.

While G protein and β -arrestin interactions are common to most GPCRs, many of the receptors additionally interact with receptor specific interaction partners and modulating proteins, these include, but are not restricted to, receptor activity modifying proteins (RAMPs) and other GPCRs through the formation of homo- or heterodimers. This allows GPCRs to signal through a large selection of pathways.

1.2.1.2 Altering GPCR signalling output: Biased signalling

For a long time, GPCRs were presumed to exist in either an active or an inactive state, described through a simple two-state receptor model (“on” and “off”). However, extensive analysis of different signalling pathways and receptor structures revealed that they rather exist as an ensemble of many different conformations, each tailored to a specific response and downstream effects. It was further discovered that by stabilising distinct conformations, some GPCR ligands can specifically activate one or a few signalling pathways, while not affecting others. While this finding vastly increases the complexity of the system, it has become one of the most exciting research areas in the field, as it opens novel possibilities to selectively block the pathways implicated in disease. This phenomenon has since been described through several different terms including biased signalling, ligand directed signalling and functional selectivity and a variety of different types of biased ligands have been identified. Some ligands preferentially promote the activation of G proteins over β -arrestins (*i.e.* G protein-biased agonists), while others favour the stimulation of β -arrestins (*i.e.* β -arrestin-biased agonists) or a specific G protein subtype (*e.g.* G_i over G_s) over others [89].

While classical biased ligands bind to the orthosteric ligand pocket, functional selectivity can also be achieved through allosterically binding modulators. In fact, there is a growing body of evidence suggesting that the intricate network of GPCR signalling pathways is regulated by a series of different endogenous allosteric modulators, including non-small molecule interactions, such as homo- and hetero-receptor dimer formation or sodium ions [90, 91]. For the angiotensin II type I receptor it could even be shown that membrane stretch can bias the receptor towards β -arrestin signalling [92].

Biased ligands have become a major focus for drug discovery efforts, because of their potential to selectively stimulate therapeutically relevant pathways and avoid on-target side effects. Despite the fact that the physical basis for biased agonism is still not fully understood, there are some success stories where SBDD assisted the discovery of novel biased drugs. One of the most exemplary systems for biased signalling is the μ -opioid receptor, where G_i protein signalling is associated with the positive analgesic effects of morphine, while β -arrestin pathways are related to certain side effects such as respiratory depression and constipation associated with opioids [93, 94, 95]. Through virtual screening, a G_i biased non-morphinan agonist could be identified, which displayed nanomolar affinity against μ -opioid receptor, and significant pain relief for mice [96].

1.2.2 Regulating GPCR signalling with ligands and drugs

The primary endogenous ligand of a GPCR is a compound naturally produced by the body that activates the receptor upon binding and as such, initiates downstream signalling responses. As already briefly discussed in the previous section, additional physiological ligands can further regulate the signalling output of a receptor. In contrast, a drug is a substance of known chemical structure that is added exogenously to a system, to provoke specific physiological receptor responses.

Classical GPCR drugs either stimulate or inhibit receptor activity by directly interacting with the orthosteric binding site. Based on their activation intensity they can be divided into four major classes. Full receptor agonists can stimulate maximal GPCR activity, leading to

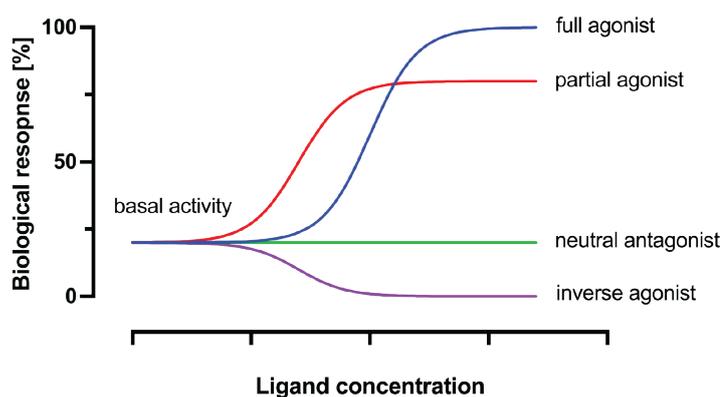


Figure 1.5: Effects of different ligands on the biological response of receptors. Antagonists and inverse agonists can only be differentiated if the basal activity of the receptor can be determined *e.g.* by using a constitutively active receptor system. Therefore, in this thesis, all inhibitors are referred to as antagonists.

the recruitment of downstream signal transducers (Figure 1.5). In contrast, partial agonists cannot activate the receptor to the same extent, only leading to a partial response, and if in the presence of a full agonist, even reducing the agonist response. Interestingly, numerous approved drugs are actually physiological agonists for GPCRs that stimulate receptor signalling, such as salmeterol targeting β_2 -adrenergic receptors (β_2 AR) as well as numerous peptide hormones [11]. Inhibitors act as agonist blockers and can be divided into neutral antagonists and inverse agonists. Neutral antagonists can bind to GPCRs without affecting the receptor's constitutive activity, whereas inverse agonists reduce the receptor response below its unstimulated levels (Figure 1.5). In practice, it is very difficult to differentiate between an antagonist and an inverse agonist, as a constitutively active receptor system is required to detect inverse agonism. Therefore, inverse agonists are often first described as antagonists and both terms might be used inconsistently to describe the same ligand. Thus, to avoid any confusion, this differentiation is not further considered throughout the rest of this thesis and all inhibitors are referred to as antagonists.

Despite being the most studied drug target, the differences in ligand binding requirements for agonists and antagonists of a given receptor are still not understood. Agonists and antagonists can have very similar structures, yet when binding to the receptor leading to different effects.

1.2.2.1 Allosteric ligands

Besides the orthosteric site, receptor activation can further be modulated indirectly by ligand binding to an allosteric pocket. Multiple allosteric sites are known to exist within a single receptor. They can be found throughout the entire receptor, on the extracellular as well as the intracellular site and also on the lipid face of transmembrane helices [97]. However, allosteric regulators are not restricted to exogenous molecules like natural products and synthetic compounds, but can also be of endogenous nature, including G proteins, β -arrestins, ions (*e.g.* Zn^{2+} , Na^+ , Cl^-), lipids, RAMPs, autoantibodies, membrane stretch, and other GPCRs through the formation of homo/heterodimers [98].

Allosteric modulators that enhance agonist binding and/or signalling responses are referred to as positive allosteric modulators (PAMs), while negative allosteric modulators (NAMs) decrease agonist binding and/or signalling efficacy. Additionally, there are some molecules that can bind to the orthosteric as well as an allosteric site, which are termed bitopic modulators [99].

Despite the increased interest in allosteric modulation, most of these evolutionarily unique sites remain largely unexplored. The binding sites for allosteric modulators are very diverse and often unique to a specific receptor, thus representing lucrative target sites for the development of more specific drugs with fewer side effects than orthosteric ligands. By screening a DEL-library using receptor with a ligand bound to the orthosteric binding site, the first NAM, and shortly after PAM, for the β_2 AR was discovered [100, 101]. Allosteric modulators have also become a focus of structure-based drug design. Using an optimised crystal-structure based model of the dopamine D₃ receptor, once in complex with and once without dopamine, a large-scale virtual screen enabled the identification of both orthosteric and allosteric ligands [102]. In another docking study non-competitive antagonists for two related class B GPCRs were identified [103].

1.3 Strategies to determine GPCR structures

Structural information is key for rational drug design. However, while membrane proteins represent more than half of all drug targets [104], the number of structures still lacks significantly behind soluble proteins. Because of their low expression levels, poor solubilisation efficacy, instability, and tendency to aggregate in detergent solubilised states, as well as their intrinsic conformational flexibility, GPCRs have historically been especially difficult targets for structural studies. The first near-atomic resolution (2.8 Å) structure of a GPCR, the one of bovine rhodopsin, was solved in the year 2000 [105]. However, rhodopsin is a very atypical GPCR, as it is exceptionally stable and highly abundant in natural sources. In contrast, the first crystal structure of a ligand-activated receptor could only be solved another seven years later [106, 107], requiring the development of complementary tools, strategies and technical developments, which will be discussed in subsequent chapters. Since then, 689 GPCR structures have been solved, including 120 unique receptor complexes. In the timeframe of this thesis, the number of unique receptor structures and receptor- ligand complexes more than doubled¹! This explosion of available GPCR structures lead to increased application of structure-based approaches in GPCR drug design. Nonetheless, so far, there are only a few examples in which structures determined between 2007–2016 have been used directly for drug discovery of clinical agents (reviewed in Congreve *et al.* [68]). One reason for this is that the drug development processes usually takes a very long time and many candidates might not have reached clinical trials yet. On the other hand, it's highly likely that many structures and results remain unpublished proprietary. Thus, the number of examples will inevitably expand over time.

¹<https://gpcrdb.org/structure/statistics>, the latest released structure being from 05.01.2022

1.3.1 Receptor stabilisation for structural studies

1.3.1.1 Ligand-induced receptor stabilisation

One of the reasons why GPCRs are hard to work with is that they are highly dynamic. By complementing structural data with biophysical techniques such as NMR, electron paramagnetic resonance (EPR) and fluorescence spectroscopy, supported by MD simulations, it became evident that GPCRs spontaneously sample a variety of different, functionally relevant conformational states [108, 109, 110, 111]. Changes between major conformational states can result in movements of up to 10 Å of the transmembrane helices [112]. The more flexible parts of the receptor such as the intracellular loops and the N- and C-termini can move even more and undergo substantial fluctuation even within a certain conformational state. This behaviour is very detrimental for any crystallisation attempts. However, ligands alter the fraction of time the GPCR spends in each conformational state and can thereby shift the equilibrium to a preferential conformational state [113, 112, 114]. This increases the likelihood of successful crystallisation. Indeed, excluding bovine opsin^{3CAP}, no crystal-structure of a native receptor could be solved in the absence of ligand [115]. In all other cases, the receptor needed to be thermostabilised, as demonstrated with the crystal structures of oligomerised β_1 -adrenergic receptor (β_1 AR) or endothelin ET_B receptor, to be solved without a ligand [116, 117]. However, the recent developments in cryo-EM allowed further advancement of structure determination of *apo*-receptors. In 2021 twelve *apo*-structures of eleven different receptors could be solved, including one family B receptor [118], several metabotropic glutamate [119, 120] and chemokine receptors [119, 121] as well as one frizzled receptor [122], most in complex with signalling proteins. Nonetheless, for the majority of solved structures, ligand-stabilisation of a specific conformational state was required to enable structure determination. Furthermore, for many receptors only stabilisation with a high-affinity ligand yielded diffracting crystals [123, 124, 48] and complexes with lower affinity ligands, often including the primary endogenous agonist, could only be obtained through additional thermostabilisation of the receptor, *e.g.* by site directed mutagenesis or additional conformational selection through the addition of intracellular binding partners [125, 126].

1.3.1.2 Receptor stabilisation by site directed mutagenesis and truncations

Site directed mutagenesis

Site directed mutagenesis can be used to introduce point mutations that help to increase expression levels [127, 128] and/or thermostability of the receptor. Such mutations can be rationally engineered by introducing intramolecular cysteine or salt bridges [129, 130] or mutations that stabilise a transmembrane helix interface [131]. Very often the design of these mutations can be inspired by comparing the receptor sequence with more stable, closely related receptors with known structure. Another approach is systematic alanine-scanning mutagenesis where each amino acid residue is mutated into an alanine. By screening these mutants in presence of a specific ligand, the receptors can be stabilised in a specific conformation (Figure 1.6). These stabilised receptors are also known as StaRsTM. The StaR approach could be generalised across GPCRs from different structural types and was largely applied in industrial drug discovery pipelines [132].

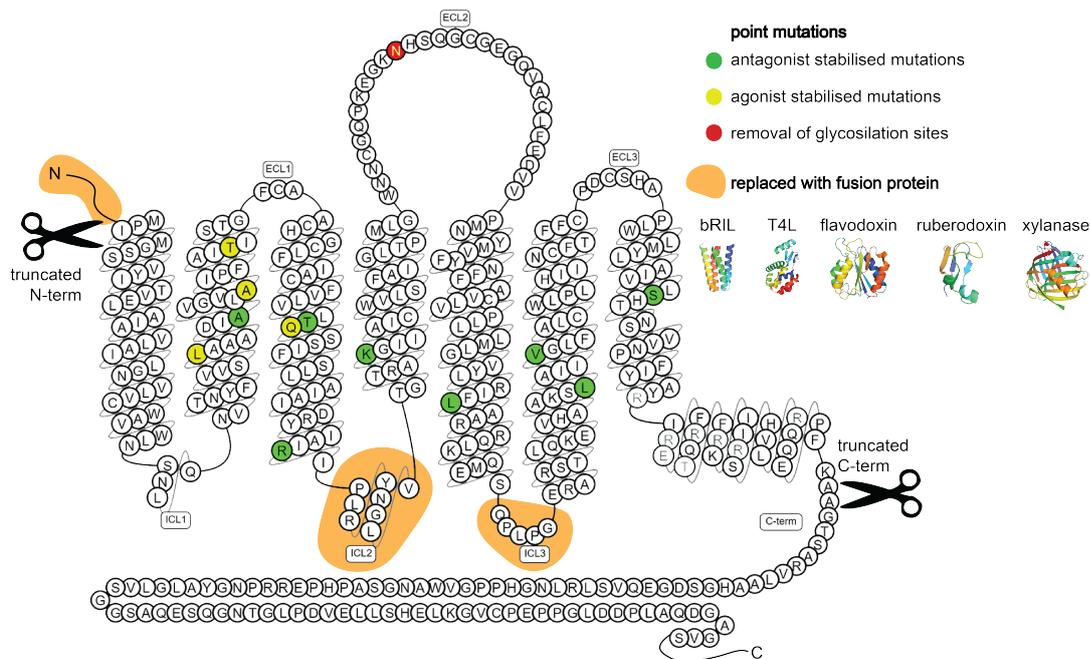


Figure 1.6: Stabilising receptor modifications for structural studies of GPCRs. Exemplary snake-plot of the $A_{2A}R$ showing regions with stabilising modifications. Residues mutated to alanine for thermostabilisation in either the agonist- or antagonist-bound conformation are highlighted in yellow and green respectively. Asparagine removed to prevent receptor glycosylation is shown in red. The most common regions where fusion proteins are inserted are shown in orange.

Truncations

Flexible parts of the GPCR usually impair crystallisation. Luckily, they can be predicted and are usually removed or replaced (Figure 1.6). Class B, C and F receptors possess large and flexible N-terminal extracellular domains, which are commonly removed and sometimes crystallised separately [133]. Also, the C-terminus displays a high degree flexibility, consequently significant parts of the C-terminus are usually truncated. In contrast to crystallography, cryo-EM is able to handle a certain degree of sample heterogeneity and inherent flexibility and might allow to resolve several distinct conformational states from a single sample [134, 135]. Thus, conformational selection and the removal of all flexible parts of the receptor is less crucial. Furthermore, the sample is usually frozen straight after purification and does not have to withstand several days in crystallisation conditions. Nevertheless, protein quality remains important for high-resolution structure determination and many of these stabilising truncations are also beneficial for cryo-EM studies.

1.3.1.3 Applications of fusion proteins for GPCR stabilisation and structure determination

The ICL3 of class A receptors is often unstructured and is therefore usually shortened or replaced by a fusion protein (Figure 1.6). In addition to conformationally stabilising the receptor by the reduction of molecular flexibility, they increase the hydrophilic surface of the receptor,

which is very helpful for the formation of crystal contacts [136]. Fusion proteins are usually small, single domain proteins. The first GPCR structure solved using a fusion protein was the human β_2 AR with a cysteine-free T4-lysozyme (T4L) mutant [106]. Since then, introduction of fusion proteins became one of the most successful strategies for GPCR structure elucidation. Insertion of T4L has proven fruitful, not only at ICL 3, but also at the N-terminus, or in some cases into ICL 2 [137, 138] (Figure 1.6). Further fusion proteins such as rubredoxin, flavodoxin, xylanase and the thermostabilised apocytochrome b562RIL (bRIL) protein have also been utilized. Fusions with the bRIL protein actually resulted in the highest resolution GPCR structures solved to date, A_{2A}AR with bRIL at ICL 3 at 1.7 Å [139, 140] and human δ -receptor with an N-terminal bRIL at 1.8 Å [91].

It is important to bear in mind that placing fusion proteins into intracellular loops might interfere with native signalling interactions. Receptors with a fusion protein inserted into ICL 3 are *e.g.* not able to interact with the G protein anymore, due to steric hindrance. Nevertheless, fusion proteins have also been very useful for the structure determination of GPCR complexes, especially of class A GPCRs that expose very little polar surface on the extracellular side. Detergents used for the solubilisation of GPCR complexes form very large micelles, making this area even less accessible to the formation of crystal contacts. Consequently, for the structure of β_2 AR-G_s complex as well as for the rhodopsin-arrestin structure, an N-terminal T4L-fusion was used to increase the polar surface for crystallisation [141, 142]. In cryo-EM, an N-terminal fusion protein can be helpful too, as, especially in complex with synthetic antibodies, they can serve as fiducial marker for receptor orientation and help to overcome size limitations [143, 144, 145].

1.3.1.4 The use of receptor stabilising antibodies

Another approach used for the stabilisation of GPCRs is the addition of antibodies or nanobodies as “crystallisation chaperones”. They can stabilise the receptor in a specific conformation and increase the crystallisation surface. A monoclonal antigen binding fragment (Fab fragment) was used to initially crystalize β_2 AR by specifically binding and thereby stabilising ICL 3 [107]. Nanobodies are a type of antibody produced by llamas and other camelids that are devoid of light chains. These single-domain antigen binding fragments are much smaller than Fab fragments (15kDa, approximately 25% the size of a Fab fragment) and can be easily cloned and expressed in *E. coli* [146]. It is possible to select for nanobodies that preferentially bind to distinct GPCR conformations using ligand-bound receptors. By immunizing llamas with purified agonist bound receptor, it was possible to generate a G protein mimetic nanobody that binds to the G protein interaction site and stabilises the receptor in its activated conformation [141]. The nanobody increased agonist affinity almost to the same extent as the full G_s protein and it is even possible to enhance the affinity of a nanobody by directed evolution, while maintaining its conformational selectivity [126]. The usage of such a high-affinity nanobody allowed high-resolution structure determination of the β_2 AR active state and also determination of a β_2 AR structure bound to an endogenous low-affinity agonist [126]. While G protein binding is abolished when using a receptor with a T4L-fusion, the nanobody can still bind to the receptor [141]. Nanobodies are thus a great surrogate to simplify crystallisation of active state GPCRs.

Furthermore, nanobodies have become a key tool to stabilise GPCR-G protein complexes. Nanobody 35 (Nb35) was developed to prevent the dissociation of the G protein by GTP γ S and is now routinely used for structure determination of ternary agonist-GPCR-G protein complexes [141, 119].

1.3.2 Crystallisation and structure determination of GPCRs

Several GPCRs have been crystallised in a micelle-like environment using direct crystallisation strategies such as vapour diffusion (also referred to as *in surfo* crystallisation). However, the most used system for GPCR crystallisation is lipidic cubic phase (LCP), which provides a more membrane-like environment (also referred to as *in meso* crystallisation) (Figure 1.7) (reviewed in Caffrey, [54]). LCP consists of a single, highly curved lipidic bilayer that separates two continuous, non-contacting aqueous channels. Transmembrane domains are inserted into the lipidic bilayer while soluble domains extend into the water channels. LCP is formed by mixing a lipid component with a water solution containing the protein of interest in a defined ratio (usually 3:2) at 20°C. Another advantage of LCP crystallisation is that it tolerates higher amounts of protein impurities compared to *in surfo* crystallisation [147]. In contrast to *in surfo* crystallisation, the size of the detergent micelles does not influence crystallisation success, but high detergent concentration can prevent LCP formation [148]. Monoacylglycerols (MAGs) such as monoolein (MAG 9.9), are most commonly used as the lipidic component; usually mixed in a 9:1 ratio with cholesterol. MAGs are comprised of a glycerol head, esterified with a cis-monosaturated fatty acid. The length of the fatty acid may be varied in order to change the properties of the resulting LCP. To crystallise the complexes of β_2 AR-G_s and rhodopsin-arrestin, a novel lipid, MAG 7.7, which contains a shorter fatty acid chain and therefore forms larger water channels, was used to better accommodate the large complexes [107, 142].

Despite the many successes archived through LCP crystallisation, and further development of the method, the process remains highly laborious, especially extracting crystals from the viscous LCP is extremely difficult, and it is not compatible with automation for high-throughput studies. Further, a common problem is that many crystallisation efforts only yield rather small crystals that diffract poorly. Nonetheless, small crystals can still be useful if data collection is performed at different sources. Serial femtosecond X-ray crystallography (SFX) using X-ray free electron lasers (XFEL) is a recent development in protein crystallography that allows high diffraction data collection from crystals of micrometre to nanometre size (Figure 1.7) [149]. The XFEL beam delivers a high density of short X-ray laser pulses that are shorter than fifty femtosecond duration, allowing to terminate protein diffraction before significant radiation damage occurs. This enables data collection with reduced radiation damage using a dose higher than usually used for cryogenically cooled crystals. In the LCP-SFX method, a stream of gel-like LCP, with fully hydrated microcrystals, runs continuously in vacuum across the XFEL beam [150, 151, 152].

Another, even more recent possibility is to investigate small LCP-grown crystals by micro-crystal electron diffraction (MicroED) [153]. To this end, the gel-like cubic phase needs to be first converted into a more liquid-like sponge phase, so it can be deposited onto a transmission electron microscope grid. Using focused ion-beam milling the crystals are reduced to thin lamellae that are suitable for MicroED data collection and structure determination.

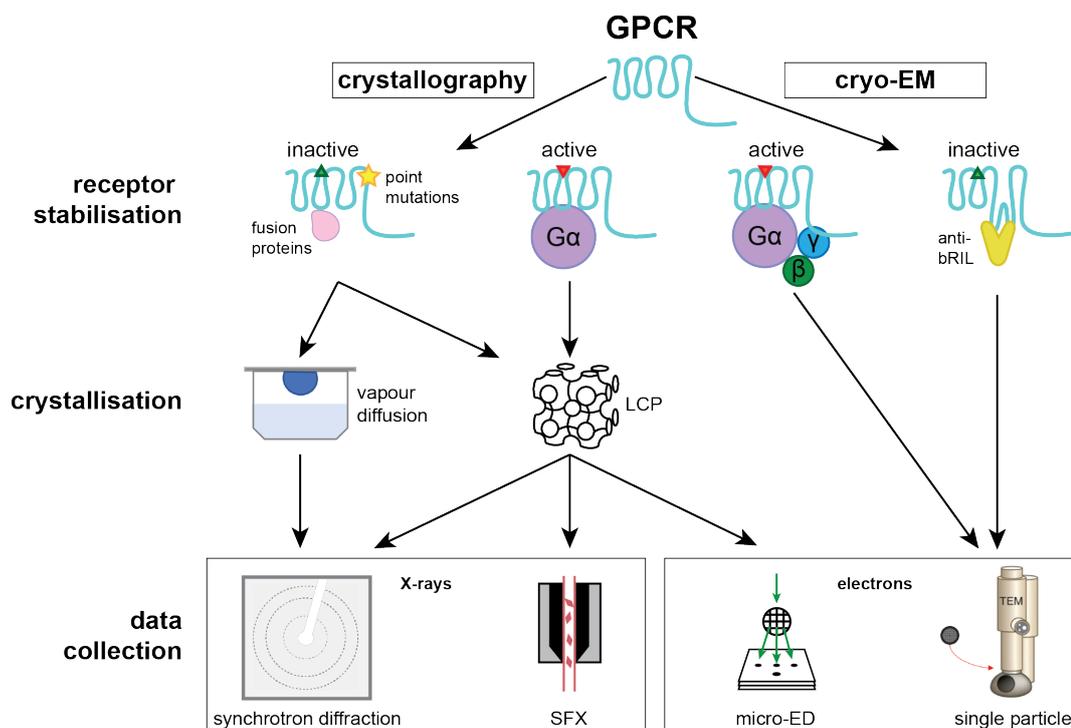


Figure 1.7: Strategies for GPCR structure determination. The most popular strategies for GPCR structure determination are X-ray crystallography and cryo-EM. Inactive state GPCR structures have mostly been solved by crystallographic means, either by *in meso*- (LCP) or classical *in surfo*- (vapour diffusion) crystallisation. To this end, the receptor usually requires to be additionally stabilised by ligands, fusion-proteins, and/or thermostabilising point mutations. Diffraction data is most commonly collected at synchrotron sources. However, more recently methods have been introduced that enable data collection from nano-crystals in LCP, either through serial femtosecond crystallography (SFX) or using electron diffraction (micro-ED). Active state structures require the presence of an intracellular transducer protein. As a result, engineered, minimal G_{α} proteins or conformation-specific nanobodies were developed. Structures of complexes with the full trimeric G protein (G_{α} , G_{β} , and G_{γ}) are very hard to crystallise, but are now routinely solved by single particle cryo-EM. Furthermore, novel strategies are established to also enable structure determination of smaller, inactive state GPCRs by cryo-EM, employing universal antibodies (e.g. anti-bRIL Fabs).

1.3.3 Cryo-EM structural approaches for GPCR drug discovery

At the onset of this project, the main method to obtain high-resolution structural information of GPCRs was X-ray crystallography. Despite being able to solve over 400 structures by crystallography, less than a quarter of these structures represent an active (or active-like) state of the receptor, almost exclusively restricted to the rhodopsin-like family (most of them actually being of visual rhodopsin). To date, the β_2 AR- G_s complex remains the only structure solved by crystallography of a fully-active receptor in complex with its G protein transducer [107]. However, the advent of high-resolution structure determination by cryo-EM has led to an explosion of novel active state structures and, since 2019, cryo-EM has overtaken X-ray crystallography for structure determination of membrane proteins (Figure 1.7) [154]. This development has been enabled mainly by improvements in electron detection and processing software, as well as the introduction of Volta phase plates [155, 156]. A big advantage of cryo-EM is that it

requires much less protein than for crystallisation; while only 0.3 mg/mL calcitonin- G_s complex were required for structure determination with cryo-EM [81], the concentration of the β_2AR-G_s complex in LCP was approximately 25 mg/mL [107]. Further, cryo-EM allows structure determination of more native-like, and even of full-length wild type receptors, in complex with a variety of signalling partners and in more native environments such as lipid nanodiscs [157, 158, 159].

While the resolution of the first couple of cryo-EM receptor structures were limited to approximately 3.3–4.1 Å, resolutions below 3 Å that can support accurate modelling of drug-receptor interaction have become more common (in one case even reaching a resolution below 2 Å [160]). However, it also has to be noted that in cryo-EM structures, density maps often show large variations in resolution throughout the receptor, often being worse around the orthosteric binding pocket [68]. Furthermore, data processing of cryo-EM micrographs still requires lots of computing power and may take weeks until the structure is solved, making the technique less amenable for routine use in commercial GPCR drug discovery and development. Conversely, in X-ray structures, the map quality is often more uniform. Once a receptor structure has been solved, new structures bound to different ligands can be rapidly determined from the diffraction data using molecular replacement. Often the crystals can grow under similar conditions and in some cases, it is even possible to soak crystals of GPCRs and determine multiple structures from one crystallisation experiment, further speeding the process up [50].

Another barrier for the routine use of cryo-EM in industrial SBDD are technology patents around tools for stabilisation of GPCR-G protein complexes, *e.g.* the previously mentioned Nb35, and the high cost of access for state-of-the-art microscope time. Therefore, recent efforts focused on developing sustainable workarounds by developing novel dominant negative G_s proteins that do not require the additional stabilisation by nanobodies [119].

Finally, there is still a lower size limit for cryo-EM structure determination of membrane proteins, which currently lays at around 60 kDa. It is therefore still unrealistic to determine the inactive state structure of a single, small 35–60 kDa GPCR. Thus, most structures are currently determined of receptor-transducer complexes, dimers or receptors with large extracellular domains. The smallest GPCR complex solved so far was of the calcitonin gene-related peptide receptor in complex with RAMP1, which amounts to 73 kDa of ordered structure [118].

One approach to overcome this limitation is by binding specific synthetic Fab fragments of a molecular mass of about 50 kDa or two nanobodies (each ~ 18 kDa in molecular mass) to increase the size of a GPCR in the inactive state. These antibodies can further be used as fiducial markers that help to identify the orientation of the particles on the grid. To circumvent having to develop specific Fabs for each novel target, synthetic antibodies against the commonly used bRIL-fusion protein have been developed, which can be used with most engineered receptor constructs [161, 144, 162]. However, this approach requires the antibody to bind the receptor in a single, specific conformation to not introduce additional inhomogeneity in the system.

1.4 The Adenosine A_{2A} Receptor

1.4.1 Adenosine signalling

Adenosine is a major metabolite and signalling nucleoside that is ubiquitously produced in almost all of the cells in the human body and has been shown to play a regulatory role in virtually every organ system studied. Endogenous purine adenosine is formed by the attachment of adenine to a ribose sugar and represents an important building block in the biogenesis of nucleic acids and 5'-adenosine triphosphate (ATP), the energy currency of the cell. Extracellularly, adenosine functions as a signalling molecule, exerting a range of cytoprotecting responses that are intended to preserve and restore tissue and organ homeostasis. Extracellular adenosine is mainly generated by the enzymatic break-down of ATP, which is released into the extracellular space in response to metabolic stress, such as hypoxia or energy depletion, tissue injury and inflammation (Figure 1.8). This can happen through a variety of different mechanisms, including transporter- or channel-mediated release, active vesicular exocytosis and direct release through mechanical stress or cell destruction [163]. The released ATP functions as a cellular danger signal that promotes innate and adaptive immune responses. In the extracellular space ATP is progressively dephosphorylated by ectonucleotidases such as CD39 and CD73, resulting in the formation of adenosine [164]. In contrast to the stimulatory properties of ATP, extracellular adenosine exerts counteractive, dampening effects, *e.g.* suppressing immune responses [165]. Besides this ATP-pathway, adenosine levels are further regulated through degradation by adenosine deaminases (ADAs) and reuptake into the cell through equilibrative nucleoside transporters (ENTs) (Figure 1.8). Under normal conditions extracellular levels are in the nanomolar range; however, upon cellular stress they can quickly rise to millimolar concentrations. As such, adenosine exerts pleiotropic functions throughout the body and is involved in the pathology of various diseases: in the central nervous system (CNS) it is involved in the modulation of neurotransmitter release, mediates depressant, anticonvulsant, sleep-promoting, antinociceptive effects, and is linked to neurodegenerative diseases [166, 167, 168]. In the cardiovascular system adenosine is associated with vasoconstriction, as well as vasodilation of veins and arteries [169]. Adenosine signalling further promotes a beneficial response in tissue injury, plays a role in controlling insulin responsiveness in adipose tissue, muscle and liver [170], and is involved in cancer growth and immune evasion [171].

The main effects of extracellular adenosine are mediated by adenosine receptors (ARs), which are located at the plasma membrane of cells and belong to the superfamily of G protein-coupled receptors. There are four subtypes of human adenosine receptors: A₁R, A_{2A}R, A_{2B}R and A₃R, classified according to sequence homology and their differential G protein coupling [172]. The A₁R and A₃R share 49% sequence identity and mainly couple to G_{i/o} proteins, thus, activation leads to AC inhibition and a reduction in cAMP levels. In contrast, A_{2A}R and A_{2B}R are almost 59% identical and are G_{s/olf} coupled. Activation of these subtypes leads to an increase in cAMP production and activation of PKA. A_{2B}R and A₃R are additionally reported to couple to G_q proteins, regulating the activity of PLC [173, 174]. While A₁R and A_{2A}R have a high affinity for adenosine, the A_{2B}R and A₃R show relatively low affinity (Table 1.1).

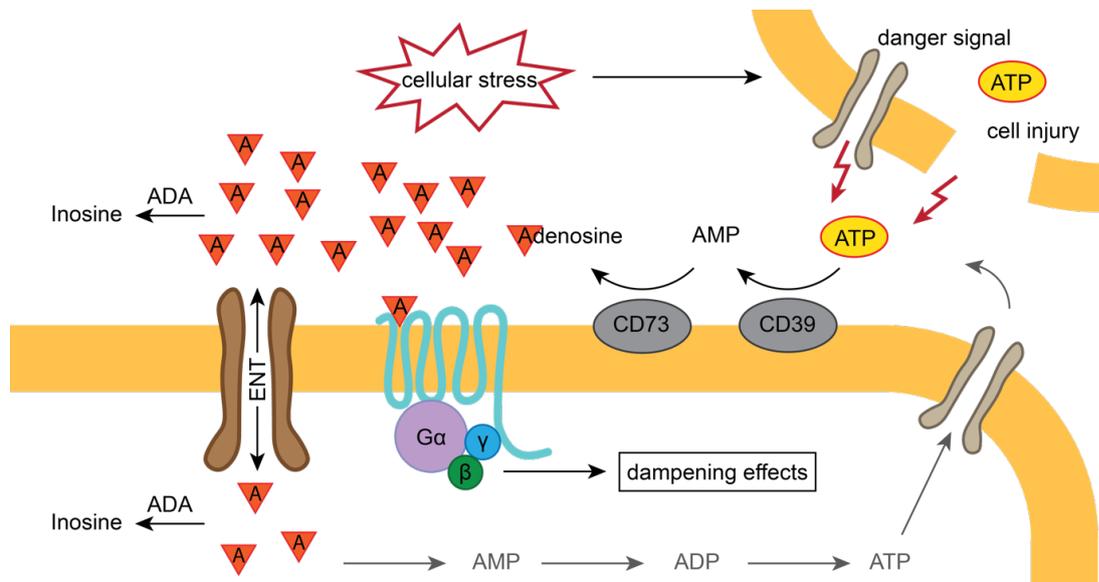


Figure 1.8: Adenosine signalling. Upon stress, cells release ATP into the extracellular space through a variety of different mechanisms. Extracellular ATP is hydrolysed by CD39 and CD73 to adenosine. Adenosine functions as signalling molecule through interaction with ARs and can further be metabolised by ADA. ENTs mediate the reuptake of adenosine into the cell.

ARs are expressed in almost every tissue; however, different subtypes are expressed in a cell and tissue specific manner. A_1R is expressed predominantly in neurons and glia cells in the CNS, spinal cord, testis, kidney and adipose tissues [175]. The majority of $A_{2A}R$ s are distributed in the liver, heart, lung, cells of the immune system, and CNS, primarily in neuronal cells in the striatum, olfactory tubercle, and the nucleus accumbens [176, 177]. The most abundant expression of A_3R is found in the lung and liver [178]. Finally, $A_{2B}R$ is distributed in the CNS and peripheral organs such as the bowel, bladder, lung and the excretory duct of the testis [179].

Because of their central role in adenosine signalling, ARs have become a main focus for targeted therapies in numerous human diseases. Despite the vast interest from both academia and industry, there is a strikingly low success rate of drug candidates reaching the market. To date, there are only two selective ligands that gained FDA approval: an $A_{2A}R$ agonist (Regadenoson) is used as coronary vasodilator in cardiac imaging [180] and an antagonist (Istradefylline), targeting the same receptor, is used in combination with another drug for the treatment of Parkinson's disease [181].

1.4.2 Efforts targeting $A_{2A}R$ signalling as therapy in human diseases

The $A_{2A}R$ is probably the best characterised out of the four human subtypes. For the pharmaceutical industry it first became of interest because of its key role in inflammation, where $A_{2A}R$ agonists showed anti-inflammatory effects, ultimately limiting tissue damage [169]. These cytoprotective effects were predominantly studied in pulmonary tissue for the treatment of respiratory diseases such as asthma. However, efforts to develop specific agonists remained fruitless, as adverse effects on the cardiovascular system, *e.g.* causing hypertension, could not be prevented [182].

Table 1.1: Comparison of the different adenosine receptor subtypes. List is not exhaustive but highlights main examples [169].

	A ₁ R	A _{2A} R	A _{2BR}	A _{3R}
G protein coupling	G _{i/o}	G _{s/olf}	G _s (G _q)	G _{i/o} (G _q)
Adenosine affinity	high	high	low	low
Signalling upon activation	↓cAMP, ↓PKA, ↑PLC, ↑MAPK, ↑Ca ₂ ⁺ , ↑PKC, can directly couple to cardiac K ⁺ and voltage gated Ca ₂ ⁺ channels	↑cAMP, ↑PKA, ↑PKC	↑cAMP, ↑PKA, ↑PLC, ↑Ca ₂ ⁺	↓cAMP, ↓PKA, →PLC, ↑Ca ₂ ⁺ , →modulation of PKC activity
Expression	CNS, spinal cord, testis, kidney and adipose tissues	liver, heart, lung, cells of the immune system, and CNS	CNS and peripheral organs such as the bowel, bladder, lung and the excretory duct of the testis	lung and liver
Function in human body	protects heart by downgrading ionotropic effects, modulation of neurotransmitter release, inflammation, circadian clock	coronary vasodilatation, decrease dopaminergic activity in CNS, Inhibition of central neuron excitation, promoting angiogenesis, sleep	role in inflammation, immunity and tissue repair, bone metabolism, wound healing, glucose homeostasis	cardiac muscle relaxation, smooth muscle contraction, inhibition of neutrophil degranulation, inflammation, regulation of cell cycle
Instigated as drug target for	neuropathic pain, respiratory diseases (asthma), arrhythmia, Huntington's disease, schizophrenia, diabetes	immuno-therapy for cancer, arthritis, asthma, heart failure, depression and anxiety, Parkinson's disease, drug addiction	cancer, diabetes, atherosclerosis, hypercholesterolemia	asthma, lung fibrosis, cancer, stroke, epilepsy, diabetic kidney diseases, angina

When it was discovered that the $A_{2A}R$ forms heterodimers with the dopamine D_2 receptor (D_2R), modulating dopaminergic responses, the research interest shifted towards its role in the CNS [183]. $A_{2A}R$ antagonists were shown to enhance D_2R -dependant signalling and were therefore further perused as an adjuvant treatment in Parkinson's disease [184], recently resulting in the approval of the first-in-class drug Istradefylline [185]. Besides directly modulating D_2R -signalling, further interactions with metabotropic glutamate receptors and cannabinoid receptors were discovered. The ability of the $A_{2A}R$ to regulate neurotransmission suggests that $A_{2A}R$ antagonists might find utility in the treatment of other neurological indications such as attention deficit hyperactivity disorder and pain [186, 187].

In recent years the $A_{2A}R$ has attracted attention for its role in immune-oncology. Cancer cells have developed a series of different strategies to evade being attacked by the immune system. One of these mechanisms is the generation of high levels of adenosine and $A_{2A}R$ activation, which leads to a suppression of immune responses. Thus, pharmacological blockade of the $A_{2A}R$ could enhance tumour regression and be used as immunotherapy for cancer [171]. These findings led to the redirection of several of the previously developed CNS small molecule inhibitors for cancer treatment. Four of these are currently in phase 1 clinical trials [66].

Despite the great potential as therapeutic target in many human disease areas, the widespread distribution of adenosine receptors, as well as the ubiquitous role of adenosine signalling in many physiological processes, also pose a major challenge for the development of targeted therapeutics. While there are countless studies investigating the use of adenosine ligands as potential therapeutics, most compounds and strategies eventually had to be abandoned because of the lack of specific function and the appearance of adverse effects.

1.4.3 Characterisation of $A_{2A}R$ structure

Human $A_{2A}R$ is a 44.7 kDa protein composed of 412 amino acids. The structure of the $A_{2A}R$ has been extensively characterised and there are several structures published (Figure 1.9). The first $A_{2A}R$ structure was published in 2008, being only the fourth GPCR structure solved, after rhodopsin and the $\beta_{1/2}$ -adrenoreceptors [188]. This was achieved through the aid of an engineered T4-lysozyme fusion protein inserted into the third intracellular loop. The fusion protein improved expression, protein stability in detergents, and also increased the rather sparse polar surface of the $A_{2A}R$, facilitating better crystal contacts. Many structures of the receptor in complex with different ligands followed, applying different strategies to stabilise the receptor for crystallisation (Figure 1.9).

The insertion of a fusion protein in an intracellular loop abolishes the interaction of the receptor with the G protein. Therefore, thermostabilised variants of the receptor were developed. By applying the StaR[®] technology, two different sets of mutations could be identified that either stabilise the receptor in the agonist- or the antagonist-bound state [189, 125, 125]. These mutations increased the melting temperature of the receptor significantly and for the first time allowed the crystallisation of $A_{2A}R$ in absence of any fusion-proteins. Furthermore, the StaR-receptor constructs relied less on stability conferred by high affinity ligands and allowed co-crystallisation with low affinity compounds such as caffeine and theophylline [189]. The structure with the highest resolution so far was obtained from a receptor construct which combines the thermostabilising mutations with a bRIL-fusion protein [140, 139].

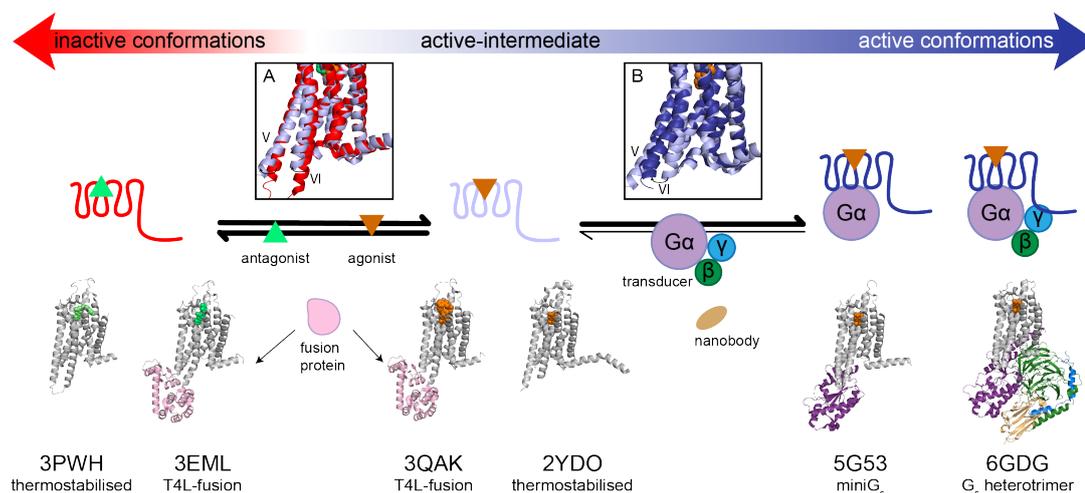


Figure 1.9: Milestone structures of A_{2A}R solved along the receptor’s conformational trajectory. Crystallisation of inactive state (red) A_{2A}R required the presence of an antagonist (green triangle), as well as stabilisation through either a fusion protein (*e.g.* T4L, light pink) or thermostabilising point mutations (*e.g.* 3PWH). Addition of an agonist (orange triangle) shifts the conformational equilibrium of the receptor towards an active-intermediate state (light blue). The main conformational changes on the intracellular side of the receptor between these two states are depicted in panel (A) and include an outward shift of helix V and VI. The fully active state of the receptor is only obtained in presence of an intracellular transducer and was first solved in complex with an engineered, miniaturised G_s protein (purple) (5G53). In 2018 the active state (blue) structure of agonist bound receptor in complex with a heterotrimeric G_s protein (G_α, G_β and G_γ) was solved too (6GDG). The heterotrimeric G_s protein was stabilised by nb35. Panel (B) shows, how in the fully active state, helix VI is moved out even further compared to the active-intermediate state.

To date, the A_{2A}R is one of the best structurally characterised GPCRs and has been crystallised 57 times, in complex with four different agonists and 17 different antagonists¹. As such, most of them represent the inactive state of the receptor. While the agonist-bound structures do display some conformational differences compared to antagonist bound structures, they do not represent the fully active state, but rather an active intermediate state [125, 190].

The first active state structure was only published in 2016, when the receptor agonist complex was stabilised with an engineered minimal protein (mini-G_s) [191]. Two years later a cryo-EM structure of the A_{2A}R, in complex with an engineered heterotrimeric G protein, was published [84]. The cryo-EM and X-ray structure were practically identical, highlighting the validity of using engineered G protein surrogates. Observed structural differences could be mostly accredited to the distinct experimental setup and as such were probably caused by crystal contacts or the difference in pH. However, the interface between mini-G_s and the receptor was highly similar in both structures. As for the other G protein complex structures, the main interactions were formed by the C-terminal α5 helix from the G_α subunit and amino acid residues in TM III, TM V, TM VI, TM VII, helix VIII and ICL 2 of A_{2A}R. Besides these interactions with the mini-G protein, an additional interaction interface between the A_{2A}R and the β subunit was observed, which was significantly more extensive than noted in the β₂AR complex. [84].

¹<https://gpcrdb.org>, accessed 28.02.2021

Interestingly, further comparisons between the different conformational states of the $A_{2A}R$ and the β_2AR revealed that these receptors have different energy landscapes. While the agonist-bound state of the β_2AR resembles the inactive state, the active intermediate state of $A_{2A}R$ is closer to the fully active state [191]. Through this large variety of different $A_{2A}R$ structures, in the active as well as inactive and intermediate states (Figure 1.9), it was possible to get mechanistic insights into receptor activation and ligand recognition.

1.4.3.1 The orthosteric binding site of $A_{2A}R$

The orthosteric binding site of all ARs is located on the extracellular surface between the transmembrane domain and ECL2 (Figure 1.10). Most $A_{2A}R$ ligands contain a heterocyclic core, which is stabilised by hydrogen bonding with the conserved asparagine Asn 253^{6.55} (numbers in superscript denote the residue position according to an extended Ballesteros-Weinstein numbering scheme [192]. The first number indicates the helix (1-7), and the second number the residue position relative to the most conserved residue of the respective helix, defined as number 50. In this extended numbering scheme proposed by Isberg *et al.* [192], residues in conserved loop regions are assigned positions as well. In this case the first two numbers indicate the two helices flanking the loop.) on helix VI and a π -stacking interaction with Phe 168^{45.52} in ECL2 (Figure 1.10). Exocyclic amino groups of the ligand are further stabilised through a hydrogen bond with Glu 169^{ECL2} (Figure 1.10 A). Deeper inside the binding pocket, agonists, as well as antagonists, are stabilised by interactions with His 250^{6.52} (Figure 1.10 A).

In the antagonist bound state, the top end of the binding pocket is closed off by a salt-bridge “lid” formed between Glu 169^{ECL2} and His 264 on ECL3. This salt bridge is observed in most antagonist structures; however, its physiological function is somewhat controversial. While some studies demonstrate a role of this salt bridge in controlling the dissociation kinetics of antagonists from the receptor [140], in structures obtained at a more physiological pH these residues were further apart and adopted different rotamers, questioning the biological relevance of this interaction [84, 193]. Towards the intracellular site, the binding pocket is confined by the bulky lining formed by Trp 246^{6.48}, which is part of the so called “toggle-switch”. Upon receptor activation movement of Trp 246^{6.48} and Phe 243^{6.44} propagates to the outward shift of TM VI that is associated with the active state of the receptor [191, 84].

Agonist specific binding seems to be mainly mediated through the interactions made by the ribose moiety (Figure 1.10 B). In the nineties, the triad Glu 13^{1.39}-Ser 277^{7.42}-His 278^{7.43} was identified as a main feature required for agonist activation [194]. Subsequent structural studies revealed that the Glu 13^{1.39}-His 278^{7.34} interaction is required to stabilise the conformation of the histidine, which in turn forms a hydrogen bond with the ribose (Figure 1.10 B) [190, 125]. The ribose moiety is further coordinated through an interaction with Ser 277^{7.42} and in the case of 5'-N-Ethylcarboxamidoadenosine (NECA) with Thr 88^{3.36}, which explains the higher affinity of this synthetic agonist compared to adenosine [125]. The importance of these polar residues in conveying an active-state conformation of the receptor is further highlighted by the fact that the Ser 277-Ala and Thr 88-Ala mutations are part of the Star2 mutations, stabilising the receptor in the inactive state [189].

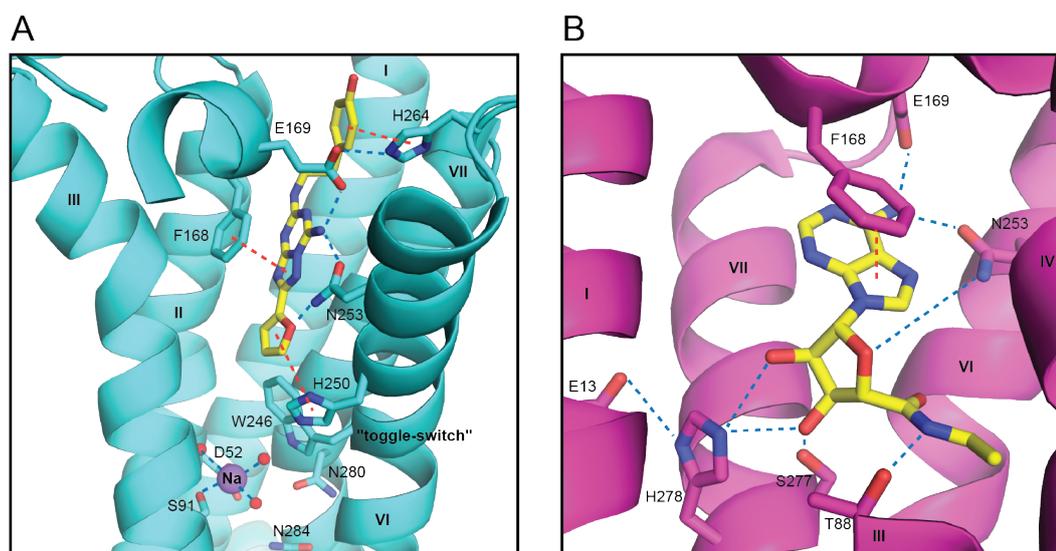


Figure 1.10: Orthosteric binding site of A_{2A}R. (A) A_{2A}R (cyan) bound to the antagonist ZM241385 shown in yellow (PDB: 4E1Y). Residues involved in key interactions with the ligand are shown as sticks. For clarity helices IV and V are hidden. Other helices are labelled with I–VII. Aromatic interactions between the ligands and the receptor are indicated in red, hydrogen bonds in blue. The top of the binding pocket is closed off by a hydrogen bond “lid” formed between Glu 169 and His 264. The bottom of the ligand binding pocket is lined by Trp 246 (toggle-switch), which separates the sodium binding pocket from the orthosteric site. The sodium ion is shown as purple sphere with the coordinated water molecules as red spheres. (B) The orthosteric binding site in the activated receptor (magenta) with NECA bound (PDB: 5G53). In addition to the hydrogen bonds with the ligand the stabilising interaction between Glu 13 and His 278 is highlighted. The upper part of helix III is hidden for clarity.

1.4.3.2 Structural mechanisms of receptor activation

As mentioned previously, the conformational changes observed upon receptor activation are mainly characterised by a 14 Å outwards movement of the cytoplasmic end of helix VI away from the receptor core and slight changes in positions of the cytoplasmic ends of helix V and VII (Figure 1.11 A) [191, 84]. The mechanisms required to induce these changes involve a series of sequence motifs, which are conserved amongst most class A GPCRs. The centre of the large helix bulge in active-like structures is the NPxxY motif on helix VII, the centre of which is formed by Asn 284^{7.49}. Upon agonist binding, Asp 52^{2.50}, another highly conserved residue across the GPCR family, gets protonated. This allows for the formation of a hydrogen bond between the backbone oxygen of Asn 284^{7.49} and Asp 52^{2.50} (Figure 1.11 B), causing the helix bulge in TM VII [195, 196]. Residue Tyr 288^{7.53} experiences an upward movement and is now positioned to form a hydrogen bond with Tyr 197^{5.58}, which is considered a possible conserved activation switch (Figure 1.11 B) [197]. Interestingly, this Tyr-Tyr interaction can only be observed in the fully active, G protein-bound state. In agonist bound intermediate states, Tyr 197^{5.58} has been shown to flip out into the membrane [190, 125, 198].

These conformational changes open up a cleft within the cytoplasmic face of the receptor, into which helix 5 of the G protein is inserted (Figure 1.11 C). The G_α subunit (respectively the mini G protein used to solve the A_{2A}R active state structures) mainly interacts through

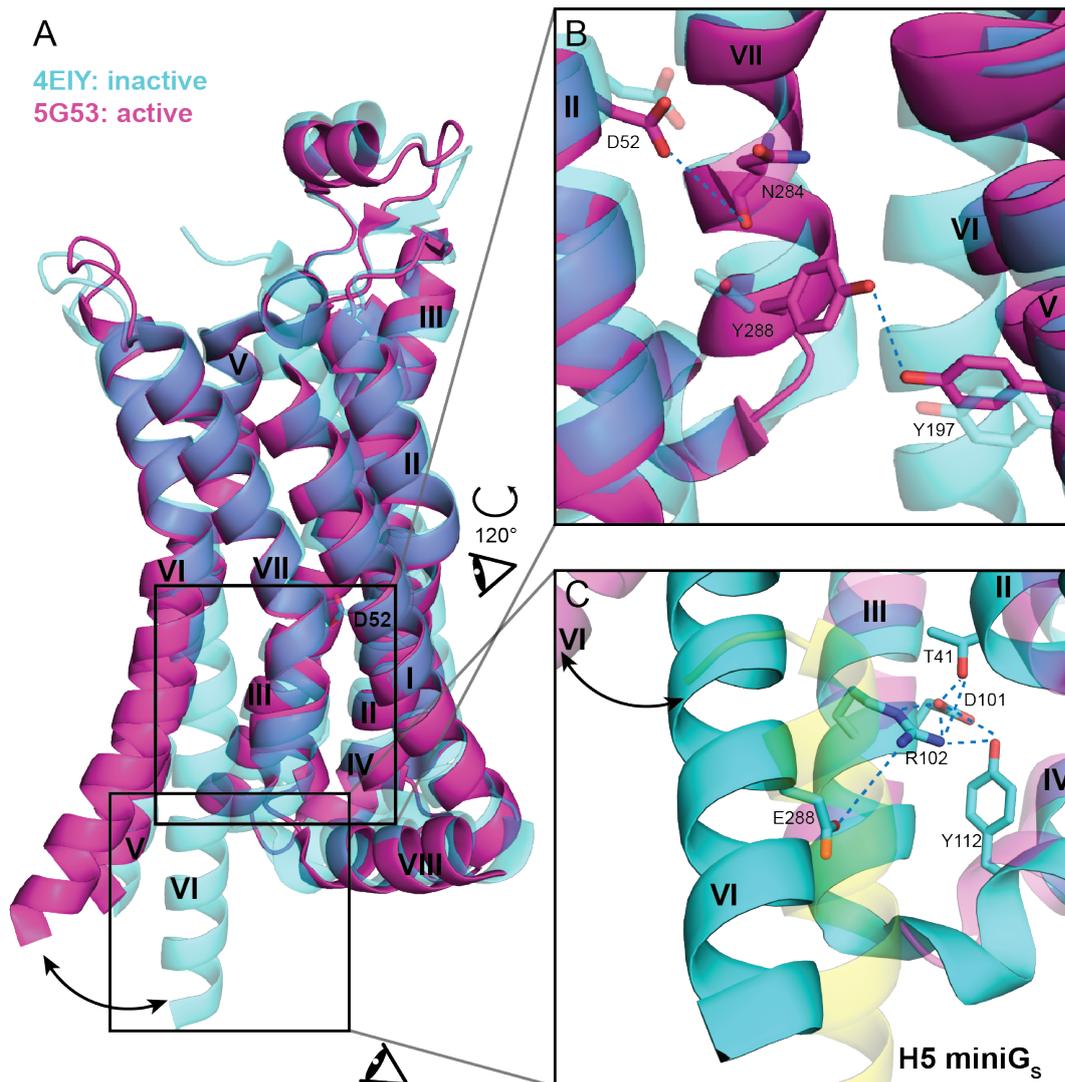


Figure 1.11: Structural changes upon receptor activation. (A) Overlay of active (magenta) and inactive (cyan) A_{2A}R structures (PDB: 4EIY and 5G53). Helices are labelled I–VIII. The large outward movement of helix VI is highlighted by an arrow. Most of the movement in helix V is hidden behind the active state helix VI. (B) Zoom-in of the NPxxY region, after a rotation of about 120° around the receptor. For orientation, Asp 52 is also labelled in panel A. Upon activation Asp 52 gets protonated and forms a hydrogen bond with Asn 284, causing helix VII to slightly bulge. Tyr 288 is rotated more towards helix V which allows the formation of an interaction with Tyr 197. These conformational changes open up a cleft on the cytoplasmic side of the receptor, highlighted in the last panel. (C) Zoom-in on the G-protein binding site. The end of helix 5 of the G_s protein is shown in transparent yellow. It can only be accommodated upon outwards movement of helix VI. In the active state Arg 102 undergoes intensive vdW interactions with the G protein (behind H5, not visible here). However, in the Inactive state the position of Arg 102 is locked by a hydrogen bond with Glu 288 (= ionic lock), which is further stabilised by interactions with the TDY triad (Thr 41, Asp 101, Tyr 112). Hydrogen bonds are shown as dashed blue lines.

extensive van der Waals interactions with residues on helix III, V, VI, VII, ICL 2, and helix VIII with the A_{2A}R; however, six polar interactions are also made across the interface [191]. The GPCR conserved DRY motif (Asp 101^{3.49}, Arg 102^{3.50}, Tyr 103^{3.51}) on helix III is involved in binding the C-terminus of helix 5 of the G_α subunit. Compared to the inactive state structure, Arg 102^{3.50} is moved towards Tyr 103^{3.51}, making space for the G protein to move in. Additionally, intensive van der Waals interactions are formed with Tyr 319 of the G_s subunit [191].

In inactive and intermediate state structures, the position of Arg 102^{3.50} is locked by residues of the TDY triad and the ionic lock. The ionic lock denotes a conserved salt bridge between Glu 228^{6.30} and Arg 102^{3.50} of the DRY motif, which is important for the deactivation of the receptor. It can be observed in several but not all inactive A_{2A}R structures [189]. The TDY triad is formed by Thr 41^{2.39}, Asp 101^{3.49} of the DRY motif, and Tyr 112^{34.53}. It is conserved between all ARs and some class A GPCRs and is involved in the formation of a network of polar interactions that stabilise the ionic lock and thus, stabilising the receptor in the inactive state (Figure 1.11 C) [199].

1.4.3.3 Allosteric regulation

A series of GPCRs have been shown to be allosterically modulated by ions binding. In the case of the A_{2A}R, sodium acts as NAM, specifically stabilising the inactive conformation and decreasing agonist affinity [188]. The first high-resolution crystal structure of A_{2A}R revealed an allosteric binding pocket formed by the highly conserved residues Asp 52^{2.50}, Ser 91^{3.39}, Asn 280^{7.45}, and Asn 284^{7.49}, which is separated from the orthosteric binding site by Trp 246^{6.48} (Figure 1.10 A, Na = purple sphere) [150]. Upon agonist binding, the pocket collapses as a consequence of concerted movement of several residues situated in TM III, TM VI, and TM VII, inducing the structural changes required for receptor activation discussed in the previous section 1.4.3.2.

The same high-resolution structure also revealed specific binding sites for lipids and cholesterol, which possibly modulate receptor activity. Direct cholesterol binding is also consistent with the observation that the addition of cholesteryl hemisuccinate (CHS) to detergent solutions stabilises solubilised receptors [188]. One of the cholesterol binding sites might stabilise parts of helix VI and thereby fixing the position of Asn 253^{6.55} in the binding site [150]. Similarly, a lipid molecule seems to stabilise the conformation of the first eight N-terminal residues of helix I, which do not make any direct contacts with the rest of the helical bundle [150].

Further allosteric binding sites are thought to locate to the extracellular loop regions. Swapping ECL 2 of the A_{2B}R with the ECL of A_{2A}R lead to increased affinity of the A_{2B}R for selective A_{2A}R agonists and therefore, suggest a role of this loop in stabilising agonist-bound receptor [200].

Finally, a crystal structure with a novel antagonist scaffold uncovered a potential allosteric site adjacent to the ligand binding site, which is not explored by the natural ligand [193]. In complex with the studied ligand, Tyr 9^{1.35}, Ala 63^{2.60}, Ile 66^{2.63}, Ser 67^{2.64}, Leu 267^{7.31}, Met 270^{7.34}, Tyr 271^{7.35}, and Ile 274^{7.39} form a compact pocket, in which the two tyrosine residues form

aromatic π -stacking interactions with aromatic substitutions of the ligand. Interestingly, the positions and rotamers of these two residues are unique and the equivalent part in other structures is very shallow or more solvent exposed.

1.4.3.4 Receptor subtype selectivity

Many of the key residues involved in ligand binding and receptor activation are conserved between the different AR subtypes. However, there are some sequence differences that are responsible for some subtype selectivity. The A₃R for example, has a serine in position 6.25, in contrast to the histidine found in the other subtypes, which makes the receptor more tolerant towards larger substitutions deeper in the binding pocket [201]. In the A₁R a threonine is found at position 7.34, while in the A_{2A}R there is a methionine. This substitution is thought to accommodate cycloalkylic groups, which are characteristic for A₁-selective xanthine derivatives [202].

Nonetheless, these specific residues in the main orthosteric ligand binding pocket only explain some of the observed subtype selectivity. Therefore, the extracellular loops, especially ECL2, are suspected to be the main region conveying subtype selectivity, as their amino acid sequence and structural composition vary significantly between the different receptors [203, 204]. One important feature for stabilising ECL2 architecture are disulphide bridges. While the cysteine-bridge between Cis 77^{3.25} and Cis 166^{45.50} is highly conserved, not only in adenosine receptors but all class A GPCRs [197], the first crystal structure of the A_{2A}R exhibited two additional disulphide bonds [188]. Mutation of these residues reduced adenosine potency, demonstrating the structural importance of these interactions [205]. Unexpectedly, the A_{2B}R also has three cysteine residues in ECL2. However, these seem not to play such an important structural role, as no effect on ligand binding was observed upon mutation of these residues [203].

1.4.4 A_{2A}R ligands and structure-based drug discovery

Historical A_{2A}R ligands closely resembled either the endogenous agonist adenosine or are derived from the purines adenine and xanthine (Figure 1.12). However, quite often these chemophores have significant shortcomings. For example, adenosine derivatives suffer from very short *in vivo* half-lives, while many antagonists incorporate toxicophores such as electron rich rings that are poorly soluble. In fact, most antagonists that entered clinical trials have been discontinued due to poor chemical properties, pharmacokinetics and/or compound-related toxicity. Even the recently approved inhibitor Istradefylline still displays suboptimal bioavailability and a poor photostability [206]. Despite the abundance of A_{2A}R ligand information it was difficult to discover novel scaffolds. This changed once the first structures were published, opening the doors for SBDD.

The novel crystal structures enabled highly successful virtual screens, yielding several novel compounds with sub-micromolar affinity and different scaffolds, including triazoles, triazines and chromones (Figure 1.12) [209, 210, 211]. While most of these screens focused on the discovery of novel antagonists, there are also several reports of modelling agonists into the A_{2A}R structure [212, 213, 214]. Nonetheless, virtual studies were not only used to identify novel chemotypes, but further MD studies were used to assess the stability of receptor-ligand

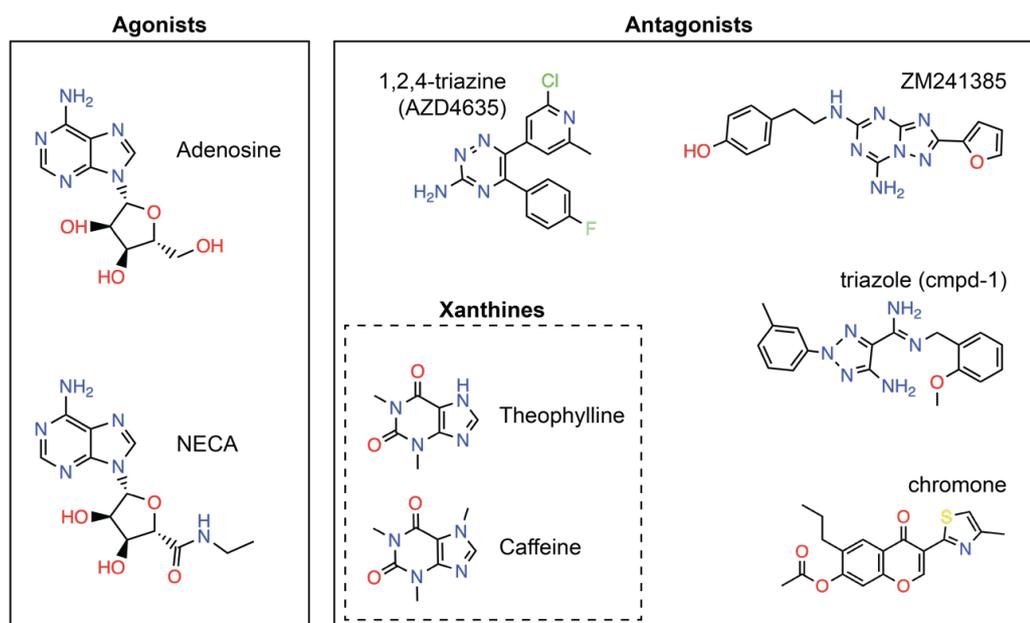


Figure 1.12: Exemplary A_{2A}R ligands relevant for drug discovery. A_{2A}R has been crystallised with four different agonists, shown here are adenosine and NECA. Most antagonists are based on a bicyclic core, *e.g.* a xanthine scaffold or a triazolotriazine core such as found in ZM241385. Through virtual screens novel scaffolds were discovered and followed up. Here shown are the 1,2,4-triazine AZD4635 [207], the triazole cmpd-1 [193] and a ligand from the chromone series [208].

complexes in lipid environments [199], to identify possible ligand recognition domains [215], and in few cases structural information was also used for the development of antagonist lead compounds.

The triazine-series is the first published example for GPCR SBDD, where the chemical scaffold was first identified through in silico screening [211]. Through rational design the initial 1,3,5-compound was optimised to a 1,2,4-scaffold, which fitted deeper into the binding pocket and therefore, displayed higher affinity and increased subtype selectivity over the A₁R [216]. X-ray structure-directed optimisation of this hit-series lead to a further improvement of drug like properties and better off-rates [217]. Subsequently, the lead compound, was validated preclinically as an agent potentially useful for the treatment of cancer [218]. While this small molecular antagonist (now called AZD4635) moved on into clinical phase 1 trials, these came to a stall by the end of 2021, because of the internal prioritisation of other oncology projects¹.

The virtual screening hit triazole has also been the subject of structure-based optimisation and, through directed substitutions, ligand efficiency could be improved; however, affinity could not be further increased [219]. Interestingly, in a different effort to identify a dual compound displaying A_{2A}R, as well as glutamate ionotropic receptor NMDA type 2B (NR2B) antagonist activities, a similar compound, with an aminotriazole core was synthesised. The co-crystal structure with this compound provided insights into antagonist binding that are not observed in other A_{2A}R structures and revealed a potential allosteric pocket [193].

¹<https://www.soseiheptares.com/news/725/129>

Similarly, the first chromone hits could be optimised into potent A_{2A}R antagonists through rational modification of the structure. A detailed druggability analysis of the entire receptor binding site further enabled the identification of energetically “unhappy” water molecules in the binding site. These could be specifically targeted by suitably placed substituents on the chromone core [220]. Subsequent studies then focused on also increasing the pharmacokinetics of this ligand series. Rigorous free energy perturbation methods have been used to explain the underlying structure-affinity relationships of the chromone scaffold. The suggested binding mode could then be further validated and some of the model parameters revised by solving the crystal structure of two chromone-A_{2A}R complexes [208].

Despite the fact that the timelines for most compounds in clinical trials are probably still too short to have been significantly affected by SBDD, these few examples show the potential of using integrated structural and pharmacological information for the optimisation of drug-like properties of lead candidates and how high-resolution views of ligand-protein interactions could drive SBDD of new agents.

1.5 Screening for novel A_{2A}R ligands

Adenosine signalling plays a very important role in many pathophysiological processes. One of the main players mediating these signals is the A_{2A}R. Therefore, this receptor represents a highly interesting drug target. Despite the significant efforts made in many disease areas, only a disappointing number of specific compounds could be developed so far. SBDD approaches promise to change this and consequently, are pursued by many players in pharma industry. Due to the vast amount of structural and pharmacological data available, the A_{2A}R also represents an excellent model for the study and application of novel techniques to other GPCR- and membrane protein-targets, as well as the refinement of current technologies. Based on this rationale, Novartis used the A_{2A}R in two different screens for the discovery of novel ligands.

On the one hand, they explored the use of focused libraries. In a large-scale virtual study, they screened over 3.5 million compounds for A_{2A}R-binding, resulting in over 700 validated hits with previously identified, as well as novel chemical scaffolds. On the other hand, Novartis pursued the implementation of large DNA encoded libraries (DELs, see section 1.1.1.1) for membrane protein targets. DELs are combinatorial libraries and as such, exponentially growing in size. They built a library of 321 million different compounds, which was screened for binding against several different targets. In these experiments, the A_{2A}R was engaged as counter target besides the investigated receptor, which allowed the identification of non-specific binders. While the assay did not yield any results for the targets of interest, it did produce a series of very interesting A_{2A}R hits, with distinct chemical properties. Hence, it became of great interest to further characterise these compounds and evaluate their potential as novel scaffolds for A_{2A}R antagonists.

1.6 Aims of this project

The aim of this project was to characterise a selection of these novel hits pharmacologically and structurally in order to determine specific binding characteristics, which might be promising for further drug development. The main focus of this study was the determination of experimental structures of ligand-receptor complexes. There are very few structures of initial hits; however, identifying the binding modalities of structurally distinct scaffolds early on in the discovery process, would allow rational, more direct structural optimisation, speeding up the discovery process and potentially helping to explore unexpected binding pockets on the receptor.

Chapter 2

Materials and Methods

2.1 Standard protocols

2.1.1 Buffers and chemicals

All buffers were made with either Milli-Q[®] or Extra Pure, deionized water (Thermo Fisher Scientific), using previously prepared stock solutions. These were prepared by dissolving the chemicals, purchased from Sigma-Aldrich, Melford Laboratories, or Thermo Fisher Scientific unless stated otherwise, at an appropriate concentration in Milli-Q[®] or Extra Pure water and were filtered using a 0.22 μm Whatman membrane filter (Cytiva).

2.1.2 Media and agar plates

Media for culturing *E. coli* was prepared using pre-made Lysogeny broth (LB) granules (Sigma-Aldrich), diluted in Milli-Q[®] water and sterilised through autoclaving. Antibiotics were added just before use and diluted to an appropriate working concentration according to Table 2.1. LB-agar plates were prepared by dissolving LB granules together with 1% agar powder (Melford Laboratories) in Milli-Q[®] water. The solution was autoclaved and, once cooled to body temperature, the appropriate antibiotic was added and the solution was distributed into plastic petri dishes, either next to a flame or in a sterile hood. Once solid, the plates were stored at 4°C, while liquid media was stored at room temperature.

Table 2.1: Working concentration for commonly used antibiotics.

Antibiotic	Working concentration
Ampicillin	100 $\mu\text{g}/\text{mL}$
Carbenicillin	100 $\mu\text{g}/\text{mL}$
Chloramphenicol	25 $\mu\text{g}/\text{mL}$
Gentamycin	10 $\mu\text{g}/\text{mL}$
Kanamycin	50 $\mu\text{g}/\text{mL}$
Spectinomycin	50 $\mu\text{g}/\text{mL}$
Tetracycline	10 $\mu\text{g}/\text{mL}$

2.1.3 SDS-PAGE

Samples for sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (PAGE) analysis were mixed with 4 or 5 × SDS-loading dye (250 mM Tris-HCl pH 6.8, 10 % SDS, 0.25 % bromophenol blue, 500 mM dithiothreitol, and 50 % glycerol) and incubated at room temperature for at least ten minutes. Meanwhile, a Mini-Protean[®] TGX (4–20 %) or Criterion[™] TGC (4–20 %) precast gel (BioRad) was pre-run for ten minutes at 160 V with SDS-running buffer (2.5 mM Tris, 19.2 mM glycine, 0.01 % SDS). Samples were then loaded onto the gel, together with a molecular mass protein ladder, such as the Color Prestained Protein Standard (NEB). The gel was run at a constant voltage of 160 V for 56 minutes, or until the dye front reached the bottom of the gel. For analysis of purified protein samples, the gel was then stained for about half an hour with Quick Coomassie Stain (Generon) and subsequently destained with deionised water. For protein analysis of non-purified samples, the gel was left unstained and used for Western-blotting. Both approaches were visualised with a G:BOX (Syngene).

2.1.4 Western blotting

Following SDS-PAGE, proteins on the gel were transferred on a 0.2 µm polyvinylidene difluoride (PVDF) or nitrocellulose-membrane using semi-dry blotting at 400 mA in a Trans-blot[®] turbo[™] blotting system (BioRad). Membranes were blocked in Tris-buffered saline with Tween-20 (TBST) and 5 % bovine serum albumin (BSA) (Sigma), either at 4 °C overnight or at room temperature for at least one hour with constant agitation, and were then probed using either an anti-FLAG-tag (1:1000) monoclonal antibody (mouse Anti-DYKDDDDDK tag 9A3, Cell Signaling Technology), followed by a horseradish peroxidase (HRP) coupled anti-mouse IgG1 secondary antibody (1:5000) or an HRP-coupled HIS-tag antibody (1:5000) (HIS-tag horseradish peroxidase-conjugated antibody monoclonal mouse IgG1 clone, R&D systems, biotechene). After each antibody incubation the membrane was washed three times with TBST for ten minutes. To develop the blot Clarity[™] Western ECL substrate (BioRad) was used and exposed in the G:BOX (Syngene).

In cases where re-blotting with a second antibody was necessary, the membrane was first stripped by washing twice in mild stripping buffer (200 mM glycine pH 2.2, 1 % SDS, 10 % Tween 20) for 5–10 minutes, followed by two washes in phosphate buffered saline (PBS) for ten minutes and two washes in TBST for five minutes. Then the membrane was ready to block and probe again.

2.1.5 Native-Page

Native page was performed following the Novex[®] Life Technologies[™] (Thermo Fisher Scientific) user guide for Bis-Tris NativePAGE[™]. Protein samples were mixed with 4 × NativePAGE[™] sample buffer and 0.125 % NativePAGE[™] G-250 Sample additive and loaded on NativePAGE Novex[®] Bis-Tris (3–12 %) Mini Gels. The inner gel chamber was then filled with 1 × NativePAGE[™] dark blue Cathode buffer and the outer chamber with 1 × NativePAGE[™] Anode buffer. Electrophoresis was performed at 150 V for 90–115 minutes and the gel was subsequently destained in 10 % acetic acid.

2.2 Molecular biology

2.2.1 Primers and genes

All receptor constructs were kindly provided from our industrial collaborator, Novartis AG, and cloned into the appropriate expression vector, where necessary. A list of constructs and primers can be found in the appendix (Table B.2, Table B.3).

Primers were purchased from Eurofins Genomics, where they were purified as standard desalted primers and underwent quality control by MALDI-TOF mass spectrometry (MS).

2.2.2 Plasmid backbones

All receptor constructs were cloned into either a pFastBacTM 1 (Invitrogen) or pcDNATM 3.1 (Thermo Fisher Scientific) plasmid backbone for insect cell or mammalian cell expression, respectively (Figure 2.1). The pFastBacTM 1 vector allows quick integration of the gene of interest into a bacmid using *E. coli* cells. The gene of interest is under the control of a strong polyhedron (PH) promoter for high levels of expression in insect cells. The pcDNATM 3.1 vector is designed for high-level protein expression in mammalian cell lines under the cytomegalovirus (CMV) enhancer-promotor.

2.2.3 Heat-shock transformation

Transformation of plasmid DNA into competent *E. coli* cells (*e.g.* One ShotTM OmniMAXTM chemically competent *E. coli* (Thermo Fisher Scientific)) was achieved following a standard heat-shock protocol. To this end, approximately 0.5 μ L of purified plasmid DNA were added to 50–100 μ L chemically competent cells and incubated on ice for 30 minutes. Then, heat shock

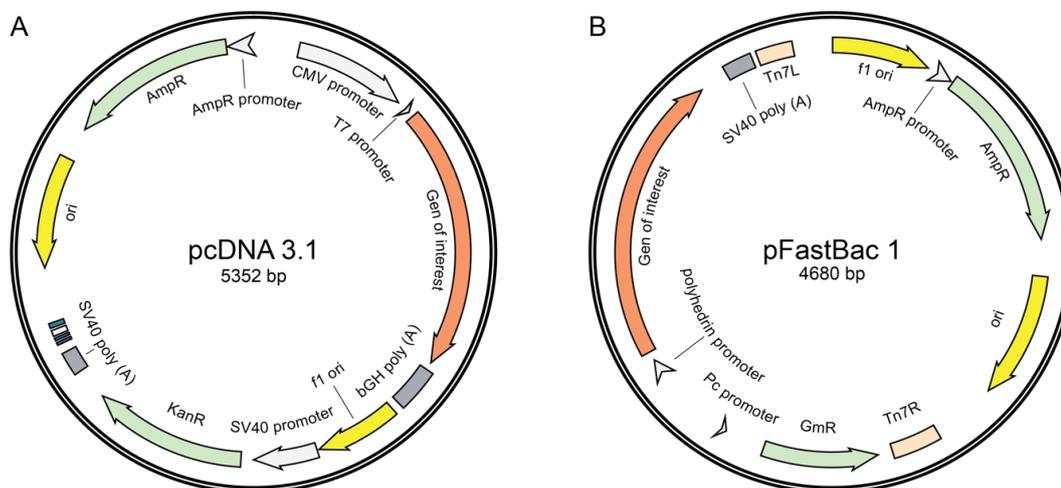


Figure 2.1: Plasmid backbone of the most commonly used vectors in this study. (A) pcDNATM 3.1 and (B) pFastBac vector. Gene of interest denotes the place where the receptor was usually cloned in. Usually this receptor construct would contain an N-terminal HA and FLAG-tag and a C-terminal HRV-cleavage site followed by a 10 \times HIS-tag. The number below the name indicates plasmid size in base pairs without the insert (see Table B.3).

was performed for 30–40 seconds at 42 °C. After another two minutes incubation on ice, 400 μL of fresh LB were added to the cells to aid recovery and they were incubated at 37 °C for one hour. For simple transformations, such as for plasmid amplification, 100 μL of cells were plated directly onto a LB-agar plate containing the appropriate antibiotic. If transforming with a cloning product, cells were first centrifuged for one minute at $3000 \times g$ and resuspended in 100 μL , before plating everything onto an appropriate agar plate.

2.2.4 Polymerase Chain Reaction

DNA amplification for cloning purposes was performed through a polymerase chain reaction (PCR) using the Q5[®] High-Fidelity 2 \times Master Mix (New England BioLabs (NEB)). Usually, reactions were performed in a 50 or 25 μL reaction, consisting of 0.5 μM forward and reverse primers, 1 ng/ μL vector, and 1 \times of the Q5[®] High-Fidelity Master Mix. A standard PCR protocol is shown in Table 2.2. The annealing temperature (step 3) was predicted using the NEB calculator¹ and the extension time (step 4) was adjusted allowing 30–40 s/kb. Generation of the correct product was monitored using agarose gel electrophoresis. If the desired product could not be detected on the gel, optimisation of annealing temperature was determined through gradient PCR.

Table 2.2: Standard PCR protocol.

	Temperature	Time
1.	98 °C	30 s
2.	98 °C	10 s
3.	62–70 °C	30 s
4.	72 °C	0.5–3 min
5.	Repeat steps 2–4	30 \times
6.	72 °C	2 min
7.	10 °C	forever

2.2.5 Colony PCR

To determine if insertion reactions were successful, colony PCR was performed. Following transformation, half of several individual colonies were picked and mixed with 10 μL of PCR master mix, consisting out of 1 \times Q5[®] High-Fidelity Master Mix and 0.5 μM forward and reverse primers. The other half of each colony was streaked onto a fresh LB-agar plate. PCR thermocycling was performed following the standard PCR protocol, but extending the initial denaturation step (step 1) to seven minutes.

2.2.6 Agarose gel electrophoresis

To visualise PCR results, and for DNA purification purposes, samples were mixed with a 6 \times Purple Gel Loading Dye (NEB) and loaded onto an 1 % agarose gel. The gel was prepared by dissolving 1 % agarose in 1 \times TEA buffer (40 mM tris-acetate pH 8.5, 1 mM ethylenediaminete-

¹<https://tmcalculator.neb.com>

traacetic acid (EDTA)), to which 0.5× SYBR safe DNA Gel Stain (Invitrogen) was added. To determine the size of the DNA fragments approximately 2 µL of GeneRuler DNA Ladder Mix (Thermo Fisher Scientific) was loaded alongside the DNA samples. The gels were run in 1× TEA buffer at 100 V for about 20 minutes and visualised using a G:BOX (Syngene).

2.2.7 DNA purification

Gel clean-up of PCR products

Most PCR products and linearized plasmids were additionally cleaned up by isolating the specific band from an agarose gel before subsequent cloning steps. To this end, the entire product was loaded on an agarose gel for gel electrophoresis. The specific DNA band was then visualised using a SafeImagerTM 2.0 Blue-Light Transilluminator (Invitrogen) and excised, paying attention to cut out as little excess agarose gel as possible. DNA extraction was then performed using a Nucleospin[®] Gel and PCR Clean Up Kit (Macherey-Nagel). To this end, the cut-out piece of gel was dissolved in 200 µL/mg of gel of guanidinium thiocyanate binding buffer at 50 °C, before loading onto a silica membrane column. The column was washed twice with 650 µL ethanolic NT3 buffer and then incubated at 70 °C for two minutes, until completely dry. The column was incubated at 70 °C with 15 µL NE elution buffer for another five minutes before eluting by centrifugation. The elution step was repeated once and the DNA concentration was determined using NanoDrop spectroscopy at 260 nm.

Plasmid preparation

For plasmid isolation the desired DNA was heat-shock transformed into One ShotTM Omni-MAXTM chemically competent *E. coli* cells (Thermo Fisher Scientific) for amplification. The next day one single colony was inoculated overnight at 37 °C in about 2 mL of LB media, supplemented with the appropriate antibiotic. The cells were harvested by centrifugation at 3000 × g for ten minutes and the plasmid DNA extracted using a Nucleospin[®] Plasmid MiniPrep kit (Macherey-Nagel). To this end, the cells were resuspended in 250 µL A1 resuspension buffer and subsequently lysed and neutralised by the addition of A2 lysis and neutralisation buffer. The supernatant was cleared by centrifugation at 11 000 × g for ten minutes and transferred to the kit supplied columns. The column was washed twice with AW wash buffer, followed by one wash with A4 buffer, supplemented with ethanol. After drying the membrane, by centrifugation at 11 000 × g for two minutes, the DNA was eluted in 50 µL elution buffer and the concentration determined using NanoDrop spectroscopy at 260 nm. Plasmid samples were stored at −20 °C.

2.2.8 Infusion cloning

To clone an insert into any site in a different vector the seamless In-Fusion HD Cloning Plus (Takara-Bio) strategy was used. Henceforth, insert and linearized plasmid were amplified with a minimum of 15 base pairs homologous overlap at their termini, that were complementary to the linearized vector or insert respectively. After gel clean-up of both fragments, they were annealed using In-Fusion cloning, forming a circular plasmid that could be used to transform *E. coli* cells. This was achieved by incubating the vector and insert at a molar ratio of 1:1–1:2 together with 1× of In-Fusion enzyme mix (Takara-Bio) at 50 °C for 20 minutes. The exact

vector to insert ratio was calculated depending on their size¹. The obtained product was cooled to 4 °C and immediately transformed into One Shot™ OmniMAX™ chemically competent *E. coli*.

2.3 Cell culture

2.3.1 Insect cell culture

For large scale protein expression *Spodoptera frugiperda* (*Sf9*) insect cells were grown in Sf-900™ II serum-free media (SFM) (Gibco®) and maintained at a density of 0.5–8 million cells/mL in shaking cultures at 27 °C and 120 rpm. Cell were split every 48–72 hours. To this end, cell density was counted manually with a hemocytometry or automatically with a Countess™ II cell counter (Invitrogen). Cell viability was determined through trypan blue staining (Invitrogen) and kept above 95 % in maintenance cultures and above 80 % in infected expression cultures.

2.3.2 Bac-to-Bac baculovirus expression

2.3.2.1 Bacmid preparation

High-level recombinant protein expression in insect cells was performed using the baculovirus expression system. To this end, bacmid and virus preparation were performed following adapted versions of the Invitrogen life technologies Bac-to-Bac® Baculovirus Expression System user guide. All expression constructs were cloned into the pFastBac™1 vector, if not already provided in this vector. To isolate recombinant bacmids, MAX Efficiency® DH10Bac cells (Thermo Fischer Scientific) were used. The expression vector was transformed following a standard heat shock protocol, but because bacmid isolation requires multiple antibiotics, cells were incubated at 37 °C for at least 8 hours shaking at 220 rpm before spreading on agar plates containing kanamycin, tetracycline, gentamycin, 5-Bromo-4-Chloro-3-Indlyl β -D-Galactopyranoside (X-Gal) and β -D-1-thiogalactopyranoside (IPTG) at 50, 10, 7 and 100 μ g/mL respectively and 0.16 mM IPTG. The plates were incubated at 37 °C for 48 hours, until blue and white colonies could clearly be distinguished. Four white colonies per construct were picked and incubated in 2 mL of LB media for 16 hours shaking at 220 rpm at 37 °C. Then the cultures were harvested by centrifugation for ten minutes at 2900 \times g. Pellets were resuspended, lysed and neutralised using the Macherey-Nagel Nucleospin® Plasmid Miniprep Kit buffers according to their manual. The cleared supernatant was then mixed with isopropanol to reach a final concentration of 40 % isopropanol. The sample was gently mixed and centrifuged at 15 000 \times g for ten minutes. The resulting transparent pellet was washed with 200 μ L 70 % ethanol. After final centrifugation at 15 000 \times g for ten minutes the ethanol was removed and the pellet air dried next to a flame to keep sterile. The pellet was carefully resuspended in 30 μ L sterile Milli-Q® water and bacmid concentration was determined using nanodrop. Transfection was usually performed straight after bacmid preparation; in few occasions the isolated bacmid was stored at –20 °C before transfection.

¹<https://takarabio.com/learning-centers/cloning/in-fusion-cloning-tools>

2.3.2.2 *Sf9* Cell Transfection

For transfection, 1×10^6 *Sf9* cells were plated on sterile 6-well plates in a total volume of 3 mL Sf-900™ II SFM. The bacmid DNA and the X-tremeGENE HP DNA Transfection Reagent (Sigma-Aldrich) were individually pre-diluted at a 1:10 ratio in Sf-900™ II SFM and then gently mixed at a 2:1 ratio. The mixture was incubated at room temperature for up to 30 minutes before adding 150 μ L to each well of cells. Transfected cells were incubated at 27 °C for about 60 hours until cultures showed clear signs of cell death and individual cells appeared enlarged and irregular shaped compared to the non-transfected control. The supernatant containing the initial baculovirus (V_0) was harvested and stored at 4 °C until further used in LightSafe tubes. The media on the cells was replaced and the cells were incubated for another 60 hours before they were harvested for expression testing.

2.3.2.3 Virus Amplification

On the day of infection 2×10^6 cells/well were plated on sterile 6-well plates in a total volume of 2 mL Sf-900™ II SFM. Cells were incubated at room temperature for 1 hour to allow for attachment, which was verified by microscopy inspection. Once the cells attached to the well, 40 μ L of V_0 virus were added to each well. 48 hours post infection, clear signs of infection were visible when looking at the cells under the microscope. 2 mL of medium containing virus from each well were collected and transferred to sterile 15 mL falcon tubes. The tubes were centrifuged at $500 \times g$ for five minutes to remove cells and large debris and to obtain clarified V_1 baculoviral stock. The supernatant was transferred to fresh LightSafe 15 mL tubes, 2% fetal bovine serum (FBS) (Gibco®) was added and stored at 4 °C.

For large-scale expression the baculovirus was further amplified to V_2 . To this end, *Sf9* cells were grown to a density of 2×10^6 cells/mL. The cells were infected with V_1 virus at a multiplicity of infection (MOI) of 0.1, as a higher MOI reduces the quality of the baculovirus. To calculate the volume of virus required the following formula was used:

$$\text{Inoculum [mL]} = \frac{\text{MOI} \times \text{density} \times \text{culture volume}}{\text{titer of viral stock}} \quad (2.1)$$

For V_1 viral stock a virus titer of 1×10^7 pfu/mL was assumed for this calculation. The infected cells were incubated at 27 °C with shaking at 120 rpm for about 72 hours, until cell viability dropped below 40%. The cells were transferred into two 50 mL falcon tubes and the supernatant was harvested by centrifuging for ten minutes at $500 \times g$. The V_2 virus solution was filtered using 0.45 μ m sterile syringe filters and was transferred into LightSafe tubes. 2% FBS was added and the viruses were stored at 4 °C. Instead of determining a virus titer for the V_2 virus using a plaque assay, the optimal virus dilution for protein expression was determined empirically through small scale expression tests.

2.3.2.4 Expression testing to determine the optimal virus dilution and time to harvest

To determine the amount of virus and expression time required for optimal protein production, without determining the virus titer by plaque assays, small scale expression tests were performed. To this end, *Sf9* insect cells were grown to a density of 2×10^6 cells/mL, and >95 % viability and distributed into 25 mL aliquots. The cultures were infected with 1:100, 1:500, 1:1000, 1:2000, and 1:5000 (v/v) dilutions of V₂ virus and incubated at 27 °C with shaking at 120 rpm. After 48, 60, and 72 hours post-infection, 2 mL aliquots were taken and cell count, viability, and, if a fluorescent expression was co-expressed, fluorescent signal, were determined. Then the cells were harvested by centrifugation at $500 \times g$ for ten minutes. Pellets were then stored at -20 °C until samples of all timepoints were collected. For expression, testing by Western blotting crude membrane preps were prepared. To this end, each 2 mL pellet was resuspended in 500 μ L PBS, supplemented with EDTA free protease inhibitor cocktail (abcam). The cells were lysed by vortexing at maximal speed for one minute and membranes were pelleted by centrifugation at $20\,598 \times g$ for one hour at 4 °C. The pellets were resuspended in 150 μ L high salt buffer (50 mM HEPES pH 7.5, 1 M NaCl 1:100 protease inhibitor cocktail). Samples were then prepared for Western blotting as described above. Expression levels were determined by densitometry analysis relative to each other. The virus dilution and expression time with the highest expression level, but for which cell viability was still above 80 % was chosen for all subsequent experiments.

2.3.2.5 Large-scale *Sf9* expression

Large scale expression was performed in either 1 or 3 L Erlenmeyer flasks with 400 mL or 1 L cultures respectively, at a density of 1.8×10^6 – 2.5×10^6 cells/mL, and >95 % viability. Cells were infected with the previously determined amount of V₂ virus and maintained at 27 °C with shaking at 120 rpm. At the determined day of best expression, cell density, viability and diameter of each culture were monitored and the cells were harvested by centrifugation at $3000 \times g$ for 15 minutes at 4 °C. Cell pellets from 1 L culture were resuspended in PBS, transferred to a 50 mL falcon tube and centrifuged again for ten minutes at $3000 \times g$ at 4 °C. After determining the wet weigh of the pellet, they were flash frozen in liquid nitrogen and stored at -80 °C until further use.

2.3.3 Protein expression in mammalian cells

2.3.3.1 Mammalian cell culture

For recombinant protein expression in mammalian cell culture, either human embryonic kidney 293 cells (HEK293) or Chinese hamster ovary cells (CHO) were cultured in Dulbecco's Modified Eagle Medium (DMEM), supplemented with 10 % FBS, at 37 °C and 5 % CO₂. Cells were maintained at 10–90 % confluency. For passaging, cells were dislodged by a three-minute incubation in $1 \times$ trypsin at room temperature for the HEK cells and 37 °C for CHO cells. Trypsin was neutralised by the addition of fresh DMEM and cells were split into new culture flasks at 10 % confluency.

2.3.3.2 DNA maxi prep

To get viable amounts and concentrations of DNA for mammalian cell transfection a Maxiprep kit (Macherey-Nagel) was utilised. 400 mL of LB supplemented with the appropriate antibiotic were inoculated with 1 mL of preculture containing the plasmid of interest. The culture was grown over night at 37°C, shaking at 200 rpm. The volume of culture corresponding to 1200 OD₆₀₀ was harvested by centrifugation at 5000 × g for 15 minutes at 4°C. Plasmid preparation was then performed following the kit manual. In brief, the cell pellet was resuspended in 12 mL Resuspension buffer (RES-EF) supplemented with RNase A and then lysed and neutralised with the corresponding buffers. The lysate was cleared by applying on an equilibrated NucleoBond[®] Xtra Column Filter. The column was subsequently washed with the provided wash buffers and the plasmid DNA eluted in 15 mL elution buffer (ELU-EF). If available, the final plasmid clean-up was then performed using either the provided Finalizer syringe filters following the provided manual. Alternatively, plasmid DNA was precipitated by adding 0.7 volumes of room-temperature isopropanol. After centrifugation at 15 000 × g for 30 minutes at 4°C the supernatant was carefully removed and the pellet washed with endotoxin-free 70% ethanol. The pellet is allowed to dry at room temperature for 30 minutes and then reconstituted into endotoxin-free buffer (10 mM Tris-HCl, 10 mM EDTA). The final plasmid concentration was determined by UV-spectrophotometry at 260 nm and adjusted to 0.6–1 µg/µL. Before transfection, all DNA samples were filter sterilised using a 0.22 µm syringe filter.

2.3.3.3 PEI transfection of HEK293 cells

HEK293 cells were transiently transfected using the polyethyleneimine (PEI) method. To this end, HEK293 cells were grown as monolayers in DMEM medium supplemented with 10% FBS at 37°C in a moist, 5% CO₂ atmosphere to about 75% confluency. Before preparing the transfection mix, the medium was replaced with fresh DMEM + 5% FBS (Table 2.3), then the flask was returned into the incubator until the transfection mix was ready. For transfection, usually either a 1:1 or 1:1.5 DNA to PEI ratio was used. Plasmid DNA and PEI were prediluted separately in DMEM with 5% FBS according to Table 2.4. After a five-minute incubation the pre-diluted components were mixed and incubated for another ten minutes, to let the complex form. The DNA: PEI-complex was added dropwise to the cells and they were incubated at 37°C in a moist, 7% CO₂ atmosphere for 48 hours. 24 hours post transfection the media was replaced with fresh DMEM supplemented with 10% FBS. To harvest, cells were scraped into an appropriate volume of PBS, collected and centrifuged at 200 × g for five minutes. Cell pellets were frozen in liquid nitrogen and stored at –80°C.

Table 2.3: Volume of media to be added before transfection. As media usually DMEM + 5% FBS was used.

	T25	T75	T175
DMEM to be added (mL)	2	6	13

Table 2.4: DNA to PEI ratios for HEK293 transfection.

	T25	T75	T175
DNA (μg)	25	75	173
PEI (μg) 1:1/1:1.5 DNA to PEI ratio	25/37.5	75/112.5	173/259.5
Total volume transfection mix	500 μL	1 mL	2 mL

The DNA to PEI ratio and expression time were determined empirically for each plasmid and batch of new cells by small scale expression trials using Western-blotting to determine the amount of expressed protein from crude membrane preps as described for the baculovirus titration.

2.3.3.4 Trans-IT transfection of CHO cells

PEI transfection of CHO cells was not successful. Therefore, the TransIT-X2[®] dynamic delivery system (Mirus) was used for transient transfection of CHO cells. Cells were seeded at 2×10^5 cells/mL the day before transfection so that they reached 85–100% confluency at the time point of transfection. The plasmid DNA was diluted with serum free medium, and the transfection reagent was added according to Table 2.5. The optimal DNA: TransIT-X2 ratio (m/v) was determined for each construct in small scale expression tests. The mixture was incubated for 20 minutes at room temperature to allow sufficient time for complex formation. Then, the complex was distributed to the cells in complete growth media. Cells were incubated at 37 °C in a moist, 7% CO₂ atmosphere for 48 hours. The optimal expression time was previously determined in small scale expression tests.

Cells were harvested by adding $1 \times$ sterile Versene[™] solution (Thermo Scientific), incubated for 10 minutes at 37 °C, and collected by centrifugation at $300 \times g$ for ten minutes. The cell pellet was resuspended in PBS, counted with a Countess[™] II cell counter (Invitrogen) and adjusted to the desired concentration in assay buffer. If not used directly for subsequent assays, cells were washed 2 times with PBS before detaching by adding $1 \times$ trypsin and incubating for three minutes at 37 °C. The trypsin was neutralised by adding fresh DMEM, the cells counted with a Countess[™] II cell counter (Invitrogen), and harvested by centrifuging at $300 \times g$ for ten minutes. The cell pellet was resuspended in freezing media (DMEM supplemented

Table 2.5: DNA to TrasIT-X2 ratios for CHO cell transfection. For A_{2A}R transfection a 1:4 DNA to TransIT-X2 ratio was used.

	24-well plate	6-well plate	T25	T75
Surface area	1.9 cm ²	9.6 cm ²	25 cm ²	75 cm ²
Complete growth medium	0.5 mL	2.5 mL	6.5 mL	19.7 mL
DNA (μg)	0.5	2.5	6.5	19
TransIT-X2 (μL)	2	10	26	78
Total volume transfection mix	50 μL	250 μL	630 μL	1.9 mL

with 18 % FBS and 10 % sterile DMSO) at a concentration of 2×10^6 cells/mL and immediately transferred into a -80°C freezer. After 48 hours the frozen cells were transferred into liquid nitrogen storage until further used.

2.4 Protein purification

2.4.1 Protein Quantification

Protein quantification of purified protein was usually performed by one or several of the three following methods: First, UV-spectrophotometry at 280 nm using a NanoDrop system. Second densitometry on SDS-PAGE. To this end, known amounts of BSA references were loaded alongside the other protein samples on the same gel. Band intensities were determined using the Fiji (ImageJ 2.0.0) software and unknown concentrations were determined by linear regression on the BSA standards. Lastly, using a colorimetric detergent compatible (DC) protein assay (BioRad). This assay was performed in 96-well plates, according to the manufacturer's manual. 5 μL of protein sample were mixed with 25 μL of reagent A' and 200 μL of reagent B. Alongside, eight dilutions of a BSA ranging from 0.2- 1.5 mg/mL were prepared to generate a protein standard curve. Each assay point was prepared in triplicate. After a 15-minute incubation at room temperature, absorbance was read at 750 nm. Unknown protein concentrations were determined by linear regression from the standard curve using Microsoft Excel. Final concentrations were calculated using the mean of sample triplicates.

Protein concentrations of membrane samples were determined using the PierceTM BCA Protein Assay Kit (Thermo Fisher Scientific). The assay was performed according to the manufacturer's manual and unknown protein concentrations were determined as described for the DC-assay.

2.4.2 A_{2A}R membrane preparation

Frozen cell pellets were thawed and resuspended in low salt buffer (50 mM HEPES pH 7.5, 10 mM MgCl₂, 20 mM KCl) and supplemented with EDTA free protease inhibitor cocktail (Abcam). Each pellet was homogenised by Dounce homogenisation, applying 25-30 strokes. Then the homogenate was centrifuged at $234\,788 \times g$ for 60 minutes at 4°C . The supernatant was discarded and the pellet resuspended in high salt buffer (50 mM HEPES pH 7.5, 1 M NaCl) with protease inhibitors and homogenised and centrifuged again for 60 minutes at $234\,788 \times g$ and 4°C . This high salt wash was repeated until each cell pellet corresponding to one litre of cell culture was washed with at least 420 mL of high salt buffer. The mass of the final membrane pellet was determined and used to calculate the amount of freezing buffer (50 mM HEPES pH 7.5, 10 mM MgCl₂, 20 mM KCl, 40 % Glycerol) according to the following formula:

$$\text{volume [mL]} = \frac{\text{pellet weight [g]} \times 10}{3} \quad (2.2)$$

After adding protease inhibitors, and careful resuspension of the pellet, the membranes were flash frozen in liquid nitrogen and stored at -80°C until further use.

If working with the A_{2A}-StaR2-bRIL construct, buffers contained 40 mM Tris pH 7.4 instead of HEPES pH 7.5 and the final membrane pellet was resuspended in 50 mL 40 mM Tris pH 7.4 per 2 L of cell pellet for freezing.

2.4.3 A_{2A}R membrane preparation for binding assays

For the radioligand binding experiments, wild type A_{2A}R was transiently expressed in HEK293 cells. Membranes were prepared according to a protocol adapted from Jaakola *et al.* [188]. Cell pellets were resuspended in 20 mL per three T175 cell pellets of ice-cold 50 mM Tris-HCl buffer pH 7.4 + 1:100 EDTA free protease inhibitor mix (abcam). A Dounce homogenizer was used to homogenise the cell suspension. The cytosolic and membrane fractions were separated using a high-speed centrifugation step of 100 000 × g at 4 °C for 30 minutes. The pellet was resuspended in 10 mL of Tris buffer and the homogenisation and centrifugation step was repeated. The resulting pellet was resuspended in 3 mL 50 mM Tris-HCl buffer, pH 7.4. BCA assay was used to determine the amount of total membrane protein and adenosine deaminase (ADA) was added to a final concentration of 0.8 IU/mL. The final membranes were distributed into 500 uL aliquots, flash frozen in liquid nitrogen and stored at −80 °C. Receptor expression was verified using Western-blotting.

When using *Sf9* insect cell membranes for the binding assay, cell pellets from 200 mL of cell culture were used, which expressed the different receptor constructs. Crude membrane preparations were prepared according to a protocol established thorough *pers. com.* with V.-P. Jaakola based on the protocols used in Jaakola *et al.* [188]. The cell pellets were resuspended in 4 mL of low salt buffer supplemented with protease inhibitors, homogenised using Dounce homogenisation, and centrifuged at 25 000 × g for 40 minutes at 4 °C. This step was repeated three times, then the pellets were resuspended in 5 mL binding buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 1 mM MgCl₂) and the total protein concentration was determined using the BCA assay. The final membranes were aliquoted flash frozen in liquid nitrogen and stored at −80 °C. Correct protein expression was verified with Western-blotting.

2.4.4 A_{2A}R purification for crystallisation

The A_{2A}R purification was adapted from previously published work. Purification of the A_{2A}-StaR2-bRIL construct was based on Rucktooa *et al.* [50], while purification of A_{2A}-bRIL was based on Liu *et al.* [150] and *pers. com.* with V.-P. Jaakola.

2.4.4.1 Purification of A_{2A}-StaR2-bRIL

Per purification, usually 2–4 L of *Sf9* insect cell culture were used (corresponding to 6 × 10⁹–10 × 10⁹ cells or 24–48 g of cell paste). Membranes were prepared as described in section 2.4.2. Membranes were thawed, resuspended in 40 mM Tris-HCl pH 7.6, supplemented with EDTA free protease inhibitor cocktail and 3 mM theophylline and incubated for 2 hours at room temperature. Then the receptor was solubilised with 1.5 % decylmaltoside (DM) for 2 hours at 4 °C. Insoluble material was separated by centrifugation at 145 000 × g for 60 minutes at 4 °C. The solubilised material was applied to a 5 mL nickel-nitrilotriacetic acid (Ni-NTA) Superflow cartridge (QuiagenP), pre-equilibrated in 10 column volumes (CV) 40 mM Tris pH 7.4, 200 mM

NaCl, 0.15 % DM, 1 mM theophylline, using a peristaltic pump at about 2–3 mL/min flow. The column was washed with 25 CV wash buffer (40 mM Tris pH 7.4, 200 mM NaCl, 25 mM imidazole, 0.15 % DM, 1 mM theophylline), then the receptor was eluted with 4 CV elution buffer (40 mM Tris pH 7.4, 200 mM NaCl, 280 mM imidazole, 0.15 % DM, 1 mM theophylline), collecting 2 mL fractions. Elution fractions were analysed on SDS-PAGE and protein containing fractions were pooled and concentrated to about 500 μ L using a Vivaspin 20, 50 000 MWCO PES (Satorius) concentrator. The protein sample was centrifuged at $20\,598 \times g$ for ten minutes at 4 °C and applied on a Superdex 200 10/300 GL (Cytiva) column, pre-equilibrated with 40 mM Tris pH 7.4, 200 mM NaCl, 0.15 % DM, 1 mM theophylline. Size exclusion chromatography (SEC) was performed at 0.3 mL/min flow, collecting 200 μ L fractions. The protein content from the monomer-peak was verified by SDS-PAGE, pooled and concentrated to 30–40 mg/mL using a Millipore Amicon[®] Ultra-4 50k MWCO concentrator (Merk). Protein concentration was determined using NanoDrop, densitometry on SDS-PAGE, and the DC assay.

Protocol adaptations aiming to optimise protein quality

To improve protein quality the following adaptations were made to the protocol stated above:

To speed up the protocol, and reduce possible proteolysis, membranes were only incubated for 30 minutes at 4 °C. During the solubilisation step, the NaCl concentration was increased to 0.8 M; other buffers were kept the same. Further, the Ni-NTA column was connected to a ÄKTA avant (Cytiva) system, allowing for automatisation. The solubilised material was applied with 0.8 mL/min flow, while the wash and elution step were performed at 2 mL/min and 1 mL/min respectively. All concentration steps were performed with Millipore Amicon[®] Ultracell concentrator with a 50k cut off. For SEC a new Superdex 2000 Increase 10/300 GL column was used. To better separate the peaks the flow rate was reduced to 0.2 mL/min.

Co-purification with FOC-2

As the amount of ligand available was limited, for co-purification with FOC-2, the purification was performed using loose nickel resin instead of a prepacked column, allowing to reduce buffer volumes as much as possible. 1 mL Ni Sepharose[™] 6 Fast flow resin (Cytiva) per litre of culture was washed in MilliQ[®] water and equilibrated with 40 mM Tris pH 7.4, 200 mM NaCl, 0.15 % DM, 1 mM theophylline before incubating with the solubilised material for 2 hours at 4 °C. The resin was transferred to an empty Ecnopac[®] Chromatography column (BioRad) and the flow through was collected. The column was washed with 15 CV 40 mM Tris pH 7.4, 200 mM NaCl, 25 mM imidazole, 0.15 % DM, 1 mM theophylline, followed by 10 CV of the same buffer but with 50 μ M FOC-2. Then, the receptor was eluted with 4 CV elution buffer containing 50 μ M FOC-2 instead of theophylline. All following steps were performed following the above stated protocol, expect of changing the SEC buffer to 40 mM Tris pH 7.4, 0.2 M NaCl 1.5 % DM, 10 μ M FOC-2.

2.4.4.2 Purification of A_{2A}-bRIL

Per purification, usually membranes from 3–9 L of *Sf9* insect cell culture were used. They were thawed and the volume was diluted 1:2 with high salt buffer to reduce the glycerol concentration to 20 %. Then, 4 mM Theophylline and 2 mg/mL iodacetamide were added and the sample was

incubated for 45 minutes, rolling at 4 °C. The receptor was solubilised using 1 % n-Dodecyl- β -D-Maltoside (DDM) and 0.2 % Cholesteryl hemisuccinate (CHS) for 4 hours at 4 °C. Solubilised material was isolated through ultracentrifugation at $145\,000 \times g$ for 60 minutes at 4 °C, diluted 1:2 with high salt buffer to further reduce the glycerol and detergent concentration and supplemented with 20 mM imidazole. 1.5 mL HisPur™ Cobalt Resin (Thermo Fisher Scientific) per litre of culture were washed in MilliQ® water and equilibrated with 50 mM HEPES pH 7.5, 800 mM NaCl, 10 mM MgCl₂, 10 % Glycerol, 20 mM imidazole before incubating with the solubilised material rolling at 4 °C overnight. The resin-supernatant mixture was centrifuged for ten minutes at $3000 \times g$ at 4 °C and most of the supernatant was decanted. The rest of the material was distributed onto empty Ecno-Pac® Chromatography columns (BioRad), so that the resulting resin volume did not exceed 3.5 mL per column. The remaining flow through was collected, and the column washed with 10 CV of wash buffer 1 (50 mM HEPES pH 7.5, 0.8 M NaCl, 10 mM MgCl₂, 10 % glycerol, 25 mM imidazole, 0.1 % DDM, 0.02 % CHS, 8 mM ATP, 25 μ M ZM241385), followed by 4 CV of wash buffer 2 (50 mM HEPES pH 7.5, 0.8 M NaCl, 10 % glycerol, 50 mM imidazole, 0.05 % DDM, 0.01 % CHS, 25 μ M ZM241385). The protein was then eluted with 4 CV elution buffer (25 mM HEPES pH 7.5, 0.8 M NaCl, 10 % glycerol, 220 mM imidazole, 0.025 % DDM, 0.005 % CHS, 25 μ M ZM241385), collecting 1.5 mL fractions. Elution fractions were analysed by SDS-PAGE and fractions containing the highest amount of protein were pooled and concentrated to 60 mg/mL using a Vivaspın 20, 100 000 MWCO PES (Satorius) concentrator. The final protein concentration was determined by NanoDrop and densitometry on SDS-PAGE. For protein quality control, size exclusion was run with an aliquot of the final sample using Superdex 200 10/300 GL (Cytiva) column, pre-equilibrated with 25 mM Tris pH 7.4, 0.8 M NaCl, 1 % glycerol, 0.91 % DDM, 0.002 % CHS, 25 μ M ZM241385.

Co-purification with different Ligands

For co-purification of the A_{2A}-bRIL construct with different ligands the ligand was exchanged while bound to the affinity column by adding the ligand of interest to the wash buffer. The used concentration of ligands in each purification step are listed in Table 2.6.

Table 2.6: Ligand concentration added to each purification step.

Co-purification with	Solubilisation	Wash 1	Wash 2	Elution
Theophylline	4 mM	1 mM	1 mM	1 mM
LUF5834	4 mM theophylline	50 μ M	50 μ M	50 μ M
ZM241385	4 mM theophylline	25 μ M	25 μ M	25 μ M
FOC-2	4 mM theophylline	1 mM theophylline	50 μ M	50 μ M

2.5 Biochemical and pharmacological assays

2.5.1 Radioligand binding

2.5.1.1 Saturation binding assays

Experimental setup

Saturation assays were performed using *Sf9* insect cell membranes overexpressing either a bRIL-wt-A_{2A} or a bRIL-StAR2-A_{2A} construct (for more information see Table B.3). Membrane aliquots containing 1 µg of protein were incubated in a total volume of 200 µL binding buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM MgCl₂) to get the best assay window between specific and non-specific binding. Nonspecific binding was determined in the presence of 100 µM theophylline and 0.5 nM [³H]ZM241385. All binding assays were performed in MultiScreenTM-HV 96-well filter plates with 0.45 µm hydrophilic Durapore[®] PVDF filters (MilliporeSigmaTM). For the saturation assay, 100 µL of cell membrane (1 µg of protein) were used. Serial dilutions of radioligand [³H]ZM241385 (American Radiolabeled Chemicals, Inc.) were prepared in binding buffer, ranging from 0.125–64 nM, of which 50 µL were added to each well, together with 50 µL of binding buffer. The plate was incubated at room temperature for 1.5 hours to reach equilibrium. Incubation was terminated by rapid vacuum filtration to separate the bound and free radioligand. Filters were subsequently washed five times with 200 µL of room temperature binding buffer, the plastic cover on the bottom of the plate was removed and each filter was carefully detached and transferred to vial with 10 mL Emulsifier-SafeTM scintillation fluid (PerkinElmer). The filters were incubated in the scintillation fluid for approximately two hours, the bound radioactivity was detected using a 1450 LSC & luminescence Microbeta TriLux Scintillation counter (Perkin Elmer), measuring each flask for five minutes. Additionally, counts per minute (cpm) of the equivalent of 200 µL 1 nM radioligand were measured, as reference to convert cpm to decays per minute (dpm).

Data analysis

A conversion factor to translate the measured cpm in dpm was calculated using the reference measurement of 1 nM radioligand and the specific activity listed in Table 2.7. With this factor, every value was converted into dpm before further analysis. Then the data was fitted in GraphPad Prism 9.2.0 using the non-linear regression function for saturation binding – one site – total and nonspecific binding (Equation 2.3), assuming one site, independent binding.

$$Y = B_{max} \times \frac{X}{K_d + X} \quad (2.3)$$

Each value was measured in triplicates. K_d and B_{max} values were calculated from the mean of three independent experiments; errors are given as standard error or the mean. Shown curves are representative curves from one single experiment.

Table 2.7: Parameters of reference measurement to convert cpm in dpm.

Specific activity	50 Ci/mmol
Volume	200 μ L
Concentration	1 nM
Amount	0.2 pmol
Theoretical dpm	22 000 dpm

2.5.1.2 Competition binding assays

Experimental setup

To determine the affinity of the different novel ligands competition binding assays were performed. These were done using membranes from HEK293 cells transiently transfected with wild-type A_{2A}R. Membrane aliquots containing 10–15 μ g of protein were incubated in a total volume of 200 μ L binding buffer (50 mM Tris-HCl pH 7.4) to get the best assay window between specific and non-specific binding. Nonspecific binding was determined in the presence of 100 μ M theophylline and represented less than 25 % of the total binding. Then, to each well were added 100 μ L of cell membrane (15 μ g protein), 50 μ L of 2 nM radioligand [³H]ZM241385 and 50 μ L of the indicated compounds in increasing concentrations in the same buffer. The mixture was incubated at room temperature for 2 hours to reach equilibrium. Incubations were terminated by rapid vacuum filtration and filters were subsequently washed 5 times with 200 μ L of ice-cold binding buffer. The bound radioactivity was determined as stated above.

Data analysis

The cpm data was fitted in GraphPad Prism 9.2.0 using the non-linear regression function using the One site – Fit logIC₅₀ Least square fit function (Equation 2.4). Each value was measured in triplicates. pIC₅₀ values were calculated from the mean of three independent experiments; errors are given as standard error or the mean. Shown curves are representative curves from one single experiment, normalised against the minimal (no radioligand) and maximal (no competing cold ligand) response of each dilution series.

$$Y = \text{Bottom} + \frac{\text{Top} - \text{Bottom}}{1 + 10^{X - \text{LogEC}_{50}}} \quad (2.4)$$

2.5.2 LANCE cAMP assay

Experimental setup

To assay compound potency, a LANCE[®] Ultra cAMP assay kit was used (PerkinElmer). To reduce variability in protein expression from different transiently-transfected cell batches, frozen cell-stocks from one single large-scale transfection were used. The cells were thawed rapidly at 37 °C and stepwise diluted in 5 mL PBS to prevent osmotic shock. They were centrifuged for ten minutes at 275 \times g and resuspended in 1 mL of fresh PBS. After automatically counting the cells they were centrifuged again for five minutes at 275 \times g and resuspended in stimulation buffer (PBS pH 7.4, 5 mM HEPES, 0.1 % BSA stabiliser, 50 μ M rolipram, 50 μ M cilostamide,

0.8 IU/mL ADA) to a final concentration of 400 cells/ μ L. To create a standard curve, concentrations ranging from 1×10^{-6} M to 1×10^{-11} M cAMP were prepared in stimulation buffer. Similarly, $4\times$ intermediate dilution series of the different antagonists were prepared in stimulation buffer, including 11 half-log dilutions starting at the highest soluble compound concentration at maximally 10% DMSO and a “no antagonist” control. The assay was performed in Optiplates-384 microplates (PerkinElmer). 2000 cells/well were pre-incubated with 2.5 μ L of antagonist dilution for 30 minutes at 37 °C before stimulating the A_{2A}R expressing cells with 0.31 μ M or 75 μ M NECA, respectively, by incubating for 30 minutes at 37 °C. Separately, a 1:50 dilution of the Eu-cAMP tracer and a 1:150 dilution of the ULight solution were prepared using the provided detection buffer. 5 μ L of each solution were pipetted separately onto the stimulated cells, resulting in a final assay volume of 20 μ L. The assay was incubated with the detection mix for at least an hour before measurements were recorded using a Victor x5 multi-label plate reader (PerkinElmer), with an excitation filter of 340 nm and 615/665 nm emission filters.

Data analysis

The time-resolved fluorescence resonance energy transfer (TR-FRET) signal was plotted in GraphPad Prism 9.2.0 by non-linear regression using the log (agonist/inhibitor) vs. response – Variable slope (four parameters) function fitting the Hill slope from the data (Equation 2.5). For fitting the agonist curves the HillSlope is >0 while for inhibitor curves HillSlope <0 . Each value was measured in triplicates. Top and bottom values were then taken to normalise and invert the data. When reporting the EC/IC₅₀ values the mean of all experiments was taken. The error is reported as standard error of the mean. To plot the means, the mean of the triplicates from each individual experiment were taken and the error bars represent the standard deviation.

$$Y = \text{Bottom} + \frac{\text{Top} - \text{Bottom}}{1 + 10^{(\text{Log}EC_{50} - X) \times \text{Hill Slope}}} \quad (2.5)$$

2.5.3 Gel shift assay

For biophysical studies with the novel ligands, an mCherry tagged wild type A_{2A}R construct was used and expressed in 3.6 L of *Sf9* insect cell culture volume. As stated previously, membranes were first washed extensively and then solubilised in 1% Lauryl Maltose Neopentyl Glycol (LMNG) with 0.2% CHS. The solubilised material was mixed with 2 mL TALON[®] metal affinity resin (TaKaRa) per litre of culture. After overnight incubation at 4 °C the resin was transferred to a Econo-Pac[®] Chromatography column, washed with 15 CV of 50 mM HEPES pH 7.5, 0.8 M NaCl, 10 mM MgCl₂, 10% glycerol, 25 mM imidazole, 0.001/0.0002% LMNG/CHS, 8 mM ATP, 1 mM theophylline, and 10 CV of 50 mM HEPES pH 7.5, 0.8 M NaCl, 10% glycerol, 75 mM imidazole, 0.001/0.0002% LMNG/CHS, 1 mM theophylline. Then, the receptor was eluted with 220 mM imidazole. Fractions were pooled into four 2.5 mL samples, and exchanged into buffer without imidazole and additional theophylline using PD10 desalting columns (Cytiva). Samples were concentrated to a protein concentration >1 mg/mL using 100K MiliPore concentrators. The protein concentration was monitored by measuring absorbance at 558 nm on the NanoDrop (for mCherry detection). The monomer in all four samples was about 60–70% pure.

For the gel shift assay, 10 mM ligands in DMSO were pre-diluted 1:100 in protein buffer. Per ligand, 8 assay points were prepared in a PCR plate, as well as one point in a separate tube. To this end, 60 μg of protein were mixed with 10 μM pre-diluted ligand in a total volume of 10 μL . The samples were equilibrated at room temperature for 30 minutes before running a thermogradient ranging from 55–80 $^{\circ}\text{C}$ on a real-time PCR machine for another 30 minutes. A separate sample was kept on ice during this time. All samples were mixed with SDS-loading dye and run on SDS-PAGE. Band intensities were determined with the BioRad ImageLab software and normalised against each respective 4 $^{\circ}\text{C}$ reference.

2.5.4 nanoDSF

For nanoDSF experiments the A_{2A}-StaR2-bRIL construct was used, purified in the same way as for crystallographic experiments. The protein was diluted 1: 200 to dilute out the ligand present from the co-purification, resulting in a final protein concentration of 0.1–0.2 mg/mL. The protein was incubated with 50 μM ligand in a total volume of 30 μL (3.5 μM). As negative-control 100 % DMSO was added instead of ligand. The samples were left for 30 minutes at room temperature to equilibrate and were then centrifuged for one minute at 700 \times g before loading in high sensitivity grade Prometheus NT.48 capillaries (Nanotemper). The capillaries were loaded into the Prometheus NT.48 and a discovery scan was performed to adjust the excitation power of the device was to give fluorescence signal at 330 nm just between 15 000 and 20 000 RFU. Melting scans were performed at 90 % intensity from 20–90 $^{\circ}\text{C}$ with a ramp rate of 2.5 $^{\circ}\text{C}/\text{min}$, recording fluorescence emission at 330 nm (F_{330}), 350 nm (F_{350}) and the 350:330 ratio. Measurements were all performed in triplicates. The apparent melting temperatures (T_m) were obtained at the minimum of the first derivative of the 350:330 ratio by the PR.ThermControl software and the ΔT_m for each ligand was calculated by using the DMSO sample as null-reference.

2.6 Crystallography

2.6.1 Crystallisation screens

For crystal screening, a series of different commercial as well as homemade crystallisation screens were used. The used commercially available sparse matrix screens included the MemGold1 and 2, the MemTrans and the MemMeso screen from Molecular Dimensions, as well as the Rigaku LCP Wizard and the JBS-LCP-HTS screens, which have been specifically designed to include a range of conditions commonly associated with crystallisation of membrane proteins. However, most of them were designed for vapor diffusion crystallisation and not for the use with LCP. Therefore, some of the screens were also used as 70 % dilution.

Additionally, homemade optimisation screens were designed around known A_{2A}R crystallisation conditions as described below.

DoE-screen

A more specific matrix screen was designed for the A_{2A}-StaR2-bRIL construct, focusing around the published crystallisation conditions. A novel screening matrix was generated combining the different crystallisation components using fractional factorial design of experiment (DoE) to

cover as much of the chemical space as possible (Ellistat^{©1}). Many of the factors used in the screen design are discrete, which increases the number of conditions to be screened. Therefore, the experiment had to be manually modified, as the screen had to fit on a 96-well plate and hence, did not fully follow the fractional factorial design. The screen was set up with two different buffers (100 mM MES and 100 mM sodium citrate) at pH 5 and 5.5, PEG400 from 25–35 % (v/v) as precipitant, 5 different salts (200 mM Na-tartrate, 100 mM NaCl, 200 mM LiSO₄, 50 mM NaSCN and 10 mM Zn-acetate) and two different additives (2 % (v/v) 2,5-hexanediol and 0.5 mM theophylline) in different combinations. A list of the complete screen can be found in the appendix (Table B.1). An xml file was generated of the screen that could be read by the RockMaker[®] software (Formulatrix), so the screen could be dispensed using the Formulator[®] screen builder liquid handler (Formulatrix).

HEPTARES-cond PEG screen

This screen was set up as an optimisation screen around the published A_{2A}-StaR2-BRIL-theophylline complex condition (9–32 % (w/v) PEG 400, 100 mM tri-sodium citrate pH 5–5.4, 0.05 M sodium thiocyanate, 2 % (v/v) 2,5-hexanediol, 0.5 mM theophylline [221]), varying the PEG400 concentration from 20–48 % in one direction and pH from pH 4–7.5 in the other direction.

SCRIPPS screens

For crystallisation of the A_{2A}-bRIL construct optimisation screens were designed around the published condition for the A_{2A}-bRIL-ZM241385 complex (25–28 % (v/v) PEG400, 0.04–0.06 M sodium thiocyanate, 2 % (v/v) 2,5-hexanediol, 100 mM sodium citrate pH 5 [150]), varying the PEG400 concentration from 20–31 % and the pH from pH 3.5–6.5 (SCRIPPS 1 and 2). The SCRIPPS 4 screen additionally covered the pH space from pH 5.5–7.5, while SCRIPPS 3 and 5 were copies of SCRIPPS 2 and 4 but set up without any 2,5-hexanediol present.

2.6.2 Ligand exchange

When attempting to exchange the ligand before crystallisation, 100 or 200 μ M ligand (depending on the concentration of the stock solution in DMSO) was added to an aliquot of purified, concentrated, and ultra-centrifuged (15 minutes 436 000 \times g at 4 °C) protein. To this end, the ligand was usually diluted 100-fold to keep the DMSO concentration as low as possible. The sample was incubated at room temperature for 30 minutes to allow for equilibration. Then, any precipitated ligand or protein was removed by centrifugation at 20 598 \times g for 15 minutes at 12 °C.

2.6.3 *In-meso* crystallisation

Before setting up crystallisation trials, the final protein sample (A_{2A}-StaR2-bRIL construct: usually at \sim 30–40 mg/mL, A_{2A}-bRIL construct: usually at \sim 60 mg/mL) was reconstituted into lipidic cubic phase (LCP) by mixing detergent solubilised receptor with monoolein (Generon), supplemented with 10 % (w/w) cholesterol (Anatrace), using the twin-syringe method, in a

¹<https://ellistat.com>

final ratio of 40 % protein, 54 % monoolein, and 6 % cholesterol. If the receptor was crystallised in complex with any ligand other than ZM241385, an additional 10 μ M ligand were supplemented to the lipid mix before mixing with the protein. Crystallisation trials were performed in 96-well glass sandwich plates (JennaBioscience or Laminex) by an NT8 crystallisation robot (Formulatrix) using 50 nL protein-laden LCP overlaid with 0.8 μ L precipitant solution in each well, and sealed with either a glass cover or a Laminex Film Cover (Molecular Dimensions). Protein reconstitution in LCP and crystallisation trials were carried out at room temperature (\sim 18–20 °C). The crystallisation plates were stored and imaged in an incubator/imager (RockImager[®] 1000, Formulatrix) at 20 °C. Each LCP bolus was automatically imaged every few days with normal and cross polarised imaging and the resulting images were viewed using the RockMaker[®] software. Crystals were harvested by cutting open the glass cover with a diamond glass cutter (or scalpel if using plastic covers) and scooping up a large fraction of the LCP bolus with a mounted mesh LithoLoop 0.2/0.4 mm -20/40 micron (Molecular Dimensions) that was then immediately flash frozen in liquid nitrogen without adding an extra cryoprotectant.

2.6.4 *In-meso* soaking

The soaking of LCP-crystals was performed following the protocol described in Rucktooa *et al.* [50]. For soaking experiments, the A_{2A}-StaR2-bRIL construct was purified in complex with theophylline, reconstituted in LCP, and crystallisation trials were set up using LCP-SBSG-100/200 Modular SBX XPO glass plates with a 100-micron UVP Cover film (Swissci). To open the well a small, cross-wise incision was made, so one corner of the cover film could be carefully lifted to add the soaking liquid. The opened well was flooded with 10 μ L of mother liquor supplemented with 1 mM ligand, leading to a final ligand concentration of 925 μ M and a theophylline concentration of 74 μ M. The well was re-sealed using crystal clear sealing tape (Hampton Research) and incubated for either one or 24 hours at 20 °C before harvesting the crystals.

2.7 Data collection and Structure solving

2.7.1 Semi-automated serial data collection

The A_{2A}-bRIL-ZM241385 data was collected using a semi-automated serial data collection strategy at the I24 micro-focus beamline at diamond light source. Since each mesh loop carried a large fraction of the LCP bolus, containing many microcrystals, crystals could not be visually identified, but had to be located using a grid scan. About 10–20 crystals per mesh were selected and their position coordinates stored in the data collection table, which was then used for semi-automated data collection. For each crystal 100 images with 0.1 degree of oscillation at 50 % transmission and 0.01 seconds exposure were collected.

2.7.2 Data processing

Indexing, scaling and multi-crystal merging was achieved by using the xia2-multiplex pipeline which is part of the DIALS project [222], run automatically at Diamond Light Source Ltd. Molecular replacement was carried out using the Phaser graphical user interface (GUI) in

Phenix 1.14-3260, employing a modified version of the 4E1Y structure, from which all additional features, such as lipids and ions, as well as the bound ZM241385 ligand were deleted to prevent model bias. An unambiguous density for ZM241385 was observed immediately after molecular replacement. To confirm the identity of the ligand, either ZM241385 or the appropriate novel ligand were manually built into the ligand density using the Crystallographic Object-Oriented Toolkit (COOT) [223].

Then, three rounds of refinement were carried out using the phenix.refine program [224] in the Phenix software package.

2.8 Electron microscopy

The preparation of negative stain and cryo-grids, as well as grid screening and data acquisition on the respective microscopes, were assisted by Alex Flynn and Bolin Wang.

2.8.1 Receptor-Fab complex preparation

To generate particles that are large enough to be suitable for cryo-EM analysis, a receptor-Fab complex was formed using the A_{2A}-bRIL receptor construct and the anti-bRIL BAK5 antigen-binding fragment (Fab) [144]. The A_{2A}-bRIL construct was purified as described above. However, 1% and 0.001% LMNG were used for solubilisation and purification respectively. The protein was concentrated to 2 mg/mL using a Vivaspin20 100K concentrator. Then purified Fab (kindly provided by Andreas Kießling) was added at a 1:1 molar ratio, as well as 100 μ M FOC-6. The sample was incubated for 30 minutes at 4 °C to allow for complex formation and equilibration and was then further concentrated using an Amicon[®] Ultra-2 100K concentrator to about 26 mg/mL. Large aggregates were removed by centrifugation at $20\,817 \times g$ for ten minutes at 4 °C, before injecting 50 μ L of sample on a Superose 6 Increase 5/150 GL column (Cytiva), pre-equilibrated with 25 mM HEPES (pH 7.5), 0.2 M NaCl, 1% Glycerol, 10 μ M ZM241385, and 20 μ M FOC-6. SEC was deliberately run without any additional detergent in the buffer to reduce the free LMNG concentration as much as possible. A flow rate of 0.2 mL/min was used, collecting 100 μ L fractions and total of 3 runs were performed, each time injecting 50 μ L of sample. All fractions containing the A_{2A}-bRIL-Fab complex were pooled and the concentration determined using NanoDrop. A 30 μ L aliquot of the unconcentrated sample (0.4 mg/mL) was taken for further analysis, the rest was concentrated using an Amicon[®] Ultra-2 100K concentrator to 7 mg/mL. While concentrating, 30 μ L aliquots were taken at 1.5 mg/mL, and 3.5 mg/mL for analysis of intermediate protein concentrations. Any leftover protein that was not directly used for EM was flash frozen in liquid nitrogen and stored at -80 °C.

2.8.2 Negative stain

To assess sample homogeneity after freezing, before setting up cryo-Grids the complex was first analysed by negative stain. To this end, the 0.4 mg/mL sample was further diluted in SEC buffer to 40 μ g/mL. Carbon coated copper grids with 200 mesh (Sigma-Aldrich) were glow discharged using a Ted Pella Easiglow system prior to applying 4 μ L of diluted sample. The

sample was incubated on the grid for one minute, then excess liquid was carefully removed with a piece blotting paper. For staining 4 μL of 1% uranyl acetate were applied twice, the second time incubating for one minute before carefully blotting off excess liquid. The grid was dried completely under a lamp before mounting on the microscope. A second grid was prepared the same way, but using 2% uranyl acetate instead. The grids were screened on a T12 FEI Tecnai G2-spirit microscope, using a Gatan Ultra Scan 4000 CCD camera at 68k \times magnification.

2.8.3 Cryo-grid preparation

The first set of cryo-grids was prepared straight after finishing the concentration step of the SEC-purified complex. For this first series of grids, either 0.6/1 or 1.2/1.3 400 mesh QuantiFoil[®] grids were glow discharged either using a GloQube (Quorum) (30 mA, 30 seconds, negative polarity) or a Tergeo Plasma cleaner (PIE Scientific). Cryo-samples were then prepared using a Vitrobot MK IV, applying 3 μL of sample, using a blotting time of six seconds and a blotting force of six. The chamber temperature was set to 8 $^{\circ}\text{C}$ and the humidity between 40–60%. Grids were clipped and stored in liquid nitrogen.

A second set of grids was prepared using thawed protein samples. Homogeneity of these samples was first assessed by negative stain before proceeding with cryo-grid preparation. For this second set, Au-Flat Protochips GF-12/1.3-3 Au-45 nm gold grids were used. Grids were glow discharged using the Tergeo Plasma cleaner and prepared using the Vitrobot as described above. However, this time, a blotting force of 6 or 0 was used and the chamber humidity was set to 90%. The 0.4 mg/mL sample was not used in this series, as previous screening showed that higher protein concentrations gave better results.

2.8.4 Cryo-EM data collection

Cryo-EM images were acquired on a TitanKrios G3i microscope (Thermo Fisher Scientific) equipped with a Selectris energy filtered FEI Falcon 4 direct electron detector in counting mode, corresponding to a pixel size of 0.91 \AA , at 300 kV and 130k \times nominal magnification. Each image was dose-fractionated into 46 frames with a dose rate of 6.05 electrons per pixel per second and total exposure time of five seconds, resulting in an accumulated dose of 36.5 electrons per \AA^2 . Data were collected over 24 hours resulting in a total number of 5757 movie stacks.

2.9 *In silico* structural analysis

2.9.1 Generating a homology model for the A_{2B}R

A_{2B}R protein structure was predicted by submitting the amino acid sequence to the GPCR-I-TASSER [225] and GPCR-SSFE 2.0 [226] servers for *ab initio* modelling. Additionally, homology modelling was performed using the SWISS-MODEL [227], and ModBase [228] server, using a series of different A_{2A}R (4UHR, 2YDV, 3EML and 4E1Y), as well as β_2 AR (2RH1 and 2R4R) and rhodopsin (1L9H and 1U19) structures as template. Each model was validated using the QMEANBrane (Qualitative Model Energy Analysis for membrane proteins) scoring function

from SWISS-MODEL [64]. The models with the highest QMEAN Z-score were further refined employing the ModRefiner algorithm on the Zhang group server [190]. Based on its final QMEAN Z-score, the ModRefined homology model generated on the ModBase server using an antagonist stabilised A_{2A}R structure (3EML) template, was selected for further applications.

2.9.2 Docking with Glide

Molecular docking was performed using the Glide component [229] in the Schrödinger Maestro suite 12.5.139. To this end, the antagonist bound 4EIY structure was imported into the Maestro suite and processed using the protein prep wizard. During this process missing hydrogens were added, lipids and all waters beyond 3 Å were deleted and hydrogen bonds were assigned using PROPKA at pH 7.0. Restrained minimization was performed using the OPLS3e (optimised potentials for liquid simulations) force field. The ligand and bound Na⁺ ion were left in place. Similarly, the ligands were prepped using the OPLS3e force field, generating possible states at pH 7.0 using Epik. Chiralities in the ligand were determined from the 3D structure. To define a search space the grid preparation wizard was used. For this, the enclosing box was defined using a centroid of the workspace ligand.

As the number of tested ligands was limited, they could be docked using extreme precision. This docking method is designed to be a more powerful and discriminating procedure than standard settings, in which the standard sampling is followed by an additional anchor and growth procedure. Additionally, the sampling of torsion was biased for amides as nonplanar conformations were penalised. Per ligand 10 poses were included in the final result output.

Protein and ligand preparation for docking the ligands into the A₁R and A_{2B}R was performed as described above, using the 5UEN structure for A₁R and the generated A_{2B}R model. As the 5UEN structure contains a covalently bound ligand, the ligand could not be automatically recognised by the program. Therefore, the SO₃-moiety of the ligand had to be removed, breaking the covalent interaction with Tyr 271, so the centroid around the workspace ligand could be drawn when preparing the screening grid. The A_{2B}R homology model does not contain any ligand at all. Therefore, the screened grid space was defined around the residues ASN 254 and PHE 173, using a box length of 10 Å. Docking was then performed using the same parameters as for the A_{2A}R. Root-mean-square-deviation (RMSD) values between docking results and ligands from crystal structures were calculated by uploading the ligand coordinates to DockRMSD [230].

2.9.3 Docking using constraints

Based on the known interaction with other ligands, two different constraints were set up during the grid preparation step and tested when docking the ligands. A positional constraint at different radii was set around the aromatic ring of the ligand interacting with the phenylaniline at position 45.52 and a hydrogen bond constraint for the asparagine at position 6.55. Analysis of docking reference compounds revealed that only the positional constraints of 2 Å around the aromatic ring of the ligand did improve docking result.

As there was no ligand for the A_{2B}R structure, the positional constraint was set at 4 Å around Phe 173^{45.52}.

2.9.4 Docking with Autodock Vina

To look for binding sites outside the orthostatic binding pocket, the AutoDock Vina software (Molecular Graphics Lab, The Scripps Research Institute) [26] was used. A search space of $1\,322\,880\text{ \AA}^3$ was defined by a fid (size (\AA): $x = 106$, $y = 120$, $z = 104$) centered on the centre of the 4E1Y model prepped in Maestro, encompassing the entire surface of the receptor. While the receptor model was kept rigid, rotation around rotatable bonds of the ligand was allowed. The search efficiency and parameters were set a default for each run and up to 20 modes were written into the result file, using an exhaustiveness of 300.

Chapter 3

Pharmacological characterisation of novel A_{2A}R drug-screening hits

3.1 A novel selection of A_{2A}R ligand hits

In order to evaluate novel screening methods for their GPCR targets, and to identify novel chemotypes of A_{2A}R inhibitors, Novartis performed two large-scale screening campaigns. The first of these campaigns entailed the virtual screening of a focused library, which led to the identification of over 700 hits. The second was focused on the application of DELs for membrane protein targets. Novartis validated the hits from both screens previous to this study and confirmed that they were actual A_{2A}R-binders. Furthermore, they were all classified as inhibitors. Six hits were selected from both screens and provided for further pharmacological and structural analysis. Ligands originating from the focused virtual screen are subsequently referred to as FOC-1 – FOC-6, while DEL-screen ligands are referred to as DEL-1 – DEL-6. None of the validation data were made available, therefore, no previous knowledge was assumed except that these ligands were all supposed to bind and inhibit the receptor.

The collection of hits consists of a variety of different chemical scaffolds, several of which had not yet been investigated as A_{2A}R ligands. They were roughly categorised in two classes according to their cyclic core (Table 3.1, Table 3.2). Most of the structurally characterised A_{2A}R ligands contain a purine-like bicyclic core. Accordingly, half of the newly identified ligands also display this known A_{2A}R-ligand feature. FOC-4, for example, is a purine bioisostere, very similar to the natural ligands theophylline and caffeine. In contrast, the other half of the selected ligands contain a monocyclic core. Nonetheless, some compounds still have similarities to known ligands, *e.g.* FOC-2 is a triazine bioisostere similar to the hit series described in Congreve *et al.* [217] and FOC-3 contains a triazole core, which is also present in the antagonist characterised by Sun *et al.* [193]. Despite not being completely novel scaffolds, further analysis of these compounds would provide a reference point for the evaluation of the used methods and results.

Table 3.1: List of novel compounds containing a monocyclic core. The ligands were identified in either a DEL screen or a focused virtual screen. The cyclic system highlighted in the structure is referred to as core of the respective compound.

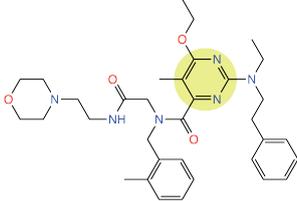
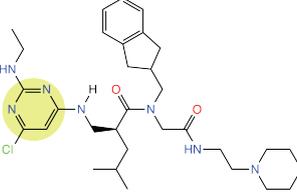
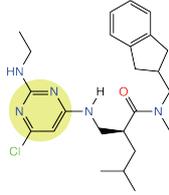
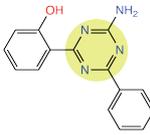
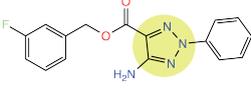
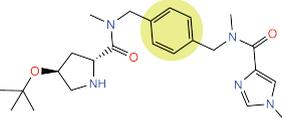
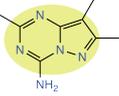
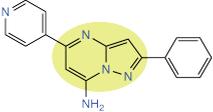
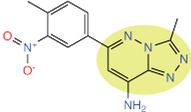
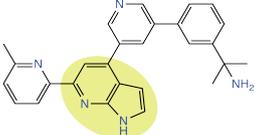
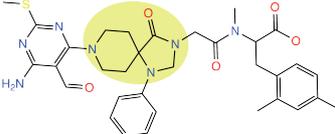
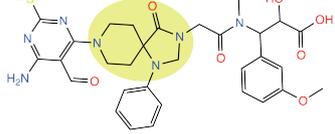
Compounds with monocyclic core				
Name	Screen	Core	MW [Da]	Structure
DEL-1	DNA encoded library	pyrimidine	634.83	
DEL-2	DNA encoded library	pyrimidine	600.21	
DEL-5	DNA encoded library	pyrimidine	558.02	
FOC-2	Focused virtual screen	triazine	264.29	
FOC-3	Focused virtual screen	triazole	312.3	
DEL-4	DNA encoded library	benzene	441.58	

Table 3.2: List of novel compounds containing a bicyclic core. The ligands were identified in either a DEL screen or a focused virtual screen. The cyclic system highlighted in the structure is referred to as core of the respective compound.

Compounds with bicyclic core				
Name	Screen	Core	MW [Da]	Structure
FOC-4	Focused virtual screen	pyrazolotriazine	177.21	
FOC-5	Focused virtual screen	pyrazolopyrimidine	287.33	
FOC-6	Focused virtual screen	triazolopyrimidine	283.3	
FOC-1	Focused virtual screen	diazaindene	419.53	
DEL-3	DNA encoded library	triazaspiro	661.3	
DEL-6	DNA encoded library	triazaspiro	663.8	

Interestingly, the selection also contained some ligands which seemed to originate from the same ligand series. DEL-5 and DEL-2 for example contain the same scaffold (Table 3.1). However, the large morphonyl-propylpropanamide moiety found in DEL-2 is replaced by a small methyl substituent in DEL-5. Similarly, DEL-6 and DEL-3 were generated from the same scaffold only differing in the substitution of a single group (Table 3.2). Before attempting any structural characterisation, the ligand receptor interactions were characterised pharmacologically. This information was then used to guide the decision on which compounds to focus subsequent efforts.

3.2 Measuring concentration-response curves and affinity

The first step in the initial refinement process, after having identified a series of hits from a drug screening assay is to generate concentration-response curves. This verifies that the compounds are showing normal competitive behaviour and specifically bind the target in a reversible manner. If this is not the case, it may indicate that the compound does not bind the target, but rather that the observed effect is caused by an unspecific interaction with another assay component. Furthermore, recording concentration-response curves allows the determination of the affinity of the compounds, which is considered one of the cornerstones of rational drug design, commonly serving as the main factor for choosing promising leads and their optimisation strategies [231, 232].

Despite significant advances in alternative technologies, such as fluorescence- and bioluminescence-based assays [233], radioligand binding assays remain one of the most sensitive and quantitative techniques to measure binding parameters between a given receptor and its ligands, given that there is a suitable radioligand available [234]. Especially for the study of GPCR-ligand interactions, radioligand binding assays are very well established; the first ever GPCR radio-ligand assay having been established during the 1960s [235].

In a saturation assay the receptor is incubated with increasing amounts of radioligand until all of the receptors that can bind the radioligand are occupied. This allows both ligand affinity and receptor density to be determined [236, 237]. If the binding constant of the radioligand is known, the affinity of unlabelled compounds can also be determined by testing their ability to compete the radioligand for the receptor binding site over a wide concentration range. This is termed a competition assay [238, 239].

For $A_{2A}R$, radioligand assays are frequently used to characterise receptor integrity and receptor-ligand interactions, employing commercially available radiolabelled agonists as well as antagonists. Thus, I choose a radioactive competition binding assay using the well characterised radioligand [3H]ZM241385 [240] to determine the affinity of this novel selection of hits.

3.2.1 Setting up a competition assay for $A_{2A}R$ from HEK293 cells

The most physiologically relevant system available was to transiently transfect human embryonic kidney cells (HEK293) with wild type (WT) human $A_{2A}R$. To this end, a WT HA-Flag- A_{2A} -10 \times HIS construct was cloned into a 3.1-DNA vector for mammalian expression. The tags

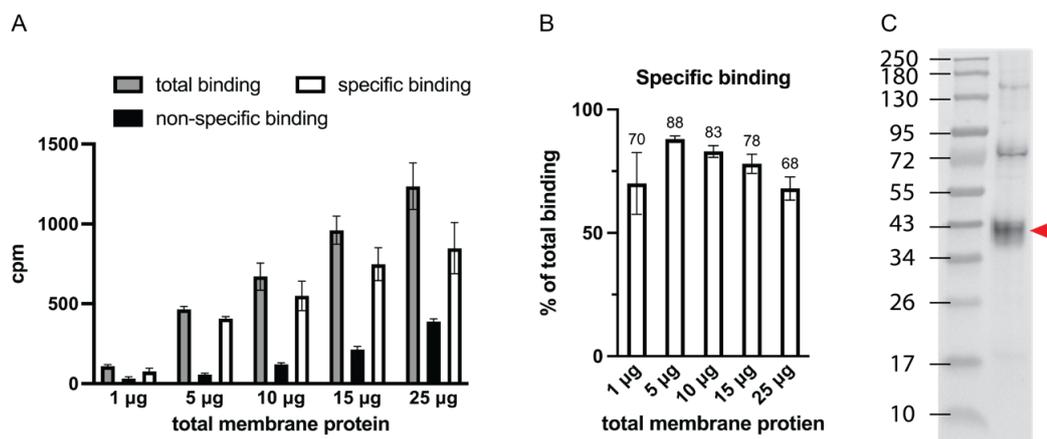


Figure 3.1: Determination of optimal amount of membrane protein to use for radioactive binding assays. Two independent experiments were performed, each in three technical replicates. Here, representative results from one experiment are shown. Error bars represent standard deviation (SD) between the technical replicates. **(A)** Cpm from varying amounts of total HEK cell membrane protein incubated with 0.5 nM [3 H]ZM241385 (total binding). Non-specific binding was determined in parallel reactions in the presence of 100 μ M theophylline. Specific binding can then be determined by subtracting the non-specific binding from the total binding. **(B)** Specific binding represented as percentage of total binding. **(C)** Western blot of HEK293 cells expressing A_{2A}R probed with a HIS-tag antibody. The red arrow head indicates the monomer of A_{2A}R.

were added for easy protein identification via Western blot. Transient transfection was done following the PEI method [241]. Best PEI: DNA ratio as well as the expression time were tested and determined by Western blot using a HIS-tag HRP-coupled antibody (data not shown).

Radioactive binding assays were performed on crude membrane preparations. To this end, the cells were lysed, and membranes were concentrated while washing off soluble cell components. The amount of total protein was determined using a BCA assay to standardise the amount of receptor used in subsequent steps. The membranes were then incubated with the respective ligands for about an hour to allow equilibrium to be established (or at least steady state). Excess ligand was washed off by filtration, leaving only the membrane bound radioligand, which could then be detected through scintillation counting. The presence of the receptor in the membrane sample was checked by Western blot (Figure 3.1 C).

Competition binding assays are usually performed at a radioligand concentration slightly below the dissociation constant K_D . As saturation as well as competition binding assays with [3 H]ZM241385 are very well described in literature, this value was taken from Jaakola *et al.* [188]. However, the amount of membranes used was determined empirically, to establish the largest assay window between total and non-specific binding. To this end, total binding was measured by incubating 1 μ g, 5 μ g, 10 μ g, and 25 μ g of total membrane protein with 0.5 nM radioligand (Figure 3.1 A). Non-specific binding was determined in parallel reactions in the presence of an excess (100 μ M) of theophylline, allowing determination of specific binding as a fraction of total binding. Using 5 μ g or 10 μ g of total membrane protein resulted in 80–90% specific binding (Figure 3.1 B). However, using more membranes led to significantly higher non-specific binding, resulting in only 68% \pm 4.7% specific binding at 25 μ g of total membrane protein. When using less membranes, only very low counts were obtained, which resulted in

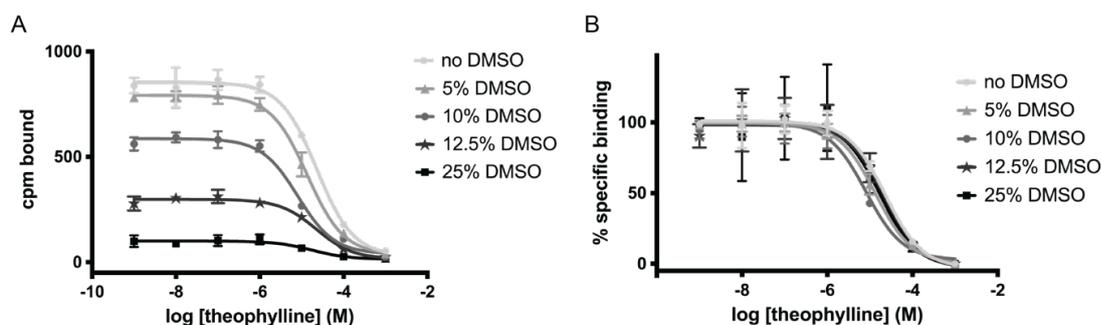


Figure 3.2: Effect of increased DMSO concentration on binding assays. (A) Competition binding assays with theophylline were performed on membranes from HEK cells transiently transfected with A_{2A}R. The ligand was dissolved in increasing concentrations of DMSO. The number of cpm recorded drastically decreased with higher DMSO concentrations. Curves from a single experiment are shown, performed in triplicates. Error bars indicate SD between technical replicates. (B) Curves from A after normalisation. Despite the effect on the measured signal the shape of the curve was not affected, by the DMSO. However, the error bars were significantly larger.

decreased specific binding and increased variation in-between the three technical replicates ($70\% \pm 12.5\%$ at $1\ \mu\text{g}$ total membrane protein). Therefore, a membrane concentration of $5\text{--}10\ \mu\text{g}$ of total membrane protein was chosen for subsequent experiments.

Dimethyl sulfoxide (DMSO) is the most common solvent for compounds in drug discovery, as its able to dissolve both polar and nonpolar compounds at very high loads [242]. However, at higher concentrations DMSO is toxic to cells and can impact the outcome of pharmacological assays. As water solubility of some of the novel compounds was limited, first the effect of increasing concentrations of DMSO on the binding assay was determined by recording concentration response curves for the reference compound theophylline at varying DMSO concentrations. This experiment showed that DMSO had a significant effect on the maximal number of counts recorded (Figure 3.2 A). At a final concentration of 25% DMSO, only 10% of the signal could be recorded compared to a curve without any DMSO. Despite this, the shape of the curve was not significantly affected, and the curves overlaid after normalisation (Figure 3.2 B). The determined IC_{50} were comparable; however, with the reduction in signal, the variation in-between the three technical replicates increased significantly. Therefore, I established that the maximal concentration of DMSO used for this assay should not exceed 5%. This was higher than the arbitrary 1% commonly used as a cut off in many pharmacological assays, allowing to probe slightly higher concentrations of the less water-soluble compounds, but still low enough to get a good number of counts and reasonable error bars. In practice, this limit meant that the highest ligand concentration that could be measured corresponded to $1/5$ of the stock solution, given the compound is soluble in 20% DMSO.

3.2.2 Determining the binding constants for the novel ligands

To validate the assay, binding constants of some well characterised reference compounds were determined and compared to literature values. To this end, theophylline and 2-chloroadenosine (Cl-adenosine) were chosen. Cl-adenosine is a more stable analogue of the natural agonist adenosine, while theophylline is a xanthine derivative, similar to caffeine, that acts as a nons-

elective adenosine antagonist. Both compounds have been extensively described in literature, are easily commercially available and have a reasonable water-solubility. To generate dose response curves, ligand dilution series were prepared covering concentrations 3 log units lower and higher than the reported K_i . Additionally, a no-ligand sample was included, which was used later on to normalise the curves. For the ligands with unknown affinity, dilution series were started with a maximal concentration of 1 mM and then adjusted according to the first results. For each ligand, at least three independent experiments were performed, each time in triplicate (except for DEL-6, as there was not enough material available). As a control, one of the reference compounds was included in each experiment.

The data were fitted in GraphPad Prism using non-linear regression, assuming the presence of a single binding site and reversible binding at equilibrium. The inhibition constant K_i of a non-labelled ligand can be calculated if the K_D of the radioligand is known [243]. However, since the K_D of the radioligand was not assayed and could only be taken from literature, it was not used to calculate K_i values for the competing ligands, as this value was not determined under the same experimental conditions. Instead, the concentration of inhibitor required to displace half of the bound radioligand was reported as IC_{50} . However, by definition (Cheng-Prusoff equation, Equation 3.1) the IC_{50} will be very similar to the K_i in the situation where the radioligand concentration $[R]$ is lower than its K_D . Therefore, the IC_{50} represents an accurate relative measure of the affinity of unlabelled ligands in competition assays where the rest of the system remains unchanged (radioligand concentration, receptor preparations, *etc.*). This should be the case in the experiments reported here, as membranes from the same batch were used in each experiment.

$$IC_{50} = K_i \left(1 + \frac{[R]}{K_D} \right) \quad (3.1)$$

The pIC_{50} values of theophylline and adenosine were determined at 4.6 ± 0.03 and 5.7 ± 0.05 , respectively. In literature, pK_i values ranging from 5.2–5.6 can be found for theophylline and from 6.8–7.4 for adenosine (Figure 3.3 A, Table 3.3) [189, 125, 244, 245]. According to the Cheng-Prusoff equation, IC_{50} values will always be slightly higher than the calculated K_i (Equation 3.1). The remaining discrepancy may be explained by differences in expression system and experimental conditions. This reasoning is further supported by the observation that the difference between recorded and literature value was slightly larger for adenosine than for theophylline. It is known that $A_{2A}R$ agonist affinity is highly sensitive to buffer conditions ($A_{2A}R$ is *e.g.* allosterically modulated by sodium chloride [188]), and therefore it is not surprising that more variability in the determined binding constants could be observed for this ligand. Nonetheless, the obtained results were very robust within the performed experiments ($n = 9$), therefore, the assay was suitable to determine the binding constant of the novel ligands.

For all tested compounds concentration response curves could be fitted, confirming that they behave as $A_{2A}R$ ligands. For DEL-3, DEL-5, and DEL-6 only a relatively low pIC_{50} of 4.2 ± 0.03 , 3.9 ± 0.12 , and 3.6 ± 0.08 were determined (Figure 3.3 B). Unfortunately, it was not possible to include enough data points from the bottom plateau of these ligands, as concentrations above 1 mM could not be tested due to solubility issues. However, since the recorded pIC_{50} of these three ligands was below the pIC_{50} of theophylline, which is regarded a low affinity compound [246], they did not represent very promising candidates.

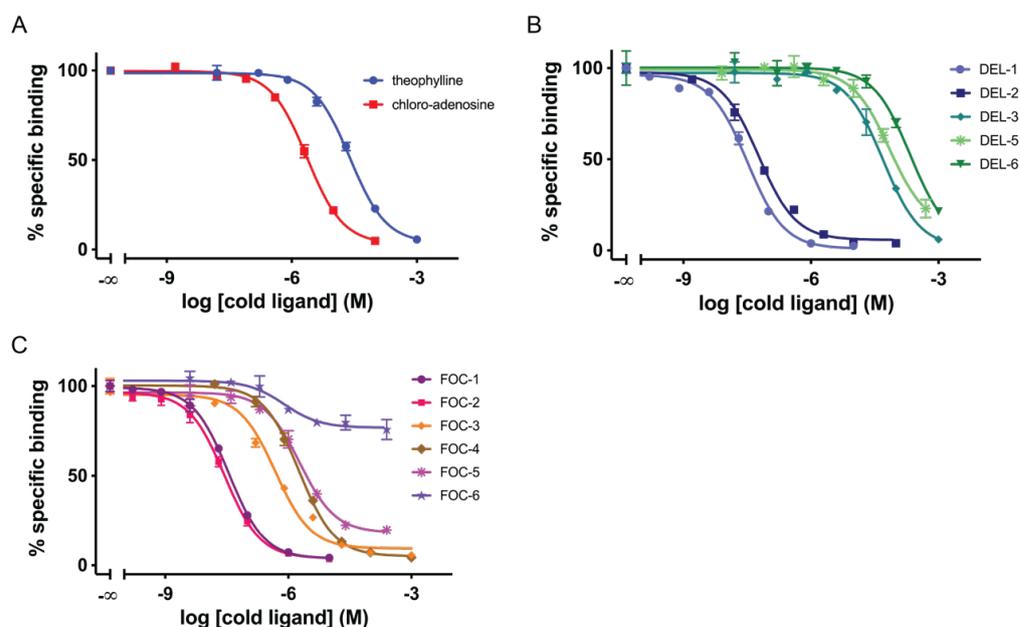


Figure 3.3: Competition binding curves recorded with different ligands. (A) Representative binding curves determined with the reference ligands adenosine and theophylline, (B) the ligands from the DEL screen and (C) the ligands from the focused virtual screen. Experiments were all performed on the same batch of membrane preparations from transiently transfected HEK-cells. Binding curves from one representative experiment are shown. Error bars indicate the SD between three technical triplicates. Curves were normalised against a no-ligand reference. Cold ligand = non-labelled ligand.

Table 3.3: Affinities of novel ligands determined through competition binding. Affinity values are shown for the novel ligands and one reference antagonist and agonist from radioligand competition binding using [^3H]ZM241385 with membranes prepped from HEK293 cells transiently transfected with $\text{A}_{2\text{A}}\text{R}$. All data are presented as mean pIC_{50} from n independent experiments. Errors are given as standard error of the mean. Literature values for the reference compounds are given in parenthesis.

Compound	pIC_{50}	SEM	n
theophylline	4.6 (5.2–5.6)	0.03	9
Cl-adenosine	5.7 (6.8–7.4)	0.05	9
DEL-1	7.3	0.19	4
DEL-2	7.2	0.05	3
DEL-3	4.2	0.03	3
DEL-5	3.9	0.12	3
DEL-6	3.6	0.08	2
FOC-1	7.2	0.15	3
FOC-2	7.3	0.13	4
FOC-3	6.1	0.08	3
FOC-4	5.8	0.03	3
FOC-5	5.7	0.03	4
FOC-6	6.0	0.07	3

FOC-4, FOC-5, and FOC-6 represent slightly better candidates with an pIC_{50} of 5.8 ± 0.03 , 5.7 ± 0.03 and 6.0 ± 0.07 , similar to theophylline and adenosine (Table 3.2). Strikingly, while the data for FOC-6 could be reliably fitted to a sigmoidal binding curve, the curve plateaued at a value significantly greater than non-specific binding, different from all the other compounds (Figure 3.3 C). This indicated that FOC-6 was not able to completely displace the radioligand, which suggests that the ligand might bind to a secondary, possibly allosteric binding site as well. Though, to validate this hypothesis, further experiments need to be performed.

The most compelling compounds according to this first analysis were DEL-1, DEL-2, FOC-1 and FOC-2, which had a pIC_{50} of 7.3 ± 0.19 , 7.2 ± 0.05 , 7.2 ± 0.15 , and 7.3 ± 0.13 . These are very good values for primary hits and therefore made them very interesting candidates for further studies.

3.2.3 Comparison to insect cell expression and effect of different constructs

In order to crystallise the $A_{2A}R$, a stabilised receptor construct is usually engineered and expressed in *Sf9* insect cells. To make sure the ligands would behave similarly under these conditions, ligand binding was assessed using differently stabilised $A_{2A}R$ constructs from insect cell membranes. However, since the amount of compound available was very limited, but structural studies usually require large amounts of compound, these additional binding studies were only performed with the reference compounds.

To be able to better compare the determined binding constants with literature, as a first step, a saturation curve for the radioligand [3H]ZM241385 was determined using a WT-like construct. This construct consisted of the WT receptor but with an N-terminal bRIL fusion to increase expression levels. Successful overexpression of the engineered construct was assessed by Western blotting using HIS-tag and FLAG-tag antibodies for the C- and N-terminal tags of the construct, respectively (Figure 3.4 B). Saturation assays require large amounts of radioligand. Therefore, no experiments were performed on non-infected control membranes. The introduced N-terminal bRIL fusion should have no effect on the pharmacology of the receptor, which could be confirmed by the assay. A pK_D of 9.2 ± 0.05 was determined (Figure 3.4 A), which is in line with what is reported in the literature [188]. This value was then used in subsequent competition assays to determine the pK_i of the competing ligands. A saturation assay further allows to determine the B_{max} of the radioligand, which can then be used to calculate receptor density. For this WT-like construct a B_{max} of 2084 ± 245 dpm could be determined, translating into 19 pmol of receptor per mg of total membrane protein.

In a next step, the same experiment was performed with a construct containing point mutations that stabilise the receptor in the antagonist bound state (StaR2 mutations: A54L, T88A, K122A, V239A, R107A, L202A L235A, S277A, (N145 removal of glycosylation site)) [189]. As for the WT-like construct, expression in *Sf9* cells was confirmed by Western blotting prior to setting up the binding assays (Figure 3.4 B). The stabilised receptor bound the radioligand with a similar affinity ($pK_D = 9.0 \pm 0.02$) to the WT-like construct and the reported literature value [189]. However, the mutations had a notable effect on the expression levels of the receptor and the B_{max} was increased to 8810 ± 1706 dpm (Figure 3.4 C). This resulted in 80 pmol/mg

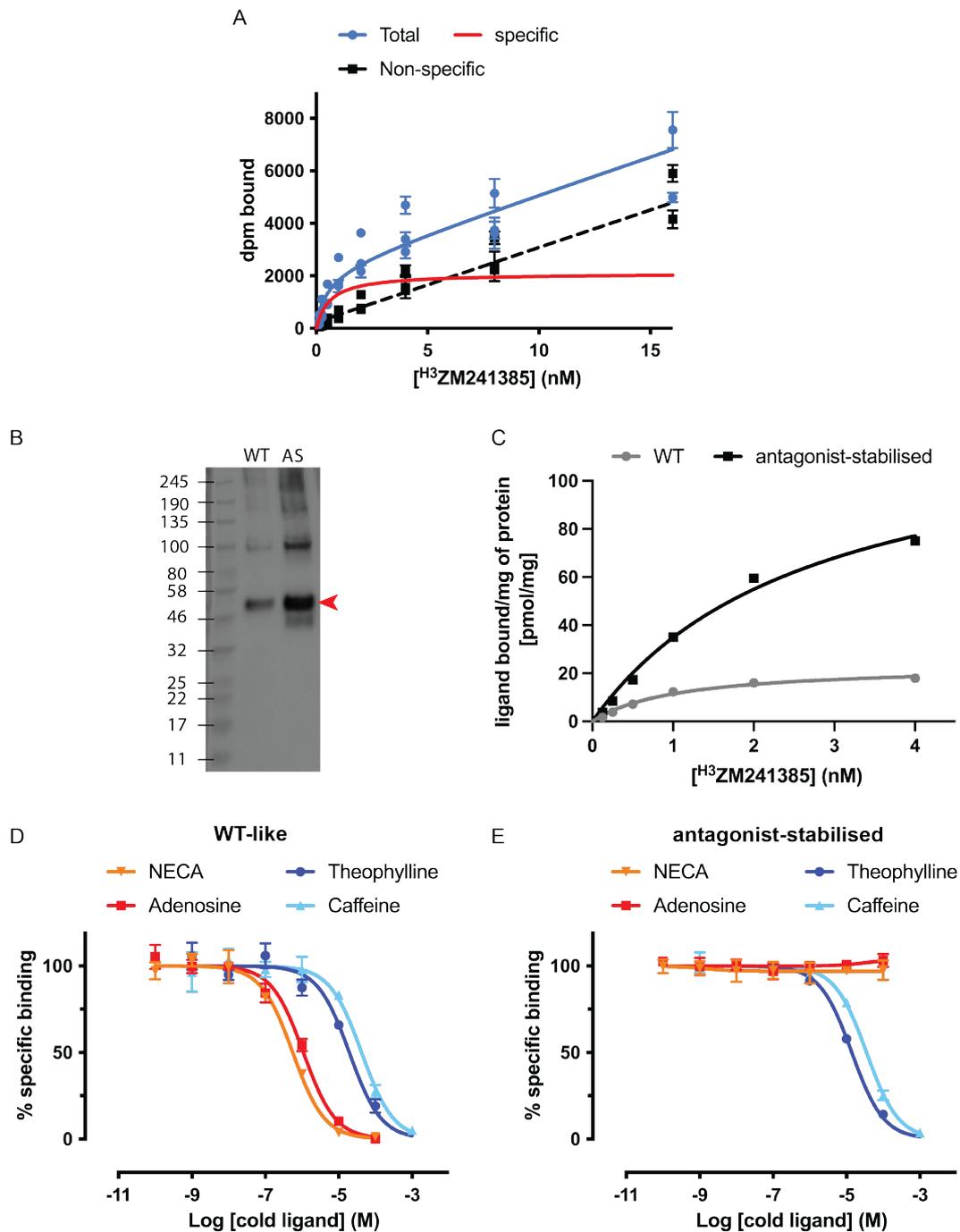


Figure 3.4: Binding assays on stabilised constructs from *Sf9* membranes. (A) Representative saturation binding curves from a single experiment for bRIL-stabilised WT construct using increasing concentrations of [³H]ZM241385. Nonspecific binding was determined in parallel reactions in the presence of 100 μ M theophylline. Specific binding was determined by subtracting the non-specific binding from the total. (B) Western blot of *Sf9* membranes used for subsequent binding assays using an anti-FLAG-tag antibody. Membranes from *Sf9* cells expressing either the WT-like construct (WT) (at 17 μ g of total membrane protein) or expressing a construct containing thermostabilising point mutations that stabilise the receptor in the antagonist bound conformation (AS) (at 21 μ g of total membrane protein) were loaded on the gel. (C) comparison of saturation curves of the WT-like construct and the antagonist-stabilised construct. (D) Competition binding with reference agonists (NECA, adenosine) and antagonists (theophylline, caffeine) using the WT-like construct and (E) the antagonist-stabilised construct. Error bars in the graphs A, C-E represent the SD from three technical replicates. To calculate the K_D and K_i values reported in the text, the mean of 3 independent experiments was taken \pm SEM.

of protein and represented a significant increase in functional receptor reaching the membrane, illustrating one of the reasons why these type of modifications are exploited for crystallisation purposes.

Finally, ligand binding of the reference agonists adenosine and NECA, and the antagonists theophylline and caffeine were assessed in competition binding assays and compared for the two constructs. Similarly, to what was observed in the HEK cell system, the pK_i determined for the reference compounds (NECA: 6.5 ± 0.04 , adenosine: 6.4 ± 0.13 , theophylline: 5.1 ± 0.18 , and caffeine: 4.8 ± 0.14) were slightly lower than reported in literature (Figure 3.4 D). This discrepancy was again higher for the measured agonists than the antagonists, suggesting they are due to differences in the assay conditions. The construct stabilised in the antagonist bound form clearly showed a reduced affinity for agonists, to an extent that no K_i could be determined. In contrast, the affinity for the antagonists remained unchanged compared to the WT-like construct (theophylline: 5.2 ± 0.22 and caffeine: 4.9 ± 0.19) (Figure 3.4 E). This is exactly what has been reported before and suggests that the novel ligands should bind the stabilised constructs expressed in insect cells with a similar affinity to what has been determined in the HEK293-system.

3.3 Determining ligand antagonism

The determined affinities confirmed that the novel ligand specifically bind $A_{2A}R$ but did not disclose anything about the induced downstream effects. Therefore, the compounds were further assessed in a secondary assay, looking at a functional response. As HEK293 cells have been shown to have a relatively high endogenous $A_{2B}R$ expression [247], these experiments were performed in Chinese hamster ovary (CHO) cells.

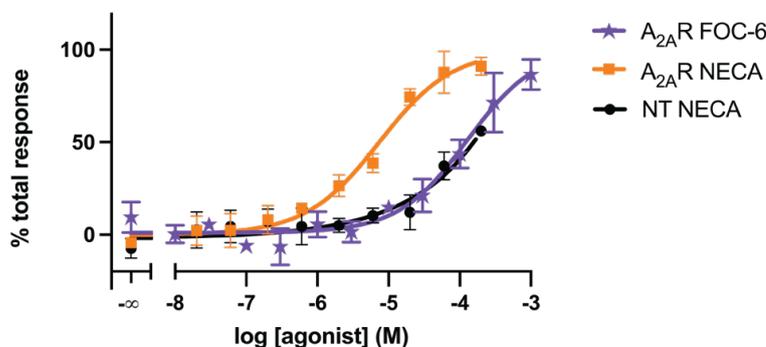


Figure 3.5: Receptor activation upon NECA stimulation of transfected and non-transfected CHO cells. LANCE assay measuring cAMP production upon NECA activation in cells transiently transfected with WT $A_{2A}R$ (orange squares, EC_{50} : $7.4 \mu M$) and non-transfected (NT) CHO cells (black dots). Stimulating the $A_{2A}R$ transfected cells with FOC-6 (purple stars) instead of NECA had no specific effect. Representative curves from a single experiment are shown. Data points were normalised against top and bottom values of the curve fit. Error bars indicate the SD between three technical triplicates.

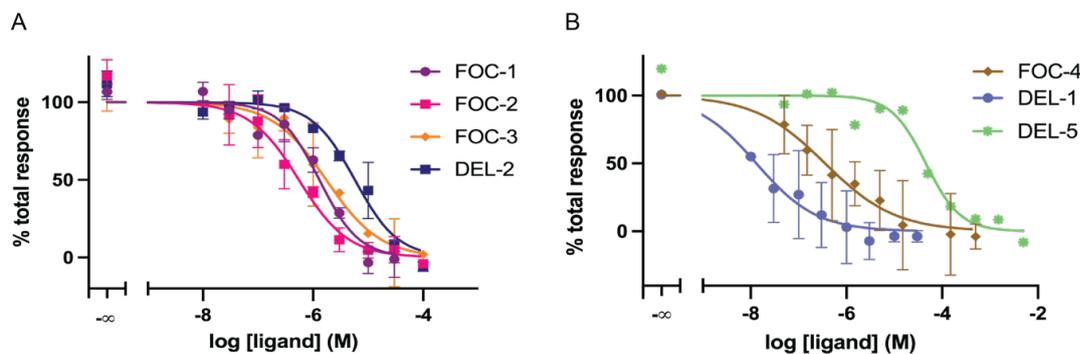


Figure 3.6: LANCE assays measuring cellular cAMP production upon NECA activation of $A_{2A}R$ transfected CHO cells after treatment with different inhibitors. (A) Mean of two or three independent experiments plotted for FOC-1, FOC-2, FOC-3, and DEL-2. Error bars depict the standard deviation between the different experiments. **(B)** Same as A but for FOC-4, DEL-1 and DEL-5, for which the response seemed to be less robust. Because of the limited amount of DEL-5 the experiment could only be performed once. Hence, the mean of triplicates from one single experiment is plotted and no errors could be determined. Experiments for DEL-1, DEL-2, and FOC-3 were performed in duplicates, for FOC-1, FOC-2, and FOC-4 in triplicates.

The $A_{2A}R$ is a prototypical G_s coupled receptor, which means that receptor activation leads to stimulation of AC and an increase in cAMP synthesis. Changes in cellular cAMP levels are routinely used to determine GPCR activation employing fluorescent reporter-based systems such as the LANCE[®] Ultra cAMP assay. This specific assay is a homogeneous time-resolved fluorescence resonance energy transfer (TR-FRET) immunoassay, based on the competition between a europium (Eu) chelate-labelled cAMP tracer and cellular cAMP for binding sites on cAMP-specific monoclonal antibodies labelled with a fluorescent dye. Binding of the labelled cAMP leads to energy transfer to the antibody and a fluorescent readout. Thus, the maximal fluorescent signal is recorded in the absence of any free cAMP, whereas high cellular cAMP production is translated into a low fluorescent readout. To measure the action of antagonists, $A_{2A}R$ was first blocked with increasing concentrations of the inhibitor before stimulating with an agonist at an EC_{90} concentration. The EC_{90} of the agonist was determined empirically for the used system by recording a concentration response curve for the non-selective agonist NECA (Figure 3.5). The assay confirmed that most of the ligands act as $A_{2A}R$ inhibitors (Figure 3.6). Despite the relative high affinities measured, the inhibitory action of most ligand determined in this assay was more moderate. However, one has to note that these results are highly assay dependent and this assay was performed on transiently transfected cells. Therefore, the determined IC_{50} cannot be compared directly with the IC_{50} from the binding assays. Furthermore, most of the ligands contained single and bicyclic aromatic ring systems, which might have interfered with the fluorescent readout to a certain degree, potentially explaining the high error bars on the results for DEL-1, DEL-5 and FOC-4 (Figure 3.6 B). Unfortunately, the appropriate ligand and cell controls could not be included, because of the limited amount of compound available.

According to this assay, DEL-1 was the most potent inhibitor of the novel ligands, with a pIC_{50} of 7.7 ± 0.18 . This value was similar to the affinity determined in the radioactive competition assay. Interestingly, the other ligands of similar affinity were weaker antagonists

Table 3.4: Antagonist potency of novel ligands determined by recording concentration response curves using the LANCE cAMP assay. pIC_{50} are calculated as mean from n independent experiments. Errors are given as standard error of the mean.

Compound	pIC_{50}	SEM	n
DEL-1	7.7	0.18	2
DEL-2	5.2	0.09	2
DEL-5	4.34	-	1
FOC-1	5.9	0.07	3
FOC-2	6.3	0.04	3
FOC-3	5.8	0.26	2
FOC-4	6.5	0.34	3

in this assay; FOC-2 being 10 times and FOC-1 and DEL-2 being even 100 times less potent than DEL-1 (Table 3.4). For FOC-3 and FOC-4 an pIC_{50} of 5.8 ± 0.26 and 6.5 ± 0.34 was determined, which was more in line with the affinity. For DEL-5, one of the ligands with the lowest affinity, the lowest antagonist potency was recorded in this assay ($pIC_{50} = 4.3$). Nonetheless, this specific result needs to be taken with extra caution, as the amount of DEL-5 available was only enough for one single experiment. For DEL-6, DEL-3 and DEL-4 there was not enough material available to perform the experiment at all.

Surprisingly, the results for FOC-5 and FOC-6 could not be fitted to a sigmoidal binding curve, indicating that these ligands either interfered with the assay or simply did not behave as antagonists. These results further added to the hypothesis that FOC-6 does not solely bind to the orthosteric binding site, but potentially might bind a secondary site. The ligand acting as an agonist could be excluded, as no positive concentration-dependent effect could be determined (Figure 3.5). However, to identify a potential allosteric effect of FOC-6 on A_{2A}R signalling more experiments will be required.

3.4 Thermostabilising effect of the compounds

In many cases protein expression and purification can be improved significantly by the addition of a specific ligand that serves to stabilise the protein and thereby reduces its propensity to unfold and aggregate. This effect can be determined as an increase in melting temperature (ΔT_m) in thermal denaturation assays [248]. For GPCRs specifically, ligand binding locks the receptor in a specific conformation and shifts the equilibrium towards a specific state, creating a more homogeneous protein sample, which is required to crystallise the receptor for structural studies. Thus, there is a correlation between a high thermal shift and a positive outcome in crystallographic experiments, which can be used to guide the selection of the most promising compounds for structural studies [249].

For purified membrane proteins, changes in thermostability can be quantified by measuring the extent of unfolding using a variety of different techniques, such as assessing protein activity as a function of temperature using radioligands [189], fluorescent size exclusion [250] or differential scanning fluorimetry (DSF) either using specific probes such as thiol-specific fluorophore N-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]maleimide, which fluoresces when

covalently bound to internal cysteines that become exposed upon denaturation [251] or intrinsic tryptophan fluorescence [252]. However, most of these techniques require very specific equipment or specifically labelled compounds or protein constructs. In contrast, SDS-gel based assays allow the study of thermal stability of membrane proteins in every standard biochemistry laboratory. If followed by Western blotting, or if using a fluorescently tagged protein construct, thermal stability can even be assayed using total cell lysates instead of purified protein [253].

3.4.1 Melting temperature of receptor-ligand complexes determined using a gel shift assay

As the gel shift assay does not require any specific equipment or labels, this assay was chosen as the first strategy to determine the thermostabilising effect of the novel ligands. Together with the affinity data, this information should help to select the best candidates for subsequent structural studies.

Despite being one of the simplest methods to assess thermostability, the gel shift assay is relatively error prone and less sensitive than other methods. Therefore, a WT-like receptor construct with minimal thermostabilising modifications was chosen, which would hopefully allow better visualisation of small thermostabilising effects. However, as *apo*-WT-A_{2A}R was too unstable for purification, the receptor had to be purified in the presence of the transient, low-affinity ligand theophylline. Before setting up the thermal unfolding experiments, affinity purified receptor was buffer exchanged using a PD10 column to remove any excess ligand. For each ligand nine aliquots were prepared, which were preincubated with 100 μ M of the novel ligands for 30 minutes at room temperature. Subsequently, the aliquots were heated to eight different temperatures, ranging from 55–80 °C; a control sample was kept at 4 °C. After temperature treatment, SDS-sample buffer was added to each condition and all samples were loaded on an SDS-PAGE. Unfolded protein formed large aggregates that were unable to run

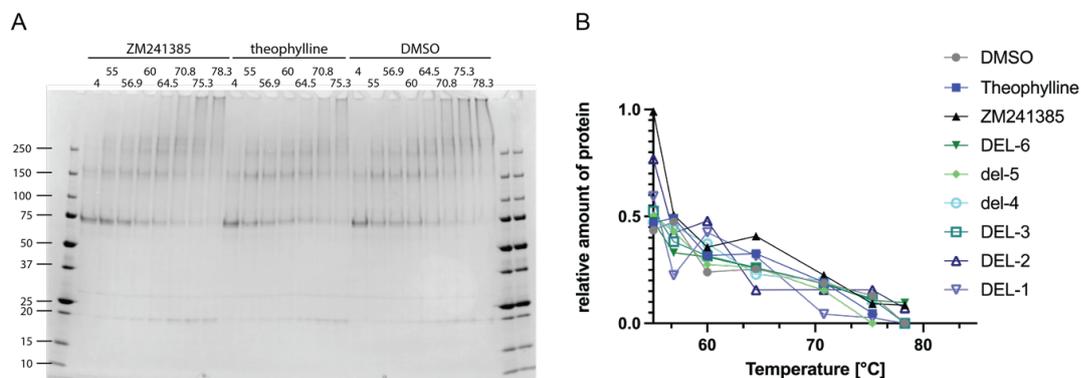


Figure 3.7: Attempt of trying to determine thermostabilising effects using the gel shift assay. (A) SDS-PAGE showing the gel-shift assay for ZM241385, theophylline and the DMSO control. (B) Despite seeing a clear decrease in signal on the gel, when plotting the relative amount of protein in a temperature dependent manner, no clear difference between the ligand-stabilised receptor and the DMSO control could be determined.

into the gel. Therefore, the intensity of the protein band after Coomassie staining decreased in a temperature dependent manner, which could be quantified through subsequent densitometry analysis of the gel (Figure 3.7).

Unfortunately, no thermostabilising effect could be determined using this method, not even for the reference compound ZM241385, for which a ΔT_m of up to 12 °C has been previously reported [188].

3.4.2 Determining the melting temperature of receptor-ligand complexes by nanoDSF

As the Gel shift assay was not suited to analyse the thermostabilising effect of the here studied A_{2A}R ligands, another method was chosen. DSF experiments record changes in fluorescence occurring upon protein unfolding. To follow this process in a label-free manner, the intrinsic fluorescence of tryptophan can be used, which is strongly dependent on its close surroundings. Tryptophan residues buried inside a protein structure and thus, located in a hydrophobic environment, have a fluorescence emission maximum at round 330 nm, while in a more polar surrounding, such as encountered upon unfolding, fluorescence is quenched by solvent molecules and the emission maximum is red-shifted to about 350 nm [254, 255]. By following the change in fluorescence emission in a temperature dependant manner protein unfolding can be monitored and thereby the thermal stability of a protein assessed. The apparent melting temperature (T_m) of a protein is defined as the temperature required for the unfolding of 50 % of a protein, corresponding to the midpoint of the transition from folded to unfolded, which can be calculated from the inflection point from a sigmoidal curve. An increased melting temperature of a specific receptor-ligand complex indicates a more stable and rigid protein conformation and allows to preselect the ligand with the highest likelihood for a positive outcome in crystallisation experiments. Therefore, the same thermostabilised protein sample prepared for crystallographic experiments was also used for DSF experiments. To generate an “*apo*-protein” sample the protein solution was diluted 200 times with ligand free buffer, thereby hopefully releasing most of the bound ligand. Then, 50 μ M ligand were preincubated with 3.5 μ M receptor for half an hour at room temperature before setting up the DSF experiment. As negative-control, DMSO alone was used. Melting scans were performed covering a temperature range from 20–90 °C at a heating rate of 2.5 °C/min. As the fluorescence emission maximum of tryptophan depends on its immediate surroundings the optimal wavelength to record changes in fluorescence emission might vary for different proteins [256]. The Prometheus nanoDSF device therefore records fluorescence emission at 330 nm (F_{330}) and 350 nm (F_{350}). When following the red-shift in fluorescent emission due to protein unfolding, the 350:330 nm ratio is often suited best to detect any changes in protein structure, and was therefore used here to calculate the T_m for the different receptor-ligand complexes. The T_m is then visualised by plotting the first derivative of each measurement, where the curve minimum corresponds to the inflection point of the melting curve (Figure 3.8 A). Without the addition of any ligand (DMSO-sample), the apparent T_m was 56 °C. The thermostabilising effect of each ligand (ΔT_m) was then determined by relating the T_m measured for the different ligand-receptor complexes to this non-ligand reference (Figure 3.8 B). As expected, ZM241385 conferred the largest thermostabilising effect, increasing the melting temperature by 4 °C. Unfortunately, only minor changes in T_m were observed for all

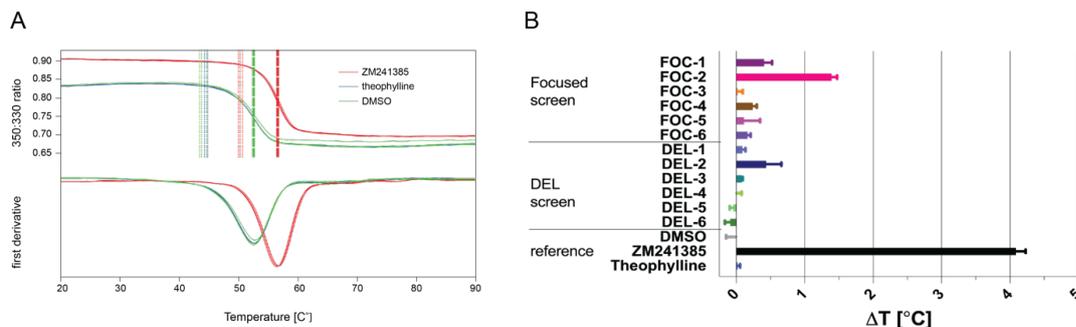


Figure 3.8: Thermostabilising effect of ligands recorded by nanoDSF. (A) Exemplary melting curves obtained for ZM241385, theophylline and the DMSO control. The top panel shows the curves obtained at the 350:330 ratio, the bottom the first derivative, where the curve minimum visualises to the T_m at the inflection point of the above melting curves. **(B)** Comparison of ΔT_m induced by the different ligands.

the novel ligands; some even displayed a slightly destabilising effect (Figure 3.8 B). According to the recorded ΔT_m , FOC-2 seemed to be the best candidate for crystallographic studies out of this ligand series, followed by DEL-2 and FOC-1. This nicely correlated with the determined affinity, as these compounds also displayed the highest pIC_{50} . Interestingly, despite having a similar affinity, DEL-1 seems to stabilise the receptor less than DEL-2. DEL-6 and DEL-5 had a slightly destabilising effect on the receptor, reducing T_m compared to the DMSO-control. Only minor changes in ΔT_m could be observed for the rest of the compounds. Albeit most of them having higher affinities than the reference compound theophylline, their ΔT_m was not significantly different.

3.5 Discussion

A set of novel $A_{2A}R$ ligands was identified through two different, novel screening approaches and was further validated to determine their potential as lead candidates. To be able to rationally enhance the ligands in subsequent optimisation experiments, as much pharmacodynamic data as possible should be gathered that helps to identify and assign the functional contributions of each structural element of the ligand. Structural characterisation of membrane protein-ligand interactions is still very difficult and usually highly laborious. Pharmacological and biophysical data helps to select the candidates that are most likely to be successful in structure determination approaches and on which efforts should be focused.

One important aspect to consider, specifically when designing inhibitors, is ligand affinity, as they are required to tightly bind to the receptor in order to interfere in the effects of the native ligand. For all but one ligand, reliable binding curves could be recorded, confirming that they specifically bind the $A_{2A}R$. Four ligands, DEL-1, DEL-2, FOC-1 and FOC-2, had an affinity in the sub-micromolar range, indicating that they are good lead candidates. By monitoring cAMP levels, I further confirmed that these ligands are indeed antagonists. Ligands with a high affinity also inhibited the receptor with a higher potency than the compounds of weaker affinity.

FOC-2 is a 1,3,5-bioisostere of previously characterised triazine ligands [217]. It was therefore not surprising that this ligand had a high affinity too. Similarly, FOC-1 contains the typical bicyclic core which is present in many A_{2A}R ligands. Together with the relatively high affinity, this suggested that the ligand probably adopts a similar binding pose as classical A_{2A}R ligands such as ZM241385. However, the FOC-1 core consists of a dideazapurine, containing significantly fewer nitrogen substitutions that could engage in hydrogen bonding, which indicates that the aromatic substituents might contribute important interactions for binding.

In contrast, DEL-1 and DEL-2 were structurally very distinct from previously characterised ligands and are thus especially interesting hits for subsequent structural characterisation. Despite having the same ligand core as DEL-2, the affinity of DEL-5 was reduced by more than one order of magnitude. Further, initial results suggested that the potency of DEL-5 was significantly reduced as well. This indicated that the bulky morphonyl-propylpropanamide moiety was important for binding and function and is therefore likely to directly interact with the receptor.

For FOC-3, FOC-4, FOC-5, and FOC-6 affinities in the lower micromolar range could be determined, suggesting they were still better binders than the nonspecific antagonist theophylline. Interestingly, while the potency of FOC-3 and FOC-4 was in a similar range than what was measured for the best ligands, FOC-5 and FOC-6 did not inhibit receptor activation by NECA. Additionally, FOC-6 was not able to completely displace the radioligand, suggesting this ligand might bind to a secondary, potentially allosteric site on the receptor.

Finally, DEL-3, DEL-5, and DEL-6 were of relatively low affinity. Because I identified stronger binders in this selection of hits, they were not further considered as interesting drug candidates. It is usually more difficult to get structures with low affinity compounds. Therefore, the high affinities determined for DEL-1, DEL-2, FOC-1 and FOC-2 suggest higher chances for getting co-crystal structures with the A_{2A}R.

A second factor that helps to judge if a receptor-ligand complex might crystallise, is the thermostabilising effect of the ligand on the receptor. Because the thermostabilising effects of this novel series of compounds was relatively low, the gel shift assay was not sensitive enough to detect these changes. Sensitivity of the assay might be increased through a series of assay-optimisation steps, but to obtain good signal intensity this would probably require larger amounts of sample and therefore also compound, which were not available. Therefore, nanoDSF measurements were performed. They facilitated the determination of ΔT_m inferred by the different compounds onto the exact receptor sample used for subsequent crystallographic experiments. This method was also quicker and more labour efficient than the gel shift assay; however, it is highly dependent on the availability of specific equipment.

The apparent melting temperature determined for the *apo*-receptor was very high. On one hand, this was because a highly stabilised receptor construct has been used. The thermostabilised StaR2-construct already has a melting temperature of 47 °C [189] and the BRIL-fusion in the third intracellular loop must have increased the stability even further. On the other hand, the protein was purified in the presence of the high-affinity ligand ZM241385. Despite diluting the protein sample 200 times to remove the ligand, some might still have remained bound to the receptor, stabilising a certain fraction of receptor molecules in the sample and therefore increasing the apparent T_m of the non-liganded control. Thus, the observed thermostabilising

effect of any added ligand would appear lower than it actually is, explaining why the ΔT_m for the reference ligand ZM241385 measured in this experiment is lower than reported elsewhere [188, 189]. Weaker binders might not have been able to completely displace residual ZM241385 from the binding site and therefore, no effect on T_m could be observed. Hence, the absolute ΔT_m values are likely to be slightly higher than recorded here. Despite being smaller than expected, the ΔT_m correlated with the affinities of the compounds. This indicates that the observed changes are real and conferred by to the specific compound. Consequently, they could be used to rank the compounds according to their likelihood of stabilising the receptor for crystallographic experiments. However, as even the ΔT_m of the best candidate was significantly lower than the effect of ZM241385, this experiment suggests that it might be quite difficult to crystallise any of these receptor-ligand complexes.

Chapter 4

In silico structural studies of A_{2A} receptor-ligand interactions

Before attempting to obtain co-crystal structures with these novel ligands, the ligand- receptor interactions were predicted using computational docking programs. Docking poses can provide initial models of interaction between small molecules and proteins at atomic levels, which can facilitate the characterisation of small molecule behaviour at a protein binding site [257, 258]. As the novel ligands studied in this research are all hits from primary screening efforts, it is likely that few of them, if any at all, will crystallise in complex with the receptor. Thus, if no structures can be obtained, *in silico* analysis can serve as the basis to derive interaction fingerprints of the ligands [259, 260], as well as to construct quantitative structure-activity relationship (QSAR) models, which will inform chemical refinement of the ligands [261, 262, 263].

4.1 Validation of docking approach

Considering the limitations of computer resources, docking was performed with a flexible ligand and a rigid receptor. As the agonist and antagonist bound state of GPCRs differ significantly for most receptors, it was important to use the appropriate receptor conformational state as the basis for ligand binding simulations. The LANCE assay confirmed that most ligands exhibited an inhibitory activity and hence an inhibitor bound structure was chosen for model based structural studies.

To validate the docking approach, the reference ligands, ZM241385 and theophylline, were docked in a benchmark test [264]. Redocking a ligand that has previously been co-crystallised into the cognate receptor is one of the most highly accepted measures of docking success, as the root mean square deviation (RMSD) from the crystal pose can be measured to determine the accuracy of the predicted pose [265, 258]. If there are several co-crystal structures available, an additional benchmark can be established by evaluating the ligand RMSD from the predicted and the “known” poses. In this case the “known” pose is simply obtained by aligning the reference structure and the ligand’s co-crystal receptor structure [266].

To dock the different ligands into the A_{2A}R structure the docking program Glide from Schrödinger was used. Glide initially generates a series of different poses, which are then passed through multiple hierarchical filters, to find the energetically most probable combination of conformation and orientation of a ligand relative to the receptor. These filters test the spatial fit of the ligand to the defined active site and examine the complementarity of ligand-receptor interactions based on empirical scoring functions [267] (See section 1.1.2). Poses that pass these initial screens enter the final stage of the algorithm, which involves the evaluation and further minimisation of ligand-receptor interaction energy. Only these energy-minimised poses are then used to calculate the final docking score, approximating the ligand binding free energy. This score is based on several factors, including electrostatic and vDW contributions, the nonbonded interaction energy, and other terms rewarding or penalising interactions known to influence ligand binding (See section 1.1.2).

For both ligands there is a crystal structure available to which the docking results could be compared. Thus, ZM241385 and theophylline were docked twice, once into a structure solved with ZM231385 (4EIY) and once into the structure determined in complex with theophylline (5MZJ). The overall protein structure of the A_{2A}R bound to these two different ligands is very similar (C α RMSD: 0.28), including all the residues in the orthosteric binding pocket [221].

To determine the RMSD between the different ligand poses, DockRMSD [230] was used. This algorithm maps the two structures without the assumption of atomic ordering and then calculates the RMSD for each atomic mapping and determines the lowest score. For most benchmark tests, RMSD values below 2 Å are considered good scores [258]. Docking ZM241385 into the 4EIY structure resulted in an almost identical ligand pose as observed in the crystal structure (RMSD: 0.22 Å) and a high docking score of -10.6 (Figure 4.1 A). The docking software Glide uses composite docking scores, combining force field-based and knowledge-based scoring functions that are highly complex. However, generally a score of -10 or lower usually represents good binding. For targets with shallow active sites, or which predominantly form hydrophobic interactions, scores of -8 are considered very good. Glide predicts that ZM241385 forms interactions with Phe 168^{45,52}, Glu 169^{45,53}, His 250^{6,52}, and Asn 253^{6,55} which aligns well with the crystal structure. Hydrophobic interactions such as with Leu 267^{7,32}, Met 270^{7,35}, and Ile 274^{7,39} are considered for the calculations but will be discussed to a much lesser extent here.

Similarly, docking theophylline into either the 4EIY or the 5MZJ structure resulted in poses that closely match the crystal structure (RMSD of 0.29 Å and 0.21 Å respectively) (Figure 4.1 B, D). The interaction of the ligand with His 278^{7,43} and the backbone carbonyl of Ala 81^{3,28} described in the crystal structure are not visible, as they are made through two water molecules that were not included in the docking structure to increase the speed of docking. Nonetheless, docking theophylline into the theophylline-bound receptor structure (5MZJ), did lead to a slightly better docking score of -8.0 compared to -7.7 when docking into the ZM251385-bound structure. The docking score obtained for theophylline was significantly poorer than for ZM241385, which corresponds well with the weaker affinity of theophylline.

The biggest difference between a docked ligand pose and a crystal structure was observed when docking ZM251385 into the theophylline structure (5MZJ), as the phenol moiety extending from the ethylamine chain of ZM241385 adopts a different orientation (Figure 4.1 C). Large substituent flexibility has been described previously for this area of the pharmacophore [268]

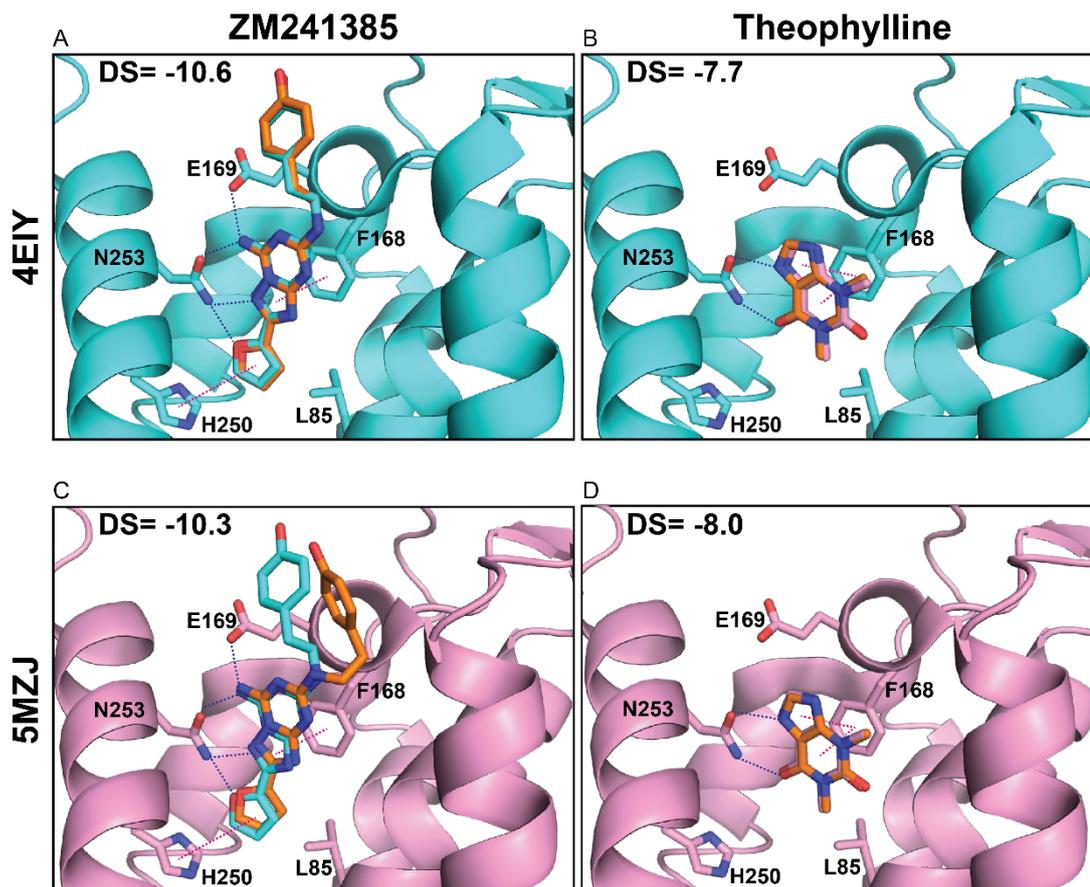


Figure 4.1: Comparison of docking poses for reference compounds to crystal structures. The docked compound positions are shown in orange, while the 4E1Y protein structure and bound ZM241385 are shown in cyan and the 5MZJ and bound theophylline from this structure in pink. Docking scores are denoted DS. Hydrogen bonds are shown as blue dashed lines and aromatic stacking interactions in red. (A) ZM241385 was docked into 4E1Y structure and compared to the bound ZM241385 (cyan). (B) Docking of theophylline into the same structure, compared to the ligand position of aligned 5MZJ (pink). (C) ZM241385 docked into 5MZJ compared to the ligand position of aligned 4E1Y (cyan). (D) Theophylline docked into the 5MZJ structure compared to the bound ligand.

and in the original crystal structure this region had higher temperature factors than other parts of the ligand, reflecting its flexibility [188]. Furthermore, an alternative position of the phenol group can be observed in another ZM241385-bound structure (3PWH) [189]. However, the bicyclic core, as well as the furan moiety of the ligand, form the same π -stacking and hydrogen bond interactions with the receptor in the docked model as in the crystal structure and are positioned very similarly (overall RMSD: 1.76 Å).

Overall, the A_{2A}R structure used for docking did not have much of an impact on the resulting ligand poses. The high similarity of the poses obtained with either protein structure was also reflected in the very similar docking scores.

4.2 Docking of novel ligands

Following the benchmarking test using known ligands, the selection of novel ligands from Novartis was docked into the prepared 4E1Y structure using Glide. The obtained docking scores roughly correlated with the pIC_{50} previously determined in the radioligand binding assays (Figure 3.2). Ligands with a tighter dissociation constant than theophylline ($pIC_{50} > 4.6$) were assigned better docking scores (< -8), while the ligands for which $pIC_{50} < 4.6$ was measured (including theophylline), a docking score > -8 was obtained (Figure 4.2). The only exception was DEL-2, for which docking resulted in one of the poorest docking scores, despite previously showing one of the tightest affinities in the experimental binding assay. The best docking score was obtained for FOC-2, which is also one of the compounds with the tightest affinity. The scores for FOC-2 and FOC-3 were even slightly better than the docking score obtained with the original ligand ZM241385, even though their affinity is more than 50 times lower. Interestingly, these compounds both contain a chemical scaffold that was identified in previous docking studies [209, 210, 211].

The receptor residues known to be important for ligand binding were also engaged in ligand interactions in the present docking study (Table 4.1). These were predominantly Phe 168^{45.52}, Glu 169^{45.53}, His 250^{6.52}, and Asn 253^{6.55}. Most of the ligands were predicted to form a π -stacking interaction between their aromatic core and Phe 168^{45.52}. Additionally, all ligands with a docking score < -8.5 were further stabilised by a second π -stacking interaction with either His 250^{6.52} or Tyr 271^{7.39}. The ligands not predicted to form this type of interaction with

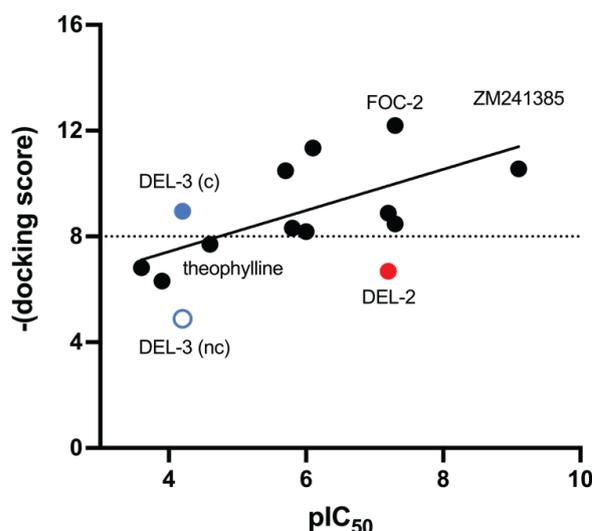


Figure 4.2: Correlation of docking scores with the previously determined affinities. Docking scores for the different ligands were plotted against their pIC_{50} values determined in radioligand competition assays ($R^2 = 0.5$). Ligands with a higher affinity than theophylline were assigned docking scores < -8 . The only exception is DEL-2, which had a relatively poor score assigned despite great affinity (filled red circle). DEL-3 was the only compound for which the score significantly changed after introducing a knowledge based positional constraint (discussed in chapter 4.2.3). The score obtained with this constraint is shown as filled blue circle (c) and the one without constraint as open blue circle (nc).

Table 4.1: Docking scores and residues predicted to form polar interactions for ligand docking into A_{2A}R (4E1Y). The list of ligands is sorted according to the highest Glide docking score. Residues predicted to be involved in polar interactions are listed according to interaction type as π -stacking interactions, hydrogen (H-) bonds, and salt bridges. If nothing else is stated, the interaction is formed by the main functional group of the residue, otherwise the involved group is stated in squared brackets (NH₂ = amino group, O = hydroxy group, bb = backbone).

Ligand	Score	Interactions		
		π -stacking	H-bonds	Salt bridge
FOC-2	-12.2	F168, Y271	N253 [O], E169	-
FOC-3	-11.3	F168, Y271	N253 [NH ₂ , O], E169	-
ZMA	-10.6	F168, H250	N253 [NH ₂ (2x), O], E169	-
FOC-5	-10.5	F168, H250, Y271	N253 [NH ₂ , O], E169	-
FOC-1	-8.9	F168 (2x), Y271	-	H264 (π -cation)
DEL-1	-8.5	-	E169 [bb NH ₂], Y271, S67 [bbO]	-
FOC-4	-8.3	F168	N253 [NH ₂ , O], E169	-
FOC-6	-8.2	F168	N253 [O], E169	-
Theo	-7.7	F168 (2x)	N253 [NH ₂ , O]	-
DEL-4	-7.5	F168	E169 [bbNH ₂]	E169
DEL-6	-6.8	-	K150, Y271, E169	K150, H264 (π -cation)
DEL-2	-6.7	F168	S67 [bbO], E169 [O, bbNH ₂], Y271	N253 [NH ₂] (halogen), K150 (π -cation)
DEL-5	-6.3	H264	E169 [O, bbNH ₂]	H264 (π -cation)
DEL-3	-4.9	-	E169 [O, bbNH ₂], F168 [bbNH ₂]	K150 (π -cation)

Phe 168^{45,52}, DEL-6, DEL-5 and DEL-3, generally had docking scores > -8 . The only exception was DEL-1, which had a docking score of -8.5 , despite being predicted not to form any aromatic stacking interactions. Furthermore, most of the ligands were predicted to form hydrogen bonds with either Glu 169^{45,53}, Asn 253^{6,55}, or both. In addition, DEL-1 was predicted to form an additional hydrogen bond with Tyr 271^{7,39}, as was DEL-6 with Lys 150^{45,34}.

Interestingly, in their docked poses, DEL-1 and DEL-2 were the only ligands that formed a polar interaction with a residue other than on helix VI and VII or in the loop region, which was mediated through the backbone carbonyl of Ser 67^{2,64} (Figure 4.3). The only ligand not predicted to form any hydrogen bonds with the receptor was FOC-1. In turn, the docked pose showed a π -cation interaction with His 264^{7,28} (Figure 4.3). Remarkably, all other ligands forming either π -cation or salt bridges were assigned a docking score lower than theophylline.

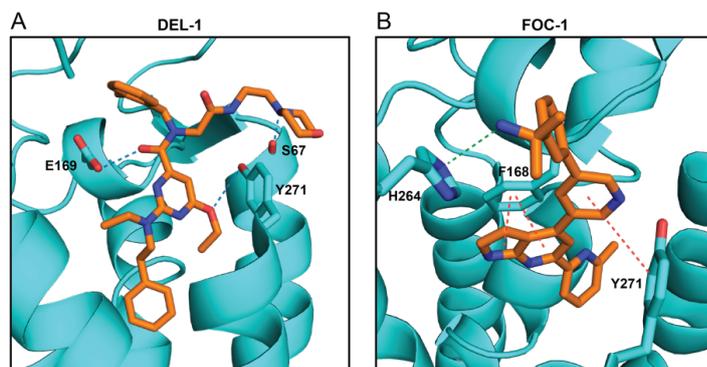


Figure 4.3: Docking poses for DEL-1 and FOC-1. (A) Docking of DEL-1 and (B) Pose of FOC-1 into the 4EIY structure. Hydrogen bonds are shown as blue, aromatic stacking interactions as red, and salt bridges as green dashed lines. Residues involved in binding are shown as sticks. For clarity helix VII is hidden above Y271.

4.2.1 Comparison of predicted binding for different bioisosteres

Several of the selected screening hits are similar to known $A_{2A}R$ ligands and result from the exchange of a few atoms, or groups of atoms, with alternative, broadly similar, atoms or atom groups. However, they likely still have very similar biological properties and are therefore regarded as bioisosteres (see chapter 3.1). For several of these $A_{2A}R$ bioisosteres, there are co-crystal structures available. Thus, while these hits do not provide new ligand scaffolds, the comparison with the bioisostere structure can be used to further validate the docking approach and increase the confidence in the obtained poses. Furthermore, this comparison might help to analyse which groups and substituent arrangements might be important for receptor interaction.

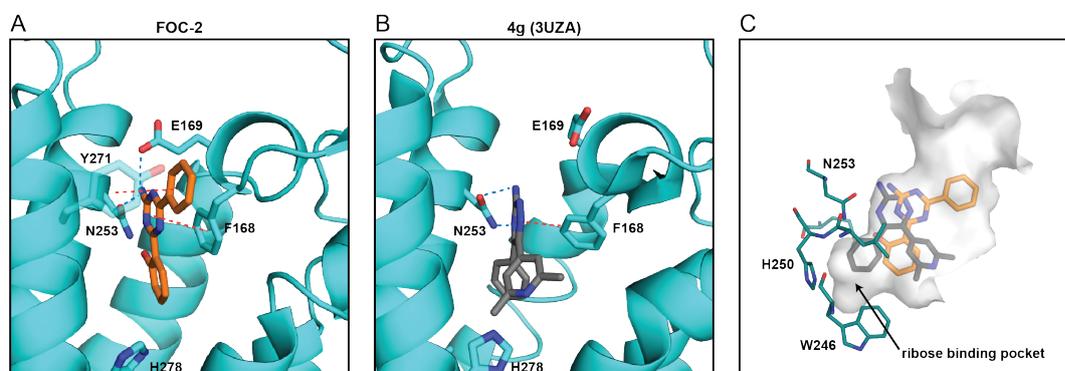


Figure 4.4: Comparison of the docked pose for FOC-2 to the binding pose of the 1,2,4-triazines 4g in the co-crystal structure 3UZA. (A) docking of FOC-2 into the 4EIY structure. (B) Conformation of 4g from the 1,2,4-triazine ligand series taken from the 3UZA structure, shown in the 4EIY structure. For clarity, helices I and II are hidden. Hydrogen bonds are shown as blue, aromatic stacking interactions as red, and salt bridges as green dashed lines. Residues involved in binding are shown as sticks. (C) Comparison of the positions of the 1,2,4-triazine 4g from the co-crystal structure 3UZA and the docked pose of FOC-2 inside the $A_{2A}R$ binding pocket. For orientation Asn 253 and His 250 are represented as sticks outside the surface. Trp 246 is lining the bottom of the binding pocket. The ribose binding pocket is indicated by an arrow.

FOC-2 is a 1,3,5-isomer of the previously published 1,2,4-triazine series, out of which two ligands, termed 4g and 4e, were co-crystallised with the receptor (3UZA, 3UZC) [217]. Comparison of the receptor conformation in the optimised 4EIY structure used for docking with the conformation in the crystal structures 3UZA and 3UZC, showed the receptor in an identical conformational state, with most ligand-binding residues in the same position, rotameric state, and an average C α RMSD between the structures of 0.688 Å and 0.685 Å respectively. Similar to the crystal structures, the amino-triazine core of FOC-2 was predicted to form a π -stacking interaction with Phe 168^{45.52} and hydrogen bonding interaction with the sidechain of Asn 253^{6.55} (Figure 4.4). However, the docking model suggested that the wider spacing of the two aryl pendants in the 1,3,5-triazine isomer hinders the ligand from penetrating deeper into the binding pocket, and therefore, from optimally accessing the region occupied by the ribose moiety of the natural ligand adenosine (Figure 4.4 C). However, the different isomeric state of the amino group on the triazine core allowed the formation of an additional hydrogen bond with Glu 169^{45.53}, which is not reported for the 1,2,4-isosteres.

The second aryl ring was predicted to point towards the solvent exposed opening of the binding site, forming a second π -stacking interaction with Tyr 271^{7.35}. The 1,2,4-triazines possess high ligand efficiencies as they occupy the region where there is a cluster of “unhappy” waters [217]. This is translated into high affinity and slow off rates of the ligands. Thus, despite these additional polar interactions of FOC-2, the less favourable positioning in the binding site is a potential cause of its lower affinity and probably faster off rates.

Similar to FOC-2, the triazine core found in FOC-3 was identified in previous docking studies [210]. Furthermore, a similar scaffold can be found in A_{2A}R co-crystal structure with Cmpd-1 (4UIG), a compound that was originally designed to have dual activity against A_{2A}R, as well as NR2B [193]. However, in FOC-3, the first moiety is a fluorobenzene instead of a methoxyphenyl, which is linked to the core by an ester and not an amide, while the methylphenyl-group is substituted by a phenyl-moiety (Figure 4.5 A). Docking of FOC-3 predicted the same aromatic stacking interaction with Phe 168^{45.52} and a hydrogen bond between E 169^{45.53} and the amino-triazole core as observed in the crystal structure of Cmpd-1 (Figure 4.5 B, D). By contrast, in its predicted pose, FOC-3 was flipped upside-down compared to Cmpd-1, with the longer fluorobenzene moiety pointing towards the receptor core. While the shorter phenyl substituent was still positioned to form a second π -stacking interaction with Phe 271^{7.35}, it could not extend into the potentially allosteric secondary pocket formed by Tyr 9^{1.35}, Ala 63^{2.61}, Ser 67^{2.65}, Leu 267^{7.32}, Met 270^{7.35}, Tyr 271^{7.36}, and Ile 274^{7.39}, which is occupied by methoxyphenyl pendant of Cmpd-1. Thus, according to this prediction, FOC-3 could not fully explore this secondary binding site, which might explain the less favourable binding properties compared to Cmpd-1.

The comparison of the docking pose obtained for FOC-2 to the co-crystallised bioisostere revealed a similar pose of the main scaffold, thus further validating the docking approach. However, FOC-3 was docked upside-down compared to the bioisostere Cmpd-1. This might indicate that the main interactions predicted between the ligand core and the orthosteric ligand binding site might not be enough to fully predict the orientation of the ligand. Though, the different orientation might also be a real effect, caused by the slight alterations in the terminal substitutions of the two ligands. In this case the comparison highlights the effects that the

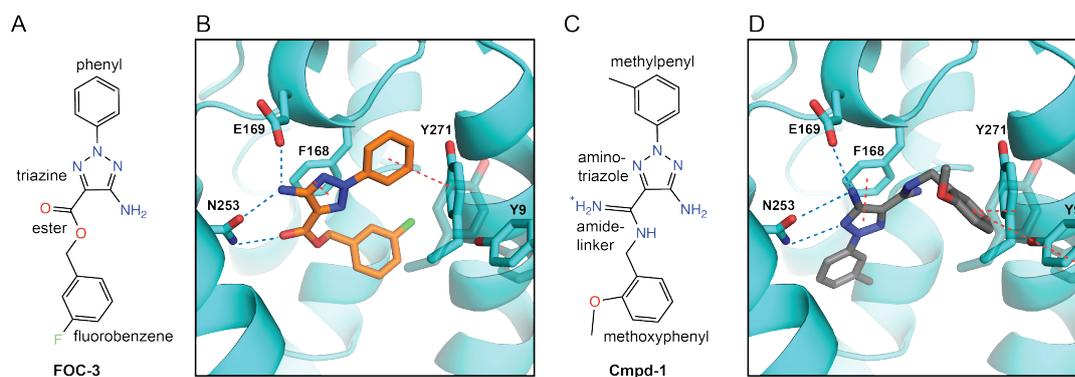


Figure 4.5: Comparison between the docking of FOC-3 and the co-crystal structure of triazole compound Cmpd-1. For clarity helix VII is hidden. Hydrogen bonds are shown as blue, aromatic stacking interactions as red, and salt bridges as green dashed lines. Residues involved in binding are shown as sticks. The alternative rotamers and positions of Y271 and Y9 adopted in the triazole co-crystal structure (5UIG) are shown in darker, transparent sticks. **(A)** Docking of FOC-3 into 4EIY structure. **(B)** ligand pose of Cmpd-1 taken from the 5UIG co-crystal structure.

chemical differences between the bioisosteres had on the interactions made with the receptor. The example of FOC-2 nicely illustrates how ligand isomerisation could be used to better exploit the deeper regions inside the binding pocket, which helped to enhance ligand properties [217]. I therefore wanted to evaluate if similar conclusions about the contributions of different ligand moieties to receptor binding could be drawn for some of the ligands with a novel scaffold.

4.2.2 Comparison of docking poses of ligands containing the same chemical scaffold

Interestingly, the investigated selection of ligand hits also contained some ligands that originate from the same ligand series, as they contain the exact same scaffold, only differing in the composition of one single moiety. The fact that a specific scaffold is picked up twice in the original screen indicates that this ligand scaffold seems to be favourable for $A_{2A}R$ binding. Furthermore, comparison of the docked poses of such ligand pairs directly allows the analysis of the effect of different substituent on receptor binding, which might contribute to subsequent ligand optimisation efforts.

DEL-2 and DEL-5 both contain the same scaffold, with a chlorinated pyrimidine moiety and an indole pendent group. However, the methyl moiety on the central nitrogen on DEL-5 has been substituted by a bulky morpholinyl-ethyl-acetamide moiety on DEL-2 (Figure 4.6 A, D). Both docked ligands were oriented the same way, with the pyrimidine moiety facing into the binding pocket (Figure 4.6). However, the bulky substituent of DEL-2 meant the molecule was predicted to orient itself in a way that the pyrimidine moiety could form a π -stacking interaction with Phe 168^{45.52}, while the chloride substitution on the pyrimidine ring engaged with Asn 253^{6.55} (Figure 4.6 B). Furthermore, the ligand core formed hydrogen bonds with Glu 169^{45.53}. The long morpholinyl-ethyl-acetamide substitution was predicted to face towards

helix I, forming contacts with almost all residues in the potential allosteric pocket described with Cmpd-1 (Figure 4.6 C). Thus, DEL-2 might display a more similar binding mode to Cmpd-1 than FOC-3, despite the different core.

In contrast to DEL-2, docked DEL-5 was rotated around the ligand axis by about 180°. The ligand was positioned closer to the entrance to the binding pocket and the characteristic π -stacking interaction with Phe 168^{45,52} was not formed. However, since the ligand was facing the other way, the amine substitution on the pyrimidine ring was predicted to form a hydrogen bond with Glu 169^{45,53} (Figure 4.6 F).

Comparing the docking results of these two ligands suggests that a large substitution on the linker connecting the pyrimidine and indole moiety is desired. Not only did it position the core to form the characteristic π -stacking interaction with Phe 168^{45,52}, but the ligand was further

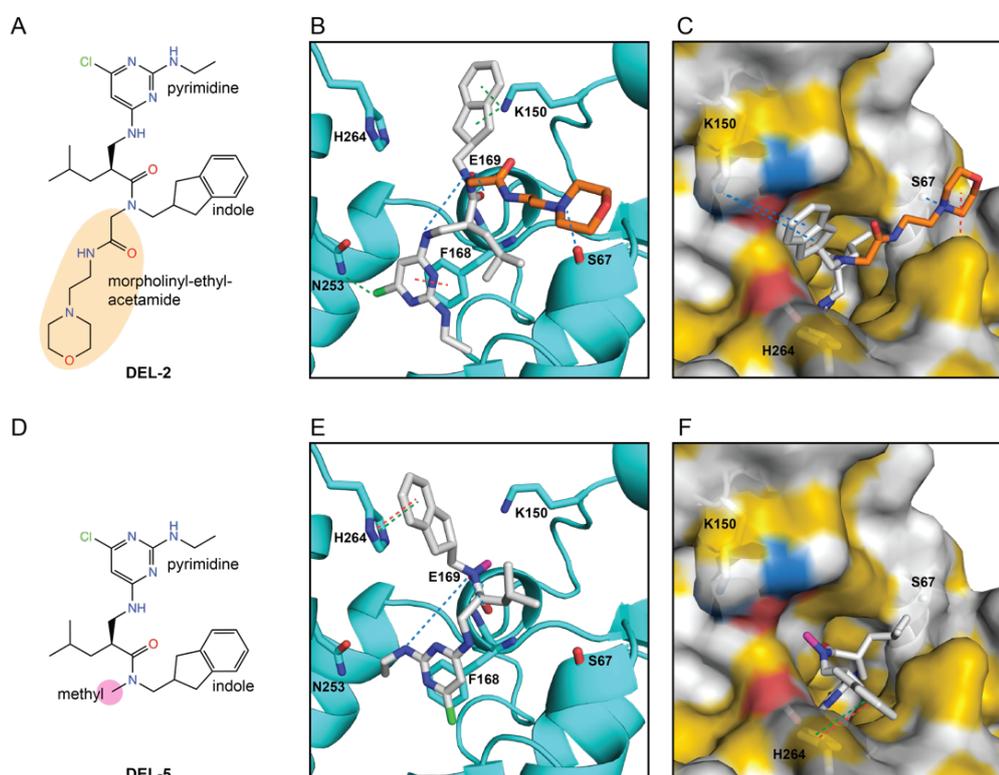


Figure 4.6: Comparison of docking poses of DEL-2 and DEL-5. (A) structure of DEL-2. The large morpholinyl moiety is highlighted in orange. (B) Binding pose of DEL-2. Helix I is hidden for clarity. Residues engaging in polar interactions are shown as sticks. Hydrogen bonds are shown as blue, aromatic stacking interactions as red, and π -cation and salt bridges as green dashed lines. (C) Surface representation of the binding pocket showing the overall orientation and position of DEL-2. The long morpholinyl moiety of DEL-2 extended into the potentially allosteric cleft that is explored by Cmpd-1. The surface is coloured according to the YRB scheme: hydrophobic regions are shown in yellow, positively charged nitrogen atoms in the side chain of arginine and lysine are blue and negatively charged oxygen atoms in the side chains of glutamate and aspartate are red. All remaining atoms are white [269]. (D) Chemical structure of DEL-5. The small methyl substitution in place of the morpholinyl group is highlighted in pink. (E) Binding pose of DEL-5 docked into the receptor. The methyl is shown as pink stick. (F) Surface representation of the binding pocket showing the overall orientation and position of DEL-5.

predicted to extend into a secondary, potentially allosteric binding pocket. Nonetheless, optimal interaction within this site probably depends on ligand-induced conformational changes, as the positions of key residues in this pocket were significantly altered in the Cmpd-1 crystal structure [193]. This might explain why the docking score obtained with DEL-2 does not correlate with the actual affinity of the ligand.

Similarly, DEL-3 and DEL-6 both contain the same ligand scaffold. They consist of a triazaspiro-core with a pyrimidine moiety and a phenyl substituent. They only differ in their exact composition of their benzenepropanoic acid substitution (Figure 4.7 A, D). Both compounds were not predicted to penetrate deeply into the binding pocket, but to mostly form interactions with residues on ECL2 (Figure 4.7). Furthermore, neither compound was predicted to form the persistent π -stacking interaction with Phe 168^{45,52} observed for most other ligands, or any other interaction of this type, despite both containing several suitable ring systems. The benzenepropanoic acid substitution in DEL-3 is less polar than the substituent in DEL-6 and was positioned directing into the binding pocket. The substitution on DEL-6 is more polar and therefore causes a higher desolvation penalty when docked into the binding cav-

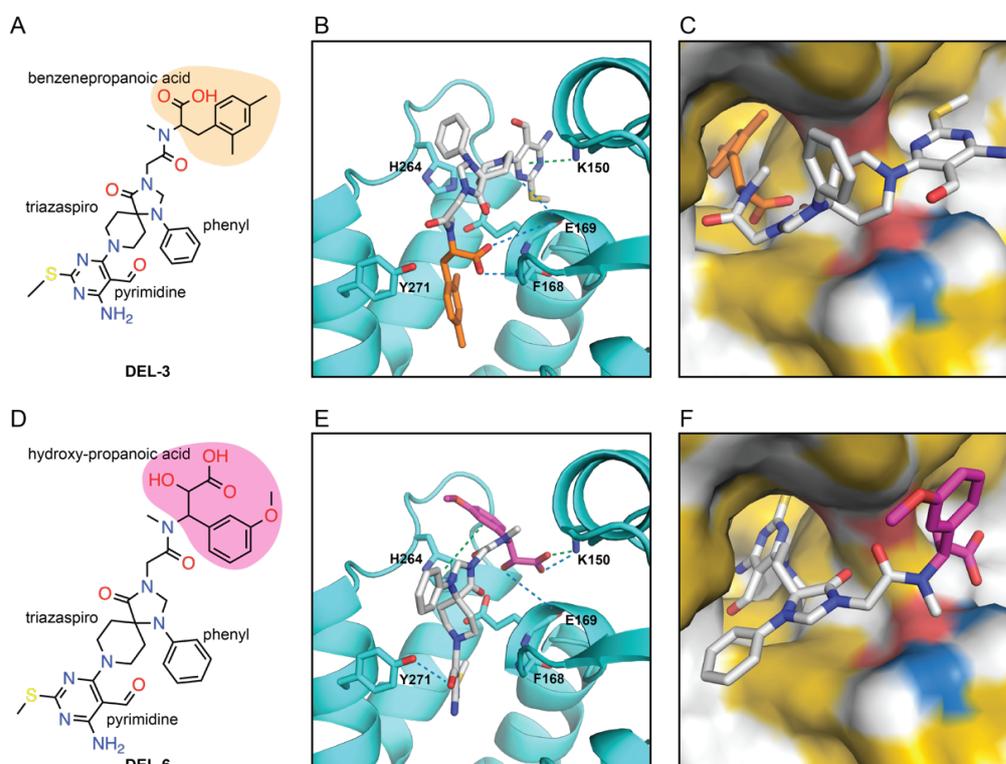


Figure 4.7: Comparison of docking poses of DEL-3 and DEL-6. (A) Chemical structure of DEL-3. The benzenepropanoic acid is highlighted in orange. (B) Docking pose of DEL-3 into the receptor. Helices I and II were hidden for clarity. Residues engaging in polar interactions are shown as sticks. Hydrogen bonds are shown as blue and π -cation and salt bridges as green dashed lines. (C) surface representation of the binding pocket showing the overall orientation and position of the ligand. Surfaces are coloured according to the YRB scheme. (D) Chemical structure of DEL-6. The benzenepropanoic acid is slightly altered to a hydroxy-propanoic acid, highlighted in pink. (E) Docking pose of DEL-6. (F) Surface representation of the binding site.

ity. As this hydroxy-propanoic acid is not predicted to form enough polar interactions with the receptor to compensate for this penalty, it was positioned at the solvent exposed surface of the binding pocket. This flipped the molecule upside down, so that the shared pyrimidine moiety pointed into the binding pocket and formed a hydrogen bond with Tyr 271^{7,35} (Figure 4.7 E, F).

This comparison suggests that a triazaspiro-core with an adjacent phenyl-substitution is probably too bulky to be efficiently positioned into the binding site the same way as other ligands. Furthermore, it seems less favourable if both terminal moieties are highly polar, as this precluded the ligand from penetrating deeper into the binding pocket, resulting in a poorer docking score and affinity of DEL-6. While this comparison allowed conclusions to be drawn about the effects of different substituents on the docking of the same ligand scaffold, it was very surprising that a scaffold such as the triazaspiro-core of DEL-3 and DEL-6, that had been picked up twice in a primary ligand screen, would not form any of the canonical interactions with the receptor that are observed in A_{2A}R-ligand co-crystal structures. I therefore wanted to evaluate whether the docking of these ligands could be improved by introducing knowledge-based constraints that would bias the docking towards ligand-receptor interactions identified in crystal structures.

4.2.3 Docking novel compounds using knowledge-based binding constraints

So far, all available crystal structures of ligand-receptor complexes show that the central core of the ligand resides in the same binding pocket, dominated by hydrogen bonds with the completely conserved Asn 253^{6,55}. Alanine scanning studies highlighted the key role of this residue, as ligand binding was completely abolished when Asn 253^{6,55} was mutated [270]. Furthermore, in all co-crystal structures of A_{2A}R with a small molecule ligand, the ligand is stabilised by a π -stacking interaction with Phe 168^{45,52}, which is highly conserved in all adenosine receptors. However, in the presented docking studies, not all of the novel ligands formed these pivotal interactions, despite the presence of suitable functional groups in the ligand structure. Therefore, the predicted poses might not represent the true binding mode. To investigate whether the docking results could be improved by considering previous knowledge about ligand binding, constraints were introduced during the docking run.

In an initial test, docking constraints were used which required the ligand to form both of the above-mentioned hydrogen bonds and π - π -interactions. However, these constraints were too restrictive and reduced the docking scores of the reference ligands (data not shown). Therefore, this set of constraints were not pursued further. Previous surface calculations on the ZM241385 bound structure revealed that Phe 168^{45,52} has the highest contact area with the ligand and that the aromatic π -stacking interaction alone contributes a quarter of the total binding energy for this specific ligand [271]. Therefore, a constraint was created that should position an aromatic moiety of each ligand at a similar position as the triazolotriazine core of ZM241385 found in the crystal structure, thereby provoking a π -stacking interaction with Phe 168^{45,52}. To this end, the aromatic ring engaged in the π -stacking interaction in ZM241385 was selected as centre of a sphere for a positional constraint. The radius of this sphere was determined empirically to be

2 Å by docking the reference ligands with differently set up constraints; lower or higher radii both worsened the docking scores compared to the ligand docked without constraints (data not shown).

Introducing the constraint had no significant impact on the scoring table and the binding poses of most ligands (supplementary Table B.4). The docking of the reference ligands, as well as FOC-2, DEL-1, FOC4, and DEL-2, remained unchanged. Despite a slightly improved docking score, DEL-6 and DEL-5 still did not form a π -stacking interaction with Phe 168^{ECL3}. However, after docking with the constraint, the hydroxypropanoic acid of DEL-6 was reoriented and the hydroxyl groups were predicted to form hydrogen bonds with Lys 150^{ECL2} and Lys 153^{ECL2}, while the benzene group formed π -stacking and π -cation interactions with His 264^{ECL3} (Figure 4.8 A, B). Interestingly, previous MD simulations proposed the involvement Lys 150^{ECL2} and His 264^{ECL3} in the stabilisation of fragment-like PAMs in an extracellular allosteric site [272], indicating that even though the interaction enforced by the constraint was not formed, the constraint might still have led to the discovery of a more favourable ligand pose.

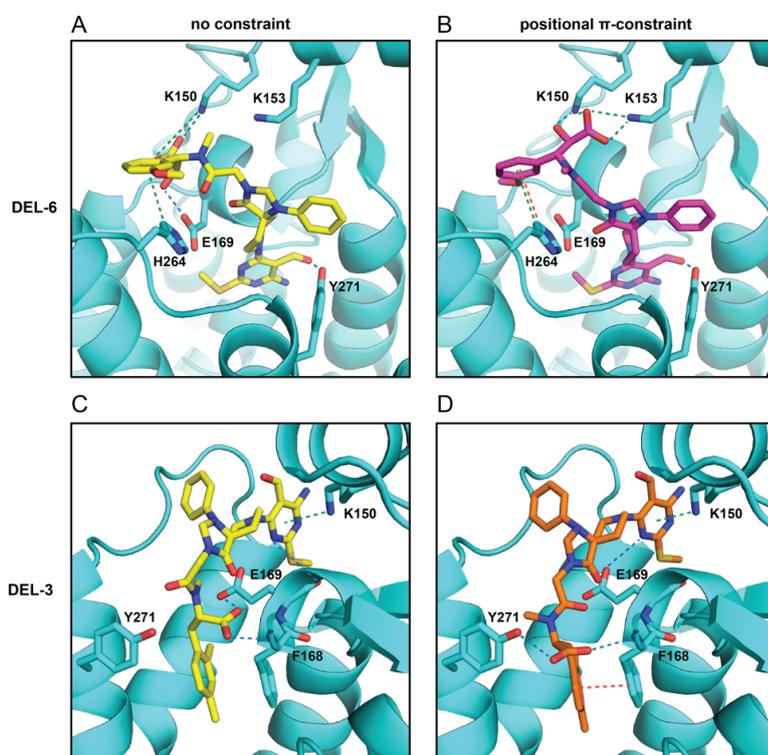


Figure 4.8: Docking of DEL-3 and DEL-6 into A_{2A}R with and without positional π -constraints. (A) Ligand conformation obtained for DEL-6 after docking without constraints (ligand in yellow) and (B) after restraining the docking using a positional constraint for an aromatic interaction with F168 (ligand in orange). (C) Docking pose for DEL-3 without constraint and (D) with the positional π -constraint. For clarity helices I and II have been hidden. Hydrogen bonds are shown as blue, aromatic stacking interactions as red, and salt bridges as green dashed lines. Residues involved in binding are shown as sticks.

In contrast, docking DEL-3 with the knowledge-based constraint did indeed enforce the formation of an aromatic stacking interaction with Phe 168^{45.52}. This improved the docking score from -4.9 to -8.9 . Strikingly, the overall position of the ligand did not change much; however, the constraint pulled the ligand 1.8 \AA deeper into the binding pocket and thereby positioned the phenyl moiety to form the required aromatic stacking interaction with Phe 168^{45.52} (Figure 4.8 C, D). This slight downwards shift propagated throughout the molecule. The adjacent carboxylic acid was shifted by 2.4 \AA allowing the formation of two additional hydrogen bonds with the backbone of Phe 168^{45.52} and Tyr 271^{7.35}. Together with the binding energy provided by the novel π -stacking interaction, these additional hydrogen bonds must have accounted for the large improvement in docking score. In contrast, the pyrimidine moiety was docked to the same site, still forming the hydrogen bond and π -cation interaction with Glu 169^{45.53} and Lys 150^{45.34}.

Thus, the knowledge-based constraint introduced seems to be very well suited to enforce more relevant ligand poses.

4.2.4 Screening for alternative binding sites for FOC-6

The pharmacological assays suggested that FOC-6 might bind to an alternative binding site as well as the orthosteric pocket. However, the docking studies performed in Glide were restricted to a grid around the orthosteric site, making it impossible to find potential binding sites outside of this pocket. Therefore, a second round of docking was performed using Autodock Vina, which allows for a wider search space compared to Glide, encompassing the entire extracellular part of the receptor. As a control, this docking was not only performed using FOC-6, but also the previously used reference compounds ZM241385 and theophylline, as well as the only approved A_{2A}R inhibitor Istradefylline, and the compounds with the highest and poorest docking scores in Glide.

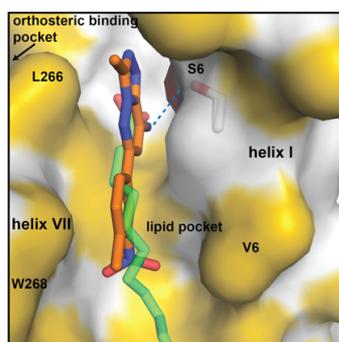


Figure 4.9: Blind docking of FOC-6 into the whole receptor. A secondary binding pose was obtained, in which FOC-6 (orange) was positioned into a hydrophobic cleft between helix I and VII. The orthosteric binding pocket would lay behind helix VII, indicated with an arrow. FOC-6 could potentially form a hydrogen bond with S6 (shown as sticks). However, in the 4E1Y crystal structure this pocket is occupied by a lipid (green). The protein surface is coloured according to the YRB scheme: hydrophobic regions are shown in yellow, positively charged nitrogen atoms are blue and negatively charged oxygen atoms are red. All remaining atoms are white.

Table 4.2: Docking scores obtained through blind docking into the whole extracellular portion of A_{2A}R using Autodock vina. The inactive state structure (4EIY) was prepared in the same way as for the docking studies in Glide; however, when selecting the region for the docking the whole extracellular part of the receptor was chosen. ”-“ indicates the compound was not docked to this pocket by the program.

Ligand	Orthosteric binding pocket score	Lipid binding pocket score
ZMA	-11	-9.2
Theophylline	-6.1	-5.7
Istradefylline	-	-8.2
DEL-3	-	-
DEL-5	-8.6	-8.4
FOC-2	-9.3	-9
FOC-6	-9	-8.3

The highest scoring pose was obtained when docking FOC-6 into the orthosteric binding pocket. However, with a slightly lower score, it could also be docked into a hydrophobic pocket defined by helix I and VII (Figure 4.9). In the predicted pose within this hydrophobic pocket, the exocyclic amine group on the central core of the ligand was positioned towards the orthosteric cavity, potentially forming a polar interaction with Ser6^{1,32}. However, in the original 4EIY crystal structure, this pocket is occupied by a lipid [150]. There are some examples of small molecule ligands that bind to conserved lipid sites of membrane proteins, displacing a bound phospholipid [273]. However, all other compounds were docked into this hydrophobic pocket too, most of them with even higher docking scores than obtained for FOC-6 (Table 4.2). The only exception was DEL-3, which surprisingly could not be docked into either pocket. Therefore, binding of FOC-6 to this lipid-site would not explain the unique behaviour of this compound observed in the binding studies.

4.3 Predicting receptor subtype selectivity of ligands by docking against A₁R and A_{2B}R models

The ligand binding site of adenosine receptors is largely conserved (Figure 4.10). While there are certain differences in amino acid sequences that are implied in conveying subtype selectivity (highlighted in bold), the main regions responsible for this selectivity are thought to be found in the extracellular loop region [201, 200]. Docking studies with the novel ligands predicted that most of them do not penetrate deep into the binding site, but form a series of interactions at the surface of the pocket. These involve Lys 150^{ECL2} and Lys 153^{ECL2} (highlighted in yellow), both of which are unique to A_{2A}R. This suggests that the some of the ligands might show some subtype selectivity.

To further investigate this hypothesis, the novel ligands were docked into an A₁R structure and an A_{2B}R homology model. A first indication for subtype selectivity can be deduced by comparing the docking scores. If a ligand has a very good score for one subtype, but a much weaker score for the other receptors, this might suggest some specificity. This notion might further be enforced by relating the scores and docking poses to a known, well characterised

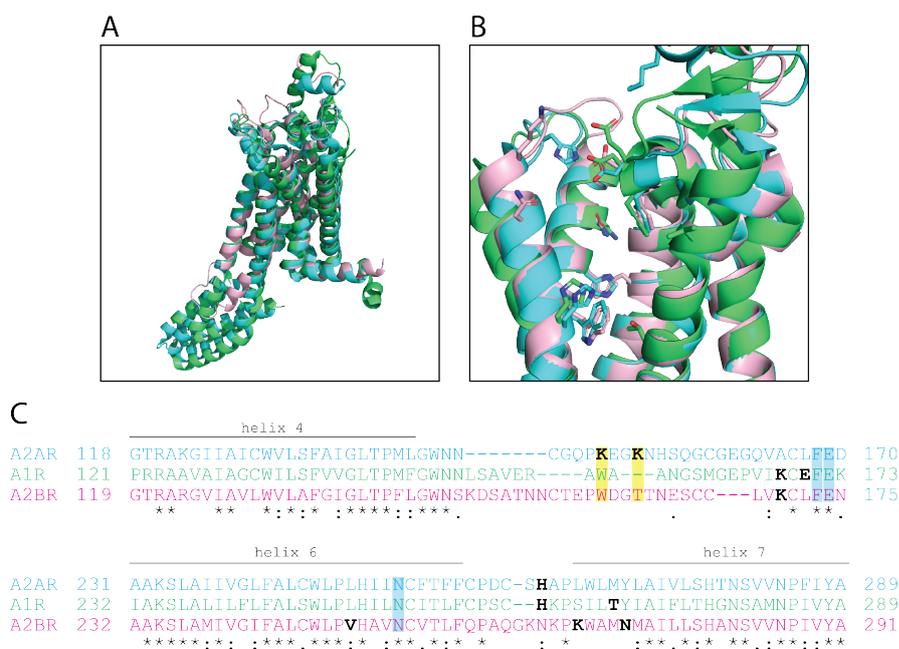


Figure 4.10: Alignment of different adenosine receptors. (A) Structural alignment of inactive state structures of A₁R (green), A_{2A}R (cyan), and a homology model of the A_{2B}R (pink). (B) Structural alignment of ligand binding sites with relevant conserved residues in sticks. (C) Sequence alignment of A₁R, A_{2A}R and A_{2B}R. Residues important for ligand binding in all subtypes are highlighted in blue. Sequence differences conveying subtype specific are flagged in bold and A_{2A}R specific residues that are involved in ligand binding in yellow. Asterisks (*) indicate positions which have a single, fully conserved residue, the colon (:) indicates conservation between groups of strongly similar properties and a period (.) indicates conservation between groups of weakly similar properties.

reference. ZM241385 is one of the best described AR antagonists with selectivity for the A_{2A}R subtype. *In vitro* pharmacological experiments revealed a low affinity at A₁R as well as low potency at A_{2B}R [274]. Thus, ligands that have a weaker docking score than ZM241385 for the other subtypes, but high A_{2A}R docking score, might display selectivity for the A_{2A}R subtype as well. Finally, a detailed comparison of the obtained docking poses will help to identify regions that are predicted to be involved in ligand recognition, but might be specific to a certain subtype, as *e.g.*, the previously mentioned Lys 150 and Lys 1153 in ECL3.

4.3.1 Docking the novel ligands into an inactive state A₁R crystal structure

While A_{2A}R has been crystallised in complex with a series of different ligands, there are currently only five structures available for A₁R, only two of which are in the antagonist bound state. The highest resolution structure (5UEN) was chosen for subsequent docking studies and the receptor was prepared as described for A_{2A}R. The novel ligands were then docked against the receptor using Glide with the same parameters as for A_{2A}R (protein preparation and docking settings are discussed in chapter 2.9.2).

Table 4.3: Docking scores and residues predicted to form polar interactions for ligand docking into the 5UEN A₁R structure. The ligands are sorted by their Glide docking score, determined without defining any additional constraints (Without constraints). These results are compared to the scores and predicted residues obtained after docking with a positional constraint of 2 Å around an aromatic ligand core that forms a π -stacking interaction with Phe 171^{45:52} (with positional π -constraint). Residues predicted to be involved in polar interactions are listed according to interaction type as π -stacking interactions, hydrogen (H-) bonds, and salt bridges. If nothing else is stated, the interaction is formed by the main functional group of the residue, otherwise the involved group is stated in squared brackets (NH₂ = amino group; O = hydroxy group; bb = backbone).

Ligand	Without constraints			With positional π -constraint				
	Score	π -stacking	H-bonds	Salt bridge	Score	π -stacking	H-bonds	Salt bridge
DEL-5	-10.9	F171	N254 [NH ₂ , O]	-	-10.8	F171	N254 [NH ₂ , O]	-
DEL-2	-10.8	F171	N254 [NH ₂ , O], E172	E172, E170	-10.7	F171	N254 [O], E172	E172, E170
FOC-3	-10.6	F171	N254 [NH ₂ , O]	-	-10.6	F171	N254 [NH ₂ , O]	-
DEL-1	-9.9	-	F171 [bbNH ₂], Y12	-	-10.1	-	F171, Y12	-
DEL-6	-9.8	F171	N254 [O], N70 [NH ₂], K168	K168	-9.9	F171	N254 [O], K168	K168
FOC-1	-9.7	F171	N254 [O], E172	E172, E170	-9.7	F171	N254 [O], E172	E172, E170
DEL-3	-9.2	F171	N254 [O], K168	K168	-9.1	F171	N254 [O], K168	K168
FOC-2	-8.7	F171	N254 [O], H278	-	-8.4	F171	N254 [O], H278	-
ZMA	-8.1	F171	N254 [NH ₂ , O], V62 [bbO]	-	-8.1	F171	N254 [NH ₂ , O], V62 [bbO]	-
FOC-4	-8.0	F171	N254 [NH ₂ , O]	-	-8.0	F171	N254 [NH ₂ , O]	-
FOC-5	-7.5	F171	-	-	-7.5	F171	N254 [O]	-
FOC-6	-7.4	F171	N254 [NH ₂ , O]	-	-7.4	F171	N254 [NH ₂ , O]	-
Theo	-7.3	F171	N254 [NH ₂ , O]	-	-7.3	F171 (2x)	N254 [NH ₂ , O]	-
DEL-4	-5.5	-	-	-	-6.7	F171, W ²⁴⁷	N254 [NH ₂], F171 [bbNH ₂]	E172

The main residues involved in ligand recognition are conserved amongst all AR subtypes, including the hydrogen bonding with the asparagine at position 6.55 and the aromatic interaction with phenylalanine at position 45.52 (Figure 4.10) [201]. As the knowledge-based positional constraint improved the predicted docking poses for A_{2A}R, it was also applied to the other subtypes. The docking scores obtained for A₁R covered a similar range as observed for A_{2A}R, ranging from -6.7 for DEL-4 to -10.8 for DEL-5 (Table 4.3). However, the best scores were slightly worse than in A_{2A}R, where the best ranked compound, FOC-2, had a score of -12.2 . Introducing the positional constraint around Phe 171^{45.52} for A₁R docking had no significant effect on the scoring and positioning of most ligands. The only exception was DEL-4, which could not be properly placed without the constraint (Figure 4.11 B). However, after applying the constraint during docking, several of the typical interactions were formed, including a π -stacking interaction with Phe 171^{45.52} (Figure 4.11 C).

Interestingly, DEL-5, which had the worst docking score when docked against A_{2A}R, had the highest score when docked against A₁R (Table 4.5). DEL-5 is a bulky ligand and therefore was predicted not to fit into the binding pocket of A_{2A}R, only forming interactions at the solvent exposed end of the pocket. In contrast, in the A₁R structure, helix VII is bent outwards by about 3 Å compared to A_{2A}R, widening the ligand binding pocket and allowing larger ligands such as DEL-5 to insert slightly deeper into the cavity. In this pose, an aromatic stacking interaction with Phe 45.52 was predicted (Figure 4.12 B). The ligand core was further stabilised by two hydrogen bonds, both involving Asn 256^{6.55} with one going to one of the ligand's pyrimidine nitrogen and the other going to the adjacent amine.

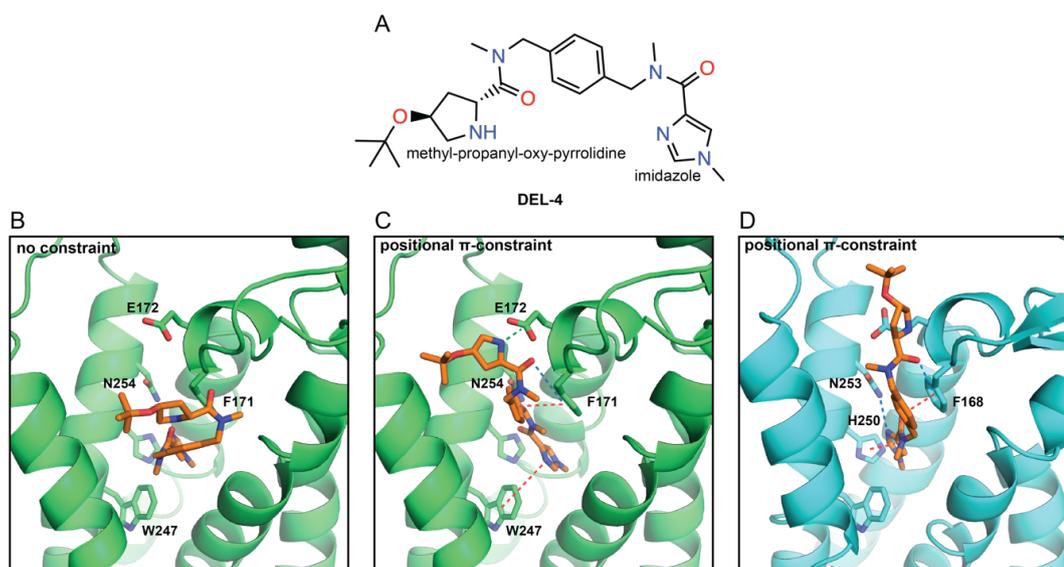


Figure 4.11: Docking of DEL-4 into the A₁R. (A) Chemical structure of DEL-4. Docking of DEL-4 into the A₁R (B) without and, (C) with the positional π -constraint. A₁R is shown as green cartoon. (C) The orientation adopted by the ligand if docked with constraints resembles the conformation adopted when docked into the A_{2A}R, shown in cyan. Important residues are shown as sticks and formed hydrogen bonds, aromatic interactions and salt bridges as dashed blue, red and green lines, respectively. Helices I and II were hidden for clarity.

The core of the bioisostere DEL-2 was positioned similarly to DEL-5 (Figure 4.12 A). Docking placed the large morpholinyl moiety at the top of the binding cavity, sandwiched in between two glutamic acid residues (Glu 170^{45,52} and Glu 172^{45,53}). However, the tight coordination of this group meant the pyrimidine core was pulled closer towards Phe 171^{54,53} and thereby abolished the hydrogen bond between Asn 256^{6,55} and the pyrimidine ring, explaining the slightly lower docking score compared to the smaller bioisostere DEL-5.

While the higher docking scores indicated the large cores of these ligands are better suited for binding into the wider binding pocket of A₁R, the predictions suggested that the long morpholinyl moiety in DEL-2 adopts quite different positions in the two subtypes. Several of the residues in the secondary pocket that accommodates this substituent in A_{2A}R are not conserved in A₁R, *e.g.* at position 7.34 there is a threonine instead of a methionine. Interestingly, this specific residue has previously been linked to subtype selectivity [202, 275, 276]. This would suggest some subtype selectivity towards A_{2A}R. In contrast, one of the glutamic acid residues tethering the morpholinyl moiety in A₁R bound state is unique to the A₁R, which might suggest A₁R selectivity. Thus, this example illustrates that binding data for the A₁R would first be required to make more definite statements about subtype selectivity of any of these compounds.

The wider binding pocket of A₁R also allowed for smaller ligands to sit deeper in the binding cavity. Therefore, in contrast to A_{2A}R, in A₁R, FOC-2 was predicted to access more buried regions. In this pose, the modified phenol moiety was pointing towards helix VII, where the hydroxyl-group was engaged by His 278^{7,42}, similar to what has been observed in the crystal structures of A_{2A}R with the 1,2,4-triazines [217]. However, since the top of helix VII is bent outwards, Tyr 271^{7,35} was moved out of the binding pocket so the second aromatic interaction observed for FOC-2 docked into A_{2A}R could not be formed. Furthermore, with FOC-2 positioned ~ 3.4 Å deeper inside the binding pocket, the hydrogen bond formed between the unsubstituted exocyclic amine group and Glu 169^{ECL2} of A_{2A}R could also not be established with A₁R. This ultimately led to a lower docking score, despite fitting deeper into the binding pocket, suggesting potential A_{2A}R selectivity of this ligand.

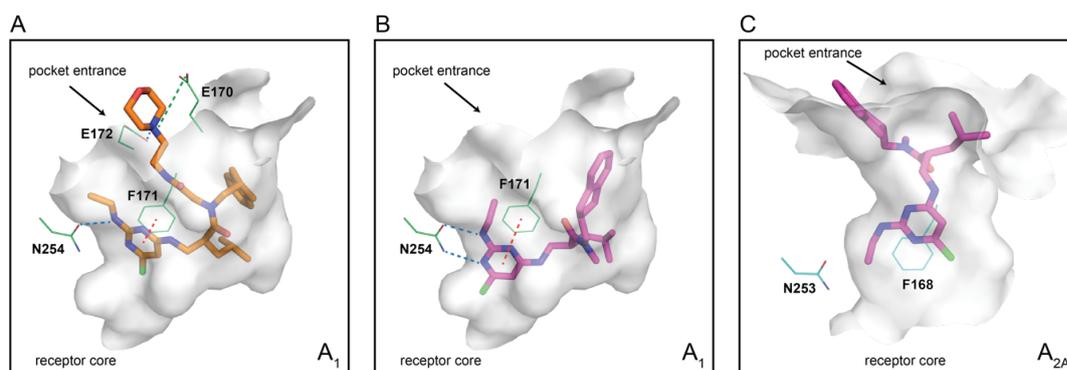


Figure 4.12: Docking of DEL-5 and DEL-2 into the wider binding pocket of A₁R compared to the docked pose of DEL-5 in A_{2A}R. (A) Docking of DEL-2 into the binding pocket of A₁R. (B) Comparison of binding poses of DEL-5 in A₁R and (C) A_{2A}R. The binding site is shown as surface representation. The position of residues important for binding is indicated by line representations.

4.3.2 Docking the novel ligands into an inactive state A_{2B}R homology model

4.3.2.1 Building a homology model for A_{2B}R

To further understand subtype selectivity of the ligands, A_{2B}R should be included into any analysis as it is the closest homologue of A_{2A}R. However, no structure is currently available for A_{2B}R, so a homology model was created to enable docking studies. A variety of different models were generated, using different servers and template models for homology model building, as well as *ab initio* modelling. The quality of the different models was assessed using the QMEANBrane scoring function (SWISS-MODEL) and the QMEAN-Z score from this function was used to compare the different models. The homology models generally scored better than the *ab initio models* and could be even further refined by using the ModRefine algorithm on the Zhang group server [190].

The two highest ranked homology models were both based on A_{2A}R templates, with one obtained using an antagonist (3EML) and one an agonist (4UHR). These models were selected for further analysis. Both models were highly similar to their template with an average C α RMSD of 0.42 Å and 0.05 Å respectively (Figure 4.13 A, B). Comparison of the two homology models revealed that they mostly differed at the periphery and in the positioning of their loops (Figure 4.13 C). The positions and rotamers of the corresponding residues important at the A_{2A}R binding pocket did not differ much, except for the residues at the very top of the binding site and on the small helical segment on ECL2 (Figure 4.13 D). The average C α RMSD between the two A_{2B}R models was similar to the average C α RMSD between the two differently stabilised A_{2A}R templates (1.44 Å and 1.36 Å, respectively), which suggested that the models might contain some state-specific features as well. Thus, as the ligands were all considered inhibitors, the model generated with the antagonist-stabilised template was chosen for docking studies.

4.3.2.2 Docking of the novel ligands into the A_{2B}R homology model

The novel ligands were docked against the A_{2B}R homology model using the same docking parameters and restraint as for the other receptors. Docking against the A_{2B}R homology model was notably less successful than the docking of the same ligands against A_{2A}R crystal structures. Three of the largest, most branched compounds, namely DEL-3, DEL-6, and DEL-2 could not be docked into the model at all. However, while the bulky morpholinyl-ethyl-acetamide moiety of DEL-2 likely prevented docking, the ligand with the smaller methyl, DEL-5, could be modelled into the binding site. The docking score of DEL-5 was still relatively poor (−4.1) and very few specific interactions with the receptor were predicted, indicating that this ligand series is not very well suited to bind A_{2B}R.

Also introducing the positional constraint to force a π -stacking interaction with Phe173 at position 45.52 was more difficult, as the homology model did not contain any ligand to select as the centre for the constraint. Thus, the constraint had to be defined from the phenyl ring of the side chain, which required the positional radius to be increase from 2 Å to 4 Å. The obtained docking scores for the A_{2B}R were significantly lower than for the other subtypes. The best score was obtained for FOC-3, which was slightly improved by applying the constraint (−8.5).

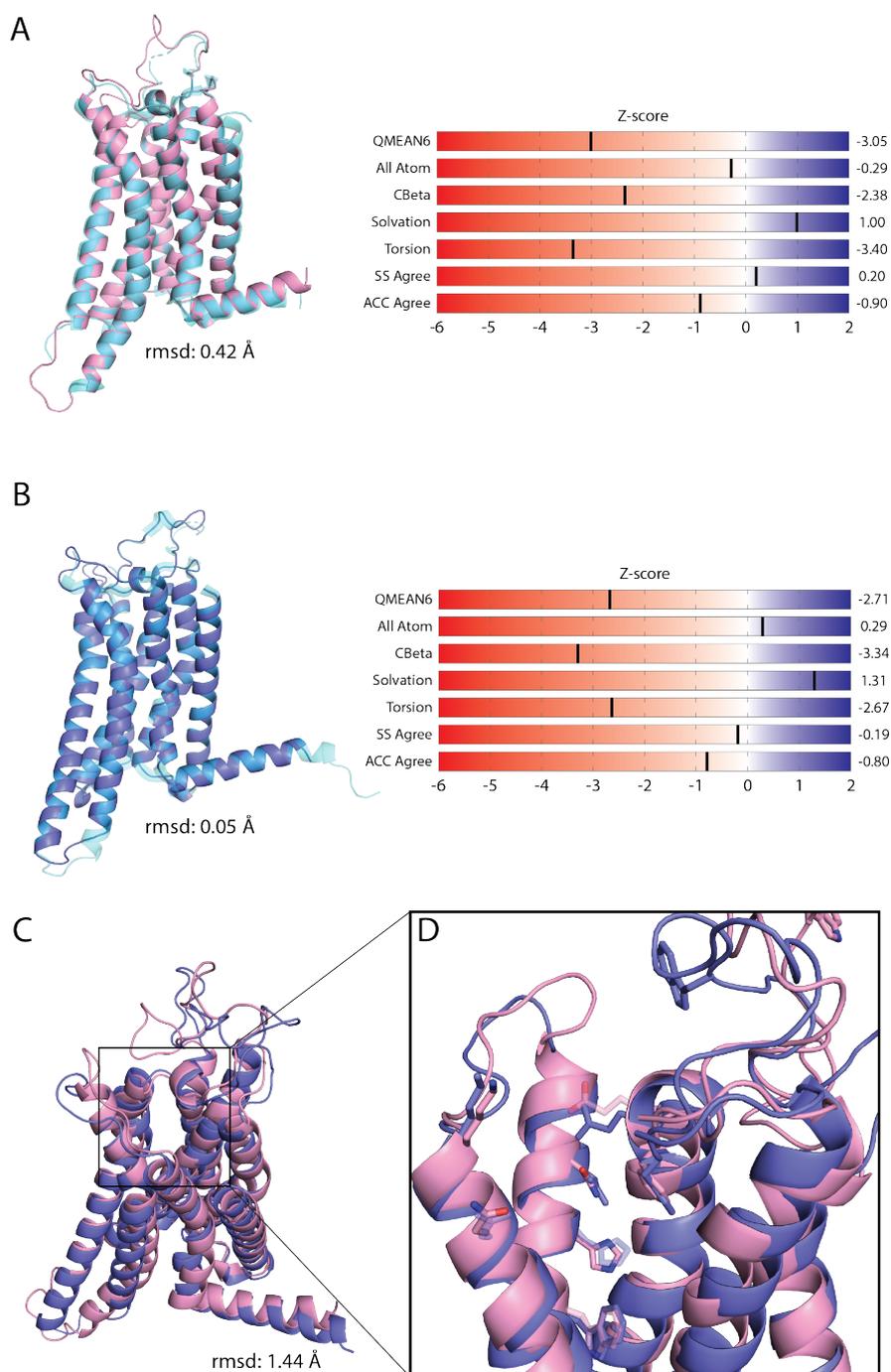


Figure 4.13: Homology models generated for A_{2B}R. Both models were validated using the SWISS-MODEL QMEANBrane server. Each model was aligned to its template (cyan) in pymol. **(A)** Model from ModBase, generated using the antagonist bound A_{2A} structure 3EML as template and refined using ModRefiner (Zhang Lab). **(B)** Model generated from SWISS-MODEL using the agonist bound A_{2A} structure 4UHR as template. **(C)** Alignment of the two A_{2B}R models. **(D)** Zoom-in on the binding pocket. Residues involved in binding of the different ligands are highlighted as sticks.

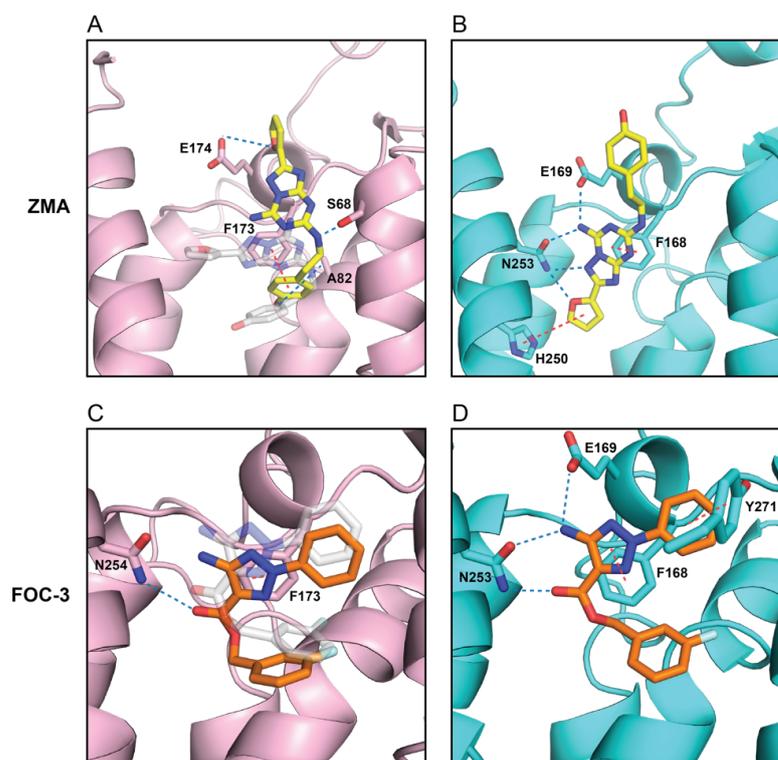


Figure 4.14: Docking of ZM241385 and FOC-3 for the $A_{2B}R$ compared to the poses in $A_{2A}R$. Ligand binding pockets in the $A_{2B}R$ and $A_{2A}R$ are shown in pink and cyan, respectively. The side chain of residues forming hydrogen bonds and aromatic interactions are shown in stick, while the interactions are depicted by red and blue lines. Docking poses obtained without constrains are shown in transparent grey. **(A)** ZM241385 binding to the $A_{2B}R$ and **(B)** to the $A_{2A}R$. **(C)** Docking poses of FOC-3 in $A_{2B}R$ and **(D)** in $A_{2A}R$. Helices I and VII are hidden for clarity.

However, overall, using the constraint generally worsened the docking scores and caused more changes in the ranking of the ligands than for the other two subtypes (Table 4.4). DEL-2, DEL-3 and DEL-6 could still not be docked.

The reference ligand ZM241385 seemed to adopt an unfavourable, folded up position when docked into $A_{2B}R$, forming no polar interactions with the receptor (Figure 4.14 E). Introducing the π -stacking restraint did seem to relax the molecule, but the typical interaction of the bicyclic triazolotriazine unit with Phe173^{45,52} was not predicted to form, instead the terminal phenyl moiety formed a π -stacking interaction with this residue, positioning the compound higher up in the binding pocket than observed in the $A_{2A}R$. Furthermore, this caused the ligand to be positioned upside down, with the phenol moiety pointing into the binding pocket, while the furan ring was positioned at the surface (Figure 4.14 A, B).

The predicted conformation and new binding position for ZM241385 in $A_{2B}R$ indicate this pose is unlikely to occur, which nicely correlates with a significantly lower docking score and is in line with the fact that ZM241385 is a subtype specific $A_{2A}R$ inhibitor. This suggests that ligands with a $A_{2B}R$ docking score lower than ZM241385, but high $A_{2A}R$ docking score might be subtype selective as well. Interestingly, this includes all the DEL-compounds, except of DEL-1.

Table 4.5: Comparison of ligand Glide scores for docking against different AR subtypes. The docking scores determined with the positional π -constraint for the three assessed AR subtypes are compared. The list of scores is sorted according to the scores for the A_{2A}R. The best docking score for each subtype is highlighted in bold and the scores for the reference compounds ZM241385 (ZMA) and theophylline (Theo) are marked in italic. Asterisks indicates compounds that could be selective over one (*) or both (**) of the other subtypes, suggested by either a very high docking score or, specifically for A_{2A}R, if the score for A_{2A}R is < -8 and for the other subtypes $>$ the score of ZMA. The scores assigned to DEL-2 would suggest subtype selectivity for the A₁R, but the predicted binding poses might suggest selectivity for the A_{2A}R. Therefore, this score is captioned with an exclamation mark.

Docking scores (with positional π -constraint)			
Ligand	A _{2A} R	A ₁ R	A _{2B} R
FOC-2	-12.2 **	-8.4	-7.7
FOC-3	-11.0	-10.6	-7.8
<i>ZMA</i>	<i>-10.6</i>	<i>-8.1</i>	<i>-5.9</i>
FOC-5	-10.2 *	-7.5	-6.8
DEL-3	-8.9	-9.1	-
FOC-1	-8.8	-9.7	-8.0
DEL-1	-8.5	-10.1 **	-7.7
FOC-4	-8.3 *	-8.0	-4.5
FOC-6	-8.2 *	-7.4	-7.0
<i>Theo</i>	<i>-7.7</i>	<i>-7.3</i>	<i>-6.1</i>
DEL-4	-7.6	-6.7	-5.0
DEL-6	-7.4	-9.9	-
DEL-2	-6.7	-10.7 !	-
DEL-5	-6.6	-10.8 **	-4.3

The positional π -constraint also caused major changes in ligand positioning of FOC-3, which lead to a weaker docking score. Regardless of the lower score, the predicted pose for FOC-3 docked into A_{2B}R with the constraint, more closely resembled the average ligand binding mode observed in the different A_{2A}R co-crystal structures, making it a more plausible pose to occur. In this pose, FOC-3 formed a π -stacking interaction with Phe 173^{45,52} through its central triazole core instead of one of the phenyl moieties (Figure 4.14 C), similar to what was observed for the triazole compound in the A_{2A}R co-crystal structure [193]. However, A_{2B}R has an asparagine in position 7.35 instead of a tyrosine and was therefore missing the second π -stacking interaction formed in A_{2A}R (Figure 4.14 D), which explains the lower docking score.

The highest ranked compound when docking into A_{2A}R was FOC-2 (Table 4.5), with a docking score of -12.194. When docking the same compound into A_{2B}R the docking score deteriorated to -7.664. Looking at the binding poses revealed that the ligand is predicted to be positioned very similarly in both receptors. However, when docked to A_{2B}R, FOC-2 sat about 1.6 Å deeper inside the binding pocket and the two aromatic pendants were flipped, with the phenol moiety at the top of the binding pocket. This abolished the second hydrogen bond, which was predicted to form between the hydroxyl group and Asn 254^{6,55} of A_{2A}R. Furthermore, FOC-2 was predicted to form an additional π -stacking interaction with A_{2A}R, between the phenyl moiety and Tyr 271^{7,35}, which likely contributed to the good docking score.

A_{2B}R does not have an aromatic residue at the position 7.35. The lack of these two additional interactions explains the weaker docking score of FCO-2 for A_{2B}R and might suggest some subtype specificity.

While the orthosteric binding site of ARs is largely conserved between subtypes, this analysis revealed that the subtle variations between these three receptors still lead to the prediction of significantly different binding poses, suggesting that some of the ligands might indeed display some subtype selectivity.

4.4 Discussion

4.4.1 Validity of the predictions

The results from this docking study are in good agreement with the experimental data previously described in chapter 3. Generally, high affinity ligands had higher docking scores than ligands with an affinity in the micromolar range and *vice versa*.

The ligands from the virtual screen (FOC-ligands) generally scored better. This is expected, since they have been identified in docking studies and thus, were selected according to their good docking scores. In contrast, the hits from the DEL-screen were generally larger and therefore probably also more difficult to dock [277, 278]. Comparing the results of scaffolds that have previously been identified with published structures of bioisosteres, *e.g.* FOC-3 with Cmpd-1, revealed similar ligand poses, further corroborating the confidence in the docking.

4.4.2 Possibilities for chemotype optimisation

Evaluating the proposed binding mode to the receptor can aid the chemical optimisation of the ligand hit. This is probably best illustrated when comparing the predicted binding pose of FOC-2 with the crystal structures of the 1,2,4-triazine bioisosteres. 1,3,5-triazines, similar to FOC-2, were already discovered in a previous virtual screen [211]. Considering the proposed binding mode of the hit, it was hypothesised that by reducing the spacing in-between the two aromatic substituents the ligand could sit more deeply in the receptor pocket and access the region normally occupied by the ribose group of adenosines. This hypothesis was subsequently validated and refined using site directed mutagenesis data from literature and biophysical mapping approaches [216] and eventually by solving the structure of two 1,2,4-triazine bioisosteres [211, 217]. This ligand optimisation finally resulted better ligand-efficiency, higher affinity and subtype selectivity [217].

While the affinity reported for the 1,2,4-triazines is significantly higher than the affinity of FOC-2, the binding properties determined for FOC-2 might still be good enough for structure determination and it would be interesting to see if the rationale described above could be supported by an actual structure of the precursor. Furthermore, such a structure could potentially show new directions for optimisation strategies of this scaffold.

Similar optimisation strategies could be applied to the other ligands characterised in this study. These might additionally benefit from the conclusions drawn by comparing ligands with similar scaffolds. Comparison of the docking poses of DEL-3 and DEL-6 *e.g.* revealed that DEL-6 binding might be improved by reducing the polarity of the ligand to reduce the

desolvation penalty of the compound. This might allow the ligand to sit more deeply into the binding pocket. However, the proper establishment of such strategies will be a task for medicinal chemists, and therefore not be further discussed here.

4.4.3 Possible allosteric ligands

Ligand optimisation might be especially interesting for compounds such as DEL-2, where the docking poses predict interaction with potentially allosteric sites. In recent years there has been a growing interest in allosteric drugs for GPCR targets, as they promise the potential for on-target selectivity and a way to finetune the actions of the medicine, as allosteric effects are usually saturable [279]. However, the list of allosteric A_{2A}R ligands is still short, and there is very little knowledge about the structural basis of allosteric modulation [280, 281]. The probably best characterised allosteric modulators are sodium ions and amiloride, that both bind to the same binding site [282, 150].

The only other potentially allosteric site that has been characterised by experimental structures is the secondary site described in complex with Cmpd-1 [193]. Interestingly, two of the selected compounds, DEL-1 and DEL-2 were also predicted to extend into this pocket. The crystal structure of Cmpd-1 revealed significant ligand-induced conformational changes in this region. Thus, it is likely the receptor is also adopting a different conformation in the DEL-1- and DEL-2-bound state compared to the ZM241385-bound state used for docking. This would explain the low docking score of DEL-2 and the few polar interactions predict for DEL-1, despite their high affinity for the receptor.

Interestingly, the corresponding residues of Leu 267^{7.31} and Met 270^{7.34} which are part of this possible allosteric A_{2A}R pocket, are also suggested to be involved in allosteric ligand binding of a biased A₁R ligand (Leu 269^{7.31} and Thr 270^{7.35} in A₁R) [283]. However, while the described pocket in A_{2A}R extends more towards the extracellular aspect of TMI, TMII, and TMVII, the biased A₁ ligand described by Wall *et al.* (2020) mostly occupies a hydrophobic pocket underneath ECL3 and further engages with Glu 172^{ECL2} that is specific to the A₁R. Remarkably, when docked into the A₁R, DEL-2 was predicted to form an interaction with Glu 172^{ECL2} as well. This makes it very difficult to judge if this ligand displays any subtype selectivity.

Therefore, further pharmacological experiments will be required to characterise possible allosteric responses. Additionally, MD simulations could be used to optimise the models of the final complexes. By accounting for receptor flexibility, MD simulations allow more accurate binding poses to be determined than the rigid docking models used here and might help to gain insights into the structural and functional role of this secondary binding site and the associated novel receptor conformations [284, 285, 286, 287]. However, MD simulations are extremely computationally intensive, requiring much more computational power and time than rigid docking studies and it would therefore have been impossible to study all the screened compounds in a similar timeframe.

Molecular dynamics approaches were also used to identify allosteric binding sites on A_{2A}R [272, 288]. Deganutti & Moro (2017) used MD simulations to characterise the binding site of fragment-like PAMs that were shown to slightly slow down the unbinding kinetics of orthosteric ligands. The biggest contribution to fragment stabilisation in this binding site was provided by

ionic interactions with Lys 153^{ECL2} and His 246^{ECL3}. Interestingly, while DEL-6 was one of the few ligands that did not form the canonical π -stacking interaction with Phe 168^{45.52}, nor any hydrogen bonds with Asn 253^{6.55}, it did form interactions with Lys 153^{ECL2} and His 246^{ECL3}, indicating that this ligand might have more allosteric character. Unfortunately, not enough material of this compound was available to perform any functional studies that could have supported this notion.

Finally, a completely different potential allosteric pocket was identified in a hydrophobic cleft formed at the top of helix I and VII by docking FOC-6 into the whole extracellular part of the receptor. However, this prediction was not specific for FOC-6 and so the obtained pose would not explain the unique behaviour of this compound observed in the binding studies. Furthermore, this pocket is occupied by a lipid in the original crystal structure [150]. A_{2A}R is known to be allosterically modulated by lipids and lipid interactions have also been shown to depend on the activation state of the receptor [289, 290]. As such, the identified region is likely to be an allosteric site, to which binding might modulate receptor activity. Increasing evidence from biophysical and structural studies suggest that binding of small-molecule drugs to binding sites at the protein-phospholipid interface is not uncommon [291, 273, 225, 292]. The best characterised lipid interactions of A_{2A}R are with cholesterol. Several specific cholesterol binding sites have been characterised structurally as well as proposed by computational studies and cell-based assays have shown that A_{2A}R-dependent cAMP production is positively correlated with membrane cholesterol [150, 293, 294, 295, 296]. However, a very recent study suggests that these cholesterol interactions are of transient nature and lack high-affinity binding [297]. Thus, despite having a rather low affinity, FOC-6 might be able to displace a transient lipid from this pocket. However, to ultimately identify the binding site of FOC-6, structural, or at least more sophisticated computational studies, will be required. Nonetheless, even if this pocket is not the binding site of FOC-6, it still represents an interesting target site for allosteric A_{2A}R modulators.

4.4.4 Possible subtype selective ligands

Many of the novel ligands are predicted to form extensive contacts with the top of helices VI and VII and ECL 2 and 3. These regions are not only involved in the formation of potentially allosteric sites but are also considered important for subtype selectivity [200, 298] and as such, these results might indicate a high degree of subtype selectivity.

In theory, ligands that were assigned a good docking score for A_{2A}R but weaker scores for the other subtypes, might preferably bind to the A_{2A}R subtype, especially if the docking score for the other subtypes compares to the score of ZM241385, which is A_{2A}R subtype specific. As such, FOC-2, which scored highest in the A_{2A}R studies, but a significantly lower score for the other subtypes, could potentially be subtype selective. However, binding studies with 1,2,4-tiazine bioisosteres showed that this compound series generally also showed affinity for the A₁R and only modest selectivity for A_{2A}R could be achieved when introducing modifications at the 3,5-positions of one of the pendant rings [217]. The series was also only about ten-fold selective over the A_{2B} subtype. It is therefore rather unlikely that the 1,3,5-bioisostere FOC-2 would

be highly subtype selective. Pharmacological assays, such as *e.g.* the LANCE assay, are thus required to determine if some of the ligands suspected to be subtype specific, indeed show a preference for one subtype over others.

It should be noted that docking was optimised for the A_{2A}R and might therefore have been somewhat biased towards this subtype. Furthermore, the models used might have contributed to this bias as well. The selected A₁R structure was crystallised in complex with a covalently bound ligand, which might introduce a slightly different conformation than noncovalent ligands. For A_{2B}R, a homology model based on the A_{2A}R was used, as there are no crystal structures available. Therefore, in some instances the specific positions and rotamers of residues in the binding site might be suboptimal and negatively influence the docking. Introducing constraints to enforce a specific type of binding pose did help to generate more realistic poses, but this might have had the consequence that apparent docking scores were reduced, due to restrained flexibility in the ligand.

Subtype selectivity over the A₃R was not assessed. However, hits from previous docking studies generally displayed a relatively high A_{2A}/A₃R selectivity [209]. This was especially true for chemotypes with unsubstituted exocyclic amine groups that form hydrogen bond interaction networks with Asn 253^{6.55} and Glu169^{ECL2} in A_{2A}R. While these residues are conserved in A₁R, A_{2A}R, and A_{2B}R, A₃R has a hydrophobic valine at the position of the glutamic acid.

Overall, these docking studies provide meaningful insights into the binding modes of the different ligands. While most ligands are predicted to bind in a similar pose than described for previously characterised A_{2A}R ligands, DEL-1, DEL-2, DEL-6, and FOC-6 might extend into possible allosteric pockets and introduce a novel receptor conformation. They are thus particularly interesting for subsequent structural work.

Chapter 5

Towards experimental structures of novel A_{2A} receptor ligand complexes

Most proteins, as well as small organic compounds, are quite flexible in solution and sample a variety of different conformations. As a result, the docking poses determined *in silico* may differ from what is observed in solution [299]. All the tested ligands (potentially excluding FOC-6) bind into the same orthosteric binding site, interacting with similar residues of the receptor. Nonetheless, it is still possible that some of them introduce conformational changes in the receptor, which would lead to different poses than obtained *in silico*. This especially might be the case for the ligands that show a significant discrepancy between predicted docking scores and the experimentally measured affinity (*e.g.* DEL-2). To fully confirm the presence of the ligand in the binding site, and elucidate the actual binding mode, experimentally obtained ligand-receptor complex structures need to be analysed. Protein X-ray crystallography is routinely used for structure determination in drug discovery and was used to determine the vast majority of the A_{2A}R high-resolution structural models obtained so far. Therefore, the aim of this chapter was to crystallise the A_{2A}R in the presence of at least one of the novel compounds and solve the structure of the receptor-ligand complex to validate the binding poses obtained *in silico*.

5.1 Crystallisation strategies for different receptor-ligand complexes

One of the main limiting factors when evaluating the binding properties of early screening hits is that the compounds are usually only available in low quantities, especially if they originate from an internal library and first need to be synthesised. This is particularly limiting for structural studies of GPCR-ligand complexes. As these receptors usually sample a large variety of conformational states, they need to be trapped in a certain conformation in order to be crystallised. This stabilisation of the receptor is particularly important if they are removed

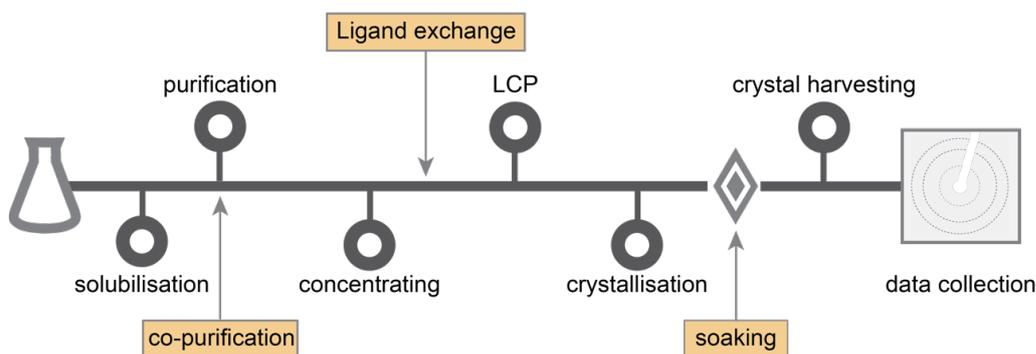


Figure 5.1: Schematic of purification and crystallisation process used for structure determination of receptor-ligand complexes. The orange boxes highlight the timepoints at which the novel ligand is added following either the co-purification, ligand exchange or soaking strategy.

from their native membrane environment, such as when solubilising in detergents, and is usually assisted by the presence of the ligand. Ligands are therefore typically added throughout the entire purification process to enhance stability. This usually further helps to increase protein yields and to reach the high protein concentrations required for crystallographic experiments [225].

$A_{2A}R$ is an example of a receptor that requires the presence of a ligand throughout the entire purification process to ensure stability of the receptor. This involves milligrams of expensive ligand that might not be available. Therefore, when conducting a co-purification with a novel ligand, the solubilisation step is usually performed in the presence of the low affinity ligand theophylline, which is available in large quantities at a reasonable price. Theophylline is then subsequently replaced by higher affinity ligands (co-purification, Figure 5.1) [188]. Nonetheless, for most compounds analysed in this study, only a few milligrams of compound were available, which in most cases was too little to perform co-purifications. Thus, this approach could only be attempted for the ligand with the highest affinity, where the concentration required to fully saturate the receptor was slightly lower. Consequently, alternative strategies such as crystal soaking as well as ligand exchange strategies were explored, where reduced masses of ligand are required (Figure 5.1). For an overview comparing these different crystallisation strategies see the introduction chapter 1.1.3.1.

5.2 Receptor construct choice

Construct design is the first and one of the most crucial steps of a successful crystallographic experiment. For novel target proteins, it is important to screen a variety of different constructs, as usually only half of the protein constructs entering expression trials finally yield suitable protein samples [300]. GPCRs, in particular, require many modifications to make them stable enough for crystallisation. For this study, a construct combining the StaR2 mutations with the bRIL-fusion (A_{2A} -StaR2-bRIL) was used. The use of this construct previously allowed high

resolution structure determination of A_{2A}R [139, 140]. Furthermore, because of the high degree of stabilisation, it also crystallises in complex with lower affinity ligands and therefore, could be successfully used for *in meso* soaking experiments [50].

5.3 Expression of A_{2A}R in *Sf9* cells for structural studies

Initial constructs were obtained from Novartis and cloned into the pFastBacTM vector, if not already provided in this vector. All expression cassettes contained an N-terminal hemagglutinin (HA) signal sequence, a FLAG epitope and a C-terminal 10x histidine (HIS) tag. For X-ray crystallography, all protein constructs were expressed in *Sf9* insect cells using the Bac-to-Bac[®] expression system. Recombinant baculoviruses were generated and amplified as described in the methods section. The optimal virus dilution and expression time for large-scale expression were then determined by performing small-scale expression tests. To this end, *Sf9* cells were grown to 2×10^6 cells/mL and infected with different virus to cell ratios (v/v), ranging from 1:500 to 1:5000. Samples were taken 48, 60, and 72 hours post-transfection, and cell viability and diameter were measured. Protein expression was measured by Western Blot on crude membrane preparations using an anti-FLAG or anti-HIS epitope antibody and compared by densitometry (Figure 5.2). The virus to cell ratio and expression time that resulted in the highest amount of protein, whilst not killing more than about 20% of the cells, was selected for subsequent large-scale expression.

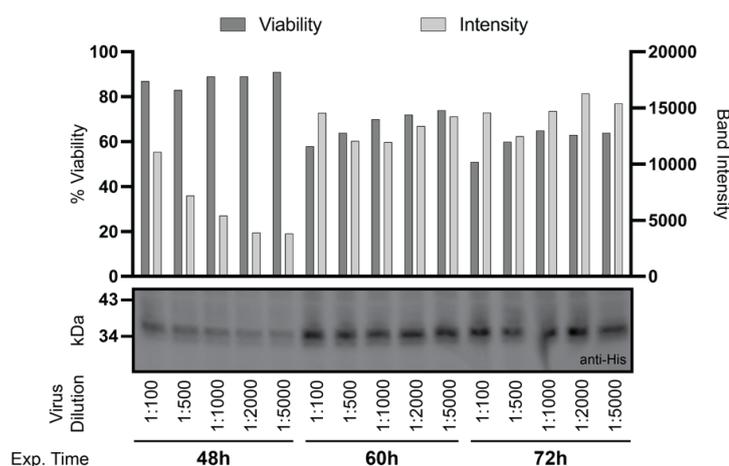


Figure 5.2: Example of an expression test to determine the best virus dilution and harvesting time for expression of a stabilised A_{2A}R construct in *Sf9* cells. For structural studies all A_{2A}R constructs were expressed in *Sf9* insect cells, using the Bac-to-Bac baculovirus expression system. Cells were infected at a density of 2×10^6 cells/mL with different dilutions of V₂ virus and aliquots were taken after 48, 60, and 72 hours post-infection. Cell viability was recorded at each time-point. A crude membrane preparation was performed for each aliquot and receptor expression was assessed by Western blotting using a HIS-tag antibody. Band intensities were quantified and compared with the cell viability from the corresponding timepoints.

5.4 Effects of C-terminal HRV-cleavage site

Each laboratory usually has their preferred construct design and backbone, and especially in an industrial setting these things are often highly standardised. Most of the constructs used in this study were kindly provided by Novartis and were therefore designed following the same structure, consisting of a N-terminal HA-tag followed by a FLAG-tag, the gene of interest, and a C-terminal HRV-cleavage site tailed by a 10× HIS-tag. In an initial proof of concept study, a A_{2A} -StaR2-bRIL^{HRV-HIS} construct, designed according to this standardised scheme, was purified and crystallised in the presence of the reference compound ZM241385. These first protein purifications for A_{2A} R structural studies were performed at Novartis, together with V.-P. Jaakola and were therefore based heavily on his experiences and protocols, including the protocols for membrane protein preparations from *Sf9* cells established by the Stevens laboratory (Joint center for innovative membrane protein technologies, SCRIPPS Institute, CA, USA) and the protein purification from Jaakola *et al.* [188] and Liu *et al.* [150]. These protocols were optimised for the purification of the A_{2A} -T4Lysozyme and A_{2A} -bRIL construct and had established the importance of high sodium chloride concentrations and adding CHS to detergent solubilised receptors to enhance protein stability. In contrast, the thermostabilised A_{2A} R constructs (StaR2) were first introduced for vapour diffusion crystallography experiments by an independent research group and the purification protocols were therefore optimised accordingly [189]. This optimisation included a reduction of the sodium chloride concentration, and the use of a shorter chain detergent (DM instead of DDM). Initially, the A_{2A} -StaR2-bRIL^{HRV-HIS} construct was therefore purified according to a hybrid protocol, based on the experience V.-P. Jaakola, but using DM as a detergent and including as size exclusion chromatography (SEC) step, which allowed the separation of aggregates and higher oligomers from the main monomer peak at approximately 14 mL (Figure 5.3 A). Fractions containing monomeric receptor were

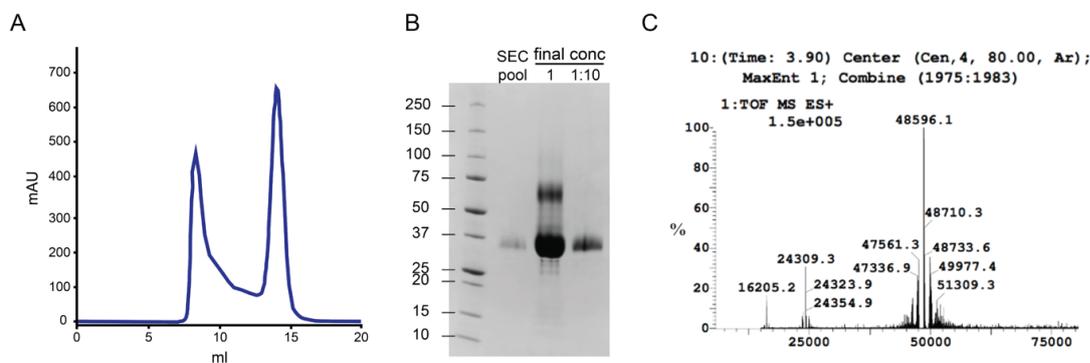


Figure 5.3: Purification of A_{2A} -StaR2-bRIL^{HRV-HIS}. (A) Size exclusion chromatogram using a Superdex 200 10/300 gel filtration column. The first peak corresponds to a higher oligomeric state or aggregates, while the peak at approximately 14 mL consists of receptor monomer. Fractions corresponding to this peak were pooled and concentrated. (B) 1 μ L of this pool (first lane), 1 μ L of the final concentrated sample, and a 1:10 dilution were loaded on a SDS-PAGE. The receptor monomer corresponds to the band slightly below 37 kDa. In the concentrated sample some dimeric protein is visible slightly below 75 kDa. (C) The identity of the final protein sample was analysed by mass spectrometry. The MW of the construct after HA-tag removal is 48 586.7 Da. The 10 Da difference potentially comes from oxidation.

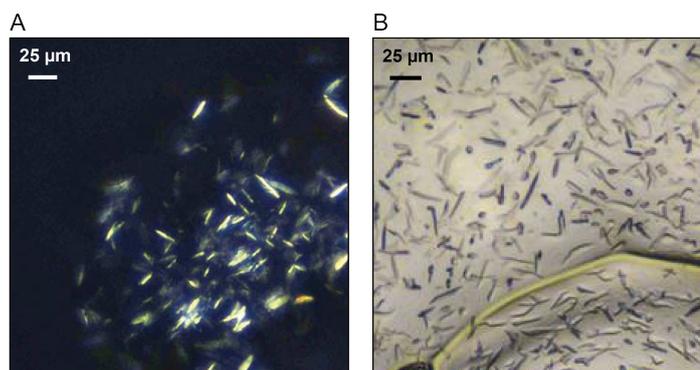


Figure 5.4: Crystals obtained for the A_{2A} -StaR2-bRIL^{HRV-HIS}-ZM241385 complex. **(A)** Crystals obtained in 30 % (v/v) PEG500 DME, 0.1 M sodium citrate tribasic dihydrate pH 5, 0.1 M sodium chloride imaged after 5 days using polarized light. **(B)** Crystals obtained in 20 % (v/v) PEG1000, 0.1 M HEPES/sodium hydroxide pH 7.5, 200 mM lithium sulphate imaged after 10 days using visible light.

concentrated to a final concentration of 34 mg/mL (according to NanoDrop measurements). SDS-PAGE-analysis revealed that after concentrating, small amounts of dimer were present in the final sample, visible when loading high quantities of protein on the gel. However, at lower concentrations, only a single protein band was visible, corresponding to the monomeric receptor (Figure 5.3 B). The identity of the receptor was further confirmed by mass spectroscopy (MS) analysis of the sample (Figure 5.3 C). The small difference from the expected size (10 Da) can be explained through oxidation or other small posttranslational modifications.

The protein was therefore used for crystallisation. To this end, it was reconstituted into LCP and initial crystallisation trials were carried out in commercially available sparse matrix screens including the MemMeso HT-96/FX-96, the Rigaku Wizard LCP, and the JBS-LCP-HTS screen. All trials were set up in 96-well LCP glass sandwich plates with 200 µm spacers.

When crystallising in LCP, it can be hard to distinguish contaminations or certain features of the cubic phase itself from protein crystals. However, crystals are known to rotate plane-polarized light, except for crystals with cubic space groups, such as sodium chloride crystals, and certain space groups when viewed down their symmetry axes. Therefore, cross-polarized images were routinely taken to assist the identification of crystal hits. Typical rod-shaped A_{2A} R crystals appeared after three days in 30 % (v/v) PEG500 DME, 0.1 M sodium citrate tribasic dihydrate pH 5, 0.1 M sodium chloride (Figure 5.4 A), and after ten days in 20 % (v/v) PEG1000, 0.1 M HEPES/sodium hydroxide pH 7.5, 200 mM lithium sulphate (Figure 5.4 B). In both conditions, crystals reached a maximal size of 22–27 µm. The first hit highly resembled published A_{2A} R crystallisation conditions (*e.g.* 3EML [188]) and therefore suggested that the crystals were indeed protein crystals. About 20 crystals were harvested and sent for data collection to the Swiss Light Source. Surprisingly, none of these crystals diffracted properly. However, there were few low-resolution diffraction spots near the beam stop, at least indicating that the crystals were protein.

The A_{2A} -StaR2-bRIL construct was specifically chosen because it allows crystallisation of receptor-ligand complexes with low affinity compounds, such as theophylline. Therefore, an identical purification of A_{2A} -StaR2-bRIL^{HRV-HIS} was repeated, but, in this instance, theo-

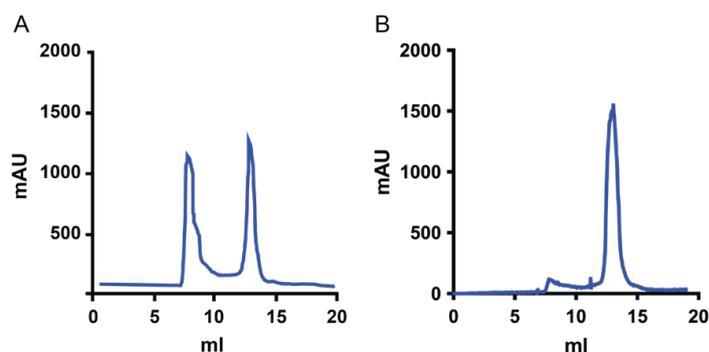


Figure 5.5: Improvement of SEC profile for A_{2A} -StaR2-bRIL^{HRV-HIS}-theophylline complex. (A) SEC profile before changing the buffer conditions, displaying a significant fraction of higher oligomers or aggregation. (B) After adapting the purification protocol the first peak was significantly reduced resulting in a single monodispersed peak.

phylline was retained throughout the entire purification process. Theophylline shows some absorption at 280 nm, interfering with nanodrop protein quantification; therefore, protein concentration was additionally estimated through densitometry on SDS-PAGE. This analysis suggested that the final protein sample was concentrated to 35 mg/mL before reconstituting into LCP. Crystallisation trials with the A_{2A} -StaR2-bRIL^{HRV-HIS}-theophylline complex were carried out using the commercially available MemMeso screen as well as two home-made optimisation screens. The HEPTARES-cond PEG screen was set up around the crystallisation condition for the A_{2A} -StaR2-bRIL^{HRV-HIS}-theophylline complex described by Cheng *et al.* [221]. The second screen contained different combinations of all published A_{2A} R crystallisation conditions compiled from a statistical design of experiment analysis in Ellistat[©]¹, further denoted as DOE screen. Full details about both of these screens can be found in the methods (see chapter 2.6.1). No crystals were obtained from any of these screens.

In an attempt to improve protein quality, the protein purification buffers were modified to match the conditions in Rucktooa *et al.* [50] exactly. The main changes consisted in switching the buffers from 50 mM HEPES pH 7.4 to 40 mM Tris pH 7.4 and reducing the sodium chloride concentration from 0.8 M to 0.2 M. These changes indeed seemed to improve protein quality, as the aggregation/multimer peak in the SEC profile was significantly reduced (Figure 5.5). This time, the protein concentration was additionally determined using the detergent compatible BCA assay from BioRad (DC-assay). While the final protein concentration determined by densitometry resulted in 36 mg/mL the DC-assay indicated that it might be even higher (68 mg/mL). There might be differences in absorbance of the Coomassie stain between the A_{2A} R and the BSA used as a standard, so densitometry might be a rather inaccurate method for determining protein amounts, explaining the discrepancy observed. 68 mg/mL would be higher than the protein concentration used for crystallisation by Rucktooa *et al.* [50], but around the concentration used with other constructs [150]. Despite the optimisation in protein quality, no crystals were obtained. There was also no heavy aggregation in any of the crystallisation conditions, which would have indicated that the protein concentration might have been too high.

¹<https://ellistat.com>

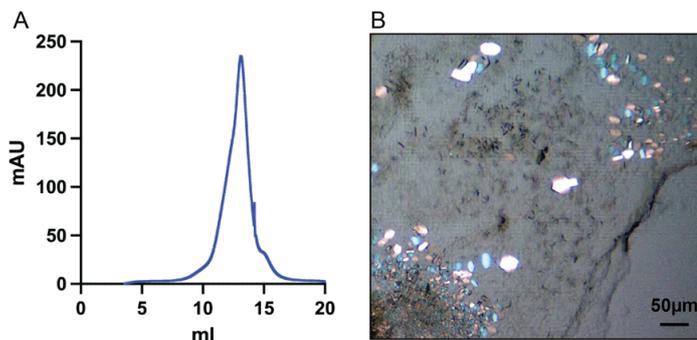


Figure 5.6: Results for A_{2A}-StaR2-bRIL purification without the HRV cleavage site. (A) SEC profile, displaying a single monodispersed peak. (B) crystals obtained in complex with the reference ligand theophylline. The best crystals had a hexagonal, plate like shape, quite different from the crystals previously obtained with ZM241385.

Because I didn't obtain any diffracting crystals, I decided to compare the protein sequence of the construct obtained from Novartis to the literature, revealing that that the original construct did not contain an HRV-cleavage site between the end of the gene and the C-terminal HIS-tag. Minuscule things can have big impacts in crystallography. Therefore, the cleavage site was removed through infusion cloning (Takara-Bio). The modified construct was reassessed by sequencing and then cloned into a pFastBacTM vector for expression in *Sf9*-cells.

The corrected A_{2A}-StaR2-bRIL construct was then purified following the *exact* protocol described in Rucktooa *et al.* [50]. In addition to the change in buffer (see above, chapter 5.4), I also used a prepacked nickel-column instead of loose cobalt-resin and did no overnight incubation. This resulted in one single, monodispersed peak in the SEC profile (Figure 5.6 A). The protein was concentrated to a final concentration of 28 mg/mL (DC-assay) before reconstituting into LCP for crystallisation. As the affinity for theophylline is rather low, 1 mM compound was also added to the monoolein as well as the crystallisation screen. Crystallisation trials were set up using the HEPTARES-cond PEG screen and a variety of different glass plates (Swissci Modular Glass Base, Swissci Modular SBS XPOL Glass base and Jenna bioscience glass base), spacers (100 μ m and 200 μ m) and covers (glass plate, 100 μ m film and 40 μ m film). Additionally, one plate using the DOE screen and the Jenna bioscience glass plate and glass cover was set up. Some small microcrystals were obtained in all of the plates. However, the best crystals grew in 25-33% (v/v) PEG400, 0.05 M NaSCN, 2% (v/v) 2,5-hexanediol, 0.1 M sodium citrate pH 6-6.5, 0.5 mM theophylline using Swissci Modular glass bases, with a 100 μ m spacer and a plastic cover film. Crystals were of a plate like, squared to hexagonal shape and grew to a maximal size of 30 μ m \times 37 μ m (Figure 5.6 B).

Now, being able to crystallise the A_{2A}-StaR2-bRIL in presence of theophylline, this construct could be used for crystallisation with the novel compounds. As I had obtained many conditions with nice large crystals in this first experiment, most of them were already used for subsequent soaking studies, before assessing protein diffraction of these crystals.

5.5 Crystallisation trials with A_{2A}-StaR2-bRIL construct

5.5.1 Crystal soaking

Crystal soaking is one of the most common approaches to solve many different receptor-ligand complexes. Rucktooa *et al.* [50] recently described a method how this should be possible for the A_{2A}R crystallised in LCP. In their approach, highly stabilised A_{2A}-StaR2-bRIL construct was purified and crystallised in the presence of the low affinity compound theophylline, which was subsequently replaced by other ligands. The novel Novartis ligands selected for crystallographic studies displayed a higher affinity than theophylline. Thus, using the same approach, it should be possible to outcompete theophylline from the binding site when adding the novel ligands in excess to the crystals. Furthermore, this strategy only uses minimal amounts of ligand and should therefore be applicable to most of the novel ligands.

The A_{2A}-StaR2-bRIL construct was purified as described above and concentrated to a final concentration of 40 mg/mL (nanoDrop) with 1 mM theophylline present and mixed with monoolein supplemented with 10 % (w/w) cholesterol and 10 µM theophylline. The final protein to lipid ratio was 2:3 (v/w), resulting in about 30 µL of protein laden LCP. Crystallisation trials were performed in 96-well glass sandwich plates and sealed with a plastic cover film. Crystals were grown for 18 days, reaching an average size of 15–25 µm (Figure 5.7 A), the largest ones reaching a maximal size of 50 µm. To soak the crystals, a cross-wise incision was made into the cover over the crystal containing wells using a sharp scalpel. This allowed a small corner of the cover to be lifted and to flood the well with 10 µL mother liquor supplemented with 1 mM ligand, either FOC-1, FOC-2, or DEL-1, resulting in a final ligand concentration of 925 µM and theophylline concentration of 74 µM. The wells were resealed with crystal clear sealing tape and incubated for either one or 24 hours.

Despite this process sounding very similar to the regular soaking process for sitting drop crystallisation experiments, it is in fact far more complex. The main reason for this is the small well depth of only 100–200 µm, depending on the spacer used. When making the incision into the cover film, the applied pressure very easily squeezes the entire well, which can cause the entire LCP drop to detach and float around the well. This effect is especially pronounced if the mesophase is more liquid and has turned into sponge phase. As soon as any of the mother liquor liquid touches the rim of the well, the entire well content gets drawn underneath the spacer and any crystals present are lost. In addition, once a well is opened further manipulations should happen as quickly as possible to prevent the drop from drying and any phase transitioning of the mesophase that could harm the crystals. All of this makes crystal handling extremely difficult and led to the loss of many promising crystals before they could actually be soaked or harvested.

Crystals of 25 µm length are still relatively small for single crystal diffraction experiments. Therefore, crystals were harvested using mesh loops, so a large fraction of the LCP with several crystals could be harvested at once, allowing for serial data collection on a single loop. However, as the mesophase gets scooped up as well in this process, it becomes almost impossible to see any crystals on the mesh once mounted on the goniometer at the beamline (Figure 5.7 B). None

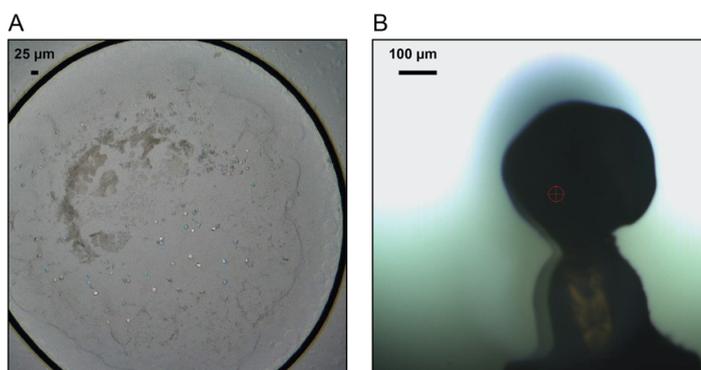


Figure 5.7: Crystal soaking of A_{2A} -Star2-bRIL-theophylline complex. (A) Example of theophylline crystals used for soaking experiments. (B) mesh loop mounted on goniometer carrying a big, dark dollop of LCP in which it is impossible to optically locate any crystals.

of the harvested samples showed any diffraction and it was impossible to judge if this was due to bad crystal quality or if crystal harvesting was unsuccessful because of all the encountered difficulties.

5.5.2 Ligand exchange

In a second approach, the ligand was added to the purified protein, just before setting up the crystallisation plates (Figure 5.1). This reduces crystal handling compared to the soaking approach and, therefore, the chances of crystals getting lost in the process, while still only using small amounts of ligand. As the protein is already highly concentrated at this point, it is important to consider the molarity of the protein, to ensure the ligand is actually added in excess. For this experiment A_{2A} -Star2-bRIL was concentrated to 35 mg/mL, which corresponds to 720 µM. This is already quite high and therefore, only the best candidate with the highest affinity and thermostabilising effect from each screening series were chosen for subsequent experiments. The affinity of FOC-2 and DEL-2 is about 500 times higher than what was measured in the same experiment for theophylline. Thus, only a small excess should be sufficient for complete ligand exchange. Further, the high affinity reference compound ZM241385 was included in the experiments as proof of principle.

The second important aspect for this approach was to keep the final DMSO concentration as low as possible, as it might otherwise affect crystallisation. Consequently, the final ligand concentration depended on its solubility in DMSO: for FOC-2 a stock solution of 200 mM could be prepared, but DEL-2 and ZM241385 could only be prepared at 100 mM, resulting in a maximal final concentration of 3.3 mM and 1.7 mM respectively, while keeping the DMSO concentration below 2%.

To further increase the chance of a complete ligand exchange, the protein sample was warmed to room temperature before incubating for 30 minutes with the novel ligand. This allows for quicker equilibration and thereby hopefully a complete ligand exchange by the time the

Table 5.1: Summary of collected data for A_{2A}-StaR2-bRIL ligand exchanged from theophylline to ZM241385. Despite some pipelines finding a result, the data was too poor for proper analysis, as there were too few observations recorded and the cell parameters did not match what was previously observed. Statistics for the highest-resolution shell are shown in parentheses.

Data set	9471-D2	9486-D2-2	9486-D2-3
Pipeline	fast_dp	xia2 dials	xia2 dials
Space group	P222	C121	P222 ₁
Cell dimensions			
a, b, c (Å)	54.09, 85.47, 95.44	155.69, 58.16, 104.34	80.32, 87.50, 143.82
α, β, γ (°)	90, 90, 90	90, 90.97, 90	90, 90, 90
Resolution (Å)	7.36–23.86	1.23–54.48	1.23–35.95
Measured reflections	191 (17)	25 007 (25)	98 161 (59)
Unique reflections	131 (11)	20 334 (24)	65 485 (59)
Completeness (Å)	18.8 (22.7)	7.5 (0.2)	22.3 (0.4)

crystals are set up. Additionally, an effort was made to keep the ligand concentration high in all subsequent steps. To this end, 100 μ M ligand was added to the lipid as well as the HEPTARES-cond PEG crystallisation screen.

For ZM241385 several conditions on the HEPTARES-cond PEG optimisation screen, as well as the commercial MemMeso screen yielded crystals. The crystal yielding conditions on the random screen were very similar to the condition used for optimisation screen, but the obtained crystals were very small. One of the reasons for this might be that no additional ligand was added to this screen. In contrast, the HEPTARES-cond optimisation was supplemented with 100 μ M ZM241385 and used to set up crystal trials on glass plates with either glass or plastic covers. Crystals were obtained in both plate types. The largest crystals grew in 22.55% PEG400, 50 mM NaSCN, 2% (v/v) 2,5-hexanediol, 0.1 M sodium citrate pH 5.5. Interestingly, in the plate with the plastic cover, the crystals grew slightly larger, reaching a maximal length of 40 μ m, while with the glass cover, they only grew up to 30 μ m (Figure 5.8 A). Crystal shape was more elongated and rod like compared to the previously obtained theophylline crystals, which is in line with what was observed for A_{2A}-ZM241385 complexes (V.-P. Jaakola, *pers. com.*). Crystals smaller than 10 μ m did not diffract. The slightly bigger crystals showed some weak protein diffraction that was not good enough for data collection. Data could only be collected and auto processed using the xia2 dial or fast dp pipeline from the largest crystals (30–40 μ m). However, for all three recorded datasets, the correct space group and cell parameters could not be determined, and overall data quality, number of measured reflections, and completeness were too poor to justify further processing (Table 5.1).

The receptor also crystallised after ligand exchange with FOC-2. The obtained crystals were short needles, similar to the ones observed for ZM241385, however, significantly smaller. The best crystal hits grew in 25.6–27.6% PEG400, 50 mM NaSCN, 2% (v/v) 2,5-hexanediol, 0.1 M sodium citrate pH 5.5, reaching a maximal length of 10–15 μ m (Figure 5.8 B). Interestingly, using the plastic cover resulted in a larger number of crystals, while crystal size was not significantly affected. Similar to the ZM241385 complex, there was also one hit on the MemMeso

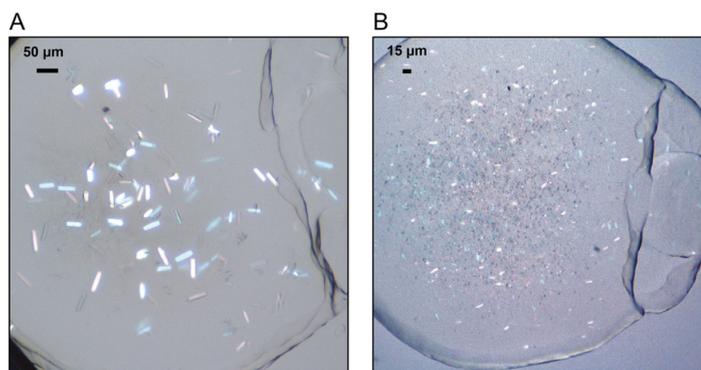


Figure 5.8: Crystals obtained after ligand exchange. The A_{2A}-StaR2-bRIL construct was copurified with theophylline. Before setting up the LCP excess of (A) ZM241385 and (B) FOC-2 was added to the protein solution.

screen, in a condition closed to the optimised one. While these crystals were even smaller than the ones on the optimisation screen, they seemed to have more of a second dimension. Unfortunately, none of the obtained crystals diffracted.

In contrast, even after 34 days, no crystals appeared in any of the plates set up with DEL-2. As the receptor does crystallise in the presence of theophylline alone, this indicates that ligand exchange did happen; however, it might have either occurred only partially, resulting in a conformational mixture which does not crystallise, or binding of DEL-2 destabilises the receptor significantly, preventing crystallisation. This is quite surprising, as the thermostabilising effect measured for DEL-2 was, despite being rather small, still significantly higher than the one of theophylline.

5.5.3 Co-purification

Ligand exchange to FOC-2 on purified protein did yield crystals, but they were of very poor quality. This might be due to an incomplete ligand exchange, introducing heterogeneity into the protein sample. To eliminate any crystal defects that might arise from a potential mixture of two ligand-receptor complexes, a co-purification with FOC-2 was performed. To this end, the purification protocol was slightly modified to minimise the volumes of buffer used once FOC-2 was added. This mainly involved changing from a pre-packed nickel column to loose resin in a gravity driven column, which allowed a reduction in total column volume and omitted the need to prepare excess buffer, as required when using an automated purification system. A_{2A}-StaR2-bRIL was still solubilised in the presence of theophylline and the ligand was then exchanged once bound to the affinity column. To keep the ligand concentration high after reconstitution of the receptor into LCP, 100 µM FOC-2 was added to the lipid, as well as the optimisation screens. Besides the HEPTARES-cond PEG optimisation screen, a MES-additive screen was prepared that was based around published crystallisation conditions for A_{2A}-StaR2-bRIL in complex with several ligands other than theophylline or ZM241385 (for more details see methods chapter 2.6.1). Some microcrystals were obtained in 25.1% (w/v) PEG400, 0.05 M sodium thiocyanate, 2% (v/v) 2,5-hexanediol, 0.1 M tri-sodium citrate pH 5.5 and 0.1 mM FOC-2 and some slightly larger ones with 37.3% (w/v) PEG400, 0.2 M potassium/sodium tartrate, 0.5% (w/v) 1,2,3-

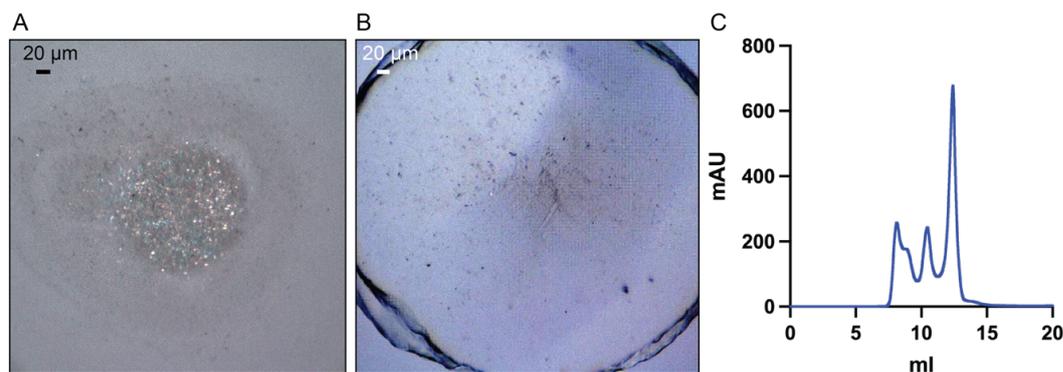


Figure 5.9: Co-purification of A_{2A} -StaR2-bRIL in complex with FOC-2. (A) Microcrystals obtained 37.3% (w/v) PEG400, 0.2 M potassium/sodium tartrate, 0.5% (w/v) 1,2,3-heptanetriol, 0.1 M MES pH 6 and 0.1 mM FOC-2. (B) The same condition but without additional FOC-2 did not yield any crystals. (C) SEC chromatogram of A_{2A} -StaR2-bRIL-FOC-2 complex. Significant amounts of higher oligomers and possibly aggregates are present (first two peaks).

heptanetriol, 0.1 M MES pH 6 and 0.1 mM FOC-2. Interestingly, the latter did not appear if using the screen without any additional compound added, suggesting that off rates of FOC-2 might be relatively high and therefore, an excess of ligand is required to maintain enough receptor in the ligand-bound state to facilitate crystal formation (Figure 5.9 A, B).

Despite being too few and too small for a successful data collection, these crystal hits provided a good starting point for further optimisation. Firstly, protein quality needed to be enhanced. The protein preparation resulted in similar yields as previous purifications with this construct. However, SEC showed that there were two populations of higher oligomers present in the sample, which could not be completely separated from the monomer peak (Figure 5.9 C). Subsequently, about half of the protein was lost on the concentrator while concentrating for crystallisation, which might have led to some aggregation in the sample as well.

5.5.4 Improving protein quality: Co-purification and crystallisation with Theophylline

One of the most important things for a successful crystallisation experiment is a high quality, homogeneous protein sample. In most protein preparations used for crystallography, SEC revealed the presence of at least one higher oligomeric species, which could not be completely separated from the monomer. Thus, the first step in optimising any crystals was to improve protein quality and reduce the amount of higher order oligomers in the sample. As the amount of ligand available was limited, these optimisation experiments had to be performed in the presence of a reference compound. Additionally, using a reference compound has the advantage that the exact conditions required for well diffraction crystals are known. ZM241385 is the most commonly used compound for crystallisation studies, as it strongly stabilises the receptor (28 out of the 59 published structures used ZM241385¹). However, getting a high-quality protein

¹<https://gpccrdb.org/structure>, latest update 31.01.2022

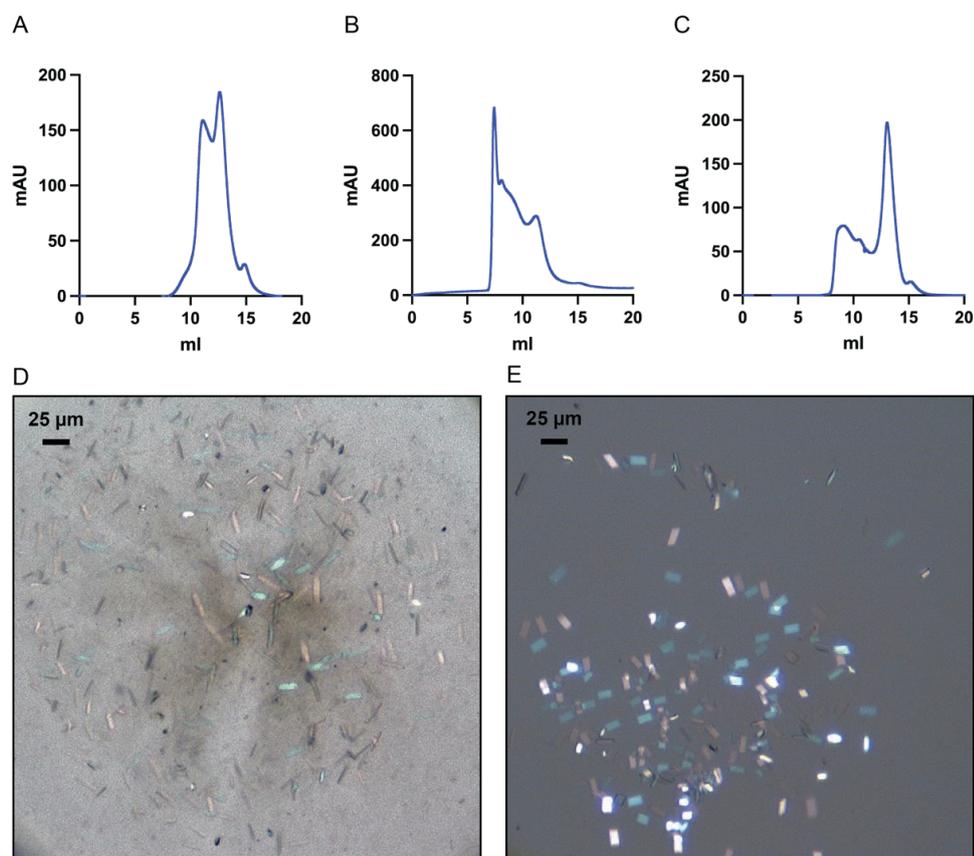


Figure 5.10: Attempt to optimise A_{2A}-StaR2-bRIL-theophylline protein quality. (A) exact repetition of first purification as performed in Figure 5.6. This time a dimer-peak was present, which could not be completely separated from the monomer peak. (B) Chromatogram after increasing the sodium chloride concentration to 0.8 M and changing the detergent to DDM with CHS. (C) Chromatogram for the purification using the same buffer conditions as in A but using the Äkta Pure system to automate the affinity purification step. (D) crystals obtained with the receptor purified in A. (E) crystals obtained with the receptor purified in C.

sample with a less stabilising ligand might be more difficult and require some additional receptor stabilisation. Therefore, theophylline was chosen as a reference ligand, as its thermostabilising effect was similar or in some cases even lower than the one measured for the novel compounds.

A single, monodispersed peak was obtained when the protein was first purified in presence of theophylline for the crystal soaking experiments (Figure 5.6 A). However, since the best crystals were used and potentially destroyed in the soaking experiment, diffraction with theophylline alone could not be verified. Hence, the first step was to reproduce this purification and validate diffraction of the crystals. Surprisingly, even after several attempts these previous results could not be reproduced. Each time there was a population of dimers present that could not be completely separated by SEC from the monomer fraction (Figure 5.10 A). Despite this, crystals could be obtained using the HEPTARES-cond PEG optimisation screen. However, they were rather small and either impossible to harvest, or did not diffract (Figure 5.10 D).

The purification protocol used to purify the A_{2A}-StaR2-bRIL construct was based on the methods published by Cheng *et al.* [221]. Interestingly, this protocol differs quite a bit from protocols used with other constructs employed for LCP crystallisation. To determine the contribution of different buffer components, a small-scale comparison of different conditions was performed (data not shown). Changing the detergent to DDM with CHS as well as using a higher sodium chloride concentration significantly increased protein yields and consequently should have helped to stabilise the receptor. Therefore, for the next large-scale experiment, the protein was solubilised in DDM/CHS and the sodium chloride concentration was increased to 0.8 M. These changes did indeed increase protein yields significantly; unfortunately, this led to even more aggregation than previously observed (Figure 5.10 B).

Consequently, buffer conditions were maintained as previously used. Instead, some technical changes were made to the protocol to reduce the purification time. This mainly included the reduction of the pre-incubation time with ligand before solubilisation to half an hour, as well as reducing the temperature for this incubation to 4 °C, and the automation of the affinity purification using an Äkta Pure system. Through these changes, the entire purification could be performed in only a single day, whilst minimising the chance for technical mistakes. Indeed, looking at the size exclusion chromatogram suggests that protein quality did improve, as the oligomeric fraction was reduced and better separated from the monomer peak (Figure 5.10 C). This was also reflected in the obtained crystals, which were significantly bigger than in previous experiments, reaching a maximal length of 45 µm (Figure 5.10 E). Crystals were harvested by two different people, to minimise the chances of all the crystals being destroyed during the harvesting process. Despite all this effort, no protein diffraction could be recorded.

5.6 Crystallisation trials with the A_{2A}-bRIL construct

As the initial results with the A_{2A}-StaR2-bRIL construct could never be reproduced and all subsequently obtained crystals did not diffract, a different receptor construct was used for further crystallisation trials. Using the A_{2A}-bRIL construct without the thermostabilising StaR2-mutations allowed structure determination of the A_{2A}R-ZM241285 complex at almost the same resolution as the A_{2A}-StaR2-bRIL construct [150]. However, no co-structures with low-affinity compounds have been determined using that construct so far. Therefore, soaking experiments are not possible with this construct. Nonetheless, in an alternative approach towards determining multiple co-crystal structures of a GPCR target in complex with novel ligands, the authors demonstrated that by using the A_{2A}-bRIL construct a transient ligand could be replaced by an excess of a high-affinity ligand added just before crystal formation [301]. As the best candidates from this study had relatively high affinity in the two-digit nanomolar range, crystallisation attempts with a less stabilised receptor might have been possible.

5.6.1 Crystallisation and data collection with ZM241385

Because I struggled with the purification of the first construct, in a first instance, crystallisation and diffraction of the A_{2A}-bRIL construct was assessed in complex with the reference compound ZM241385. This would allow the effects caused by protein quality to be distinguished from the ones caused by insufficient stabilisation by the ligand. The original A_{2A}-bRIL construct was

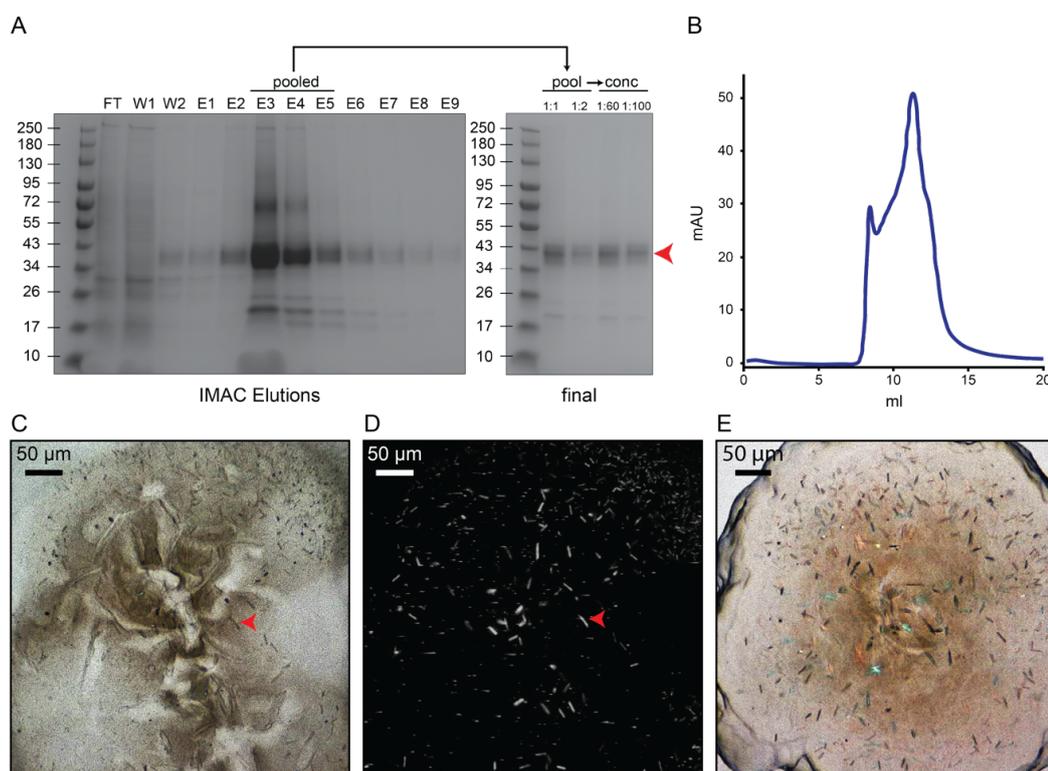


Figure 5.11: Purification and crystallisation of A_{2A}-bRIL construct in complex with ZM241385. (A) SDS-PAGE after IMAC purification. FT = flow-through after loading the cobalt resin into gravity column (9 μ L of 1:10 dilution loaded), W1/W2 = column washes with different concentrations of imidazole (9 μ L loaded) and E1–E9 = imidazole elutions (9 μ L loaded). Fractions E3–E5 were pooled (1 μ L of undiluted and 1:2 diluted pool loaded on the gel on the right) and concentrated (samples labelled conc on the right-hand side gel, 1 μ L of 1:60 and 1:100 diluted final sample). (B) SEC chromatogram recorded the next day using a small aliquot of the final, concentrated sample. (C) Crystals obtained using fresh protein imaged after 3 days with visible light and (D) second harmonic generation (SHG) imaging. The red arrow heads point out the same position of one single crystal. (E) crystals obtained with protein after storage at -80°C .

obtained in a pFastBac vector from Novartis, while the bacmid and subsequent virus were produced and titered as described previously. The receptor was then expressed and purified following the protocol reported in Liu *et al.* [150]. The most striking difference to the protocol used for purification of the A_{2A}-StaR2-bRIL construct is that no size exclusion step is performed. Instead affinity chromatography was performed using more selective cobalt resin. Then, only the highest concentrated elution fractions were pooled to minimize the amount of detergent accumulating in the subsequent concentration step. The receptor was concentrated to 60 mg/mL (NanoDrop). As observed with the first construct, the determined protein concentration varied depending on the quantification method. Interestingly, for this construct in complex with ZM241385, the concentration determined by densitometry on the Coomassie stained gel, as well as the DC assay, was higher compared to that determined on NanoDrop, resulting in about 100 mg/mL, while for the A_{2A}-StaR2-bRIL-theophylline complex it was usually lower. Because there was no equipment for analytical size exclusion available at the time, protein purity was initially only monitored by SDS-PAGE (Figure 5.11 A), and size exclusion analysis could only

be performed 24 hours after finishing the purification and setting up crystallisation trials. The chromatogram revealed that, similarly to what was observed before, there was a fraction of dimeric receptor present that could not be separated from the monomeric peak (Figure 5.11 B). However, in this case, the dimer peak was significantly smaller than the monomer peak. Further, some of the oligomerization might have occurred during the 24 hours the protein was stored on ice before it could be analysed.

Crystallisation trials were set up using a home-made optimisation screen (subsequently called SCRIPPS 1) based on the published conditions for the A_{2A}-bRIL-ZM241385 complex [150], while varying the PEG400 concentration and the pH. Small, rod-like crystals appeared after twelve hours. Due to limited beamtime, crystals were harvested after only three days, when they were only 20 μm long and sent to P14 at the German Electron Synchrotron DESY for data collection. A few crystals showed clear protein diffraction out to about 3.7 \AA and data collection was attempted for several crystals (Figure 5.11 C, D). Unfortunately, these crystals did not survive long enough in the beam to collect an entire data set. To get better data, more and bigger crystals would be required. Providentially, this construct proved to crystallise even after freezing and storage at -80°C , therefore, protein from the same purification could be used to optimise crystallisation. Using the same screen and protein, but preparing the LCP freshly and leaving the crystals to grow for longer was already enough to get several conditions with crystals at higher density. After eight days they reached a maximal length of about 40 μm (Figure 5.11 E). Furthermore, they also grew in width and their shape could now be best described as “Z-shaped” compared to the initial more rod-like crystals. These improved crystals were likely to give better diffraction, indicating that there was no problem with the construct or the protein preparation. As there was still a lot of protein from this preparation available, ligand exchange trials were set up before sending off any crystals for data collection.

5.6.2 Attempts to exchange ZM241385 with novel ligands

Despite the A_{2A}R having a very high affinity for ZM241385, it may still have been possible to exchange it for slightly lower-affinity ligands if they are supplemented in excess. To this end, ligand in DMSO at the highest concentration possible was added to the final protein in a 1:100 ratio, to not exceed 1 % final DMSO and then incubated at room temperature for half an hour to allow for equilibration. FOC-2 and DEL-1 were soluble at 200 mM while for FOC-1 only 100 mM could be reached in DMSO, resulting in a final ligand concentration of 2 mM and 1 mM respectively, which is 20–40 times higher than the ZM241385 concentration.

To expand the PEG400 concentration- and pH-range tested, two fresh optimisation screens were set up around the conditions published for 4E1Y (SCRIPPS 2, 4). Furthermore, both screens were set up additionally without any hexanediol (SCRIPPS 3, 5). Because of the limited amount of ligand available no additional ligand could be added to the screens. Nonetheless, crystals were obtained in almost all of the plates. Ligand exchange with FOC-1 yielded the fewest conditions with crystals and the obtained crystals were relatively small compared to the trials set up with the other ligands (Figure 5.12 A). Slightly more crystals were obtained on the screens without hexanediol. Similarly, the sample with added DEL-1 seemed to crystallise better without hexanediol. For this sample, crystals were generally slightly larger than for FOC-1, but the crystal density in a single drop was usually lower (Figure 5.12 B). The largest crystals

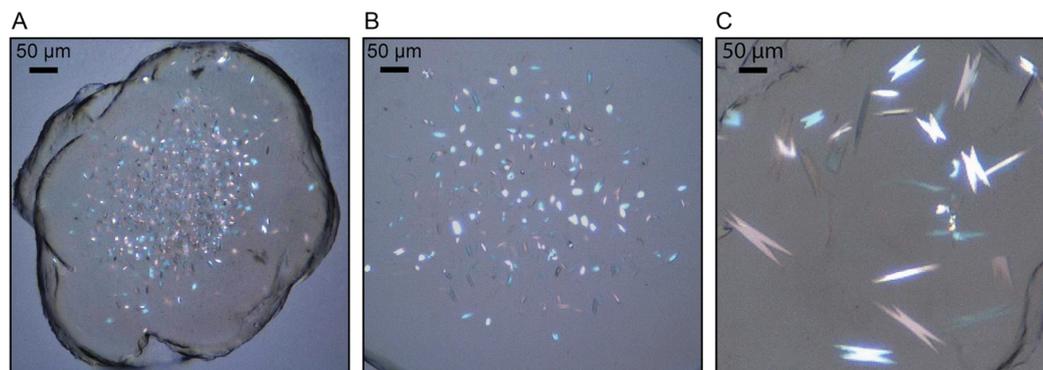


Figure 5.12: Ligand exchange attempting to displace ZM241385 with three novel compounds. Excess of FOC-1 (A), DEL-2 (B) and FOC-2 (C) was added to concentrated receptor bound to ZM241385. Crystals were imaged after eight days using polarized light.

were obtained after ligand exchange with FOC-2, reaching a maximal size of $10\ \mu\text{m} \times 60\ \mu\text{m}$ (Figure 5.12 C). Interestingly, this sample seemed to crystallise better with hexanediol present in the screen. However, for all three samples, crystal size varied a lot depending on the exact mother liquor composition and it was therefore difficult to judge if the observed differences were ligand induced or due to general variability. Hence, it was not possible to judge if the ligand exchange was successful at this point.

As the crystal density in most drops was very high, but most of the crystals still relatively small ($<30\ \mu\text{m}$), they were harvested using mesh loops, which facilitated serial data collection from a single loop. Data were collected at the I24 beamline at Diamond Light Source using a semi-automated serial approach. To this end, crystals were located in the loop using a grid scan. For data collection 10° (100 images) from many different crystals were recorded at 50% transmission and 0.01 s exposure. Data were collected for all ligand-exchanged samples, with the best crystals diffracting to $2\ \text{\AA}$ resolution. The diffraction data was auto-processed using the Xia2.multiplex pipeline. For most of the recorded data sets the obtained space group and cell dimensions were practically identical to previously deposited data [150] (Supplementary Table B.5). Molecular replacement solutions were found in XDS using the 4E1Y structure with additional features such as the ligand, Na^+ and lipids removed. Immediately after molecular replacement an unambiguous electron density in the orthosteric binding site could be identified. In all cases, this density clearly fitted ZM241385 and not any of the novel ligands (Figure 5.13). Each data set underwent three cycles of limited initial refinement, further corroborating the identity of the ligand as ZM241385. Therefore, a full refinement was not carried out, since exactly the same structure has already been solved previously at even higher resolution.

To enhance the chance of ligand exchange, the ZM241385 concentration of the protein sample was reduced through several rounds of diluting and re-concentrating to $0.3\ \text{nM}$. Then $2\ \text{mM}$ FOC-2 were added, so that the final ligand concentration was over 6000 times higher than the ZM241385 concentration. Unfortunately, this process also led to significant protein losses, so that the final protein concentration was only $35\ \text{mg/mL}$, and the volume was only large enough to set up one crystal plate. No crystals were formed, indicating that either protein concentration was too low or the newly formed FOC-2 complex not stable enough to crystallise.

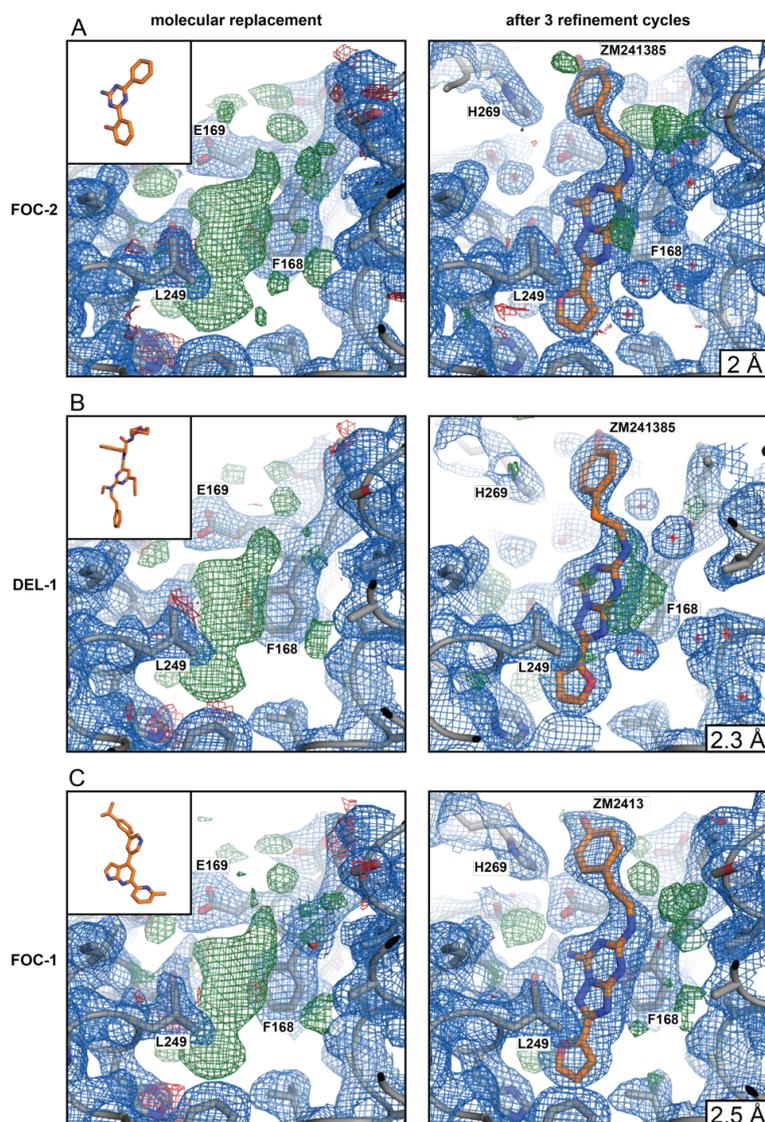


Figure 5.13: Electron density maps for the ligand binding site of A_{2A} -bRIL-ZM241385 complex after the attempted ligand exchange, showing that ZM241384 could not be displaced by the novel ligands. The $2F_o-F_c$ map for key residues in the ligand binding pocket is shown in blue ($\sigma = 1$) and F_o-F_c maps in green ($\sigma = 3$). Maps are shown straight after molecular replacement (left panel) and after 3 rounds of refinement (right panel) for datasets obtained from crystals after ligand exchange with (A) FOC-2, (B) DEL-1, and (C) FOC-1. The structure of the added ligand is depicted in the overlaid box for each panel. In all datasets the difference density clearly fits ZM241385 and not the novel ligand. This is corroborated after refinement with ZM241385. While there is some extra difference density remaining in the central section of the DEL-1 exchanged sample after refinement with ZM241385, it is unlikely this comes from DEL-1. DEL-1 is significantly larger and more branched than ZM241385 and would therefore not fit the rest of the density.

5.6.3 Ligand exchange LUF5834 and theophylline

As pointed out previously, ZM241385 is a very high-affinity ligand, capable of tightly binding to the receptor and significantly stabilising it in the antagonist bound conformation [188]. Thus, exchanging ZM241385 might be challenging. A more transient ligand is therefore the obvious choice when looking for a ligand to stabilise the receptor during the purification process, before attempting crystallisation with novel compounds. An additional advantage of using such transient ligands is that they do not stabilise the A_{2A}-bRIL construct enough to allow for crystal formation. Thus, if crystals appear, this is a strong indication that the ligand exchange was successful. One of the ligands used in previously reported ligand-exchange experiments with A_{2A}R is LUF5834, a high-affinity but transient partial agonist [301]. Therefore, in a next step, the receptor was purified in the presence of 50 μ M LUF5834 and concentrated to a final concentration of 66 mg/mL (Figure 5.14 B). Subsequently, the protein sample was incubated with 2 mM ligand for 30 minutes. In contrast to the ligand-exchanged experiments described by Ishchenko *et al.* [301], no additional ligand could be added to the crystallisation screen, as not enough material was available. As negative control, crystallisation trials were also set up with the LUF5834-complex, to test that no other component in either the protein sample or the crystallisation conditions was prone to the formation of salt or detergent crystals. Unfortunately, all ligand-exchange experiments set up with the novel ligands were unsuccessful. However, addition of 2 mM ZM241385 did yield very small, rod-like crystals (Figure 5.14 C). These crystals were significantly smaller and their shape was less well defined than any of the previously observed ZM241385 crystals.

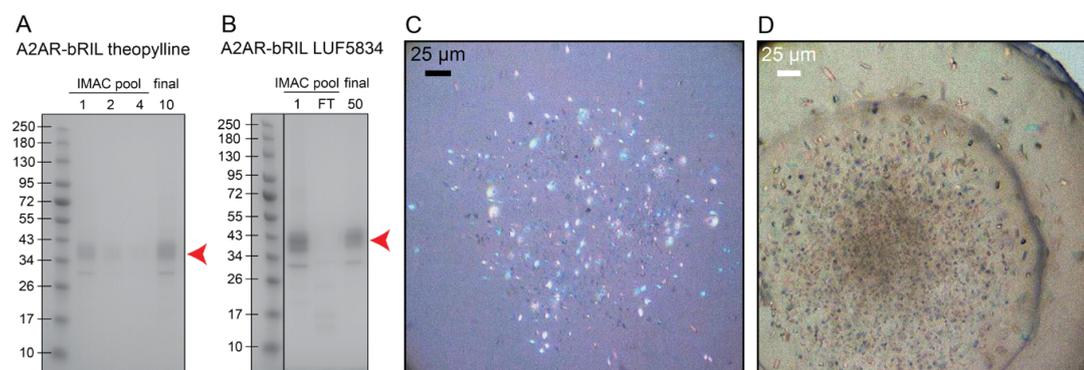


Figure 5.14: Ligand exchange experiments using A_{2A}-bRIL purified in complex with theophylline or LUF5834. (A) SDS-PAGE of receptor-theophylline complex after IMAC purification. Protein containing elutions were pooled (IMAC pool) and 1 μ L of the undiluted and 1:2 and 1:4 diluted sample were loaded on the gel. The pooled fractions were concentrated and 1 μ L of 1:10 diluted final sample was loaded as well (numbers above the band indicate the dilution factor). (B) SDS-PAGE of receptor-LUF5834 complex after IMAC purification. Protein containing elutions were pooled (IMAC pool) and 1 μ L of the undiluted sample was loaded on the gel. The pooled fractions were concentrated and 1 μ L of 1:50 diluted final sample as well as 9 μ L the flow through (FT) from the concentrator were loaded as well. Monomeric receptor is pointed out by the red arrow heads. (C) Crystals obtained after ligand exchange of LUF5834 and (D) theophylline protein with ZM241385.

Despite being regarded as transient ligand, the reported K_i of LUF5834 is still in the low nanomolar range [302], which is lower than the affinity determined for most of the novel ligands. In contrast, the affinity determined in the radioligand-binding experiments for the reference compound theophylline was significantly lower. Therefore, this ligand was further explored for subsequent ligand-exchange experiments. To this end, I used the identical purification protocol as before, but maintained 1 mM theophylline throughout the entire process (Figure 5.14 A). For the ligand-exchange experiment the two ligands with the highest affinity and solubility in DMSO, FOC-2 and DEL-1, were chosen, as well as ZM241385 as a control. In each case, 2 mM of ligand was added to the final protein sample, which was concentrated to 36 mg/mL (DC assay). Besides the 4E1Y-optimisation screens (SCRIPPS 2–5), the commercial MemMeso and the home-made DoE screen were set up with each ligand-complex sample. Again, no crystal-hits were obtained for any of the novel ligands. However, for the ZM241385 exchanged sample, several conditions with microcrystalline hits were observed with all of the SCRIPPS optimisation screens (Figure 5.14 D). Similar to when replacing LUF5834 with ZM241385, these crystals remained significantly smaller than the ones obtained after directly co-purifying the receptor with ZM241385. Therefore, diffraction was not assessed. Because of their small size, it is highly likely that for proper data acquisition a brighter X-ray source such as XFEL would have been required, similar to what is suggested in the ligand-exchange approach by Ishchenko *et al.* [301].

5.6.4 ZM241385–FOC-6 complex

The ligand-exchange experiments suggested that none of the novel ligands stabilised the receptor well enough to facilitate crystal formation. Therefore, co-purification of the A_{2A} -bRIL construct in complex with any of these novel ligands was not attempted. In contrast, ZM241385's stabilising effects are significant enough that crystallisation is even possible under less ideal conditions, *e.g.* with some dimer present in the protein sample, after freezing, or after ligand exchange. The data from the competition assay shows that FOC-6 is not able to completely

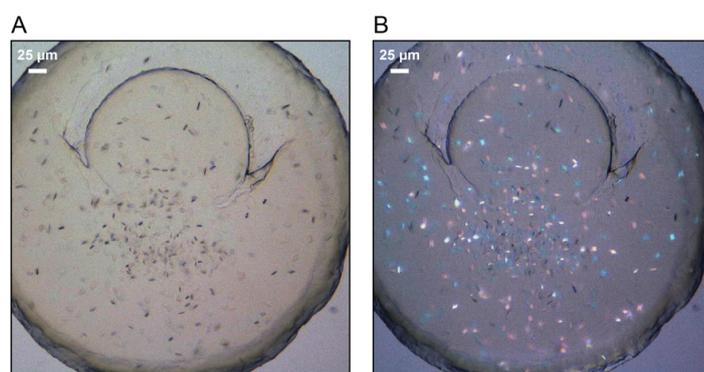


Figure 5.15: Significantly smaller crystals were obtained after adding FOC-6 to A_{2A} -bRIL–ZM241385 complex. 200 μ M FOC-6 were added to the same protein as used in the ligand exchange experiments in Figure 5.12. (A) Crystals imaged after 5 days using visible and (B) using polarized light. Crystals did not grow larger than about 10 μ m in length and did not diffract beyond 8 Å .

displace ZM241385 from the binding site. However, a sigmoidal binding curve could still be fitted, suggesting FOC-6 might bind to a secondary binding site as well (see chapter 3.2.2). Assuming this hypothesis is true, and FOC-6 binds to a secondary, allosteric site, it might either additionally stabilise the active or the inactive form. If it stabilises the inactive form, the ZM241385 stabilised receptor could be used for crystallisation to locate this secondary binding site of FOC-6. To investigate this idea, an aliquot of the same frozen protein-ZM241385 complex used for previous successful crystallisation was mixed with 200 μM of FOC-6 and incubated for 30 minutes to allow for equilibration. When setting up the LCP and crystallisation trials, an additional 100 μM of ligand was added to the lipid mixture as well as the crystallisation screens. Small crystals appeared after only 24 hours. However, even after two weeks they did not grow larger than about 10 μm in length and did not diffract beyond 8 \AA (Figure 5.15).

As the excess of ligand might have affected crystallisation, potentially displacing some of the ZM241385, the experiment was repeated with freshly purified receptor-ZM241385 complex. If FOC-6 does not have to compete with the bound ZM241385, smaller amounts of ligand should be enough to just occupy the secondary binding site. Thus, this time FOC-6 was added in a 1:1 molar ligand to receptor ratio, not adding any additional ligand in the subsequent crystallisation steps. Surprisingly, under these conditions, no crystals at all were obtained. These results indicate that FOC-6 does have some effect on the receptor that is preventing the formation of well-ordered crystals. Either FOC-6 outcompetes well enough to remove some ZM241385 from the orthosteric binding site, creating a mixed population that is less prone to crystallisation, or it binds to a second site, inducing a different conformation than ZM241385 binding, thereby destabilising the receptor. The first rationale seems to be less likely, as adding excess of higher affinity ligands did not displace ZM241385 and prevent the formation of diffracting crystals.

5.7 Alternatives to crystallisation: Cryo-EM

Protein crystallisation is a very sensitive system that can be easily disturbed by small changes, such as the introduction of a novel ligand. This study illustrates how, especially for the analysis of early lead candidates, the investigated ligands might not have the right properties to promote receptor crystallisation, hampering structure directed optimisation of a hit series. Cryo-EM has the potential to provide a valid alternative for structure determination in these cases. Nonetheless, despite the explosion of active state GPCR structures determined by cryo-EM in the recent years, the study of inactive state class A GPCRs has been limited by the small size of the receptor [303]. To circumvent this size limit and introduce additional fiducial markers for single-particle cryo-EM structure determination of inhibitor complexes synthetic universal antibodies can be used [144, 162].

The A_{2A}R mainly consists of the 7TM core, which is embedded in the membrane or detergent micelles and does not contain many extracellular features. Therefore, the construct used here for crystallography introduced an intracellular bRIL fusion in ICL3, which provides an additional surface to form crystal contacts. Now, with respect to cryo-EM studies, this construct was ideal to try and form a complex with a universal anti-bRIL antibody that would not only help to increase the size of the complex, but also introduce some additional asymmetry (Figure 5.16 A). In light of previous cryo-EM experience with membrane proteins of

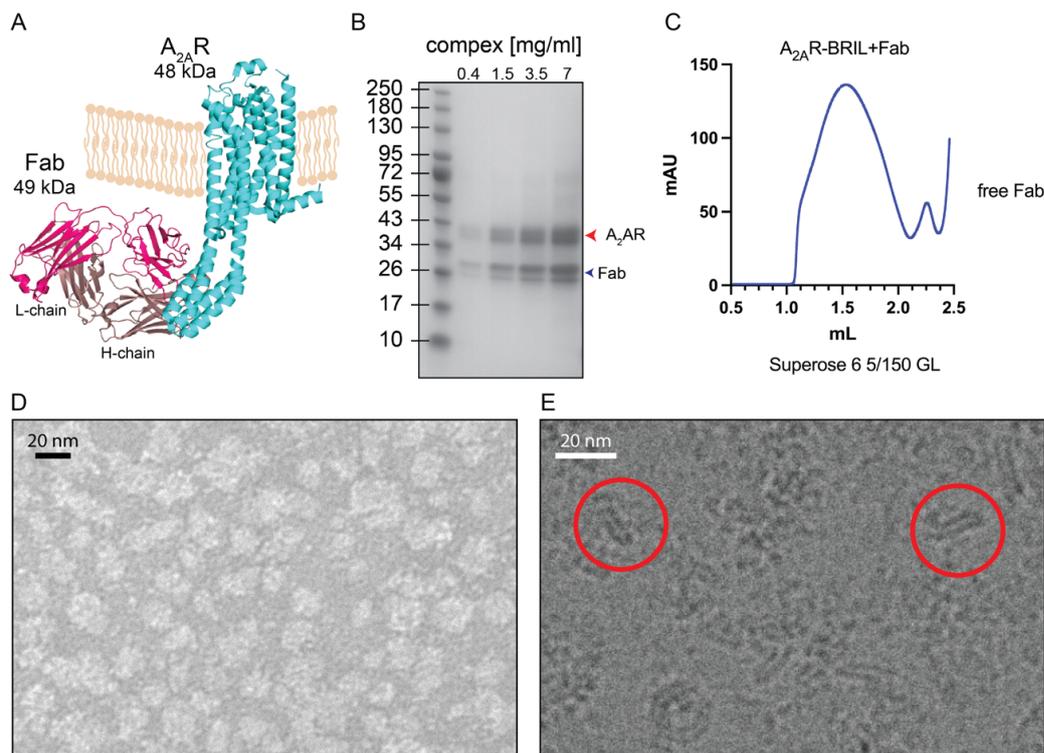


Figure 5.16: Cryo-EM studies on the A_{2A}R-bRIL-Fab complex. (A) Model of Fab binding to the bRIL-fusion in ICL3 of the receptor. A_{2A}-bRIL is about 48 kDa, while the Fab adds another 49 kDa. (B) SDS-Page of SEC purified complex, fractions were taken along the concentration process. The 0.4 mg/mL sample is straight from the SEC column. (C) SEC chromatogram of the complex purification. (D) Negative stain micrograph of 40 μ g/mL complex after storage at -80°C . (E) Micrograph of 7 mg/mL protein on Au-Flat protochips with gold film 130k \times magnification. Red circles highlight side view of LMNG micelle with complex.

other lab members, the detergent to solubilise the receptor was changed to LMNG instead of DDM. LMNG has a significantly lower off-rate, which means that buffers can be prepared with a detergent concentration below the critical micelle concentration without disrupting the formed micelles, leading to a reduction of excess detergent and empty micelles in the system. The A_{2A}-bRIL construct was purified as described for crystallography and then incubated with purified anti-bRIL BAK5 Fab [144] in a 1:1 molar ratio and 100 μM FOC-6. Afterwards, the complex was further purified by size exclusion (Figure 5.16 C). GPCR complexes are often applied onto cryo-EM grids at relatively high concentrations [304]. Therefore, the peak fractions were further concentrated to 7 mg/mL, taking several aliquots along the way (Figure 5.16 B). Cryo-grids were set up straight away and excess complex was flash frozen and stored at -80°C . Initial screening of these grids revealed that the complex fell apart when applied onto quantifoil grids. Only the Fab particles were visible in the holes, clustering around the edges, while the centre of the holes were essentially clear. The micelles with the receptor were drawn to the carbon support of the grid. Looking at the distribution of the Fab particles, the grids prepared with the higher concentrated protein sample looked generally better, having more particles towards the centre of the hole. Therefore, the effect of a different grid material was investigated. Before preparing new grids complex integrity after freezing was checked by nega-

tive stain EM. The sample still looked quite homogeneous and particle size suggested that the complex was still sound (Figure 5.16 D). Using the higher concentrated protein samples gold cryo-grids were prepared. Indeed, particle distribution could be improved compared to the first grids, and more particles were visible towards the middle of the holes. On the grids with the highest protein concentration some elongated disc shaped particles could be observed, which might represent LMNG micelles, possibly even with a Fab sticking out of some (Figure 5.16 E). This type of particle could not be identified on the grids prepared first, where the particles of the disintegrated complex were much smaller and more globular. Nevertheless, the grids were still very inhomogeneous, containing aggregates, as well as disintegrated complex. The fragile gold grids were also strongly damaged, only leaving a few squares for data collection. Consequently, the micrographs will be challenging to analyse, probably not allowing the generation of a high-resolution model. If the observed particles do indeed correspond to the intact receptor-Fab complex can only be judged after a first data analysis and 2D classification of the particles. Thus, it is too early to say if this strategy will be successful.

Nonetheless, these results suggest that cryo-EM might be an alternative approach for structure elucidation of inhibitor bound A_{2A}R. However, substantial sample optimisation as well as optimisation of grid preparation will be required to improve the micrographs enough to obtain structural data at a resolution that is high enough to identify an additional ligand binding site or an alternate ligand conformation in the orthosteric binding site.

5.8 Discussion

Protein crystallisation is a multidimensional problem dependant on fragile equilibria. Even in a well-established system, small changes can have a detrimental effect on the crystallisation process of a protein, making it hard to reproduce results under different conditions. Bucher *et al.* [305] demonstrated that the amino-acid sequence of the used purification tag can have a major effect on both the formation of crystals and their ability to diffract X-rays. A similar observation was made when I attempted to crystallise the slightly altered expression construct A_{2A}-StaR2-bRIL^{HRV-HIS} (see chapter 5.4). The C-terminal HRV-site only added an additional eight amino acids before the HIS-tag; however, this was enough to abolish the formation of crystals of the theophylline bound receptor and had a detrimental effect on the ability of ZM241385-complex crystals to diffract. Unfortunately, the C-terminus with the HIS-tag is not resolved in any published crystal structure, making it impossible to give structural reasoning for this observation. However, the study by Bucher *et al.* [305] suggests, that such effects are not related to steric hindrance in crystal packing and thus, the additional eight amino acids might rather have an effect on receptor stability. The C-terminus of the A_{2A}R is known to be flexible and impede protein crystallisation and is therefore truncated after residue 316 for structural studies. Thus, it might be possible that these eight amino acids re-introduce some of the innate flexibility of the C-terminus. Alternatively, they might be provoking protein oligomerisation and/or aggregation under specific conditions such as the used crystallisation conditions.

Similar effects can be observed when changing the ligands, as they might induce different receptor conformations along the activation spectrum. Furthermore, different charges on the ligands may also strongly influence the crystallisation process [306, 307]. Therefore, the co-

crystal structures of A_{2A}R with several important ligand series have remained elusive so far (*e.g.* Regadenoson, Preladenant, and Istradefylline). When I co-purified A_{2A}R with FOC-2, only small microcrystals could be obtained. Thermal unfolding experiments revealed that the stabilising effect of FOC-2 on the receptor is rather modest and previous examples in literature, using ligands with even higher affinity, showed that for some ligands it is very hard to get a co-crystal structure following a co-purification protocol [221]. Nonetheless, since at least some small crystals were obtained, there are still several aspects that can be optimised, *e.g.* trying to decrease the heterogeneity of the protein sample. In addition to protein quality, protein quantity needs to be optimised. Determining the protein concentration by different techniques suggested that the actual protein concentration might have been lower than assumed. In LCP, A_{2A}R usually crystallises at about 35–60 mg/mL [308, 150, 188], therefore, a higher protein concentration might favour protein crystallisation. It must be noted at this point that all efforts to further improve protein quality were unsuccessful and the monodispersed SEC peaks obtained at the beginning of the project could not be reproduced. This cannot be explained, except by differences in fundamental supplies such as chemicals or even water. Following the initial Covid19 lockdown, a significant decrease in water quality was reported in the faculty, which also affected the MilliQ[®] systems. These impurities had clear negative impacts on membrane protein purifications in our laboratory. The missing of some specific batches of sodium chloride on protein purification outcomes were also observed.

Finally, crystallisation conditions need to be optimised. In the current screen, a lot of salt crystallisation could be observed, possibly derived from the potassium/sodium tartrate present in the crystallisation condition. A slightly reduced salt concentration might avoid the formation of salt crystals and benefit protein crystallisation. However, screening the effect of different crystallisation conditions will be very time consuming, as these initial crystal hits could not be reproduced with frozen protein and thus, each round of optimisation would require a fresh protein preparation. Furthermore, the limited amount of compound constrains the number of different screens that can be set up with the crucial ligand additive.

Because of the limited quantity of ligands available, co-purification could only be attempted for FOC-2. For the other ligands, alternative strategies had to be explored, which required less compound. It has been previously reported that soaking of A_{2A}R-theophylline crystals can be a successful alternative path to get receptor-ligand co-crystal structures, if other crystallisation strategies did not result in diffracting crystals [221]. Despite being a widely used technique in drug discovery of soluble proteins, crystal soaking is not a straight-forward approach for crystals grown in LCP [50, 51, 52]. On the one hand, this method poses technical challenges, entailing manual handling and low throughput. Furthermore, to facilitate this process, a comprehensive knowledge of how different crystallisation conditions can affect the LCP is crucial. Sponge-phase like conditions make crystal handling steps even more difficult, as they are more liquid and should therefore, if possible, be avoided. On the other hand, high quality crystals of the receptor-theophylline complex are required [50]. Theophylline is a transient ligand with a relatively low affinity. This characteristic of theophylline facilitates its replacement by another ligand during the crystal soaking process, but it also means that perfect protein samples and

crystallisation conditions are required, as the ligand is not able to assist much in stabilising the receptor. Optimisation of all of these factors was difficult to achieve in the timeframe of this project; therefore, a third strategy was explored.

To use less ligand than required for a full co-purification, but circumvent the difficulties of soaking the crystals, I attempted to exchange the ligand just before setting up the LCP. No crystals were obtained when trying to exchange theophylline with DEL-2, even though the affinity of DEL-2 was about 40 times higher than the one of theophylline (see chapter 3.2.2). One reason for this could be that ligand exchange was not complete, leading to a mixture of different receptor conformations. This would imply that DEL-2 binding to the receptor induces a different conformation than theophylline binding and is in line with the observations made in the docking experiments described in chapter 4. Docking predicted the ligand to extend into a secondary pocket, which might introduce a different receptor conformation. It is interesting to note that the structure of the Cmpd-1-A_{2A}R complex, for which this secondary pocket was described first, could not be solved in LCP [193]. The authors observed that the ligand-induced structural changes around Tyr⁹¹⁻³⁵ propagated to the top of helix I. The extracellular part of this helix is involved in the formation of crystal contacts and the slight changes in conformation prevented the formation of type I crystals. However, LCP crystallisation is known to mainly produce type I crystal packing [309]. The complex therefore had to be crystallised by vapour diffusion. Consequently, it is likely that structure determination by crystal soaking would also not work for DEL-2, instead, vapour diffusion crystallisation strategies should be explored.

The affinity of ZM241385 was about 50 times higher than the affinity of the novel ligands selected for crystallographic experiments. Therefore, in the experiments where I attempted to displace ZM241385 with these compounds, ZM241385 was not outcompeted completely from the binding site, but probably a mixed population of A_{2A}R bound to different ligands was present. For DEL-1, a similar ligand pose was predicted by docking as for DEL-2, where the ligand formed contacts with all residues defining the secondary binding pocket. Thus, DEL-1 binding in this experiment might have induced a slightly different receptor conformation, which may have hampered the crystallisation process, leading to smaller crystals. It is likely that the majority of crystals formed were in fact ZM241385-bound A_{2A}R. However, the presence of some DEL-1 in the crystals cannot be excluded and could potentially explain the extra density observed around the ligand core in the solved A_{2A}R structure.

In contrast, much larger crystals were obtained when attempting to exchange ZM241385 with FOC-2. There are two factors that could contribute to this observation. Firstly, less FOC-2-receptor complex might have been present at equilibrium, which would also explain the unambiguous density for ZM241385 in the solved structure. However, as the affinity of FOC-2 was similar to DEL-1, this would further imply that the ligand has unfavourable kinetics, *i.e.* very fast dissociation rates. Secondly, FOC-2 binding might have induced a highly similar receptor conformation as ZM241385, as observed for the triazine compounds 4e and 4g [217]. This is also suggested by the very high score obtained when docking FOC-2 into a ZM241385-derived A_{2A}R structure. Thus, small amounts of FOC-2-bound receptor present in the system would not distort crystallisation.

When I attempted to exchange the more transient and lower affinity ligand, theophylline for FOC-2, small crystals were obtained. They displayed a different morphology than the previously obtained theophylline-complex crystals, indicating that the ligand was exchanged. If the FOC-2-induced conformation was similar to the ZM241385 conformation, it would also be more similar to the theophylline-bound state, as ZM241385- and theophylline-bound structures are practically identical [221]. Therefore, a mixed population would not fully prevent crystal formation. However, if a relatively large theophylline-bound A_{2A}R fraction was present in the protein sample used for crystallisation, this might explain why the crystals seemed to be rather inhomogeneous and did not diffract.

In contrast, ZM241385 was clearly able to displace more theophylline from the receptor binding site, as typical ZM241385-complex crystals were formed. Furthermore, these crystals did also diffract. The higher quality of these crystals can be explained by the 50 times higher affinity of ZM241385 for A_{2A}R which drives the receptor into a specific conformation [188]. Ligand exchange with ZM241385 was even possible using a less stabilised receptor construct, further validating the rationale of this approach. Nonetheless, protein diffraction at a synchrotron source was still too poor to solve the structure. This could be due to residual theophylline, introducing defects in the crystal lattice or an inherent problem with the protein sample. In both cases, further optimising the protein and crystallisation conditions will certainly help to push the equilibrium more towards the ZM241385-bound state and lead to better crystals. If the crystals remain small, it may be possible to use a brighter X-ray source as described by Ishchenko *et al.* [301]. Alternatively, these tiny crystals could also be analysed by cryo-electron diffraction [153].

At the end of the day, the receptor could not be crystallised with any of the novel compounds. All the ligands that were used in these crystallisation trials had a much higher affinity and an increased thermostabilising effect on A_{2A}R compared to theophylline. Thus, further knowledge, about the binding kinetics of the ligand would be required to better predict the crystallisability of these complexes. This knowledge would be beneficial especially when attempting ligand displacement strategies. It is likely that the dissociation rate (k_{off}) of the assessed ligands is very high and therefore, the receptor adopts multiple conformations that impair formation of well-ordered crystals. In ligand exchange experiments, this would lead to a mixture of receptor conformations, stabilised by the two different ligands. Furthermore, it is also possible that the ligands induce structural changes necessitating a different form of crystal packing [193]. Thus, exploring different crystallisation techniques, such as vapour diffusion, might be required to crystallise these receptor-ligand complexes.

Chapter 6

Overall conclusions and future directions

Using two different HT-screening methods, Novartis identified a set of novel A_{2A}R ligands that included compounds with known, as well as novel, chemical scaffolds. The main goal of this study was to characterise the interactions of these novel compounds with A_{2A}R, both pharmacologically and structurally, in order to gain insights into structure-activity related mechanisms. The results presented here show that the HT-screens were highly successful, yielding high affinity compounds with potentially novel binding modes. Furthermore, they could form the basis for rational chemical optimisation of selected hits, which might lead to the discovery of new allosteric ligands.

6.1 Key findings of the project

6.1.1 Identification of novel A_{2A}R binders

One of the screens Novartis used for the identification of novel A_{2A}R ligands was a large-scale virtual screen using a focused library. Out of the selection of compounds made available for further analysis, FOC-1, FOC-2, and FOC-3 had sub-micromolar affinity, while the others displayed affinities in the low micromolar range. Furthermore, all of these compounds inhibited cAMP production and thus, receptor activation.

In virtual screens, large compound libraries are docked into a rigid receptor structure. This biases the results towards ligands that are likely to bind to or induce a receptor conformation similar to the ligand in the original structure. As such, several of the hits contained a bicyclic core similar to the purine core in adenosine. However, hits with monocyclic cores were discovered too. Interestingly, the triazine and triazole scaffolds found in FOC-2 and FOC-3 have been identified in previous docking screens as well [209, 210, 211], indicating that these are very potent scaffolds. Subsequent *in silico* studies were performed to further analyse the predicted binding poses of these ligands. Most of the ligands were predicted to form hydrogen bonds and aromatic interactions with residues that are known to be involved in ligand recognition. Comparison of

FOC-3 and FOC-2 with bioisosteres that have been further optimised and crystallised showed that the predicted poses were similar to the experimental structures, validating the docking approach.

In contrast, using the DEL-screen, more novel scaffolds could be identified. Similarly to the virtual screen, this primary hit selection resulted in the identification of two ligands, DEL-1 and DEL-2, that had a high affinity in the low nanomolar range. Interestingly, DEL-1 also displayed the highest antagonist potency of all ligands tested. In a DEL screen, the receptor conformation is more flexible and therefore ligands with more diverse binding modes were identified. Docking studies of these hits suggested that DEL-1 and DEL-2 both extend into a secondary binding pocket that is not occupied by the natural ligand.

Interestingly, this selection of ligands contained two pairs of compounds, DEL-2/DEL-5 and DEL-3/DEL-6, that only differed in the composition of one single moiety. Comparison of the pharmacological data with the predicted docking poses facilitated predictions about the structural role of these chemical groups. The large morpholinyl-ethyl-acetamide moiety on DEL-2 was important for positioning the ligand core in the binding site and further interacted with a secondary binding pocket, greatly increasing the affinity of the ligand. In DEL-6 the highly polar hydroxypropanoic acid moiety was identified to hinder the ligand from properly binding into the orthosteric pocket. Targeted chemical substitutions in these regions might facilitate further improvements the binding properties of these compounds.

6.1.2 Different strategies for the determination of early stage ligand co-structures

One of the main issues when evaluating the binding properties of early screening hits is that they are usually only available in low quantities, especially if they are not available commercially and need to be synthesised internally. This is particularly limiting for structural studies of GPCR-ligand complexes, where the presence of a ligand is usually required throughout the purification process to stabilise the receptor [225]. Therefore, I explored three different approaches to generate co-crystal structures with the best candidates out of this novel ligand selection. The first strategy was to co-purify the receptor with the novel ligand. This was the most straightforward approach, but used the largest amounts of ligand and therefore could only be attempted with FOC-2. The second approach was to soak ligands into crystals obtained in complex with theophylline. While this approach was successfully used for several A_{2A}R ligands in Sosei Heptares Therapeutics' drug discovery pipeline [50, 221, 208, 308], it requires a lot of manual practice and optimal sample quality. The final strategy was to exchange the ligand once the protein was purified and concentrated, but before setting up LCP. In this way minimal quantities of compound were used, while reducing the crystal handling. The validity of this approach could be shown by exchanging the low affinity ligand theophylline with the reference ligand ZM241385. Unfortunately, no diffracting crystals were obtained with any of the three approaches for any of the novel ligands. One reason for this might have been the low thermostabilising effect of the ligands, making the complex more difficult to crystallise [307]. Another reason could be that the ligands induced a different conformational state, which is less amendable for crystal formation. While the vast majority of A_{2A}R crystal structures were

solved in LCP, this might not always be the optimal methodology. LCP crystallisation favours type I crystal packing [309]. Thus, if the ligand induces an altered receptor conformation that requires different crystal packing, classical *in surfo* methods might be preferred [193].

6.1.3 Identification of novel potential subtype-selective and allosteric ligands

Surprisingly, out of the 12 selected ligands, four might bind to an allosteric binding site. DEL-1 and DEL-2 were predicted to extend into a previously identified pocket extending towards the extracellular aspect of TM I, TM II, and TM VII that is not explored by the natural ligand [193]. Both ligands inhibited receptor activation by NECA. While displaying similar affinity than other ligands, DEL-1 had a clearly higher potency. Thus, interaction with this secondary region might potentially increase ligand potency. Furthermore, docking of the ligands into an A₁R structure and an A_{2B}R homology model predicted that these ligands adopt a different binding pose in these subtypes. Several residues in this pocket are subtype specific *e.g.* Thr 270^{7,34} in A₁R (Met 270^{7,34} in A_{2A}R) has been previously associated with subtype selectivity [202, 275, 276]. Consequently, interaction with this secondary site might increase A_{2A}R subtype specificity. Therefore DEL-1 and DEL-2 might be good starting points for subtype selective ligands.

Despite the fact that no protein diffraction was obtained with any of the novel ligands, the results from the crystallisation trials further corroborate the hypothesis that DEL-1 and DEL-2 induce a different conformational state in A_{2A}R. The crystal structure with the triazole Cmpd-1 showed that ligand binding induced a slightly different conformation that was not compatible with type I crystal packing [193]. If DEL-2 would induce a similar conformation this would explain why no crystals were obtained in LCP. Furthermore, when I tried to displace ZM241385 with DEL-1, this led to smaller crystals compared to FOC-2 or when crystallised with ZM241385 alone. This indicates that the subpopulation of DEL-1-bound receptor hampers the crystallisation process and suggests the presence of multiple conformational states. In contrast, FOC-2 was predicted to bind in a similar pose as ZM241285 and ligand binding would therefore likely induce a similar receptor conformation, which did not affect crystal formation.

The third ligand that could potentially bind to a secondary site is FOC-6. FOC-6 displayed a very unusual behaviour in the binding assays: despite binding to the receptor in a competitive manner, the ligand was not able to fully displace the radioligand, even at millimolar concentrations. Furthermore, it neither inhibited nor stimulated cAMP production. This indicates that FOC-6 might bind to a secondary site. However, the functional relevance of this interaction still needs to be established.

To identify possible alternative binding sites on A_{2A}R, blind docking studies were performed, scanning the entire extracellular part of the receptor. The docking studies suggested that FOC-6 binds to a previously identified lipid-pocket at the extracellular end of helix I and VII [150]. While some studies suggest lipid (especially cholesterol) interactions to be highly specific, a more recent study showed an absence of high-affinity binding and a more transient nature of interactions [310, 294, 297]. This suggests that FOC-6 might indeed be able to compete with a lipid for binding, even though only micromolar affinity for the receptor could be determined.

However, other ligands were predicted to bind to this pocket as well, so binding to this pocket would not explain the unusual behaviour of FOC-6. Nonetheless, lipid interaction sites might be an interesting target for the allosteric modulation of A_{2A}R and other GPCRs.

Overall, DEL1, DEL-2 and FOC-6 are very exciting hits that display interesting characteristics that should be investigated further. On the one hand these ligands could give rise to novel ligand series with improved subtype selectivity or allosteric function. On the other hand, identifying the binding sites for these ligands will provide further insights into the conformational landscape of A_{2A}R and potentially reveal novel target sites for allosteric modulators.

6.2 Future directions

The *in silico* studies performed, predicted that several of the ligands might adopt distinct binding poses and introduce novel receptor conformations. However, this hypothesis could not be fully validated, and additional studies are required to properly characterise these receptor-ligand interactions.

6.2.1 Strategies to further advance structural work

The most valuable result would still be to generate a high-resolution structure of the ligand bound to A_{2A}R, which would allow the most detailed characterisation of the ligand binding site. As discussed previously, one of the first things to try would be *in surfo* crystallisation, to allow the formation of different crystal packing. However, LCP crystallisation might also be successful. The poor thermostabilising effects of the ligands suggest that the complex might not have been stable enough for crystallisation and therefore require some additional stabilising measures. One strategy for this, without having to further modify the receptor construct, could be the addition of stabilising lipids. When the first A_{2A}R structure was solved, addition of CHS to the detergent solubilised A_{2A}R was shown to be crucial for successful crystallisation [188]. Also cholesterol is well known to affect the function and stability of GPCRs and is therefore routinely mixed with monoolein for LCP crystallisation [310]. I therefore also added these lipids during my purifications and crystallisation trials. However, we recently discovered that 1-palmitoyl-2-oleoyl-*sn*-glycero-3phosphate (POPA), a phosphatic acid lipid, also stabilises A_{2A}R by +5.8 °C [3]. This is slightly more than half of the effect recorded with CHS. It would be interesting to see if these effects are cumulative and a lipid mixture could be used to further stabilise the receptor facilitating the crystallisation of less stable receptor-ligand complexes.

FOC-2 crystals could be obtained but were too small for standard structure determination at a synchrotron X-ray source. Diffraction by XFELs can improve data quality of small crystals given the advantages of the intense and very short XFEL pulses. This was exploited for structure determination of the β -arrestin-rhodopsin complex [142]. Subsequent studies also showed the application of SFX for the determination of multiple co-crystal structures in SBDD of A_{2A}R ligands [301]. Unfortunately, XFEL access is still highly limited. Similar serial approaches for room-temperature structure determination of GPCRs have been developed; however, the obtained data still remains at slightly lower resolution [139]. An alternative strategy to analyse

these small crystals might be to explore the emerging potential of micro ED. A recent study reported structure determination of A_{2A}R at 2.8 Å resolution from LCP grown crystals using this technique [153].

Over the course of this project, cryo-EM has become a powerful technique for high-resolution structure determination, and since 2019 the number of membrane protein structures determined by cryo-EM has overtaken the number of X-ray structures [154]. While most of the crystal structures represent an inactive receptor state, the vast majority of cryo-EM structures are solved in the active state; however, structure determination of inactive state GPCRs, especially class A GPCRs with small extracellular domains, still represents a major challenge. Recently, universal anti-bRIL synthetic antibodies have been developed to overcome this hurdle [144, 162]. The Fab helps to increase the size of the complex and serves as a fiducial marker assisting in determining particle orientation. In preliminary studies I could show that a stable complex of the A_{2A}R-bRIL and the anti-bRIL Fab could be formed and applied onto cryo-EM grids. Elongated disc shaped particles could be observed, which are typical for LMNG micelles and the sideview of the Fab sticking out of the micelle could be identified. However, these grids were still very heterogeneous and further optimisation of sample and grid preparation will be required to improve the micrographs enough to obtain structural data at a resolution that is high enough for the identification of ligand densities. The main challenge will be to ensure that the complex does not dissociate when applied onto the grid. This might require testing several different detergents, several rounds of size exclusion purification and/ or technical changes to the protocol, such as complex formation on the grid. Nonetheless, these results suggest that cryo-EM might become an alternative approach for structure elucidation of inhibitor bound A_{2A}R.

6.2.2 Biophysical characterisation of receptor-ligand interactions

Besides structural approaches the ligand-receptor interactions and associated conformational changes can further be analysed by using biophysical and advanced *in silico* techniques. Assuming DEL-1 and DEL-2 binding induces a similar receptor conformation as Cmpd-1, docking scores and pose predictions should be improved when docking these ligands into the 5UIG structure, where the secondary binding pocket is more established and relevant residues should be in a more favourable rotamer conformation. Furthermore, the effect of structurally relevant waters should be examined and MD simulations can help to account for ligand induced changes in the receptor [284, 286]. These computational simulations can further be improved by providing more experimental knowledge about the ligand-receptor interactions [216].

Another approach that has been shown increasingly valuable to study the conformational landscape of GPCRs and other membrane proteins and the changes induced upon ligand binding is double electron-electron resonance (DEER), also known as pulsed electron-electron double resonance spectroscopy (PELDOR) [311, 312]. To this end, a combination of site-directed spin labelling and electron paramagnetic resonance (EPR) spectroscopy is used to resolve distances between a pair of spin labels in the range of 15–80 Å, allowing to resolve conformational changes at sub-angstrom resolutions [313, 314, 109]. Katie Hardman is in the process of creating a cysteine free-A_{2A}R construct, and specific cysteine-mutants, for the site-directed spin labelling.

With a suitable protein sample in hand, she aims to investigate the effect of different ligands, as well as varying lipid environments and even membrane tension on the conformational space of A_{2A}R by DEER spectroscopy.

As mentioned in previous chapters, an in-depth knowledge of binding kinetics at GPCRs is required to successfully target this class of proteins [315]. As such, it would be helpful to determine the rate constants k_{off} and k_{on} for the different ligands. If rate constants are determined for the WT-receptor, as well as for a set of active site mutations, this information can be used in combination with molecular modelling and docking to map the small molecule binding site, even in the absence of a co-crystal structure [216]. Ultimately, understanding the binding kinetics of the different ligands, as well as structure kinetic relationships, will help to increase ligand residence time and help to further optimise the ligands [316, 317]. In order to determine these rate constants for the novel ligands, I started setting up SPR experiments during my secondment at Novartis. Unfortunately, the assay could not be established during the timeframe of the secondment and further experiments will be required to obtain the desired information.

6.2.3 Functional aspects of ligand binding

The docking studies allowed some speculation about subtype selectivity of the ligands. These hypotheses should be verified experimentally. One way to do this would be to apply the LANCE[®] assay described in chapter 3.3 on the different subtypes. HEK cells have been shown to have relatively high endogenous A_{2B}R expression [247] which would make it more difficult to assign responses to a specific AR subtype. Therefore, the LANCE[®] assay was established using CHO cells rather than HEK cells. Constructs for A₁R and A_{2B}R were established and could be expressed in CHO cells. However, the assay still needs to be optimised for the different subtypes.

The LANCE[®] assay could further be used to explore the possible allosteric function of the ligands. FOC-6 showed no direct inhibitory or activating effect; however, it might have an effect on agonist binding. This could be detected if a curve shift can be observed when recording an agonist curve in presence of FOC-6. Alternatively, other assays could be used to investigate the effect on alternative signalling pathways. Bioluminescence resonance energy transfer (BRET) based biosensors have been extensively used to study protein-protein interactions of GPCRs in living cells, such as the interactions between receptors, G proteins, β -arrestins, and their many binding partners. Furthermore, these biosensors can also be used to measure the formation and accumulation of second messengers following receptor activation and signal transduction [318]. To further characterise the functional effects of the here studied novel A_{2A}R ligands, BRET-based studies could *e.g.* be used to directly monitor the interaction between the A_{2A}R and the G protein [319, 320]. The potentially allosteric ligands might affect this interaction differently than previously studied agonists or antagonists and might therefore give valuable information on signalling bias. Similarly, β -arrestin requirement assays might help to further complete the picture [321].

6.3 Significance of the work

A_{2A}R is one of the main mediators of adenosine signalling and as such represents an interesting drug target for neurodegenerative disease, as well as in immunotherapy for cancer treatment. Despite significant efforts in developing specific agonist and antagonistic drugs, most projects eventually come to a halt, either because of a lack of efficacy or adverse effects resulting from A_{2A}R activation. One strategy to circumvent such effects might be the use of allosteric or potentially even biased ligands [280, 322, 323].

SBDD has become an integral part of most drug discovery projects. Since the release of the first (non-rhodopsin) GPCR structure, SBDD has also become more established in GPCR drug discovery. However, the quick generation of many different receptor-ligand co-crystal structures still remains a major hurdle to fully take advantage of SBDD.

This project explored a variety of different strategies to characterise a series of novel A_{2A}R ligands, pharmacologically and structurally, which led to the identification of several potentially allosteric ligands. These hits would probably not have been identified if only one single parameter, such as affinity or docking score would have been considered, which shows the importance of employing complementary techniques and in-depth analysis of screening hits.

Despite not leading to a co-crystal structure, the crystallographic efforts further corroborated that these ligands induce a conformational change of the receptor, and the presented results form a basis for rational chemical optimisation of the selected hits. Upon optimisation these hits might become interesting lead candidates for the discovery of novel allosteric A_{2A}R drugs.

Appendix A

Exploring novel ways in which crystallography might assist drug discovery

Production of microcrystals for serial and time-resolved
crystallography

Preface

During a three-month secondment in the laboratory of Prof. Arwen Pearson, I had the opportunity to learn and develop new methods for the generation of protein microcrystals for serial and time-resolved crystallographic studies. The results from this work contributed to a publication on homogeneous batch micro-crystallisation of proteins from ammonium sulphate, one of the most common crystallisation agents used [1]. The following chapter provides background on the topic of this publication and discusses the results I obtained.

Abstract

Time-resolved crystallography is an exciting new avenue to study dynamic processes in proteins. However, the generation of large volumes of homogeneous microcrystal suspensions required for these types of studies is still a major bottleneck. While batch crystallisation is typically the method of choice for this purpose, it usually requires information about the phase diagram of the protein, which can only be determined experimentally in extensive crystallisation experiments that are very sample consuming. Therefore, the aim of this project was to establish some general guidelines for homogeneous batch micro-crystallisation from ammonium sulphate conditions. To this end, a robust micro-crystallisation protocol for the soluble model protein aspartate α -decarboxylase (ADC) was established. Using this protocol, crystal size and size distribution could be fine-tuned according to the experimental requirements given by the envisioned study, rather than having to adjust the experiment according to the crystals obtained. By employing data from two additional model enzymes, it was further explored how the ammonium sulphate concentration can be used to navigate the crystallisation diagram to quickly establish conditions for micro-crystallisation.

Furthermore, the ADC micro-crystallisation protocol developed here was used to generate crystals for the characterisation of a novel 3D printed and X-ray compatible microfluidic chip that allows low-consumption serial synchrotron data collection in flow [2].

A.1 Introduction

To explain how proteins perform their biological function, knowledge about the protein structure, as well as the dynamic motions that underlie the chemical mechanism, is required. One of the few techniques that can provide insights into protein structure and dynamics, at near atomic resolutions, is time-resolved crystallography. The idea of directly capturing structural dynamics by crystallography is almost as old as macromolecular crystallography itself, but has received a tremendous boost since the establishment of free electron lasers (XFELs) and 3rd and 4th generation synchrotrons, all of which facilitate serial data collection [149, 324, 325]. The traditional approach to time-resolved crystallography is to trap and solve the structure in different intermediate states. This can be achieved by soaking crystals of an *apo*-protein with substrate and trapping the intermediates by snap-cooling the crystals after a defined time delay. However, this limits the achievable time resolution to the speed at which the crystal can be cooled. In contrast, serial approaches are based on collecting a single diffraction pattern from many crystals in random orientations and merging these still images to give a complete dataset. This enables time-resolved data collection based on the “pump-probe” method, in which reaction initiation (pump) is followed by an X-ray probe. Besides enabling greater time resolution for time-resolved studies, serial data collection allows room-temperature (RT) data collection with limited radiation damage [326, 139, 327] and the structure solution of more challenging proteins, such as integral membrane proteins, in cases where only small crystals can be obtained [150].

A.1.1 Strategies for serial time-resolved crystallography

To study protein function with time-resolved crystallography there are two main requirements: first, the studied protein needs to be active in the crystalline state, and second, a controlled initiation of the reaction needs to be possible. Furthermore, for effective reaction initiation, the process needs to be triggered in a significant fraction of the sample, and the triggering event needs to be faster than the process of interest. Therefore, the simplest biological processes to investigate are those that can be triggered by short laser pulses, for example, in inherently light sensitive proteins such as myoglobin [328, 329], photoactivatable yellow protein [330, 331] and photosystem I & II [149, 332]. For non-natural photoactivatable systems, photolabile protecting groups can be chemically added to the natural substrate [333], or photocaged unnatural amino acids can be introduced into the protein during translation [334, 335]. A second protein activation approach employs the delivery of substrates or ligands by diffusion [336, 337, 338]. This can be achieved in flow using T-shaped mixing devices [2] or on fixed targets *e.g.* employing liquid application strategies [336]. Even though these “rapid-mixing” strategies are applicable to a wider range of targets, they are limited in time-resolution, due to the diffusion rate of the ligand, and are only possible if ligand diffusion is fast compared to the timescale of the reaction of interest.

Serial data collection can be performed either at synchrotron or XFEL X-ray sources. Data collection at an XFEL is based on the “diffraction before destruction” principle and therefore inherently requires a serial approach. The XFEL beam delivers high intensity X-ray laser pulses of shorter than fifty femtosecond duration, which allows recording of protein diffraction

before significant radiation damage occurs (serial femtosecond crystallography (SFX)) [149]. This means that only a single “snapshot” per crystal is collected, requiring many, differently oriented crystals for a complete data set. Because of the high beam intensity, this strategy allows diffraction data collection from crystals of micrometre to nanometre size and femtosecond time resolution for time-resolved studies.

Serial data collection approaches have also become increasingly popular at synchrotrons. They follow the same idea of distributing the radiation dose over thousands of crystals to maximise the signal. However, because the brightness of synchrotron beams is considerably weaker than that of XFELs, longer exposure times are required. Nonetheless, today's 3rd and 4th generation synchrotrons provide sufficient photon flux to perform time-resolved experiments at timescales as short as 100 ps (using polychromatic radiation) [339].

A.1.2 Crystal requirements for time-resolved studies

In contrast to conventional single-crystal rotation crystallography, where a complete data set can be collected from a single large crystal, serial data collection requires many small crystals. In fact, for most time-resolved applications, crystal size is minimised to ensure uniform activation of reactions (*e.g.* by laser or substrate diffusion) across the whole crystal. The desired crystal size is dependent on the experimental setup used, *i.e.* the delivery system, the activation strategy (for time-resolved experiments), as well as the used X-ray source.

The most popular crystal delivery system at XFELs are liquid jets [51]. For these and other microfluidic experiments, such as microfluidic mixing devices (*e.g.* 3DMiX), crystal size is limited by the size of the nozzle or channel of the system. Clogging and jet instability are significant issues when working with liquid jets, causing substantial delays and reduced hit rates [340]. These effects are further increased by a polydisperse size distribution of the crystals [341]. Additionally, a homogeneous size distribution is important to tune the exposure time required to obtain good diffraction, without the appearance of overloads, and to reduce the challenge for good scaling and merging of the data, as large variation in crystal size will lead to discrepancies in the measured intensities [342]. Similarly, for time-resolved experiments, minimising any variations in crystal size is important to allow for homogeneous reaction initiation throughout the sample.

Time-resolved experiments generally demand small crystals. Light activated processes require consideration of the laser penetration depth relative to the size of the crystal under investigation [339]. Larger crystals lead to increased absorption of the pump laser pulse, resulting in inhomogeneous activation throughout the crystal and thus, a lower fraction of the sample ultimately activated. This results in reduced signal from the desired activated state relative to the inactive state, making the interpretation of structural changes more difficult. Likewise, the time required to activate a reaction by a ligand diffusing into the crystal depends on the thickness of the crystal and the diffusion coefficient of the ligand [343]. Thus, varying crystal sizes will lead to uncertain diffusion times, burring the final time-resolution.

Because of their brightness, XFELs allow the use of very small crystals ($<10\ \mu\text{m}$). Conventionally, fully hydrated microcrystals are delivered in a jet of liquid or viscous media, which continuously runs in vacuum across the XFEL beam. The crystals in the stream are randomly oriented and produce a diffraction pattern whenever the XFEL pulse coincides with a crystal

being in the focal region [324, 149]. These strategies of data collection require a continuous supply of crystals throughout the entire process and are thus associated with a very high sample consumption. To address this issue, a series of different delivery systems have been developed, including drop on demand dispensing and fixed targets [344, 345]. Nonetheless, for serial crystallography experiments sample consumption and availability is a major concern, especially when challenging samples are investigated, such as membrane proteins or protein complexes [51]. When sample consumption is high and sample availability limited, time-resolved experiments become impossible as multiple complete data sets need to be collected, each covering a different timepoint along the reaction coordinate.

Due to the limited availability of XFEL beamtime, a variety of sample delivery strategies have been adapted and developed for serial experiments at synchrotron sources. [346, 347, 348, 325, 336, 349]. Because of the reduced beam intensity available, many serial synchrotron (SSX) experiments can benefit from slightly larger crystals ($\sim 10\text{--}50\ \mu\text{m}$) [325].

A.1.3 Micro-crystallisation strategies

Despite the high demand of crystal samples for serial crystallography experiments, methods for their generation are still sparsely explored. Usually, the large volumes of microcrystalline sample required for serial experiments dictate the type of crystallisation used. About one third of all microcrystal samples are produced in batch, as it is usually easier to scale up to the large volumes required compared to vapour diffusion [350]. Furthermore, if the crystal density required is not reached, crystals can be concentrated by low speed centrifugation. In batch crystallisation, the protein and precipitant solution are usually mixed directly, aiming for a ratio that places the protein within the nucleation zone of the crystallisation phase diagram after mixing [351]. However, finding this starting point usually first requires the determination of the crystallisation phase diagram. This can be achieved by determining the solubility limit of the protein either through stepwise dissolution of protein crystals or by small scale batch experiments with added crystal seeds [340]. Alternatively, the crystallisation space can be mapped by a series of experiments with varied protein and precipitant concentrations [350]. All of these strategies are very time-consuming and require several rounds of optimisation. Furthermore, the obtained microcrystal solutions often display a rather broad crystal size distribution. To a certain extent, filtering of the crystal slurries can help to obtain a more homogeneous crystal sample. However, this strategy is very inefficient with a continuous size distribution. Furthermore, filtering can damage the crystals, which can lead to complications in the downstream analysis of the diffraction data [352].

The second most popular technique is lipidic cubic phase (LCP) crystallisation [350]. LCP crystallisation first became popular for the crystallisation of membrane proteins, as it mimics their native lipidic environment and therefore benefits protein stability, producing highly ordered crystals [353, 354, 355]. Even though these crystals are often limited in size, they are frequently obtained at high density. Because LCP is highly viscous, it turned out to be an ideal carrier media for slow flowing jets, not only allowing high-resolution structure data of challenging membrane protein samples to be collected, but also significantly reducing sample consumption compared to liquid jets [152]. However, the lower flow rate comes at the price of broader stream diameters, causing increased background scattering. As proteins are directly

crystallised in LCP they don't need to be transferred to a different carrier media, simplifying crystal handling procedures and reducing mechanical stress imposed on the crystals. However, because of the high density of LCP, time-resolved experiments initiated through rapid mixing are not possible and thus, only light activated reactions can be studied.

There are only a few crystallisation protocols which specifically focus on the generation of high quality micro- and nanocrystals [350, 356, 352, 341, 340]. To produce microcrystals exhibiting a monodisperse, uniform size distribution, Kupitz & Grotjohann *et al.* [340] suggest the use of free interface diffusion. This strategy is based on the formation of crystals at the interface of a highly concentrated protein solution with precipitant solution. The larger the interface, the more nuclei are formed. Therefore, the authors suggest an approach in which the precipitant is slowly dropped into the protein solution. As the protein solution has a lower density than the precipitant solution, the small precipitant droplets move through the protein solution to form two distinct layers, creating a large transient interface area at which protein nucleation occurs [340]. Another promising approach is the use of crystal seeds. Seeding is a well described method, regularly used for the growth of large single crystals. However, it can also be extremely helpful for the growth of nano- and microcrystals in batch, as crystal size usually depends on the number of seeds. Many seeds lead to smaller crystals, as there remains less protein available in solution for crystal growth. While producing microcrystals, by physically crushing large crystals, might still result in suitable diffraction for some proteins, the resulting crystal slurry is usually heterogeneous in size and crystal damage manifests as a high mosaic spread, leading to the streaking of spots and issues while indexing [352]. Better results can be obtained by spiking these crushed crystals into new crystallisation drops. Additional rounds of crushing and seeding further improve sample homogeneity [352]. Similarly, Ibrahim *et al.* [341] proposed different microseeding protocols, in which either the crystals of the first seeding experiment are used to produce a highly concentrated seed stock for the second round, or a second spike of the initial seeds is added to the supernatant after the removal of the first microcrystals.

More recently a supersaturation-controlled micro-crystallisation method was described [356]. Nucleation in protein crystallisation is known to be dependent on the level of supersaturation, therefore, in conventional crystallisation, slowing of the rate of supersaturation is used to grow large single crystals. By allowing increased evaporation at specific time delays, the reverse approach can be applied, increasing supersaturation and facilitating rapid and continuous nucleation. This results in increased microcrystal density and improved reproducibility [356].

Despite the success of these methods for some protein systems, the generation of microcrystals remains rather challenging, as most strategies require prior knowledge of the crystallisation phase diagram and/or do not suggest how to scale up to the amounts required for an SSX/SFX experiment, and therefore require additional optimisation steps.

A.1.4 Aim

The main goal of this project was to establish a robust batch micro-crystallisation protocol for *E. coli* L-aspartate α -decarboxylase (ADC) (60 kDa homotetramer), which could then be used to generate enough microcrystals for proof-of-principle experiments with a novel, X-ray-compatible microfluidic device for serial and time-resolved crystallography.

There are only a few crystallisation protocols which specifically focus on the generation of high quality micro- and nanocrystals. However, these protocols are often very time consuming and cannot be readily upscaled. Therefore, in order to establish the wide application of the micro-crystallisation protocols developed, two other proteins were investigated and the micro-crystallisation results compared between the three. This comparison aimed to identify novel, more straight forward approaches for transitioning from vapour diffusion to batch crystallisation, based on the precipitating properties of the frequently used crystallisation additive ammonium sulphate.

A.2 Methods

The full methods are described in Stohrer *et al.* [1]. Described here is only the micro-crystallisation protocol that was developed for ADC.

To screen different crystallisation conditions a large amount of cell mass was prepared, using 11 L of *E. coli* culture. This allowed protein to be prepared from the same batch of cells throughout the optimisation process. ADC was purified as described previously [357, 358].

All crystallisation solutions were prepared by dissolving the powders or liquids from the corresponding salts and buffer components (from Sigma Aldrich or Roth) and making up to the total volume with MilliQ[®] water. The pH was adjusted using the corresponding acid-base pair. Before use, all solutions were filtered using 0.1 µm syringe filters. The ammonium sulphate stock solution was prepared freshly before each crystallisation experiment, as the hygroscopic characteristics of the precipitant quickly led to a change in concentration of the stock solution.

Before crystallisation, ADC was buffer exchanged into 50 mM Tris pH 7.5, 100 mM NaCl, 0.1 mM DTT and concentrated to 25 mg/mL. Crystallisation trials were set by mixing the protein with precipitant solution (1.85–2.10 M (NH₄)₂SO₄, 65 mM citric acid pH 3.8, 71 mM Na₂HPO₄) in a 1:3 protein: precipitant ratio. Initial crystallisation was performed in 0.5 mL Eppendorf tubes. Firstly, 20 µL of protein were added, followed by 60 µL of mother liquor, then the tube was vortexed for ten seconds. The samples were then incubated at 18 °C for one to two days. To have control over ambient humidity, reactions were set up inside a glove bag containing Drierite[™] as a moisture absorbent. To accelerate the production of large amounts of crystals, reactions were scaled up using up to 200 µL of protein.

A.3 Micro-crystallisation of ADC

A.3.1 Using changes in ammonium sulphate concentration to control crystal size

In a simplified approach, crystal growth can be described with a two-dimensional phase diagram, as a function of two crystallisation parameters, usually protein and precipitant concentration. For protein crystallisation to occur, protein concentration needs to be above its solubility limit, in a supersaturated state. Depending on the level of supersaturation, this region of the diagram can be divided into three different zones: at very high supersaturation, precipitation occurs, leading to the formation of amorphous aggregates. At slightly lower supersaturation, protein

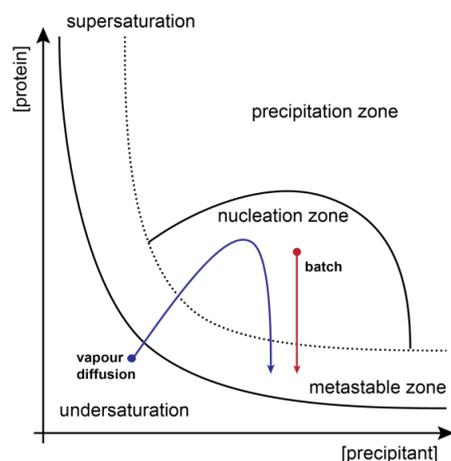


Figure A.1: Simplified protein crystallisation diagram. The different routes of reaching nucleation and metastable zones are shown for vapour diffusion in blue and batch crystallisation in red. The boundaries between undersaturation/metastable and nucleation/precipitation zone are shown as black lines and the boundary between metastable and nucleation zone as dashed line. Figure adapted from Stohrer *et al.* [1].

crystal nucleation is initiated, while in the metastable zone, crystal growth is facilitated [359]. Drop equilibration in a vapour diffusion experiment leads to a diagonal move across the phase diagram from a point of undersaturation into the nucleation zone, followed by a vertical drop in protein concentration and a shift into the metastable zone, where crystal growth occurs (Figure A.1). To enable micro-crystallisation in batch, the initial crystallisation conditions must be inside the nucleation zone of the crystallisation phase diagram. If conditions for vapour-diffusion crystallisation are known, it is possible to estimate the required conditions based on a few simplifying assumptions. Firstly, it must be assumed that full equilibration of a vapour-diffusion drop is reached before crystallisation, and secondly, that it then remains constant. In this case, the maximal precipitant concentration approximately equals the precipitant concentration of the reservoir. Thus, crystal nucleation should also occur at this precipitant concentration. Since batch experiments are usually set up by mixing the protein and precipitant in a 1:1–1:3 ratio, the concentration of the precipitant for batch crystallisation should be chosen to be 1.3–2 times higher than in the vapour-diffusion conditions. Similarly, the protein solution is also diluted by this mixing step, and consequently, the sample concentration needs to be higher than in vapour-diffusion.

Using this assumption as a starting point, a robust protocol could be generated for the micro-crystallisation of ADC. Precipitant concentrations in the range of 1.7–2.2 M ammonium sulphate usually gave microcrystals, concentrations above this range led to precipitation and concentrations below resulted in only few, very large crystals. The protein concentration was kept constant at 6.3 mg/mL. Furthermore, exchanging the protein from an imidazole containing elution buffer into a tris-buffer and using a citrate-phosphate buffer in the mother liquor helped to grow crystals of more uniform size.

Interestingly, the exact precipitant concentration could be used as a means to control crystal size. While 2.1 M ammonium sulphate gave the smallest crystals of about 5–10 μm size; with decreasing ammonium sulphate concentration crystals increased in size up to 50 μm at 1.8 M

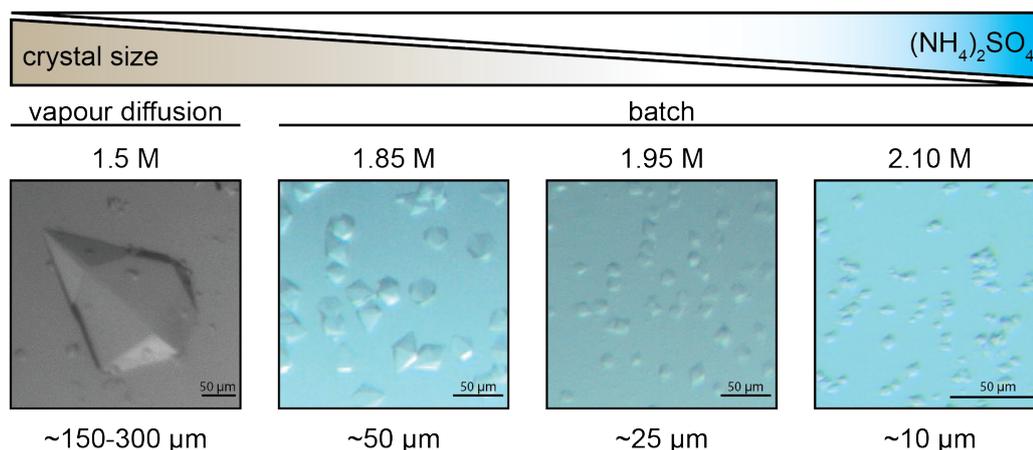


Figure A.2: Effect of increasing ammonium sulphate concentration on ADC micro-crystallisation. Vapour-diffusion crystallisation required the smallest concentration of ammonium sulphate. In micro-crystallisation experiments higher concentrations led to smaller crystals. Figure from Stohrer *et al.* [1].

ammonium sulphate (Figure A.2). Changes in concentration as small as 0.05 M ammonium sulphate were tested, which still resulted in differences to crystal size. This allows fine-tuning of crystal size according to the experimental requirements, rather than having to choose the experiment to be compatible with the obtained crystals.

A.3.2 Reproducibility of the crystallisation protocol

The tight dependency of ADC crystallisation on the precipitant concentration might explain why the system was highly sensitive to the age of the used buffers and ambient air humidity. The observation was made that dry air led to smaller crystals of a more homogeneous size distribution, while higher ambient humidity levels resulted in larger variations in crystal size (Figure A.3). Ammonium sulphate is known to be highly hygroscopic and, therefore, elevated air humidity might lead to changes in concentration of the stock solutions, resulting in differences in concentration in the final crystallisation conditions and consequently, to a higher variability in crystal size. Therefore, it was important to set up the crystallisation experiments in a humidity-controlled environment, such as inside a glove bag.

Similarly, protein quality had a major impact on crystallisation success. When using fresh protein, crystallisation was robust and could be finely controlled by adjusting the precipitant concentration. However, when using protein that had been stored at 4 °C for several weeks, polydispersion occurred, and some very large crystals appeared. This indicates non-uniform nuclei formation, possibly from some small aggregates acting as initial crystal growth sites. If the protein was re-purified by size exclusion chromatography, this effect could be slightly diminished, but an increased propensity to crystal size polydispersity remained an issue.

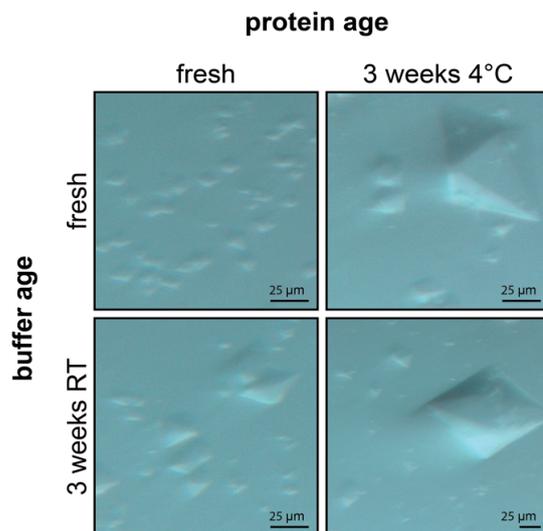


Figure A.3: Effects of buffer and protein age on ADC micro-crystallisation. Comparison of crystals obtained in batch crystallisation using the same conditions (2 M ammonium sulphate) but either set up with freshly purified protein, protein stored at 4 °C for three weeks, buffer prepared from fresh stock solutions and buffer prepared from stock solutions that were stored at RT for three weeks. The age of both, buffer and protein, caused more polydisperse crystal sizes. This effect was even more pronounced when combined. Figure from Stohrer *et al.* [1].

Nonetheless, respecting these limitations and setting up crystallisation trials inside a humidity-controlled glove bag and using freshly purified protein, allowed reproducible micro-crystallisation, generating crystals of defined size. Furthermore, the crystallisation reactions were successfully scaled up using protein volumes of up to 200 μL . Even at these volumes, crystal size was still very reproducible.

A.4 Micro-crystallisation of other model enzymes

When using ammonium sulphate as the precipitant, the complexity of a crystallisation experiment can be simplified by assuming that crystallisation is mainly driven by the precipitant. However, the effects of changing the ammonium sulphate concentration will very much depend on the crystallisation phase diagram of the investigated protein. Therefore, the micro-crystallisation results of *A. cycloclastes* copper nitrite reductase (AcNiR) and *E. coli* copper amine oxidase (ECAO) were analysed. The corresponding crystallisation experiments for AcNiR and ECAO were carried out by Dr. Sam Horrell and Susanne Meier, respectively.

For the generation of ECAO microcrystals, besides an increase in ammonium sulphate concentration, it was essential to omit sodium citrate from the batch crystallisation conditions. In contrast to ADC, even though an increase in precipitant concentration was required to obtain microcrystals, ECAO showed less sensitivity to the exact ammonium sulphate concentration. Indeed, concentrations ranging from 2.9–4 M all resulted in crystals of about the same size, going below or above this range resulted in no crystals at all. Even though crystal size was not as tuneable as for ADC, micro-crystallisation of ECAO was highly reproducible and large amounts of crystals could be produced.

In contrast to the other proteins, AcNiR crystals emerged from an initial precipitation phase. The amount of precipitation formed was dependent on the protein to precipitant ratio, and as such, also on the ammonium sulphate concentration. Microcrystals were obtained using either a 1:3 or a 1:2 ratio, accompanied by heavy or moderate precipitation, respectively. At a 1:1 protein to precipitant ratio, only very light precipitation was observed, and no crystals appeared. Interestingly, in the condition with the higher ammonium sulphate concentration, stronger precipitation was observed. However, under these conditions, crystals formed overnight, while at the lower ammonium sulphate concentration crystallisation took one to two days.

A.5 Discussion

As previously mentioned, batch crystallisation needs to be started inside the nucleation zone of the crystallisation phase diagram [351]. For all three systems tested, this required a shift to higher precipitant concentration compared to vapour diffusion crystallisation. The observed effects of altering the ammonium sulphate concentration for the three model enzymes can be explained by differences in the phase diagram of the specific proteins (Figure A.4). Increasing the ammonium sulphate concentration translates into a right shift on the phase diagram. For ADC this resulted in more, smaller crystals, which can be explained with a rather steep boundary between the metastable and the nucleation zone (Figure A.4, line a). At lower precipitant concentrations, the starting point (Figure A.4, 1i) will be closer to this boundary; thus, little time will be spent in the nucleation zone before entering the growth zone. This results in fewer, larger crystals. Shifting the starting point towards the right will increase the amount of time spent in the nucleation zone and therefore, smaller crystals are obtained (Figure A.4, 2). In contrast, for ECAO no effect on crystal size could be observed in a similar concentration range. This can be explained by a flatter boundary between the metastable and nucleation zone (Figure A.4, line b). In this case the time spent inside the nucleation zone is not significantly altered when shifting the starting point towards the right (Figure A.4, 1ii \rightarrow 2).

Unlike ADC and ECAO, AcNiR crystals emerged from an initial precipitation phase. To explain the observations made for this protein, not only the phase diagram but also other properties of the precipitant need to be considered. Ammonium sulphate is highly water soluble and its respective ions have great ability to “salt out” proteins [360]. Therefore, small changes in concentration might have a strong effect on protein solubility and thus on the speed with which nuclei are formed. The observed precipitation in AcNiR crystallisation implies that the precipitation process is fast. The initial precipitated phase is probably a mixture of amorphous aggregate and the ordered nuclei required for crystallisation. At a higher ammonium sulphate concentration, when mixing the protein and precipitant in a 1:3 ratio, protein is driven out of solution more quickly and crystals appear faster. However, this also leads to more amorphous aggregation. In contrast, if the protein to precipitant ratio is 1:2, and accordingly the ammonium sulphate concentration is lower, this process is slower and, therefore, more ordered nuclei are formed and less aggregation is observed. However, crystals will also take longer to appear.

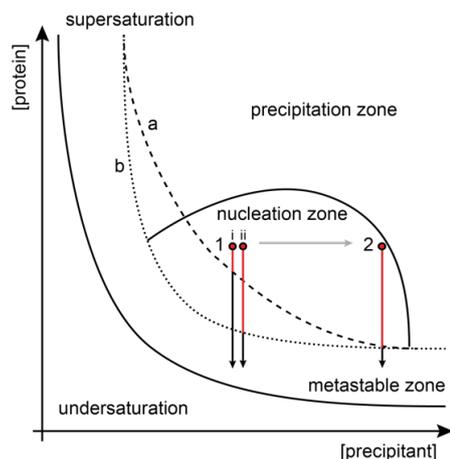


Figure A.4: Schematic of a generic crystallisation phase diagram. The proposed boundaries between the metastable and the nucleation zone are shown for ADC (a) and ECAO (b). The arrows represent time spent in the metastable zone for batch crystallisation at a lower ammonium sulphate concentration for ADC (1i) and ECAO (1ii) and a higher concentration (2). The red part of the arrow illustrates the difference in time spent in the nucleation zone, depending on the starting point and borderline between the metastable and the nucleation zone. Figure and legend from Stohrer *et al.* [1]

The observations made for these three model proteins demonstrate how the precipitating properties of ammonium sulphate can be exploited to quickly transition from vapour diffusion conditions to batch, allowing faster and easier generation of microcrystals required for time-resolved crystallographic experiments. Furthermore, depending on the phase diagram of the studied protein, adjusting the ammonium sulphate concentration can lead to a controlled way to adjust crystal size.

Ammonium sulphate is a very successful precipitant for protein crystallisation [361, 362] that has been used in about 7.3% of the crystallisation conditions reported in the protein data base (pdb) [1]. Therefore, this approach should also be applicable to many other proteins.

A.6 Application of microcrystals: testing 3D printed microfluidic chips for serial synchrotron experiments in flow

The produced ADC microcrystals were used to test a newly developed 3D printed microfluidic chip (3D-MiXD) [2]. The design of this 3D-MiXD device incorporated a 3D flow focusing geometry through a simple cross-shaped junction. Through this geometry, the sample is focused in the centre of the channel by buffer or water that is fed in through the cross channels. Using the ADC and other sample crystals, it was possible to generate a stable flow of crystals of around 30 μm size inside the chip. Through the 3D flow focusing the crystal flow was centred in the 3DMiXD and a more uniform crystal speed was obtained. Additionally, this helped to reduce sample waste and prevented fouling of the device.

Furthermore, the design of the chip allows time-resolved experiments to be performed by mixing in substrate through the side channels. The X-ray interaction region begins at the channel intersection. Thus, by collecting data at specific distances from the mixing point, different timepoints after the mixing event can be resolved.

Merging of about 10 000 still diffraction images allowed the structure of ADC to be solved at about 2 Å by molecular replacement, where it was possible to clearly identify a difference-density in the active site. Further analysis revealed that this density does not fit the substrate, but must come from a small molecule component in the crystallisation buffer. It was therefore not possible to assess whether substrate is able to diffuse into the crystal upon mixing.

Consequently, the next step will be to determine, if it is possible to exchange the crystals into a different buffer, removing this unknown component and leaving the active site free for the mixed in substrate.

Appendix B

Supplementary material

Table B.1: Composition of the DoE screen. The PEG400 concentration is given in % (v/v). Both buffers were added at 0.1 M concentration. pH refers to the pH of the used buffer. Hex = 2% (v/v) 2,5-hexanediol, Theo = 0.5 mM theophylline.

Well	PEG400	Buffer	pH	Salt	Hex	Theo
1 A	25	Na-citrate	5	LiSO ₄ (0.2 mM)	no	yes
2 A	25	Na-citrate	5	LiSO ₄ (0.2 mM)	yes	no
3 A	25	Na-citrate	5	NaCl (0.1 M)	no	yes
4 A	25	Na-citrate	5	NaCl (0.1 M)	yes	yes
5 A	25	Na-citrate	5	NaCl (0.1 M)	yes	no
6 A	25	Na-citrate	5	Na-tartrate (0.2 M)	no	yes
7 A	25	Na-citrate	5	Na-tartrate (0.2 M)	yes	no
8 A	25	Na-citrate	5	NaSCN (0.05 M)	no	yes
9 A	25	Na-citrate	5	NaSCN (0.05 M)	yes	no
10 A	25	Na-citrate	5	Zn-acetate (0.01 M)	no	no
11 A	25	Na-citrate	5	Zn-acetate (0.01 M)	yes	yes
12 A	25	Na-citrate	5.5	LiSO ₄ (0.2 mM)	no	no
1 B	25	Na-citrate	5.5	LiSO ₄ (0.2 mM)	yes	yes
2 B	25	Na-citrate	5.5	NaCl (0.1 M)	no	no
3 B	25	Na-citrate	5.5	NaCl (0.1 M)	yes	yes
4 B	25	Na-citrate	5.5	Na-tartrate (0.2 M)	no	yes
5 B	25	Na-citrate	5.5	Na-tartrate (0.2 M)	yes	no
6 B	25	Na-citrate	5.5	Na-tartrate (0.2 M)	yes	yes
7 B	25	Na-citrate	5.5	NaSCN (0.05 M)	no	yes
8 B	25	Na-citrate	5.5	NaSCN (0.05 M)	yes	no
9 B	25	Na-citrate	5.5	Zn-acetate (0.01 M)	no	no
10 B	25	Na-citrate	5.5	Zn-acetate (0.01 M)	yes	yes
11 B	25	MES	5	LiSO ₄ (0.2 mM)	no	no
12 B	25	MES	5	LiSO ₄ (0.2 mM)	yes	yes
1 C	25	MES	5	NaCl (0.1 M)	no	no

Table B.1: Composition of the DoE screen. (Continued)

Well	PEG400	Buffer	pH	Salt	Hex	Theo	
2	C	25	MES	5	NaCl (0.1 M)	no	yes
3	C	25	MES	5	NaCl (0.1 M)	yes	no
4	C	25	MES	5	Na-tartrate (0.2 M)	no	no
5	C	25	MES	5	Na-tartrate (0.2 M)	yes	yes
6	C	25	MES	5	NaSCN (0.05 M)	no	no
7	C	25	MES	5	NaSCN (0.05 M)	yes	yes
8	C	25	MES	5	Zn-acetate (0.01 M)	no	yes
9	C	25	MES	5	Zn-acetate (0.01 M)	yes	no
10	C	25	MES	5.5	LiSO ₄ (0.2 mM)	no	yes
11	C	25	MES	5.5	LiSO ₄ (0.2 mM)	yes	no
12	C	25	MES	5.5	NaCl (0.1 M)	no	yes
1	D	25	MES	5.5	NaCl (0.1 M)	yes	no
2	D	25	MES	5.5	NaCl (0.1 M)	yes	yes
3	D	25	MES	5.5	Na-tartrate (0.2 M)	no	no
4	D	25	MES	5.5	Na-tartrate (0.2 M)	no	yes
5	D	25	MES	5.5	Na-tartrate (0.2 M)	yes	no
6	D	25	MES	5.5	NaSCN (0.05 M)	no	no
7	D	25	MES	5.5	NaSCN (0.05 M)	yes	yes
8	D	25	MES	5.5	Zn-acetate (0.01 M)	no	yes
9	D	25	MES	5.5	Zn-acetate (0.01 M)	yes	no
10	D	30	Na-citrate	5	LiSO ₄ (0.2 mM)	yes	yes
11	D	30	Na-citrate	5	NaCl (0.1 M)	no	no
12	D	30	Na-citrate	5	Na-tartrate (0.2 M)	no	no
1	E	30	Na-citrate	5.5	NaSCN (0.05 M)	no	yes
2	E	30	Na-citrate	5.5	ZN-acetate (0.01 M)	yes	no
3	E	30	MES	5	NaSCN (0.05 M)	yes	no
4	E	30	MES	5	ZN-acetate (0.01 M)	no	yes
5	E	30	MES	5.5	LiSO ₄ (0.2 mM)	no	no
6	E	30	MES	5.5	Na-tartrate (0.2 M)	yes	yes
7	E	35	Na-citrate	5	LiSO ₄ (0.2 mM)	no	yes
8	E	35	Na-citrate	5	LiSO ₄ (0.2 mM)	yes	no
9	E	35	Na-citrate	5	NaCl (0.1 M)	no	no
10	E	35	Na-citrate	5	NaCl (0.1 M)	yes	yes
11	E	35	Na-citrate	5	Na-tartrate (0.2 M)	no	no
12	E	35	Na-citrate	5	Na-tartrate (0.2 M)	yes	yes
1	F	35	Na-citrate	5	NaSCN (0.05 M)	no	no
2	F	35	Na-citrate	5	NaSCN (0.05 M)	yes	yes
3	F	35	Na-citrate	5	Zn-acetate (0.01 M)	no	yes
4	F	35	Na-citrate	5	Zn-acetate (0.01 M)	yes	no
5	F	35	Na-citrate	5.5	LiSO ₄ (0.2 mM)	no	no

Table B.1: Composition of the DoE screen. *(Continued)*

	Well	PEG400	Buffer	pH	Salt	Hex	Theo	
	6	F	35	Na-citrate	5.5	LiSO ₄ (0.2 mM)	yes	yes
	7	F	35	Na-citrate	5.5	NaCl (0.1 M)	no	yes
	8	F	35	Na-citrate	5.5	NaCl (0.1 M)	yes	no
	9	F	35	Na-citrate	5.5	Na-tartrate (0.2 M)	no	no
	10	F	35	Na-citrate	5.5	Na-tartrate (0.2 M)	no	yes
	11	F	35	Na-citrate	5.5	Na-tartrate (0.2 M)	yes	no
	12	F	35	Na-citrate	5.5	NaSCN (0.05 M)	no	no
	1	G	35	Na-citrate	5.5	NaSCN (0.05 M)	yes	yes
	2	G	35	Na-citrate	5.5	Zn-acetate (0.01 M)	no	yes
	3	G	35	Na-citrate	5.5	Zn-acetate (0.01 M)	yes	no
	4	G	35	MES	5	LiSO ₄ (0.2 mM)	no	no
	5	G	35	MES	5	LiSO ₄ (0.2 mM)	yes	yes
	6	G	35	MES	5	NaCl (0.1 M)	no	yes
	7	G	35	MES	5	NaCl (0.1 M)	yes	no
	8	G	35	MES	5	NaCl (0.1 M)	yes	yes
	9	G	35	MES	5	Na-tartrate (0.2 M)	no	yes
	10	G	35	MES	5	Na-tartrate (0.2 M)	yes	no
	11	G	35	MES	5	NaSCN (0.05 M)	no	yes
	12	G	35	MES	5	NaSCN (0.05 M)	yes	no
	1	H	35	MES	5	Zn-acetate (0.01 M)	no	no
	2	H	35	MES	5	Zn-acetate (0.01 M)	yes	yes
	3	H	35	MES	5.5	LiSO ₄ (0.2 mM)	no	yes
	4	H	35	MES	5.5	LiSO ₄ (0.2 mM)	yes	no
	5	H	35	MES	5.5	NaCl (0.1 M)	no	no
	6	H	35	MES	5.5	NaCl (0.1 M)	yes	yes
	7	H	35	MES	5.5	Na-tartrate (0.2 M)	no	no
	8	H	35	MES	5.5	Na-tartrate (0.2 M)	yes	yes
	9	H	35	MES	5.5	NaSCN (0.05 M)	no	yes
	10	H	35	MES	5.5	NaSCN (0.05 M)	yes	no
	11	H	35	MES	5.5	Zn-acetate (0.01 M)	no	no
	12	H	35	MES	5.5	Zn-acetate (0.01 M)	yes	yes

Table B.2: List of used primers. Primer sequences are given in 5' → 3' direction. T_m indicates the melting temperature for the annealing part of the primer. FW = forward primer, RV = reverse primer, mam. expression = for expression in mammalian cell culture.

ID	Oligo Name	Sequence (5' → 3')	T_m	Purpose
Cloning primers				
037	2411-FW	ATCGGGCGCGGATCCATGAAGACCATC ATCGC	59	HRV removal, transfer into pFastBac vector
038	2411-RV	CTCGACAAGCTTCTATTAGTGGTGATG GTGATGGTGATGGTGATGGTGGGCAGC TGCCTTGAAAGGTTCTTGC	63	HRV removal, transfer into pFastBac vector
039	pFastbac739-FW	TAGAAGCTTGTCGAGAAGTACTAGAGG	67	vector linearisation
040	pFastbac739-RV	GGATCCGCGCCCGAT	69	vector linearisation
042	A _{2A} pcDNA3.1-FW	CTAGCGTTTAAACTTAAGCCCATGAAG ACCATCATC	56	cloning into pcDNA3.1 vector for mam. expression
043	A _{2A} pcDNA3.1-RV	TAAACGGGCCCTCTATTAGTGGTGGTG GTG	57	cloning into pcDNA3.1 vector for mam. expression
044	pcDNA3.1-lin-FW	CACCACCACCACTAATAGAGGGCCCGT TTA	58	vector linearisation
045	pcDNA3.1-lin-RV	GATGATGGTCTTCATGGGCTTAAGTTT AAACGCTAG	55	vector linearisation
Sequencing primers				
	pFASTBAC F	TCCGGATTATTCATACCGTCCC	60.3	eurofins standard vector primer
	pFASTBAC R	CCTCTACAAATGTGGTATGGCTG	60.6	eurofins standard vector primer
	A _{2A} seq 2	CTTCAACTTCTTTGCCTGTGTGCTGG	59.5	A _{2A} sequencing (seq start at ~100 bp)
	CMVf	CGCAAATGGGCGGTAGGCGTG	75	sequencing primer FW pcDNA3.3
	bGHrev	ACACCTACTCAGACAATGCG	64	sequencing primer RV pcDNA3.3

Table B.3: List of used A_{2A}R constructs. The backbone indicates the used expression vector and antibiotic resistance selection marker of the vector (resistance). The pACB-pp64-3 vector additionally carries a fluorescent expression marker (exp. marker) that is co-expressed with the receptor.

Name	A _{2A} R-bRIL	A _{2A} -StaR2-bRIL _{HRV-HIS}	A _{2A} -StaR2-bRIL	A _{2A} -StaR2-bRIL	WT-A _{2A} -mCherry	FL-A _{2A} -StaR2-N-bRIL	FL-WT-A _{2A} -N-bRIL	WT-A _{2A} (mam)
Backbone	pFsatBac	pACB-gp64-3	pFsatBac	pACB-PH	pFastBac	pFastBac	pFastBac	pcDNA3.3
Resistance	AmpR	AmpR	AmpR	AmpR	AmpR	AmpR	AmpR	Amp/Kan
Exp. marker	none	EGFP (p 6:9)	none	none	none	none	none	none
Promoter	PH	GP-64	PH	PH	PH	PH	PH	CMV
Exp. host	Insect cells, <i>Sf9</i>	Insect cells, <i>Sf9</i>	Insect cells, <i>Sf9</i>	Insect cells, <i>Sf9</i>	Insect cells, <i>Sf9</i>	Insect cells, <i>Sf9</i>	Insect cells, <i>Sf9</i>	HEK293, CHO
N-tag	HA-Flag	HA-Flag	HA-Flag	HA-Flag	HA-Flag	HA-Flag	HA-Flag	HA-Flag
Insert	ADORA2A P2-A316	ADORA2A, P2-A316	ADORA2A, P2-A316	ADORA2A P2-S412 (FL)	ADORA2A P2-S412 (FL)	ADORA2A P2-S412 (FL)	ADORA2A P2-S412 (FL)	ADORA2A P2-S412 (FL)
PDB	4EIY	5IU4	(A54L, T88A, R107A, K122A, N154A, L202A, L235A, V239A, S277A), bRIL insertion at L208/E219	10 × HIS	10 × HIS	(A54L, T88A, R107A, K122A, L202A, L235A, V239A, S277A, N154A), N-term-bRIL	WT	
Modifications	bRIL insertion at L208/E219			N-term-bRIL			WT, N-term-bRIL	
C-tag	10 × HIS	PreSc-10 × HIS	10 × HIS	mCherry-PreSc-10 × HIS	PreSc-10 × HIS	PreSc-10 × HIS	PreSc-10 × HIS	PreSc-10 × HIS
Application	crystallography	crystallography	crystallography	gel shift assay	binding assay	binding assay	binding assay	pharmacology

Table B.4: Docking scores and residues predicted to form polar interactions for ligand docking into the A_{2A}R (4EIY) determined using a positional π -constraint. The positional constraint was defined by a 2 Å radius around an aromatic ligand core that forms a π -stacking interaction with Phe168^{45,52}. The list of ligands is sorted according to the highest Glide docking score. Residues predicted to be involved in polar interactions are listed according to interaction type as π -stacking interactions, hydrogen (H-) bonds, and salt bridges. If nothing else is stated, the interaction is formed by the main functional group of the residue, otherwise the involved group is stated in squared brackets (NH₂ = amino group, O = hydroxy group, bb = backbone).

Ligand	Score	Interactions		
		π -stacking	H-bonds	Salt bridge
FOC-2	-12.194	F168, Y271	N253 [O], E169	-
FOC-3	-11.023	F168, Y271	N253 [NH ₂ , O], E169	-
ZMA	-10.56	F168, H250	N253 [NH ₂ (2×), O], E169	-
FOC-5	-10.167	F168, H250, Y271	N253 [NH ₂ , O], E169	-
DEL-3	-8.948	F168	F168 [bbO], Y271, E169	K150 (π -cation)
FOC-1	-8.843	F168 (2×), Y271	-	H264 (π -cation)
DEL-1	-8.472	-	E169 [bbNH ₂], Y271, S67 [bbO]	-
FOC-4	-8.311	F168	N253 [NH ₂ , O], E169	-
FOC-6	-8.18	F168	N253 [O], E169	-
Theo	-7.704	F168 (2×)	N253 [NH ₂ , O]	-
DEL-4	-7.603	F168, H250	N253 [NH ₂], F168 [bbNH ₂]	-
DEL-6	-7.376	H264	K153, K150, Y271	K153, K150, H264 (π -cation)
DEL-2	-6.683	F168	S67 [bbO], E169[O, bbNH ₂], Y271	N253 [NH ₂] (halogen), K150 (π -cation)
DEL-5	-6.598	H264	E169 [O, bbNH ₂]	H264 (π -cation)

Table B.5: Data collection and refinement statistics. Statistics are given for the A_{2A}-bRIL-ZM241385 complex after ligand exchange with either FOC-2, DEL-2, or FOC-1. * Refinement was not carried out fully. Statistics are reported after three rounds of Phenix-refinement. ** Only some water were molecules modelled during these initial refinement rounds. Statistics for the highest-resolution shell are shown in parentheses. MR = molecular replacement.

Data collection	FOC-2	DEL-1	FOC-1
Space group	C 2 2 2 ₁	C 2 2 2 ₁	C 2 2 2 ₁
Cell dimensions			
a (Å)	39.57	39.66	39.62
b (Å)	179.19	179.59	179.78
c (Å)	140.03	139.86	139.89
α (°)	90.0	90.0	90.0
β (°)	90.0	90.0	90.0
γ (°)	90.0	90.0	90.0
Source	DLS: I24	DLS: I24	DLS: I24
Wavelength (Å)	0.9686	0.9686	0.9686
Resolution (Å)	41.39–2.05 (2.123–2.05)	37.78–2.34 (2.424–2.34)	41.39–2.39 (2.475–2.39)
Measured reflections	271 337 (25 537)	164 529 (15 914)	119 034 (11 156)
Unique reflections	31 606 (3141)	21 645 (2116)	20 204 (2022)
Completeness (%)	97.11 (90.98)	99.09 (96.46)	98.28 (97.00)
CC _{1/2}	0.964 (0.326)	21 645 (2116)	0.818 (0.239)
Mean I/σ(I)	3.37 (0.65)	2.50 (0.46)	3.21 (0.82)
Multiplicity	8.6 (8.1)	7.6 (7.5)	5.9 (5.5)
B-Factors (Å ²)	25.18	30.72	26.03
R _{merge}	0.4149 (2.099)	0.7019 (2.679)	0.5426 (1.118)
R _{mes}	0.4416 (2.24)	0.7524 (2.873)	0.593 (1.229)
R _{pim}	0.1452 (0.7587)	0.265 (1.014)	0.2325 (0.4964)
Refinement*	FOC-2	DEL-1	FOC-1
MR Search model	4EIY- <i>apo</i>	4EIY- <i>apo</i>	4EIY- <i>apo</i>
R _{work} (%)	22.19	22.81	23.71
R _{free} (%)	26.19	29.39	29.78
No. of atoms	3304	3234	3130
Protein	3105	3105	3105
Ligand	25	25	25
Water**	174	104	-
R. M. S. Deviations			
Bond length(Å)	0.008	0.009	0.009
Bond angle (°)	1.03	1.05	1.16
Ramachandran statistic (%)			
Favoured	98.17	97.38	97.64
Allowed	1.83	2.62	2.36
Outliers	0	0	0

Bibliography

- [1] Stohrer, C., Horrell, S., Meier, S., Sans, M., Von Stetten, D., Hough, M., Goldman, A., Monteiro, D.C. and Pearson, A.R. Homogeneous batch micro-crystallization of proteins from ammonium sulfate. *Acta Crystallographica Section D: Structural Biology*, **77**(2), pp. 194–204 (2021).
- [2] Monteiro, D.C., Von Stetten, D., Stohrer, C., Sans, M., Pearson, A.R., Santoni, G., Van Der Linden, P. and Trebbin, M. 3D-MiXD: 3D-printed X-ray-compatible microfluidic devices for rapid, low-consumption serial synchrotron crystallography data collection in flow. *IUCrJ*, **7**(2), pp. 207–219 (2020).
- [3] Cecchetti, C., Strauss, J., Stohrer, C., Naylor, C., Pryor, E., Hobbs, J., Tanley, S., Goldman, A. and Byrne, B. A novel high-throughput screen for identifying lipids that stabilise membrane proteins in detergent based solution. *PLoS ONE*, **16**(7 July), pp. 1–20 (2021).
- [4] DiMasi, J.A., Grabowski, H.G. and Hansen, R.W. Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of Health Economics*, **47**, pp. 20–33 (2016).
- [5] Matthews, H., Hanison, J. and Nirmalan, N. "Omics"-informed drug and biomarker discovery: Opportunities, challenges and future perspectives. *Proteomes*, **4**(3) (2016).
- [6] Swinney, D.C. and Anthony, J. How were new medicines discovered? *Nature Reviews Drug Discovery*, **10**(7), pp. 507–519 (2011).
- [7] Singer, J.A. and Purcell, W.P. Relationships Among Current Quantitative Structure-Activity Models. *Journal of Medicinal Chemistry*, **10**(6), pp. 1000–1002 (1967).
- [8] Guha, R. On exploring structure-activity relationships. *Methods in Molecular Biology*, **993**, pp. 81–94 (2013).
- [9] Afshar, M., Prescott, C.D. and Varani, G. Structure-based and combinatorial search for new RNA-binding drugs. *Current Opinion in Biotechnology*, **10**(1), pp. 59–63 (1999).
- [10] Gallego, J. and Varani, G. Targeting RNA with Small-molecule drugs: Therapeutic promise and chemical challenges. *Accounts of Chemical Research*, **34**(10), pp. 836–843 (2001).

- [11] Sriram, K. and Insel, P.A. G protein-coupled receptors as targets for approved drugs: How many targets and how many drugs? In *Molecular Pharmacology*, vol. 93, pp. 251–258. American Society for Pharmacology and Experimental Therapy (2018).
- [12] Hauser, A.S., Attwood, M.M., Rask-Andersen, M., Schiöth, H.B. and Gloriam, D.E. Trends in GPCR drug discovery: New agents, targets and indications. *Nature Reviews Drug Discovery*, **16**(12), pp. 829–842 (2017).
- [13] Enyedy, I.J., Huang, Y., Long, Y.Q., Roller, P.P., Yang, D., Wang, S., Ling, Y., Nacro, K., Tomita, Y., Wu, X., Cao, Y., Guo, R., Li, B. and Zhu, X. Discovery of small-molecule inhibitors of Bcl-2 through structure-based computer screening. *Journal of Medicinal Chemistry*, **44**(25), pp. 4313–4324 (2001).
- [14] Cano, P. and Fan, B. A geometric and algebraic view of MHC-peptide complexes and their binding properties. *BMC Structural Biology*, **1**, pp. 1–14 (2001).
- [15] Verlinde, C.L. and Hol, W.G. Structure-based drug design: progress, results and challenges. *Structure*, **2**(7), pp. 577–587 (1994).
- [16] Lipinski, C.A., Lombardo, F., Dominy, B.W. and Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, **46**(1-3), pp. 3–26 (2001).
- [17] Erlanson, D.A., Fesik, S.W., Hubbard, R.E., Jahnke, W. and Jhoti, H. Twenty years on: The impact of fragments on drug discovery. *Nature Reviews Drug Discovery*, **15**(9), pp. 605–619 (2016).
- [18] Hopkins, A.L. and Groom, C.R. The druggable genome. *Nature Reviews Drug Discovery*, **1**(9), pp. 727–730 (2002).
- [19] Li, G., Zheng, W., Chen, Z., Zhou, Y., Liu, Y., Yang, J., Huang, Y. and Li, X. Design, preparation, and selection of DNA-encoded dynamic libraries. *Chemical Science*, **6**(12), pp. 7097–7104 (2015).
- [20] Reddavid, F.V., Lin, W., Lehnert, S. and Zhang, Y. DNA-Encoded Dynamic Combinatorial Chemical Libraries. *Angewandte Chemie - International Edition*, **54**(27), pp. 7924–7928 (2015).
- [21] Urbina, H.D., Debaene, F., Jost, B., Bole-Feysot, C., Mason, D.E., Kuzmic, P., Harris, J.L. and Winssinger, N. Self-Assembled Small-Molecule Microarrays for Protease Screening and Profiling. *ChemBioChem*, **7**(11), pp. 1790–1797 (2006).
- [22] Harris, J., Mason, D.E., Li, J., Burdick, K.W., Backes, B.J., Chen, T., Shipway, A., Van Heeke, G., Gough, L., Ghaemmaghami, A., Shakib, F., Debaene, F. and Winssinger, N. Activity Profile of Dust Mite Allergen Extract Using Substrate Libraries and Functional Proteomic Microarrays. *Chemistry & Biology*, **11**(10), pp. 1361–1372 (2004).
- [23] Wrenn, S.J., Weisinger, R.M., Halpin, D.R. and Harbury, P.B. Synthetic ligands discovered by in vitro selection. *Journal of the American Chemical Society*, **129**(43), pp. 13137–13143 (2007).

- [24] Franzini, R.M. and Randolph, C. Chemical space of DNA-encoded libraries: Miniperspective. *Journal of Medicinal Chemistry*, **59**(14), pp. 6629–6644 (2016).
- [25] Cheng, T., Li, Q., Zhou, Z., Wang, Y. and Bryant, S.H. Structure-based virtual screening for drug discovery: A problem-centric review. *AAPS Journal*, **14**(1), pp. 133–141 (2012).
- [26] Trott, O. and Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, **31**(2), pp. 455–461 (2009).
- [27] Coleman, R.G., Carchia, M., Sterling, T., Irwin, J.J. and Shoichet, B.K. Ligand Pose and Orientational Sampling in Molecular Docking. *PLoS ONE*, **8**(10), p. e75992 (2013).
- [28] Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Mainz, D.T., Repasky, M.P., Knoll, E.H., Shelley, M., Perry, J.K., Shaw, D.E., Francis, P. and Shenkin, P.S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry*, **47**(7), pp. 1739–1749 (2004).
- [29] Huang, S.Y., Grinter, S.Z. and Zou, X. Scoring functions and their evaluation methods for protein-ligand docking: Recent advances and future directions. *Physical Chemistry Chemical Physics*, **12**(40), pp. 12899–12908 (2010).
- [30] Huang, N., Kalyanaraman, C., Irwin, J.J. and Jacobson, M.P. Physics-based scoring of protein - Ligand complexes: Enrichment of known inhibitors in large-scale virtual screening. In *Journal of Chemical Information and Modeling*, vol. 46, pp. 243–253. J Chem Inf Model (2006).
- [31] Muegge, I. PMF scoring revisited. *Journal of Medicinal Chemistry*, **49**(20), pp. 5895–5902 (2006).
- [32] Pason, L.P. and Sotriffer, C.A. Empirical Scoring Functions for Affinity Prediction of Protein-ligand Complexes. *Molecular Informatics*, **35**(11-12), pp. 541–548 (2016).
- [33] Sotriffer, C.A., Sanschagrin, P., Matter, H. and Klebe, G. SFCscore: Scoring functions for affinity prediction of protein-ligand complexes. *Proteins: Structure, Function, and Bioinformatics*, **73**(2), pp. 395–419 (2008).
- [34] Ballester, P.J. and Mitchell, J.B. A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking. *Bioinformatics*, **26**(9), pp. 1169–1175 (2010).
- [35] Ballester, P.J., Schreyer, A. and Blundell, T.L. Does a more precise chemical description of protein-ligand complexes lead to more accurate prediction of binding affinity? *Journal of Chemical Information and Modeling*, **54**(3), pp. 944–955 (2014).
- [36] Ain, Q.U., Aleksandrova, A., Roessler, F.D. and Ballester, P.J. Machine-learning scoring functions to improve structure-based binding affinity prediction and virtual screening. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, **5**(6), pp. 405–424 (2015).

- [37] Ericksen, S.S., Wu, H., Zhang, H., Michael, L.A., Newton, M.A., Hoffmann, F.M. and Wildman, S.A. Machine Learning Consensus Scoring Improves Performance Across Targets in Structure-Based Virtual Screening. *Journal of Chemical Information and Modeling*, **57**(7), pp. 1579–1590 (2017).
- [38] Palacio-Rodríguez, K., Lans, I., Cavasotto, C.N. and Cossio, P. Exponential consensus ranking improves the outcome in docking and receptor ensemble docking. *Scientific Reports*, **9**(1), pp. 1–14 (2019).
- [39] Batool, M., Ahmad, B. and Choi, S. A structure-based drug discovery paradigm. *International Journal of Molecular Sciences*, **20**(11) (2019).
- [40] Ferreira, L.G., Dos Santos, R.N., Oliva, G. and Andricopulo, A.D. Molecular docking and structure-based drug design strategies. *Molecules*, **20**(7), pp. 13384–13421 (2015).
- [41] Roberts, N.A., Martin, J.A., Kinchington, D., Broadhurst, A.V., Craig, J.C., Duncan, I.B., Galpin, S.A., Handa, B.K., Kay, J., Kröhn, A., Lambert, R.W., Merrett, J.H., Mills, J.S., Parkes, K.E., Redshaw, S., Ritchie, A.J., Taylor, D.L., Thomas, G.J. and Machin, P.J. Rational design of peptide-based HIV proteinase inhibitors. *Science*, **248**(4953), pp. 358–361 (1990).
- [42] Fritz, T.A., Tondi, D., Finer-Moore, J.S., Costi, M.P. and Stroud, R.M. Predicting and harnessing protein flexibility in the design of species-specific inhibitors of thymidylate synthase. *Chemistry and Biology*, **8**(10), pp. 981–995 (2001).
- [43] Shoichet, B.K., McGovern, S.L., Wei, B. and Irwin, J.J. Lead discovery using molecular docking. *Current Opinion in Chemical Biology*, **6**(4), pp. 439–446 (2002).
- [44] Salsbury, F.R. Molecular dynamics simulations of protein dynamics and their relevance to drug discovery. *Current Opinion in Pharmacology*, **10**(6), pp. 738–744 (2010).
- [45] Alonso, H., Bliznyuk, A.A. and Gready, J.E. Combining docking and molecular dynamic simulations in drug design. *Medicinal Research Reviews*, **26**(5), pp. 531–568 (2006).
- [46] Wlodawer, A. and Vondrasek, J. Inhibitors of HIV-1 Protease: A Major Success of Structure-Assisted Drug Design. *Annual Review of Biophysics and Biomolecular Structure*, **27**(1), pp. 249–284 (1998).
- [47] Marrakchi, H., Lanéelle, G. and Quémard, A. InhA, a target of the antituberculous drug isoniazid, is involved in a mycobacterial fatty acid elongation system, FAS-II. *Microbiology*, **146**(2), pp. 289–296 (2000).
- [48] Wang, L., Gu, Q., Zheng, X., Ye, J., Liu, Z., Li, J., Hu, X., Hagler, A. and Xu, J. Discovery of new selective human aldose reductase inhibitors through virtual screening multiple binding pocket conformations. *Journal of Chemical Information and Modeling*, **53**(9), pp. 2409–2422 (2013).
- [49] Ren, J.X., Li, L.L., Zheng, R.L., Xie, H.Z., Cao, Z.X., Feng, S., Pan, Y.L., Chen, X., Wei, Y.Q. and Yang, S.Y. Discovery of novel Pim-1 kinase inhibitors by a hierarchical multistage virtual screening approach based on svm model, pharmacophore, and molecular docking. *Journal of Chemical Information and Modeling*, **51**(6), pp. 1364–1375 (2011).

- [50] Rucktooa, P., Cheng, R.K., Segala, E., Geng, T., Errey, J.C., Brown, G.A., Cooke, R.M., Marshall, F.H. and Doré, A.S. Towards high throughput GPCR crystallography: In Meso soaking of Adenosine A2A Receptor crystals. *Scientific Reports*, **8**(1), p. 41 (2018).
- [51] Cheng, R.K. Towards an Optimal Sample Delivery Method for Serial Crystallography at XFEL. *Crystals*, **10**(3), p. 215 (2020).
- [52] Healey, R.D., Basu, S., Humm, A.S., Leyrat, C., Cong, X., Golebiowski, J., Dupeux, F., Pica, A., Granier, S. and Márquez, J.A. An automated platform for structural analysis of membrane proteins through serial crystallography. *Cell Reports Methods*, **1**(6), p. 100102 (2021).
- [53] Singh, O., Shillings, A., Craggs, P., Wall, I., Rowland, P., Skarzynski, T., Hobbs, C.I., Hardwick, P., Tanner, R., Blunt, M., Witty, D.R. and Smith, K.J. Crystal structures of ASK1-inhibitor complexes provide a platform for structure-based drug design. *Protein Science*, **22**(8), pp. 1071–1077 (2013).
- [54] Caffrey, M. A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes. *Acta Crystallographica. Section F, Structural biology communications*, **71**(Pt 1), pp. 3–18 (2015).
- [55] Huang, C.Y., Olieric, V., Ma, P., Panepucci, E., Diederichs, K., Wang, M. and Caffrey, M. In meso in situ serial X-ray crystallography of soluble and membrane proteins. *Acta Crystallographica Section D: Biological Crystallography*, **71**, pp. 1238–1256 (2015).
- [56] Huang, C.Y., Olieric, V., Ma, P., Howe, N., Vogeley, L., Liu, X., Warshamanage, R., Weinert, T., Panepucci, E., Kobilka, B., Diederichs, K., Wang, M. and Caffrey, M. In meso in situ serial X-ray crystallography of soluble and membrane proteins at cryogenic temperatures. *Acta Crystallographica Section D: Structural Biology*, **72**(1), pp. 93–112 (2016).
- [57] Huang, C.Y., Olieric, V., Howe, N., Warshamanage, R., Weinert, T., Panepucci, E., Vogeley, L., Basu, S., Diederichs, K., Caffrey, M. and Wang, M. In situ serial crystallography for rapid de novo membrane protein structure determination. *Communications Biology*, **1**(1) (2018).
- [58] Anderson, A.C. The process of structure-based drug design. *Chemistry and Biology*, **10**(9), pp. 787–797 (2003).
- [59] Klebe, G. Virtual ligand screening: strategies, perspectives and limitations. *Drug Discovery Today*, **11**(13-14), pp. 580–594 (2006).
- [60] Bissantz, C., Kuhn, B. and Stahl, M. A medicinal chemist's guide to molecular interactions. *Journal of Medicinal Chemistry*, **53**(14), pp. 5061–5084 (2010).
- [61] Michel, J., Tirado-Rives, J. and Jorgensen, W.L. Energetics of displacing water molecules from protein binding sites: Consequences for ligand optimization. *Journal of the American Chemical Society*, **131**(42), pp. 15403–15411 (2009).

- [62] Amadasi, A., Surface, J.A., Spyralis, F., Cozzini, P., Mozzarelli, A. and Kellogg, G.E. Robust classification of "relevant" water molecules in putative protein binding sites. *Journal of Medicinal Chemistry*, **51**(4), pp. 1063–1067 (2008).
- [63] Harner, M.J., Frank, A.O. and Fesik, S.W. Fragment-based drug discovery using NMR spectroscopy. *Journal of Biomolecular NMR*, **56**(2), pp. 65–75 (2013).
- [64] Studer, G., Biasini, M. and Schwede, T. Assessing the local structural quality of transmembrane protein models using statistical potentials (QMEANBrane). In *Bioinformatics*, vol. 30. Bioinformatics (2014).
- [65] Bartesaghi, A., Merk, A., Banerjee, S., Matthies, D., Wu, X., Milne, J.L. and Subramaniam, S. 2.2 Å resolution cryo-EM structure of β -galactosidase in complex with a cell-permeant inhibitor. *Science*, **348**(6239), pp. 1147–1151 (2015).
- [66] Zhang, X., Belousoff, M.J., Zhao, P., Kooistra, A.J., Truong, T.T., Ang, S.Y., Underwood, C.R., Egebjerg, T., Šenel, P., Stewart, G.D., Liang, Y.L., Glukhova, A., Venugopal, H., Christopoulos, A., Furness, S.G., Miller, L.J., Reedtz-Runge, S., Langmead, C.J., Gloriam, D.E., Danev, R., Sexton, P.M. and Wootten, D. Differential GLP-1R Binding and Activation by Peptide and Non-peptide Agonists. *Molecular Cell*, **80**(3), pp. 485–500.e7 (2020).
- [67] Ceska, T., Chung, C.W., Cooke, R., Phillips, C. and Williams, P.A. Cryo-EM in drug discovery. *Biochemical Society Transactions*, **47**(1), pp. 281–293 (2019).
- [68] Congreve, M., de Graaf, C., Swain, N.A. and Tate, C.G. Impact of GPCR Structures on Drug Discovery. *Cell*, **181**(1), pp. 81–91 (2020).
- [69] Fredriksson, R., Lagerström, M.C., Ludin, L.G. and Schiöth, H.B. The G-Protein-Coupled Receptors in the Human Genome Form Five Main Families. Phylogenetic Analysis, Paralogon Groups, and Fingerprints. *Molecular Pharmacology*, **63**(6), pp. 1256–1272 (2003).
- [70] Wu, H., Wang, C., Gregory, K.J., Han, G.W., Cho, H.P., Xia, Y., Niswender, C.M., Katritch, V., Meiler, J., Cherezov, V., Conn, P.J. and Stevens, R.C. Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science*, **344**(6179), pp. 58–64 (2014).
- [71] Dann, C.E., Hsieh, J.C., Rattner, A., Sharma, D., Nathans, J. and Leahy, D.J. Insights into Wnt binding and signalling from the structures of two Frizzled cysteine-rich domains. *Nature*, **412**(6842), pp. 86–90 (2001).
- [72] Schiöth, H.B. and Lagerström, M.C. Structural diversity of g proteincoupled receptors and significance for drug discovery. *Nature Reviews Drug Discovery*, **7**(4), pp. 339–357 (2008).
- [73] Hamm, H.E., Deretic, D., Arendt, A., Hargrave, P.A., Koenig, B. and Hofmann, K.P. Site of G protein binding to rhodopsin mapped with synthetic peptides from the α subunit. *Science*, **241**(4867), pp. 832–835 (1988).

- [74] Syrovatkina, V., Alegre, K.O., Dey, R. and Huang, X.Y. Regulation, Signaling, and Physiological Functions of G-Proteins. *Journal of Molecular Biology*, **428**(19), pp. 3850–3868 (2016).
- [75] Zhao, D.Y., Pöge, M., Morizumi, T., Gulati, S., Van Eps, N., Zhang, J., Miszta, P., Filipek, S., Mahamid, J., Plitzko, J.M., Baumeister, W., Ernst, O.P. and Palczewski, K. Cryo-EM structure of the native rhodopsin dimer in nanodiscs. *Journal of Biological Chemistry*, **294**(39), pp. 14215–14230 (2019).
- [76] Cherfils, J. and Chabre, M. Activation of G-protein $G\alpha$ subunits by receptors through $G\alpha$ - $G\beta$ and $G\alpha$ - $G\gamma$ interactions. *Trends in Biochemical Sciences*, **28**(1), pp. 13–17 (2003).
- [77] Van Eps, N., Preininger, A.M., Alexander, N., Kaya, A.I., Meier, S., Meiler, J., Hamm, H.E. and Hubbell, W.L. Interaction of a G protein with an activated receptor opens the interdomain interface in the alpha subunit. *Proceedings of the National Academy of Sciences of the United States of America*, **108**(23), pp. 9420–9424 (2011).
- [78] Pierce, K.L., Premont, R.T. and Lefkowitz, R.J. Seven-transmembrane receptors. *Nature Reviews Molecular Cell Biology*, **3**(9), pp. 639–650 (2002).
- [79] Mizuno, N. and Itoh, H. Functions and regulatory mechanisms of Gq-signaling pathways. *NeuroSignals*, **17**(1), pp. 42–54 (2009).
- [80] Suzuki, N., Hajicek, N. and Kozasa, T. Regulation and physiological functions of G12/13-mediated signaling pathways. *NeuroSignals*, **17**(1), pp. 55–70 (2009).
- [81] Liang, Y.L., Khoshouei, M., Radjainia, M., Zhang, Y., Glukhova, A., Tarrasch, J., Thal, D.M., Furness, S.G.B., Christopoulos, G., Coudrat, T., Danev, R., Baumeister, W., Miller, L.J., Christopoulos, A., Kobilka, B.K., Wootten, D., Skiniotis, G. and Sexton, P.M. Phase-plate cryo-EM structure of a class B GPCR–G-protein complex. *Nature*, **546**(7656), pp. 118–123 (2017).
- [82] Liang, Y.L., Khoshouei, M., Glukhova, A., Furness, S.G., Zhao, P., Clydesdale, L., Koole, C., Truong, T.T., Thal, D.M., Lei, S., Radjainia, M., Danev, R., Baumeister, W., Wang, M.W., Miller, L.J., Christopoulos, A., Sexton, P.M. and Wootten, D. Phase-plate cryo-EM structure of a biased agonistbound human GLP-1 receptor-Gs complex. *Nature*, **555**(7694), pp. 121–125 (2018).
- [83] Zhang, H., Qiao, A., Yang, D., Yang, L., Dai, A., De Graaf, C., Reedtz-Runge, S., Dharmarajan, V., Zhang, H., Han, G.W., Grant, T.D., Sierra, R.G., Weierstall, U., Nelson, G., Liu, W., Wu, Y., Ma, L., Cai, X., Lin, G., Wu, X., Geng, Z., Dong, Y., Song, G., Griffin, P.R., Lau, J., Cherezov, V., Yang, H., Hanson, M.A., Stevens, R.C., Zhao, Q., Jiang, H., Wang, M.W. and Wu, B. Structure of the full-length glucagon class B G-protein-coupled receptor. *Nature*, **546**(7657), pp. 259–264 (2017).
- [84] García-Nafría, J., Nehmé, R., Edwards, P.C. and Tate, C.G. Cryo-EM structure of the serotonin 5-HT1B receptor coupled to heterotrimeric Go. *Nature*, **558**(7711), pp. 620–623 (2018).

- [85] Campbell, A.P. and Smrcka, A.V. Targeting G protein-coupled receptor signalling by blocking G proteins. *Nature Reviews Drug Discovery*, **17**(11), pp. 789–803 (2018).
- [86] Pitcher, J.A., Freedman, N.J. and Lefkowitz, R.J. G protein-coupled receptor kinases. *Annual Review of Biochemistry*, **67**(1), pp. 653–692 (1998).
- [87] DeWire, S.M., Ahn, S., Lefkowitz, R.J. and Shenoy, S.K. β -Arrestins and cell signaling. *Annual Review of Physiology*, **69**, pp. 483–510 (2007).
- [88] Jean-Charles, P.Y., Kaur, S. and Shenoy, S.K. G Protein-Coupled Receptor Signaling Through β -Arrestin-Dependent Mechanisms. *Journal of Cardiovascular Pharmacology*, **70**(3), pp. 142–158 (2017).
- [89] Tan, L., Yan, W., McCorvy, J.D. and Cheng, J. Biased Ligands of G Protein-Coupled Receptors (GPCRs): Structure-Functional Selectivity Relationships (SFSRs) and Therapeutic Potential. *Journal of Medicinal Chemistry*, **61**(22), pp. 9841–9878 (2018).
- [90] Smith, M.A., Pirrat, P., Pearson, A.R., P Kurtis, C.R., Trinh, C.H., Gaule, T.G., Knowles, P.F., V Phillips, S.E. and McPherson, M.J. Exploring the Roles of the Metal Ions in Escherichia coli Copper Amine Oxidase. *Biochemistry*, **49**, pp. 1268–1280 (2010).
- [91] Fenalti, G., Giguere, P.M., Katritch, V., Huang, X.P., Thompson, A.A., Cherezov, V., Roth, B.L. and Stevens, R.C. Molecular control of δ -opioid receptor signalling. *Nature*, **506**(7487), pp. 191–196 (2014).
- [92] Tang, W., Strachan, R.T., Lefkowitz, R.J. and Rockman, H.A. Allosteric modulation of β -arrestin-biased angiotensin II type 1 receptor signaling by membrane stretch. *Journal of Biological Chemistry*, **289**(41), pp. 28271–28283 (2014).
- [93] Bohn, L.M., Lefkowitz, R.J., Gainetdinov, R.R., Peppel, K., Caron, M.G. and Lin, F.T. Enhanced morphine analgesia in mice lacking β -arrestin 2. *Science*, **286**(5449), pp. 2495–2498 (1999).
- [94] Bohn, L.M., Gainetdinov, R.R., Lin, F.T., Lefkowitz, R.J. and Caron, M.G. μ -opioid receptor desensitization by β -arrestin-2 determines morphine tolerance but not dependence. *Nature*, **408**(6813), pp. 720–723 (2000).
- [95] Raehal, K.M., Walker, J.K. and Bohn, L.M. Morphine side effects in β -arrestin 2 knockout mice. *Journal of Pharmacology and Experimental Therapeutics*, **314**(3), pp. 1195–1201 (2005).
- [96] Jeong, P., Kim, S.K., Li, Q., jin Oh, S., Son, S., Chen, G., Tan, H., Kim, S., Park, J.H., Park, K.D., Kim, Y.O., Yoon, M.H., Kim, Y.C. and Goddard, W.A. Discovery of Novel Biased Opioid Receptor Ligands through Structure-Based Pharmacophore Virtual Screening and Experiment. *ChemMedChem*, **14**(20), pp. 1783–1794 (2019).
- [97] Gentry, P.R., Sexton, P.M. and Christopoulos, A. Novel allosteric modulators of G protein-coupled receptors. *Journal of Biological Chemistry*, **290**(32), pp. 19478–19488 (2015).

- [98] Van Der Westhuizen, E.T., Valant, C., Sexton, P.M. and Christopoulos, A. Endogenous allosteric modulators of G protein-coupled receptors. *Journal of Pharmacology and Experimental Therapeutics*, **353**(2), pp. 246–260 (2015).
- [99] Congreve, M., Oswald, C. and Marshall, F.H. Applying Structure-Based Drug Design Approaches to Allosteric Modulators of GPCRs. *Trends in Pharmacological Sciences*, **38**(9), pp. 837–847 (2017).
- [100] Ahn, S., Kahsai, A.W., Pani, B., Wang, Q.T., Zhao, S., Wall, A.L., Strachan, R.T., Staus, D.P., Wingler, L.M., Sun, L.D., Sinnaeve, J., Choi, M., Cho, T., Xu, T.T., Hansen, G.M., Burnett, M.B., Lamerdin, J.E., Bassoni, D.L., Gavino, B.J., Husemoen, G., Olsen, E.K., Franch, T., Costanzi, S., Chen, X. and Lefkowitz, R.J. Allosteric "beta-blocker" isolated from a DNA-encoded small molecule library. *Proceedings of the National Academy of Sciences of the United States of America*, **114**(7), pp. 1708–1713 (2017).
- [101] Ahn, S., Pani, B., Kahsai, A.W., Olsen, E.K., Husemoen, G., Vestergaard, M., Jin, L., Zhao, S., Wingler, L.M., Rambarat, P.K., Simhal, R.K., Xu, T.T., Sun, L.D., Shim, P.J., Staus, D.P., Huang, L.Y., Franch, T., Chen, X. and Lefkowitz, R.J. Small-molecule positive allosteric modulators of the β_2 -adrenoceptor isolated from DNA-encoded libraries. *Molecular Pharmacology*, **94**(2), pp. 850–861 (2018).
- [102] Lane, J.R., Chubukov, P., Liu, W., Canals, M., Cherezov, V., Abagyan, R., Stevens, R.C. and Katritch, V. Structure-based ligand discovery targeting orthosteric and allosteric pockets of dopamine receptors. *Molecular Pharmacology*, **84**(6), pp. 794–807 (2013).
- [103] DeGraaf, C., Rein, C., Piwnicka, D., Giordanetto, F. and Rognan, D. Structure-Based Discovery of Allosteric Modulators of Two Related ClassB G-Protein-Coupled Receptors. *ChemMedChem*, **6**(12), pp. 2159–2169 (2011).
- [104] Overington, J.P., Al-Lazikani, B. and Hopkins, A.L. How many drug targets are there? *Nature Reviews Drug Discovery*, **5**(12), pp. 993–996 (2006).
- [105] Palczewski, K., Kumasaka, T., Hori, T., Behnke, C.A., Motoshima, H., Fox, B.A., Le Trong, I., Teller, D.C., Okada, T., Stenkamp, R.E., Yamamoto, M. and Miyano, M. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science*, **289**(5480), pp. 739–745 (2000).
- [106] Cherezov, V., Rosenbaum, D.M., Hanson, M.A., Rasmussen, S.G.F., Thian, F.S., Kobilka, T.S., Choi, H.J., Kuhn, P., Weis, W.I., Kobilka, B.K. and Stevens, R.C. High-resolution crystal structure of an engineered human β_2 -adrenergic G protein-coupled receptor. *Science*, **318**(5854), pp. 1258–65 (2007).
- [107] Rasmussen, S.G.F., Choi, H.J., Rosenbaum, D.M., Kobilka, T.S., Thian, F.S., Edwards, P.C., Burghammer, M., Ratnala, V.R.P., Sanishvili, R., Fischetti, R.F., Schertler, G.F.X., Weis, W.I. and Kobilka, B.K. Crystal structure of the human β_2 adrenergic G-protein-coupled receptor. *Nature*, **450**(7168), pp. 383–387 (2007).

- [108] Gusach, A., Maslov, I., Luginina, A., Borshchevskiy, V., Mishin, A. and Cherezov, V. Beyond structure: emerging approaches to study GPCR dynamics. *Current Opinion in Structural Biology*, **63**, pp. 18–25 (2020).
- [109] Wingler, L.M., Elgeti, M., Hilger, D., Latorraca, N.R., Lerch, M.T., Staus, D.P., Dror, R.O., Kobilka, B.K., Hubbell, W.L. and Lefkowitz, R.J. Angiotensin Analogs with Divergent Bias Stabilize Distinct Receptor Conformations. *Cell*, **176**(3), pp. 468–478.e11 (2019).
- [110] Manglik, A., Kim, T.H., Masureel, M., Altenbach, C., Yang, Z., Hilger, D., Lerch, M.T., Kobilka, T.S., Thian, F.S., Hubbell, W.L., Prosser, R.S. and Kobilka, B.K. Structural insights into the dynamic process of β 2-adrenergic receptor signaling. *Cell*, **161**(5), pp. 1101–1111 (2015).
- [111] Torrens-Fontanals, M., Stepniewski, T.M., Aranda-García, D., Morales-Pastor, A., Medel-Lacruz, B. and Selent, J. How do molecular dynamics data complement static structural data of GPCRs. *International Journal of Molecular Sciences*, **21**(16), pp. 1–26 (2020).
- [112] Latorraca, N.R., Venkatakrisnan, A.J. and Dror, R.O. GPCR dynamics: Structures in motion. *Chemical Reviews*, **117** (1)(1), pp. 139–155 (2017).
- [113] Hilger, D. The role of structural dynamics in GPCR-mediated signaling. *FEBS Journal*, **288**(8), pp. 2461–2489 (2021).
- [114] Ye, L., Van Eps, N., Zimmer, M., Ernst, O.P. and Scott Prosser, R. Activation of the A2A adenosine G-protein-coupled receptor by conformational selection. *Nature*, **533**(7602), pp. 265–268 (2016).
- [115] Park, J.H., Scheerer, P., Hofmann, K.P., Choe, H.W. and Ernst, O.P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature*, **454**(7201), pp. 183–187 (2008).
- [116] Huang, J., Chen, S., Zhang, J.J. and Huang, X.Y. Crystal structure of oligomeric β 1-adrenergic G protein-coupled receptors in ligand-free basal state. *Nature Structural and Molecular Biology*, **20**(4), pp. 419–425 (2013).
- [117] Shihoya, W., Nishizawa, T., Okuta, A., Tani, K., Dohmae, N., Fujiyoshi, Y., Nureki, O. and Doi, T. Activation mechanism of endothelin ET B receptor by endothelin-1. *Nature*, **537**(7620), pp. 363–368 (2016).
- [118] Josephs, T.M., Belousoff, M.J., Liang, Y.L., Piper, S.J., Cao, J., Garama, D.J., Leach, K., Gregory, K.J., Christopoulos, A., Hay, D.L., Danev, R., Wootten, D. and Sexton, P.M. Structure and dynamics of the CGRP receptor in apo and peptide-bound forms. *Science*, **372**(6538) (2021).
- [119] Zhang, J., Qu, L., Wu, L., Tang, X., Luo, F., Xu, W., Xu, Y., Liu, Z.J. and Hua, T. Structural insights into the activation initiation of full-length mGlu1. *Protein and Cell*, **12**(8), pp. 662–667 (2021).

- [120] Du, J., Wang, D., Fan, H., Xu, C., Tai, L., Lin, S., Han, S., Tan, Q., Wang, X., Xu, T., Zhang, H., Chu, X., Yi, C., Liu, P., Wang, X., Zhou, Y., Pin, J.P., Rondard, P., Liu, H., Liu, J., Sun, F., Wu, B. and Zhao, Q. Structures of human mGlu2 and mGlu7 homo- and heterodimers. *Nature*, **594**(7864), pp. 589–593 (2021).
- [121] Tsutsumi, N., Qu, Q., Mavri, M., Baggesen, M.S., Maeda, S., Waghray, D., Berg, C., Kobilka, B.K., Rosenkilde, M.M., Skiniotis, G. and Garcia, K.C. Structural basis for the constitutive activity and immunomodulatory properties of the Epstein-Barr virus-encoded G protein-coupled receptor BILF1. *Immunity*, **54**(7), pp. 1405–1416.e7 (2021).
- [122] Xu, L., Chen, B., Schihada, H., Wright, S.C., Turku, A., Wu, Y., Han, G.W., Kowalski-Jahn, M., Koziulewicz, P., Bowin, C.F., Zhang, X., Li, C., Bouvier, M., Schulte, G. and Xu, F. Cryo-EM structure of constitutively active human Frizzled 7 in complex with heterotrimeric Gs. *Cell Research*, **31**(12), pp. 1311–1314 (2021).
- [123] Xiang, J., Chun, E., Liu, C., Jing, L., Al-Sahouri, Z., Zhu, L. and Liu, W. Successful Strategies to Determine High-Resolution Structures of GPCRs. *Trends in Pharmacological Sciences*, **37**(12), pp. 1055–1069 (2016).
- [124] Hua, T., Vemuri, K., Pu, M., Qu, L., Han, G.W., Wu, Y., Zhao, S., Shui, W., Li, S., Korde, A., Laprairie, R.B., Stahl, E.L., Ho, J.H., Zvonok, N., Zhou, H., Kufareva, I., Wu, B., Zhao, Q., Hanson, M.A., Bohn, L.M., Makriyannis, A., Stevens, R.C. and Liu, Z.J. Crystal Structure of the Human Cannabinoid Receptor CB1. *Cell*, **167**(3), pp. 750–762.e14 (2016).
- [125] Lebon, G., Bennett, K., Jazayeri, A. and Tate, C.G. Thermostabilisation of an agonist-bound conformation of the human adenosine A2A receptor. *Journal of Molecular Biology*, **409**(3), pp. 298–310 (2011).
- [126] Ring, A.M., Manglik, A., Kruse, A.C., Enos, M.D., Weis, W.I., Garcia, K.C. and Kobilka, B.K. Adrenaline-activated structure of β 2-adrenoceptor stabilized by an engineered nanobody. *Nature*, **502**(7472), pp. 575–579 (2013).
- [127] Parker, E.M., Kameyamas, K., Higashijima, T. and Ross, E.M. Reconstitutively Active G Protein-coupled Receptors Purified from Baculovirus-infected Insect Cells. *The Journal of Biological Chemistry*, **266**(1), pp. 519–527 (1991).
- [128] Warne, T., Serrano-Vega, M.J., Tate, C.G. and Schertler, G.F. Development and crystallization of a minimal thermostabilised G protein-coupled receptor. *Protein Expression and Purification*, **65**(2), pp. 204–213 (2009).
- [129] Xie, G., Gross, A.K. and Oprian, D.D. An opsin mutant with increased thermal stability. *Biochemistry*, **42**(7), pp. 1995–2001 (2003).
- [130] Miller, J.L. and Tate, C.G. Engineering an ultra-thermostable β 1-adrenoceptor. *Journal of Molecular Biology*, **413**(3), pp. 628–638 (2011).
- [131] Roth, C.B., Hanson, M.A. and Stevens, R.C. Stabilization of the Human β 2-Adrenergic Receptor TM4-TM3-TM5 Helix Interface by Mutagenesis of Glu1223.41, A Critical Residue in GPCR Structure. *Journal of Molecular Biology*, **376**(5), pp. 1305–1319 (2008).

- [132] Robertson, N., Jazayeri, A., Errey, J., Baig, A., Hurrell, E., Zhukov, A., Langmead, C.J., Weir, M. and Marshall, F.H. The properties of thermostabilised G protein-coupled receptors (StaRs) and their use in drug discovery. *Neuropharmacology*, **60**(1), pp. 36–44 (2011).
- [133] Culhane, K.J., Liu, Y., Cai, Y. and Yan, E.C.Y. Transmembrane signal transduction by peptide hormones via family B G protein-coupled receptors. *Frontiers in Pharmacology*, **6**(NOV), pp. 1–23 (2015).
- [134] Serna, M. Hands on methods for high resolution cryo-electron microscopy structures of heterogeneous macromolecular complexes. *Frontiers in Molecular Biosciences*, **6**(MAY), p. 33 (2019).
- [135] Basak, S., Gicheru, Y., Rao, S., Sansom, M.S. and Chakrapani, S. Cryo-EM reveals two distinct serotonin-bound conformations of full-length 5-HT_{3A} receptor. *Nature*, **563**(7730), pp. 270–274 (2018).
- [136] Chun, E., Thompson, A.A., Liu, W., Roth, C.B., Griffith, M.T., Katritch, V., Kunken, J., Xu, F., Cherezov, V., Hanson, M.A. and Stevens, R.C. Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure*, **20**(6), pp. 967–976 (2012).
- [137] Zou, Y., Weis, W.I. and Kobilka, B.K. N-Terminal T4 Lysozyme Fusion Facilitates Crystallization of a G Protein Coupled Receptor. *PLoS ONE*, **7**(10), p. e46039 (2012).
- [138] Hollenstein, K., Kean, J., Bortolato, A., Cheng, R.K.Y., Doré, A.S., Jazayeri, A., Cooke, R.M., Weir, M. and Marshall, F.H. Structure of class B GPCR corticotropin-releasing factor receptor 1. *Nature*, **499**(7459), pp. 438–443 (2013).
- [139] Weinert, T., Olieric, N., Cheng, R., Brünle, S., James, D., Ozerov, D., Gashi, D., Vera, L., Marsh, M., Jaeger, K., Dworkowski, F., Panepucci, E., Basu, S., Skopintsev, P., Doré, A.S., Geng, T., Cooke, R.M., Liang, M., Prota, A.E., Panneels, V., Nogly, P., Ermler, U., Schertler, G., Hennig, M., Steinmetz, M.O., Wang, M. and Standfuss, J. Serial millisecond crystallography for routine room-temperature structure determination at synchrotrons. *Nature Communications*, **8**(1), pp. 1–11 (2017).
- [140] Segala, E., Guo, D., Cheng, R.K.Y., Bortolato, A., Deflorian, F., Doré, A.S., Errey, J.C., Heitman, L.H., Ijzerman, A.P., Marshall, F.H., Cooke, R.M., Doré, A.S., Errey, J.C., Heitman, L.H., Ijzerman, A.P., Marshall, F.H. and Cooke, R.M. Controlling the Dissociation of Ligands from the Adenosine A_{2A} Receptor through Modulation of Salt Bridge Strength. *Journal of Medicinal Chemistry*, **59**(13), pp. 6470–6479 (2016).
- [141] Rasmussen, S.G.F., DeVree, B.T., Zou, Y., Kruse, A.C., Chung, K.Y., Kobilka, T.S., Thian, F.S., Chae, P.S., Pardon, E., Calinski, D., Mathiesen, J.M., Shah, S.T.A., Lyons, J.A., Caffrey, M., Gellman, S.H., Steyaert, J., Skiniotis, G., Weis, W.I., Sunahara, R.K. and Kobilka, B.K. Crystal structure of the β ₂ adrenergic receptor–Gs protein complex. *Nature*, **477**(7366), pp. 549–555 (2011).

- [142] Kang, Y., Zhou, X.E., Gao, X., He, Y., Liu, W., Ishchenko, A., Barty, A., White, T.A., Yefanov, O., Han, G.W., Xu, Q., de Waal, P.W., Ke, J., Tan, M.H.E., Zhang, C., Moeller, A., West, G.M., Pascal, B.D., Van Eps, N., Caro, L.N., Vishnivetskiy, S.A., Lee, R.J., Suino-Powell, K.M., Gu, X., Pal, K., Ma, J., Zhi, X., Boutet, S., Williams, G.J., Messerschmidt, M., Gati, C., Zatsepin, N.A., Wang, D., James, D., Basu, S., Roy-Chowdhury, S., Conrad, C.E., Coe, J., Liu, H., Lisova, S., Kupitz, C., Grotjohann, I., Fromme, R., Jiang, Y., Tan, M., Yang, H., Li, J., Wang, M., Zheng, Z., Li, D., Howe, N., Zhao, Y., Standfuss, J., Diederichs, K., Dong, Y., Potter, C.S., Carragher, B., Caffrey, M., Jiang, H., Chapman, H.N., Spence, J.C.H., Fromme, P., Weierstall, U., Ernst, O.P., Katritch, V., Gurevich, V.V., Griffin, P.R., Hubbell, W.L., Stevens, R.C., Cherezov, V., Melcher, K. and Xu, H.E. Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. *Nature*, **523**(7562), pp. 561–567 (2015).
- [143] Shukla, A.K., Westfield, G.H., Xiao, K., Reis, R.I., Huang, L.Y., Tripathi-Shukla, P., Qian, J., Li, S., Blanc, A., Oleskie, A.N., Dosey, A.M., Su, M., Liang, C.R., Gu, L.L., Shan, J.M., Chen, X., Hanna, R., Choi, M., Yao, X.J., Klink, B.U., Kahsai, A.W., Sidhu, S.S., Koide, S., Penczek, P.A., Kossiakoff, A.A., Woods Jr, V.L., Kobilka, B.K., Skiniotis, G. and Lefkowitz, R.J. Visualization of arrestin recruitment by a G-protein-coupled receptor. *Nature*, **512**(7513), pp. 218–222 (2014).
- [144] Mukherjee, S., Erramilli, S.K., Ammirati, M., Alvarez, F.J., Fennell, K.F., Purdy, M.D., Skrobek, B.M., Radziwon, K., Coukos, J., Kang, Y., Dutka, P., Gao, X., Qiu, X., Yeager, M., Eric Xu, H., Han, S. and Kossiakoff, A.A. Synthetic antibodies against BRIL as universal fiducial marks for single-particle cryoEM structure determination of membrane proteins. *Nature Communications*, **11**(1), pp. 1–14 (2020).
- [145] Coscia, F., Estrozi, L.F., Hans, F., Malet, H., Noirclerc-Savoye, M., Schoehn, G. and Petosa, C. Fusion to a homo-oligomeric scaffold allows cryo-EM analysis of a small protein. *Scientific reports*, **6**(August), p. 30909 (2016).
- [146] Domanska, K., Vanderhaegen, S., Srinivasan, V., Pardon, E., Dupeux, F., Marquez, J.A., Giorgetti, S., Stoppini, M., Wyns, L., Bellotti, V. and Steyaert, J. Atomic structure of a nanobody-trapped domain-swapped dimer of an amyloidogenic β 2-microglobulin variant. *Proceedings of the National Academy of Sciences of the United States of America*, **108**(4), pp. 1314–1319 (2011).
- [147] Kors, C.A., Wallace, E., Davies, D.R., Li, L., Laible, P.D. and Nollert, P. Effects of impurities on membrane-protein crystallization in different systems. *Acta Crystallographica Section D: Biological Crystallography*, **65**(10), pp. 1062–1073 (2009).
- [148] Ai, X. and Caffrey, M. Membrane protein crystallization in lipidic mesophases: detergent effects. *Biophysical Journal*, **79**(1), pp. 394–405 (2000).
- [149] Chapman, H.N., Fromme, P., Barty, A., White, T.A., Kirian, R.A., Aquila, A., Hunter, M.S., Schulz, J., DePonte, D.P., Weierstall, U., Doak, R.B., Maia, F.R.N.C., Martin, A.V., Schlichting, I., Lomb, L., Coppola, N., Shoeman, R.L., Epp, S.W., Hartmann, R., Rolles, D., Rudenko, A., Foucar, L., Kimmel, N., Weidenspointner, G., Holl, P., Liang, M.,

- Barthelmeß, M., Caleman, C., Boutet, S., Bogan, M.J., Krzywinski, J., Bostedt, C., Bajt, S., Gumprecht, L., Rudek, B., Erk, B., Schmidt, C., Hömke, A., Reich, C., Pietschner, D., Strüder, L., Hauser, G., Gorke, H., Ullrich, J., Herrmann, S., Schaller, G., Schopper, F., Soltau, H., Kühnel, K.U., Messerschmidt, M., Bozek, J.D., Hau-Riege, S.P., Frank, M., Hampton, C.Y., Sierra, R.G., Starodub, D., Williams, G.J., Hajdu, J., Timneanu, N., Seibert, M.M., Andreasson, J., Rucker, A., Jönsson, O., Svenda, M., Stern, S., Nass, K., Andrichke, R., Schröter, C.D., Krasniqi, F., Bott, M., Schmidt, K.E., Wang, X., Grotjohann, I., Holton, J.M., Barends, T.R.M., Neutze, R., Marchesini, S., Fromme, R., Schorb, S., Rupp, D., Adolph, M., Gorkhover, T., Andersson, I., Hirsemann, H., Potdevin, G., Graafsma, H., Nilsson, B. and Spence, J.C.H. Femtosecond X-ray protein nanocrystallography. *Nature*, **470**(7332), pp. 73–77 (2011).
- [150] Liu, W., Wacker, D., Gati, C., Han, G.W., James, D., Wang, D., Nelson, G., Weierstall, U., Katritch, V., Barty, A., Zatsepin, N.A., Li, D., Messerschmidt, M., Boutet, S., Williams, G.J., Koglin, J.E., Seibert, M.M., Wang, C., Shah, S.T.T., Basu, S., Fromme, R., Kupitz, C., Rendek, K.N., Grotjohann, I., Fromme, P., Kirian, R.A., Beyerlein, K.R., White, T.A., Chapman, H.N., Caffrey, M., Spence, J.C.C., Stevens, R.C. and Cherezov, V. Serial femtosecond crystallography of G protein-coupled receptors. *Science*, **342**(6165), pp. 1521–1524 (2013).
- [151] Zhou, X.E., Gao, X., Barty, A., Kang, Y., He, Y., Liu, W., Ishchenko, A., White, T.A., Yefanov, O., Han, G.W., Xu, Q., de Waal, P.W., Suino-Powell, K.M., Boutet, S., Williams, G.J., Wang, M., Li, D., Caffrey, M., Chapman, H.N., Spence, J.C., Fromme, P., Weierstall, U., Stevens, R.C., Cherezov, V., Melcher, K. and Xu, H.E. X-ray laser diffraction for structure determination of the rhodopsin-arrestin complex. *Scientific Data*, **3**, p. 160021 (2016).
- [152] Weierstall, U., James, D., Wang, C., White, T.A., Wang, D., Liu, W., Spence, J.C., Bruce Doak, R., Nelson, G., Fromme, P., Fromme, R., Grotjohann, I., Kupitz, C., Zatsepin, N.A., Liu, H., Basu, S., Wacker, D., Won Han, G., Katritch, V., Boutet, S., Messerschmidt, M., Williams, G.J., Koglin, J.E., Marvin Seibert, M., Klinker, M., Gati, C., Shoeman, R.L., Barty, A., Chapman, H.N., Kirian, R.A., Beyerlein, K.R., Stevens, R.C., Li, D., Shah, S.T., Howe, N., Caffrey, M. and Cherezov, V. Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nature Communications*, **5**, p. 3309 (2014).
- [153] Martynowycz, M.W., Shiriaeva, A., Ge, X., Hattne, J., Nannenga, B.L., Cherezov, V. and Gonen, T. MicroED structure of the human adenosine receptor determined from a single nanocrystal in LCP. *Proceedings of the National Academy of Sciences of the United States of America*, **118**(36) (2021).
- [154] Choy, B.C., Cater, R.J., Mancina, F. and Pryor, E.E. A 10-year meta-analysis of membrane protein structural biology: Detergents, membrane mimetics, and structure determination techniques. *Biochimica et Biophysica Acta - Biomembranes*, **1863**(3), p. 183533 (2021).

- [155] Wu, M. and Lander, G.C. How low can we go? Structure determination of small biological complexes using single-particle cryo-EM. *Current Opinion in Structural Biology*, **64**, pp. 9–16 (2020).
- [156] Nygaard, R., Kim, J. and Mancina, F. Cryo-electron microscopy analysis of small membrane proteins. *Current Opinion in Structural Biology*, **64**, pp. 26–33 (2020).
- [157] Velazhahan, V., Ma, N., Pándy-Szekeres, G., Kooistra, A.J., Lee, Y., Gloriam, D.E., Vaidehi, N. and Tate, C.G. Structure of the class D GPCR Ste2 dimer coupled to two G proteins. *Nature*, **589**(7840), pp. 148–153 (2021).
- [158] García-Nafriá, J. and Tate, C.G. Cryo-EM structures of GPCRs coupled to G_s, G_i and G_o. *Molecular and Cellular Endocrinology*, **488**, pp. 1–13 (2019).
- [159] Staus, D.P., Hu, H., Robertson, M.J., Kleinhenz, A.L., Wingler, L.M., Capel, W.D., Latorraca, N.R., Lefkowitz, R.J. and Skiniotis, G. Structure of the M2 muscarinic receptor– β -arrestin complex in a lipid nanodisc. *Nature*, **579**(7798), pp. 297–302 (2020).
- [160] Mobbs, J.I., Belousoff, M.J., Harikumar, K.G., Piper, S.J., Xu, X., Furness, S.G., Venugopal, H., Christopoulos, A., Danev, R., Wootten, D., Thal, D.M., Miller, L.J. and Sexton, P.M. Structures of the human cholecystokinin 1 (CCK1) receptor bound to G_s and G_q mimetic proteins provide insight into mechanisms of G protein selectivity. *PLoS Biology*, **19**(6), p. e3001295 (2021).
- [161] Tsutsumi, N., Mukherjee, S., Waghray, D., Janda, C.Y., Jude, K.M., Miao, Y., Burg, J.S., Aduri, N.G., Kossiakoff, A.A., Gati, C. and Garcia, K.C. Structure of human frizzled5 by fiducial-assisted cryo-EM supports a heterodimeric mechanism of canonical wnt signaling. *eLife*, **9**, pp. 1–20 (2020).
- [162] Bloch, J.S., Mukherjee, S., Kowal, J., Filippova, E.V., Niederer, M., Pardon, E., Steyaert, J., Kossiakoff, A.A. and Locher, K.P. Development of a universal nanobody-binding Fab module for fiducial-assisted cryo-EM studies of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America*, **118**(47) (2021).
- [163] Layland, J., Carrick, D., Lee, M., Oldroyd, K. and Berry, C. Adenosine: Physiology, pharmacology, and clinical applications. *JACC: Cardiovascular Interventions*, **7**(6), pp. 581–591 (2014).
- [164] Zimmermann, H. Extracellular metabolism of ATP and other nucleotides. *Naunyn-Schmiedeberg's Archives of Pharmacology*, **362**(4-5), pp. 299–309 (2000).
- [165] Cekic, C. and Linden, J. Purinergic regulation of the immune system. *Nature Reviews Immunology*, **16**(3), pp. 177–192 (2016).
- [166] Nazario, L.R., da Silva, R.S. and Bonan, C.D. Targeting adenosine signaling in Parkinson's disease: From pharmacological to non-pharmacological approaches. *Frontiers in Neuroscience*, **11**(NOV), p. 658 (2017).

- [167] Cellai, L., Carvalho, K., Faivre, E., Deleau, A., Vieau, D., Buée, L., Blum, D., Mériaux, C. and Gomez-Murcia, V. The adenosinergic signaling: A complex but promising therapeutic target for Alzheimer's disease. *Frontiers in Neuroscience*, **12**, p. 520 (2018).
- [168] Stockwell, J., Jakova, E. and Cayabyab, F.S. Adenosine A1 and A2A receptors in the brain: Current research and their role in neurodegeneration. *Molecules*, **22**(4) (2017).
- [169] Effendi, W.I., Nagano, T., Kobayashi, K. and Nishimura, Y. Focusing on Adenosine Receptors as a Potential Targeted Therapy in Human Diseases. *Cells*, **9**(3), p. 785 (2020).
- [170] Antonioli, L., Blandizzi, C., Csóka, B., Pacher, P. and Haskó, G. Adenosine signalling in diabetes mellitus-pathophysiology and therapeutic considerations. *Nature Reviews Endocrinology*, **11**(4), pp. 228–241 (2015).
- [171] Leone, R.D. and Emens, L.A. Targeting adenosine for cancer immunotherapy. *Journal for ImmunoTherapy of Cancer*, **6**(1), pp. 1–9 (2018).
- [172] Fredholm, B.B., IJzerman, A.P., Jacobson, K.A., Linden, J. and Müller, C.E. International union of basic and clinical pharmacology. LXXXI. Nomenclature and classification of adenosine receptors - An update. *Pharmacological Reviews*, **63**(1), pp. 1–34 (2011).
- [173] Haskó, G., Csóka, B., Németh, Z.H., Vizi, E.S. and Pacher, P. A2B adenosine receptors in immunity and inflammation. *Trends in Immunology*, **30**(6), pp. 263–270 (2009).
- [174] Baraldi, P.G., Preti, D., Borea, P.A. and Varani, K. Medicinal chemistry of A3 adenosine receptor modulators: Pharmacological activities and therapeutic implications. *Journal of Medicinal Chemistry*, **55**(12), pp. 5676–5703 (2012).
- [175] Fredholm, B.B., Arslan, G., Halldner, L., Kull, B., Schulte, G. and Wasserman, W. Structure and function of adenosine receptors and their genes. *Naunyn-Schmiedeberg's Archives of Pharmacology*, **362**(4-5), pp. 364–374 (2000).
- [176] Dunwiddie, T.V. and Masino, S.A. The role and regulation of adenosine in the central nervous system. *Annual Review of Neuroscience*, **24**, pp. 31–55 (2001).
- [177] Fredholm, B.B., Chern, Y., Franco, R. and Sitkovsky, M. Aspects of the general biology of adenosine A2A signaling. *Progress in Neurobiology*, **83**(5), pp. 263–276 (2007).
- [178] Salvatore, C.A., Jacobson, M.A., Taylor, H.E., Linden, J. and Johnson, R.G. Molecular cloning and characterization of the human A3 adenosine receptor. *Proceedings of the National Academy of Sciences of the United States of America*, **90**(21), pp. 10365–10369 (1993).
- [179] Gessi, S., Merighi, S., Fazzi, D., Stefanelli, A., Varani, K. and Borea, P.A. Adenosine receptor targeting in health and disease. *Expert Opinion on Investigational Drugs*, **20**(12), pp. 1591–1609 (2011).

- [180] Mustafa, S.J., Morrison, R.R., Teng, B. and Pelleg, A. Adenosine receptors and the heart: Role in regulation of coronary blood flow and cardiac electrophysiology. In C. Wilson and S. Mustafa, eds., *Adenosine Receptors in Health and Disease*, vol. 193 of *Handbook of Experimental Pharmacology*, chap. Adenosine, pp. 161–188. Springer Verlag, Berlin, Heidelberg (2009).
- [181] Chen, J.F. and Cunha, R.A. The belated US FDA approval of the adenosine A2A receptor antagonist istradefylline for treatment of Parkinson’s disease. *Purinergic Signalling*, **16**(2), pp. 167–174 (2020).
- [182] Trevethick, M.A., Mantell, S.J., Stuart, E.F., Barnard, A., Wright, K.N. and Yeadon, M. Treating lung inflammation with agonists of the adenosine A2A receptor: Promises, problems and potential solutions. *British Journal of Pharmacology*, **155**(4), pp. 463–474 (2008).
- [183] Pollack, A.E. and Fink, J.S. Adenosine antagonists potentiate D2 dopamine-dependent activation of Fos in the striatopallidal pathway. *Neuroscience*, **68**(3), pp. 721–728 (1995).
- [184] Diaz-Cabiale, Z., Hurd, Y., Guidolin, D., Finnman, U.b., Zoli, M., Agnati, L.F., Vanderhaeghen, J.j., Fuxe, K. and Ferré, S. Adenosine A2A agonist CGS 21680 decreases the affinity of dopamine D2 receptors for dopamine in human striatum. *Neuroreport*, **12**(9), pp. 1831–1834 (2001).
- [185] Dungo, R. and Deeks, E.D. Istradefylline: First global approval. *Drugs*, **73**(8), pp. 875–882 (2013).
- [186] Pires, V.A., Pamplona, F.A., Pandolfo, P., Fernandes, D., Prediger, R.D. and Takahashi, R.N. Adenosine receptor antagonists improve short-term object-recognition ability of spontaneously hypertensive rats: A rodent model of attention-deficit hyperactivity disorder. *Behavioural Pharmacology*, **20**(2), pp. 134–145 (2009).
- [187] Sawynok, J. Adenosine receptor targets for pain. *Neuroscience*, **338**, pp. 1–18 (2016).
- [188] Jaakola, V.P., Griffith, M.T., Hanson, M.A., Cherezov, V., Ellen, Y.T., Lane, J.R., Ijzerman, A.P. and Stevens, R.C. The 2.6 Å Crystal Structure of a Human A2A Adenosine Receptor Bound to an Antagonist. *Science*, **322**(5905), pp. 1211–1217 (2008).
- [189] Doré, A.S., Robertson, N., Errey, J.C., Ng, I., Hollenstein, K., Tehan, B., Hurrell, E., Bennett, K., Congreve, M., Magnani, F., Tate, C.G., Weir, M. and Marshall, F.H. Structure of the Adenosine A2A Receptor in Complex with ZM241385 and the Xanthines XAC and Caffeine. *Structure*, **19**(9), pp. 1283–1293 (2011).
- [190] Xu, D. and Zhang, Y. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophysical Journal*, **101**(10), pp. 2525–2534 (2011).
- [191] Carpenter, B. and Tate, C.G. Engineering a minimal G protein to facilitate crystallisation of G protein-coupled receptors in their active conformation. *Protein Engineering Design and Selection*, **29**(12), pp. 583–593 (2016).

- [192] Isberg, V., De Graaf, C., Bortolato, A., Cherezov, V., Katritch, V., Marshall, F.H., Mordalski, S., Pin, J.P., Stevens, R.C., Vriend, G. and Gloriam, D.E. Generic GPCR residue numbers - Aligning topology maps while minding the gaps. *Trends in Pharmacological Sciences*, **36**(1), pp. 22–31 (2015).
- [193] Sun, B., Bachhawat, P., Chu, M.L.H., Wood, M., Ceska, T., Sands, Z.A., Mercier, J., Lebon, F., Kobilka, T.S. and Kobilka, B.K. Crystal structure of the adenosine A2A receptor bound to an antagonist reveals a potential allosteric pocket. *Proceedings of the National Academy of Sciences of the United States of America*, **114**(8), pp. 2066–2071 (2017).
- [194] IJzerman, A.P., Von Frijtag Drabbe Künzel, J.K., Kim, J., Jiang, Q. and Jacobson, K.A. Site-directed mutagenesis of the human adenosine A2A receptor. Critical involvement of Glu13 in agonist recognition. *European Journal of Pharmacology*, **310**(2-3), pp. 269–272 (1996).
- [195] Jiang, Q., Van Rhee, A.M., Kim, J., Yehle, S., Wess, J. and Jacobson, K.A. Hydrophilic side chains in the third and seventh transmembrane helical domains of human A2A adenosine receptors are required for ligand recognition. *Molecular Pharmacology*, **50**(3), pp. 512–521 (1996).
- [196] Massink, A., Gutiérrez-De-Terán, H., Lenselink, E.B., Zacarías, N.V., Xia, L., Heitman, L.H., Katritch, V., Stevens, R.C. and IJzerman, A.P. Sodium ion binding pocket mutations and adenosine A2A receptor function. *Molecular Pharmacology*, **87**(2), pp. 305–313 (2015).
- [197] Isberg, V., Mordalski, S., Munk, C., Rataj, K., Harpsøe, K., Hauser, A.S., Vroiling, B., Bojarski, A.J., Vriend, G. and Gloriam, D.E. GPCRdb: an information system for G protein-coupled receptors. *Nucleic Acids Research*, **44**(D1), pp. D356–D364 (2016).
- [198] Lebon, G., Edwards, P.C., Leslie, A.G. and Tate, C.G. Molecular determinants of CGS21680 binding to the human adenosine A2A receptor. *Molecular Pharmacology*, **87**(6), pp. 907–915 (2015).
- [199] Rodríguez, D., Piñeiro, Á. and Gutiérrez-De-Terán, H. Molecular dynamics simulations reveal insights into key structural elements of adenosine receptors. *Biochemistry*, **50**(19), pp. 4194–4208 (2011).
- [200] Seibt, B.F., Schiedel, A.C., Thimm, D., Hinz, S., Sherbiny, F.F. and Müller, C.E. The second extracellular loop of GPCRs determines subtype-selectivity and controls efficacy as evidenced by loop exchange study at A2A adenosine receptors. *Biochemical Pharmacology*, **85**(9), pp. 1317–1329 (2013).
- [201] Jespers, W., Schiedel, A.C., Heitman, L.H., Cooke, R.M., Kleene, L., van Westen, G.J., Gloriam, D.E., Müller, C.E., Sotelo, E. and Gutiérrez-de Terán, H. Structural Mapping of Adenosine Receptor Mutations: Ligand Binding and Signaling Mechanisms. *Trends in Pharmacological Sciences*, **39**(1), pp. 75–89 (2018).

- [202] Tucker, A.L., Robeva, A.S., Taylor, H.E., Holeton, D., Bockner, M., Lynch, K.R. and Linden, J. A1 adenosine receptors. Two amino acids are responsible for species differences in ligand recognition. *Journal of Biological Chemistry*, **269**(45), pp. 27900–27906 (1994).
- [203] Schiedel, A.C., Hinz, S., Thimm, D., Sherbiny, F., Borrmann, T., Maaß, A. and Müller, C.E. The four cysteine residues in the second extracellular loop of the human adenosine A2B receptor: Role in ligand binding and receptor function. *Biochemical Pharmacology*, **82**(4), pp. 389–399 (2011).
- [204] Knight, A., Hemmings, J.L., Winfield, I., Leuenberger, M., Frattini, E., Frenguelli, B.G., Dowell, S.J., Lochner, M. and Ladds, G. Discovery of Novel Adenosine Receptor Agonists That Exhibit Subtype Selectivity. *Journal of Medicinal Chemistry*, **59**(3), pp. 947–964 (2016).
- [205] De Filippo, E., Namasivayam, V., Zappe, L., El-Tayeb, A., Schiedel, A.C. and Müller, C.E. Role of extracellular cysteine residues in the adenosine A2A receptor. *Purinergic Signalling*, **12**(2), pp. 313–329 (2016).
- [206] Wang, Y., Xu, H., Wang, H., Zheng, Z., Meng, Z., Xu, Z., Li, J. and Xue, M. Design, Synthesis, and Biological Activity Studies of Istradefylline Derivatives Based on Adenine as A2A Receptor Antagonists. *ACS Omega*, **6**(6), pp. 4386–4394 (2021).
- [207] Borodovsky, A., Barbon, C.M., Wang, Y., Ye, M., Prickett, L., Chandra, D., Shaw, J., Deng, N., Sachsenmeier, K., Clarke, J.D., Linghu, B., Brown, G.A., Brown, J., Congreve, M., Cheng, R.K., Dore, A.S., Hurrell, E., Shao, W., Woessner, R., Reimer, C., Drew, L., Fawell, S., Schuller, A.G. and Mele, D.A. Small molecule AZD4635 inhibitor of A2AR signaling rescues immune cell function including CD103+ dendritic cells enhancing anti-tumor immunity. *Journal for immunotherapy of cancer*, **8**(2) (2020).
- [208] Jespers, W., Verdon, G., Azuaje, J., Majellaro, M., Keränen, H., García-Mera, X., Congreve, M., Deflorian, F., de Graaf, C., Zhukov, A., Doré, A.S., Mason, J.S., Åqvist, J., Cooke, R.M., Sotelo, E., Gutiérrez-de Terán, H., García-Mera, X., Congreve, M., Deflorian, F., Graaf, C., Zhukov, A., Doré, A.S., Mason, J.S., Åqvist, J., Cooke, R.M., Sotelo, E. and Gutiérrez-de-Terán, H. X-Ray Crystallography and Free Energy Calculations Reveal the Binding Mechanism of A2A Adenosine Receptor Antagonists. *Angewandte Chemie - International Edition*, **59**(38), pp. 16536–16543 (2020).
- [209] Katritch, V., Jaakola, V.P., Lane, J.R., Lin, J., Ijzerman, A.P., Yeager, M., Kufareva, I., Stevens, R.C. and Abagyan, R. Structure-based discovery of novel chemotypes for adenosine A2A receptor antagonists. *Journal of Medicinal Chemistry*, **53**(4), pp. 1799–1809 (2010).
- [210] Carlsson, J., Yoo, L., Gao, Z.G., Irwin, J.J., Shoichet, B.K. and Jacobson, K.A. Structure-based discovery of A2A adenosine receptor ligands. *Journal of Medicinal Chemistry*, **53**(9), pp. 3748–3755 (2010).

- [211] Langmead, C.J., Andrews, S.P., Congreve, M., Errey, J.C., Hurrell, E., Marshall, F.H., Mason, J.S., Richardson, C.M., Robertson, N., Zhukov, A. and Weir, M. Identification of novel adenosine A2A receptor antagonists by virtual screening. *Journal of Medicinal Chemistry*, **55**(5), pp. 1904–1909 (2012).
- [212] Deflorian, F., Kumar, T.S., Phan, K., Gao, Z.G., Xu, F., Wu, H., Katritch, V., Stevens, R.C. and Jacobson, K.A. Evaluation of molecular modeling of agonist binding in light of the crystallographic structure of an agonist-bound A2A adenosine receptor. *Journal of Medicinal Chemistry*, **55**(1), pp. 538–552 (2012).
- [213] Keränen, H., Åqvist, J. and Gutiérrez-De-Terán, H. Free energy calculations of A2A adenosine receptor mutation effects on agonist binding. *Chemical Communications*, **51**(17), pp. 3522–3525 (2015).
- [214] Tosh, D.K., Phan, K., Gao, Z.G., Gakh, A.A., Xu, F., Deflorian, F., Abagyan, R., Stevens, R.C., Jacobson, K.A. and Katritch, V. Optimization of adenosine 5'-carboxamide derivatives as adenosine receptor agonists using structure-based ligand design and fragment screening. *Journal of Medicinal Chemistry*, **55**(9), pp. 4297–4308 (2012).
- [215] Sabbadin, D., Ciancetta, A., Deganutti, G., Cuzzolin, A. and Moro, S. Exploring the recognition pathway at the human A2A adenosine receptor of the endogenous agonist adenosine using supervised molecular dynamics simulations. *MedChemComm*, **6**(6), pp. 1081–1085 (2015).
- [216] Zhukov, A., Andrews, S.P., Errey, J.C., Robertson, N., Tehan, B., Mason, J.S., Marshall, F.H., Weir, M. and Congreve, M. Biophysical mapping of the adenosine A2A receptor. *Journal of Medicinal Chemistry*, **54**(13), pp. 4312–4323 (2011).
- [217] Congreve, M., Andrews, S.P., Doré, A.S., Hollenstein, K., Hurrell, E., Langmead, C.J., Mason, J.S., Ng, I.W., Tehan, B., Zhukov, A., Weir, M. and Marshall, F.H. Discovery of 1,2,4-triazine derivatives as adenosine A2A antagonists using structure based drug design. *Journal of Medicinal Chemistry*, **55**(5), pp. 1898–1903 (2012).
- [218] Congreve, M., Brown, G.A., Borodovsky, A. and Lamb, M.L. Targeting adenosine A2A receptor antagonism for treatment of cancer. *Expert Opinion on Drug Discovery*, **13**(11), pp. 997–1003 (2018).
- [219] Carlsson, J., Tosh, D.K., Phan, K., Gao, Z.G. and Jacobson, K.A. Structure-activity relationships and molecular modeling of 1,2,4-triazoles as adenosine receptor antagonists. *ACS Medicinal Chemistry Letters*, **3**(9), pp. 715–720 (2012).
- [220] Mason, J.S., Bortolato, A., Congreve, M. and Marshall, F.H. New insights from structural biology into the druggability of G protein-coupled receptors. *Trends in Pharmacological Sciences*, **33**(5), pp. 249–260 (2012).
- [221] Cheng, R.K., Segala, E., Robertson, N., Deflorian, F., Doré, A.S., Errey, J.C., Fiez-Vandal, C., Marshall, F.H. and Cooke, R.M. Structures of Human A1 and A2A Adenosine Receptors with Xanthines Reveal Determinants of Selectivity. *Structure*, **25**(8), pp. 1275–1285.e4 (2017).

- [222] Winter, G., Waterman, D.G., Parkhurst, J.M., Brewster, A.S., Gildea, R.J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I.D., Sauter, N.K. and Evans, G. DIALS: Implementation and evaluation of a new integration package. *Acta Crystallographica Section D: Structural Biology*, **74**, pp. 85–97 (2018).
- [223] Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography*, **66**(4), pp. 486–501 (2010).
- [224] Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H. and Adams, P.D. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallographica Section D: Biological Crystallography*, **68**(4), pp. 352–367 (2012).
- [225] Zhang, D., Zhao, Q. and Wu, B. Structural Studies of G Protein-Coupled Receptors. *Molecules and cells*, **38**(10), pp. 836–42 (2015).
- [226] Worth, C.L., Kreuchwig, F., Tiemann, J.K., Kreuchwig, A., Ritschel, M., Kleinau, G., Hildebrand, P.W. and Krause, G. GPCR-SSFE 2.0 - A fragment-based molecular modeling web tool for Class A G-protein coupled receptors. *Nucleic Acids Research*, **45**(W1), pp. W408–W415 (2017).
- [227] Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., De Beer, T.A., Rempfer, C., Bordoli, L., Lepore, R. and Schwede, T. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, **46**(W1), pp. W296–W303 (2018).
- [228] Pieper, U., Webb, B.M., Dong, G.Q., Schneidman-Duhovny, D., Fan, H., Kim, S.J., Khuri, N., Spill, Y.G., Weinkam, P., Hammel, M., Tainer, J.A., Nilges, M. and Sali, A. ModBase, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Research*, **42**(D1) (2014).
- [229] Friesner, R.A., Murphy, R.B., Repasky, M.P., Frye, L.L., Greenwood, J.R., Halgren, T.A., Sanschagrin, P.C. and Mainz, D.T. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of Medicinal Chemistry*, **49**(21), pp. 6177–6196 (2006).
- [230] Bell, E.W. and Zhang, Y. DockRMSD: An open-source tool for atom mapping and RMSD calculation of symmetric molecules through graph isomorphism. *Journal of Cheminformatics*, **11**(1), pp. 1–9 (2019).
- [231] Ferenczy, G.G. and Keser, G.M. The impact of binding thermodynamics on medicinal chemistry optimizations. *Future Medicinal Chemistry*, **7**(10), pp. 1285–1303 (2015).
- [232] Kairys, V., Baranauskiene, L., Kazlauskiene, M., Matulis, D. and Kazlauskas, E. Binding affinity in drug design: experimental and computational techniques. *Expert Opinion on Drug Discovery*, **14**(8), pp. 755–768 (2019).

- [233] Stoddart, L.A., White, C.W., Nguyen, K., Hill, S.J. and Pflieger, K.D. Fluorescence- and bioluminescence-based approaches to study GPCR ligand binding. *British Journal of Pharmacology*, **173**(20), pp. 3028–3037 (2016).
- [234] Flanagan, C.A. Chapter 10 - GPCR-radioligand binding assays. In A. K. Shukla, ed., *G Protein-Coupled Receptors*, vol. 132 of *Methods in Cell Biology*, chap. GPCR-radio, pp. 191–215. Academic Press (2016).
- [235] Paton, W.D.M. and Rang, H.P. The uptake of atropine and related drugs by intestinal smooth muscle of the guinea-pig in relation to acetylcholine receptors. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, **163** (1965).
- [236] Bylund, D.B. and Toews, M.L. Radioligand binding methods practical guide and tips. *American Journal of Physiology - Lung Cellular and Molecular Physiology*, **265**(5 9-5) (1993).
- [237] Dong, C., Liu, Z. and Wang, F. Radioligand saturation binding for quantitative analysis of ligand-receptor interactions. *Biophysics Reports*, **1**(3), pp. 148–155 (2015).
- [238] Hulme, E.C. and Trevethick, M.A. Ligand binding assays at equilibrium: Validation and interpretation. *British Journal of Pharmacology*, **161**(6), pp. 1219–1237 (2010).
- [239] Maguire, J.J., Kuc, R.E. and Davenport, A.P. Radioligand binding assays and their analysis. *Methods in Molecular Biology*, **897**, pp. 31–77 (2012).
- [240] Alexander, S.P. and Millns, P.J. [³H]ZM241385 - An antagonist radioligand for adenosine A_{2A} receptors in rat brain. *European Journal of Pharmacology*, **411**(3), pp. 205–210 (2001).
- [241] Longo, P.A., Kavran, J.M., Kim, M.S. and Leahy, D.J. Transient mammalian cell transfection with polyethylenimine (PEI). In *Methods in Enzymology*, vol. 529, pp. 227–240. NIH Public Access (2013).
- [242] Balakin, K., Savchuk, N. and Tetko, I. In Silico Approaches to Prediction of Aqueous and DMSO Solubility of Drug-Like Compounds: Trends, Problems and Solutions. *Current Medicinal Chemistry*, **13**(2), pp. 223–241 (2005).
- [243] Yung-Chi, C. and Prusoff, W.H. Relationship between the inhibition constant (K_i) and the concentration of inhibitor which causes 50 per cent inhibition (I₅₀) of an enzymatic reaction. *Biochemical Pharmacology*, **22**(23), pp. 3099–3108 (1973).
- [244] Varani, K., Gessi, S., Merighi, S., Vincenzi, F., Cattabriga, E., Benini, A., Klotz, K.N., Baraldi, P.G., Tabrizi, M.A., Mac Lennan, S., Leung, E. and Borea, P.A. Pharmacological characterization of novel adenosine ligands in recombinant and native human A_{2B} receptors. *Biochemical Pharmacology*, **70**(11), pp. 1601–1612 (2005).
- [245] Kull, B., Arslan, G., Nilsson, C., Owman, C., Lorenzen, A., Schwabe, U. and Fredholm, B.B. Differences in the order of potency for agonists but not antagonists at human and rat adenosine A_{2A} receptors. *Biochemical Pharmacology*, **57**(1), pp. 65–75 (1999).

- [246] Wang, J., Guo, Z., Fu, Y., Wu, Z., Huang, C., Zheng, C., Shar, P.A., Wang, Z., , Xiao, W. and Wang, Y. Weak-binding molecules are not drugs?-toward a systematic strategy for finding effective weak-binding drugs. *Briefings in Bioinformatics*, **18**(2), pp. 321–332 (2017).
- [247] Cooper, J., Hill, S.J. and Alexander, S.P. An endogenous A(2B) adenosine receptor coupled to cyclic AMP generation in human embryonic kidney (HEK 293) cells. *British Journal of Pharmacology*, **122**(3), pp. 546–550 (1997).
- [248] Senisterra, G., Chau, I. and Vedadi, M. Thermal denaturation assays in chemical biology. *Assay and Drug Development Technologies*, **10**(2), pp. 128–136 (2012).
- [249] Vedadi, M., Niesen, F.H., Allali-Hassani, A., Fedorov, O.Y., Finerty, P.J., Wasney, G.A., Yeung, R., Arrowsmith, C., Ball, L.J., Berglund, H., Hui, R., Marsden, B.D., Nordlund, P., Sundstrom, M., Weigelt, J. and Edwards, A.M. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proceedings of the National Academy of Sciences*, **103**(43), pp. 15835–15840 (2006).
- [250] Hattori, M., Hibbs, R.E. and Gouaux, E. A fluorescence-detection size-exclusion chromatography-based thermostability assay for membrane protein precrystallization screening. *Structure*, **20**(8), pp. 1293–1299 (2012).
- [251] Alexandrov, A.I., Mileni, M., Chien, E.Y., Hanson, M.A. and Stevens, R.C. Microscale Fluorescent Thermal Stability Assay for Membrane Proteins. *Structure*, **16**(3), pp. 351–359 (2008).
- [252] Strutz, W. Exploring Protein Stability by NanoDSF. *Biophysical Journal*, **110**(3), p. 393a (2016).
- [253] Ashok, Y., Nanekar, R. and Jaakola, V.P. Defining thermostability of membrane proteins by western blotting. *Protein Engineering Design and Selection*, **28**(12), pp. 539–542 (2015).
- [254] Duy, C. and Fitter, J. How aggregation and conformational scrambling of unfolded states govern fluorescence emission spectra. *Biophysical Journal*, **90**(10), pp. 3704–3711 (2006).
- [255] Zahid, N.I., Abou-Zied, O.K., Hashim, R. and Heidelberg, T. Fluorescence probing of the temperature-induced phase transition in a Glycolipid self-assembly: Hexagonal \leftrightarrow micellar and cubic \leftrightarrow lamellar. *Langmuir*, **28**(11), pp. 4989–4995 (2012).
- [256] Vivian, J.T. and Callis, P.R. Mechanisms of tryptophan fluorescence shifts in proteins. *Biophysical Journal*, **80**(5), pp. 2093–2109 (2001).
- [257] Meng, X.Y., Zhang, H.X., Mezei, M. and Cui, M. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Current Computer Aided-Drug Design*, **7**(2), pp. 146–157 (2012).

- [258] McNutt, A.T., Francoeur, P., Aggarwal, R., Masuda, T., Meli, R., Ragoza, M., Sunseri, J. and Koes, D.R. GNINA 1.0: molecular docking with deep learning. *Journal of Cheminformatics*, **13**(1), pp. 1–20 (2021).
- [259] Deng, Z., Chuaqui, C. and Singh, J. Structural Interaction Fingerprint (SIFt): A Novel Method for Analyzing Three-Dimensional Protein-Ligand Binding Interactions. *Journal of Medicinal Chemistry*, **47**(2), pp. 337–344 (2004).
- [260] Singh, J., Deng, Z., Narale, G. and Chuaqui, C. Structural interaction fingerprints: A new approach to organizing, mining, analyzing, and designing protein-small molecule complexes. *Chemical Biology and Drug Design*, **67**(1), pp. 5–12 (2006).
- [261] Guasch, L., Sala, E., Valls, C., Mulero, M., Pujadas, G. and Garcia-Vallvé, S. Development of docking-based 3D-QSAR models for PPAR γ full agonists. *Journal of Molecular Graphics and Modelling*, **36**, pp. 1–9 (2012).
- [262] Muñoz-Gutiérrez, C., Cáceres-Rojas, D., Adasme-Carreño, F., Palomo, I., Fuentes, E. and Caballero, J. Docking and quantitative structure–activity relationship of bi-cyclic heteroaromatic pyridazinone and pyrazolone derivatives as phosphodiesterase 3A (PDE3A) inhibitors. *PLoS ONE*, **12**(12), p. e0189213 (2017).
- [263] Velázquez-Libera, J.L., Rossino, G., Navarro-Retamal, C., Collina, S. and Caballero, J. Docking, Interaction Fingerprint, and Three-Dimensional Quantitative Structure-Activity Relationship (3D-QSAR) of Sigma1 Receptor Ligands, Analogs of the Neuroprotective Agent RC-33. *Frontiers in Chemistry*, **7**, p. 496 (2019).
- [264] Kim, R. and Skolnick, J. Assessment of programs for ligand binding affinity prediction. *Journal of Computational Chemistry*, **29**(8), pp. 1316–1331 (2008).
- [265] Koes, D.R., Baumgartner, M.P. and Camacho, C.J. Lessons learned in empirical scoring with smina from the CSAR 2011 benchmarking exercise. *Journal of Chemical Information and Modeling*, **53**(8), pp. 1893–1904 (2013).
- [266] Wierbowski, S.D., Wingert, B.M., Zheng, J. and Camacho, C.J. Cross-docking benchmark for automated pose and ranking prediction of ligand binding. *Protein Science*, **29**(1), pp. 298–305 (2020).
- [267] Eldridge, M.D., Murray, C.W., Auton, T.R., Paolini, G.V. and Mee, R.P. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *Journal of Computer-Aided Molecular Design*, **11**(5), pp. 425–445 (1997).
- [268] Mantri, M., De Graaf, O., Van Veldhoven, J., Göblyös, A., Von Frijtag Drabbe Künzel, J.K., Mulder-Krieger, T., Link, R., De Vries, H., Beukers, M.W., Brussee, J. and Ijzerman, A.P. 2-Amino-6-furan-2-yl-4-substituted nicotinonitriles as A2A adenosine receptor antagonists. *Journal of Medicinal Chemistry*, **51**(15), pp. 4449–4455 (2008).
- [269] Hagemans, D., van Belzen, I.A., Luengo, T.M. and Rüdiger, S.G. A script to highlight hydrophobicity and charge on protein surfaces. *Frontiers in Molecular Biosciences*, **2**(OCT), p. 56 (2015).

- [270] Kim, J., Wess, J., Van Rhee, A.M., Schoneberg, T. and Jacobson, K.A. Site-directed mutagenesis identifies residues involved in ligand recognition in the human A2A adenosine receptor. *Journal of Biological Chemistry*, **270**(23), pp. 13987–13997 (1995).
- [271] Jaakola, V.P., Lane, J.R., Lin, J.Y., Katritch, V., Ijzerman, A.P. and Stevens, R.C. Ligand binding and subtype selectivity of the human A2A adenosine receptor: identification and characterization of essential amino acid residues. *The Journal of biological chemistry*, **285**(17), pp. 13032–44 (2010).
- [272] Deganutti, G., Zhukov, A., Deflorian, F., Federico, S., Spalluto, G., Cooke, R.M., Moro, S., Mason, J.S. and Bortolato, A. Impact of protein–ligand solvation and desolvation on transition state thermodynamic properties of adenosine A2A ligand binding kinetics. *In Silico Pharmacology*, **5**(1), p. 16 (2017).
- [273] Wright, D.J., Simmons, K.J., Johnson, R.M., Beech, D.J., Muench, S.P. and Bon, R.S. Human TRPC5 structures reveal interaction of a xanthine-based TRPC1/4/5 inhibitor with a conserved lipid binding site. *Communications Biology*, **3**(1), pp. 1–11 (2020).
- [274] Poucher, S.M., Keddie, J.R., Singh, P., Stoggall, S.M., Caulkett, P.W., Jones, G. and Collis, M.G. The in vitro pharmacology of ZM 241385, a potent, non-xanthine, A2A selective adenosine receptor antagonist. *British Journal of Pharmacology*, **115**(6), pp. 1096–1102 (1995).
- [275] Ashton, T.D., Aumann, K.M., Baker, S.P., Schiesser, C.H. and Scammells, P.J. Structure-activity relationships of adenosines with heterocyclic N6-substituents. *Bioorganic and Medicinal Chemistry Letters*, **17**(24), pp. 6779–6784 (2007).
- [276] Petrelli, R., Torquati, I., Kachler, S., Luongo, L., Maione, S., Franchetti, P., Grifantini, M., Novellino, E., Lavecchia, A., Klotz, K.N. and Cappellacci, L. 5'-C-ethyl-tetrazolyl-N 6-substituted adenosine and 2-chloro-adenosine derivatives as highly potent dual acting A1 adenosine receptor agonists and A3 adenosine receptor antagonists. *Journal of Medicinal Chemistry*, **58**(5), pp. 2560–2566 (2015).
- [277] Dhanik, A., McMurray, J.S. and Kavraki, L.E. DINC: A new AutoDock-based protocol for docking large ligands. *BMC Structural Biology*, **13**(SUPPL.1), pp. 1–14 (2013).
- [278] Devaurs, D., Antunes, D.A., Hall-Swan, S., Mitchell, N., Moll, M., Lizée, G. and Kavraki, L.E. Using parallelized incremental meta-docking can solve the conformational sampling issue when docking large ligands to proteins. *BMC Molecular and Cell Biology*, **20**(1), pp. 1–15 (2019).
- [279] Thal, D.M., Glukhova, A., Sexton, P.M. and Christopoulos, A. Structural insights into G-protein-coupled receptor allostery. *Nature*, **559**(7712), pp. 45–53 (2018).
- [280] Korkutata, M., Agrawal, L. and Lazarus, M. Allosteric Modulation of Adenosine A2A Receptors as a New Therapeutic Avenue. *International Journal of Molecular Sciences*, **23**(4), p. 2101 (2022).

- [281] Giorgi, I., Biagi, G., Bianucci, A.M., Borghini, A., Livi, O., Leonardi, M., Pietra, D., Calderone, V. and Martelli, A. N6-1,3-Diphenylurea derivatives of 2-phenyl-9-benzyladenines and 8-azaadenines: Synthesis and biological evaluation as allosteric modulators of A2A adenosine receptors. *European Journal of Medicinal Chemistry*, **43**(8), pp. 1639–1647 (2008).
- [282] Göblyös, A. and IJzerman, A.P. Allosteric modulation of adenosine receptors. *Purinergic Signalling*, **5**(1), pp. 51–61 (2009).
- [283] Wall, M.J., Hill, E., Huckstepp, R., Barkan, K., Deganutti, G., Leuenberger, M., Preti, B., Winfield, I., Wei, H., Imlach, W., Dean, E., Hume, C., Hayward, S., Oliver, J., Zhao, F.Y., Spanswick, D., Reynolds, C.A., Lochner, M., Ladds, G. and Frenguelli, B.G. A biased adenosine A1R agonist elicits analgesia without cardiorespiratory depression. *bioRxiv*, p. 2020.04.04.023945 (2020).
- [284] Marino, K.A. and Filizola, M. Investigating small-molecule ligand binding to G protein-coupled receptors with biased or unbiased molecular dynamics simulations. In *Methods in Molecular Biology*, vol. 1705, pp. 351–364. NIH Public Access (2018).
- [285] Harmalkar, A. and Gray, J.J. Advances to tackle backbone flexibility in protein docking. *Current Opinion in Structural Biology*, **67**, pp. 178–186 (2021).
- [286] Salmaso, V. and Moro, S. Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: An overview. *Frontiers in Pharmacology*, **9**(AUG), p. 923 (2018).
- [287] Araki, M., Matsumoto, S., Bekker, G.J., Isaka, Y., Sagae, Y., Kamiya, N. and Okuno, Y. Exploring ligand binding pathways on proteins using hypersound-accelerated molecular dynamics. *Nature Communications*, **12**(1), pp. 1–10 (2021).
- [288] Caliman, A.D., Miao, Y. and McCammon, J.A. Mapping the allosteric sites of the A2A adenosine receptor. *Chemical Biology and Drug Design*, **91**(1), pp. 5–16 (2018).
- [289] Bruzzese, A., Dalton, J.A. and Giraldo, J. Insights into adenosine A2A receptor activation through cooperative modulation of agonist and allosteric lipid interactions. *PLoS Computational Biology*, **16**(4), p. e1007818 (2020).
- [290] Song, W., Yen, H.Y., Robinson, C.V. and Sansom, M.S. State-dependent Lipid Interactions with the A2A Receptor Revealed by MD Simulations Using In Vivo-Mimetic Membranes. *Structure*, **27**(2), pp. 392–403.e3 (2019).
- [291] Payandeh, J. and Volgraf, M. Ligand binding at the protein–lipid interface: strategic considerations for drug design (2021).
- [292] Ottosson, N.E., Ejneby, M.S., Wu, X., Yazdi, S., Konradsson, P., Lindahl, E. and Elinder, F. A drug pocket at the lipid bilayer–potassium channel interface. *Science Advances*, **3**(10) (2017).
- [293] Lee, J.Y. and Lyman, E. Predictions for cholesterol interaction sites on the A2A adenosine receptor. *Journal of the American Chemical Society*, **134**(40), pp. 16512–16515 (2012).

- [294] Rouviere, E., Arnarez, C., Yang, L. and Lyman, E. Identification of Two New Cholesterol Interaction Sites on the A2A Adenosine Receptor. *Biophysical Journal*, **113**(11), pp. 2415–2424 (2017).
- [295] Charalambous, C., Gsandtner, I., Keuerleber, S., Milan-Lobo, L., Kudlacek, O., Freissmuth, M. and Zezula, J. Restricted collision coupling of the A2A receptor revisited: Evidence for physical separation of two signaling cascades. *Journal of Biological Chemistry*, **283**(14), pp. 9276–9288 (2008).
- [296] McGraw, C., Yang, L., Levental, I., Lyman, E. and Robinson, A.S. Membrane cholesterol depletion reduces downstream signaling activity of the adenosine A2A receptor. *Biochimica et Biophysica Acta - Biomembranes*, **1861**(4), pp. 760–767 (2019).
- [297] Huang, S.K., Almurad, O., Pejana, R.J., Morrison, Z.A., Pandey, A., Picard, L.P., Nitz, M., Sljoka, A. and Prosser, R.S. Allosteric modulation of the adenosine A2A receptor by cholesterol. *eLife*, **11**, pp. 1–24 (2022).
- [298] Glukhova, A., Thal, D.M., Nguyen, A.T., Vecchio, E.A., Jörg, M., Scammells, P.J., May, L.T., Sexton, P.M. and Christopoulos, A. Structure of the Adenosine A1 Receptor Reveals the Basis for Subtype Selectivity. *Cell*, **168**(5), pp. 867–877.e13 (2017).
- [299] Anderson, A.C., O’Neil, R.H., Surti, T.S. and Stroud, R.M. Approaches to solving the rigid receptor problem by identifying a minimal set of flexible residues during ligand docking. *Chemistry and Biology*, **8**(5), pp. 445–457 (2001).
- [300] Savitsky, P., Bray, J., Cooper, C.D., Marsden, B.D., Mahajan, P., Burgess-Brown, N.A. and Gileadi, O. High-throughput production of human proteins for crystallization: The SGC experience. *Journal of Structural Biology*, **172**(1), pp. 3–13 (2010).
- [301] Ishchenko, A., Stauch, B., Han, G.W., Batyuk, A., Shiriaeva, A., Li, C., Zatsepin, N., Weierstall, U., Liu, W., Nango, E., Nakane, T., Tanaka, R., Tono, K., Joti, Y., Iwata, S., Moraes, I., Gati, C. and Cherezov, V. Toward G protein-coupled receptor structure-based drug design using X-ray lasers. *IUCrJ*, **6**, pp. 1106–1119 (2019).
- [302] Beukers, M.W., Chang, L.C.W., Mulder-krieger, T., Spanjersberg, R.F., Brussee, J. and Ijzerman, A.P. New, Non-Adenosine, High-Potency Agonists for the Human Adenosine A2B Receptor with an Improved Selectivity Profile Compared to the Reference Agonist N-Ethylcarboxamidoadenosine. *Journal of Medicinal Chemistry*, **47**(15), pp. 8–10 (2004).
- [303] García-Nafria, J. and Tate, C.G. Structure determination of GPCRs: Cryo-EM compared with X-ray crystallography. *Biochemical Society Transactions*, **49**(5), pp. 2345–2355 (2021).
- [304] Danev, R., Belousoff, M., Liang, Y.L., Zhang, X., Eisenstein, F., Wootten, D. and Sexton, P.M. Routine sub-2.5 Å cryo-EM structure determination of GPCRs. *Nature Communications*, **12**(1), pp. 1–10 (2021).
- [305] Bucher, M.H., Evdokimov, A.G. and Waugh, D.S. Differential effects of short affinity tags on the crystallization of *Pyrococcus furiosus* maltodextrin-binding protein. *Acta Crystallographica Section D: Biological Crystallography*, **58**(3), pp. 392–397 (2002).

- [306] Hassell, A.M., An, G., Bledsoe, R.K., Bynum, J.M., Carter, H.L., Deng, S.J.J., Gampe, R.T., Grisard, T.E., Madauss, K.P., Nolte, R.T., Rocque, W.J., Wang, L., Weaver, K.L., Williams, S.P., Wisely, G.B., Xu, R. and Shewchuk, L.M. Crystallization of protein-ligand complexes. In *Acta Crystallographica Section D: Biological Crystallography*, vol. 63, pp. 72–79. International Union of Crystallography (2006).
- [307] Müller, I. Guidelines for the successful generation of protein-ligand complex crystals. *Acta Crystallographica Section D: Structural Biology*, **73**(2), pp. 79–92 (2017).
- [308] Amelia, T., Van Veldhoven, J.P., Falsini, M., Liu, R., Heitman, L.H., Van Westen, G.J., Segala, E., Verdon, G., Cheng, R.K., Cooke, R.M., Van Der Es, D. and Ijzerman, A.P. Crystal Structure and Subsequent Ligand Design of a Nonriboside Partial Agonist Bound to the Adenosine A2A Receptor. *Journal of Medicinal Chemistry*, **64**(7), pp. 3827–3842 (2021).
- [309] Ishchenko, A., Wacker, D., Kapoor, M., Zhang, A., Han, G.W., Basu, S., Patel, N., Messerschmidt, M., Weierstall, U., Liu, W., Katritch, V., Roth, B.L., Stevens, R.C. and Cherezov, V. Structural insights into the extracellular recognition of the human serotonin 2B receptor by an antibody. *Proceedings of the National Academy of Sciences of the United States of America*, **114**(31), pp. 8223–8228 (2017).
- [310] Sejdiu, B.I. and Tieleman, D.P. Lipid-Protein Interactions Are a Unique Property and Defining Feature of G Protein-Coupled Receptors. *Biophysical Journal*, **118**(8), pp. 1887–1900 (2020).
- [311] Glaenger, J., Peter, M.F. and Hagelueken, G. Studying structure and function of membrane proteins with PELDOR/DEER spectroscopy – The crystallographers’ perspective. *Methods*, **147**, pp. 163–175 (2018).
- [312] Peter, M.F., Bountra, K., Beis, K. and Hagelueken, G. PELDOR/DEER: An Electron Paramagnetic Resonance Method to Study Membrane Proteins in Lipid Bilayers. In *Methods in Molecular Biology*, vol. 2168, pp. 313–333. Methods Mol Biol (2020).
- [313] Pliotas, C., Ward, R., Branigan, E., Rasmussen, A., Hagelueken, G., Huang, H., Black, S.S., Booth, I.R., Schiemann, O. and Naismith, J.H. Conformational state of the MscS mechanosensitive channel in solution revealed by pulsed electron-electron double resonance (PELDOR) spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(40) (2012).
- [314] Elgeti, M. and Hubbell, W.L. DEER Analysis of GPCR Conformational Heterogeneity. *Biomolecules*, **11**(6), p. 778 (2021).
- [315] Sykes, D.A., Stoddart, L.A., Kilpatrick, L.E. and Hill, S.J. Binding kinetics of ligands acting at GPCRs. *Molecular and Cellular Endocrinology*, **485**(January), pp. 9–19 (2019).
- [316] Copeland, R.A., Pompliano, D.L. and Meek, T.D. Drug-target residence time and its implications for lead optimization. *Nature Reviews Drug Discovery*, **5**(9), pp. 730–739 (2006).

- [317] Finlay, D.B., Duffull, S.B. and Glass, M. 100 years of modelling ligand–receptor binding and response: A focus on GPCRs. *British Journal of Pharmacology*, **177**(7), pp. 1472–1484 (2020).
- [318] Salahpour, A., Espinoza, S., Masri, B., Lam, V., Barak, L.S. and Gainetdinov, R.R. BRET biosensors to study GPCR biology, pharmacology, and signal transduction. *Frontiers in Endocrinology*, **3**(AUG), p. 105 (2012).
- [319] Galés, C., Rebois, R.V., Hogue, M., Trieu, P., Breit, A., Hébert, T.E. and Bouvier, M. Real-time monitoring of receptor and G-protein interactions in living cells. *Nature Methods*, **2**(3), pp. 177–184 (2005).
- [320] Ayoub, M.A., Damian, M., Gespach, C., Ferrandis, E., Lavergne, O., De Wever, O., Banères, J.L., Pin, J.P. and Prévost, G.P. Inhibition of heterotrimeric G protein signaling by a small molecule acting on G α subunit. *Journal of Biological Chemistry*, **284**(42), pp. 29136–29145 (2009).
- [321] Barak, L.S., Ferguson, S.S., Zhang, J. and Caron, M.G. A β -arrestin/green fluorescent protein biosensor for detecting G protein-coupled receptor activation. *Journal of Biological Chemistry*, **272**(44), pp. 27497–27500 (1997).
- [322] Draper-Joyce, C.J., Bhola, R., Wang, J., Bhattarai, A., Nguyen, A.T., Cowie-Kent, I., O’Sullivan, K., Chia, L.Y., Venugopal, H., Valant, C., Thal, D.M., Wootten, D., Panel, N., Carlsson, J., Christie, M.J., White, P.J., Scammells, P., May, L.T., Sexton, P.M., Danev, R., Miao, Y., Glukhova, A., Imlach, W.L. and Christopoulos, A. Positive allosteric mechanisms of adenosine A1 receptor-mediated analgesia. *Nature*, **597**(7877), pp. 571–576 (2021).
- [323] Valant, C., May, L.T., Aurelio, L., Chuo, C.H., White, P.J., Baltos, J.A., Sexton, P.M., Scammells, P.J. and Christopoulos, A. Separation of on-target efficacy from adverse effects through rational design of a bitopic adenosine receptor agonist. *Proceedings of the National Academy of Sciences of the United States of America*, **111**(12), pp. 4614–4619 (2014).
- [324] Neutze, R., Wouts, R., Van Der Spoel, D., Weckert, E. and Hajdu, J. Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature*, **406**(6797), pp. 752–757 (2000).
- [325] Martin-Garcia, J.M., Conrad, C.E., Nelson, G., Stander, N., Zatsepin, N.A., Zook, J., Zhu, L., Geiger, J., Chun, E., Kissick, D., Hilgart, M.C., Ogata, C., Ishchenko, A., Nagarathnam, N., Roy-Chowdhury, S., Coe, J., Subramanian, G., Schaffer, A., James, D., Ketwala, G., Venugopalan, N., Xu, S., Corcoran, S., Ferguson, D., Weierstall, U., Spence, J.C.H., Cherezov, V., Fromme, P., Fischetti, R.F. and Liu, W. Serial millisecond crystallography of membrane and soluble protein microcrystals using synchrotron radiation. *IUCrJ*, **4**(Pt 4), pp. 439–454 (2017).

- [326] Suga, M., Akita, F., Hirata, K., Ueno, G., Murakami, H., Nakajima, Y., Shimizu, T., Yamashita, K., Yamamoto, M., Ago, H. and Shen, J.R. Native structure of photosystem II at 1.95Å resolution viewed by femtosecond X-ray pulses. *Nature*, **517**(7532), pp. 99–103 (2015).
- [327] Ebrahim, A., Appleby, M.V., Axford, D., Beale, J., Moreno-Chicano, T., Sherrell, D.A., Strange, R.W., Hough, M.A. and Owen, R.L. Resolving polymorphs and radiation-driven effects in microcrystals using fixed-target serial synchrotron crystallography. *Acta Crystallographica Section D: Structural Biology*, **75**, pp. 151–159 (2019).
- [328] Schotte, F., Lim, M., Jackson, T.A., Smirnov, A.V., Soman, J., Olson, J.S., Phillips, G.N., Wulff, M. and Anfinrud, P.A. Watching a protein as it functions with 150-ps time-resolved x-ray crystallography. *Science*, **300**(5627), pp. 1944–1947 (2003).
- [329] Barends, T.R., Foucar, L., Ardevol, A., Nass, K., Aquila, A., Botha, S., Doak, R.B., Falahati, K., Hartmann, E., Hilpert, M., Heinz, M., Hoffmann, M.C., Köfner, J., Koglin, J.E., Kovacsova, G., Liang, M., Milathianaki, D., Lemke, H.T., Reinstein, J., Roome, C.M., Shoeman, R.L., Williams, G.J., Burghardt, I., Hummer, G., Boutet, S. and Schlichting, I. Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation. *Science*, **350**(6259), pp. 445–450 (2015).
- [330] Jung, Y.O., Lee, J.H., Kim, J., Schmidt, M., Moffat, K., Šrajer, V. and Ihee, H. Volume-conserving trans-cis isomerization pathways in photoactive yellow protein visualized by picosecond X-ray crystallography. *Nature Chemistry*, **5**(3), pp. 212–220 (2013).
- [331] Tenboer, J., Basu, S., Zatsepin, N., Pande, K., Milathianaki, D., Frank, M., Hunter, M., Boutet, S., Williams, G.J., Koglin, J.E., Oberthuer, D., Heymann, M., Kupitz, C., Conrad, C., Coe, J., Roy-Chowdhury, S., Weierstall, U., James, D., Wang, D., Grant, T., Barty, A., Yefanov, O., Scales, J., Gati, C., Seuring, C., Srajer, V., Henning, R., Schwander, P., Fromme, R., Ourmazd, A., Moffat, K., Van Thor, J.J., Spence, J.C., Fromme, P., Chapman, H.N. and Schmidt, M. Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science*, **346**(6214), pp. 1242–1246 (2014).
- [332] Kupitz, C., Basu, S., Grotjohann, I., Fromme, R., Zatsepin, N.A., Rendek, K.N., Hunter, M.S., Shoeman, R.L., White, T.A., Wang, D., James, D., Yang, J.H., Cobb, D.E., Reeder, B., Sierra, R.G., Liu, H., Barty, A., Aquila, A.L., Deponte, D., Kirian, R.A., Bari, S., Bergkamp, J.J., Beyerlein, K.R., Bogan, M.J., Coleman, C., Chao, T.C., Conrad, C.E., Davis, K.M., Fleckenstein, H., Galli, L., Hau-Riege, S.P., Kassemeyer, S., Laksmono, H., Liang, M., Lomb, L., Marchesini, S., Martin, A.V., Messerschmidt, M., Milathianaki, D., Nass, K., Ros, A., Roy-Chowdhury, S., Schmidt, K., Seibert, M., Steinbrener, J., Stellato, F., Yan, L., Yoon, C., Moore, T.A., Moore, A.L., Pushkar, Y., Williams, G.J., Boutet, S., Doak, R.B., Weierstall, U., Frank, M., Chapman, H.N., Spence, J.C. and Fromme, P. Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser. *Nature*, **513**(7517), pp. 261–265 (2014).

- [333] Monteiro, D.C., Amoah, E., Rogers, C. and Pearson, A.R. Using photocaging for fast time-resolved structural biology studies. *Acta Crystallographica Section D: Structural Biology*, **77**(10), pp. 1218–1232 (2021).
- [334] Lee, H.M., Larson, D.R. and Lawrence, D.S. Illuminating the chemistry of life: Design, synthesis, and applications of "caged" and related photoresponsive compounds. *ACS Chemical Biology*, **4**(6), pp. 409–427 (2009).
- [335] Lang, K. and Chin, J.W. Cellular incorporation of unnatural amino acids and bioorthogonal Labeling of Proteins. *Chemical Reviews*, **114**(9), pp. 4764–4806 (2014).
- [336] Mehrabi, P., Schulz, E.C., Agthe, M., Horrell, S., Bourenkov, G., von Stetten, D., Leimkohl, J.P., Schikora, H., Schneider, T.R., Pearson, A.R., Tellkamp, F. and Miller, R.J. Liquid application method for time-resolved analyses by serial synchrotron crystallography. *Nature Methods*, **16**(10), pp. 979–982 (2019).
- [337] Beyerlein, K.R., Dierksmeyer, D., Mariani, V., Kuhn, M., Sarrou, I., Ottaviano, A., Awel, S., Knoska, J., Fuglerud, S., Jönsson, O., Stern, S., Wiedorn, M.O., Yefanov, O., Adriano, L., Bean, R., Burkhardt, A., Fischer, P., Heymann, M., Horke, D.A., Jungnickel, K.E., Kovaleva, E., Lorbeer, O., Metz, M., Meyer, J., Morgan, A., Pande, K., Panneerselvam, S., Seuring, C., Tolstikova, A., Lieske, J., Aplin, S., Roessle, M., White, T.A., Chapman, H.N., Meents, A. and Oberthuer, D. Mix-and-diffuse serial synchrotron crystallography. *IUCrJ*, **4**(6), pp. 769–777 (2017).
- [338] Olmos, J.L., Pandey, S., Martin-Garcia, J.M., Calvey, G., Katz, A., Knoska, J., Kupitz, C., Hunter, M.S., Liang, M., Oberthuer, D., Yefanov, O., Wiedorn, M., Heyman, M., Holl, M., Pande, K., Barty, A., Miller, M.D., Stern, S., Roy-Chowdhury, S., Coe, J., Nagaratnam, N., Zook, J., Verburgt, J., Norwood, T., Poudyal, I., Xu, D., Koglin, J., Seaberg, M.H., Zhao, Y., Bajt, S., Grant, T., Mariani, V., Nelson, G., Subramanian, G., Bae, E., Fromme, R., Fung, R., Schwander, P., Frank, M., White, T.A., Weierstall, U., Zatsepin, N., Spence, J., Fromme, P., Chapman, H.N., Pollack, L., Tremblay, L., Ourmazd, A., Phillips, G.N. and Schmidt, M. Enzyme intermediates captured "on the fly" by mix-and-inject serial crystallography. *BMC Biology*, **16**(1), p. 59 (2018).
- [339] Levantino, M., Yorke, B.A., Monteiro, D.C., Cammarata, M. and Pearson, A.R. Using synchrotrons and XFELs for time-resolved X-ray crystallography and solution scattering experiments on biomolecules. *Current Opinion in Structural Biology*, **35**, pp. 41–48 (2015).
- [340] Kupitz, C., Grotjohann, I., Conrad, C.E., Roy-Chowdhury, S., Fromme, R. and Fromme, P. Microcrystallization techniques for serial femtosecond crystallography using photosystem II from *Thermosynechococcus elongatus* as a model system. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **369**(1647) (2014).
- [341] Ibrahim, M., Chatterjee, R., Hellmich, J., Tran, R., Bommer, M., Yachandra, V.K., Yano, J., Kern, J. and Zouni, A. Improvements in serial femtosecond crystallography of photosystem II by optimizing crystal uniformity using microseeding procedures. *Structural Dynamics*, **2**(4) (2015).

- [342] White, T.A., Kirian, R.A., Martin, A.V., Aquila, A., Nass, K., Barty, A. and Chapman, H.N. CrystFEL: A software suite for snapshot serial crystallography. *Journal of Applied Crystallography*, **45**(2), pp. 335–341 (2012).
- [343] Schmidt, M. Mix and inject: Reaction initiation by diffusion for time-resolved macromolecular crystallography. *Advances in Condensed Matter Physics*, **2013** (2013).
- [344] Fuller, F.D., Gul, S., Chatterjee, R., Sethe Burgie, E., Young, I.D., Lebrette, H., Srinivas, V., Brewster, A.S., Michels-Clark, T., Clinger, J.A., Andi, B., Ibrahim, M., Pastor, E., De Lichtenberg, C., Hussein, R., Pollock, C.J., Zhang, M., Stan, C.A., Kroll, T., Fransson, T., Weninger, C., Kubin, M., Aller, P., Lassalle, L., Bräuer, P., Miller, M.D., Amin, M., Koroidov, S., Roessler, C.G., Allaire, M., Sierra, R.G., Docker, P.T., Glowonia, J.M., Nelson, S., Koglin, J.E., Zhu, D., Chollet, M., Song, S., Lemke, H., Liang, M., Sokaras, D., Alonso-Mori, R., Zouni, A., Messinger, J., Bergmann, U., Boal, A.K., Martin Bollinger, J., Krebs, C., Högbom, M., Phillips, G.N., Vierstra, R.D., Sauter, N.K., Orville, A.M., Kern, J., Yachandra, V.K. and Yano, J. Drop-on-demand sample delivery for studying biocatalysts in action at X-ray free-electron lasers. *Nature Methods*, **14**(4), pp. 443–449 (2017).
- [345] Lee, D., Baek, S., Park, J., Lee, K., Kim, J., Lee, S.J., Chung, W.K., Lee, J.L., Cho, Y. and Nam, K.H. Nylon mesh-based sample holder for fixed-target serial femtosecond crystallography. *Scientific Reports*, **9**(1) (2019).
- [346] Grünbein, M.L. and Kovacs, G.N. Sample delivery for serial crystallography at free-electron lasers and synchrotrons. *Acta Crystallographica Section D: Structural Biology*, **75**, pp. 178–191 (2019).
- [347] Monteiro, D.C., Vakili, M., Harich, J., Sztucki, M., Meier, S.M., Horrell, S., Josts, I. and Trebbin, M. A microfluidic flow-focusing device for low sample consumption serial synchrotron crystallography experiments in liquid flow. *Journal of Synchrotron Radiation*, **26**(2), pp. 406–412 (2019).
- [348] Mehrabi, P., Muller-Werkmeister, H.M., Leimkohl, J.P., Schikora, H., Ninkovic, J., Kri-vokuca, S., Andricek, L., Epp, S.W., Sherrell, D., Owen, R.L., Pearson, A.R., Tellkamp, F., Schulz, E.C. and Dwayne Miller, R.J. The HARE chip for efficient time-resolved serial synchrotron crystallography. *Journal of Synchrotron Radiation*, **27**, pp. 360–370 (2020).
- [349] Davy, B., Axford, D., Beale, J.H., Butryn, A., Docker, P., Ebrahim, A., Leen, G., Orville, A.M., Owen, R.L. and Aller, P. Reducing sample consumption for serial crystallography using acoustic drop ejection. *Journal of Synchrotron Radiation*, **26**(5), pp. 1820–1825 (2019).
- [350] Beale, J.H., Bolton, R., Marshall, S.A., Beale, E.V., Carr, S.B., Ebrahim, A., Moreno-Chicano, T., Hough, M.A., Worrall, J.A.R., Tews, I. and Owen, R.L. Successful sample preparation for serial crystallography experiments. *Journal of Applied Crystallography*, **52**(6), pp. 1385–1396 (2019).

- [351] Russo Krauss, I., Merlino, A., Vergara, A. and Sica, F. An overview of biological macromolecule crystallization. *International journal of molecular sciences*, **14**(6), pp. 11643–91 (2013).
- [352] Dods, R., Båth, P., Arnlund, D., Beyerlein, K.R., Nelson, G., Liang, M., Harimoorthy, R., Berntsen, P., Malmerberg, E., Johansson, L., Andersson, R., Bosman, R., Carbajo, S., Claesson, E., Conrad, C.E., Dahl, P., Hammarin, G., Hunter, M.S., Li, C., Lisova, S., Milathianaki, D., Robinson, J., Safari, C., Sharma, A., Williams, G., Wickstrand, C., Yefanov, O., Davidsson, J., DePonte, D.P., Barty, A., Brändén, G. and Neutze, R. From Macrocrystals to Microcrystals: A Strategy for Membrane Protein Serial Crystallography. *Structure*, (25), pp. 1461–1468 (2017).
- [353] Rummel, G., Hardmeyer, A., Widmer, C., Chiu, M.L., Nollert, P., Locher, K.P., Pedruzzi, I., Landau, E.M. and Rosenbusch, J.P. Lipidic cubic phases: New matrices for the three-dimensional crystallization of membrane proteins. *Journal of Structural Biology*, **121**(2), pp. 82–91 (1998).
- [354] Nollert, P., Qiu, H., Caffrey, M., Rosenbusch, J.P. and Landau, E.M. Molecular mechanism for the crystallization of bacteriorhodopsin in lipidic cubic phases. In *FEBS Letters*, vol. 504, pp. 179–186. FEBS Lett (2001).
- [355] Caffrey, M. and Cherezov, V. Crystallizing membrane proteins using lipidic mesophases. *Nature Protocols*, **4**(45), pp. 706–731 (2009).
- [356] Lee, C.H. and MacKinnon, R. Activation mechanism of a human SK-calmodulin channel complex elucidated by cryo-EM structures. *Science*, **360**(6388), pp. 508–513 (2018).
- [357] Monteiro, D.C., Patel, V., Bartlett, C.P., Nozaki, S., Grant, T.D., Gowdy, J.A., Thompson, G.S., Kalverda, A.P., Snell, E.H., Niki, H., Pearson, A.R. and Webb, M.E. The structure of the PanD/PanZ protein complex reveals negative feedback regulation of pantothenate biosynthesis by coenzyme a. *Chemistry and Biology*, **22**(4), pp. 492–503 (2015).
- [358] Arnott, Z.L., Nozaki, S., Monteiro, D.C., Morgan, H.E., Pearson, A.R., Niki, H. and Webb, M.E. The Mechanism of Regulation of Pantothenate Biosynthesis by the PanD-PanZ·AcCoA Complex Reveals an Additional Mode of Action for the Antimetabolite N-Pentyl Pantothenamide (N5-Pan). *Biochemistry*, **56**(37), pp. 4931–4939 (2017).
- [359] Rupp, B. Origin and use of crystallization phase diagrams. *Acta Crystallographica Section F: Structural Biology Communications*, **71**(3), pp. 247–260 (2015).
- [360] Hofmeister, F. Zur Lehre von der Wirkung der Salze. *Archiv für Experimentelle Pathologie und Pharmakologie*, **24**(4-5), pp. 247–260 (1888).
- [361] Gilliland, G.L., Tung, M. and Ladner, J. The biological macromolecule crystallization database and NASA protein crystal growth archive. *Journal of Research of the National Institute of Standards and Technology*, **101**(3), pp. 309–320 (1996).
- [362] Gilliland, G.L. A biological macromolecule crystallization database: A basis for a crystallization strategy. *Journal of Crystal Growth*, **90**, pp. 51–59 (1988).