

DISS. ETH NO. 30146

# **Biochemical and structural characterization of GPCR-G protein complexes**

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES

(Dr. sc. ETH Zurich)

presented by

**OLIVER MIGUEL TEJERO**

M. Sc. In Biochemistry, University of Zurich, Switzerland

born on 28.01.1995

accepted on the recommendation of:

Prof. Dr. Frédéric Allain, examiner

Prof. Dr. Gebhard F.X. Schertler, co-examiner

Prof. Dr. Adriano Aguzzi, co-examiner

Dr. Ching-Ju Tsai, co-examiner

2024



# Table of Contents

<b>Acknowledgment</b> .....	<b>vii</b>
<b>Abbreviations</b> .....	<b>viii</b>
<b>Summary</b> .....	<b>xii</b>
<b>Zusammenfassung</b> .....	<b>xv</b>
<b>1 Introduction</b> .....	<b>1</b>
<b>1.1 G protein coupled receptor superfamily</b> .....	<b>1</b>
<b>1.2 Adhesion GPCR family</b> .....	<b>4</b>
1.2.1 Adgrg6 and Adgrd1 .....	8
<b>1.3 Opsins</b> .....	<b>10</b>
1.3.1 Jumping spider rhodopsin-1 .....	11
<b>1.4 Heterotrimeric G protein</b> .....	<b>13</b>
<b>2 Aims of the thesis</b> .....	<b>16</b>
<b>2.1 Experimental strategy</b> .....	<b>17</b>
2.1.1. Adhesion GPCR project .....	17
2.1.2. Jumping spider rhodopsin-1 project .....	18
<b>3 Material and Methods</b> .....	<b>20</b>
<b>3.1 Bacterial cell transformation</b> .....	<b>20</b>
<b>3.2 Bacterial cell expression</b> .....	<b>20</b>
<b>3.3 G<math>\alpha</math>i purification</b> .....	<b>20</b>
<b>3.4 3C protease purification</b> .....	<b>21</b>
<b>3.5 GFP Nanobody purification</b> .....	<b>21</b>
<b>3.6 Mini-G<math>\alpha</math>s 393/399 purification</b> .....	<b>21</b>
<b>3.7 G<math>\beta</math><math>\gamma</math><sub>1</sub> purification from bovine retinas</b> .....	<b>22</b>
<b>3.8 Insect cell lines</b> .....	<b>22</b>
<b>3.9 Baculovirus production</b> .....	<b>22</b>
<b>3.10 Expression in insect cells</b> .....	<b>22</b>
<b>3.11 G<math>\beta</math><math>\gamma</math><sub>2</sub> expression and purification from insect cells</b> .....	<b>23</b>
<b>3.12 Peptide synthesis</b> .....	<b>23</b>
<b>3.13 Adhesion GPCR Construct Design</b> .....	<b>23</b>
<b>3.14 G protein construct design for insect cell expression</b> .....	<b>24</b>
<b>3.15 Expression in HEK293F cells</b> .....	<b>24</b>
<b>3.16 In-gel fluorescence scanning</b> .....	<b>24</b>

3.17 Western blot analysis .....	24
3.18 Detergent screening .....	25
3.19 Expression of Adgrd1-NTF.....	25
3.20 Purification of Adgrd1-NTF.....	25
3.21 Thermal stability measurements Adrgd1-NTF.....	25
3.22 Crystallization Adgrd1-NTF.....	26
3.23 Complex formation and analytical SEC .....	26
3.24 Complex purification Adrgd1-mini Gs heterotrimer .....	26
3.25 Construct design Adgrd1-mini-Gαs fusion protein .....	27
3.26 Purification Adgrd1-mini-Gαs fusion protein .....	27
3.27 Construct design Adgrd1-NTF-CTF fusion protein.....	27
3.28 Expression and purification Adgrd1-NTF-CTF fusion protein .....	27
3.29 Negative stain grid preparation and analysis .....	28
3.30 Cryo-EM grid preparation and data acquisition aGPCR-G protein complex.....	28
3.31 Cryo-EM data processing aGPCR-G protein complex.....	29
3.32 Retinal analogue synthesis .....	29
3.33 JSR1 Expression and Purification .....	29
3.34 G protein heterotrimer expression in insect cells .....	30
3.34.1 hGq, hG11, JSGi <sub>q</sub> and JSGq expression.....	30
3.35 hGq, hG11, JSGi <sub>q</sub> and JSGq heterotrimer purification .....	30
3.36 Analytical SEC experiments.....	31
3.37 Complex formation and SEC for the JSR1-bGt complex .....	31
3.38 Cryo-EM grid preparation and data acquisition JSR1-bGt complex.....	31
3.39 Complex formation and SEC for the JSR1-hGi (-scFv16) complex.....	31
3.40 Cryo-EM grid preparation and data acquisition JSR1-hGi complex .....	32
3.41 Cryo-EM grid preparation and data acquisition JSR1-hGi-scFv16 complex.....	32
3.42 Complex formation and SEC for the JSR1-JSGi <sub>q</sub> complex .....	32
3.43 Cryo-EM grid preparation and data acquisition JSR1-JSGi <sub>q</sub> complex .....	32
3.44 Cryo-EM data processing JSR1-JSGi <sub>q</sub> complex .....	33
3.45 Model building and structure refinement.....	33
3.46 Structure comparison with JSR1-G protein complexes.....	33
3.47 GTPase Glo assay® (Promega, Madison, USA).....	34
<b>4 Results .....</b>	<b>35</b>



<b>4.1 Adhesion GPCR project</b> .....	<b>35</b>
4.1.1 AGPCR-G protein complex .....	35
4.1.1.1 Construct design and expression screening .....	35
4.1.1.2 Detergent screening .....	40
4.1.1.3 Purification of Adgrd1 and Adgrg6 .....	43
4.1.1.4 Co-expression screening .....	47
4.1.1.5 Complex formation .....	48
4.1.1.6 Complex purification and cryo-EM grid preparation .....	49
4.1.1.7 Cryo-EM data collection and analysis .....	52
4.1.1.8 Summary $\alpha$ GPCR-G protein complex .....	55
4.1.2 Adgrd1-N terminal fragment .....	56
4.1.2.1 Expression and purification of the Adgrd1-NTF .....	56
4.1.2.2 Thermal stability assay for Adgrd1 <sup>NTF</sup> .....	58
4.1.2.3 Crystallization of Adgrd1 <sup>NTF</sup> .....	60
4.1.2.4 Summary Adgrd1 <sup>NTF</sup> .....	61
4.1.3 Adgrd1 fusion proteins .....	62
4.1.3.1 Expression and purification of Adgrd1-mini-G $\alpha_s$ fusion protein .....	62
4.1.3.2 Expression and purification of Adgrd1 <sup>GAIN-CTF</sup> -mini-G $\alpha_s$ fusion protein .....	63
4.1.3.3 Negative staining EM analysis of Adgrd1 <sup>GAIN-CTF</sup> -mini-G $\alpha_s$ fusion protein .....	66
4.1.3.4 Summary Adgrd1 fusion proteins .....	67
<b>4.2 Jumping spider rhodopsin-1 project</b> .....	<b>68</b>
4.2.1 Biochemical characterization of JSR1-G protein complexes .....	68
4.2.1.1 JSR1 purification .....	68
4.2.1.2 Human Gi heterotrimer purification .....	70
4.2.1.3 hGq/hG11/JSGiq/JSgq1 heterotrimer purification .....	71
4.2.1.4 GTPase Glo assay JSR1-G proteins .....	72
4.2.1.5 Analytical SEC JSR1-G protein complexes .....	77
4.2.1.6 Preparation of JSR1-G protein complex samples for cryo-EM .....	82
4.2.1.7 Summary of JSR1-G protein complex biochemistry .....	84
4.2.2 Cryo-EM grid preparation and data analysis .....	85
4.2.2.1 Grid preparation and data collection .....	85
4.2.2.2 Cryo-EM data analysis of the JSR1-bGt complex .....	86
4.2.2.3 Cryo-EM data analysis of the JSR1-hGi complexes .....	87
4.2.2.4 Cryo-EM data analysis of the JSR1-JSGiq complex .....	90
4.2.2.5 Summary cryo-EM analysis of JSR1-G protein complexes .....	94

4.2.3 JSR1-JSGiq structure analysis .....	95
4.2.3.1 JSR1-JSGiq complex overall structure.....	96
4.2.3.2 The retinal binding pocket .....	97
4.2.3.3 Micro-switch domains in JSR1.....	101
4.2.3.4 JSR1-G protein interactions.....	104
4.2.3.5 Summary JSR1-JSGiq structure analysis.....	110
<b>5 Discussion .....</b>	<b>111</b>
<b>5.1 Adhesion GPCR project.....</b>	<b>111</b>
5.1.1 AGPCR-G protein complex .....	111
5.1.1.1 Biochemistry.....	111
5.1.1.2 Cryo-EM analysis of the Adgrd1 <sup>CTF</sup> -mini-Gs complex .....	114
5.1.2 Adgrd1-NTF .....	117
5.1.2.1 Biochemistry and crystallization .....	117
5.1.3 Adgrd1 fusion proteins .....	118
5.1.3.1 Adgrd1 <sup>CTF</sup> -mini-Gα <sub>s</sub> fusion protein .....	118
5.1.3.2 Adgrd1 <sup>GAIN-CTF</sup> -mini-Gα <sub>s</sub> fusion protein .....	118
<b>5.2 Jumping spider rhodopsin-1 project.....</b>	<b>120</b>
5.2.1 JSR1-G protein complexes.....	120
5.2.1.1 Purifications .....	120
5.2.1.2 GTPase Glo assay .....	122
5.2.1.3 Analytical SEC experiments .....	123
5.2.1.4 Cryo-EM analysis of JSR1-G protein complexes .....	125
5.2.2 JSR1-JSGiq structure.....	127
5.2.2.1 Active state JSR1 .....	127
5.2.2.2 N-terminal extracellular domain.....	127
5.2.2.3 The retinal binding pocket .....	128
5.2.2.4 Micro-switch domains.....	131
5.2.2.5 JSR1-G protein interactions.....	133
<b>6 Conclusion .....</b>	<b>136</b>
<b>6.1 Adhesion GPCR project.....</b>	<b>136</b>
<b>6.2 Jumping spider rhodopsin-1 project.....</b>	<b>139</b>
<b>7 Outlook.....</b>	<b>145</b>
<b>8 Appendix .....</b>	<b>148</b>
<b>8.1 Supplementary Tables .....</b>	<b>148</b>

<b>8.2 Presentations at scientific events.....</b>	<b>168</b>
8.2.1 Poster presentations.....	168
8.2.2 Scientific talk.....	168
<b>8.3 List of publications .....</b>	<b>168</b>
<b>9 Bibliography.....</b>	<b>169</b>
<b>10 Annexed Documents .....</b>	<b>177</b>
10.1 Curriculum Vitae.....	177

## Acknowledgment

This work would not have been possible without the contribution of numerous people, and I am truly thankful to each one of them. First, I would like to express my deepest gratitude to my doctoral supervisor, Prof. Gebhard Schertler, for his unwavering support, guidance, and encouragement throughout my doctoral journey. His wisdom and mentorship have been instrumental in shaping my scientific growth. I am immensely grateful to my scientific supervisor, Dr. Ching-Ju Tsai, whose expertise and insightful guidance have been invaluable to the success of my research. Dr. Tsai's dedication to scientific excellence has been a constant source of inspiration. I extend my sincere appreciation to the members of my doctoral committee, Prof. Frédéric Allain and Prof. Adriano Aguzzi, for their valuable feedback, constructive criticism, and the thoughtful discussions that enriched my research. I would like to thank all past and present members of the Schertler group for creating a cooperative, supportive and pleasant working environment at the Paul Scherrer Institute (PSI). Special thanks to Jonas Mühle, whose contributions in the wet-lab work and engaging discussions significantly enhanced the quality of this work and my time at PSI. I am also indebted to Dr. Matthew Rodrigues for his significant contributions to the success of this research project. In addition to our scientific discussions, the informal conversations and non-work-related events within the Schertler group and the Laboratory of Biomolecular Research have been instrumental in creating a dynamic and enriching work environment. These moments of camaraderie and shared experiences have fostered a sense of unity and collaboration among colleagues, contributing significantly to the overall work experience. I would also like to acknowledge the collaborative spirit fostered by Prof. Adriano Aguzzi and his group. Our collaborative efforts have been a cornerstone of this research, and I am grateful for the enriching experience and insights into their field of research. I am deeply thankful to Dr. Pavel Afanasyev, whose invaluable guidance and expertise in cryo-electron microscopy were instrumental to the success of this work. Dr. Afanasyev's mentorship and insights significantly contributed to my understanding of cryo-EM techniques, shaping the trajectory of this research. Administrative support provided by Francine Weber and Dr. Gregor Chiccheti is duly recognized and appreciated. Their efficiency and dedication ensured the smooth progress of the administrative aspects of this work. I extend my heartfelt thanks to everyone in the biology division of PSI for fostering an environment of intellectual curiosity and stimulating discussions. The camaraderie and shared passion for research have been a driving force.

Last but not least, I am deeply thankful to my friends and family for their unconditional support and understanding. Their encouragement has been a constant source of strength throughout this challenging yet rewarding journey.

## Abbreviations

<b>7TM</b>	-	Seven transmembrane domain
<b>AC</b>	-	Adenylyl cyclase
<b>aGPCR</b>	-	Adhesion G protein-coupled receptor
<b>AH</b>	-	$\alpha$ -helical domain
<b>ATP</b>	-	Adenosine-5'-triphosphate
<b>ATR</b>	-	All-trans retinal
<b>ATR611</b>	-	All-trans retinal 6.11
<b><math>\beta</math>-ME</b>	-	$\beta$ -mercaptoethanol
<b><math>\beta</math>-OG</b>	-	N-octyl- $\beta$ -D-Glucoside
<b>cAMP</b>	-	Cyclic adenosine monophosphate
<b>CHAPS</b>	-	3-[(3- <b>cholamidopropyl</b> )dimethylammonio]-1- <b>propanesulfonate</b>
<b>CHS</b>	-	Cholesterylhemisuccinat
<b>CI</b>	-	Counterion
<b>Cryo-EM/EM</b>	-	Cryogenic electron microscopy/electron microscopy
<b>CTF</b>	-	aGPCR context: C-terminal fragment
<b>CTF</b>	-	EM context: contrast transfer function
<b>CV</b>	-	Column volume
<b>C-W-x-P</b>	-	Cysteine-Tryptophan-X-Proline conserved motif in GPCRs
<b>DAG</b>	-	Diacylglycerol
<b>DDM</b>	-	N-dodecyl $\beta$ -D-maltoside
<b>DM</b>	-	N-decyl- $\beta$ -D-maltopyranoside
<b>DNA</b>	-	Deoxyribonucleic acid
<b>DTT</b>	-	Dithiothreitol
<b>D/E-R-Y</b>	-	Aspartate/Glutamate - Arginine – Tyrosine conserved motif in GPCRs
<b>ECD</b>	-	Extracellular domain
<b>ECL</b>	-	Extracellular loop
<b>EDTA</b>	-	Ethylenediaminetetraacetic acid
<b>EGF</b>	-	Epidermal growth factor
<b>ETHZ</b>	-	Eidgenössische Technische Hochschule Zürich
<b>FSC</b>	-	Fourier shell correlation
<b>GAIN</b>	-	GPCR autoproteolysis-inducing domain
<b>GAP</b>	-	GTPase-activating protein
<b>GBM</b>	-	Glioblastoma multiforme

<b>GDN</b>	-	Glyco-diosgenin
<b>GDP</b>	-	Guanosine diphosphate
<b>GEF</b>	-	Guanidine exchange factor
<b>GFP</b>	-	Green fluorescent protein
<b>GPCR</b>	-	G protein coupled receptor
<b>GPS</b>	-	GPCR proteolytic site
<b>GRAFS</b>	-	Glutamate-like, Rhodopsin-like, Adhesion, Frizzled/taste2, Secretin receptors
<b>GRK</b>	-	G protein-coupled receptor kinase
<b>GTP</b>	-	Guanosine triphosphate
<b>GTP<math>\gamma</math>S</b>	-	Guanosine S-( $\gamma$ -thio) triphosphate
<b>HA</b>	-	Hemagglutinin
<b>HEK 293</b>	-	Human embryonic kidney 293 cells
<b>HEK 293 GnTI<sup>-</sup></b>	-	HEK293 cells lacking N-acetylglucosaminyltransferase I
<b>HEPES</b>	-	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
<b>hGx</b>	-	human G protein x; x indicates the subtype
<b>HRP</b>	-	Horseradish peroxidase
<b>ICD</b>	-	Intracellular domain
<b>ICL</b>	-	Intracellular loop
<b>IMAC</b>	-	Immobilized metal affinity chromatography
<b>IP<sub>3</sub></b>	-	Inositol 1,4,5-trisphosphate
<b>IPTG</b>	-	Isopropyl $\beta$ -D-1-thiogalactopyranoside
<b>JSGiq</b>	-	Chimeric human/jumping spider G <sub>iq</sub> protein
<b>JSGq</b>	-	Jumping spider visual G <sub>q</sub> protein 1
<b>JSR1</b>	-	Jumping spider rhodopsin 1
<b>kDa</b>	-	Kilo Dalton
<b>LB</b>	-	Lysogeny broth
<b>LMNG</b>	-	Lauryl maltose neopentyl glycol
<b>LN-7TM</b>	-	Long N-terminal region seven transmembrane receptor
<b>LNB-7TM</b>	-	Long N-terminal region seven transmembrane receptor with similarities to class B Secretin receptors
<b>MD</b>	-	Molecular dynamics
<b>mMeOp</b>	-	Mouse melanopsin
<b>MW</b>	-	Molecular weight

<b>Ni-NTA</b>	-	Nickel-nitrilotriacetic acid
<b>N-P-x-x-Y</b>	-	Asparagine-Proline-X-X-Tyrosine conserved motif in GPCRs
<b>NTF</b>	-	N-terminal fragment
<b>OD</b>	-	Optical density
<b>P-I-F</b>	-	Proline-Isoleucine-Phenylalanine conserved motif in GPCRs
<b>PCR</b>	-	Polymerase chain reaction
<b>PDB</b>	-	Protein databank
<b>PEI</b>	-	Polyethylenimine
<b>PHENIX</b>	-	Python-based Hierarchical environment for integrative crystallography
<b>PIP<sub>2</sub></b>	-	Phosphatidylinositol 4,5-diphosphate
<b>PKA</b>	-	Phosphokinase A
<b>PKC</b>	-	Phosphokinase C
<b>PLC<math>\beta</math></b>	-	Phospholipase C $\beta$
<b>PNS</b>	-	Peripheral nervous system
<b>PrP<sup>C</sup></b>	-	Cellular prion protein
<b>PSB</b>	-	Protonated Schiff base
<b>PTK7</b>	-	Protein tyrosine kinase 7
<b>PTM</b>	-	posttranslational modification
<b>PSI</b>	-	Paul Scherrer Institute
<b>Rcf</b>	-	Relative centrifugal force
<b>RGS</b>	-	Regulators of G protein signaling
<b>RhoGEF</b>	-	Rho guanine nucleotide exchange factors
<b>ROS</b>	-	Rod outer segment
<b>Rpm</b>	-	Rotations per minute
<b>RT</b>	-	Room temperature
<b>SB</b>	-	Schiff base
<b>SC</b>	-	Schwann cell
<b>SDS-PAGE</b>	-	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
<b>SEC</b>	-	Size exclusion chromatography
<b>SNARE</b>	-	Soluble N-ethylmaleimide-sensitive-factor attachment receptor
<b>SPA</b>	-	Single particle analysis
<b>TB</b>	-	Terrific broth
<b>TCEP</b>	-	Tris(2-carboxyethyl)phosphine

<b>TEV</b>	-	Tobacco etch virus
<b>T<sub>m</sub></b>	-	Melting temperature
<b>TMX</b>	-	Transmembrane helix number X
<b>TRP</b>	-	Transient receptor potential channel
<b>TRPL</b>	-	Transient receptor potential-like channel
<b>UV</b>	-	Ultraviolet
<b>UV-Vis</b>	-	Ultraviolet-Visible
<b>VOI</b>	-	Virality of infection
<b>WHO</b>	-	World health organization
<b>WT</b>	-	Wild-type
<b>YFP</b>	-	Yellow fluorescent protein



## Summary

G protein-coupled receptors (GPCR) play an important role in sensing extracellular stimuli and transmitting these signals inside of the cell. Humans have more than 800 genes encoding GPCRs which are involved in diverse physiological processes like vision, smell, taste, neurotransmission or metabolism. In my PhD project I was working with GPCRs from two different families, called the adhesion GPCR family (class B2 GPCRs) and the opsin family (class A GPCRs).

Adhesion GPCRs (aGPCRs) are characterized by large extracellular N-terminal domains and the presence of a tethered agonist. aGPCRs show a great variety of domain types in the extracellular region but many are implicated in cell-cell or cell-matrix interactions. Despite the diversity of domains, they all possess a GPCR autoproteolysis-inducing domain (GAIN). The last beta strand of the GAIN domain just after the autoproteolysis site was found to act as a tethered agonist, termed Stachel sequence. Many aGPCRs have been shown to be activated by the tethered agonist but mechanistic and structural insights were missing. As part of my PhD project, I was working towards determining the structure of an aGPCR bound to the Stachel peptide and a G protein to gain insights into the binding pose of the Stachel peptide and activation mechanism of aGPCRs. For this purpose, we were working with two different aGPCRs called *Adgrg6* and *Adgrd1*. Our collaborators, Prof. Adriano Aguzzi and his group (USZ), showed that both receptors bind to the cellular prion protein ( $\text{PrP}^{\text{C}}$ ) with opposite effects. For *Adgrg6* the  $\text{PrP}^{\text{C}}$  is an agonist and for *Adgrd1* the  $\text{PrP}^{\text{C}}$  acts as an antagonist. *Adgrg6* has been shown to be important for the myelination in the peripheral nervous system not only during development but also to maintain the myelination. *Adgrd1* was shown to be involved in glioblastoma growth where *Adgrd1* is *de novo* expressed in glioblastoma cells. Knockout or inhibition of *Adgrd1* lead to decreased tumor growth *ex vivo*, hinting that *Adgrd1* might be a valuable target for treatment of glioblastomas. While Prof. Adriano Aguzzi and his group were working on physiological questions, I was working on the structure determination of *Adgrd1* and *Adgrg6*.

I screened constructs in different expression hosts to express the receptors and did several screenings to optimize expression and stability of the receptors. I showed that both receptors can be activated by the Stachel peptide and form a complex with the Gs heterotrimer. Because *Adgrd1* showed favorable biochemical characteristics, the focus was on *Adgrd1* for structure determination. Furthermore, the experiments showed that complex formation is necessary to stabilize *Adgrd1* in order to purify the receptor. Co-expression of the receptor and the G protein, followed by complex formation, extraction of the complex from the plasma membrane by detergents and purification of the complex allowed preparation of a sample for cryo-EM studies. Promising cryo-EM data was collected but before we could solve the structure, other research groups published structures of several aGPCRs (*Adgrd1* was among them) activated by the Stachel peptide. For this reason, we did

not continue with the structure determination of the Adgrd1-Gs complex. Nevertheless, my work provides biochemical insights for Adgrg6 and Adgrd1 which can be useful for future studies on aGPCRs. In my second project, I was working with the jumping spider rhodopsin 1 (JSR1) which is part of the opsin family. Rhodopsins (Opsins bound to the retinal ligand) sense light through the covalently bound retinal molecule. The retinal molecule is covalently bound to the receptor via a Schiff base link. Rhodopsins bind the cis-retinal which acts as an inverse agonist and upon photon absorption, the retinal isomerizes to the all-trans conformation leading to structural changes in the receptor, thereby activating the receptor. Opsins are further classified into monostable and bistable opsins based on the thermal stability of the all-trans retinal. While monostable rhodopsins lose their all-trans retinal by hydrolyzing the Schiff base bond, bistable rhodopsins keep the retinal covalently bound during the whole photocycle and the retinal can switch back to the cis isomer after a second photon absorption event. Bistable rhodopsins are of great interest to researchers because they can be used as an optogenetic tool where signaling through this type of receptors can be controlled by specific wavelengths of light.

JSR1 is a bistable invertebrate rhodopsin important for the depth perception in jumping spiders (*Hasarius adansoni*). JSR1 was used successfully in optogenetic applications, showing its potential as an optogenetic tool. Our research group previously determined the crystal structure of JSR1 bound to the inverse agonist 9-cis retinal. The structure provided insights into the architecture of JSR1 in the inactive state. However, structural insights into the active state were still missing.

In my project I determined the three-dimensional structure of JSR1 bound to the non-natural agonist all-trans retinal 6.11 in complex with a G protein. This structure provided us with valuable insights into the JSR1 activation mechanism and how bistability is achieved in JSR1. Comparison to other active state GPCRs showed that the major conserved motifs in JSR1 follow the typical conformational changes seen in other GPCRs. Interestingly, these changes are more similar to non-photosensitive class A GPCRs than the photosensitive monostable vertebrate rhodopsin, meaning that JSR1 might also be a good receptor to study general mechanisms of class A GPCR activation. Furthermore, our structure provided insights how the protonated Schiff base link is stabilized in the active state. In both the inactive and active state a water mediated network is important to stabilize the protonated Schiff base. The active JSR1 structure showed that the water mediated network changes compared to the inactive state JSR1, highlighting residues that are involved in this network. JSR1 is not yet optimal for optogenetic applications because the excitation wavelengths ( $\lambda_{\max}$  value) of the inactive and active states of JSR1 overlap, meaning that illumination will lead to a mixed population of inactive and active states. Our structures will allow rational engineering of JSR1 in order to change the  $\lambda_{\max}$  value of the inactive and/or active state JSR1 to make JSR1 a more suitable optogenetic tool.

In addition, our JSR1-G protein structures highlight several features at the G protein binding site. A former PhD student Dr. Filip Pamula solved a lower resolution structure of JSR1 bound to a human Gi heterotrimer (JSR1-hGi). In my case I used a chimeric G protein where we mutated key residues in the human G $\alpha$ i subunit to match the sequence of the jumping spider visual G $\alpha$ q. The JSR1-hGi structure showed a different binding pose of the G protein than the chimeric Giq bound to JSR1. Moreover, the binding pose of the hGi showed differences to other GPCR-hGi structures, hinting that invertebrate GPCRs might engage human G proteins differently. Together these structures highlight the different interactions of JSR1 with the hGi and chimeric Giq which can be important for optogenetic applications, but further research is needed.

To summarize, this structure presents the first high resolution structure of an active bistable rhodopsin, providing insights into the mechanism of how bistability is achieved. Additionally, our active JSR1-G protein complex structures also highlight differences on how JSR1 interacts with a human G protein and a chimeric human/jumping spider G protein. This will help to rationally engineer JSR1 for optogenetic applications by changing the  $\lambda_{\max}$  to a desirable wavelength and change the G protein specificity.

## Zusammenfassung

G-Protein-gekoppelte Rezeptoren (GPCRs) spielen eine wichtige Rolle bei der Wahrnehmung extrazellulärer Reize und der Übertragung dieser Signale innerhalb der Zelle. Der Mensch besitzt etwa 800 Gene, welche GPCRs kodieren, die an verschiedenen physiologischen Prozessen wie Sehen, Riechen, Schmecken, Neurotransmission oder Stoffwechsel beteiligt sind. In meinem Promotionsprojekt habe ich mit GPCRs aus zwei verschiedenen Gruppen gearbeitet, der Adhäsions-GPCR-Gruppe und der Opsin-Gruppe.

Adhäsions-GPCRs (aGPCRs) zeichnen sich durch große extrazelluläre N-terminale Domänen und die Existenz eines gebundenen Agonisten aus. aGPCRs weisen eine große Vielfalt an Domärentypen in der extrazellulären Region auf, von denen viele an Zell-Zell- oder Zell-Matrix-Interaktionen beteiligt sind. Trotz der Vielfalt an Domänen besitzen sie alle eine GPCR-Autoproteolyse-induzierende Domäne (GAIN). Es wurde festgestellt, dass der letzte Beta-Strang der GAIN-Domäne unmittelbar nach der Autoproteolysestelle als gebundener Agonist fungiert. Diese Sequenz wurde später als Stachel-Sequenz bezeichnet. Bei vielen aGPCRs wurde gezeigt, dass sie durch den gebundenen Agonisten aktiviert werden. Es fehlten jedoch mechanistische und strukturelle Einblicke.

Im Rahmen meines Promotionsprojekts habe ich daran gearbeitet, die dreidimensionale Struktur eines aGPCRs, welcher an das Stachel-Peptid und ein G-Protein gebunden ist, zu bestimmen. Dies ermöglicht, Einblicke in die Bindungsposition des Stachel-Peptids und den Aktivierungsmechanismus von aGPCRs zu erhalten. Zu diesem Zweck haben wir mit zwei verschiedenen aGPCRs namens Adgrg6 und Adgrd1 gearbeitet. Unsere Kollaborateure, Prof. Adriano Aguzzi und seine Gruppe (USZ), zeigten, dass beide Rezeptoren an das zelluläre Prion Protein (PrP<sup>C</sup>) binden. Für Adgrg6 ist das PrP<sup>C</sup> ein Agonist und für Adgrd1 wirkt das PrP<sup>C</sup> als Antagonist. Es wurde gezeigt, dass Adgrg6 für die Myelinisierung im peripheren Nervensystem nicht nur während der Entwicklung, sondern auch zur Aufrechterhaltung der Myelinisierung wichtig ist. Für Adgrd1 wurde gezeigt, dass Adgrd1 am Wachstum von Glioblastomen beteiligt ist, wobei Adgrd1 *de novo* in Glioblastomzellen exprimiert wird. Das Ausschalten oder die Hemmung von Adgrd1 führte zu einem verminderten Tumorwachstum *ex vivo*, was darauf hindeutet, dass Adgrd1 ein wertvolles Ziel für die Behandlung von Glioblastomen sein könnte. Während Prof. Adriano Aguzzi und seine Gruppe an physiologischen Fragen arbeiteten, beschäftigte ich mich mit der Strukturbestimmung von Adgrd1 und Adgrg6.

Ich habe Konstrukte in verschiedenen Expressionssystemen getestet, um die Rezeptoren zu exprimieren, und verschiedene Experimente zur Optimierung der Expression und Stabilität der Rezeptoren durchgeführt. Ich konnte zeigen, dass beide Rezeptoren durch das Stachel-Peptid aktiviert werden können und einen Komplex mit dem Gs-Heterotrimer bilden. Da Adgrd1 günstigere biochemische Eigenschaften zeigte als Adgrg6, lag der Fokus auf der Strukturbestimmung von Adgrd1.

Darüber hinaus zeigten die Experimente, dass die Bildung eines Komplexes erforderlich ist, um Adgrd1 zu stabilisieren und den Rezeptor aufzureinigen zu können. Die Koexpression des Rezeptors und des G-Proteins, gefolgt von der Komplexbildung und Aufreinigung, ermöglichte die Vorbereitung einer Probe zur Analyse mittels Kryo-Elektronenmikroskopie (Kryo-EM). Vielversprechende Kryo-EM-Daten wurden gesammelt, aber bevor wir die Struktur lösen konnten, veröffentlichten mehrere andere Forschungsgruppen die Struktur von Adgrd1 und anderen aGPCRs, die durch das Stachel-Peptid aktiviert wurden. Interessanterweise stimmen unsere Daten gut überein mit den publizierten Daten. Trotzdem liefert meine Arbeit biochemische Erkenntnisse zu Adgr6 und Adgrd1, die für zukünftige Studien an aGPCRs nützlich sein können.

In meinem zweiten Projekt habe ich mit dem Rhodopsin 1 der Gewächshausspringspinne (*Hasarius adansoni*) (englisch: jumping spider rhodopsin 1; (JSR1)) gearbeitet, welches zur Opsin-Familie gehört. Rhodopsine (Opsine, die an den Retinal Ligand gebunden sind) nehmen Licht durch das kovalent gebundene Retinal Molekül wahr. Das Retinal Molekül ist über eine Schiff'sche Base kovalent an den Rezeptor gebunden. Rhodopsine binden das cis-Retinal, welches als inverser Agonist wirkt. Nach der Absorption eines Photons isomerisiert das Retinal in die all-trans-Konformation, was zu strukturellen Veränderungen im Rezeptor und damit zur Aktivierung führt. Opsine werden weiter in monostabile und bistabile Opsine unterteilt, basierend auf der thermischen Stabilität des all-trans-Retinals. Während monostabile Rhodopsine ihr all-trans-Retinal durch Hydrolyse der Schiff'schen Base verlieren, behalten bistabile Rhodopsine das Retinal während des gesamten Fotozyklus gebunden, welches nach einer zweiten Photonenaufnahme zum cis-Retinal zurückisomerisiert. Bistabile Rhodopsine sind für Forscher von großem Interesse, da sie als optogenetisches Werkzeug verwendet werden können, um die Aktivität des Rezeptors durch spezifische Lichtwellenlängen zu kontrollieren. JSR1 ist ein bistabiles Rhodopsin, das für die Tiefenwahrnehmung bei der Gewächshausspringspinne wichtig ist. JSR1 wurde erfolgreich in optogenetischen Anwendungen eingesetzt, was sein Potenzial als optogenetisches Werkzeug zeigt. Unsere Forschungsgruppe hat zuvor die Kristallstruktur von JSR1, gebunden an den inversen Agonisten 9-cis-Retinal, bestimmt. Die Struktur lieferte Einblicke in die Architektur von JSR1 im inaktiven Zustand. Strukturelle Einblicke in den aktiven Zustand fehlten jedoch noch.

In dieser Arbeit habe ich die dreidimensionale Struktur von JSR1 gebunden an den nicht-natürlichen Agonisten all-trans-Retinal 6.11 in Komplex mit einem G-Protein bestimmt. Diese Struktur lieferte wertvolle Einblicke in den Aktivierungsmechanismus von JSR1 und welche Faktoren zu der Bistabilität bei JSR1 führen. Der Vergleich mit anderen aktiven GPCRs zeigte, dass die wichtigsten konservierten Motive in JSR1 den typischen konformationellen Veränderungen in anderen GPCRs folgen. Interessanterweise ähneln diese Veränderungen eher nicht-photosensitiven GPCRs der Klasse A als

dem photosensitiven monostabilen Wirbeltier-Rhodopsin, was bedeutet, dass JSR1 sich auch als Rezeptor eignen könnte um allgemeine Mechanismen der Aktivierung von GPCRs der Klasse A zu studieren. Darüber hinaus lieferte unsere Struktur Einblicke, wie die protonierte Schiff'sche Base im aktiven Zustand stabilisiert wird. Sowohl im inaktiven als auch im aktiven Zustand ist ein wasser-vermitteltes Netzwerk wichtig, um die protonierte Schiff'sche Base zu stabilisieren. Die Struktur vom aktiven JSR1 zeigte, dass sich das wasser-vermittelte Netzwerk im Vergleich zum inaktiven Zustand von JSR1 verändert. JSR1 ist bislang nicht optimal für optogenetische Anwendungen geeignet, da die Anregungswellenlängen ( $\lambda_{\max}$  Wert) des inaktiven und aktiven Zustandes von JSR1 überlappen. Dies führt dazu, dass nach der Bestrahlung des Rezeptors eine gemischte Population von inaktiven und aktiven Zuständen präsent ist. Unsere Strukturen werden es ermöglichen, JSR1 rational zu modifizieren, um den  $\lambda_{\max}$  Wert des inaktiven und/oder aktiven Zustands von JSR1 zu ändern und JSR1 zu einem geeigneteren optogenetischen Werkzeug zu machen.

Zusätzlich heben unsere JSR1-G-Protein-Strukturen mehrere Merkmale an der Bindungsstelle des G-Proteins hervor. Ein ehemaliger Promotionsstudent, Dr. Filip Pamula, löste eine Struktur mit niedrigerer Auflösung von JSR1, die an ein menschliches Gi-Heterotrimer gebunden ist (JSR1-hGi). In dieser Arbeit habe ich ein chimäres G-Protein verwendet, bei dem wir Aminosäuren in der menschlichen G $\alpha$ i-Untereinheit mutiert haben, um sie an die Sequenz des visuellen G $\alpha$ q der Springspinne anzupassen. Die JSR1-hGi-Struktur zeigte eine andere Bindungsposition des G-Proteins als das an JSR1 gebundene chimäre Giq. Darüber hinaus zeigte die Bindungsposition des hGi Unterschiede zu anderen GPCR-hGi-Strukturen, was darauf hindeutet, dass GPCRs von wirbellosen Tieren möglicherweise auf eine andere Art mit menschlichen G-Proteinen interagieren. Zusammen zeigen diese Strukturen die unterschiedlichen Wechselwirkungen von JSR1 mit dem hGi und dem chimärem Giq auf, die für optogenetische Anwendungen wichtig sein können, jedoch ist weitere Forschung nötig, um die Wechselwirkungen genauer zu verstehen.

Zusammenfassend stellt diese Struktur die erste hochauflösende Struktur eines bistabilen Rhodopsins dar, die Einblicke in das Erreichen der Bistabilität bietet. Darüber hinaus heben unsere aktiven JSR1-G-Protein-Komplexstrukturen auch Unterschiede in der Interaktion von JSR1 mit einem menschlichen G-Protein und einem chimären menschlich-springspinnenartigen G-Protein hervor. Dies wird dazu beitragen, JSR1 rational zu modifizieren, um eine Änderung des  $\lambda_{\max}$  Wertes zu einer gewünschten Wellenlänge und eine Änderung der G-Protein-Spezifität zu erreichen.

# 1 Introduction

## 1.1 G protein coupled receptor superfamily

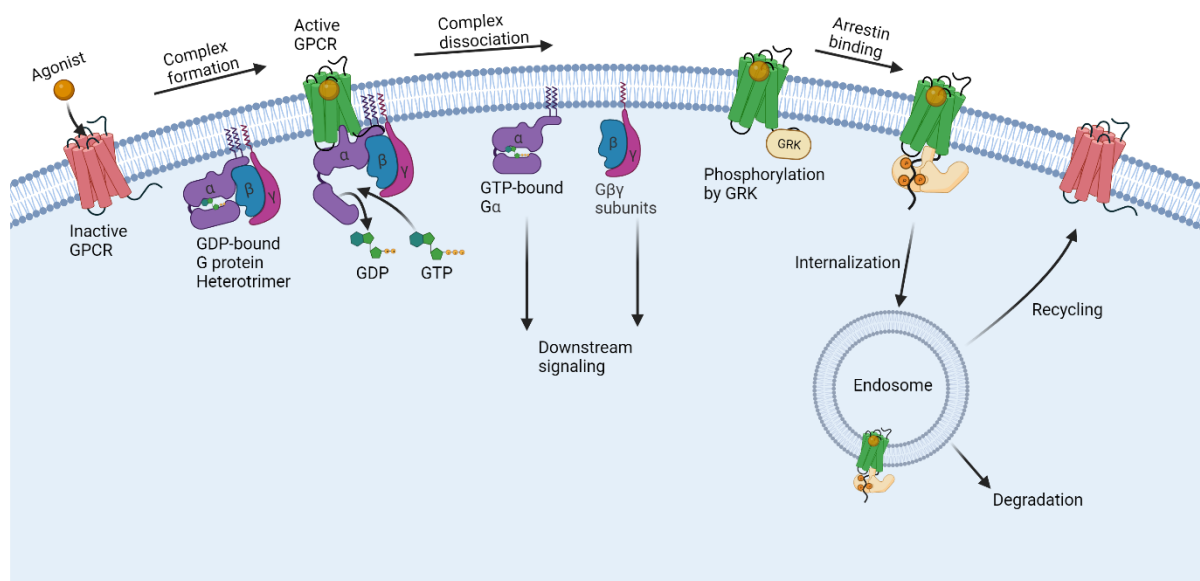
G protein coupled receptors (GPCRs) are one of the most studied receptor families because they play crucial roles in many physiological processes, such as smell, taste, vision, neurotransmission, metabolism, cellular differentiation and growth, inflammatory and immune response [1-3]. Disregulation or malfunction of GPCRs can lead to a variety of diseases, including cancer, nervous system disorders, diabetes, inflammation and cardiac dysfunction. From all approved drugs, approximately 35% target GPCRs, making them major drug targets [4]. However, there are still many orphan GPCRs, meaning GPCRs where no ligand has been identified yet. Only about 12% of all GPCRs encoded in the human genome are targeted by the current drugs but they are still the most popular drug target class for the discovery of new drugs [4-6].

The GPCR superfamily is one of the biggest protein families with around 800 GPCRs encoded in the human genome [7]. The common seven transmembrane (7TM) domain and the association with heterotrimeric guanine nucleotide-binding proteins (G proteins) characterize GPCRs [8]. GPCRs are classified into families based on their amino acid sequence similarity and functional similarities in the widely used A-F classification system [9, 10]. In this system the GPCR superfamily consists of six main families. The class A Rhodopsin-like family is the largest family with around 80% of all GPCRs grouping into this family including hormone, neurotransmitter and light-sensitive receptors [11]. In humans, there are 284 members plus 380 olfactory receptors which are structurally characterized by the 7TM in combination with an eighth helix, associating to the cytoplasmic membrane with one or a few palmitoylated cysteines at the C Terminus [11, 12]. The class B secretin receptor family consists of about 70 receptors that, besides their 7TM domain, have large N-terminal domains. Only after the release of the human genome, researchers realized that the class B can be further divided into two families, the secretin family (class B1) and the adhesion family (class B2) [11, 13, 14]. Class C glutamate receptors include metabotropic glutamate receptors, GABA receptors, taste receptors and calcium-sensing receptors [11]. These receptors contain a characteristic clamp-shaped extracellular domain to which ligands bind and it is connected to the transmembrane helix 1 (TM1) by a cysteine-rich loop [11]. Additionally, class C GPCRs are the only family which form obligatory dimers [14]. The class D includes fungal mating pheromone receptors, class E cAMP receptors and class F the frizzled/smoothed receptors [11].

For the approximately 800 human GPCRs another classification system, called the "GRAFS", was proposed based on the phylogenetic tree of the human GPCRs [14, 15]. In this system the superfamily is divided into five main families named Glutamate (G), Rhodopsin (R), Adhesion (A), Frizzled/Taste2 (F) and Secretin (S) [14, 15]. The main difference between the two systems is the division of the class

B into the Secretin and Adhesion families based on the finding that these two families have a distinct evolutionary history from each other [14, 15].

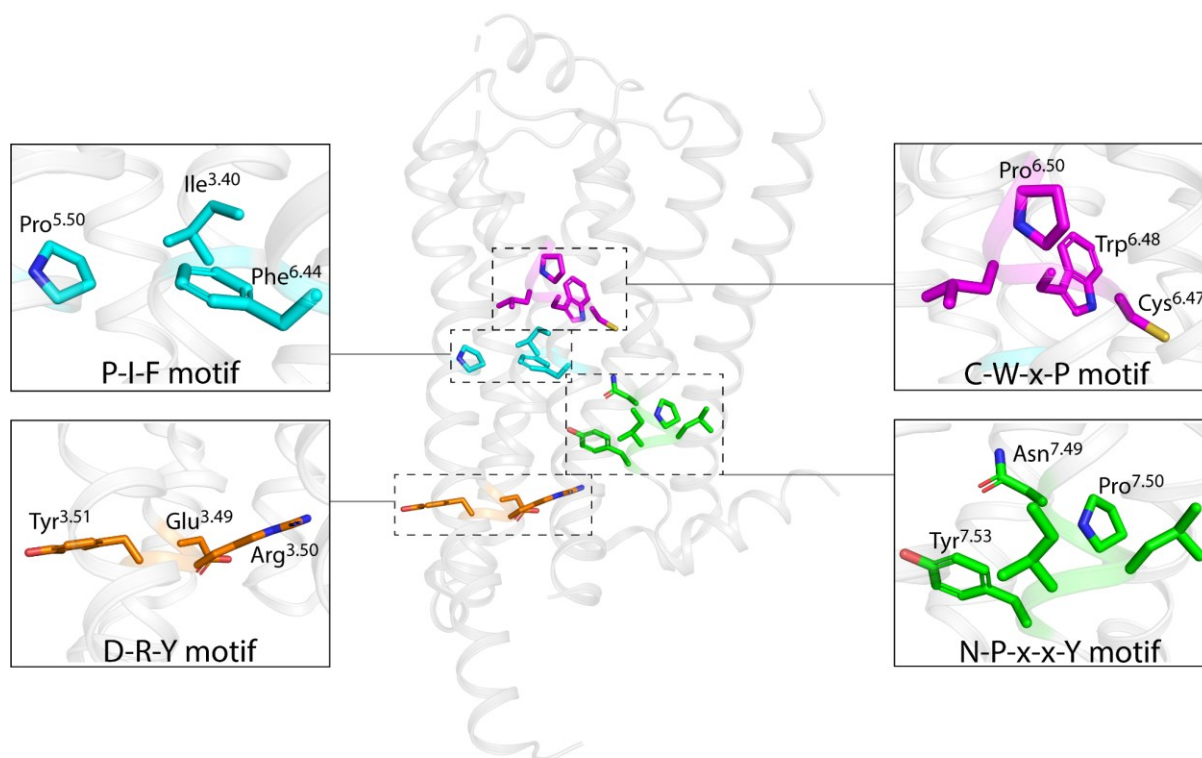
GPCRs govern cellular signaling transduction by sensing extracellular signals that modulate the activity of the receptors. These extracellular stimuli include peptides, proteins, ions, small molecules, mechanical forces, light or odorants [16]. As the name suggests, the main intracellular binding partner of GPCRs are G proteins. Activation of a GPCR leads to conformational changes and the formation of a cytoplasmic cleft that allows binding of the heterotrimeric G protein. The activated receptor induces an exchange of GDP to GTP in the alpha subunit of the G protein ( $G\alpha$ ). Next, the heterotrimer dissociates into the  $G\alpha$  subunit and the  $G\beta\gamma$  dimer. Each of them can activate downstream signaling cascades. The active receptor can either activate another G protein to amplify the signal or its C-terminus can get phosphorylated by a G protein-coupled receptor kinase (GRK). Phosphorylation of the C-terminus can be recognized by arrestin, leading to termination of the G protein activation, followed by internalization of the GPCR and removal of the bound ligand [17]. GPCRs are then either recycled back to the plasma membrane or degraded. **Figure 1** shows a schematic overview of GPCR signaling.



**Figure 1: Schematic overview of GPCR signal transduction.** Inactive GPCR (red) binds an agonist, leading to conformational changes and the opening of a cytoplasmic cleft, allowing the GDP-bound G protein heterotrimer to bind. The active GPCR (green) triggers the exchange of GDP with GTP in the  $G\alpha$  subunit. The GTP-bound G protein dissociates and  $G\alpha$  and  $G\beta\gamma$  activate downstream signaling cascades. Activation of the receptor can lead to C-terminal phosphorylation by GRKs and subsequent binding of arrestin. Arrestin activation can lead to internalization of the receptor into endosomes where the receptor is degraded or recycled back to the plasma membrane. Figure created with Biorender.

Through sequence and structural studies on GPCRs, conserved microswitch domains have been identified which show common structural changes during the activation process of GPCRs. Four such microswitch domains are present in the majority of GPCRs and are called the C-W-x-P, the P-I-F, the N-P-x-x-Y and the (E)D-R-Y motif (**Figure 2**) [18-23].





**Figure 2: The micro-switch domains of GPCRs.** The four major micro-switch domains represented by the crystal structure of the  $\beta_2$ -adrenergic receptor (PDB: 3SN6[24])

The C-W-x-P motif is composed of the highly conserved Cys<sup>6.47</sup>-Trp<sup>6.48</sup>-x-Pro<sup>6.50</sup> (Ballesteros-Weinstein numbering [25]; the first number indicates the transmembrane helix number and the second number indicates the residue position relative to the most conserved residue) in TM6 and plays a critical role in the activation mechanism[18, 21, 26]. The Pro<sup>6.50</sup> acts as a hinge for the outward movement of TM6. The conformational change of the adjacent Trp<sup>6.48</sup> initiates a cascade of structural changes involved in the activation process, leading to the outward movement of TM6. Directly below the C-W-x-P motif is the P-I-F motif composed of the residues Pro<sup>5.50</sup>-Ile<sup>3.40</sup>-Phe<sup>6.44</sup> in TM5, TM3 and TM6 respectively [18]. Phe<sup>6.44</sup> is one helix turn below the Trp<sup>6.48</sup> and is also involved in the outward movement of TM6 upon receptor activation. The third microswitch is the N<sup>7.49</sup>-P<sup>7.50</sup>-x-x-Y<sup>7.53</sup> motif [20]. Tyr<sup>7.53</sup> rearranges upon receptor activation following the TM7 movement towards the transmembrane core. In the active state this residue packs against residues at position 3.44 and 3.47, leading to a hydrogen bond network between Tyr<sup>7.53</sup> and Tyr<sup>5.58</sup> that stabilizes the active state [26, 27]. Tyr<sup>7.53</sup> in the active state is located just above the conserved Arg<sup>3.50</sup> of the E/D<sup>3.49</sup>-R<sup>3.50</sup>-Y<sup>3.51</sup> motif. In the inactive state the E/D-R-Y motif is in a “closed” ionic lock conformation where the Asp/Glu<sup>3.49</sup> forms an intra-helical ionic interaction with Arg<sup>3.50</sup>[22]. Arg<sup>3.50</sup> forms an additional inter-helical ionic interaction with Glu<sup>6.30</sup>. This ionic lock is characteristic of the inactive state of class A GPCRs although the inter-helical ionic interaction is less conserved than the intra-helical ionic interaction [18, 19]. Upon activation of the receptor this ionic lock is broken and the Arg<sup>3.50</sup> adopts a different rotamer, leading to an interaction with the G protein.

This rotamer change is further stabilized by an interaction with Tyr<sup>5.58</sup> [28, 29]. In bovine rhodopsin and the  $\beta$ 2-adrenergic receptor it was shown that this interaction is critical in forming the active state [28, 29].

These structural changes have been well characterized in many class A GPCRs but structural information of bistable rhodopsins was so far limited to the inactive state [30, 31]. It will be interesting to understand if bistable rhodopsins show similar conformational changes upon activation. Additionally, structural information on adhesion GPCRs is very limited until now and understanding if adhesion GPCRs show the same microswitch domains as class A GPCRs will be essential for the broader understanding of the GPCR superfamily.

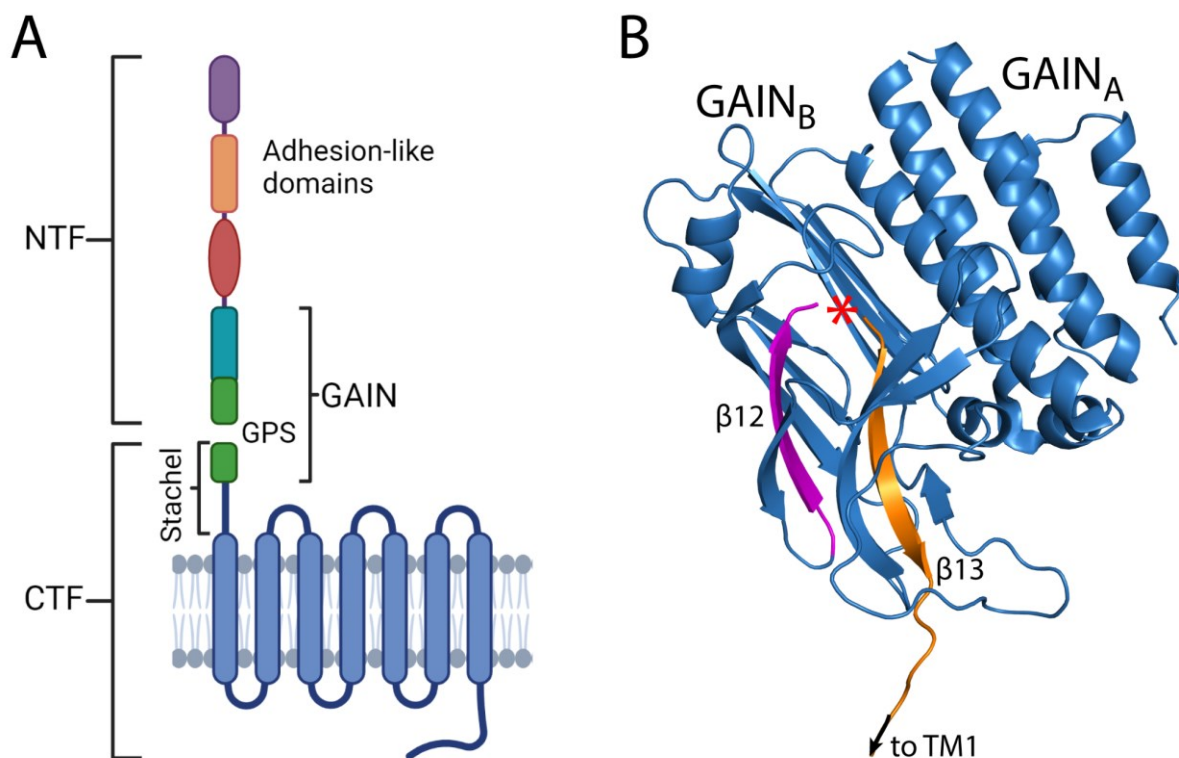
## 1.2 Adhesion GPCR family

Within the GPCR superfamily, the adhesion GPCR (aGPCR) family forms the second largest family, after the rhodopsin family, with 33 receptors encoded in the human genome [15, 32]. The human aGPCR repertoire is divided into nine distinct subfamilies, based on the molecular signature of their 7TM domain [14]. One of the unique features of aGPCRs is that they possess large N-terminal domains that can extend up to nearly 6000 amino acids [32]. Adhesion GPCRs show a great variety of domain types in the extracellular region but many of those are implicated in cell-cell or cell-matrix interactions, which is also the reason for the “adhesion” GPCR family name. The epidermal growth factor (EGF)-like receptor was one of the first aGPCRs to be cloned and similar receptors were subsequently identified, terming these receptors EGF-TM7-like receptors because they contain EGF-domains in their extracellular region [33-35]. Other names that were given for aGPCRs were LN-7TM (for long N-terminal regions) or LNB-7TM (for long N-terminal regions and the similarity to class B secretin-like receptors).

Adhesion GPCRs are an evolutionary old group of receptors as they were found in the genome of *Amoebozoa* (*D. discoideum*) and *Alveolata* (*P. tetraurelia*) based on aGPCR-specific sequence signatures in the 7TM domain, dating the origin of these receptors back to about 1200 million years ago [36]. Even though these ancient receptors lack large N-termini, unicellular non-metazoa choanoflagellate *Monosiga brevicollis* and filasterea *Capsaspora owczarzaki* already contain several aGPCR genes with the GPCR proteolysis site (GPS) site [36]. After the origin of metazoans, a diversification and numerical expansion of the aGPCR family took place that is probably attributed to cell-cell and cell-matrix contacts, a major factor of multicellularity [36]. Subsequently, adhesion GPCRs are found in the genomes of the most ancient metazoan phyla [37], primitive animals [38, 39], the closest relatives of the vertebrates [40, 41] and in all vertebrates [42]. In addition, metazoan GPCR

genome analysis indicates that aGPCRs are ancestral to the Secretin family, which probably diverged from a specific class of GPCRs [43].

After the human genome was published, it became clear that aGPCRs constitute a separate family with distinct features [14]. Besides their large extracellular domains (ECD), every aGPCR (except Adgr1) contains a juxtamembrane GPS, part of the highly conserved GPCR autoproteolysis-inducing (GAIN) domain [44, 45]. During maturation, in the endoplasmic reticulum, the autoproteolysis occurs which results in an N-terminal fragment (NTF) containing the GAIN (except the  $\beta$ 13-strand) and additional adhesion associated domains and a C-terminal fragment (CTF) with the 7TM domain and the intracellular domain (ICD) (**Figure 3A**) [45, 46].



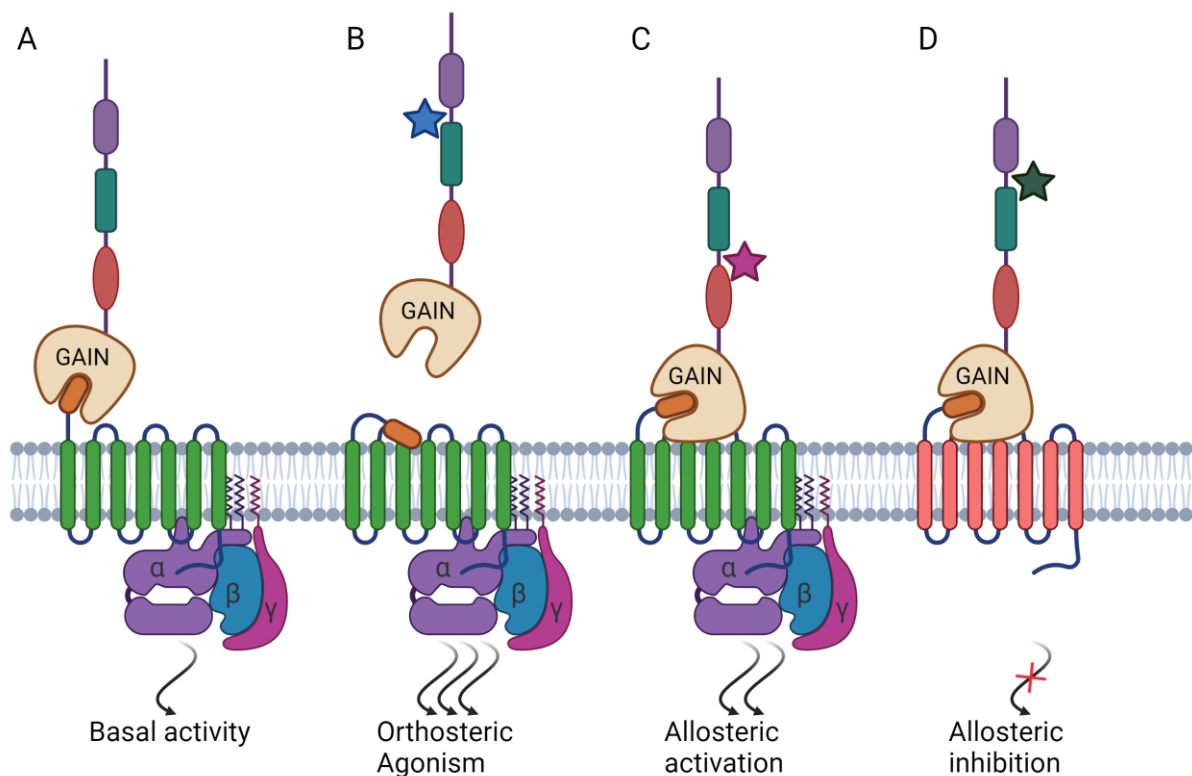
**Figure 3:** **A:** Schematic representation of a generic aGPCR. Important structural features are labelled. **B:** Crystal structure of the GAIN domain from Adgr1 shown as cartoon (PDB: 4DLQ[45]). Subdomain A is labelled  $GAIN_A$ , consisting of six  $\alpha$  helices. Subdomain B is labelled  $GAIN_B$  and consists of 13  $\beta$ -sheets and two small  $\alpha$  helices. Autoproteolysis site is labelled with a red asterisk between the  $\beta$ -strand 12 and 13. The  $\beta$ -strand 13 corresponds to the tethered agonist, also called Stachel sequence. Figure 3A was created with Biorender.

Until 2012 the GAIN domain fold was unidentified, then Araç *et al.* solved the crystal structures of the GAIN domain from the Adgr1 (latrophilin 1) and Adgrb3 (brain angiogenesis inhibitor 3) receptors [45]. These crystal structures revealed that the GAIN domain can be divided into a N-terminal subdomain A, consisting of six alpha helices and a C-terminal subdomain B, consisting of a twisted beta sandwich with 13  $\beta$ -strands and two small helices (**Figure 3B**) [45]. The GPS motif constitutes the last 5  $\beta$ -strands and is highly conserved [45]. Comparing the sequence and structure of the two receptors revealed that the three-dimensional structure of the GAIN domain is more closely conserved than the sequence (24% similarity) [45]. Even though the GPS motif is the most conserved part of the

GAIN domain and necessary for autoproteolysis to happen, the GPS motif alone is not sufficient for the cleavage but the whole GAIN domain is needed and sufficient for autoproteolysis [45]. The cleavage occurs in the short loop between  $\beta$ -strand 12 and  $\beta$ -strand 13 (**Figure 3B**) resulting in a NTF and a CTF that stay associated during receptor trafficking to the cell membrane through non-covalent interactions of the  $\beta$ -strand 13 with the surrounding  $\beta$ -strands [45]. The  $\beta$ -strand 13 also corresponds to the Stachel peptide that can act as a tethered agonist and activates aGPCRs [47]. Although all aGPCRs, except Adgra1, contain the GAIN domain, not all receptors are predicted to be cleavable due to the lack of a consensus catalytic triad within the GPS motif that is a prerequisite for cleavage to happen [48]. Whereas some receptors have been shown experimentally to be uncleavable, others are only predicted to be uncleavable due to the lack of the catalytic triad sequence [49-51]. The catalytic triad has a consensus sequence of Hx↓T/S/C and detailed analysis confirmed that proteolysis happens as an autocatalytic reaction [44].

For a long time, the search for an agonist for aGPCRs was unsuccessful because of the limited information about the signal transduction of aGPCRs. Without agonists, overexpression of receptors in heterologous expression systems was used to show G protein-coupling abilities for several receptors [52-56]. An important contribution to understanding the activation mechanism of aGPCRs, was the finding that deletion of the NTF of Adgrb2 [56], Adgrg1 [54, 57] and Adgre5 [58] resulted in an activated receptor. A milestone study was published in 2014, where it was shown that the  $\beta$ -strand 13 acts as a tethered agonist and the receptors can be activated by chemically synthesized peptides mimicking this sequence [47]. Subsequently, the tethered agonist sequences for several other aGPCRs, Adgrg1 and Adgrf1 [59], Adgrg2 [60], Adgrg5 [61], and Adgrl1 [62], have been identified.

The existence of a tethered agonistic sequence for adhesion GPCRs is widely accepted but the mechanism how aGPCRs are activated is still unclear. There are two general models suggested (**Figure 4**). The first one involves the interaction of the NTF with a ligand in combination with a mechanical force which leads to the removal of the NTF and subsequently the liberated tethered agonist sequence can bind and activate the receptor (orthosteric agonism) (**Figure 4**) [51, 59, 63-66]. The second model includes the interaction of the NTF with a ligand, leading to structural rearrangements that tune the activity of the receptor (allosteric regulation) (**Figure 4**) [59]. In the second model, NTF dissociation is not required.



**Figure 4: Models of aGPCR regulation.** Schematic representations of aGPCRs showing the 7TM domain in green (active receptor) or salmon (inactive receptor), GAIN domain in light orange with the Stachel peptide in orange and N-terminal subdomains in various colors, representing the potential variety within aGPCR ECDs. **A:** Generic aGPCR without a ligand bound. aGPCRs possess varying levels of basal activity. **B:** Orthosteric agonism model. An agonist (blue star) binds to the NTF, leading to the dissociation of the NTF. The exposed Stachel peptide can bind to the orthosteric site within the 7TM, thereby activating the receptor. **CD:** Allosteric regulation model. A ligand (magenta or dark green star) interacts with one of the N-terminal adhesion domains, leading to either activation of the receptor (C) or to stabilization of the inactive state (D). Allosteric regulation mechanisms are unknown but might be mediated through interactions of the GAIN domain with the 7TM domain. Figure created with Biorender.

During my PhD time, four different studies solved structures of several aGPCRs activated by the Stachel sequence, shedding light on the activation mechanism [67-70]. Overall, these studies showed that the binding mode of the tethered agonist is very similar between the different aGPCRs. This allowed the identification of a conserved hydrophobic motif in the Stachel sequence, crucial for binding to the 7TM domain and activation of the receptor. Interestingly, the Stachel sequence adopts an  $\alpha$ -helical structure when bound to the 7TM in contrast to the  $\beta$ -sheet fold in the GAIN domain. Surprisingly, the structure of a cleavage-deficient full-length aGPCR showed the same position of the Stachel sequence [70].

In addition, these structures provide an update for the tethered agonism model of aGPCR activation. Activation by the tethered agonist could happen in two different scenarios: (1) the Stachel sequence is buried in the GAIN domain as a  $\beta$ -sheet in the inactive state. Ligand binding to the NTF or mechanical force lead to dissociation of the NTF, thereby exposing the buried Stachel sequence. The Stachel sequence can interact with the 7TM and adopts an  $\alpha$ -helical structure, leading to receptor activation; (2) the Stachel sequence is prebound to the 7TM and adopts an  $\alpha$ -helical structure. Interactions of

ligands with the NTF result in conformational changes that lead to the activation of the receptor. In this scenario, NTF removal or even autoproteolysis are not required.

In summary, these structures provide further evidence to elucidate the activation mechanism of aGPCRs, but additional experiments are needed to determine which scenario – or both – is correct. The big open question is how the GAIN domain interacts with the 7TM. Solving a structure of an aGPCR together with the GAIN domain could shed light on the activation mechanism.

### 1.2.1 *Adgrg6* and *Adgrd1*

In this work, we focused on two members of the aGPCR family, called *Adgrd1* and *Adgrg6*. While *Adgrg6* is one of the best studied aGPCRs, especially in terms of physiological functions, not so much is known about *Adgrd1*.

*Adgrg6* was shown to be important for the development of the peripheral nervous system (PNS). First in vivo studies demonstrated that *Adgrg6* is important for the normal progression of promyelinating Schwann cells (SC) to mature myelinating SC in zebrafish and mice [71, 72]. It was shown before that a cAMP level increase induces SC maturation, demonstrated by the upregulation of myelin-associated gene expression [73]. Remarkably, the myelination defects in *Adgrg6* deficient zebrafish could be rescued by forskolin, an adenylate cyclase activator that increases cAMP levels [71]. *Adgrg6* was shown to directly couple to  $G_s$  and  $G_i$ , suggesting that *Adgrg6* is a major factor controlling the SC development and proper myelination of axons in the PNS [55]. Moreover, it was shown that *Adgrg6* has NTF and CTF specific functions. Mice, where the whole receptor was deleted, showed additional defects in radial sorting and axon degeneration whereas zebrafish, where only the CTF was deleted, did not show this phenotype [55, 71, 72]. It was shown that these defects can be rescued by adding *Adgrg6* NTF “back” to the *Adgrg6* deficient mice and zebrafish [74]. These findings suggest that the NTF of *Adgrg6* has CTF-independent functions in the development of the PNS. To elucidate the mechanism how *Adgrg6* regulates these functions, it is important to identify ligands of *Adgrg6*. Collagen IV and laminin-211 are both part of the SC basal lamina and can bind to the NTF of *Adgrg6* [74, 75]. Collagen IV was shown to induce cAMP elevation when adding it to *Adgrg6* [75] but interestingly laminin-211 only increased cAMP levels under dynamic conditions and under static conditions even suppressed cAMP elevation [74]. Additionally, overexpression of *lama2* (gene for laminin-211) is able to rescue myelination defects in hypomorphic *Adgrg6* zebrafish mutants [74]. The myelination rescue was not observed when a polymerization-deficient *lama2* mutant was used, hinting at an important function for the laminin-211 polymerization [74]. The current data support a model where, during early development, laminin-211 binds to the NTF, stabilizes the receptor in an inactive state, preventing SC differentiation and favors SC proliferation [74]. During basal lamina

maturation, laminin-211 polymerizes which leads to NTF shedding, probably through mechanical modulation, and the exposed Stachel peptide can bind to the orthosteric binding site, leading to the activation of Adgrg6 and increase of cAMP levels through  $G_s$  signaling [76]. Adgrg6 not only plays a pivotal role in the SC maturation and initiation of the myelination but also in the maintenance of the myelin sheets [77]. Our collaborators on this project, Prof. Adriano Aguzzi and his group (University hospital Zurich), found that the cellular prion protein (PrP<sup>C</sup>) is an agonist for Adgrg6 and induces cAMP elevation [77]. Further, they showed that the PrP<sup>C</sup> promotes myelin homeostasis through the interaction and activation of Adgrg6 [77]. Together, these studies show that Adgrg6 has vital function in both the development and homeostasis of the myelination in the PNS.

For Adgrd1, less is known about the physiological functions of the receptor. Nevertheless, Adgrd1 has been found to be involved in the tumor growth of glioblastoma (GBM), an aggressive brain malignancy [78]. Adgrd1 was shown to signal through the  $G_s$  pathway, elevating cAMP levels [78, 79]. Interestingly, Adgrd1 is not expressed in non-malignant brain tissue, meaning it is *de novo* expressed in GBM cells [80]. Furthermore, the expression levels correlate with the WHO grade of gliomas, showing that Adgrd1 expression is directly linked to the severity of this brain malignancy with all test samples from GBM cells (most advanced form of glioma family) showing Adgrd1 expression [80]. Knockdown of Adgrd1 in GBM cells of mice brains significantly reduced the tumor xenograft formation and increased the host survival, making it an interesting potential target for treatment [78]. However, until now high affinity small molecule ligands for Adgrd1 are still missing. The only known natural ligand of Adgrd1 is the transmembrane protein tyrosine kinase 7 (PTK7) that was shown to allosterically activate Adgrd1 when their respective extracellular domains interact in trans [81]. To evoke a response from Adgrd1, PTK7 needs to be membrane anchored, present on another cell as Adgrd1 (trans interaction) and Adgrd1 needs to be autocatalytically cleaved [81]. Knockdown of either Adgrd1 or PTK7 in glioblastomas, where they are expressed in adjacent cells, leads to a significant reduction of tumor growth [81]. The same group also developed antibodies that bind to the NTF of Adgrd1, leading to activation and an increase of cAMP levels in cells [82]. This mechanism is again cleavage dependent and an uncleavable mutant of Adgrd1 did not show activation by the same antibodies [82].

Overall, aGPCRs are a diverse group of receptors with a variety of different ECD. Because of this variety they are involved in diverse physiological processes. Here two examples were presented from the receptors we are working with in this project. The functional and mechanistic understanding of aGPCRs is still limited, indicating the need for biochemical, structural, and physiological studies of aGPCRs. My PhD work focused on the structural and biochemical characterization of the CTF from Adgrg6 and Adgrd1.



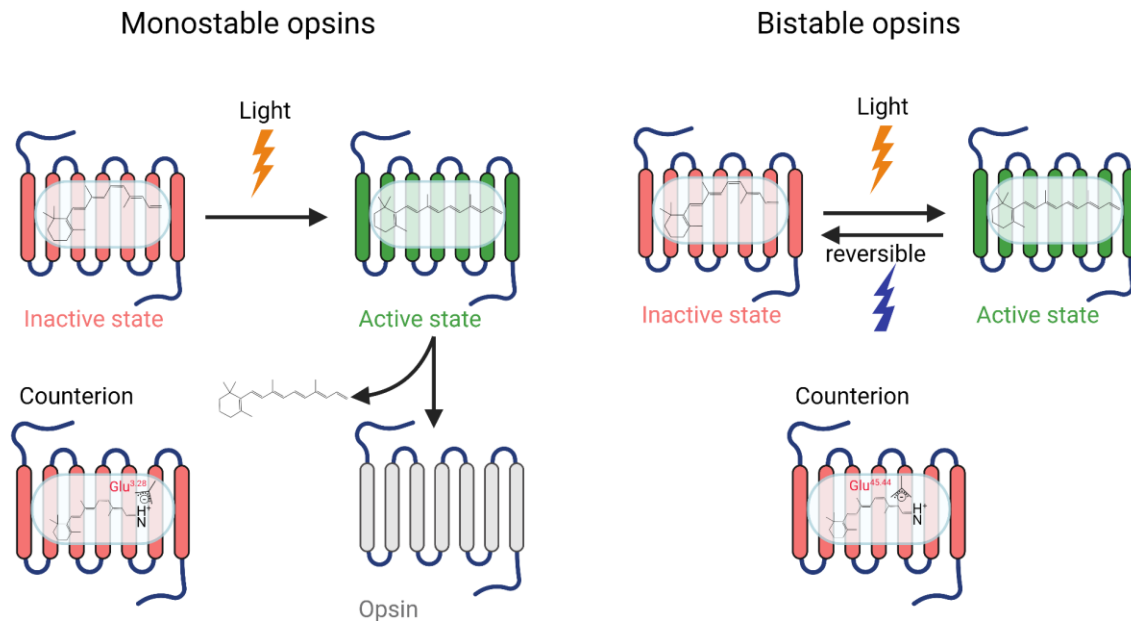
## 1.3 Opsins

Perception of light is one of the fundamental abilities in the animal kingdom. Even though various organisms evolved different vision apparatuses, the ability to sense photons is achieved by the same protein family, called opsins. Opsins can be grouped into eight subfamilies based on phylogenetics and their downstream G protein specificity: Group 1 – Gt coupled opsins (vertebrate visual and non-visual opsins); Group 2 – Gq coupled opsins (invertebrate visual opsins and melanopsin); Group 3 – invertebrate Go coupled opsins; Group 4 – Gi/Go coupled Opn3 (encephalopsin and TMT opsin); Group 5 – Gi coupled Opn5 (neuropsin); Group 6 – Gs coupled cnidarian opsins; Group 7 – retinochrome; and Group 8 – peropsins [83, 84]. While groups 1-6 act as light sensing GPCRs, groups 7 and 8 are retinal photoisomerases that produce 11-cis retinal [83, 84]. They can be further classified based on the functionality and distribution in two major cell types. Groups 1, 3, 4, 5 and 6 opsins are localized in ciliary-type photoreceptor cells and affect cyclic nucleotide signaling [84]. Group 2 opsins are localized in rhabdomeric-type photoreceptor cells and modulate the phosphoinositol signaling [84].

The ability of these receptors to sense light, stems from the light-sensitive ligand called retinal. Retinal is covalently bound to the receptors via a Schiff base linkage to the conserved lysine at position 7.43 (Ballesteros-Weinstein numbering [25]). The retinal molecule is a derivative of vitamin A. The majority of opsins bind retinal in the 11-cis conformation [85]. Opsins that have a retinal bound are called rhodopsins. The 11-cis retinal is an inverse agonist and rhodopsins bound to 11-cis retinal are in their inactive state. Upon photon absorption, the retinal isomerizes from its 11-cis form to the all-trans form which induces conformational changes in the receptor that lead to G protein binding [84-86]. Another important feature of rhodopsins is the presence of a counterion, a negatively charged residue in the orthosteric binding pocket, that stabilizes the protonated Schiff base [87, 88]. Monostable rhodopsins possess a proximal counterion at position 3.28 [87-89]. In bistable invertebrate rhodopsins, this position is occupied by a tyrosine and the negative charge is located at the distal counterion position 45.44 [30, 31, 88].

Mono- and Bistability are however not primarily classified based on the position of the counterion but based on the thermal stability of the Schiff base in the photoproduct. Monostable rhodopsins are “bleached”, meaning that they lose the all-trans retinal after photoactivation because the Schiff base linkage is deprotonated and hydrolyzed [84, 90]. On the other hand, bistable rhodopsins can revert to the 11-cis state after a second photon-isomerization event because of the thermal stability of the Schiff base in both the active and inactive states [84, 91]. **Figure 5** shows a schematic representation of monostable and bistable pigments.





**Figure 5: Biochemical properties of monostable and bistable pigments.** Monostable pigments, such as vertebrate visual opsins, lose their retinal after photoactivation because the photoproduct is thermally unstable. Monostable pigments have a proximal counterion (Glu) at position 3.28. The photoproduct of bistable opsins, such as invertebrate visual rhodopsins, is thermally stable and can revert to the dark state by a second photon absorption. Bistable pigments have a distal counterion at position 45.44. Figure created with Biorender.

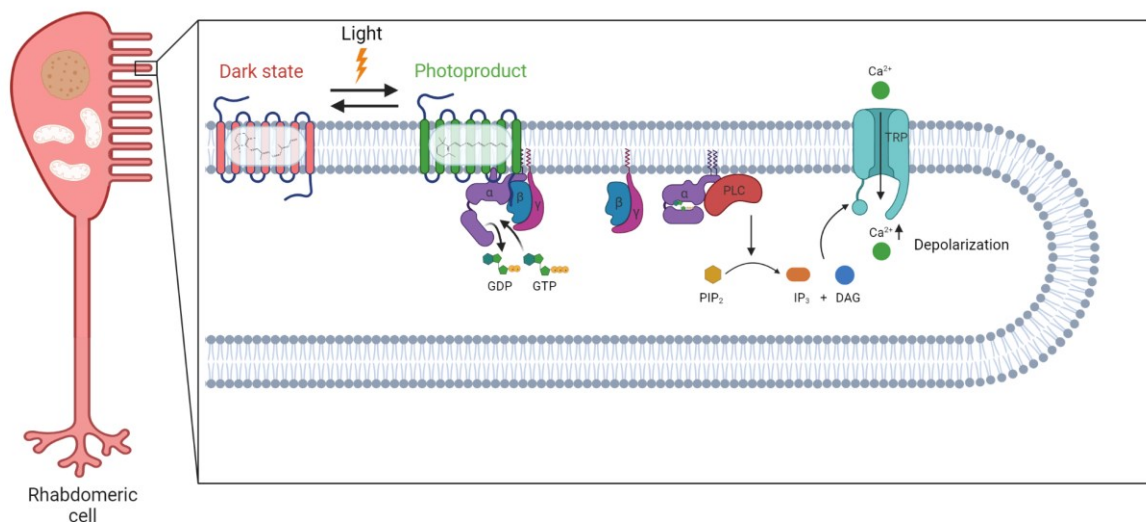
Vertebrate visual rhodopsins belong to the monostable pigments [83]. Bistable rhodopsins are essential not just for invertebrate vision but also for vertebrate non-visual light perception, which regulates for example the circadian rhythm (Melanopsin) [83, 84, 91]. In addition, bistable rhodopsins have a potential to be used as optogenetic tools to control G protein signaling pathways [84, 92-95]. Bistable rhodopsins allow precise manipulation by light because they can be switched between the active and inactive states. Mechanical and functional understanding of bistable rhodopsins is important for rational engineering of bistable rhodopsins in order to develop diverse optogenetic tools.

In my PhD work, I focused on a bistable invertebrate rhodopsin from the jumping spider (*Hasarius adansoni*) called jumping spider rhodopsin-1 (JSR1).

### 1.3.1 Jumping spider rhodopsin-1

The jumping spider rhodopsin-1 was discovered to be essential for the depth estimation of the jumping spider [96]. Nagata *et al.* further showed that JSR1 absorbs green light which is needed for accurate depth perception [96]. JSR1 is part of Group 2 – Gq coupled opsins and signals through the Gq pathway as confirmed by cell-based activity assays [83, 97, 98]. The best studied phototransduction pathway in invertebrates is from *Drosophila* [99]. Invertebrate visual rhodopsins are located in rhabdomeric photoreceptor cells and signal through the Gq pathway [91, 99]. Photoactivation results

in isomerization of 11-cis retinal to all-trans retinal, leading to structural changes in the receptors that results in the opening of a cytoplasmic cleft that allows the Gq protein to bind to the receptor. The active receptor catalyzes the exchange of GDP with GTP in the G $\alpha$  subunit of the Gq heterotrimer. The GTP bound G $\alpha$ q will subsequently dissociate from the G $\beta\gamma$  subunits and activate phospholipase C $\beta$  (PLC) which in turn hydrolyzes phosphatidylinositol 4,5-bisphosphate (PIP<sub>2</sub>) to diacylglycerol (DAG) and inositol 1,4,5 triphosphate (IP<sub>3</sub>). This leads to the opening of transient receptor potential (TRP) and TRP-like (TRPL) channels, allowing a Ca<sup>2+</sup> influx which leads to the depolarization of the cell [99] (**Figure 6**).



**Figure 6: Schematic representation of the proposed phototransduction process in invertebrate rhabdomeric cells.** Photon absorption leads to isomerization of the 11-cis retinal to the all-trans retinal, resulting in an active receptor. Active rhodopsin can activate the Gq protein which triggers a PLC mediated conversion of PIP<sub>2</sub> to IP<sub>3</sub> and DAG. These lead to the opening of the TRP/TRPL channels which allows an influx of Ca<sup>2+</sup>, leading to the depolarization of the membrane. Figure created with Biorender.

In contrast to the vertebrate phototransduction cascade where photoreceptor cells hyperpolarize because the channels close in response to light, the invertebrate photoreceptor cells depolarize in response to light [99].

As mentioned before, bistable opsins have the potential to be used as optogenetic tool. JSR1 has been shown to work as an optogenetic tool in zebrafish reticulospinal V2a neurons where activation of JSR1 lead to an increase of Ca<sup>2+</sup> and evoked a change in swimming behaviour [92]. The main issue with using JSR1 as an optogenetic tool is that the  $\lambda_{max}$  of the active and inactive states overlap, meaning that there will be a mixture of active and inactive states upon illumination [100]. One JSR1 mutant (JSR1 S199F) has been shown to have a  $\lambda_{max}$  of 380 nm in the inactive state (versus 505 nm for WT JSR1) and 540 nm in the active state (similar to WT JSR1) [98]. Mutational studies showed that a S199A mutation lead to a substantial decrease in the pKa of the Schiff base in the dark state but not in the photoproduct [98]. This hints that Ser199 is involved in the stabilization of the protonated Schiff base

in the dark state but not in the photoproduct. The disruption of the Schiff base-counterion link in the dark state of the S199F mutant leads to a deprotonated Schiff base, explaining the shift of  $\lambda_{\max}$  [98](Personal communication with Dr. Matthew Rodrigues, PSI). UV-light activation is not very desirable for in vivo applications because UV-light does not penetrate tissues well. Additional insights into the active state structure and interactions with coupling partners will be important for the rational engineering of JSR1. This could help designing JSR1 variants that can be activated with a desired wavelength and ideally can also be tailored to activate a specific G protein family.

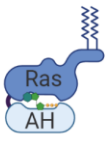
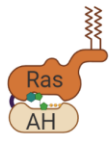
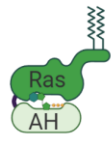



## 1.4 Heterotrimeric G protein

Today, it is known that GPCRs can interact with many downstream signaling partners. However, canonical GPCR signaling involves binding of the heterotrimeric G proteins. The G protein consists of three subunits called  $G\alpha$ ,  $G\beta$  and  $G\gamma$ . In the inactive state, the  $G\alpha$  is bound to GDP which stabilizes the heterotrimer [101]. An active GPCR acts as a guanine nucleotide exchange factor (GEF), meaning that the active receptor binds the GDP bound G protein and catalyzes the exchange of GDP with GTP in the  $G\alpha$  subunit. Binding of GTP induces the dissociation of the G protein from the receptor and subsequently also the dissociation of the  $G\alpha$  subunit and the  $G\beta\gamma$  dimer. These can activate downstream signaling partners that differ depending on the G protein type. The  $G\alpha$  subunit can hydrolyze GTP to GDP and the GDP bound form reassembles with the  $G\beta\gamma$  subunits to form the heterotrimeric G protein, which is ready for the next activation. Interestingly, humans have more than 800 genes encoding GPCRs but only 35 genes encoding G protein subunits, more specifically 16 genes encoding  $G\alpha$ , five genes encoding  $G\beta$  and 14 genes encoding  $G\gamma$  [101].

The 16 different  $G\alpha$  subunits can be divided into 4 subfamilies based on sequence similarity and functional properties:  $G\alpha_s$ ,  $G\alpha_i/o$ ,  $G\alpha_q/11$  and  $G\alpha_{12/13}$  (**Figure 7**) [101]. The  $G\alpha$  subunit is composed of two domains: the Ras and the  $\alpha$ -helical (AH) domain. The AH domain is flexible, allowing opening of the nucleotide binding site and the nucleotide can be exchanged from GDP to GTP [102]. In the nucleotide bound form the  $G\alpha$  is in a closed conformation even though it was recently shown that the local environment can lead to a range of different conformations (**Figure 1**) [102]. Huang *et al.* showed that the presence of the  $G\beta\gamma$  subunits and/or the plasma membrane can lead to different conformations than the fully closed state of the  $G\alpha$  subunit [102]. Each  $G\alpha$  subfamily is involved in binding specific downstream effector proteins that lead to different signaling outcomes.  $G\alpha_s$  mostly stimulates adenylyl cyclases (AC), leading to an increase of cytosolic cAMP level whereas the  $G\alpha_i/o$  inhibits ACs, therefore lowering the cytosolic cAMP level [101].  $G\alpha_q/11$  activates the phospholipase  $C\beta$  which produces DAG and  $IP_3$ , leading to an increase of the cytosolic  $Ca^{2+}$  level [101].  $G\alpha_{12/13}$  modulate the activity of small G-proteins from the Rho family, ultimately leading to remodeling of the

cytoskeleton [101]. The activity of the  $G\alpha$  subunit is not only regulated by GPCRs but also post-translational modifications or regulatory proteins. One example of a regulatory protein family is called regulator of G protein signaling (RGS). These proteins act as GTPase accelerating proteins (GAP), meaning they promote GTP hydrolysis in the  $G\alpha$  subunit thereby deactivating the  $G\alpha$  [103-105].  $G\alpha$  activity can also be inhibited by proteins that inhibit the guanine nucleotide exchange and there are also examples of proteins that promote  $G\alpha$  activity by acting as a GEF [105].

An example of a post-translational modification (PTM) is the phosphorylation of the  $G\alpha$ . Phosphorylation can happen by different kinases like the cAMP-dependent Ser/Thr protein kinase PKA, the Ser/Thr kinase PKC or tyrosine kinases [106-109]. Phosphorylation happens in different regions of the  $G\alpha$ , leading to specific modulation of the activity. For example, phosphorylation of Y356 in  $G\alpha_{11}$  leads to reduced receptor (M1) interactions whereas phosphorylation of S336 in  $G\alpha_{15}$  has an opposite effect and promotes coupling to the M2 or  $\beta_2$  adrenergic receptor [106, 107]. Besides phosphorylation there are other modifications like lipid modifications that can alter the activity [110].

Family	Gai/o	Gas	Gaq	Ga12/13	G $\beta$	G $\gamma$
						
Subtypes	Gai <sub>1-3</sub> , Gao, Gat <sub>1-3</sub> , Gaz	Gas, Gaolf	Gaq, Ga11, Ga14, Ga15/16	Ga12, Ga13	G $\beta_{1-5}$	G $\gamma_{1-13}$
Effector	Adenylyl cyclase Phosphodiesterase Phospholipase [cAMP] ↓	Adenylyl cyclase Axin [cAMP] ↑ PKA	PLC $\beta$ Lbc [Ca <sup>2+</sup> ] ↑ PKC Rho	P115-RhoGEF LARG PDZ-RhoGEF AKAP-Lbc Rho	Adenylyl cyclase PLC $\beta$ PI3K $\gamma$ Ion channels	

**Figure 7: Classification of  $G\alpha$ ,  $G\beta$  and  $G\gamma$  subunits.**  $G\alpha$  subunits are schematically depicted in different colors for each subfamily with the Ras and AH domains labelled. For each subfamily the  $G\alpha$  subtypes and downstream effector proteins are indicated.  $G\beta$  has five subtypes and  $G\gamma$  13 subtypes. Downstream effector proteins influenced by the  $G\beta\gamma$  dimers are indicated. Figure created with Biorender.

The  $G\beta\gamma$  subunits are obligatory dimers, believed to be assembled co-translationally [111]. Initially, the  $G\alpha$  subunit was the main focus of researchers and the  $G\beta\gamma$  subunits were thought to be less important for signaling. However, studies showed that the  $G\beta\gamma$  dimer can modulate the activity of various proteins. The  $G\beta\gamma$  dimer has been shown to influence downstream effectors also targeted by  $G\alpha$  subunits like ACs or the PLC $\beta$  [111-113]. In addition, the  $G\beta\gamma$  subunits interact with other downstream effectors like ion channels (Kir3 potassium channel or N-type calcium channel) or the soluble N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) [111, 114-116]. These are only a few examples of downstream effector proteins and in recent years the  $G\beta\gamma$  dimer gained more attention again from researchers. Because of the different subtypes and combinations of  $G\beta$  and

G $\gamma$  subunits there are more than 60 possible combinations. Depending on the subtype, the G $\beta\gamma$  dimer can interact with different effectors and render different signaling outcomes.

In GPCR-G protein structural studies the G $\alpha$  subunit is highlighted more often because most interactions with the GPCR are made with the G $\alpha$  subunit and only very few cases showed interactions with the G $\beta\gamma$  subunits [117, 118].

## 2 Aims of the thesis

Even though, I have been working with GPCRs from two different families with very different functions, the overall goal was similar in both cases. The main goal was to determine the structure of the active state receptor coupled to a G protein with single particle cryo-EM. In the adhesion GPCR project I was working with Adgrg6 and Adgrd1 and in the JSR1 project I was working with the jumping spider rhodopsin 1 (JSR1).

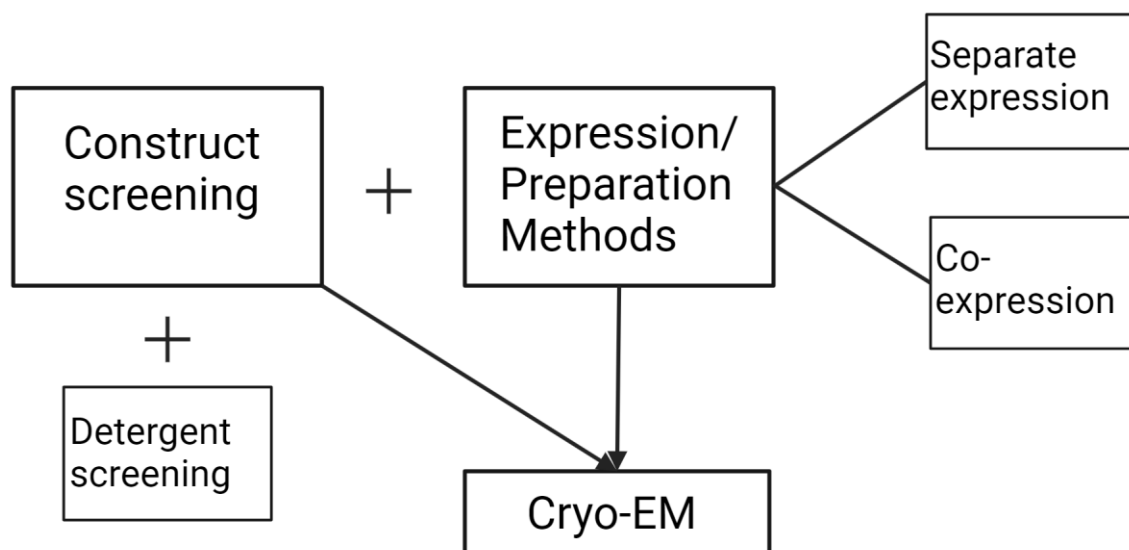
In the aGPCR project, this involves many smaller intermediate goals, all leading towards the biochemical characterization of Adgrg6 and Adgrd1 which is a prerequisite to prepare a sample for structure determination. A suitable sample will be analyzed by single particle cryo-EM for structure determination.

In the JSR1 project, the main focus is to optimize the preparation of the signaling complex of JSR1 with a G protein heterotrimer. Structural analysis is performed with single particle cryo-EM. The structure is analyzed with the most recent resources and in the context of the latest literature to form comprehensive comparisons and conclusions. This structure presents the first active state structure of a bistable rhodopsin.

## 2.1 Experimental strategy

### 2.1.1. Adhesion GPCR project

Structural information about aGPCRs was limited to the extracellular domains at the time when I started my PhD. The goal is to solve a structure of the CTF of the aGPCRs Adgrd1 and Adgrg6, activated by the tethered agonist. This project is a new project in our research group. As mentioned before, this means that many experiments have to be done before a sample for cryo-EM analysis can be produced. Our experimental strategy starts with designing several constructs for Adgrd1 and Adgrg6 focused on the CTF. Constructs are designed for either HEK or insect cell expression. Expression is evaluated by in-gel fluorescence imaging, western blot analysis or analytical fluorescence size exclusion chromatography. For selected constructs a detergent screening is done to identify a suitable detergent for both Adgrd1 and Adgrg6. In the next step, different strategies to express and purify a receptor-G protein complex are explored. In one method, we express the receptor and G protein separately and purify each component of the complex separately. The purified components are then mixed to form the complex and prepare a sample for cryo-EM analysis. In the other method, we co-express the receptor and G protein in the same cell. Activating the receptor in the cell leads to complex formation in the plasma membrane. The receptor is therefore stabilized in the active state by the interaction with the G protein. Next, the complex is extracted from the plasma membrane by detergents and purified in order to prepare cryo-EM grids. **Figure 8** shows a schematic with the different steps.



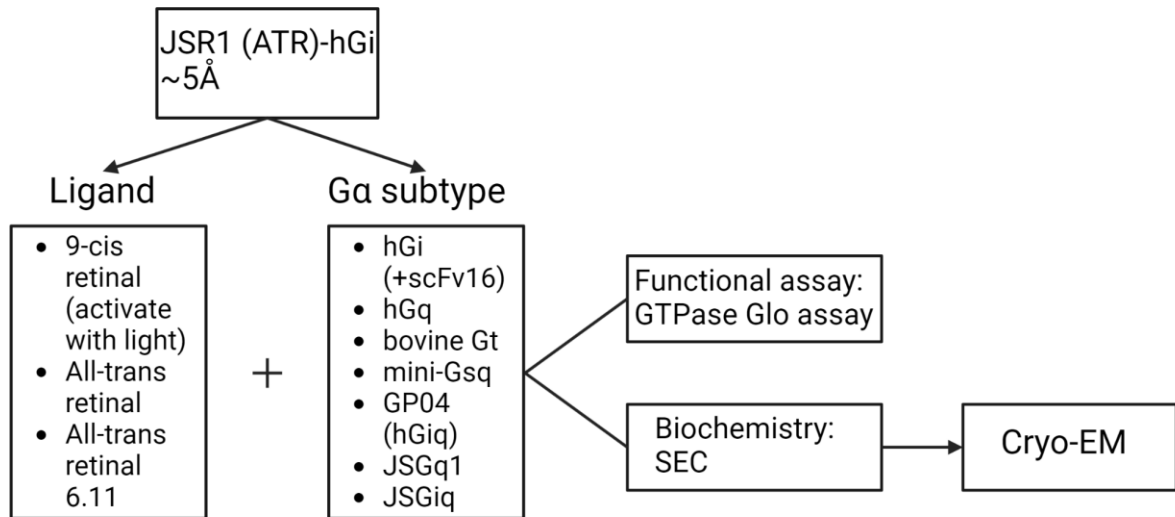
*Figure 8: Schematic showing the different steps to produce a sample for structure determination by cryo-EM.*

### 2.1.2. Jumping spider rhodopsin-1 project

Our group showed before that JSR1 can be expressed in HEK293GnTI<sup>-</sup> cells, possesses high thermal stability in detergents and can be crystallized in the inactive state bound to the inverse agonist 9-cis retinal [30]. The 2.1 Å crystal structure provided insights into the inactive state but insights into the active state structure are needed to elucidate the mechanism how the bistability of retinal is achieved and gain information on the G protein coupling. Additional insights into the protonated Schiff base (PSB) – counterion (CI) link in the active and inactive states of JSR1 were discovered by Nagata *et al.* [98]. Dr. Filip Pamula, a former PhD student in our lab, worked towards the structure determination of a JSR1-human G<sub>i</sub> complex (JSR1-hG<sub>i</sub>). He succeeded in reconstructing a cryo-EM map and determining a structure to ~5 Å resolution. Several reasons prevented determination of a higher resolution structure. First, JSR1 bound to the 9-cis retinal needs to be light activated to couple to the G protein and even though the  $\lambda_{\max}$  of the active and inactive states are shifted (9-cis vs all-trans JSR1), they still overlap, leading to a mixed population of active/inactive state JSR1 upon illumination [100]. Second, the low stability of the JSR1-hG<sub>i</sub> complex resulted in dissociation during freezing for grid preparation, lowering the number of complex particles in the EM data. Lastly, the flexibility of the complex resulted in heterogeneity in the particle set, lowering the overall resolution of the cryo-EM map.

My goal is to prepare an optimized sample of a JSR1-G protein complex to determine a high resolution structure. To achieve this goal we can change two main components of the complex. First, we can change the ligand for JSR1. As mentioned already, the  $\lambda_{\max}$  values for the active and inactive state overlap. To overcome this problem, we try to purify JSR1 bound to either the all-trans retinal or the non-natural agonist all-trans retinal 6.11. This would allow us to have a homogeneous population of active state JSR1 and could overcome the issue with the overlapping  $\lambda_{\max}$  values. The cryo-EM analysis of the JSR1-hG<sub>i</sub> complex showed that the complex is not very stable, resulting in many dissociated particles in the cryo-EM dataset. I will test if JSR1 can couple and activate different G $\alpha$  subtypes from the G<sub>i/o</sub> and G<sub>q/11</sub> families which might form a more stable complex than the JSR1-hG<sub>i</sub> complex. Complex formation is analyzed by analytical size exclusion chromatography. Furthermore, G protein subtypes showing binding to JSR1 will further be analyzed for activation by JSR1 with an *in vitro* activity assay called GTPase Glo assay. In addition, I will try to add the scFv16 to the JSR1-hG<sub>i</sub> complex, which was shown to stabilize the bovine rhodopsin-hG<sub>i</sub> complex [119]. Promising JSR1-G protein complexes are then analyzed by cryo-EM for structure determination. **Figure 9** shows a summary of the experimental strategy to prepare a optimized JSR1 signaling complex for structure determination by cryo-EM.





**Figure 9: Schematic summarizing the experimental strategy to produce an optimized JSR1 signaling complex for structure determination by single particle cryo-EM.** Abbreviations: ATR= all-trans retinal; hGx= human G protein from subtype x; JSGq1= jumping spider visual Gq1; JSGiq= chimeric human/jumping spider Giq; SEC= size exclusion chromatography.

## 3 Material and Methods

### 3.1 Bacterial cell transformation

To copy and isolate plasmid DNA or generate single colonies for expression, Mach-1 or BL21 (DE3) strains of *Escherichia coli* (*E. coli*) cells were used respectively. One  $\mu\text{l}$  of plasmid DNA of interest at a concentration of 100 ng/ $\mu\text{l}$  was added to 100  $\mu\text{l}$  of competent cells. After 30 min incubation on ice, heat shock was performed in a water bath at 42°C for 45 seconds. Afterwards, the mixture was incubated for 1 min on ice, followed by the addition of 400  $\mu\text{l}$  of SOC medium and incubation for 1 hour at 37°C while shaking at 600 rpm in a Infors HT shaker (Infors AG). Cells were subsequently plated on LB Agar plates supplemented with either kanamycin or ampicillin, depending on the antibiotic resistance gene on the plasmid DNA. Plates were incubated overnight at 37°C and single colonies were picked to inoculate liquid cultures for further steps.

### 3.2 Bacterial cell expression

Expression of Gai (Uniprot ID: P36096), mini-Gas 393/399 [120], 3C protease (Uniprot ID: P03303 and chain ID: PRO\_0000426549) and the Anti-GFP Nanobody (NbGFP) (Gift from Prof. Volodymyr Korkhov) was done in BL21 (DE3) cells. Single colonies from LB Agar plates were picked to inoculate 100 ml of LB (or TB cultures for NbGFP) supplemented with either Ampicillin (100  $\mu\text{g}/\text{ml}$ ) or Kanamycin (50  $\mu\text{g}/\text{ml}$ ). Cultures were grown overnight at 37°C while shaking at 150 rpm. Cell suspensions were used to inoculate large volume LB or TB cultures. Cells were grown at 37°C until they reached an Optical Density (OD) of 0.6-0.8 and protein expression was induced by addition of 0.5 mM isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG). Cells were further grown overnight at 23°C and shaking at 150 rpm. Cells were harvested by centrifugation at 4000 rpm for 15 min and subsequent removal of excess liquid. Cell pellets were flash frozen in liquid nitrogen before storing them at -80°C.

### 3.3 Gai purification

The Gai subunit contained a C-terminal TEV cleavage site followed by a 10xHis-tag. The *E. coli* cell pellet was thawed and resuspended in buffer A (25 mM Tris-HCl pH 7.5, 500 mM NaCl, 50 mM Imidazole, 10% glycerol and 5 mM 2-mercaptoethanol ( $\beta$ -ME)). Cells were disrupted by sonication (2 sec on/ 9 sec off) for two times 3 min at 100 % amplitude on ice (Sonics VCX-500-220 Vibra-Cell). Cell lysate was centrifuged for 30 min at 185'000 rcf and 4°C. The supernatant was loaded on 2x5 ml HisTrap Fast Flow (FF) columns. The columns were washed with 30 column volumes (CV) of buffer A and elution was done with 5x0.5 CV of buffer B (25 mM Tris-HCl pH 7.5, 500 mM NaCl, 500 mM Imidazole, 10% glycerol and 5 mM  $\beta$ -ME). Eluent was collected and supplemented with TEV protease at a ratio of 20:1 (Gai:TEV; w/w) before being dialysed against 1500 ml of buffer C (25 mM Tris-HCl pH 7.5, 150 mM NaCl, 35 mM Imidazole, 10% glycerol and 10 mM  $\beta$ -ME) using a 10 kDa cut-off dialysis membrane. After dialysis, the sample was passed through another 5 ml HisTrap FF column to remove His-tagged TEV protease and other protein contaminants. The flow through was concentrated with a 10 kDa cut-off Amicon Ultra Centrifugal Filter (Millipore) and injected to a Superdex 200 HiLoad 16/60 column (GE Healthcare) equilibrated in buffer D (25 mM HEPES pH 7.5, 150 mM NaCl, 2mM DTT). Peak fractions containing Gai were collected, pooled and concentrated before adding glycerol to a final concentration of 20% and flash freezing in liquid nitrogen. The final yield was  $\sim$ 15 mg/L and the protein was stored at -80°C until use.

### 3.4 3C protease purification

The 3C protease contained a C-terminal 10xHis-tag. Bacterial cell pellet was thawed and resuspended in buffer A (25 mM HEPES pH 7.5, 500 mM NaCl, 50 mM Imidazole, 10 % glycerol and 5 mM  $\beta$ -ME). Cells were disrupted by sonication (1 sec on/1.5 sec off) for 10 min at 35% amplitude on ice (Sonic VCX-500-220 Vibra-Cell). Cell lysate was centrifuged at 208'000 rcf for 30 min. Supernatant was loaded onto 2x5 ml HisTrap FF columns and columns were washed with 20 CV of buffer A. Elution was done with 6 CV of buffer B (25 mM HEPES pH 7.5, 500 mM NaCl, 200 mM Imidazole, 10 % glycerol and 5 mM  $\beta$ -ME) and the eluent was dialysed against 2 L of buffer C (25 mM HEPES pH 7.5, 500 mM NaCl, 10 % glycerol and 5 mM  $\beta$ -ME) overnight at 4°C. The sample was concentrated, supplemented with 20% glycerol and flash frozen in liquid nitrogen. The protein was stored at -80°C until use.

### 3.5 GFP Nanobody purification

The GFP Nanobody contained a 10xHis-tag. The cell pellet was thawed and resuspended in buffer A (25 mM HEPES pH 7.5, 200 mM NaCl and 20 mM Imidazole). The cell suspension was homogenized with a blender and cells were mechanically disrupted with a Microfluidizer (Microfluidics, IDEX). Cell lysate was centrifuged at 200'000 rcf for 1 h at 4°C and the supernatant was loaded onto a 5 ml HisTrap FF column. The column was washed with 10 CV buffer A and 10 CV buffer B (25 mM HEPES pH 7.5, 200 mM NaCl and 40 mM Imidazole) before eluting the protein with 15 CV of buffer C (25 mM HEPES pH 7.5, 200 mM NaCl and 500 mM Imidazole). The eluent was concentrated and injected to a Superdex75 16/60 (GE Healthcare) equilibrated in buffer D (25 mM HEPES pH 7.5, 200 mM NaCl). Peak fractions containing the GFP nanobody were collected and concentrated before being flash frozen in liquid nitrogen and stored at -80°C until use.

### 3.6 Mini-G $\alpha$ s 393/399 purification

Mini-G $\alpha$ s 393/399 contain a N-terminal 6xHis-tag followed by a TEV cleavage site. Purification was done according to a published protocol [121]. In short, the cell pellet was resuspended in buffer A (40 mM HEPES pH 7.5, 100 mM NaCl, 10 mM Imidazole, 10% glycerol, 5 mM MgCl<sub>2</sub> and 50  $\mu$ M GDP) supplemented with cOmplete protease inhibitor cocktail tablets (Roche) and 100  $\mu$ M DTT. The cell suspension was stirred at 4°C for 30 min before lysing the cells by sonication for 10 min with pulses of 2 sec on and 4 sec off at an amplitude of 70%. The lysate was centrifuged at 40'000 rcf for 45 min at 4°C. Supernatant was filtered using a 0.22  $\mu$ m Steritop filter before loading the filtered supernatant onto a HisTrap FF column, pre-equilibrated with buffer A. The column was washed with 10 CV of buffer B (20 mM HEPES pH 7.5, 500 mM NaCl, 40 mM Imidazole, 10% glycerol, 1 mM MgCl<sub>2</sub> and 50  $\mu$ M GDP) and the protein was eluted with 3 CV of buffer C (20 mM HEPES pH 7.5, 100 mM NaCl, 500 mM Imidazole, 10% glycerol, 1 mM MgCl<sub>2</sub> and 50  $\mu$ M GDP). TEV protease was added to the eluent at a ratio of 1:20 (w/w, TEV:mini-G $\alpha$ s) and subsequently dialyzed against buffer D (20 mM HEPES pH 7.5, 100 mM NaCl, 10% glycerol, 1 mM MgCl<sub>2</sub> and 10  $\mu$ M GDP) overnight. The dialysate was supplemented with Imidazole to a final concentration of 20 mM and loaded onto a HisTrap FF column. The flow through was collected and pooled with 1 CV wash fraction with buffer D. The protein solution was concentrated and injected into a Superdex 200 HiLoad 16/600 (GE Healthcare) pre-equilibrated with buffer E (10 mM HEPES pH 7.5, 100 mM NaCl, 10% glycerol, 1 mM MgCl<sub>2</sub>, 1  $\mu$ M GDP and 0.1 mM TCEP). Peak fractions containing the mini-G $\alpha$ s were pooled, concentrated and flash frozen in liquid nitrogen before storage at -80°C.

### 3.7 G $\beta_1\gamma_1$ purification from bovine retinas

The G $\beta_1\gamma_1$  subunits were separated from the transducin G protein heterotrimer which was purified from the rod outer segment membranes of bovine retinas as previously described [122, 123].

In short, bovine retinas (W L Lawson Company, Omaha, NE) were exposed to light in order to activate rhodopsin in the rod outer segment (ROS) membranes. Activated rhodopsin binds the transducin G protein. Bovine retinas are homogenized in a buffer containing 47% sucrose to apply an osmotic shock. ROS membranes are isolated with a sucrose gradient where the ROS membranes are located and collected at the interface of the 30% sucrose bottom layer and 25% sucrose top layer. ROS membranes were washed and the transducin G protein was finally dissociated from the ROS membranes by adding GTP to the buffer. The protein was dialysed to remove the GTP and frozen until further use. Three samples from 150 bovine retinas each were combined, and the combined sample was applied to a Blue Sepharose column in order to separate the G $\beta_1\gamma_1$  subunits from the G $\alpha$ t subunit. The relevant fractions were combined and concentrated before being flash frozen in liquid nitrogen and stored at -80°C.

### 3.8 Insect cell lines

Sf9 (*Spodoptera frugiperda*) and High Five (*Trichoplusia ni*) insect cells were obtained from Thermo Fisher Scientific. Both cell lines were cultured in Sf-900™ III serum-free medium (Thermo Fisher Scientific) at 27°C and 120 rpm. HEK293F cells were cultured in suspension with freestyle medium (Thermo Fisher Scientific) at 37°C with 5 % CO<sub>2</sub>.

### 3.9 Baculovirus production

High-titer recombinant baculovirus stocks were generated using the flashBAC™ expression system in Sf9 cells. 1.5 µg plasmid DNA was mixed with 500 ng linearized BAC10:KO1629 DNA (Oxford Expression Technologies Ltd.) and 8 µl Cellfectin™ II reagent (Thermo Fisher Scientific) in 200 µl Sf-900™ III serum-free medium. After a 20 min incubation at RT, the transfection mix was added to 1 x 10<sup>6</sup> Sf9 cells in an adherent culture plate. The plate was incubated at 27°C for 120 h. The P<sub>0</sub>-generation virus was harvested from the supernatant and used to infect a 10 ml suspension cell culture at 1 x 10<sup>6</sup> cells/ml density. Supernatant with P<sub>1</sub>-generation virus was harvested by centrifugation after 72 h incubation while shaking at 27°C. The procedure was repeated with a 100 ml cell suspension volume to generate the final P<sub>2</sub>-generation virus for protein expression.

### 3.10 Expression in insect cells

For small scale expressions, Sf9 and High Five cells were cultured to a density of 3 x 10<sup>6</sup> cells/ml and transferred to a 96 deep-well plate (DWP) and infected with virus. Different virality of infection (VOI; v/v) were tested for each construct. The 96 DWP was incubated at 27°C at 1200 rpm on a shaker plate (Eppendorf MixMate) and cells were harvested after 48 h or 72 h by centrifugation at 800 rcf and stored at -80°C until further use.

For large scale expressions, Sf9 or High Five cells were cultured to a density of 4 x 10<sup>6</sup> cells/ml. Cells were harvested by centrifugation at 800 rcf and 15 min and cells were resuspended in fresh Sf-900™ III serum-free medium to the same density. Baculovirus for the protein of interest was added to infect the cells. In case of the co-expression of the aGPCR and G protein, two baculoviruses (one for the

aGPCR and one for the G protein) were added with the same VOI. Cells were harvested after 48 h or 72 h by centrifugation at 800 rcf, and cell pellets were stored at -80°C.

### 3.11 G $\beta_1\gamma_2$ expression and purification from insect cells

The G $\beta_1\gamma_2$  subunits were cloned into the pAC8REDNK vector (Kind gift of Arnaud Poterszman [124]) for insect cell expression. The G $\beta_1$  subunit contained a N-terminal 10xHis-tag followed by a 3C protease cleavage site. The G $\gamma_2$  subunit was not further modified. Expression was done in High Five insect cells at a VOI of 3% for 48h. The cell pellet was harvested by centrifugation at 800 rcf for 15 min, and stored at -80°C until use.

The cell pellet was thawed and resuspended in buffer A (20 mM HEPES pH 7.5, 100 mM NaCl, 3 mM MgCl<sub>2</sub>, 100  $\mu$ M EDTA and 5 mM  $\beta$ -ME) supplemented with cOmplete protease inhibitor cocktail tablets (Roche) and DNase 1 (10  $\mu$ g/ml) (Roche). Cells were mechanically disrupted using a Microfluidizer (Microfluidics, IDEX) and subsequently the suspension was centrifuged at 185'000 rcf for 30 min. The pellet was resuspended in a buffer containing 20 mM HEPES, 100 mM NaCl, 20 mM Imidazole, 3 mM MgCl<sub>2</sub>, 5 mM  $\beta$ -ME and 20  $\mu$ M GDP supplemented with cOmplete protease inhibitor cocktail tablets and solubilized by adding 1% Sodium cholate (Sigma-Aldrich) and 0.05 % (w/v) DDM (Anatrace). After incubation at 4°C for 1 h, insoluble material was removed by centrifugation at 185'000 rcf for 45 min. The clarified supernatant was incubated with 5 ml of TALON resin (Takara Bio) for 90 min at 4°C. The resin was collected in a gravity flow column and the detergent was gradually exchanged to 0.1 % DDM. The resin was washed with 10 CV of IMAC buffer A (20 mM HEPES pH 7.5, 100 mM NaCl, 20 mM Imidazole, 1 mM MgCl<sub>2</sub>, 5 mM  $\beta$ -ME, 20  $\mu$ M GDP and 0.01 % DDM) and 10 CV of IMAC buffer B (IMAC buffer A with 300 mM NaCl). The protein was eluted with IMAC buffer A supplemented with 300 mM Imidazole. His-tagged HRV 3C protease was added (1:50 w/w; HRV 3C protease:G $\beta_1\gamma_2$ ) to the eluent and dialyzed against dialysis buffer (20 mM HEPES pH 7.5, 100 mM NaCl, 1 mM MgCl<sub>2</sub>, 100  $\mu$ M TCEP, 20  $\mu$ M GDP and 0.05 % DDM) overnight at 4°C. Dialysate was collected and supplemented with Imidazole to a final concentration of 20 mM before incubation with 5 ml TALON resin for 1 h at 4°C. The resin was transferred to a gravity flow column and the flow through was collected. The resin was washed with one CV of dialysis buffer supplemented with 20 mM Imidazole. The flow through and wash were combined and concentrated using a 30 kDa cut-off Amicon Ultra Centrifugal Filter (Millipore) before adding 20 % glycerol (v/v). The sample was flash frozen in liquid nitrogen and stored at -80°C.

### 3.12 Peptide synthesis

The Stachel-sequence derived Adgrd1 peptide (TNFAILMQVVPLE) and Adgrg6 peptide (THFGVLMDLPRSASQL) were synthesized at Peptide 2.0 (Chantilly, VA, USA). Additionally, FLAG epitope peptide (DYKDDDK) was synthesized at Peptide 2.0 as well.

### 3.13 Adhesion GPCR Construct Design

The cDNA encoding the human Adgrd1 (Isoform 1, Uniprot ID: Q6QNK2-1) was used to generate all truncated Adgrd1 sequences except for the D1.05 (full-length isoform 2, Uniprot ID: Q6QNK2-2) and D1.06 (full-length isoform 3, Uniprot ID: Q6QNK2-3) constructs. The cDNA encoding the human Adgrg6 (Isoform 1, Uniprot ID: Q86SQ4-1) was used to generate all Adgrg6 constructs. In section "8.1 Tables" the sequences of all constructs can be seen. The HEK cell constructs were C-terminally fused to a HRV 3C cleavage sequence (LEVLFG) followed by a YFP (Uniprot ID: A0A059PIR9), a 8xHis tag and a

TwinStrep tag (SAWSHPQFEKGGGSGGGGSSAWSHHPQFEK). The insect cell expression constructs started with a N-terminal haemagglutinin (HA) signal sequence (MKTIIALSYIFCLVFA) followed by a FLAG-tag (DYKDDDDA), a tobacco etch virus (TEV) protease cleavage site (ENLYFQG) and the truncated Adgrd1/Adgrg6 sequence. The C-terminus of Adgrd1/Adgrg6 was fused to a TEV protease cleavage site followed by a GFP (Uniprot ID: P42212) and a 8xHis tag. G6.10 and D1.12, designed for the co-expression experiments, started with a HA signal sequence followed by the Adgrd1/Adgrg6 CTF, a TEV protease cleavage site and a FLAG-tag (DYDKDDDDK). Constructs G6.11/.12 and D1.13/.14/.15, designed for co-expression experiments, started with the HA signal sequence followed by a FLAG-tag, a TEV protease cleavage site and the corresponding Adgrd1/Adgrg6 sequence. All insect cells constructs were cloned into the pAC8REDNK vector (Kind gift of Arnaud Poterszman [124]).

### 3.14 G protein construct design for insect cell expression

An engineered shortened human G $\alpha$ s subunit (mini-G $\alpha$ s 399)[120], which lacks the  $\alpha$ -helical domain, was further modified with an N-terminal Strep-tag (WSHPQFEK) followed by a GFP and a HRV 3C protease cleavage sequence (LEVLFG). The human G $\beta_1$  was modified with an N-terminal 10xHis tag followed by a HRV 3C cleavage sequence. The human G $\gamma_2$  was not further modified. All three G protein subunits were cloned into the pAC8REDNK vector for insect cell expression. For the human Gq (hGq), human G11(hG11), jumping spider Giq (JSGiq) chimera and the jumping spider visual Gq1 (JSGq) proteins the same vector was used as for the mini-Gs heterotrimer but the G $\alpha$  sequence of mini-G $\alpha$ s-GFP was replaced with hG $\alpha$ q, hG11, JSG $\alpha$ iq or JSG $\alpha$ q without a GFP fusion.

### 3.15 Expression in HEK293F cells

HEK293F cells were cultured to a density of  $2 \times 10^6$  cells/ml in Freestyle 293 Expression Medium (Thermo Fisher Scientific). For the transient transfection of a 1 L cell culture, a mixture of 50 ml of Opti-MEM™ (Gibco) with 2.5 mg of PEI (Sigma-Aldrich) was created, homogenized and incubated at 37°C until the mixture is completely transparent. This solution was mixed with a mixture of 50 ml Opti-MEM™ with 1 mg sterile DNA. After incubation of 20 minutes at room temperature, the solution was directly added to the cell suspension. After incubation at 37°C and 5% CO<sub>2</sub> for 48 or 72 h, the cells were harvested by centrifugation and frozen at -80°C.

### 3.16 In-gel fluorescence scanning

SDS-PAGE gels were imaged using a ChemiDoc Imaging System (Bio-Rad Laboratories, Inc.) with the Alexa 488 option to visualize the YFP or GFP coupled to the protein of interest.

### 3.17 Western blot analysis

SDS-PAGE was run with the samples of interest. Instead of Coomassie staining the gels, the gels were subjected to a semi-dry blot using a Trans-Blot® SD Semi-Dry Transfer Cell (Bio-Rad Laboratories, Inc.) with Polyvinylidene difluoride membranes. The membrane was placed on top of the anode and the gel was placed on top of the membrane, sandwiched between two pre-soaked filter papers. The cathode was placed on top, and the transfer was run at 15 V for 35 min. After the transfer, the membrane was soaked in 3% bovine serum albumin made with 1xTBST (Tris-buffered saline, 0.1% Tween20) buffer for at least 1h. The membrane was washed with 1xTBST buffer three times for 3min. The primary antibody solution was prepared using an anti-FLAG or anti-Strep tag antibody coupled to



HRP (0.1-0.2 µg/ml) in a 1xTBST buffer with 3 % BSA. The membrane was incubated with this solution for at least 2h at 4°C. The membrane was then washed three times with 1xTBST buffer for 3 min each time. The membrane was then incubated for 5 min with SuperSignal West Pico PLUS solution (Thermo Fisher Scientific) and then imaged for chemiluminescence. For the detection of the Adgrd1-NTF a primary antibody was used against the Adgrd1-NTF and a secondary antibody coupled to HRP was added, recognizing the heavy chain of the primary antibody.

### 3.18 Detergent screening

Cell pellets from the 96 DWP were resuspended in buffer containing 20 mM HEPES pH 7.5, 100 mM NaCl, 1 mM MgCl<sub>2</sub>, DNase I (10 µg/ml) (Roche) supplemented with cOmplete protease inhibitor cocktail tablets (Roche). Next 1% (v/w) of the respective detergent used for the screening experiment was added and the samples were incubated for 1 h at 4°C. Insoluble material was removed by centrifugation for 10 min at 200'000 rcf. The clarified supernatants were injected to a Superdex 200 increase 5/150 column (GE Healthcare) pre-equilibrated with 20 mM HEPES pH 7.5, 100 mM NaCl, 1 mM MgCl<sub>2</sub> and 0.05% DDM. The GFP fluorescence signal was measured and the chromatograms analyzed using Excel (Microsoft).

### 3.19 Expression of Adgrd1-NTF

The cDNA of the N-terminal fragment (NTF) of Adgrd1 (residues 1-544; Uniprot ID: Q6QNK2-1) was fused to a C-terminal 8x-His tag and cloned into the pAC8REDNK vector for insect cell expression. High Five insect cells were grown to a density of 3x10<sup>6</sup> cells/mL and baculovirus for Adgrd1-NTF was added to reach a VOI of 3%. Because the Adgrd1-NTF is a secreted protein, the supernatant of the cell suspension was either harvested after 30 h, 48 h or 72 h and directly used for purification.

### 3.20 Purification of Adgrd1-NTF

The pH of the supernatant was adjusted by adding Tris pH 7.5 to a final concentration of 50 mM and incubating the solution at 4°C for 30 min while stirring. The supernatant was then clarified by filtration using a 0.45 µm filter paper and the clarified supernatant was loaded onto a HisTrap Excel (Cytiva) column overnight in a loop. The column was washed with 10 CV of buffer A (20 mM HEPES pH 7.5, 500 mM NaCl and 20 mM imidazole) and 15 CV buffer B (20 mM HEPES pH 7.5, 100 mM NaCl and 20 mM Imidazole). The protein was eluted with 10 CV of elution buffer (20 mM HEPES pH 7.5, 100 mM NaCl and 500 mM Imidazole). The eluent was concentrated using a 30 kDa cut-off Amicon Ultra Centrifugal Filter (Millipore) and injected into a Superdex200 10/300 (GE Healthcare) column equilibrated with SEC buffer (20 mM HEPES pH 7.5 and 100 mM NaCl). Appropriate fractions were pooled and concentrated before flash freezing the protein in liquid nitrogen or direct usage in crystallization trials and thermal stability measurements.

### 3.21 Thermal stability measurements Adrgd1-NTF

Thermal stability assays were performed with a Rotor-Gene Q RT-PCR machine (QIAGEN) and the SYPRO orange dye (Sigma Aldrich). The following protocol was used for the assays:

1. Hold at 25°C for 2 min
2. Melt – 25°C to 90°C, 0.4 degrees per cycle, 10 seconds each cycle

The final reaction volume is 25  $\mu$ l in PCR tubes. First, a screen was done to find an optimal protein and SYPRO orange concentrations. 1-10  $\mu$ M protein concentrations were tried and 1x, 2x, 5x and 10x SYPRO concentrations. For all subsequent assays a protein concentration of 7  $\mu$ M and a SYPRO orange concentration of 2x was used. For the buffer screening the protein was concentrated that only a small volume (2.6  $\mu$ l;  $\sim$ 10% of total volume) has to be added to the tubes and the volume is mainly made of the buffer to screen. Various buffers were screened and can be seen in **Figure 28**.

### 3.22 Crystallization Adgrd1-NTF

Two samples were used for crystallization. One sample had a protein concentration of 12 mg/ml and the other 17 mg/ml. Crystallization screens were set up for the sitting drop vapor diffusion technique using a mosquito crystallization robot (SPT Labtech Ltd). Drops were set up at a 1:1 ration of protein solution to reservoir buffer solution (100 nl : 100 nl). The following 96-well plate crystallization screens were set up: Morpheus, Morpheus II, PACT premier, SG1, JCSG plus (Molecular Dimensions, Calibre Scientific) and XP screen (Jena Bioscience GmbH). The plates were stored at either 20°C or 4°C in a ROCK IMAGER<sup>®</sup> (FORMULATRIX) and regularly checked for crystal growth.

### 3.23 Complex formation and analytical SEC

Cells expressing Adgrd1, Adgrg6 or mini-Gs heterotrimer were resuspended in 20 mM HEPES pH 7.5, 100 mM NaCl, 1 mM MgCl<sub>2</sub>, DNase I (10  $\mu$ g/ml) (Roche) supplemented with cOmplete protease inhibitor cocktail tablets (Roche). Resuspended cells expressing receptor were mixed with cells expressing mini-Gs heterotrimer and the suspension was supplemented with 1% (w/v) lauryl maltose neopentyl glycol (LMNG, Anatrace), 0.1% (w/v) cholesteryl hemisuccinate (CHS, Anatrace) and 25 mU/ml apyrase. The samples with synthesized Stachel peptide also contained 200  $\mu$ M of the corresponding peptide. Samples were incubated for 2h at 4°C and insoluble material was removed by centrifugation at 200'000 rcf for 15 min. Solubilized samples were injected into a TSKgel<sup>®</sup> column (Tosoh Bioscience LLC) pre-equilibrated with 20 mM HEPES pH 7.5, 100 mM NaCl, 1 mM MgCl<sub>2</sub>, 0.01% LMNG/CHS (10:1) and the GFP fluorescence was measured.

### 3.24 Complex purification Adrgd1-mini Gs heterotrimer

Frozen cell pellet was resuspended in buffer containing 20 mM HEPES pH 7.5, 100 mM NaCl, 2 mM MgCl<sub>2</sub> and 1 mM DTT supplemented with DNase I (10  $\mu$ g/ml) (Roche) and cOmplete protease inhibitor cocktail tablets (Roche). Complex formation was initiated by adding synthesized Stachel peptide (100  $\mu$ M) and kept constant during the purification procedure. Incubation for 2 h at room temperature with the addition of apyrase (25 mU/ml) allowed complex formation on the cell membrane. Solubilization was done by supplementing 1% (w/v) LMNG (Anatrace) and 0.1% (w/v) CHS (Anatrace) and incubation for 2 h at 4°C. Insoluble material was removed by centrifugation at 136'000 rcf for 45 min. The clarified supernatant was incubated overnight with 2 ml anti-DYKDDDDK G1 Affinity Resin (Genscript) at 4°C by batch binding. The resin was collected in a gravity flow column and washed with 30 CV of wash buffer (20 mM HEPES pH 7.5, 100 mM NaCl, 2 mM MgCl<sub>2</sub>, 10% glycerol, 1mM DTT, 0.01% LMNG, 0.002% CHS and  $\pm$  0.002% glycol-diosgenin (GDN, Anatrace)). The protein complex was eluted using wash buffer supplemented with 0.2 mg/ml Flag peptide (Genscript or Peptide 2.0). The eluent was concentrated and injected into a Superdex200 increase 10/300 GL column (GE Healthcare) pre-equilibrated with 20 mM HEPES pH 7.5, 100 mM NaCl, 2mM MgCl<sub>2</sub>, 0.001% LMNG, 0.0002% CHS and  $\pm$  0.0002% GDN. The fractions containing receptor-G protein complex were combined and either



directly used at a concentration of 1 mg/ml for preparing cryo-EM grids or concentrated to 2.2 mg/ml using a 100 kDa cut-off Amicon Ultra Centrifugal Filter (Millipore) and then used to prepare cryo-EM grids.

### 3.25 Construct design Adgrd1-mini-G $\alpha$ s fusion protein

The cDNA for the Adgrd1-CTF (D1.12) was fused to mini-G $\alpha$ s 399 [120] at the C-terminus. Additionally, a HA signal sequence followed by a double FLAG tag and a TEV cleavage site was fused to the N-terminus of Adgrd1-CTF. The sequence was cloned into the pAC8REDNK vector for insect cell expression.

### 3.26 Purification Adgrd1-mini-G $\alpha$ s fusion protein

The cell pellet was resuspended in buffer A (20 mM HEPES pH 7.5, 100 mM NaCl, 2 mM MgCl<sub>2</sub> and 1 mM DTT supplemented with DNase I (10  $\mu$ g/ml) (Roche) and cOmplete protease inhibitor cocktail tablets (Roche)). Complex formation was initiated by adding synthesized Stachel peptide (100  $\mu$ M) which was kept constant during the purification procedure. Incubation for 2 h at room temperature with the addition of apyrase (25 mU/ml) allowed complex formation on the cell membrane. Solubilization was done by supplementing 1% (w/v) LMNG (Anatrace) and 0.1% (w/v) CHS (Anatrace) and the suspension was incubated for 2 h at 4°C. Insoluble material was removed by centrifugation at 136'000 rcf for 45 min at 4°C. The clarified supernatant was incubated overnight with 2 ml anti-DYKDDDDK G1 Affinity Resin (Genscript) at 4°C by batch binding. The resin was collected in a gravity flow column and washed with 30 CV of wash buffer (20 mM HEPES pH 7.5, 100 mM NaCl, 2 mM MgCl<sub>2</sub>, 10% glycerol, 1mM DTT, 0.01% LMNG, 0.002% CHS and 0.002% GDN (Anatrace)). Complex was eluted using wash buffer supplemented with 0.2 mg/ml Flag peptide (Genscript or Peptide 2.0). Eluent was concentrated using a 50 kDa cut-off Amicon Ultra Centrifugal Filter (Millipore) and injected into a Superdex 200 increase 10/300 GL column (GE Healthcare) pre-equilibrated with 20 mM HEPES pH 7.5, 100 mM NaCl, 2mM MgCl<sub>2</sub>, 1 mM DTT, 0.001% LMNG, 0.0002% CHS and 0.0002% GDN. The fractions containing the protein of interest were combined and concentrated using a 50 kDa cut-off Amicon Ultra Centrifugal Filter (Millipore). The protein was flash frozen in liquid nitrogen and stored at -80°C.

### 3.27 Construct design Adgrd1-NTF-CTF fusion protein

The cDNA encoding the human Adgrd1 (isoform 1, Uniprot ID: Q6QNK2-1) was used to generate the truncated Adgrd1<sup>GAIN-CTF</sup> receptor (residues 275-874). The constructs started with a HA signal sequence, followed by a quadruple FLAG epitope and a 3C protease cleavage site fused to the N-terminus. The C-terminus of Adgrd1<sup>GAIN-CTF</sup> was fused to a glycine-serine linker (3xGGGGGS) followed by a TEV protease cleavage site and mini-G $\alpha$ s 399 [120]. The sequence was cloned into the pAC8RED vector for insect cell expression. Point mutations were introduced to potentially form disulfide bonds between the GAIN and CTF domains. Locations of point mutations can be seen in the section “8.1 Tables” together with the full sequences of the proteins.

### 3.28 Expression and purification Adgrd1-NTF-CTF fusion protein

Expression tests were done in High Five and Sf9 insect cells in 96-DWP as described “3.10 Expression in insect cells”. After 48 h expression, 20  $\mu$ l of the cell suspension was mixed with 10  $\mu$ l of SDS buffer

and the samples were loaded to SDS-PAGE for subsequent western blot analysis. The rest of the cells were harvested by centrifugation and stored at  $-80^{\circ}\text{C}$ .

Large scale expression was done in High Five cells with a VOI of 3% as previously described in “3.10 Expression in insect cells”. Cells were harvested after 48h by centrifugation and stored at  $-80^{\circ}\text{C}$ .

The cell pellet was thawed and resuspended in buffer A (20 mM HEPES pH 7.5, 100 mM NaCl, 2 mM  $\text{MgCl}_2$ , 1 mM DTT) supplemented with DNase I (10  $\mu\text{g}/\text{ml}$ ) (Roche) and cComplete protease inhibitor cocktail tablets (Roche). The cells were mechanically disrupted with a Dounce homogenizer, followed by centrifugation at 185'000 rcf at  $4^{\circ}\text{C}$  for 45 min. Pelleted membranes were resuspended in buffer A and supplemented with 1 % LMNG (v:v) and 0.1 % CHS (v:v) for the solubilization step. After incubation for 2h at  $4^{\circ}\text{C}$  while stirring the suspension, insoluble material was removed by centrifugation at 200'000 rcf for 45 min at  $4^{\circ}\text{C}$ . The clarified supernatant was mixed with anti-DYKDDDDK G1 Affinity Resin (Genscript) and incubated overnight at  $4^{\circ}\text{C}$  on a roller mixer. The mix was transferred to an empty pre-washed gravity flow column to collect the resin. The resin was washed with 30 CV of buffer B (20 mM HEPES pH 7.5, 100 mM NaCl, 2 mM  $\text{MgCl}_2$ , 1 mM DTT, 0.01% LMNG, 0.002% CHS and 200  $\mu\text{M}$  Stachel peptide). Elution was done in three steps. In each step, two CV of buffer C (20 mM HEPES pH 7.5, 100 mM NaCl, 2 mM  $\text{MgCl}_2$ , 1 mM DTT, 0.01% LMNG, 0.002% CHS, 0.4 mg/ml FLAG peptide (Peptide 2.0)) were added to the column and the column was incubated for 45 min at  $4^{\circ}\text{C}$  on a roller mixer. The three elution fractions were combined and concentrated. The concentrated sample was injected into a Superdex200 increase 10/300 column pre-equilibrated with buffer D (20 mM HEPES pH 7.5, 100 mM NaCl, 2 mM  $\text{MgCl}_2$ , 1 mM DTT, 0.001% LMNG, 0.0002% CHS). Peak fractions were collected and a 200  $\mu\text{l}$  sample was taken, supplemented with 100 mM DTT, incubated for 1h at  $4^{\circ}\text{C}$  and again injected to the SEC column. The rest of the sample from the peak fractions was used for making negative stain grids.

### 3.29 Negative stain grid preparation and analysis

Grids were prepared from two samples with different protein concentrations. One sample had a concentration of 3.3 mg/ml and the other sample had a concentration of 0.0066 mg/ml. 3  $\mu\text{l}$  of sample was applied to a Quantifoil R1.2/1.3 Cu 200 mesh grid and blotted with a Whatman<sup>®</sup> filter paper (Sigma-Aldrich). Next, the grid was dipped in a water drop and again blotted. This was repeated four times to reduce the detergent concentration. Lastly, the grid was dipped into a drop of 1% uranyl acetate for 10 seconds before blotting and air drying the grid. Grids were stored at room temperature. Negative stain grids were analyzed with a JEOL JEM 2200FS microscope at room temperature and micrographs were taken manually with a K2 camera (Gatan) at various magnifications.

### 3.30 Cryo-EM grid preparation and data acquisition aGPCR-G protein complex

For sample A (see “5.1.1 aGPCR-G protein complex”): A volume of 3  $\mu\text{l}$  purified complex at a concentration of 1 mg/ml was applied to glow-discharged holey carbon grids (Quantifoil R1.2/1.3 Cu 200 mesh). The grids were blotted with a Vitrobot Mark IV (Thermo Fisher Scientific) and flash frozen in liquid ethane. The sample was blotted for 2s at  $4^{\circ}\text{C}$  and 100% humidity. Data was acquired on a 300kV Titan Krios (Thermo Fisher Scientific) at a nominal magnification of 165'000x, corresponding to a pixel size of 0.51  $\text{\AA}$ . A GIF Quantum LS energy filter (Gatan) was operated with an energy slit width of 20 eV. Movies were recorded with a K3 direct electron detector (Gatan) with defocus values from -0.8  $\mu\text{m}$  to -2.0  $\mu\text{m}$ . A total number of 40 frames per micrograph were recorded with an accumulated dose of 55  $\text{e}^-/\text{\AA}^2$ . EPU (FEI, Thermo Fisher Scientific) was used for automated data acquisition.

For sample B (see “5.1.1 AGPCR-G protein complex”): A volume of 3  $\mu\text{l}$  purified complex at a concentration of 2.2 mg/ml was applied to glow-discharged holey carbon grids (Quantifoil R1.2/1.3 Cu 200 mesh). The grids were blotted with a Vitrobot Mark IV (Thermo Fisher Scientific) and flash frozen in liquid ethane. The sample was blotted for 3s at 4°C and 100% humidity. Data was acquired on a 300kV Titan Krios (Thermo Fisher Scientific) at a nominal magnification of 165'000x, corresponding to a pixel size of 0.51 Å. A GIF Quantum LS energy filter (Gatan) was operated with an energy slit width of 20 eV. Movies were recorded with a K3 direct electron detector (Gatan) with defocus values from -1.2  $\mu\text{m}$  to -3.0  $\mu\text{m}$ . A total number of 40 frames per micrograph were recorded with an accumulated dose of 80 e<sup>-</sup>/ Å<sup>2</sup>. EPU (FEI, Thermo Fisher Scientific) was used for automated data acquisition.

### 3.31 Cryo-EM data processing aGPCR-G protein complex

For sample A (see “5.1.1 AGPCR-G protein complex”): Cryo-EM data was processed using cryoSPARC [125]. Raw movies were subjected to beam-induced motion correction using cryoSPARC's patch motion correction (multi). For each micrograph, the contrast transfer function (CTF) parameters were determined by patch CTF estimation (multi). Initial particles were picked on a subset of all micrographs using Topaz [126] with the general model ResNet8 and the particles were subjected to reference-free 2D classification. 2D classes showing the Adgrd1-mini-Gs complex with the corresponding particles were selected and Topaz was used to train a model for automated particle picking. Automated picking on the full dataset yielded 915'922 particles. Particles were again subjected to reference-free 2D classification. A total of 95'413 particles grouped into 2D averages showing the Adgrd1-Gs heterotrimer complex. Initial models were generated with cryoSPARC.

The sample B (see “5.1.1 AGPCR-G protein complex”): The cryo-EM data was processed with the same pipeline as described for sample A. Automated picking yielded 1'941'387 particles and a total of 33'278 particles grouped into 2D class averages showing the complex.

### 3.32 Retinal analogue synthesis

All-trans Retinal 6.11 was provided by Mordechai Sheves (Weizmann Institute of Science) and was synthesized as described previously [127].

### 3.33 JSR1 Expression and Purification

Wild-type jumping spider rhodopsin isoform-1 (JSR1) (Uniprot ID: B1B1U5) from *Hasarius adansoni* tagged with a C-terminal 1D4 epitope [96], was recombinantly expressed in HEK293 GnTi<sup>-</sup> cells as described previously [100]. Cells were harvested by centrifugation at 500 rcf and the pellets were stored at -80°C. The cells were later thawed in lysis buffer (50 mM HEPES pH 6.5, 150 mM NaCl, 3 mM MgCl<sub>2</sub>) supplemented with cOmplete protease inhibitor cocktail tablets (Roche) and mechanically disrupted in a Dounce homogenizer. All following steps were performed under dim red light conditions. Retinal was added to the lysate (all-trans retinal 6.11 (see section above)) to a final concentration of 50  $\mu\text{M}$  (or 25  $\mu\text{M}$  and 12.5  $\mu\text{M}$ ) and the suspension was incubated at 4°C overnight. DDM powder (Solgrade; Anatrace) was then added to a final concentration of 1% (w/v) to solubilize the receptor and insoluble material was removed after a 2h incubation at 4°C by centrifugation at 100'000 rcf for one hour. The supernatant was incubated with Sepharose resin (GE Healthcare Life Science) coupled to anti-1D4 antibody (Cell Essentials) at 4°C overnight. The resin was then loaded into a glass Econo-Column (Bio-Rad Laboratories), and washed with 30 CV of wash buffer (50 mM HEPES pH 6.5, 150 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.01% DDM (w/v)) before the addition of 1.5 CV of elution

buffer (50 mM HEPES pH 6.5, 150 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.01% DDM (w/v), 800 μM 1D4 peptide (Peptide 2.0)) to the resin. The slurry was incubated at 4°C overnight and the protein was eluted from the column the following day. Residual protein was eluted from the resin after the addition of three more CV of wash buffer. All eluted fractions were pooled and concentrated using a 50 kDa cut-off Amicon Ultra Centrifugal Filter (Millipore). The protein was either directly used for subsequent experiments or flash frozen in liquid nitrogen and stored at -80°C.

## 3.34 G protein heterotrimer expression in insect cells

### 3.34.1 hGq, hG11, JSGiq and JSGq expression

The design of the expression construct is described in “3.14 G protein construct design for insect cell expression”. For the G<sub>aiq</sub> chimera, mutations were introduced to the human G<sub>ai</sub> to match the sequence of the jumping spider visual G<sub>αq1</sub> sequence (A31R; D193S; L194I; [resi 337-354=DAVTDVVIKNNLKDGLF] G<sub>ai1</sub> to [resi 337-354=CAVKDTILQNNLKECNLV]) (**Figure 40**).

The G protein heterotrimer was expressed in High Five insect cells (Invitrogen). Cells were grown to a density of 4.0 x 10<sup>6</sup> cells/ml in SF900II SFM medium using a 5L Erlenmeyer flasks (Corning). Prior to expression, the medium was exchanged, and cells were resuspended to 4 x 10<sup>6</sup> cells/ml in fresh SF900II SFM medium. Then baculovirus was added to infect the cells and cells were cultured at 27°C and 120 rpm and harvested 48 hours post-infection by centrifugation at 800 rcf for 15 min. Cell pellets were flash frozen in liquid nitrogen and stored at -80°C.

## 3.35 hGq, hG11, JSGiq and JSGq heterotrimer purification

The frozen cell pellet was thawed and resuspended in a buffer containing 20 mM HEPES pH 7.5, 100 mM NaCl, 3 mM MgCl<sub>2</sub>, 100 μM EDTA, 5 mM β-ME and 20 μM GDP supplemented with DNase I (10 μg/ml) (Roche) and cOmplete protease inhibitor cocktail tablets (Roche). Cells were mechanically disrupted using a Microfluidizer (Microfluidics, IDEX) and subsequently centrifuged at 185'000 rcf for 45 min. The pellet was resuspended in 20 mM HEPES, 100 mM NaCl, 20 mM Imidazole, 3 mM MgCl<sub>2</sub>, 5 mM β-ME and 20 μM GDP supplemented with cOmplete protease inhibitor cocktail tablets and solubilized by adding 1% Sodium cholate (Sigma-Aldrich) and 0.05% (w/v) DDM (Anatrace). After incubation at 4°C for 1 h, insoluble material was removed by centrifugation at 185'000 rcf for 45 min. The clarified supernatant was incubated with 5 ml of TALON resin (Takara Bio) for 90 min at 4°C. The resin was collected in a pre-washed gravity flow column and the detergent was gradually exchanged to 0.1 % DDM. The resin was washed with 10 CV of IMAC buffer A (20 mM HEPES pH 7.5, 100 mM NaCl, 20 mM Imidazole, 1 mM MgCl<sub>2</sub>, 5 mM β-ME, 20 μM GDP and 0.01 % DDM) and 10 CV of IMAC buffer B (IMAC buffer A with 300 mM NaCl). The sample was eluted with 6 CV IMAC buffer A supplemented with 300 mM Imidazole. His-tagged HRV 3C protease was added (1:50 w/w; HRV 3C protease:Giq heterotrimer) to the eluent and the eluent was dialyzed against dialysis buffer (20 mM HEPES pH 7.5, 100 mM NaCl, 1 mM MgCl<sub>2</sub>, 100 μM TCEP, 20 μM GDP and 0.05 % DDM) overnight at 4°C. Dialysate was collected and supplemented with Imidazole to a final concentration of 20 mM before incubation with 5 ml TALON resin for 1 h at 4°C. The resin was transferred to a gravity flow column and the flow through was collected. The resin was washed with one CV of dialysis buffer supplemented with 20 mM Imidazole. The flow through and wash were combined and injected into a HiTrap Q FF 1 ml column (Cytiva). The column was washed with 15 CV of Q buffer A (20 mM HEPES pH 7.5, 100 mM NaCl, 1 mM MgCl<sub>2</sub>, 100 μM TCEP, 20 μM GDP and 0.05 % DDM) and the sample was eluted with a linear gradient over 30 CV with Q buffer B (Q buffer A with 1000 mM NaCl). The relevant

fractions were pooled and concentrated using a 50 kDa cut-off Amicon Ultra Centrifugal Filter (Millipore). The concentrated sample was injected into a SRT-C 300 SEC column (Sepax Technologies) pre-equilibrated with SEC buffer (20 mM HEPES pH 7.5, 100 mM NaCl, 1 mM MgCl<sub>2</sub>, 100 μM TCEP, 20 μM GDP and 0.02% DDM). Fractions containing the G heterotrimer were pooled and concentrated using a 50 kDa cut-off Amicon Ultra Centrifugal Filter (Millipore) before adding 10% glycerol (v/v). The sample was flash frozen in liquid nitrogen and stored at -80°C.

### 3.36 Analytical SEC experiments

A Zenix-C SEC-300 column was equilibrated with SEC buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 1 mM MgCl<sub>2</sub> and 0.01% DDM). Absorbance at 280 nm (protein) and 505 nm (retinal) was monitored. Each sample vial contained 60 μl protein solution and 50 μl were injected into the column. JSR and G protein was added to a final concentration of 8 μM, supplemented with 25 U/ml apyrase, and the samples were incubated for at least 1h at 4°C. For samples containing an antibody, the respective antibody was added to reach 8 μM final concentration. For dissociation experiments, GDP or GTPγS was added to the solution to reach a final concentration of 100 μM. Reference samples contained only one component at a concentration of 8 μM.

### 3.37 Complex formation and SEC for the JSR1-bGt complex

The 1D4 purified receptor was mixed with purified bovine Gt (bGt) heterotrimer (Purified by Dr. Ching-Ju Tsai) at a molar ratio of 1:1.25 and supplemented with apyrase (25 mU/mL) to degrade any GTP and GDP present in the solution. After at least one hour incubation at 4°C, the sample was concentrated using a 100 kDa cut-off Amicon Ultra Centrifugal Filter (Millipore) and injected into a SRT-C 300 SEC column (Sepax Technologies) pre-equilibrated with SEC buffer (20 mM HEPES pH 6.5, 100 mM NaCl and 1 mM MgCl<sub>2</sub>). Relevant fractions were pooled and the sample was concentrated to 6.4 mg/ml for preparing cryo-EM grids.

### 3.38 Cryo-EM grid preparation and data acquisition JSR1-bGt complex

A volume of 3 μl purified complex at a concentration of 6.4 mg/ml was applied to glow-discharged holey carbon grids (Quantifoil R1.2/1.3 Cu 200 mesh). The grids were blotted with a Vitrobot Mark IV (Thermo Fisher Scientific) and flash frozen in liquid ethane. The sample was blotted for 3 s at 4°C and 100 % humidity. Data was acquired on a 300kV Titan Krios (Thermo Fisher Scientific) at a nominal magnification of 165'000x, corresponding to a pixel size of 0.51 Å. A GIF Quantum LS energy filter (Gatan) was operated with an energy slit width of 20 eV. Movies were recorded with a K3 direct electron detector (Gatan) with a dose rate of 15 e<sup>-</sup>/px/s and defocus values from -1.0 to -2.2 μM in 0.2 μM intervals. A total number of 40 frames per micrograph were recorded with an accumulated dose of 50 e<sup>-</sup>/ Å<sup>2</sup>. EPU (FEI, Thermo Fisher Scientific) was used for automated data acquisition.

### 3.39 Complex formation and SEC for the JSR1-hGi (-scFv16) complex

The 1D4 purified receptor was mixed with purified hGi heterotrimer (hGαi and Gβγ subunits were mixed at an equimolar ration and incubated for at least 30 min) at a molar ratio of 1:1.25 (scFv16 was added at an equimolar ratio to hGi heterotrimer) and supplemented with apyrase (25 mU/mL) to degrade any GTP and GDP present in the solution. After one hour incubation at 4°C, the sample was concentrated using a 100 kDa cut-off Amicon Ultra Centrifugal Filter (Millipore) and injected into a SRT-



C 300 SEC column (Sepax Technologies) pre-equilibrated with SEC buffer (20 mM HEPES pH 6.5, 100 mM NaCl and 1 mM MgCl<sub>2</sub>). Relevant fractions were pooled and the samples for the JSR1-hGi and JSR1-hGi-scFv16 complexes were concentrated to ~4 mg/ml and ~3.5 mg/ml respectively and used to prepare grids for cryo-EM analysis.

### 3.40 Cryo-EM grid preparation and data acquisition JSR1-hGi complex

A volume of 3  $\mu$ l purified complex at a concentration of 4 mg/ml was applied to glow-discharged holey carbon grids (Quantifoil R1.2/1.3 Cu 200 mesh). The grids were blotted with a Vitrobot Mark IV (Thermo Fisher Scientific) and flash frozen in liquid ethane. The sample was blotted for 4 s at 4°C and 100 % humidity. Data was acquired on a 300kV Titan Krios (Thermo Fisher Scientific) at a nominal magnification of 165'000x, corresponding to a pixel size of 0.51 Å. A GIF Quantum LS energy filter (Gatan) was operated with an energy slit width of 20 eV. Movies were recorded with a K3 direct electron detector (Gatan) with a dose rate of 19 e<sup>-</sup>/px/s and defocus values from -1.0 to -2.4  $\mu$ M in 0.2  $\mu$ M intervals. A total number of 40 frames per micrograph were recorded with an accumulated dose of 70 e<sup>-</sup>/Å<sup>2</sup>. EPU (FEI, Thermo Fisher Scientific) was used for automated data acquisition.

### 3.41 Cryo-EM grid preparation and data acquisition JSR1-hGi-scFv16 complex

A volume of 3  $\mu$ l purified complex at a concentration of 3.5 mg/ml was applied to glow-discharged holey carbon grids (Quantifoil R1.2/1.3 Cu 200 mesh). The grids were blotted with a Vitrobot Mark IV (Thermo Fisher Scientific) and flash frozen in liquid ethane. The sample was blotted for 6 s at 4°C and 100 % humidity. Data was acquired on a 300kV Titan Krios (Thermo Fisher Scientific) at a magnification of 165'000x, corresponding to a pixel size of 0.51 Å. A GIF Quantum LS energy filter (Gatan) was operated with an energy slit width of 20 eV. Movies were recorded with a K3 direct electron detector (Gatan) with a dose rate of 15 e<sup>-</sup>/px/s and defocus values from -1.0 to -2.4  $\mu$ M in 0.2  $\mu$ M intervals. A total number of 40 frames per micrograph were recorded with an accumulated dose of 54 e<sup>-</sup>/Å<sup>2</sup>. EPU (FEI, Thermo Fisher Scientific) was used for automated data acquisition.

### 3.42 Complex formation and SEC for the JSR1-JSGiq complex

The 1D4 purified receptor was mixed with purified JSGiq heterotrimer at a molar ratio of 1:1.25 and supplemented with apyrase (25 mU/mL) to degrade any GTP and GDP present in the solution. After two hours incubation at 4°C, the sample was concentrated using a 100 kDa cut-off Amicon Ultra Centrifugal Filter (Milipore) and injected into a SRT-C 300 SEC column (Sepax Technologies) pre-equilibrated with SEC buffer (20 mM HEPES pH 6.5, 100 mM NaCl and 1 mM MgCl<sub>2</sub>). Relevant fractions were pooled and concentrated to 8 mg/ml for preparing cryo-EM grids.

### 3.43 Cryo-EM grid preparation and data acquisition JSR1-JSGiq complex

A volume of 3  $\mu$ l purified complex at a concentration of 8 mg/ml was applied to glow-discharged holey carbon grids (Quantifoil R1.2/1.3 Cu 200 mesh). The grids were blotted with a Vitrobot Mark IV (Thermo Fisher Scientific) and flash frozen in liquid ethane. The sample was blotted for 6 s at 4°C and 100 % humidity. Data was acquired on a 300kV Titan Krios (Thermo Fisher Scientific) at a nominal magnification of 165'000x, corresponding to a pixel size of 0.51 Å. A GIF Quantum LS energy filter (Gatan) was operated with an energy slit width of 20 eV. Movies were recorded with a K3 direct electron detector (Gatan) with a dose rate of 19 e<sup>-</sup>/px/s and defocus values from -1.0 to -2.4  $\mu$ M in 0.2

$\mu\text{M}$  intervals. A total number of 40 frames per micrograph were recorded with an accumulated dose of  $70 \text{ e}^- / \text{\AA}^2$ . EPU (FEI, Thermo Fisher Scientific) was used for automated data acquisition.

### 3.44 Cryo-EM data processing JSR1-JSGiq complex

The cryo-EM data processing for the JSR1-Gt and JSR1-hGi (-scFv16) complexes followed the same processing pipeline as described here for the JSR1-JSGiq complex. Cryo-EM data was processed using cryoSPARC (Structura Biotechnology Inc, [125]) and Relion [128]. Pre-processing was done on-the-fly using MotionCor2 [129] with dose-weighting and micrographs were binned by two, resulting in a pixel size of  $1.02 \text{ \AA}$ . Motion corrected micrographs were then imported to cryoSPARC and CTF parameters were determined by patch CTF estimation (multi). Micrographs with a CTF fit resolution  $>8 \text{ \AA}$  were discarded. Initial particles were picked on a subset of all micrographs using Topaz [126] with the general model ResNet8 and the particles were subjected to reference-free 2D classification. 2D classes showing the JSR-JSGiq complex with the corresponding particles were selected and Topaz was used to train a model for automated particle picking. Automated picking on the full dataset yielded 2'964'739 particles. Particles were again subjected to reference-free 2D classification. A total of 768'100 particles grouped into 2D averages showing the JSR-JSGiq heterotrimer complex. Initial models were made with cryoSPARC and 3D classification was done to separate different conformations of the complex. Particles from the two best classes were used for non-uniform refinement and yielded two maps with  $4.06 \text{ \AA}$  (159'665 particles) and  $4.17 \text{ \AA}$  (134'167 particles) resolution. Local CTF refinement and subsequent non-uniform refinement was done and yielded maps with the same resolution according to the Fourier Shell Correlation (FSC) at  $\text{FSC} = 0.143$  but visually the maps looked better than before. CryoSPARC's 3D variability analysis [130] was used to visualize motion in the particles set.

### 3.45 Model building and structure refinement

The initial mode of JSR was obtained from the crystal structure of the inactive state (6I9K) and the initial model for the JSGiq heterotrimer was obtained from the crystal structure of the human Gi heterotrimer (6CRK) where residues in the G $\alpha$ i subunit were mutated to match our Giq sequence. The models were docked into the electron density maps as rigid bodies using Chimera [131]. Initial refinement was done with phenix.real\_space\_refine in the Phenix suite [132] and models were manually manipulated in Coot [133] before running phenix.real\_space\_refine again. The H8 was manually built in coot because it was not modeled in the crystal structure (6I9K).

### 3.46 Structure comparison with JSR1-G protein complexes

Structure alignments and comparisons were done in PyMOL [134]. Structures were aligned with conserved residues inside the transmembrane region of the receptors. The following residues were used for alignment (Ballesteros-Weinstein numbering [25]): 1.45-1.54, 2.46-2.56, 3.34-3.44, 4.48-4.56, 7.38-7.46.

Accession codes for the structures used for comparison are listed below:

Inactive GPCRs: JSR1 (6I9K), Bovine Rhodopsin (1GZM), Squid rhodopsin (2Z73)

Active GPCR-G protein complexes: Bovine rhodopsin-G $\alpha$ t peptide (4A4M), Bovine rhodopsin-G $\alpha$ t peptide (3PQR), Bovine rhodopsin-Gi (6CMO), Bovine rhodopsin-mini-Go (6FUF), Bovine rhodopsin-G $\alpha$ t peptide (5EN0),  $\beta$ 2-adrenergic receptor-Gs (3SN6), M1-receptor-G11 (6OIJ), NTSR1-Gi NC (6OSA)

For the AHD comparison, the Ras domain of the G $\alpha$  subunits (Residues 31-58+181-354) was aligned.

Accession codes for the structures used for comparison are listed below:

GABA (B) receptor-human Gi complex (7EB2), NTSR1-human Gi complex (7LOS), cannabinoid receptor 2-human Gi complex (6PT0)

### 3.47 GTPase Glo assay® (Promega, Madison, USA)

Reactions were set up in 384-well, low-volume, PS, white, solid-bottom plates (Greiner Bio-One, Kremsmünster) and carried out at 20°C under dim-red light conditions. Each condition was prepared in octuplicates. 2.5 µl of protein solution were mixed with 2.5 µl of 2 µM GTP solution. Reference samples contained either 2 µM G protein or 0.2 µM JSR1 with either 9-cis retinal, 9-cis retinal 6.11 or all-trans retinal 6.11 (dark and light activated). Test samples contained 2 µM G protein and 0.2 µM JSR with the respective ligand (dark and light activated). For assays done with the JSGiQ, three different G protein concentrations were used (0.5 µM, 0.1 µM and 0.05 µM G protein). A control was done with no proteins. The plate was sealed and incubated for 120 minutes while shaking at 500 rpm. Once the assay with the JSGiQ chimera was incubated for 60 min or 30 min. Subsequently, 5 µl of GTP-Glo reagent was added to each well and the reaction was carried out for 30 minutes while shaking at 500 rpm. Finally, 10 µl of Detection Reagent was added to each well. Luminescence was measured after 10 minutes in a PheraStar FSX (BMG labtech) plate reader (Optical module: LUM plus, Gain: 3600, focal height: 14.5, 1 second/well). Outliers were identified using the Grubbs test [135]. GTP hydrolysis was calculated in relation to the control sample:

$$GTP\ hydrolysis\ [\%] = \left( \sum_1^n 100 - \frac{Lum(x)_n}{Mean\ Lum_0} \times 100 \right) \div n$$

where  $Lum_0$  is the luminescence signal of the control sample and  $Lum(x)$  the luminescence of each condition ( $x$ ) and  $n$  the number of replicates.



## 4 Results

### 4.1 Adhesion GPCR project

#### 4.1.1 AGPCR-G protein complex

In the last 10 years aGPCR research has gained a lot of attention because several studies showed that these receptors possess unique properties (see Introduction “1.2 Adhesion GPCRs”). Structural information remains scarce compared to other GPCR families. At the start of this project, structure determination was limited to extracellular domains (ECD) of aGPCRs. In this project we aimed to determine the structure of the transmembrane (TM) domain of an aGPCR. This work focused on Adrg6, also called Gpr126, and Adrd1, also called Gpr133. While our collaborators, Prof. Adriano Aguzzi and his group, were focused on the physiological roles of these receptors in context of the prion disease, our work focused on biochemical and structural characterization of Adrd1 and Adrg6. This project started with construct design for recombinant expression of these receptors, followed by biochemical characterization, complex formation of the active receptor with a G protein and structural characterization of an aGPCR-G protein complex. Construct design and experimental designs were created together with Dr. Ching-Ju Tsai and Prof. Gebhard Schertler (Paul Scherrer Institut). Experiments were performed by the author unless stated otherwise.

##### *4.1.1.1 Construct design and expression screening*

For initial expression screening experiments, we performed two rounds of construct design and initial test expressions. One round was aimed to express the receptors in human HEK293F cells and the other was aimed to use the insect cell expression system.

The construct design for HEK cell expression is detailed in **Table 1** and **Figure 11A**. For both Adrg6/Adrd1 several truncated/mutated receptor constructs were designed that are inactive in theory and can be activated by adding chemically synthesized Stachel peptide. From here constructs for Adrg6 will be abbreviated with G6, followed by the number of the construct. Likewise, constructs for Adrd1 contain the D1 prefix followed by the number of the construct. Constructs that contain the Stachel sequence had mutations converting the Stachel sequence to an antagonist (G6.03; D1.03/.04) or had a truncated Stachel sequence that does not act as an agonist anymore (G6.02/.04; D1.01/.02). In addition to these constructs, one Adrg6 construct also included the GAIN domain (G6.01) and two Adrd1 constructs included the full-length isoform 2 or isoform 3 (D1.05/.06). These sequences were fused to a C-terminal HRV 3C protease cleavage site followed by a YFP, a 10xHis tag, and a TwinStrep tag and were cloned into the pcDNA4/TetO vector and protein expression was tested in HEK293F cells.

	Signal peptide		Ecto domain	Stachel sequence	C-terminus	Cleavage site/ Purification tags	Activity
G6.01	Its own		GAIN domain	THFGVLMDLPR	-	HRV3C protease - YFP - HIS - TwinStrep	Activatable (?)
G6.02			-	GVLMDLPR	-		Inactive
G6.03			-	THAGVLMDLPR	truncated		Inactive
G6.04			-	GVLMDLPR	truncated		Inactive
D1.01	Its own		-	AILMQV	-		Inactive
D1.02			-	AILMQV	truncated		Inactive
D1.03			-	TNFAIAMQV	truncated		Inactive
D1.04			-	TNFAIAMQV	-		Inactive
D1.05			Full-length isoform 2	TNFAILMQV	-		Activatable
D1.06			Full-length isoform 3	TNFAILMQV	-		Activatable

Table 1: Construct design for ADGRG6 (G6.01-.04) and ADGRD1 (D1.01-.06).

Expression was done by transient transfection of HEK293F cells. Cells were harvested 72h after transfection and expression was evaluated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) and analytical fluorescent size-exclusion chromatography (SEC), monitoring the fluorescent signal from the fused YFP. Samples for SDS-PAGE were taken either directly from the cell suspension or cells were lysed in buffer containing dodecyl maltoside (DDM). Insoluble material was removed by centrifugation and supernatant was loaded to SDS-PAGE and analytical SEC column. Protein bands in SDS-PAGE were detected using fluorescence imaging (**Figure 10**).

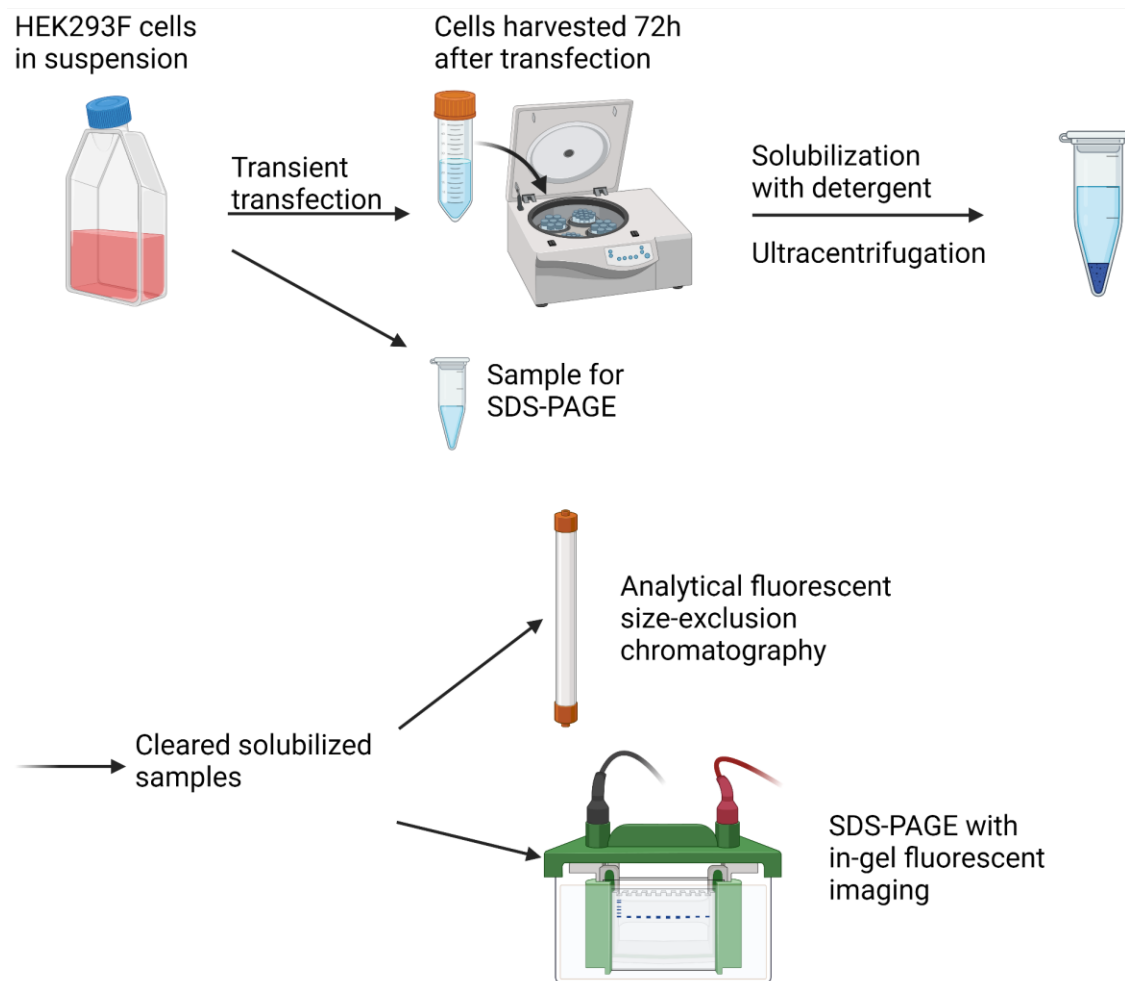
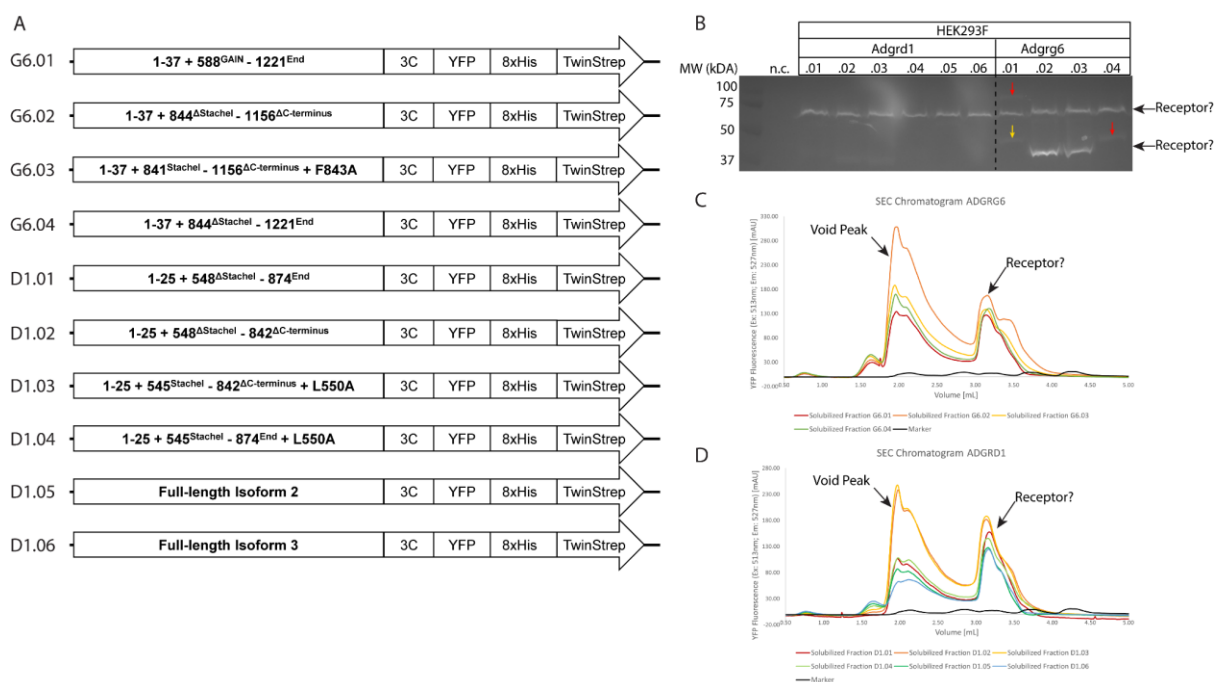


Figure 10: HEK293F cell protein expression strategy and protein expression analysis. Figure created with Biorender.

Unfortunately, all the constructs from both the cell suspension samples and DDM-solubilized samples showed bands at the same molecular weight (MW) regardless of the length of the protein constructs designed (**Figure 11B**). There were faint bands that correspond to the theoretical MW of constructs G6.01 and G6.04 (red arrows in **Figure 11B**) and lower MW bands for G6.02 and G6.03 that could be degradation products. Further, fluorescence signal was detected in the stacking gel, indicating strong aggregation in presence of SDS buffer. The analytical SEC experiments showed similar results. In all constructs there was a peak in the void volume, indicating aggregated proteins and a second peak at a retention volume that is slightly higher than the marker peak for the MW of 44kDa (**Figure 11C/D**). For all constructs the potential receptor peak was at the same position regardless of the MW of the construct, indicating that the expression did not work correctly (**Figure 11C/D**).



**Figure 11: Expression screening Adgrd1 and Adgrg6 in HEK293F cells. A:** Construct design for Adgrg6 (G6.01-.04) and Adgrd1 (D1.01-.06). **B:** In gel fluorescence image of SDS-PAGE loaded with solubilized samples from all constructs. White arrows indicate potential full-length proteins and yellow arrow indicates a potentially truncated protein. N.c. is an abbreviation of negative control, meaning uninfected HEK293F cells. **C:** Fluorescence detection SEC for all Adgrg6 constructs from solubilized samples. Void peak and potential receptor peaks are labeled. **D:** Fluorescence detection SEC for all Adgrd1 constructs from solubilized samples. Void peak and potential receptor peaks are labelled.

To further test expression, we screened different detergents for solubilization in case the receptor was not solubilized in DDM. Decyl maltoside (DM) and a lauryl maltose neopentyl glycol (LMNG)/cholesteryl hemisuccinate (CHS) mixture (10:1) were tested. However, the results showed no difference to the DDM samples. Additionally, small scale affinity purification trials have been performed, but these experiments showed negative results. Although both SDS-PAGE and analytical SEC showed fluorescence signals, we could not confirm expression of any of those constructs. Therefore, we switched our strategy and designed constructs for expression in insect cells.

Because HEK cell expression was not successful, we investigated published work of class B1 GPCR complexes studied by cryo-EM [136, 137]. Unlike class A GPCR complexes, for which individual protein components were purified and then mixed in a test tube for complex formation, class B1 GPCR complexes were formed on the cell membrane by expressing the receptor and G proteins in the same insect cell, followed by activation of the receptor to form the GPCR-G protein complex, detergent solubilization and purification of the complex. The main difference is if the complex is formed in the detergent micelle or the lipid bilayer and this is primarily due to the stability of the empty or agonist bound receptor in the detergent micelle.

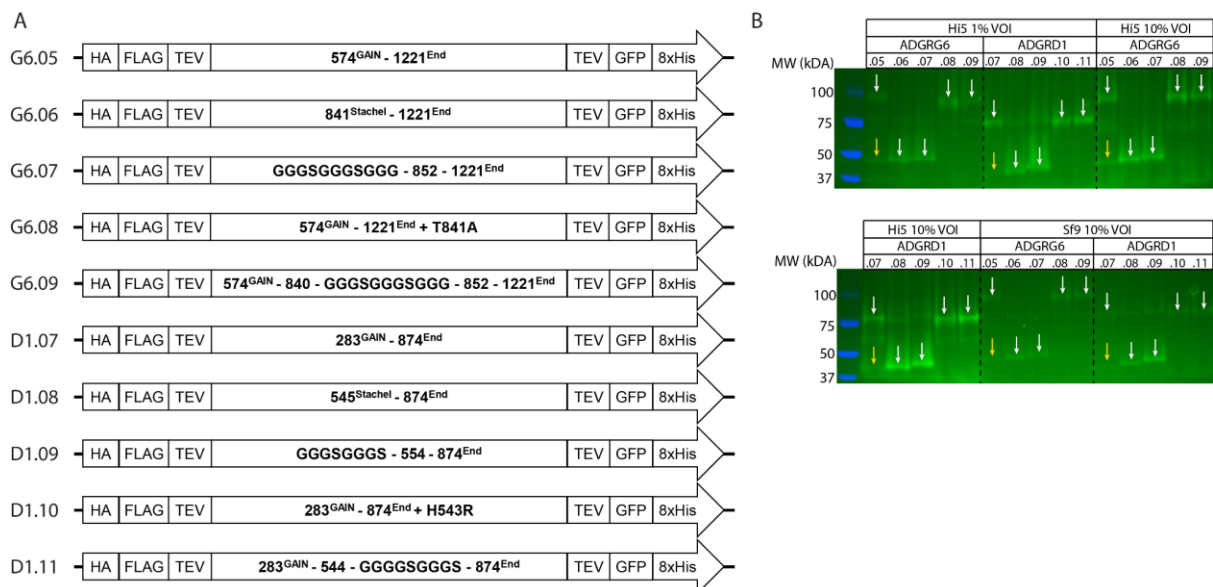
The constructs designed for insect cell expression contained a hemagglutinin signal peptide, followed by a FLAG peptide tag and a TEV cleavage site at the N-terminus. The C-terminus of the receptors was fused to a TEV cleavage site, followed by a GFP and a histidine tag. Adgrg6 and Adgrd1 receptor constructs were designed with variations of containing the GAIN domain and the Stachel sequence, covering a range of possibilities to activate the receptors in different manners (**Table 2, Figure 12A**). These constructs were cloned into the pAC8RED vector [124], where a DsRed marker was used to indicate baculovirus infection in insect cells using the flashBAC™ system.

	Signal peptide	N-term. tag and cleavage site	Ecto domain	Stachel sequence	C-term. cleavage site and tag	Activity
<b>G6.05</b>	Hemagglutinin tag (HA)	FLAG-TEV	GAIN domain	THFGVLMDLPR	TEV-GFP-His	Activatable
<b>G6.06</b>			-	THFGVLMDLPR		Active
<b>G6.07</b>			-	GGGSGGGSGGG		Activatable
<b>G6.08</b>			GAIN domain	AHFGVLMDLPR		Uncleavable Activatable
<b>G6.09</b>			GAIN domain	GGGSGGGSGGG		Activatable
<b>D1.07</b>	Hemagglutinin tag (HA)	FLAG-TEV	GAIN domain	<i>HL</i> - TNFAILMQV	TEV-GFP-His	Activatable
<b>D1.08</b>			-	TNFAILMQV		Active
<b>D1.09</b>			-	GGGSGGGSG		Activatable
<b>D1.10</b>			GAIN domain	<i>RL</i> - TNFAILMQV		Uncleavable Activatable
<b>D1.11</b>			GAIN domain	GGGSGGGSG		Activatable

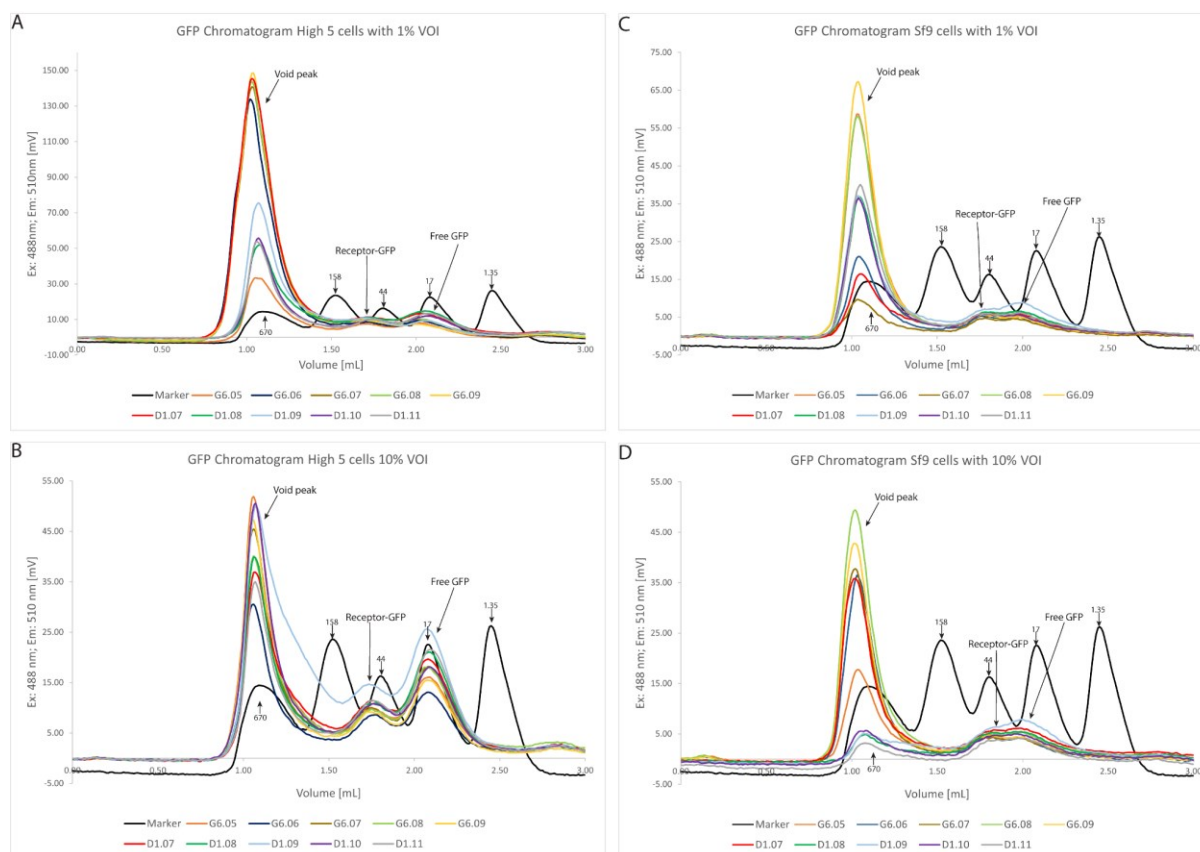
**Table 2:** Construct design for Adgrg6 (G6.05-.09) and Adgrd1 (D1.07-.11) for insect cell expression.

Small scale expression tests were set up in a 96-well format to test many conditions in parallel. Two insect cell lines, Sf9 and High Five cells, were tested with two different virus concentrations, 1% and 10% virality of infection (VOI; v/v) and two harvesting time points, 48 and 72 hours. Harvested cells were lysed and loaded onto SDS-PAGE to monitor expression using in-gel fluorescence imaging. The in-gel fluorescence data showed clear expression for both receptors in both cell lines, both VOIs and both time points. On the contrary to the first round of expression screening in HEK293F cells, bands corresponding to the MW of the constructs were detected (**Figure 12B**). For Adgrg6, intense fluorescence signal was detected in the stacking gel of the SDS-PAGE, indicating protein aggregation. For constructs G6.05 and D1.07 both the GAIN-CTF and the CTF were detected, indicating that autocleavage happens and both species can be detected. In contrast, for both theoretically

uncleavable GAIN domains (G6.08/.09; D1.10/.11) no CTF was detected, meaning the changes in the cleavage site resulted in an uncleavable GAIN domain (**Figure 12B**). For *Adgrd1* constructs that include the GAIN domain, bands with larger MW were detected that might correspond to oligomers. Generally, High Five cells showed higher expression levels than Sf9 cells and will therefore be used for expression in all future experiments. For analytical SEC experiments cells were lysed by adding 1% DDM and insoluble material was removed by centrifugation. When analyzing the solubilized samples by fluorescence SEC, the chromatograms did not show clear peaks corresponding to the MW of the constructs (**Figure 13**). All samples showed a clear peak in the void volume, indicating protein aggregation (**Figure 13**). Because the SEC experiments did not look very promising, but expression could be detected in the cells by SDS-PAGE, the next step was to do a detergent screening to find a suitable detergent for solubilization of the receptors.



**Figure 12: Expression screening for *Adgrg6* and *Adgrd1* in insect cells. A: Construct design for *Adgrg6* (G6.05-.09) and *Adgrd1* (D1.07-.11). B: In gel fluorescence images of SDS-PAGE. White arrows show full length proteins and yellow arrows show cleavage products of the corresponding receptors.**



**Figure 13: Fluorescence SEC experiments for insect cell expression of *Adrg6* and *Adrd1*.** Marker peaks are labeled with a number indicating the MW in kDa. **A:** Fluorescence SEC chromatogram for all constructs expressed in Hi5 cells with a VOI of 1%. **B:** Fluorescence SEC chromatogram for all constructs expressed in Hi5 cells with a VOI of 10%. **C:** Fluorescence SEC chromatogram for all constructs expressed in Sf9 cells with a VOI of 1%. **D:** Fluorescence SEC chromatogram for all constructs expressed in Sf9 cell with a VOI of 10%

#### 4.1.1.2 Detergent screening

For the detergent screening we decided to focus on G6.05/.06 and D1.07/.08/.10 because they contain less modifications than the other constructs. For each of these constructs a 500 mL expression was done and for each construct 1.5 grams of cell pellet was resuspended and split equally in different Eppendorf tubes containing different buffer conditions. In total six different detergents (DDM, DDM/CHS (10:1), DM, CHAPS, LMNG, LMNG/CHS (10:1)) were tested and three different additives (high NaCl concentration (300 mM), glycerol (10%; v:v), DTT (2 mM)) (**Table 3**). After solubilization, insoluble material was removed by centrifugation and the supernatants were injected into a SEC column for analysis. All samples showed a peak in the void volume, indicating that the receptors in all conditions tend to aggregate upon extraction from the lipid bilayer with detergents (**Figure 14/15**). Generally, *Adrg6* constructs showed higher signals for the void peaks, indicating less stability of *Adrg6* in detergent micelles. For most constructs peaks between the 158kDa and 44kDa marker peaks were detected that correspond to the receptor in detergent micelles (**Figure 14/15**). *Adrd1* constructs showed higher signals for the receptor peaks and better void peak signal to receptor peak signal ratios, indicating that the stability of *Adrd1* is higher in detergent micelles than *Adrg6* (**Table**

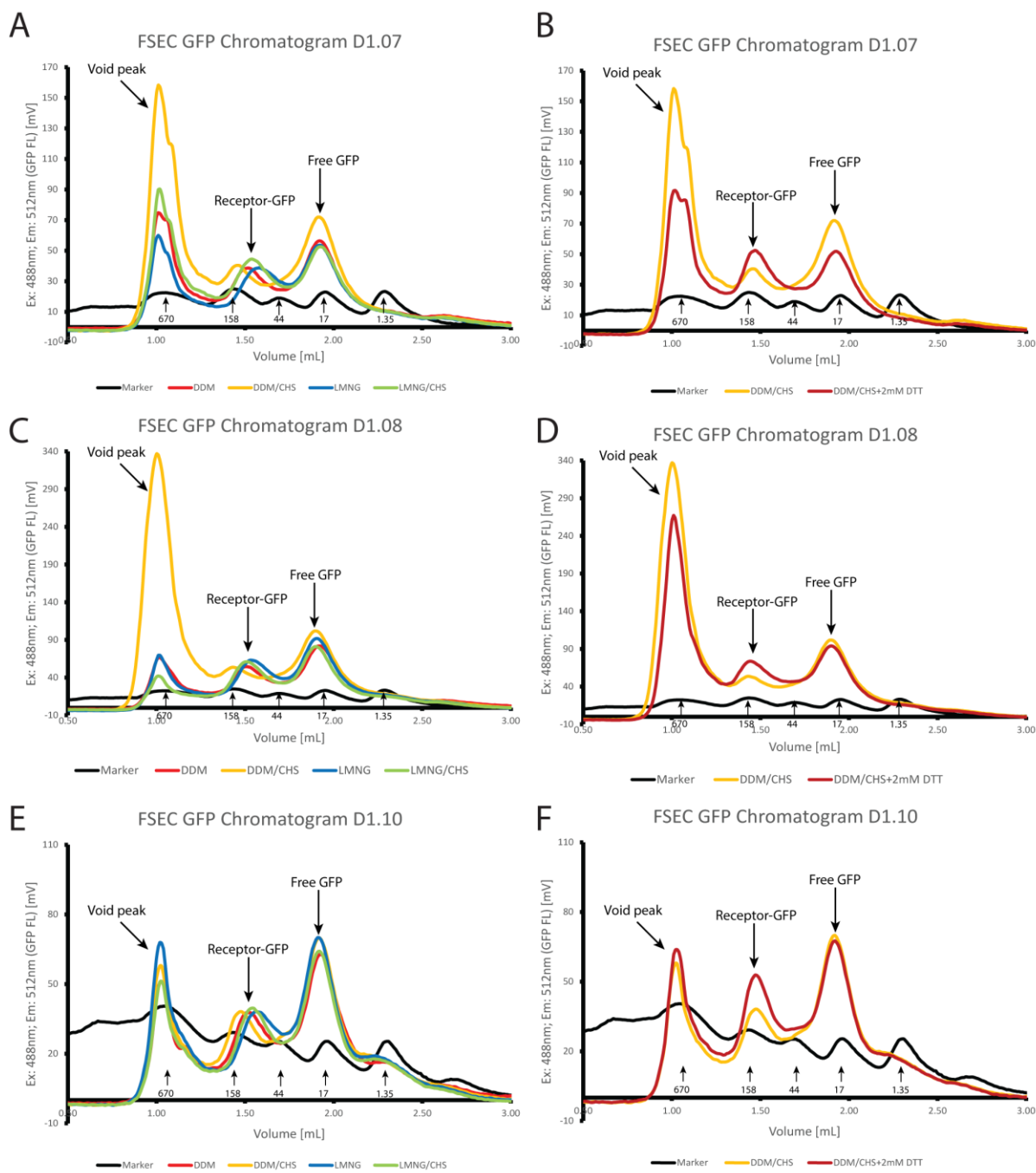
**3; Figure 14/15).** Samples solubilized in LMNG/CHS showed a promising combination of high receptor signals paired with good void to receptor peak ratios and will be used as detergent for subsequent experiments. In addition, beneficial effects were seen upon addition of 2 mM DTT to the buffer. On the contrary, addition of glycerol and 300 mM NaCl did not show any beneficial effects.

Receptor peak	DDM	DDM+ 10% glycerol	DDM+ 300mM NaCl	DDM+ 2mM DTT	DDM+Glycerol+ NaCl	DDM+Glycerol+ DTT	DDM+NaCl+DTT	DDM+Glycerol+ NaCl+DTT	DDM/CHS	DDM/C HS+10% glycerol	DDM/C HS+300 mM NaCl	DDM/C HS+2mM DTT	DDM/C HS+Glycerol+NaCl	DDM/C HS+Glycerol+DTT	DDM/C HS+NaCl+DTT	DDM/C HS+Glycerol+NaCl+DTT	LMNG	LMNG/CHS	DM	CHAPS
G6.05	14.2	15.4	27.6	14.6	27.4	19.9	24.3	26.8	25.1	26.4	34.0	25.3	29.4	25.8	32.6	31.6	17.1	19.2	18.1	14.2
G6.06	24.8	21.2	26.3	25.2	26.1	22.8	25.8	25.3	31.9	29.5	31.5	36.0	32.2	31.8	32.4	34.1	20.9	25.9	27.8	19.4
D1.07	38.7	36.5	34.5	51.9	37.0	41.2	53.4	51.5	40.4	37.6	37.2	52.5	39.0	52.4	67.4	61.4	38.8	44.5	39.0	20.9
D1.08	54.5	49.6	61.1	75.9	55.4	63.3	81.2	68.5	53.7	46.7	49.8	73.6	54.0	67.9	68.6	73.0	63.1	61.0	61.1	28.2
D1.10	37.7	31.8	35.5	45.5	37.1	42.6	46.5	43.8	38.1	33.7	35.5	52.9	38.2	42.8	47.1	46.6	38.2	39.6	33.9	26.5
Aggregate/Oligomer peak	DDM	DDM+ 10% glycerol	DDM+ 300mM NaCl	DDM+ 2mM DTT	DDM+Glycerol+ NaCl	DDM+Glycerol+ DTT	DDM+NaCl+DTT	DDM+Glycerol+ NaCl+DTT	DDM/CHS	DDM/C HS+10% glycerol	DDM/C HS+300 mM NaCl	DDM/C HS+2mM DTT	DDM/C HS+Glycerol+NaCl	DDM/C HS+Glycerol+DTT	DDM/C HS+NaCl+DTT	DDM/C HS+Glycerol+NaCl+DTT	LMNG	LMNG/CHS	DM	CHAPS
G6.05	58.0	154.5	300.1	81.0	335.2	216.7	267.2	337.5	133.2	209.4	270.9	133.1	264.9	191.4	234.7	296.5	69.9	61.0	81.3	35.3
G6.06	161.4	193.8	198.3	120.9	317.2	199.0	222.1	280.8	165.8	244.6	197.1	151.5	243.5	227.4	194.6	296.1	86.6	125.3	157.4	95.4
D1.07	74.8	171.3	220.0	88.7	296.9	199.3	180.5	257.8	158.3	175.3	212.5	91.7	178.7	150.2	108.3	197.6	60.0	90.3	50.4	31.0
D1.08	66.7	153.8	108.5	77.8	407.1	342.2	436.5	393.8	336.9	352.4	275.3	267.1	287.7	306.6	293.0	379.2	69.6	42.1	61.3	44.3
D1.10	51.1	129.0	108.4	50.4	191.1	149.0	63.4	145.8	58.0	166.1	78.2	63.8	128.2	122.9	83.9	164.5	67.9	51.2	62.7	50.1
Ratio Aggregate:Receptor peak	DDM	DDM+ 10% glycerol	DDM+ 300mM NaCl	DDM+ 2mM DTT	DDM+Glycerol+ NaCl	DDM+Glycerol+ DTT	DDM+NaCl+DTT	DDM+Glycerol+ NaCl+DTT	DDM/CHS	DDM/C HS+10% glycerol	DDM/C HS+300 mM NaCl	DDM/C HS+2mM DTT	DDM/C HS+Glycerol+NaCl	DDM/C HS+Glycerol+DTT	DDM/C HS+NaCl+DTT	DDM/C HS+Glycerol+NaCl+DTT	LMNG	LMNG/CHS	DM	CHAPS
G6.05	4.1	10.0	10.9	5.6	12.2	10.9	11.0	12.6	5.3	7.9	8.0	5.3	9.0	7.4	7.2	9.4	4.1	3.2	4.5	2.5
G6.06	6.5	9.1	7.5	4.8	12.2	8.7	8.6	11.1	5.2	8.3	6.3	4.2	7.6	7.2	6.0	8.7	4.1	4.8	5.7	4.9
D1.07	1.9	4.7	6.4	1.7	8.0	4.8	3.4	5.0	3.9	4.7	5.7	1.7	4.6	2.9	1.6	3.2	1.5	2.0	1.3	1.5
D1.08	1.2	3.1	1.8	1.0	7.3	5.4	5.4	5.7	6.3	7.5	5.5	3.6	5.3	4.5	4.3	5.2	1.1	0.7	1.0	1.6
D1.10	1.4	4.1	3.1	1.1	5.1	3.5	1.4	3.3	1.5	4.9	2.2	1.2	3.4	2.9	1.8	3.5	1.8	1.3	1.8	1.9

**Table 3: Detergent and Additive screening of Adrg6 and Adgrd1. Top:** Table showing the peak signal values for all constructs and conditions tested. Values are color coded with high values shown in red to low values shown in green. **Middle:** Table showing the values for the void volume peak (Aggregates/Oligomers) for all constructs and conditions tested. Coloring scheme is the same as the top table. **Bottom:** Table showing values of the ration between the void peak signal and receptor peak signal for all constructs and conditions tested. Coloring scheme is the same as the top table.

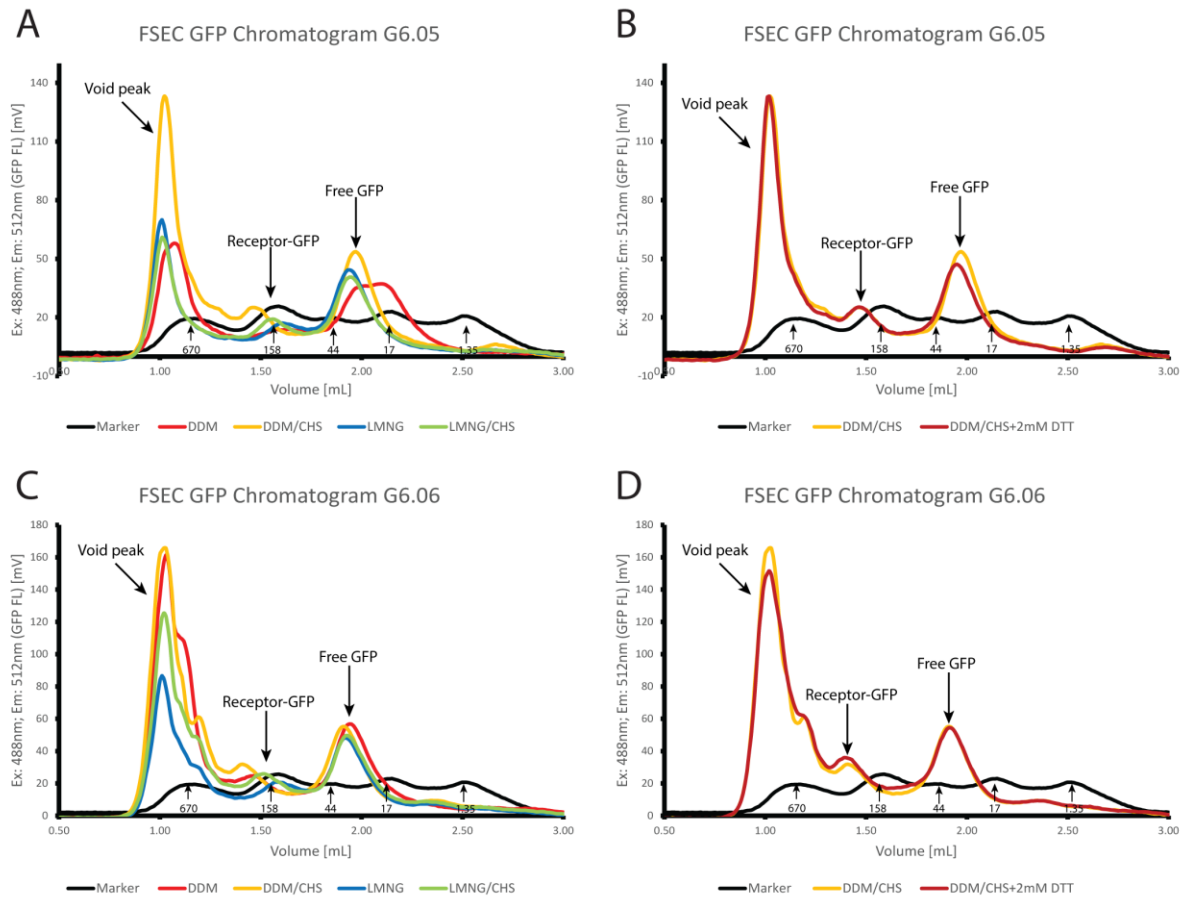
The detergent screening showed that all constructs and conditions tested suffer from aggregation of the receptor when extracted from the lipid bilayer, indicating that complex formation in the cell membrane might be essential to stabilize the receptor for purification. Having found suitable detergent conditions for solubilization, the next steps focused on purification of the receptor, complex formation and purification of the complex by co-expression for cryo-EM studies.





**Figure 14: Fluorescence SEC chromatograms from the detergent screening of D1.07/.08/.10.** X-Axis is the volume in mL and Y-axis is GFP fluorescence signal in mV. Different peaks are labeled in the figures. Peaks from the marker are labeled with arrows and numbers that show the MW in kDa. **A/B:** SEC chromatograms of construct D1.07. **C/D:** SEC chromatograms of construct D1.08. **E/F:** SEC chromatograms of construct D1.10.



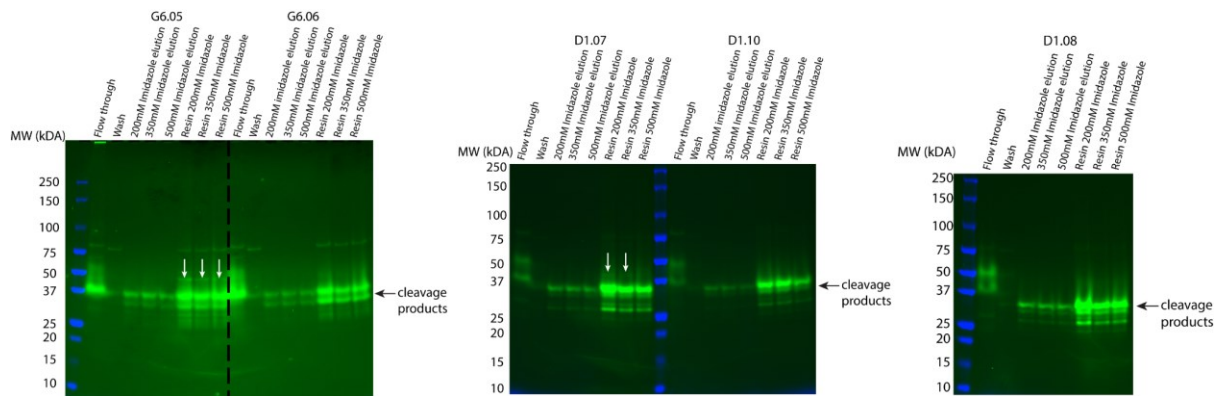


**Figure 15:** Fluorescence SEC chromatograms from the detergent screening of G6.05/.06. X-Axis is the volume in mL and Y-axis is GFP fluorescence signal in mV. Different peaks are labeled in the figures. Peaks from the marker are labeled with arrows and numbers that show the MW in kDa. A/B: SEC chromatograms for construct G6.05. C/D: SEC chromatograms for construct G6.06.

#### 4.1.1.3 Purification of Adgrd1 and Adgrg6

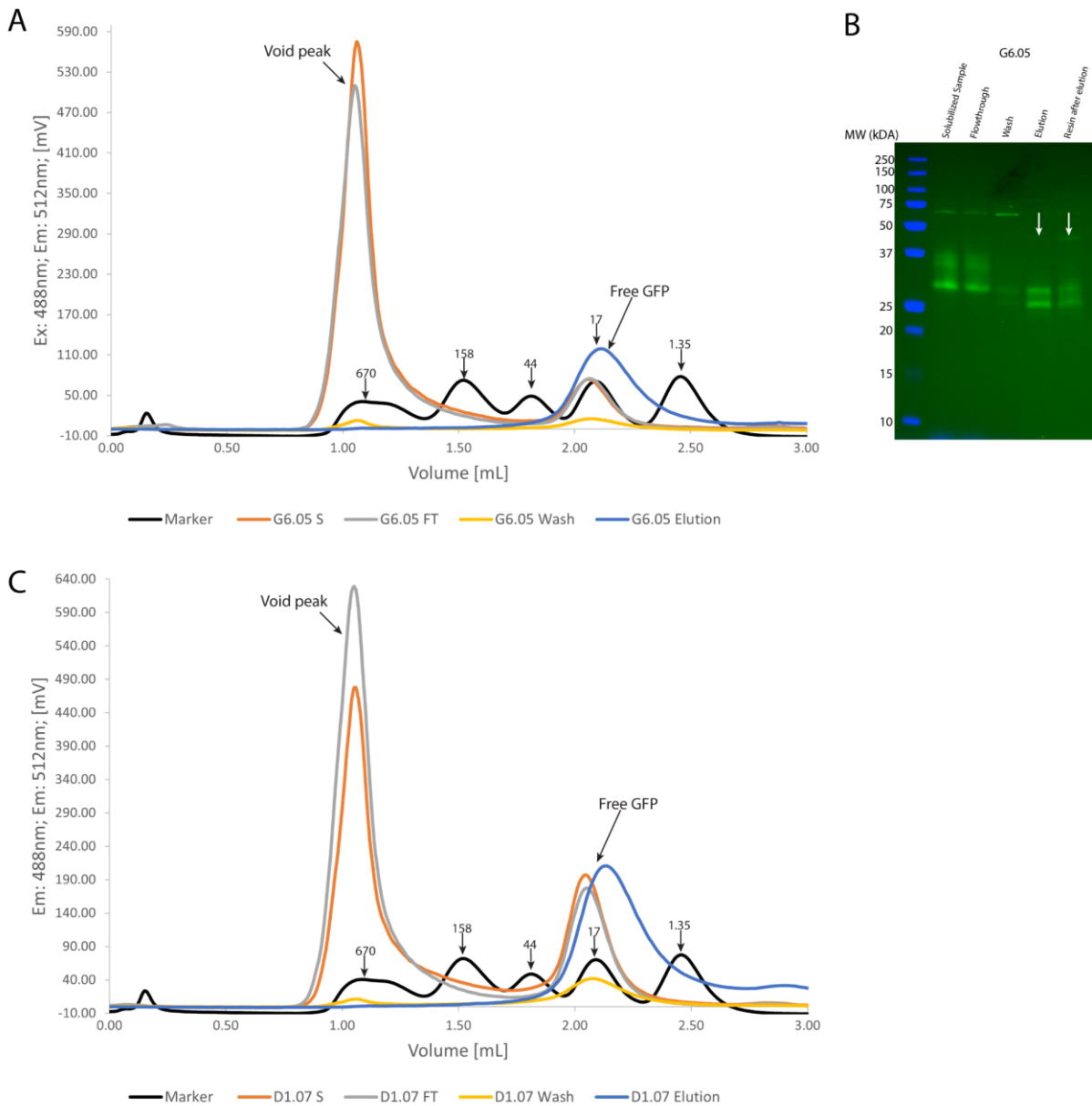
Our constructs for expression in insect cells showed clear expression of the receptors, but the FSEC experiments showed that there is a lot of aggregation, indicating that the receptors were not stable when extracted from the plasma membrane. Even with the best detergent conditions concluded from the detergent screening experiment, there was still usually a high signal for the void peak.

Despite that, we tried to purify the receptors using different strategies. Our constructs included a FLAG tag and His tag that can be used for purification. For the first purification trial, TALON resin was used to capture His tagged proteins. Constructs G6.05/.06 and D1.07/.08/.10 were used for this experiment using only a small cell pellet (0.5 g). The elution step was done with different imidazole concentrations (200, 350 and 500 mM) and the resin was also loaded on SDS-PAGE for analysis in case the protein can not be eluted. The purification trial was not successful and only for G6.05 and D1.07 (white arrows in **Figure 16**) a band could be seen for the receptor in the resin samples, but the elution samples did not show a band, meaning that the protein could not be eluted for any construct. In all of the samples a strong fluorescence band could be detected around ~30 kDa, that could correspond to free GFP with His tag (cleavage products) (**Figure 16**).



**Figure 16: Small scale purification trial with Adgrd1 and Adgrg6. A:** SDS-PAGE with in-gel fluorescence imaging of G6.05 and G6.06 showing samples taken from different purification steps. **B:** SDS-PAGE with in-gel fluorescence imaging of D1.07 and D1.10 showing samples taken from different purification steps. **C:** SDS-PAGE with in-gel fluorescence imaging of D1.08 showing samples taken from different purification steps.

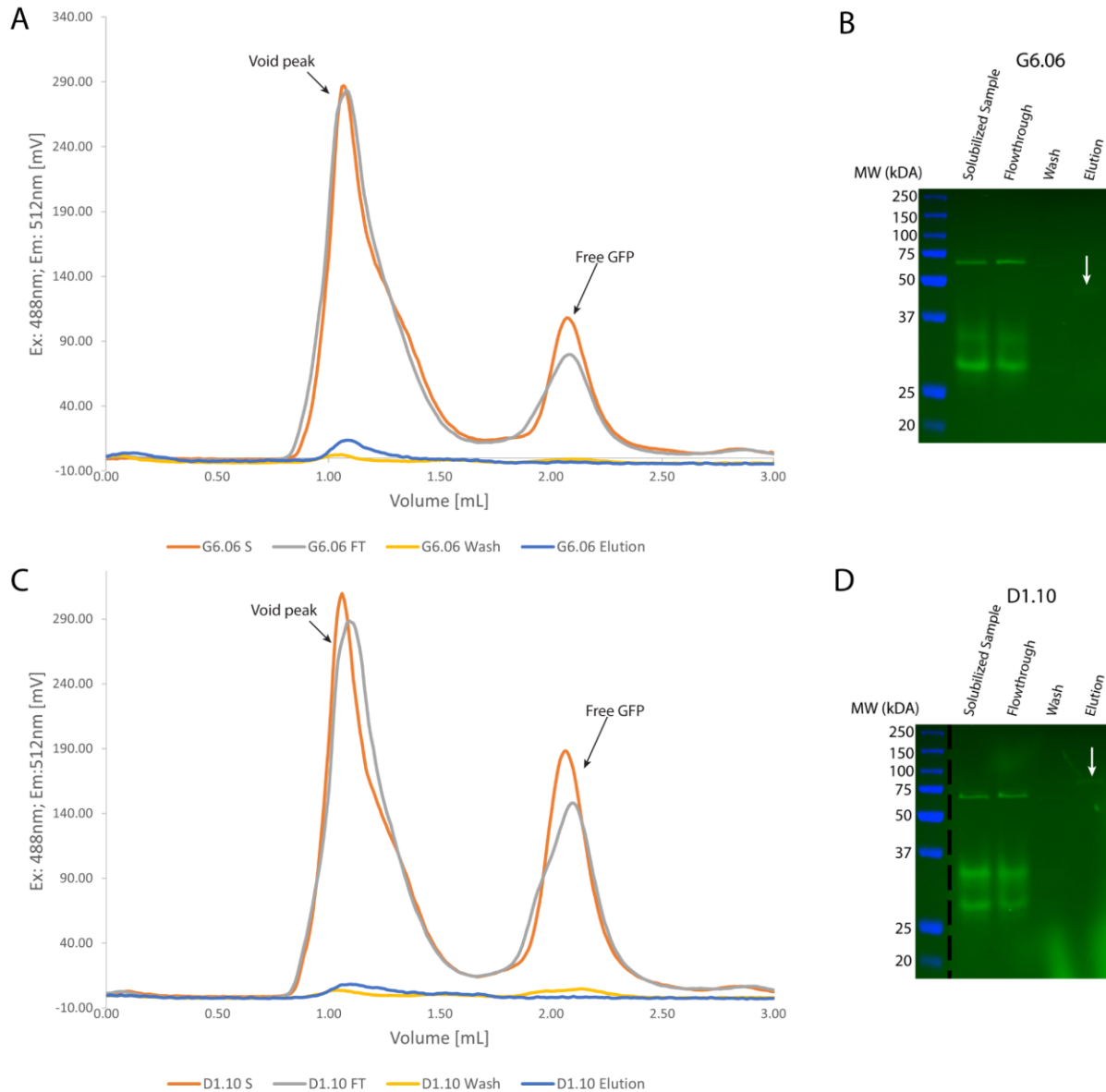
For the first purification trial only a small amount of cell pellet was used. In the next trial a bigger cell pellet (~12g) was used for constructs G6.05 and D1.07 where a band was detected in the resin samples in the previous experiment. Unfortunately, the results were similar to the previous small-scale purification test. However, the SDS-PAGE for G6.05 did show a faint band in the resin sample and even in the elution a very faint band was detected (**Figure 17B**). Additionally, the samples were analyzed with FSEC. The solubilized samples for both constructs did not show a clear peak like we saw in the detergent screening (**Figure 17A/C**). The elution sample only showed a peak for the free GFP but not for the receptor. Overall, the FSEC results of both constructs looked very similar (**Figure 17A/C**), indicating that purification of either receptor in the apo-state is difficult or not possible.



**Figure 17: Purification trial for G6.05 and D1.07 using TALON resin. Abbreviations: S=Solubilized sample; FT=Flow-through sample. A:** FSEC chromatogram for G6.05 showing samples taken from different purification steps. **B:** SDS-PAGE of G6.05 with in-gel fluorescence imaging showing samples from different purification steps. **C:** FSEC chromatogram for D1.07 showing samples taken from different purification steps.

The purification trials with TALON resin did not look promising. To exclude any negative effects from the purification method or the selected constructs, purification with another affinity column and constructs were done. This time anti-FLAG resin was used for purification tests with constructs G6.06 and D1.10. Different purification steps were again assessed using FSEC and SDS-PAGE. A cell pellet of ~12g was used for the experiment. Elution of the protein from the FLAG column was done by adding chemically synthesized FLAG peptide to the elution buffer. Similar to the purification trials with G6.05 and D1.07 and TALON resin, no clear peak (or band on SDS-PAGE) for the receptor was detected in the solubilized samples. FSEC analysis showed again a large void peak for the solubilized sample, indicating that most of the proteins are aggregated (**Figure 18A/B**). On SDS-PAGE the elution samples of both

constructs showed faint bands for the receptor, but the FSEC data shows that the peak is in the void volume (**Figure 18A/C**). In contrast to the purification with the TALON resin no fluorescent bands between 37-25 kDa could be detected, hinting that these bands correspond to free GFP (and cleavage products) that is his tagged and bound to the TALON resin.



**Figure 18: Purification of G6.06 and D1.10 using anti-FLAG resin.** Abbreviations: S=Solubilized sample; FT=Flow-through sample. **A:** FSEC chromatogram of G6.06 showing samples taken at different purification steps. **B:** SDS-PAGE with in-gel fluorescence imaging of G6.06 samples. Samples are taken from different purification steps. **C:** FSEC chromatogram of D1.10 showing samples taken at different purification steps. **D:** SDS-PAGE with in-gel fluorescence imaging of D1.10 samples. Samples were taken from different purification steps.

To summarize, these purification trials showed that purification of the receptor alone in the apo-state is not feasible. For this reason, we decided to change the strategy and focus on the co-expression of the receptor with the G protein.

#### 4.1.1.4 Co-expression screening

Several small-scale purifications of Adrg6 and Adrd1 were not successful as the protein was precipitating/aggregating during the purification procedure. Therefore, we decided to follow the strategy that other groups used for structure determination of GPCR-G protein complexes. In this strategy the receptor and G protein are expressed in the same cells, the receptor is activated, and the complex formed on the lipid bilayer. Subsequently, the complex is extracted from the cell membrane by detergents and purified for cryo-EM studies. Both Adrg6 and Adrd1 were reported to couple and activate the G<sub>s</sub> protein [47, 138].

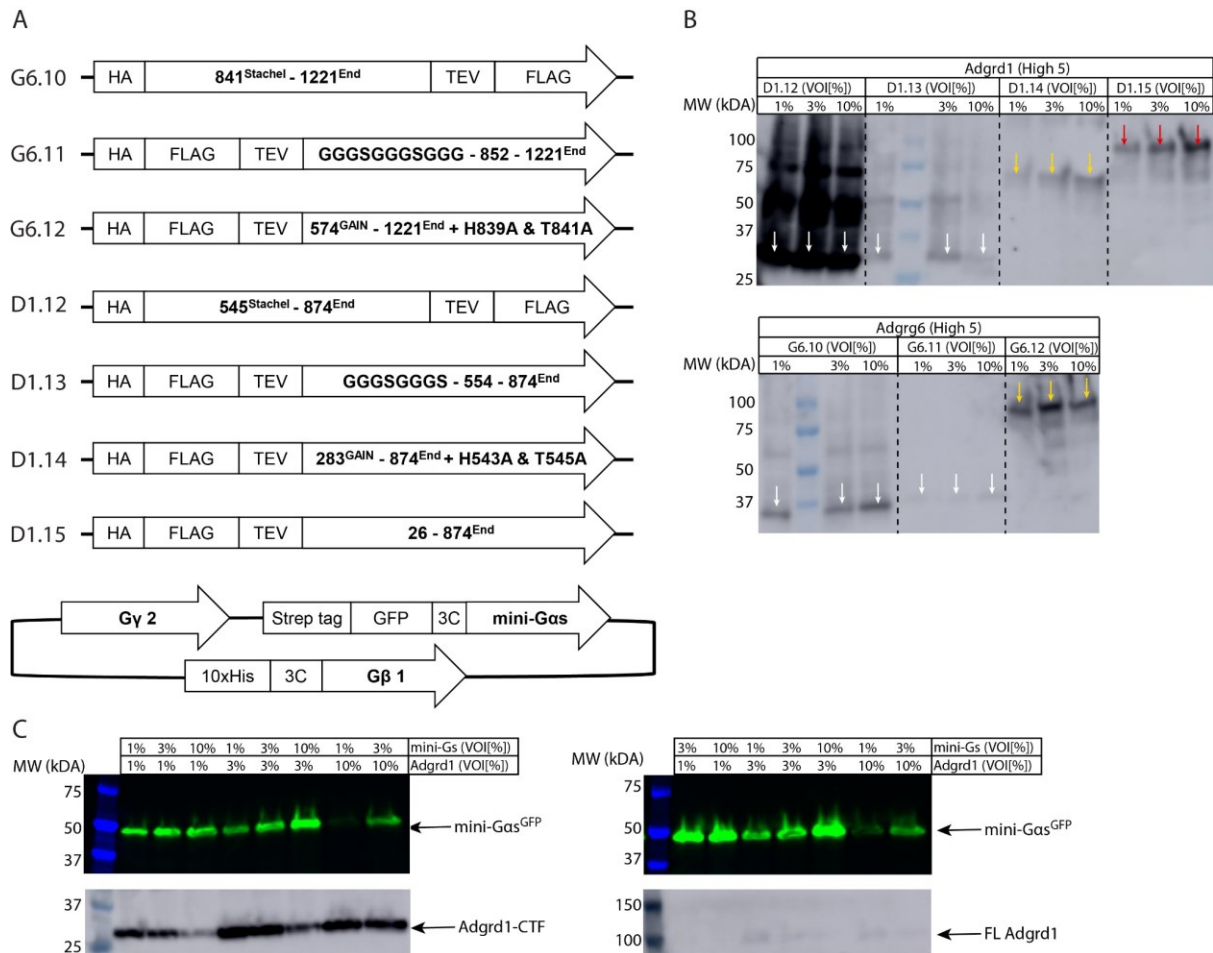
First, we designed new expression constructs for both receptors and a G<sub>s</sub> heterotrimer using a minimal G<sub>α<sub>s</sub></sub> subunit (mini-G<sub>s</sub>) [120]. Design of the receptor constructs followed the same rational as before with G6.10/.11 and D1.12/.13 consisting of the CTF but for G6.11 and D1.13 the Stachel sequence was replaced with a flexible Gly/Ser linker (**Table 4, Figure 19A**). G6.12 and D1.14 included the GAIN domain in addition to the CTF but the GPS motif was mutated to result in an uncleavable GAIN domain (**Table 4, Figure 19A**). Lastly, for Adrd1 a full-length receptor was also tested for expression (D1.15) (**Table 4, Figure 19A**).

	Signal peptide	N-term. tag and cleavage site	Ecto domain	Stachel sequence	C-term. cleavage site and tag	Activity
<b>G6.10</b>	Hemagglutinin tag (HA)	FLAG-TEV (G6.11/.12)		THFGVLMDLPR	TEV-FLAG (Only for G6.10 and D1.12)	Activatable
<b>G6.11</b>			-	GGGSGGGSGGG		Active
<b>G6.12</b>			GAIN domain-	AF--AHFGVLMDLPR		Activatable
<b>D1.12</b>	Hemagglutinin tag (HA)	FLAG-TEV (D1.13-.15)		TNFAILMQV		Activatable
<b>D1.13</b>			-	GGGSGGGSGGG		Active
<b>D1.14</b>			GAIN domain-	AL- ANFAILMQV		Activatable
<b>D1.15</b>			Full length	TNFAILMQV	Activatable	

**Table 4:** Construct design for Adrg6 (G6.10-.12) and Adrd1 (D1.12-.15) for co-expression with G protein.

For the G protein, we designed a plasmid for the expression of the G<sub>s</sub> heterotrimer with the mini-G<sub>α<sub>s</sub></sub> fused to a N-terminal GFP and Strep tag, Gβ<sub>1</sub> containing a C-terminal decahistidin tag, and an unmodified Gγ<sub>2</sub>. First, expression was tested for the new receptor constructs in insect cells. Western blot analysis probed with an anti-FLAG antibody showed expression for all constructs, but G6.11 and D1.13 with the mutated Stachel sequence showed very low expression levels (**Figure 19B**). D1.12 (Adrd1<sup>CTF</sup>) was clearly the best expressing construct and therefore the most promising for co-expression and subsequent purification of the Adrd1<sup>CTF</sup>-mini-G<sub>s</sub> complex. Co-expression screening was tested for G6.10/.12 and D1.12/.15 in combination with the G<sub>s</sub> heterotrimer in insect cells. Different virus concentrations and ratios (Receptor [VOI] : mini-G<sub>s</sub> [VOI]) were tested and analysis was done with SDS PAGE and in gel fluorescence imaging for the GFP fused to the alpha subunit and western blot analysis using an anti-FLAG antibody for detecting the receptor. Expression of the mini-G<sub>α<sub>s</sub></sub> was detected in almost all conditions tested with varying levels depending on the virus concentration and the ratio to the receptor. Co-expression with the Adrg6 receptor was not

successful because only very few conditions showed expression of the receptor. For D1.15 only very faint bands could be detected in the western blot (**Figure 19C**). In contrast, D1.12 showed clear expression in all conditions but with varying levels (**Figure 19C**). The best condition for both expression of the G protein and receptor was using a VOI of 3% for both baculoviruses.



**Figure 19: Co-expression screening of Adgrg6/Adgrd1 in combination with a Gs heterotrimer. A:** Construct design for Adgrg6 (G6.10-12), Adgrd1 (D1.12-15) and the Gs heterotrimer. **B:** Western blot analysis using an anti-FLAG antibody probing for expression of the receptor constructs. White arrows indicate CTF receptor constructs, yellow arrows indicate CTF+GAIN receptor constructs, and red arrow indicates the full length receptor. **C:** Co-expression screening of D1.12 (Adgrd1 CTF) and D1.15 (Full length Adgrd1). In gel fluorescence imaging show mini-Gas expression and western blot using an anti-FLAG antibody show receptor expression. Left shows the results for the D1.12 co-expression with mini-Gs and the right panel shows the results for the D1.15 co-expression with mini-Gs.

The co-expression screening looked promising for D1.12 but not as much for the other constructs. Nevertheless, in the next steps the focus was on forming the complex and purifying it for cryo-EM.

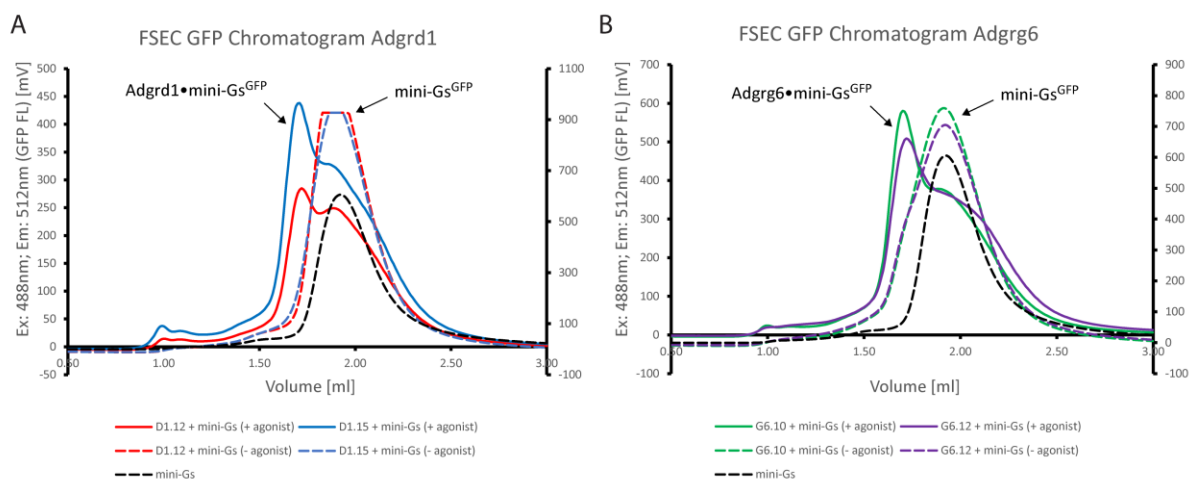
#### 4.1.1.5 Complex formation

Next, we tested if our receptor constructs are active and couple to the G protein or if we must add synthesized Stachel peptide to the buffer in order to activate the receptors. Additionally, our collaborators, Prof. Adriano Aguzzi and his group, screened cortisol and other potential agonists from Ping *et al.* [139] for activation of Adgrg6 or Adgrd1. Unfortunately, none of these ligands showed an



activating effect for Adgrg6 or Adgrd1. Therefore, the only agonist we know of are the corresponding Stachel peptides.

Instead of co-expression and subsequent complex formation in the cell membrane, we decided to use another strategy for this experiment. We expressed the G protein and receptor separately and then mixed the cells and added detergent for solubilization. The solubilization buffer either included synthesized Stachel peptide or not. The solubilized samples were then analyzed with FSEC to check for coupling to the mini-G<sub>s</sub>. Interestingly, all constructs tested needed synthesized Stachel peptide to efficiently couple to mini-G<sub>s</sub> (**Figure 20**). Especially for the CTF constructs, this was surprising because in cellular signaling assays expressing the CTF leads to robust increase in cAMP levels [47]. In our experimental conditions, we work with proteins in detergents and an even more artificial system, compared to the cell assays, which can explain the need for synthesized Stachel peptide to efficiently couple to mini-G<sub>s</sub> and form a stable complex.



**Figure 20: Complex formation experiments with Adgrg6/Adgrd1 and mini-G<sub>s</sub>.** **A:** Analytical SEC chromatogram for complex formation of Adgrd1 and mini-G<sub>s</sub>. Black dashed line represents a mini-G<sub>s</sub> only sample, representing the control. Red lines represent samples with D1.12 and blue lines samples with D1.15. Solid lines include samples with synthesized Stachel peptide in the buffer and dashed lines represent samples without additional Stachel peptide. **B:** Analytical SEC chromatogram for complex formation of Adgrg6 and mini-G<sub>s</sub>. Black dashed line represents a mini-G<sub>s</sub> only sample, representing the control. Green lines represent samples with G6.10 and purple lines samples with G6.12. Solid lines include samples with synthesized Stachel peptide in the buffer and dashed lines represent samples without additional Stachel peptide

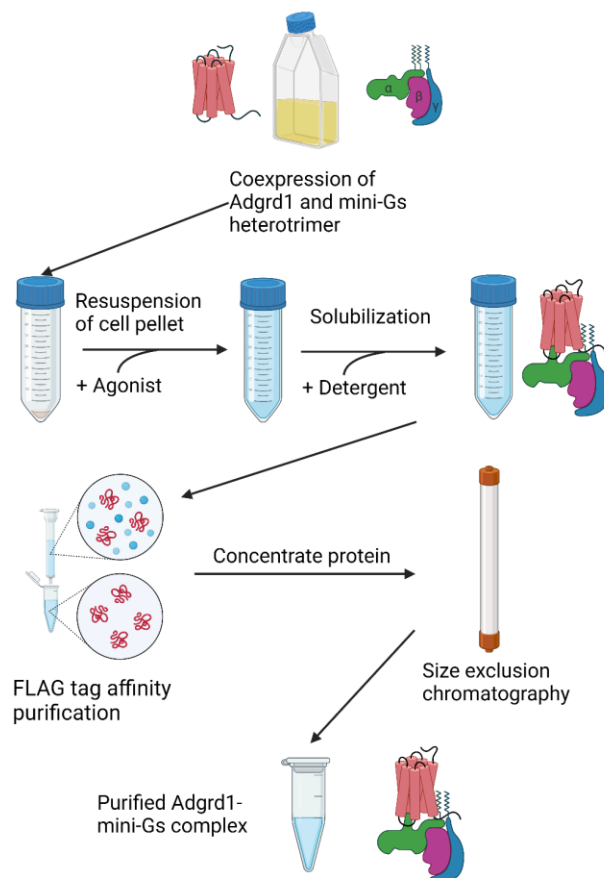
To summarize, this experiment showed that addition of chemically synthesized Stachel peptide is needed for activation of Adgrg6 and Adgrd1 and in order to couple to mini-G<sub>s</sub>. Moreover, coupling to G<sub>s</sub> was clearly shown for both receptors, meaning that a complex with G<sub>s</sub> can be formed for cryo-EM studies.

#### 4.1.1.6 Complex purification and cryo-EM grid preparation

The previous experiments showed that Adgrd1/Adgrg6 can be activated and couple to the G<sub>s</sub> protein. Co-expression experiments were not very promising for Adgrg6 but for Adgrd1, one construct (D1.12)

comprising the Adgrd1<sup>CTF</sup> showed clear expression. For complex purification we decided to focus on this construct.

To purify the complex, we co-expressed Adgrd1<sup>CTF</sup> (D1.12) with the mini-G<sub>s</sub> heterotrimer in High Five cells. Complex formation was done on the cell membrane by adding synthesized Stachel peptide to the resuspension buffer. After incubation, cells were solubilized with LMNG/CHS and the supernatant was clarified by centrifugation. Using a Flag tag affinity column and a size exclusion column step, the complex was purified, and the purified sample was used to prepare cryo-EM grids for single particle analysis (SPA) (**Figure 21**). The SEC profile of the first purification was not ideal as it showed a lot of aggregation (**Figure 22A**). We further analyzed the SEC fractionized samples by SDS-PAGE in order to identify the fraction containing the most pure and complete complex, i.e. showing all the bands of the protein components with least protein contaminant (**Figure 22B**). All components of the complex were clearly visible on SDS-PAGE and we continued with these samples to prepare cryo-EM grids (**Figure 22B**). Cryo-EM grids were prepared together with Dr. Ching-Ju Tsai from the Schertler group (PSI). SEC fractions used for grid preparation had a concentration of 1 mg/ml and were not further concentrated after SEC. Several grids were prepared with different grid types, blotting force and blotting time. All grids were prepared with a Vitrobot Mark IV at 4°C and 100 % humidity. These grids were analyzed by



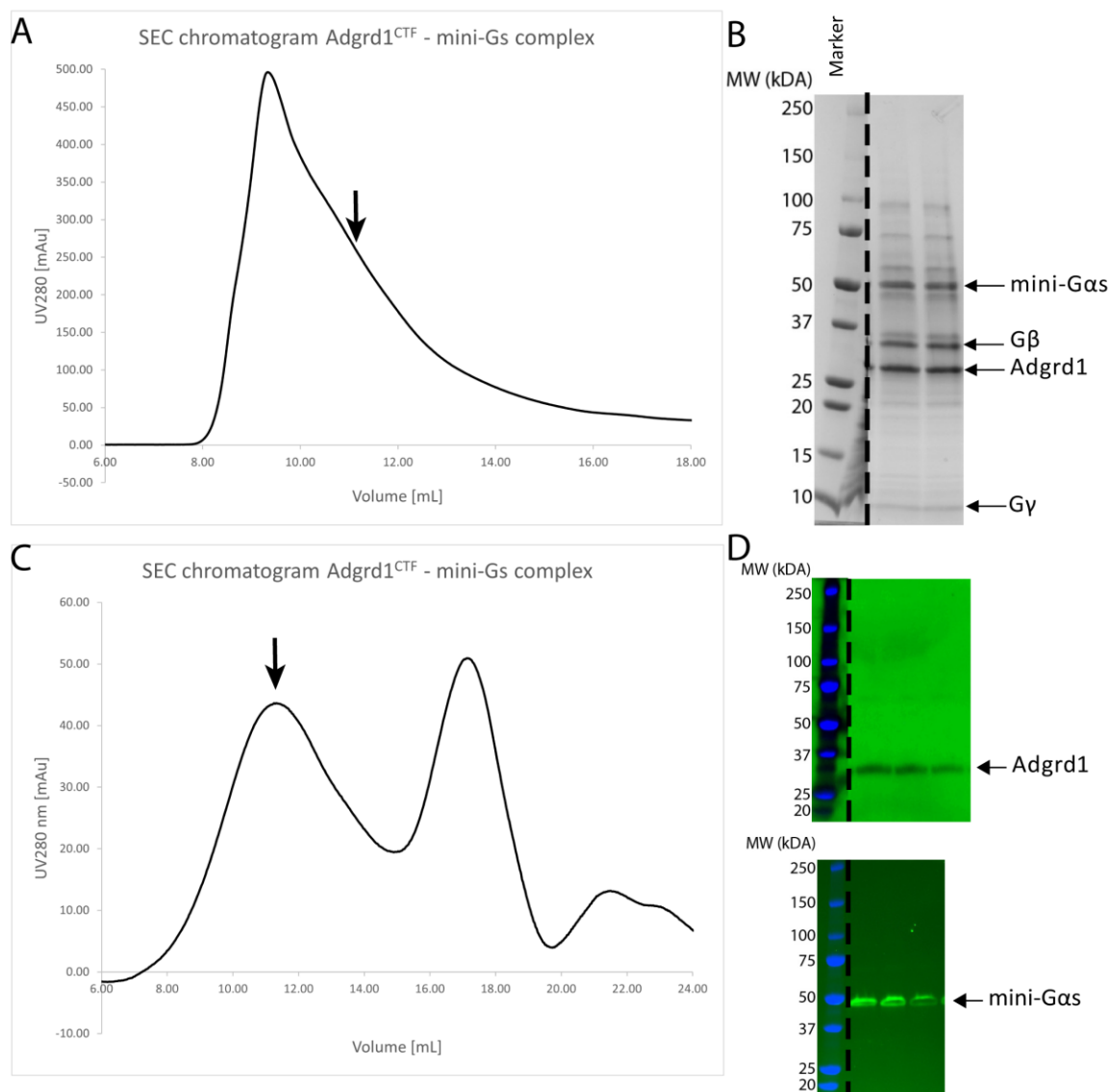
**Figure 21: Purification scheme of the Adgrd1<sup>CTF</sup>-mini-G<sub>s</sub> complex.** Figure created with Biorender.

cryo-EM (see next section).

The first purification did not look very pure, and a lot of aggregation was detected on SEC. Based on the GPR97 structure [139]), we decided to change the buffer compositions and included GDN in the detergent mix with LMNG/CHS and 10% glycerol in all buffers (except SEC buffer). The SEC profile looked better with a single peak (**Figure 22C**). However, the yield was significantly lower than in the first purification based on the signal from the SEC. This



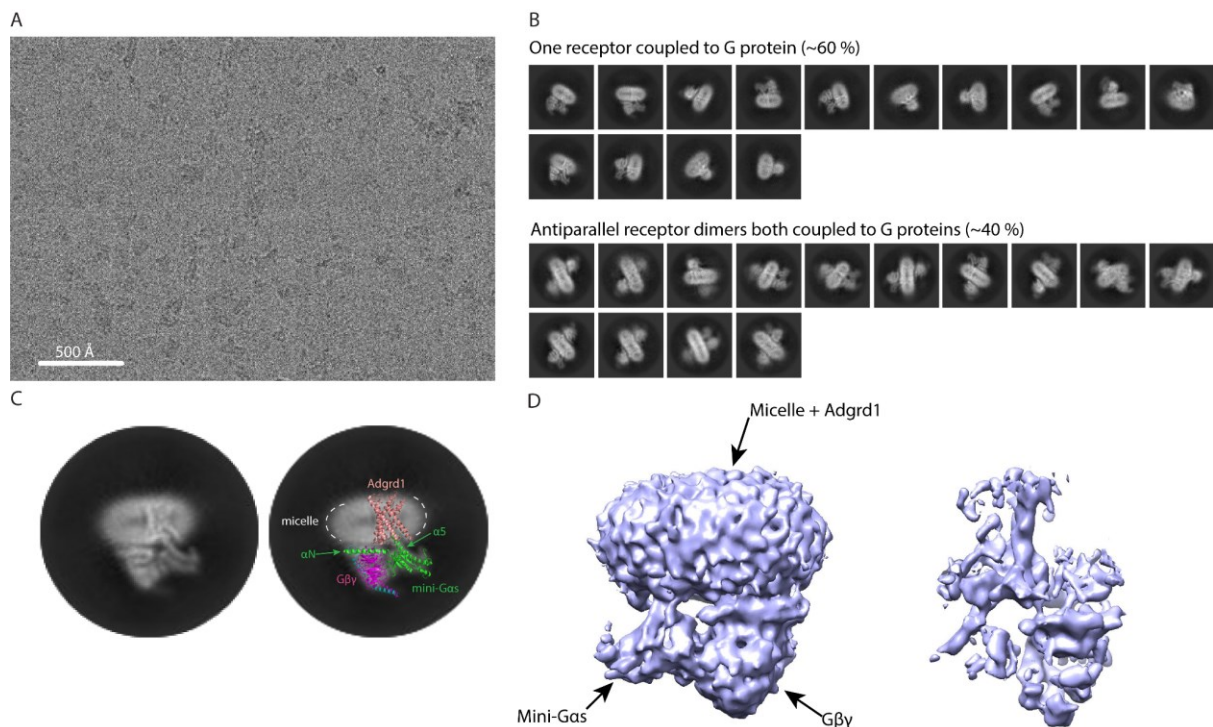
could have two reasons. First, the baculovirus was older than the one used for the first purification, meaning that it is likely not as strong anymore, reducing the expression levels of the proteins of interest. Second, the change in the buffer composition might have influenced the final yield. Because of the low yield, SDS-PAGE with Coomassie staining did not show clear bands. In gel fluorescence imaging was done to detect the mini-G $\alpha_s$  and western blot using an anti-FLAG antibody to detect Adgrd1<sup>CTF</sup> (Figure 22D). Nevertheless, enough protein was purified to prepare cryo-EM grids for analysis. Peak fractions were pooled and concentrated to 2.2 mg/ml. Cryo-EM grids were prepared together with Dr. Ching-Ju Tsai. Grids were again prepared with a Vitrobot Mark IV at 4°C and 100 % humidity. Different grid types, blotting force and blotting times were used. These grids were analyzed by cryo-EM.



**Figure 22: Purification of the Adgrd1<sup>CTF</sup>-mini-Gs complex.** **A:** SEC chromatogram of the first purification. The black arrow indicates the fraction collected for cryo-EM grid preparation. **B:** SDS-PAGE with Coomassie staining. Individual subunits of the complex are marked. **C:** SEC chromatogram of an optimized purification. Black arrow indicates the complex peak from which fractions were collected and used for cryo-EM grid preparation. **D: Upper panel:** Western blot analysis using an anti-FLAG antibody to detect Adgrd1 receptor. **Lower panel:** In gel fluorescence imaging of SDS-PAGE detecting the mini-Gs coupled to GFP.

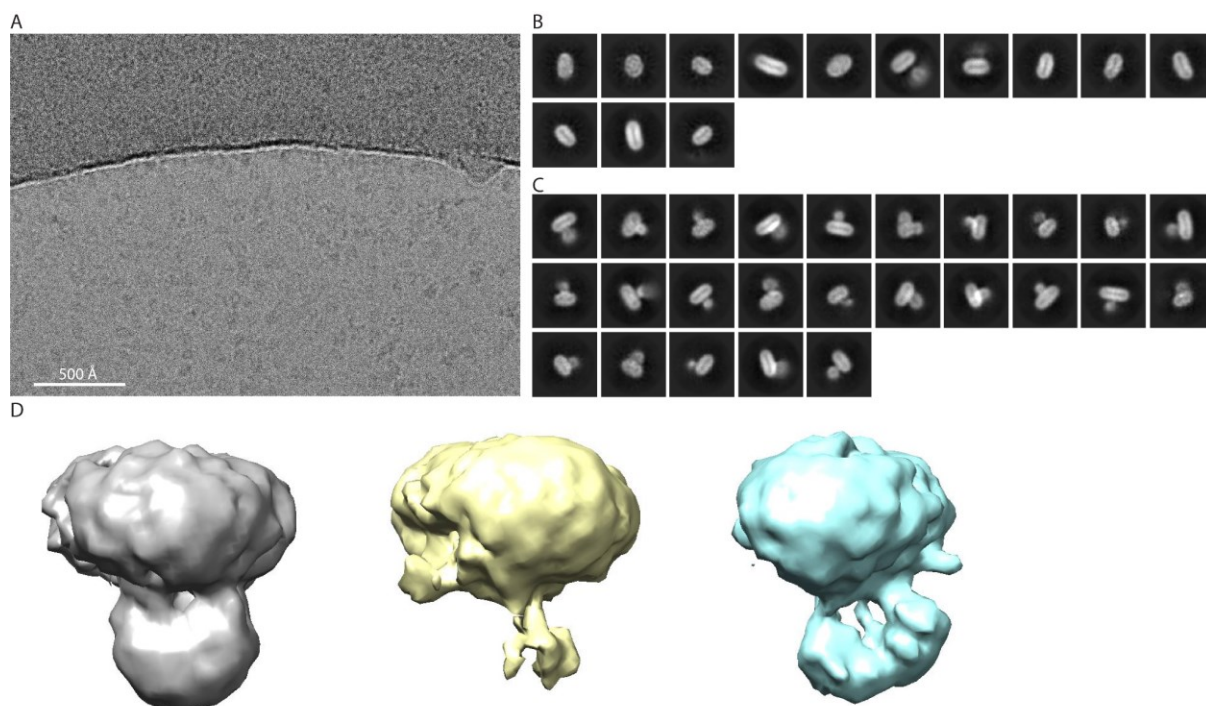
#### 4.1.1.7 Cryo-EM data collection and analysis

The first dataset we collected was from the first purification where the sample showed much aggregation (**Figure 22A/B**). Cryo-EM single particle analysis allows picking and sorting of the particles from the data collected, meaning that the sample does not have to be as pure as for crystallization because bad or undesired particles can be classified and discarded during processing. Together with Dr. Jacopo Marino (Schertler group, PSI), grids were screened, and data was collected on a Titan Krios at ScopeM ETHZ. In total two datasets were collected from two different grids. In both cases the ice thickness was suboptimal, indicating that grid preparation can be improved with longer blotting times for example. **Figure 23A** shows an example of an EM micrograph. The two datasets were combined and processed together. After particle picking, extraction and 2D classification, 2D class averages showing our complex of interest were seen (**Figure 23B**). In the best 2D class averages all components of the complex can clearly be seen and some show secondary structure features like the  $\alpha 5$  or  $\alpha N$  helix of the mini- $G\alpha_s$  (**Figure 23C**). Overlaying a 2D representation of the now published Adgrd1<sup>CTF</sup>-mini-  $G_s$  complex from Ping *et al.* [68] shows good agreement with our 2D class averages (**Figure 23C**). In addition to monomeric complex particles, receptor dimers, both coupled to G proteins, were present in the sample (**Figure 23B**). The 2D class averages revealed that the dimers were in an antiparallel orientation, indicating that these dimers are not of physiological relevance but more likely an artifact of the detergent, respectively the solubilization step (**Figure 23B**). Because the protein concentration of this sample was low with 1 mg/ml only few particles per micrograph were present (**Figure 23A**) and less than 100'000 particles classified into the best 2D class averages. From this low number of particles, it was not possible to generate a high-resolution 3D map. The ab-initio maps showed density for all components, mini- $G\alpha_s$ ,  $G\beta\gamma$  and the micelle with the receptor (**Figure 23D**) but because of the low number of good particles, the density for the transmembrane helices was not well resolved or the density was not present (**Figure 23D**). The cryo-EM analysis of this sample indicated that the sample and grid preparation need further improvement.



**Figure 23: Cryo-EM analysis of the first Adgrd1<sup>CTF</sup>-mini-Gs complex sample.** **A:** Example of an EM micrograph. **B:** 2D class averages showing one receptor coupled to G protein (top) or antiparallel receptor dimers both coupled to G protein (bottom). Box size is 224x224 pixels with a pixel size of 1.02 Å. **C:** Example of a 2D class average with secondary structure elements visible (left). Overlay of the Adgrd1-mini-Gs model from Ping et al. [68] with the 2D class average (right). Components of the complex and visible secondary structure elements are labelled. **D:** 3D experimental map of the complex with a lower contour level (left) and a higher contour level (right). Individual subunits are labelled in the map on the left.

The sample prepared for the next cryo-EM study showed an improved SEC profile (**Figure 22C**) but the yield was lower and the protein sample had to be concentrated to 2.2 mg/ml for grid preparation. Data was collected with Dr. Pavel Afanasyev (Cryo-EM Knowledge Hub, ETHZ) at ScopeM ETHZ on a Titan Krios. Grids showed improved ice thickness compared to the grids from the first sample. Particles showed improved contrast because of thinner ice layers (**Figure 24A**). However, after particle picking and 2D classification it became clear that many complex particles dissociated because many 2D classes showed only the detergent/receptor micelle (**Figure 24B**). The 2D classes that showed the complex were not as good as the ones from the first sample. They did not show any high frequency information like secondary structure features (**Figure 24C**). Additionally, the 2D class averages also did not yield clear visualization of all the subunits. After several rounds of 2D classification only ~35'000 particles were present in these 2D classes. It was not possible to perform 3D reconstruction to high resolution from this data. Initial models did not show the individual subunits clearly (**Figure 24D**).



**Figure 24: Cryo-EM analysis of the second sample for the Adgrd1<sup>CTF</sup>-mini-Gs complex.** **A:** Example of an EM micrograph. **B:** 2D class averages showing only the micelle from potentially dissociated complex particles. Box size is 240x240 pixels with a pixel size of 1.02 Å. **C:** 2D class averages showing potential complex particles. Box size is 240x240 pixels with a pixel size of 1.02 Å. **D:** Ab initio experimental maps generated from the particles in the 2D class averages shown in C.

It was surprising that the cryo-EM results from the second sample were worse than from the first sample because the SEC profile and grids overall looked more promising. There were however differences after SEC compared to the first sample preparation. After SEC the protein was concentrated whereas the first sample was used directly after SEC for grid preparation. In addition, the settings for blotting the grids were different which might also account for some of the differences. Our results showed that it might have been better to stay with the first purification protocol and just try to improve the grid preparation to get better data for high-resolution structure determination.

We were in the process of optimizing the sample and grid preparation when several groups published structures of aGPCRs activated by the Stachel peptide [67-70]. The structure of Adgrd1<sup>CTF</sup> coupled to mini-G<sub>s</sub> heterotrimer was solved by two groups and for this reason we decided to not continue with sample improvement and further data collection. Overall, these two groups used a very similar approach like we did, validating that our strategy was correct.

#### 4.1.1.8 Summary aGPCR-G protein complex

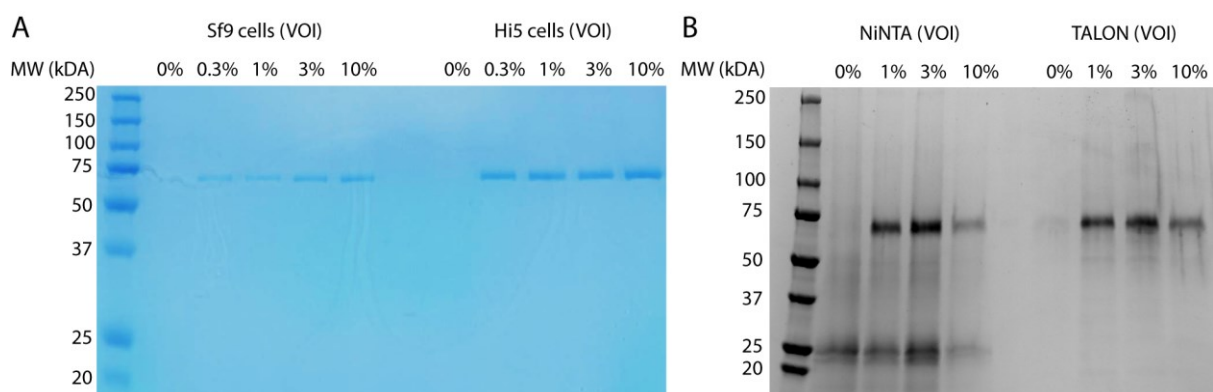
- Expression of Adgrg6 or Adgrd1 was not successful in HEK293F cells.
- Adgrg6 and Adgrd1 can be expressed in Sf9 and High Five insect cells with High Five cells showing higher expression levels.
- Adgrg6 and Adgrd1 are both unstable and aggregate when extracted from the cell membrane by detergents, making purification of the receptor alone not possible.
- Co-expression of mini-G<sub>s</sub> heterotrimer with Adgrd1, but not Adgrg6, worked well.
- Our constructs needed additional chemically synthesized Stachel peptide in order to be active and couple to mini-G<sub>s</sub>.
- Purification of an Adgrd1<sup>CTF</sup>-mini-G<sub>s</sub> complex was successful.
- Cryo-EM analysis showed 2D class averages showing the complex of interest and initial models showed clear density for all the subunits of the complex.
- Purification and grid preparation needed further improvement for high resolution structure determination.

### 4.1.2 Adgrd1-N terminal fragment

Adhesion GPCRs possess large extracellular domains that are involved in interactions with other partners. For Adgrg6 there was a crystal structure released of the N-terminal fragment (NTF) that gave insights into the regulation of Adgrg6 activity [140]. The information about the NTF of Adgrd1 was scarce when the project started in 2020. Through sequence analysis it was clear that Adgrd1 possess the GAIN domain and a PTX domain in the NTF. If there are additional domains or how the whole NTF is structured and can interact with potential ligands was unclear. We attempted to express the NTF in order to get structural and functional insights.

#### 4.1.2.1 Expression and purification of the Adgrd1-NTF

The full-length human Adgrd1<sup>NTF</sup> was cloned into the pAC8RED vector for insect cell expression. Since the NTF included a signal peptide for trafficking through the plasma membrane, the NTF is expressed as a secreted protein in insect cells. First expression tests were done in Sf9 and High Five cells with different VOIs. The supernatant was collected and used for small scale pull downs using either Ni-NTA resin or TALON resin. After precipitating the chelating reagents present in the expression medium with NiCl<sub>2</sub> and CaCl<sub>2</sub>, the clarified supernatant was mixed with the resin. After the wash steps, SDS buffer was directly added to the resin and the samples were analyzed with SDS-PAGE. The expression tests showed clear expression of the Adgrd1<sup>NTF</sup> in both cell lines and all VOIs tested. Generally, expression levels were higher in High Five cells compared to Sf9 cells (**Figure 25A**). There was not much difference in the expression levels when using different VOI for High Five cells (**Figure 25A**). For Sf9 cells the VOI had a bigger influence and expression levels increased with increasing VOI (**Figure 25A**). For the pull downs and comparison between the Ni-NTA and TALON resins, expression was done in High Five cells. The samples from the Ni-NTA resin showed a prominent contamination that was co-purified with the Adgrd1<sup>NTF</sup> and was also seen in the negative control (**Figure 25B**). On the other hand, the samples from the TALON resin showed a clear single band for the Adgrd1<sup>NTF</sup> accompanied with only minor contaminants (**Figure 25B**).

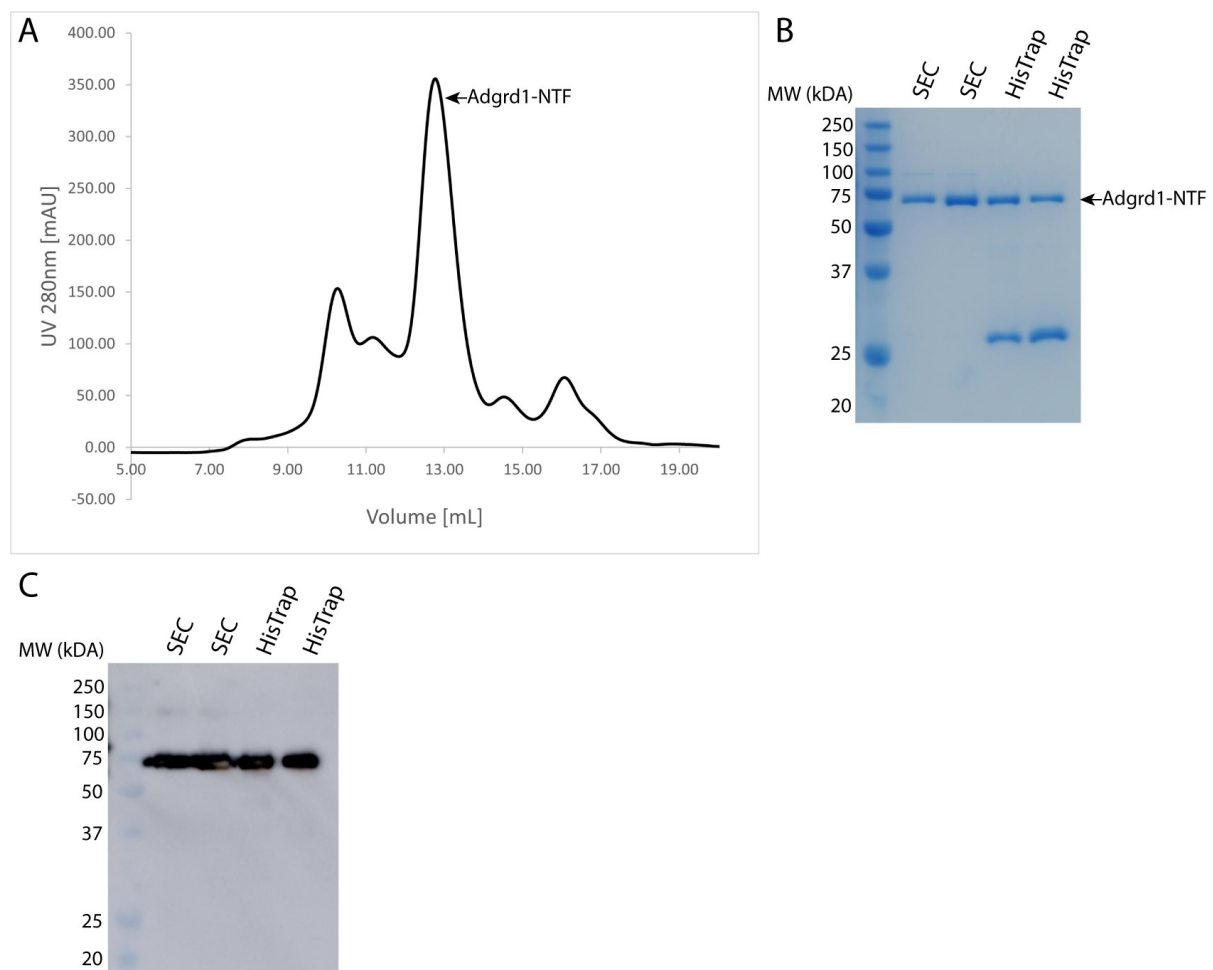


**Figure 25: Small scale expression and purification tests for Adgrd1<sup>NTF</sup>.** **A:** SDS-PAGE (Coomassie stained) with samples from test expressions in Sf9 and High Five cell using different VOI. **B:** SDS-PAGE (Coomassie stained) with samples from High Five cell expression and purified with either Ni-NTA resin or TALON resin.



The results showed that expression in High Five cells with a VOI of 3% works best. Even though the small scale pull downs with TALON resin were purer, we decided to use a Ni-NTA resin because we could use a pre-packed HisTrap Excel column where the Ni<sup>+</sup> ions can not be stripped by the chelating reagents from the medium, resulting in a simplified purification protocol. The purification protocol started with centrifugation of the cell suspension in order to harvest the supernatant from the cell culture, followed by filtration of the supernatant. Next, the supernatant was loaded onto a HisTrap Excel column. The column was washed and the protein was eluted with a buffer containing a high imidazole concentration. Eluted protein was then concentrated and injected into a SEC column. After SEC, the protein fractions corresponding to the right MW were pooled, concentrated and used for assays.

Several purifications were performed with different expression times (30h, 48h, 52h, 72h) which were all successful with yields ranging from 1-2 mg/L of cell culture (**Figure 26A/B/C**). The best expression time was 48 h. To confirm that the protein purified is really the Adgrd1<sup>NTF</sup>, a Western Blot was done using an antibody targeting the Adgrd1<sup>NTF</sup> (**Figure 26C**).

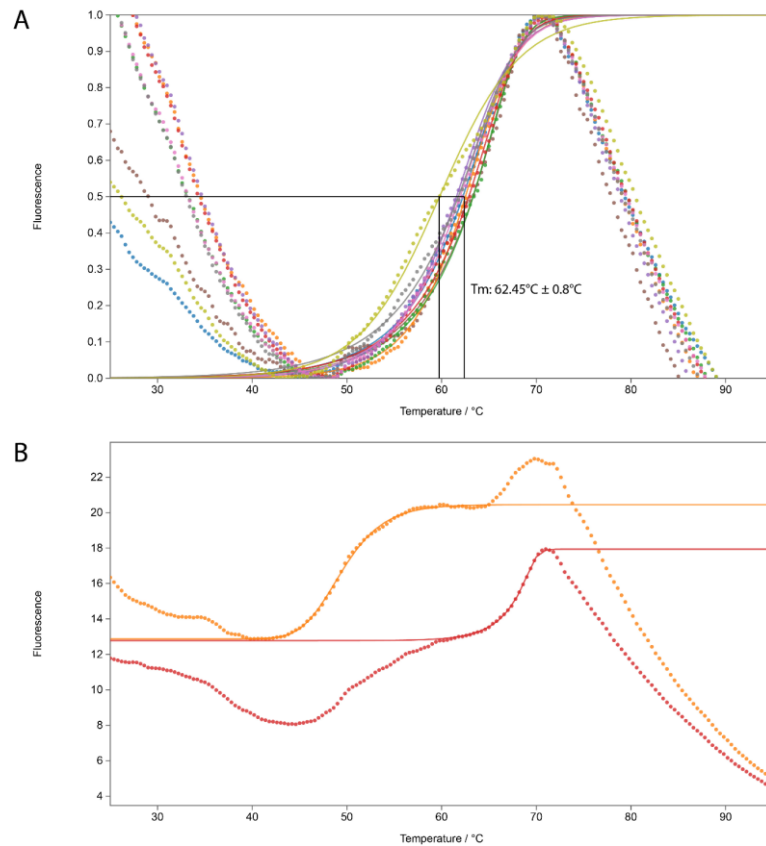


**Figure 26: Purification of the Adgrd1<sup>NTF</sup>.** **A:** SEC chromatogram from a purification of the Adgrd1<sup>NTF</sup>. **B:** SDS-PAGE (Coomassie stained) with samples from the SEC peak and from the HisTrap excel column elution fractions. **C:** Western blot using an antibody detecting the Adgrd1<sup>NTF</sup> showing the same samples as the SDS-PAGE in B.

The purified protein was used for thermal stability assays and for crystallization screens.

#### 4.1.2.2 Thermal stability assay for Adgrd1<sup>NTF</sup>

To characterize the stability of the Adgrd1<sup>NTF</sup> in different buffers, a thermal stability assay was done. The assay was done using a SYPRO orange dye that binds to hydrophobic moieties and exerts a fluorescence signal upon binding. Upon thermal unfolding of the protein, the dye binds to these exposed hydrophobic patches and fluorescence signal will increase which is measured.



**Figure 27: Thermal stability assay of Adgrd1<sup>NTF</sup>.** **A:** Thermal stability assay with Adgrd1<sup>NTF</sup> at different concentrations (1-10 μM). X-axis shows the temperature in °C and Y-axis shows fluorescence signal (normalized). **B:** Thermal stability assay with frozen and thawed protein at two different concentrations. Two melting curves can be fitted (red and orange). X-axis shows temperature in °C and Y-axis shows fluorescence signal.

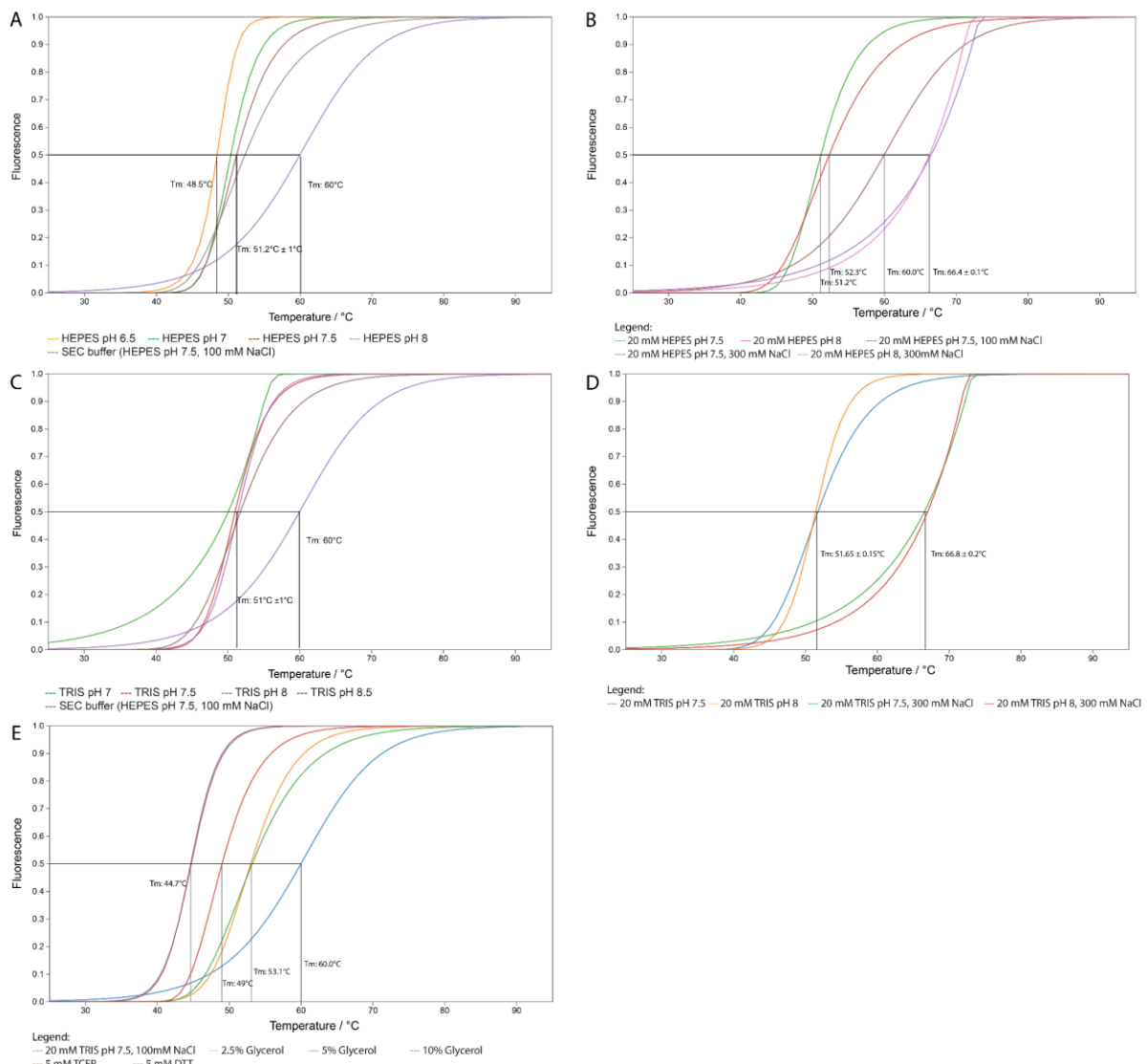
Initially, several samples were run in the SEC buffer (HEPES pH 7.5, 100 mM NaCl) with different protein concentrations to find the optimal concentration for the subsequent assays (**Figure 27A**). Protein concentrations from 1 μM to 10 μM were tested. All samples showed very similar melting temperatures (T<sub>m</sub>) of 62.45°C ± 0.8°C, except for the sample with 9 μM protein (light green curve in **figure 27A**) that presented an outlier with a T<sub>m</sub> of 59.8°C. The sample with a concentration of 7 μM protein gave the best signal to noise ratio and this concentration was used for all following assays. Additionally, the influence of freezing and thawing the protein sample was tested. When the

assay was performed with protein that was frozen and thawed, two separate melting curves were seen (**Figure 27B**). It is possible that one of the domains of the ECD is less stable after freezing and thawing the protein and that is the reason two melting curves were detected. For the following assays fresh protein was directly used after purification.

For the assay, HEPES and Tris buffers at different pH values were tested in combination with different NaCl concentrations. In addition, additives like glycerol, TCEP and DTT were also tested.



For samples in either Tris or HEPES buffer the  $T_m$  values were similar with  $51^\circ\text{C} \pm 1^\circ\text{C}$  and the pH value did not significantly affect the  $T_m$  (**Figure 28A/C**). Only HEPES buffer at pH 6.5 showed a lower  $T_m$  with  $48.5^\circ\text{C}$  (**Figure 28A**). Comparing samples with different NaCl concentrations showed that the salt concentration has a substantial influence on the thermal stability of the protein (**Figure 28B/D**). Samples with the highest concentration of NaCl tested (300 mM) had a  $T_m$  of  $66.4^\circ\text{C}$  and  $66.8^\circ\text{C}$  in HEPES or Tris buffer respectively (**Figure 28B/D**). On the other hand, samples without any NaCl had an approximately  $15^\circ\text{C}$  lower  $T_m$ , showing that NaCl stabilizes the protein (**Figure 28B/D**). The addition of reducing agents (DTT or TCEP) resulted in markedly decreased melting temperatures with  $44.7^\circ\text{C}$  (**Figure 28E**). Moreover, addition of glycerol reduced the  $T_m$  as well with  $53.1^\circ\text{C}$  (2.5% and 5% glycerol) and  $49^\circ\text{C}$  (10% glycerol) (**Figure 28E**).



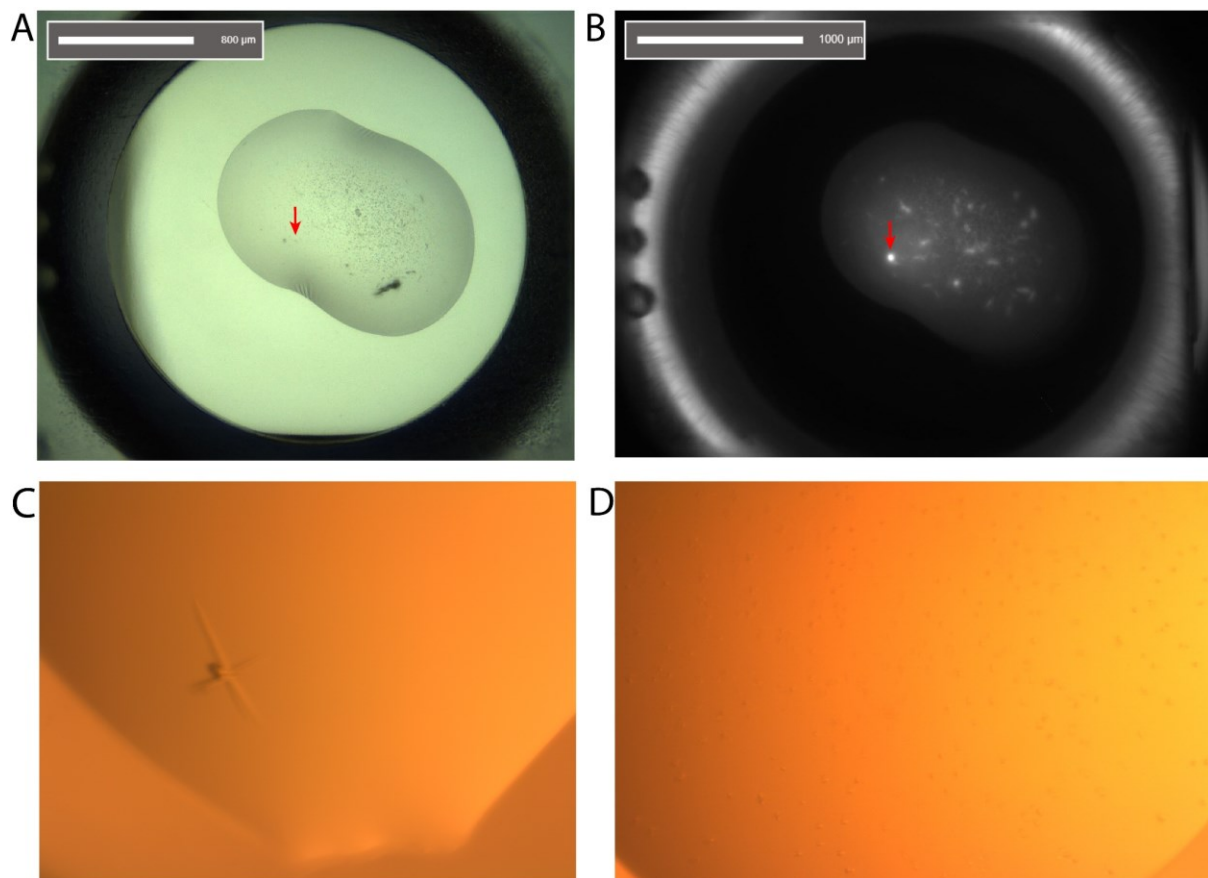
**Figure 28: Thermal stability assay of the Adgrd1<sup>NTF</sup>.** All plots show the fitted curves based on the raw data. Y-axis shows the normalized fluorescence signal and X-axis shows the temperature in  $^\circ\text{C}$ . **A:** Thermal stability screen testing HEPES buffer at different pH. **B:** Thermal stability screen testing different NaCl concentration in combination with HEPES buffer. **C:** Thermal stability screen testing TRIS buffer at different pH. **D:** Thermal stability screen testing different NaCl concentrations in combination with TRIS buffer. **E:** Thermal stability screen testing different additives.

Together, the results showed that the pH in the buffer systems tested does not influence the  $T_m$  greatly in the range from pH 6.5-8.5 whereas the salt concentration had a big influence. After successful purification and thermal stability analysis of the Adgrd1<sup>NTF</sup>, we decided to try and crystallize the protein for structure determination.

#### 4.1.2.3 Crystallization of Adgrd1<sup>NTF</sup>

Purified Adgrd1<sup>NTF</sup> was concentrated to either 12 mg/ml or 17 mg/ml and several crystallization screens were set up. For crystallization, the sitting drop vapor diffusion method was used. Crystallization screens were set up with commercially available 96-well plate screens. Screens included the Morpheus, Morpheus II, PACT premier, SG1, JCSG plus (Molecular Dimensions, Calibre Scientific) and XP screen (Jena Bioscience GmbH).

Unfortunately, almost all conditions did not show any signs of crystal growth. Only one condition from the SG1 screen showed clear microcrystal growth, as confirmed by UV imaging (**Figure 29A/B**). The reservoir buffer in this well included 0.2 M MgCl<sub>2</sub>, 0.1 M TRIS pH 8.5 and 30% w/v PEG 4K. Two other conditions in the Morpheus II and XP screen showed potential crystal growth but all of them were extremely small (**Figure 29C/D**).



**Figure 29: Crystallization of the Adgrd1<sup>NTF</sup>.** **A:** Drop from the SG1 screen imaged with normal light. Red arrow indicates position of microcrystal in the drop. **B:** UV imaging of the same drop as in A. Red arrow indicates position of microcrystal in the drop. **C:** Image of a needle that potentially could be a protein crystal from a drop in the Morpheus II screen. **D:** Image of potential microcrystals in a drop from the XP screen.

The crystallization trials were ultimately abandoned because of two reasons. First, the co-expression and purification of the Adgrd1-G protein complex started working which was the main focus of this project. Second, molecular dynamics simulations performed by Dr. Ramon Guixa Gonzalez (PSI) provided us with an idea for an alternative approach to solve a structure of the full length Adgrd1.

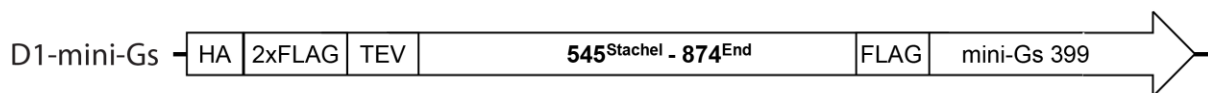
#### *4.1.2.4 Summary Adgrd1<sup>NTF</sup>*

- Expression of the Adgrd1<sup>NTF</sup> as a secreted protein in insect cells
- Purification of the Adgrd1<sup>NTF</sup> with yields ranging from 1-2 mg/L of cell culture
- Thermal stability assay showed that the T<sub>m</sub> of the Adgrd1<sup>NTF</sup> in the SEC buffer is ~60°C
- NaCl has a big influence on the thermal stability of the Adgrd1<sup>NTF</sup>
- Very few crystallization conditions showed growth of microcrystals

### 4.1.3 Adgrd1 fusion proteins

#### 4.1.3.1 Expression and purification of Adgrd1-mini-G $\alpha_s$ fusion protein

After the first purification trials of the Adgrd1 receptor, it became clear that purification of the receptor in the apo-state would not work because of the low stability of the receptor in detergents. Besides trying co-expression of Adgrd1 and the mini-G $\alpha_s$  heterotrimer, followed by complex formation on the lipid bilayer and subsequent purification of the complex, we also had an idea to design a fusion protein of Adgrd1 and mini-G $\alpha_s$ . These constructs were designed together with Dr. Ching-Ju Tsai. The construct started with a HA signal sequence, followed by a double FLAG-tag and TEV cleavage site, fused to the N-terminus of the Adgrd1<sup>CTF</sup>. The C-terminus of Adgrd1<sup>CTF</sup> was fused to mini-G $\alpha_s$  399 [120] without an additional linker since the C-terminus of Adgrd1 is relatively long with almost 50 amino acids (**Figure 30**). This construct was cloned into the pAC8RED vector for insect cell expression.



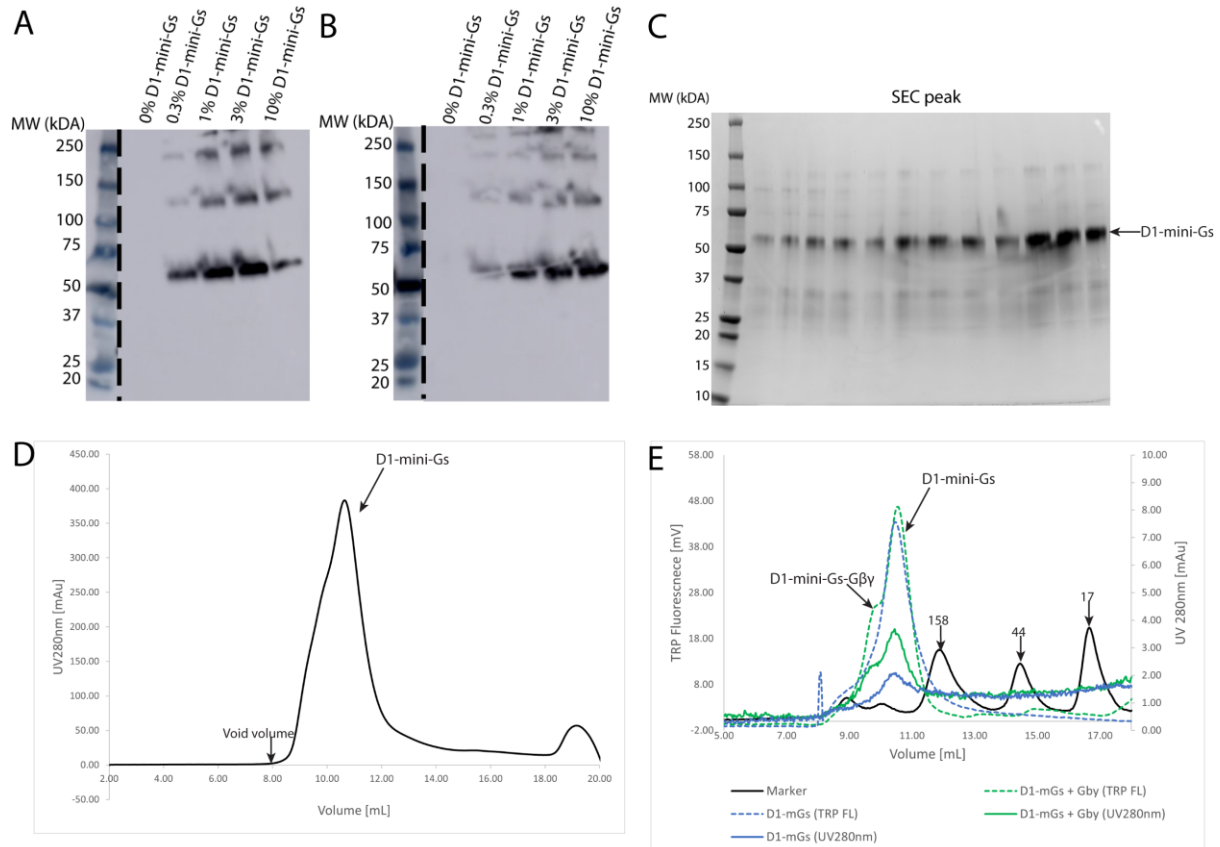
*Figure 30: Schematic representation of the D1-mini-G $\alpha_s$  fusion construct.*

Initially, expression was tested in Sf9 and High Five insect cells with different VOIs. To confirm expression, a western blot using an anti-FLAG antibody was used. The western blot analysis showed clear expression in both cell lines at similar levels (**Figure 31A/B**). Larger MW bands could also be detected that could hint towards oligomer formation.

A first large-scale expression was done in Sf9 cells. The purification protocol included addition of chemically synthesized Stachel peptide in order to activate the receptor and allow coupling to mini-G $\alpha_s$ , followed by solubilization in LMNG/CHS. Insoluble material was removed by centrifugation and the clarified supernatant was mixed with FLAG resin. After incubation the resin was washed, and the protein was eluted in three steps. The eluent was concentrated and injected into a SEC column. The purification worked well, even though there were still minor impurities after SEC, leaving room for optimization (**Figure 31C/D**). The yield was ~1.5 mg per liter of cell culture which is significantly higher than the complex purifications from the co-expression of the receptor and G protein (~200  $\mu$ g/L). This suggests that this approach presents a valuable alternative to the co-expression approach.

Next, it was tested if the Adgrd1-mini-G $\alpha_s$  fusion protein can bind the G $\beta\gamma$  subunits. Analytical SEC was used to check complex formation. The reference sample included only the Adgrd1-mini-G $\alpha_s$  fusion protein and for the test sample G $\beta_1\gamma_1$ , purified from bovine retinas, was added at an equimolar amount. The test sample did not show a clear shift of the peak but only a shoulder in the higher MW range was detected. This suggests that the complex formation is not very efficient and needs further optimization. Alternatively, a cryo-EM sample could be prepared only with the Adgrd1-mini-G $\alpha_s$  fusion protein.

As already mentioned before, during our work on the aGPCR project, four Nature papers [67-70] were published that showed the activation mechanism of aGPCRs by the Stachel peptide. For this reason, the work on this project was not further pursued because in two of the Nature papers they already solved the Adgrd1 structure.



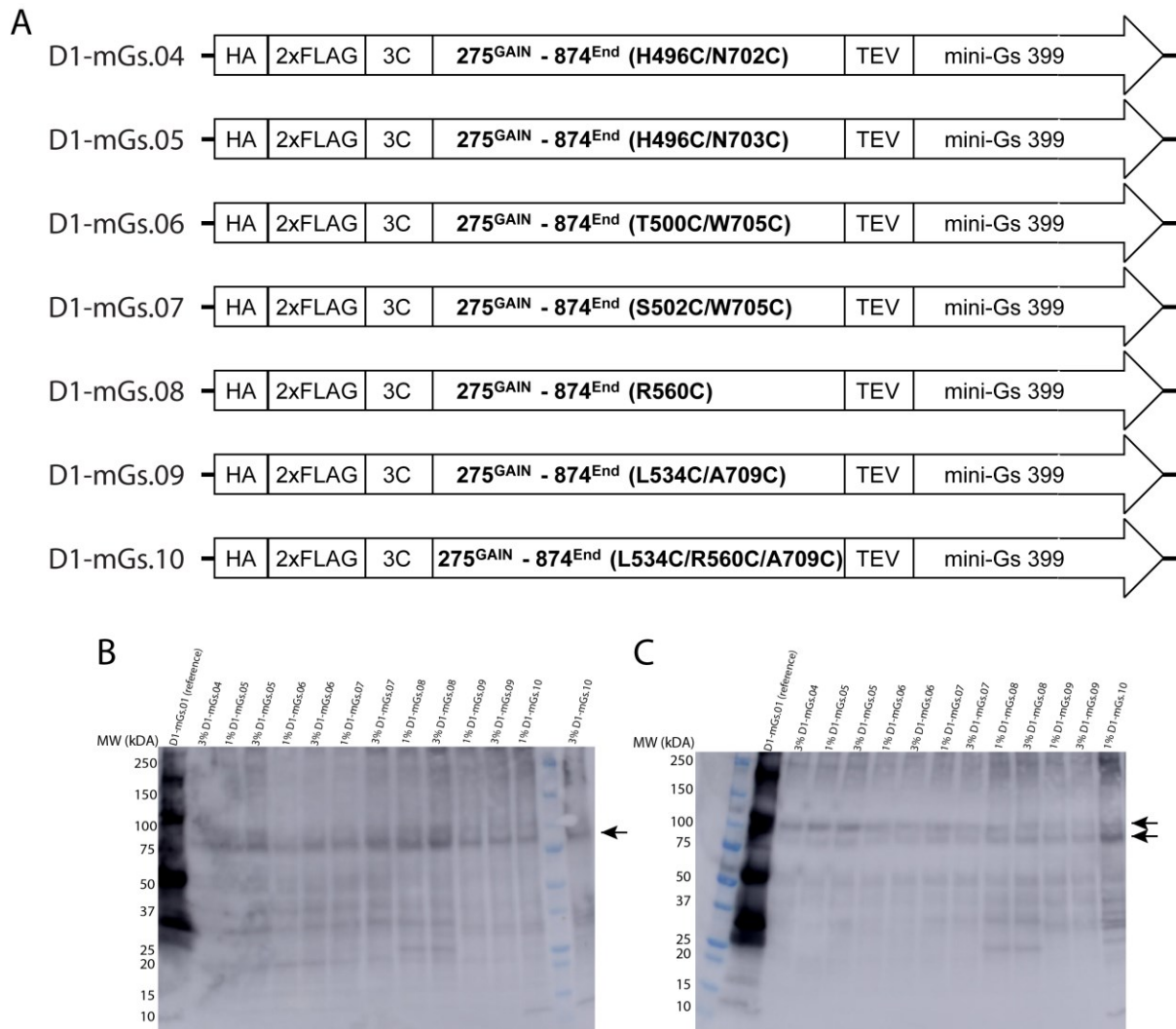
**Figure 31: Expression and purification of the Adgrd1-mini-G $\alpha_s$  fusion protein.** **A:** Western blot using an anti-FLAG antibody detecting the protein of interest from Sf9 expression. **B:** Western blot using an anti-FLAG antibody detecting the protein of interest from High Five expression. **C:** SDS-PAGE of the peak from the SEC column, stained with Coomassie blue. **D:** SEC chromatogram from the purification of the fusion protein. X-axis shows volume in mL and Y-axis shows UV 280nm signal in mAu. **E:** Analytical SEC chromatogram checking complex formation of the fusion protein with the G $\beta\gamma$  subunit. X-axis shows volume in mL, the left Y-axis show tryptophane fluorescence in mV and the right Y-axis show UV 280nm absorption in mAu. MW of the marker (black line) peaks are indicated in kDa.

#### 4.1.3.2 Expression and purification of Adgrd1<sup>GAIN-CTF</sup>-mini-G $\alpha_s$ fusion protein

Although the structure of the Adgrd1<sup>CTF</sup> in complex with a G protein has been solved, showing the activation mechanism by the Stachel peptide [68, 70], it remains unclear how the GAIN domain can interact with the 7TM domain. Barros-Álvarez *et al.* showed that the GAIN domain is highly flexible and can dissociate [67], hinting that additional modifications have to be done in order to elucidate a structure with the GAIN domain and the 7TM domain.

Our idea was to restrict the flexibility of the GAIN domain by introducing cysteine mutations in the GAIN domain and extracellular loops of the 7TM domain which could form disulfide bridges. In this work we focused on Adgrd1 and designed constructs that start from the GAIN domain until the end. A FLAG tag followed by a 3C protease cleavage site was fused at the N-terminus and the C-terminus is

extended with a glycine-serine linker, followed by a TEV protease cleavage site and a mini-Gα<sub>s</sub> 399 [120]. Dr. Ramon Guixa Gonzalez (PSI) performed molecular dynamics (MD) simulations with the Adgrd1 GAIN and 7TM domain in order to judge where cysteine mutations can be introduced. Dr. Ching-Ju Tsai additionally used Alphafold [141] to check if our construct designs give reasonable results. Together with Ramon and Ching-Ju, we designed seven constructs with either one pair (D1-mGs.04-.09) or two pairs of disulfide bridge cysteines (D1-mGs.10) (**Figure 32A**). These constructs were cloned into the pAC8RED vector for insect cell expression.



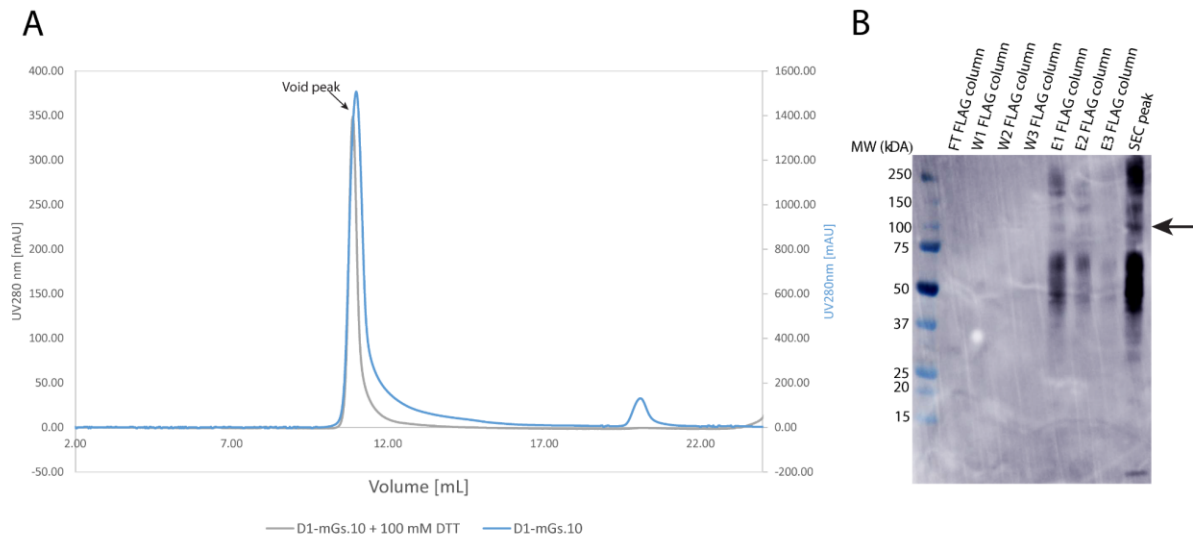
**Figure 32: Construct design and expression screening for the Adgrd1-mini-Gα<sub>s</sub> fusion constructs. A:** Schematic representation of the construct design for the Adgrd1-mini-Gα<sub>s</sub> fusion proteins. **B:** Western blot analysis using an anti-FLAG antibody probing expression in High Five insect cells. Band for the fusion protein is indicated with a black arrow. **C:** Western blot analysis using an anti-FLAG antibody probing expression in Sf9 insect cells. Bands for the fusion protein are indicated with black arrows.

Expression was tested in both High Five and Sf9 insect cells with two different baculovirus concentrations. Expression of the constructs was probed using western blot with an anti-FLAG antibody coupled to HRP (**Figure 32B/C**). The theoretical molecular weight of the constructs is ~101 kDa. Membrane proteins are very hydrophobic and are known to run faster than soluble proteins, meaning that the band on a SDS-PAGE shows a lower MW than the theoretical MW. In samples from

both cell lines a band appears between the 100 kDa and 75 kDa markers (**Figure 32B/C**). In the samples from the Sf9 cells (**Figure 32C**), another band was detected with slightly higher MW just above the 100 kDa marker, and it might correspond to a differentially modified version (e.g. glycosylation). In the High Five samples, the second larger MW band can also be detected in some of the samples but it is more faint than in the Sf9 samples (**Figure 32B**). Overall, the expression screen did not look very promising as only faint bands were detected and additionally also several lower MW bands were detected that might correspond to cleavage products.

Because most of the constructs had similar expression levels based on the intensity of the western blot bands, we decided to make a large-scale expression using High Five cells and try to purify one of the constructs (D1-mGs.10). D1-mGs.10 was purified using an anti-FLAG resin column and SEC. The SEC profile showed one single peak but unfortunately the elution volume is equivalent to the void volume, meaning that the protein is likely to form aggregates (**Figure 33A**). These aggregates could be formed from denatured proteins, or it might be that the cysteine mutations we introduced led to the formation of intermolecular disulfide bridges, resulting in the formation of oligomeric structures. In the western blot analysis of the different purification steps, we can recognize a band around the 100 kDa marker in both E1 and E2 elution fractions of the FLAG column and the SEC peak (**Figure 33B**). However, there are also strong signals in the higher and lower MW range in the western blot. The band in the higher MW range could be aggregates or oligomers. The bands in the lower MW range might be degradation or cleavage products. Because it was not clear if the proteins are aggregated or oligomerized, we collected the SEC peak fractions and analyzed them further. Part of the sample was incubated with 100 mM DTT to break potential disulfide bridges and then injected again into a SEC column. The SEC profile looked almost identical to the previous SEC profile, indicating that this protein construct is unlikely to form correct intermolecular disulfide bridges, and the protein is likely aggregated (**Figure 33A**). If the protein was correctly folded, we would also expect to see several peaks in the SEC chromatogram because the GAIN domain would dissociate upon addition of the reducing agent DTT that breaks the disulfide bridges between the GAIN and CTF domains. Since the results were not conclusive, we decided to prepare negative staining grids for analysis.

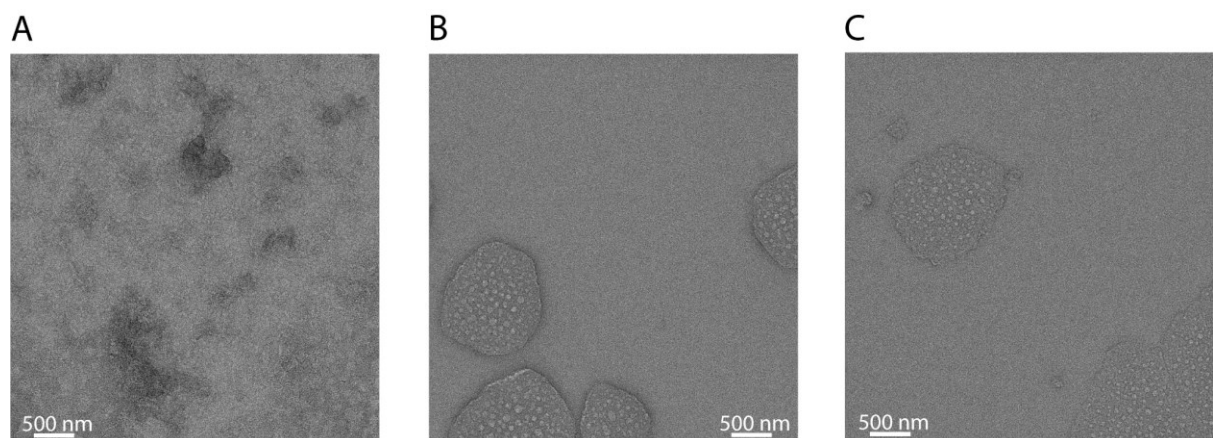




**Figure 33: Purification of D1-mGs.10.** **A:** SEC chromatogram of the purification (D1-mGs.10) and a sample from the SEC peak that was incubated with 100 mM DTT and re-injected to the SEC column (D1-mGs.10 + 100 mM DTT). **B:** Western blot analysis using an anti-FLAG antibody for different purification steps. Abbreviations: FT = flow through; W = wash step; E = elution step. Possible band for the D1-mGs.10 protein is marked with a black arrow.

#### 4.1.3.3 Negative staining EM analysis of *Adgrd1*<sup>GAIN-CTF</sup>-mini-*Gα<sub>s</sub>* fusion protein

We prepared negative staining grids and analyzed them by room-temperature EM. Together with Dr. Ching-Ju Tsai we prepared grids with two different concentrations. Once, we prepared grids with a sample taken directly from the SEC peak at a concentration of 3.3 mg/ml and the other sample was diluted to a concentration of 0.0066 mg/ml. Grids from both samples were analyzed with a JEOL JEM 2200FS microscope. The 3.3 mg/ml sample was very crowded with material and showed aggregation which is not unexpected because of the high concentration of the sample (**Figure 34A**). The sample with 0.0066 mg/ml showed less crowding but the sample also showed aggregation (**Figure 34B/C**). Together, the data implicates that our fusion protein is either not correctly folded because of the mutations or the introduced cysteines form intermolecular disulfide bridges that could not be broken by the addition of 100 mM DTT (**Figure 33A**).



**Figure 34: Negative stain grid analysis.** The three micrographs were collected with a nominal magnification of 3'000 and a K2 camera (Gatan). **A:** Example of a micrograph collected from the sample with a concentration of 3.3 mg/ml. **BC:** Examples of micrographs taken from the sample with a concentration of 0.0066 mg/ml.



It would be necessary now to screen other constructs to the stage of SEC and negative staining EM. The key is to introduce the cysteine mutations for immobilizing the GAIN domain while also maintaining correct folding. Because of time reasons this project was not further pursued as part of my PhD thesis.

#### *4.1.3.4 Summary Adgrd1 fusion proteins*

- Expression of an Adgrd1<sup>CTF</sup>-mini-G $\alpha_s$  fusion protein was successful in insect cells
- Purification of the Adgrd1<sup>CTF</sup>-mini-G $\alpha_s$  fusion protein was possible and gave a higher yield than purifying the Adgrd1<sup>CTF</sup>-mini-G $\alpha_s$  complex from co-expression.
- The Adgrd1<sup>CTF</sup>-mini-G $\alpha_s$  fusion protein did not show clear coupling to the G $\beta\gamma$  subunits.
- Expression of the Adgrd1<sup>GAIN-CTF</sup>-mini-G $\alpha_s$  fusion proteins was detectable but expression levels were low.
- Purification of the construct D1-mGs.10 was not successful, and the SEC peak appeared in the void volume, indicating aggregation or formation of oligomeric structures.
- Analysis of the negative staining grids confirmed that the protein was aggregated or formed large oligomeric structures.

## 4.2 Jumping spider rhodopsin-1 project

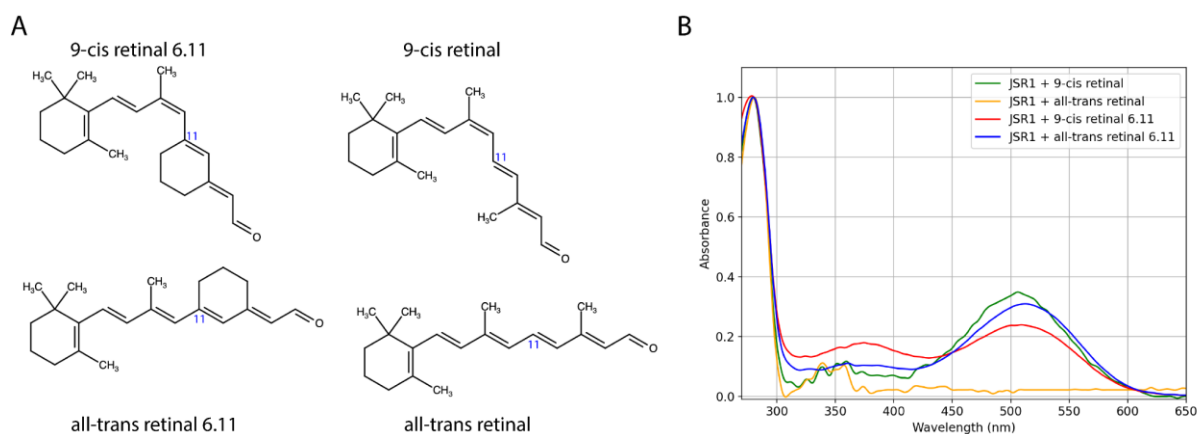
Jumping spider rhodopsin isoform 1 (JSR1) is a bistable rhodopsin found in the jumping spider [96]. Because of the ability to switch JSR1 on/off by light, this receptor is of high interest to researchers in the optogenetics field. It was recently shown that JSR1 can be used as an optogenetic tool in zebrafish [92]. However, JSR1 has an issue that the wavelength for activation/inactivation are overlapping, meaning that illumination will lead to a mixed population of active and inactive receptors [100]. Moreover, JSR1 has been shown to signal through both the human Gq (hGq) and human Gi (hGi) protein subtypes [30, 84, 98](unpublished cell signaling data for hGi from Deborah Walter (Schertler lab)) and ideally an optogenetic tool would signal through only one G protein subtype. Engineering of a JSR1 mutant, where the  $\lambda_{\max}$  of the active and inactive states are not overlapping and only one G protein pathway is activated, would allow precise spatial and temporal control of the receptor activity and G protein signaling. This would result in a valuable optogenetic tool. Rational engineering of JSR1 can be aided by mechanistic and structural insights. As mentioned before, our group already established an expression and purification protocol for JSR1 bound to the inverse agonist 9-cis retinal [30]. In addition, the crystal structure of the inactive state JSR1 (9-cis) [30] provided us with important insights but crucial aspects of the active state JSR1 were still missing. Dr. Filip Pamula, a former PhD student in our lab, worked towards the structure determination of a JSR1-human Gi complex (JSR1-hGi). He succeeded in reconstructing a cryo-EM map and determining a structure to  $\sim 5$  Å resolution. Although the resolution was not good enough to see side chain densities, a density for the all-trans retinal was visible and the structure showed interesting features in the JSR1-hGi binding interface (discussed in “4.2.3.4 JSR1-G protein interactions”).

In my work, the focus was on improving sample and grid preparation of a JSR1-G protein complex to determine a higher resolution cryo-EM map. This would allow us to gain insights into how JSR1 achieves bistability and how JSR1 interacts with the G protein.

### 4.2.1 Biochemical characterization of JSR1-G protein complexes

#### 4.2.1.1 JSR1 purification

An expression and purification protocol for JSR1 was established by Varma *et al.* [30]. Dr. Elena Lesca expressed JSR1 in HEK293GnTl<sup>-</sup> cells and the cell pellet was frozen at  $-80^{\circ}\text{C}$ . Varma *et al.* purified JSR1 bound to the inverse agonist 9-cis retinal [30]. Shortly before I started working on this project, we received retinal analogues from Prof. Mordechai Sheves (Weizmann Institute of Science). We received analogues that, in theory, are either locked in the all-trans configuration (all-trans retinal 6.11) or a 9-cis retinal 6.11 that should be locked in the all-trans configuration after light activation (**Figure 35A**). The name 6.11 comes from the carbon ring around the double bond at position 11 that should not allow isomerization to the cis configuration at this double bond.



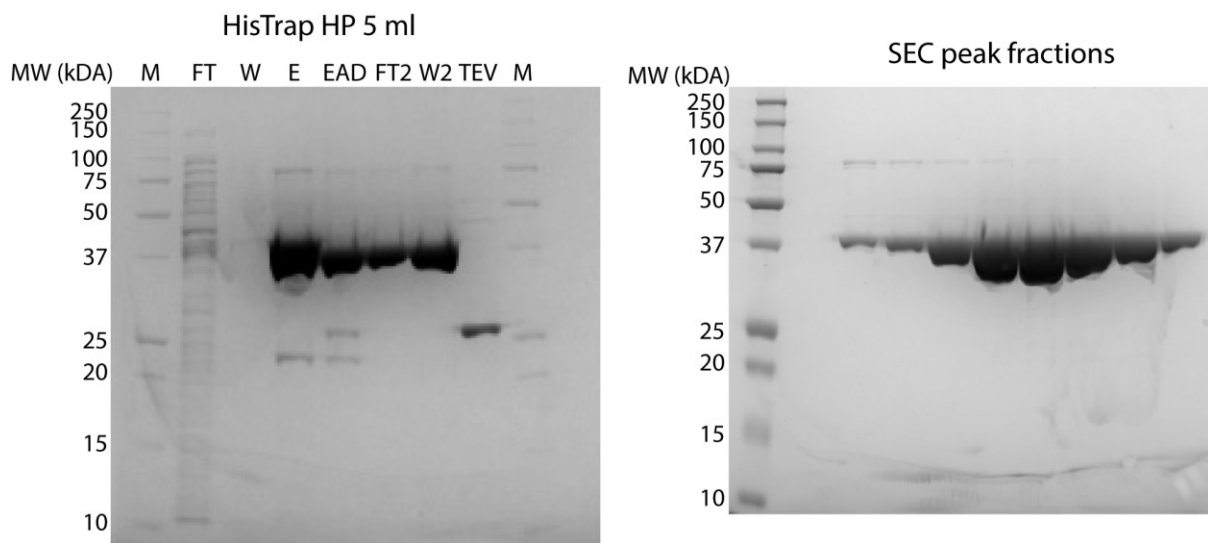
**Figure 35: Retinal variants and their UV-Vis spectra when incubated with JSR1.** **A:** Chemical structures of 9-cis retinal, all-trans retinal and the analogues 9-cis retinal 6.11 and all-trans retinal 6.11. **B:** UV-Vis spectra for JSR1 incubated with either of the four retinals, solubilized and 1D4 affinity purified. Spectra provided by Dr. Matthew Rodrigues.

As already mentioned before, the  $\lambda_{\max}$  value of the inactive and active states of JSR1 overlap, meaning that light activation will lead to a mixed population of states [30, 100]. Using the inverse agonist 9-cis retinal instead of the natural 11-cis retinal splits the  $\lambda_{\max}$  values slightly. The  $\lambda_{\max}$  for JSR1 bound to 11-cis retinal is 535 nm while JSR1 bound to 9-cis retinal is blue shifted with a  $\lambda_{\max}$  of 505 nm [100]. Photoisomerization of either isoform yields the all-trans retinal with a  $\lambda_{\max}$  of 535 nm [100]. Because of this overlap and the technical difficulties of preparing a cryo-EM sample and grids under dim-light conditions, we aimed to purify JSR1 bound to the all-trans retinal 6.11 (ATR611). An additional reason to use this analogue was that previous attempts to purify active JSR1 by adding the all-trans retinal were not successful, as confirmed by the UV-Vis spectra (**Figure 35B**). The presence of the carbon ring around the double bond at position 11 enabled binding of the ATR611 to JSR1 and subsequent purification, shown by Dr. Matthew Rodrigues from the Schertler group (**Figure 35B**). We found that ATR611 needs to be added before solubilization because incorporation of ATR611 to JSR1 did not work in detergent environment. Purification of JSR1 bound to the 9-cis retinal 6.11 was also possible (**Figure 35B**).

For the *in vitro* activity assays and preparation of a cryo-EM sample, I either purified JSR1 (ATR611 bound) in DDM or in LMNG. The solubilization step was done with DDM in both cases and for the LMNG sample a detergent exchange to LMNG was performed during the 1D4 affinity purification. Normally, JSR1 was incubated with 50  $\mu\text{M}$  of ATR611 but because our stock ran low, the incubation was once done with 25  $\mu\text{M}$  and once with 12.5  $\mu\text{M}$  ATR611 retinal. Lowering the retinal concentration resulted in lower yields. The yield from the purification where 50  $\mu\text{M}$  retinal was used, was approximately  $\sim 750 \mu\text{g}$  from 10 grams of cell pellet while the yield for the purifications with 25  $\mu\text{M}$  or 12.5  $\mu\text{M}$  were 650  $\mu\text{g}$  and 530  $\mu\text{g}$  respectively. Nevertheless, these quantities of pure protein are enough for the *in vitro* activity assays and preparation of a JSR1-G protein complex for cryo-EM studies.

#### 4.2.1.2 Human Gi heterotrimer purification

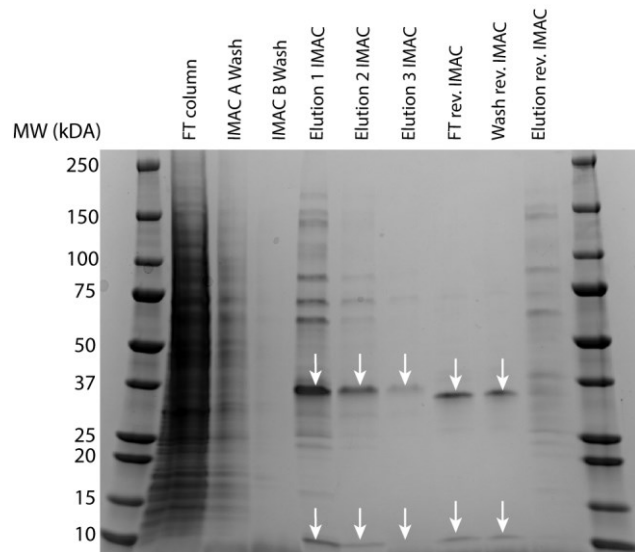
For the human Gi (hGi) heterotrimer, the G $\alpha$ i and G $\beta\gamma$  subunits were purified separately [117]. The hG $\alpha$ i subunit was expressed in BL21 (DE3) cells and cells were harvested by Dr. Ching-Ju Tsai. She also developed a purification protocol already. The protein was His-tagged and the purification involved a HisTrap column followed by dialysis with TEV protease cleavage to cleave of the His-tag and another HisTrap column before a SEC step was done, resulting in pure hG $\alpha$ i subunit with a yield of ~15 mg/L (**Figure 36**). The purification worked without any problems and pure hG $\alpha$ i subunit was detected on SDS-PAGE (**Figure 36**).



**Figure 36: Purification of G $\alpha$ i.** In both SDS-PAGE gels the intense band represents the G $\alpha$ i subunit. Left shows Coomassie stained SDS-PAGE from the HisTrap column. Abbreviations: M=Marker; FT=Flow-through; W= wash; E=Elution; EAD: Eluent after dialysis; FT2=Flow-through reverse HisTrap column; W2= Wash reverse HisTrap column; TEV= TEV protease. Right show Coomassie stained SDS-PAGE from the peak fractions of the SEC step.

The G $\beta_1\gamma_1$  subunits were purified from the transducin G protein heterotrimer which was purified from bovine retinas as described by Maeda *et al.* [122] (see Methods for details). First, bovine rhodopsin in the rod outer segment (ROS) membranes was light activated that it forms a complex with the Gt heterotrimer and the ROS membranes were isolated from bovine retinas. Addition of GTP leads to the dissociation of the Gt heterotrimer from bovine rhodopsin and the trimer was collected. To separate the G $\alpha$ t from G $\beta\gamma$  a blue Sepharose was used and pure G $\beta\gamma$  was eluted from the column. The yield was ~8.7 mg of protein from 450 retinas.

As an alternative strategy to the purification of G $\beta\gamma$  from bovine retinas, I once expressed the G $\beta_{1\gamma_2}$  subunits in insect cells and purified it. In short, High Five cells that expressed the protein were mechanically disrupted, the membranes harvested and G $\beta\gamma$  was extracted from the membranes using detergent. Next, an IMAC step was done with TALON resin. The detergent-solubilized fraction was mixed with TALON resin, allowing the his-tagged G $\beta\gamma$  to bind to the resin. The protein was eluted with Imidazole at high concentration and the his-tag was cleaved by the 3C protease during dialysis. The dialysate was again mixed with TALON resin for the reverse IMAC step. The flow-through and wash containing His-tag-free G $\beta\gamma$  were collected and concentrated. After the two affinity column steps the protein was



**Figure 37: G $\beta\gamma$  purification from insect cells.** FT=Flow-through. Samples were collected from different steps of the purification. The elution samples of the IMAC and the FT and Wash samples from the reverse IMAC show clear band for G $\beta$  (~38 kDa) and G $\gamma$  (~8 kDa).

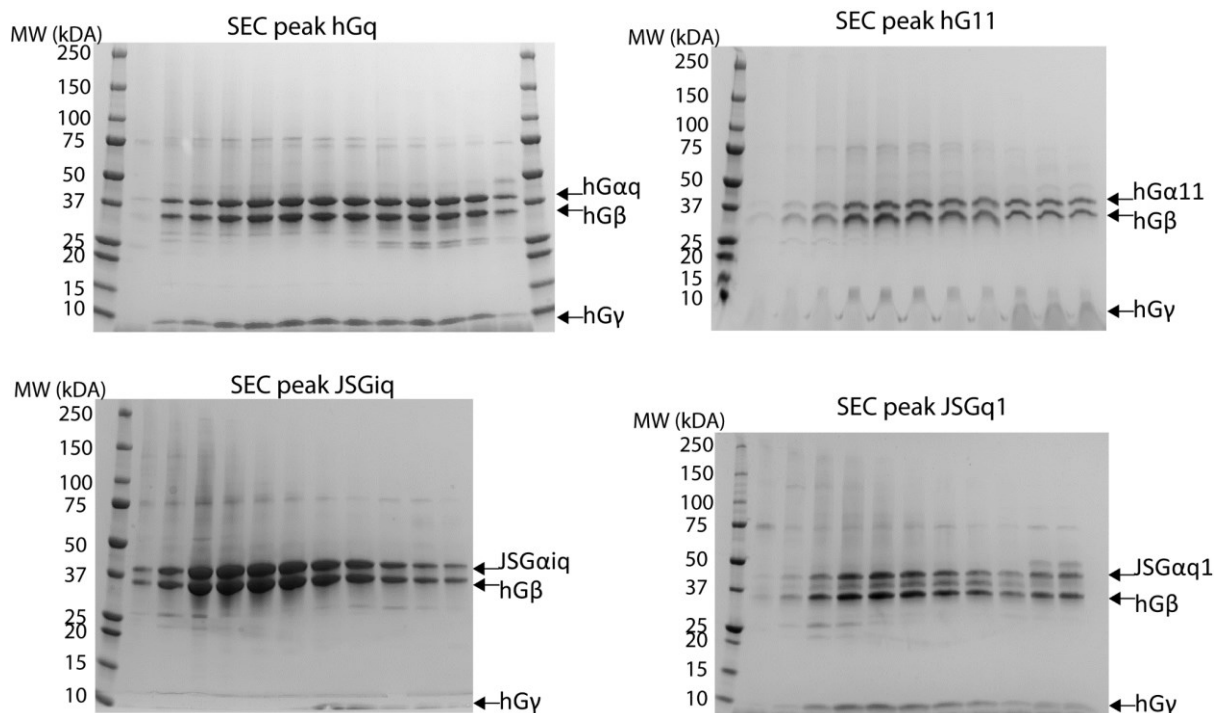
already pure and it was flash frozen until further use (**Figure 37**). The purification worked well, although the yield was not high with ~100  $\mu\text{g/g}$  of cell pellet. This could also be because the cell culture was contaminated which likely lowered the expression efficiency.

HG $\alpha_i$  and G $\beta_{1\gamma_1}$  subunits were then mixed at an equimolar ratio and incubated for at least 30 min to form the G $i$  heterotrimer. This sample was then used for the *in vitro* activity assays and complex formation experiments with JSR1.

#### 4.2.1.3 hGq/hG11/JSGiq/JSGq1 heterotrimer purification

The human Gq (hGq), human G11 (hG11), jumping spider Gq1 (JSGq1) and the human/jumping spider Giq chimera (JSGiq) heterotrimers were all expressed in insect cells. Purification details are described in the Methods section. In short, the cells were harvested by centrifugation, mechanically disrupted and the G protein heterotrimer was extracted from the cell membranes with detergent. The G protein was bound to TALON resin via the his-tag on the G $\beta$  subunit. The G protein was eluted with high Imidazole concentration and the his-tag was cleaved by the 3C protease during dialysis. Reverse IMAC purification was done with fresh TALON resin to get rid of the his-tagged 3C protease and other protein contaminants. The sample was then loaded onto a HiTrap Q Sepharose column and the G protein was eluted with high NaCl concentration. The eluent was then concentrated and injected into a SEC column. After SDS-PAGE analysis, fractions containing pure G protein were collected, concentrated and flash frozen in liquid nitrogen until further use.

The purification worked well for all different G proteins but the yield varied. The highest yield was achieved with the JSGi<sub>q</sub> (~2.5 mg/L) followed by the hG<sub>q</sub> and hG<sub>11</sub> (~750 µg/L) and the JSG<sub>q1</sub> (~500 µg/L). From previous experiences in our lab, we know that the hGi protein is well expressed and stable whereas the hG<sub>11</sub> and hG<sub>q</sub> are less stable which is reflected in the yield. In our lab we always worked with human G proteins or with the bovine transducin G protein but never with invertebrate G proteins. The JSG<sub>q1</sub> might need different buffer conditions which could stabilize the protein and give a higher yield. Interestingly, a clear band on SDS-PAGE was detected between the JSG $\alpha$ <sub>q1</sub> and G $\beta$  subunits which might be a differentially modified version of either subunit (**Figure 38**). This band was only seen for the JSG<sub>q1</sub> but not for the other G proteins (**Figure 38**).



**Figure 38: G protein purifications.** Each Coomassie stained SDS-PAGE shows samples from the SEC peak of the corresponding G protein. Subunits are labelled on the right side of the gel. The first lane in each gel is the molecular marker and the bands are labelled with the molecular weight in kDa. The gel for the hG<sub>11</sub> was run with the wrong buffer resulting in wavy bands.

Nevertheless, enough protein of each G protein subtype was purified to conduct *in vitro* activity assays and complex formation experiments with JSR1.

#### 4.2.1.4 GTPase Glo assay JSR1-G proteins

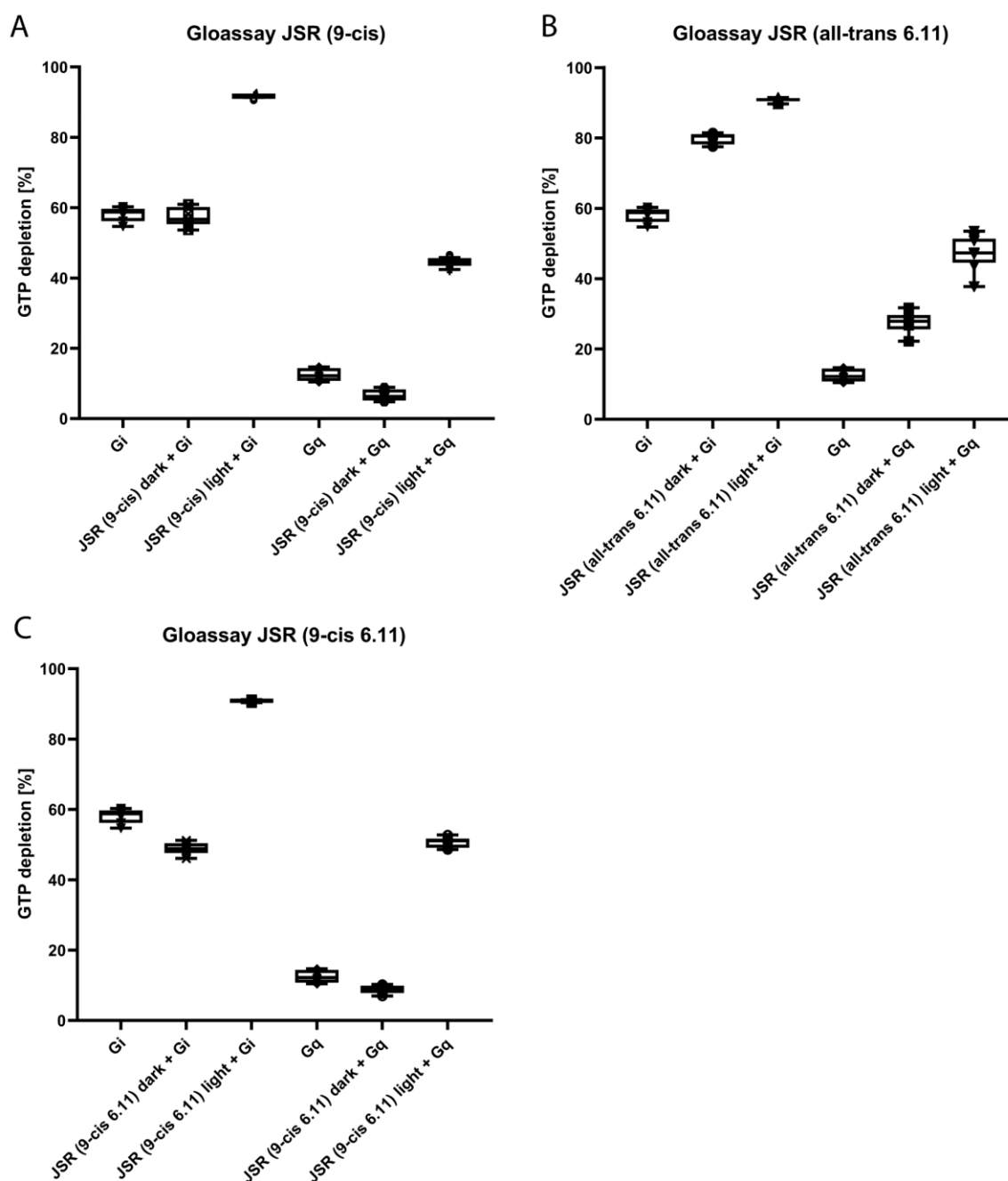
After we established a purification protocol for JSR1 bound to either of the two analogues (ATR611 and 9-cis retinal 6.11), we sought out to determine their effect on JSR1 activity. JSR1 has been shown to couple and activate hGi and hG<sub>q</sub> [30, 98] when light activated. The G $\alpha$  subunit of the heterotrimeric G protein is a GTPase. GTPase activity is based on two processes: 1. The exchange of GDP with GTP; 2. The hydrolysis of GTP to GDP. An active GPCR catalyzes the exchange of GDP for GTP, acting as a guanine-nucleotide exchange factor (GEF). In the GTPase Glo assay we measure the GTPase activity in

different samples. The remaining amount of GTP, after incubation with JSR1 and either G protein, is converted to ATP by the GTPase Glo reagent and adding a luciferase allows measurement of a light signal. This means that a low light signal corresponds to high GTPase activity and a strong light signal corresponds to low GTPase activity. To make it more intuitive we calculate the GTP depletion in percentage compared to the “GTP only” reference sample (see details in Methods section).

Other than the “GTP only” reference sample, we also include “JSR1 only” and “G protein only” reference samples. Especially, the G protein sample is important because the G proteins have basal GTPase activity without any GPCR present. In this assay we tested JSR1 (9-cis retinal), JSR1 (9-cis retinal 6.11) and JSR1 (ATR6.11) in combination with either hGi or hGq (**Figure 39**). JSR1 bound to either the 9-cis or 9-cis retinal 6.11 was purified by Dr. Matthew Rodrigues from our group.

The reference samples that included only JSR1 bound to either retinal did not show any GTPase activity as expected. In the absence of any receptor, we observed that hGi has a high basal activity with approximately 60% GTP depletion which is markedly higher than the basal activity of the hGq with 15% GTP depletion (**Figure 39**). Addition of JSR1 bound to either the 9-cis or 9-cis retinal 6.11 (dark samples) did not result in an increase in GTP depletion or even a small decrease compared to the G proteins only samples. This is expected because 9-cis retinal is an inverse agonist and stabilizes the inactive state of JSR1 [30].

Illumination of 9-cis retinal bound JSR1 results in a dynamic equilibrium of JSR states because of the overlapping  $\lambda_{\max}$  of the 9-cis, 11-cis and all-trans retinal bound JSR1 [100]. Up to 73% of JSR1 molecules adopt the active state after illumination [100]. Indeed, for both the 9-cis and 9-cis retinal 6.11 bound JSR1 we see an increase in GTP depletion in combination with the hGi and hGq heterotrimers after illumination (light samples) (**Figure 39A/C**). In combination with hGi the GTP depletion increases from ~57% to ~92% for the 9-cis bound JSR1 and from ~49% to ~91% for the 9-cis retinal 6.11 bound JSR1 (**Figure 39A/C**). Similarly, we detected an increase in GTP depletion in combination with hGq from ~7% to ~44% for the 9-cis bound JSR1 and from ~9% to ~50% for the 9-cis retinal 6.11 bound JSR1 (**Figure 39A/C**). This clearly shows light-induced receptor activation for both the 9-cis bound JSR1 and the 9-cis 6.11 bound JSR1. The 9-cis retinal 6.11 likely isomerizes to form the all-trans retinal 6.11 conformation but this needs further confirmation by future experiments.

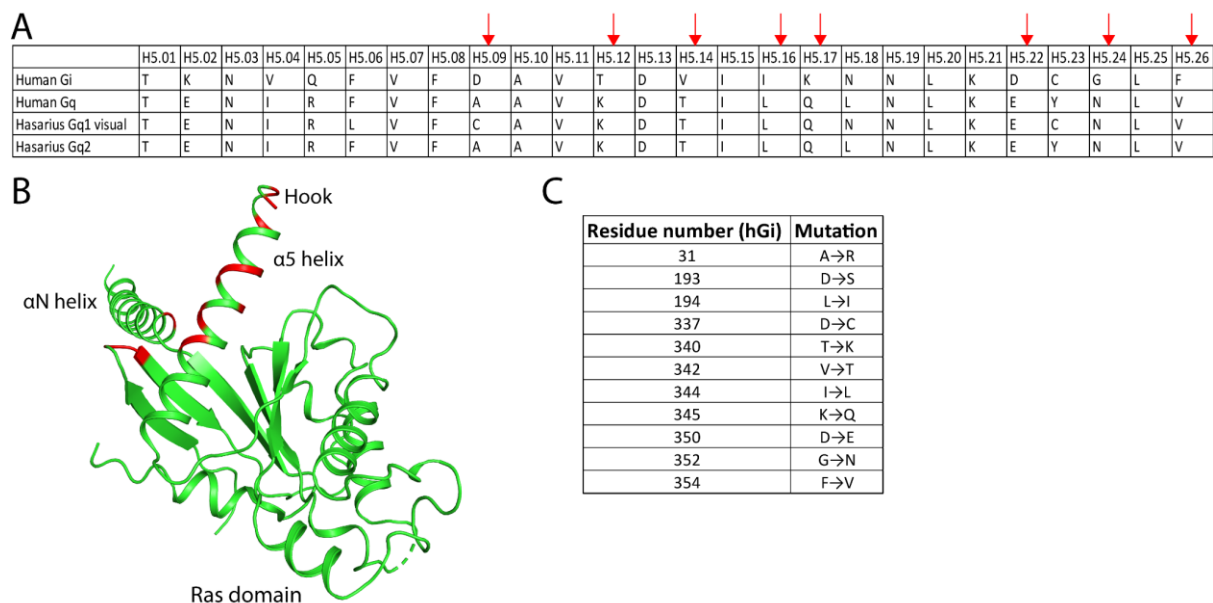


**Figure 39:** GTPase Glo assay with JSR1 (9-cis; 9-cis 6.11; all-trans 6.11) in combination with hGi and hGq. Plots show the GTP depletion in percentage on the Y-Axis by the hGi and hGq proteins in the presence of (A) JSR1 (9-cis retinal), (B) JSR1 (all-trans 6.11 retinal) and (C) JSR1 (9-cis 6.11 retinal).

In contrast to the dark samples with 9-cis or 9-cis retinal 6.11 bound JSR1, the all-trans retinal 6.11 bound JSR1 catalyzes the nucleotide exchange without any illumination (**Figure 39B**). GTP depletion is increased from ~58% to ~80% with hGi and from ~13% to ~28% with Gq (**Figure 39B**). However, illumination of the all-trans retinal 6.11 bound JSR1 leads to a further increase in GTP depletion to ~91% and ~47% for hGi and hGq respectively (**Figure 39B**). While the C<sub>11</sub>=C<sub>12</sub> double bond can not isomerize, illumination might induce conformational changes in the chromophore and its binding pocket which increase the agonistic activity for JSR1.



During the time I worked on this project, we received the sequence of the jumping spider visual Gq1 (JSGq1) from Prof. Akihisa Terakita (Osaka Metropolitan University). Making a complex with JSR1 and the JSGq1 might result in a more stable complex than JSR1 with the hGi as we have done it before. Because we know from experiences in our lab that the hGq protein is not very stable and difficult to work with, we decided to create a chimeric G protein. We used the hGai as the template and mutated residues in the C-terminal  $\alpha 5$ -helix and at contact points between JSR1 and the hGi (based on the cryo-EM map from Dr. Filip Pamula) to match the sequence of the JSGq1 (**Figure 40**). The JSGq1 and the chimeric JSGiq were expressed and purified. To make sure they are functional, I performed the GTPase Glo assay in combination with JSR1 bound to the all-trans retinal 6.11.

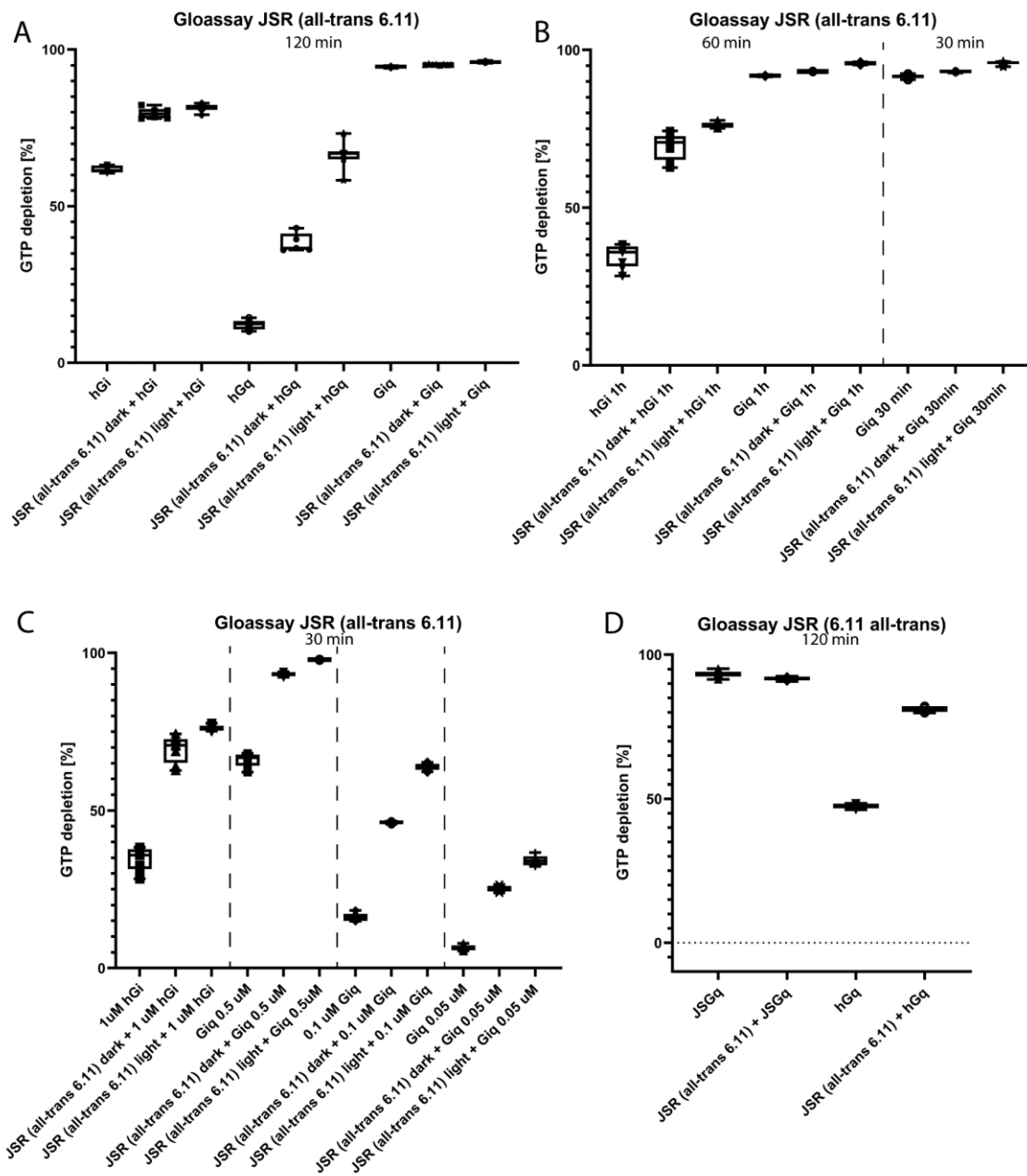


**Figure 40: Sequence comparison and mutations introduced for the JSGiq chimera.** **A:** Table comparing the  $\alpha 5$  helix sequences of the human Gai, human Gq, jumping spider Gq1 visual and jumping spider Gq2. Red arrows indicate where mutations were introduced in hGai. **B:** Human Gai structure from our JSR1-JSGiq\_1 complex. Locations where mutations were introduced are indicated in red color. **C:** Table with the mutations introduced to make the JSGiq chimera.

The first assay was done with the same conditions that were used for the assays with the hGi and hGq. Surprisingly, the JSGq1 and the JSGiq showed a drastically increased basal GTPase activity compared to the hGi and hGq (**Figure 41A/D**). In both cases the G protein reference sample showed a GTP depletion of over 90%, meaning that the samples including JSR1 (ATR611) did not show a significant difference because almost all the GTP was already depleted in the control samples. Moreover, the JSGq1 did not show clear coupling to JSR1 in analytical SEC experiments whereas the JSGiq clearly showed complex formation (see next section). For this reason, the conditions for the GTPase Glo assay were only optimized for the JSGiq.

In the next assay the incubation time was reduced from 120 min to 60 min or 30 min. Remarkably, even with the reduced reaction times the JSGiq reference sample showed a GTP depletion of ~92% for both incubation times, meaning that almost all GTP is hydrolyzed by the JSGiq in the first 30 min

(Figure 41B). The samples in combination with JSR1 (ATR611) showed a small increase in GTP depletion of ~93% and ~96% for the dark and illuminated samples respectively (Figure 41B).



**Figure 41: GTPase Glo assay with JSR1 (all-trans 6.11) in combination with hGi, hGq, JSGq1 and JSGiq chimera.** Plots show the GTP depletion in percentage on the Y-Axis by the hGi, hGq, JSGq1 or JSGiq proteins. **A:** Plot showing GTP depletion by hGi, hGq or JSGiq with/without JSR1. Incubation time was 120 min and the G protein concentration 1  $\mu$ M. **B:** Plot showing GTP depletion by hGi or JSGiq with/without JSR1 and different incubation times of 60 min or 30 min. G protein concentration was 1  $\mu$ M. **C:** Plot showing GTP depletion by hGi or JSGiq with/without JSR1. hGi had a concentration of 1  $\mu$ M and JSGiq either 0.5  $\mu$ M, 0.1  $\mu$ M or 0.05  $\mu$ M with an incubation time of 30 min. **D:** Plot showing GTP depletion by hGq or JSGq1 with/without JSR1. Incubation time was 120 min and the G protein concentration 1  $\mu$ M.

Because the JSGiq basically hydrolyzed the available GTP within 30 min, in the next assay a dilution series was done for the JSGiq. Instead of the usual 1  $\mu$ M concentration, 0.5  $\mu$ M, 0.1  $\mu$ M and 0.05  $\mu$ M JSGiq was used for the assay. The incubation time was kept at 30 min and a reference sample with 1

$\mu\text{M}$  of the hGi was done. The dilution series successfully increased the assay window, clearly showing that JSR1 (ATR611) catalyzes the exchange of GDP with GTP leading to a higher GTP depletion in combination with the JSGiq (**Figure 41C**). Compared to the hGi, the JSGiq shows higher basal activity even at half of the concentration of hGi but drops below the basal activity of hGi when taking ten times less JSGiq.

Together, these assays resulted in three major discoveries: 1. The all-trans retinal 6.11 analogue is an agonist for JSR1 in the dark and illumination results in an even more active JSR1; 2. JSR1 can activate hGi, hGq and the chimeric JSGiq; 3. The mutations introduced for creating the JSGiq chimera lead to a markedly increased basal activity compared to the hGi.

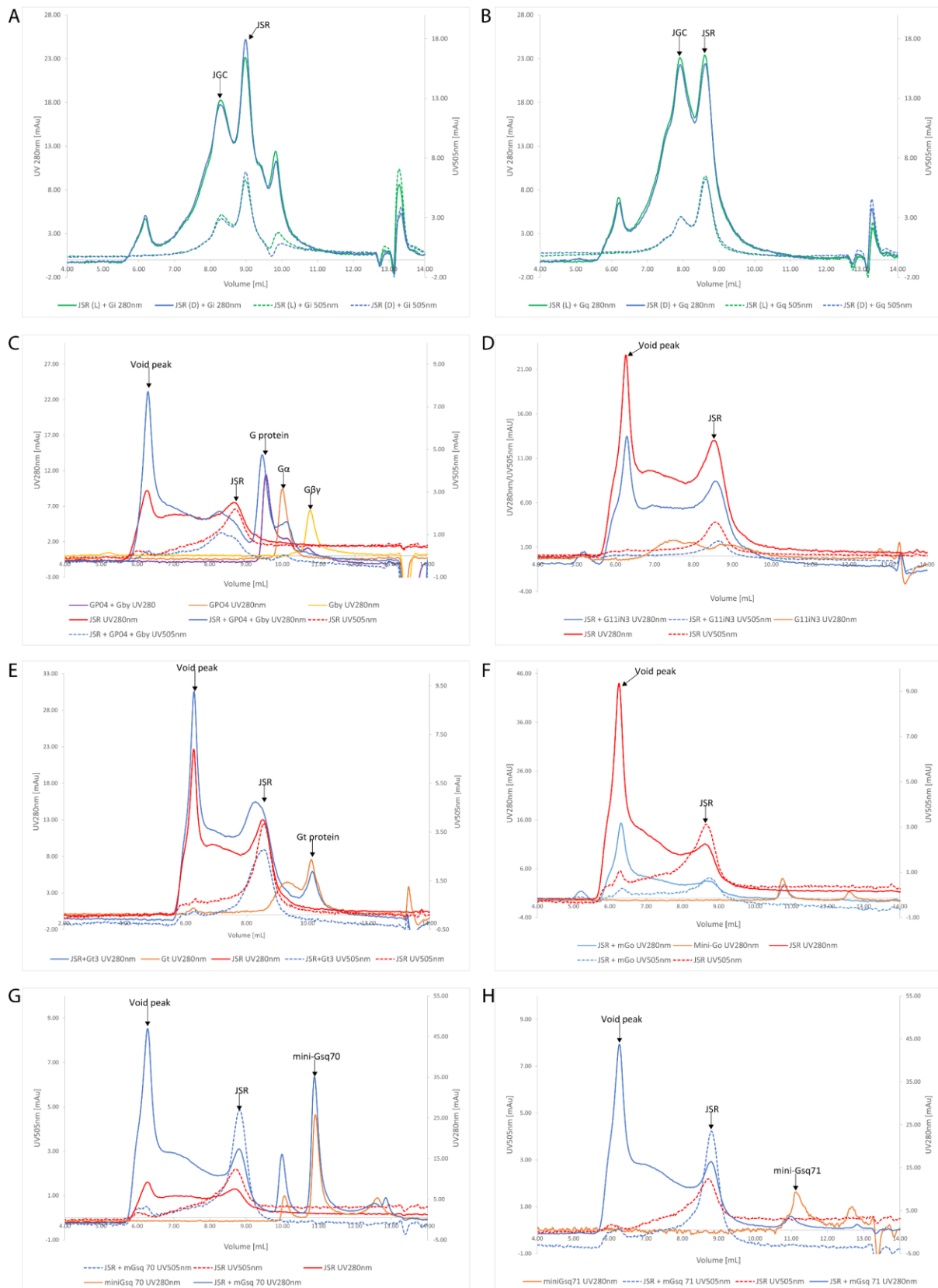
#### *4.2.1.5 Analytical SEC JSR1-G protein complexes*

The GTPase Glo assay showed that JSR1 can activate hGi, hGq and the JSGiq chimera (**Figure 41**). Activation of a G protein does not necessarily mean that a complex with JSR1 is stable and suitable for preparation of a cryo-EM sample. Before we had the JSGq1 sequence, we tested many different human G proteins from the Gi/o and Gq/11 family for complex formation with JSR1 (ATR611). From previous data by Dr. Filip Pamula we knew that complex formation with the hGi is possible. In his case JSR1 was reconstituted with 9-cis retinal, followed by illumination with 495 nm long-pass filtered light to activate JSR1 and induce G protein binding. ATR611 can be directly used to reconstitute JSR1 into the active state without any illumination.

First, I tested if JSR1 (ATR611) also forms a complex with the hGi heterotrimer. JSR1 (ATR611) was incubated with the hGi heterotrimer in the presence of apyrase for 2h and then injected into a SEC column for analysis. Apyrase is added to degrade and GDP or GTP present in the sample. JSR1 (ATR611) showed clear complex formation with the hGi, as expected, although there was still free JSR1 and G protein, hinting that the incubation time may need to be increased or the affinity of the hGi to JSR1 (ATR611) is relatively low (**Figure 42A**). Cryo-EM samples of the JSR1 (ATR611)-hGi complex were then prepared (see next section).

The JSR1 (ATR611)-hGi complex was not very stable as we have seen a lot of dissociated particles in the cryo-EM data, even with the addition of the scFv16 that can help stabilizing GPCR-G protein complexes [119] (See section “JSR-hGi complex cryo-EM data analysis”). For this reason, we tested other G proteins for complex formation with JSR1.

JSR1 has been shown to activate hGq in cell assays [98] and our GTPase Glo assay data confirmed this. Naturally, the hGq was tested for complex formation with JSR1. On the SEC profile we could detect a clear peak for the JSR1-hGq complex similar to the JSR1-hGi complex samples (**Figure 42B**). Samples



**Figure 42: Analytical SEC experiments testing complex formation of JSR1 (ATR611) with different G proteins.** Abbreviations: JGC=JSR1-G protein complex; JSR=JSR1 (ATR611). UV280nm absorption is shown as solid lines and absorption at 505 nm is shown as dashed lines. JSR1 samples in DDM are labelled with (D) and JSR1 samples in LMNG are labelled with (L) in A and B. SEC chromatograms are shown for complex formation of JSR1 (ATR611) with hGi (A), hGq (B), GP04 (C), G11iN (D), Gt (E), mini-Go (F), miniGsq70 (G) and mini-Gsq71 (H). In (C)-(H), JSR1 reference samples are shown in red, G protein reference samples in orange and complex samples in blue.

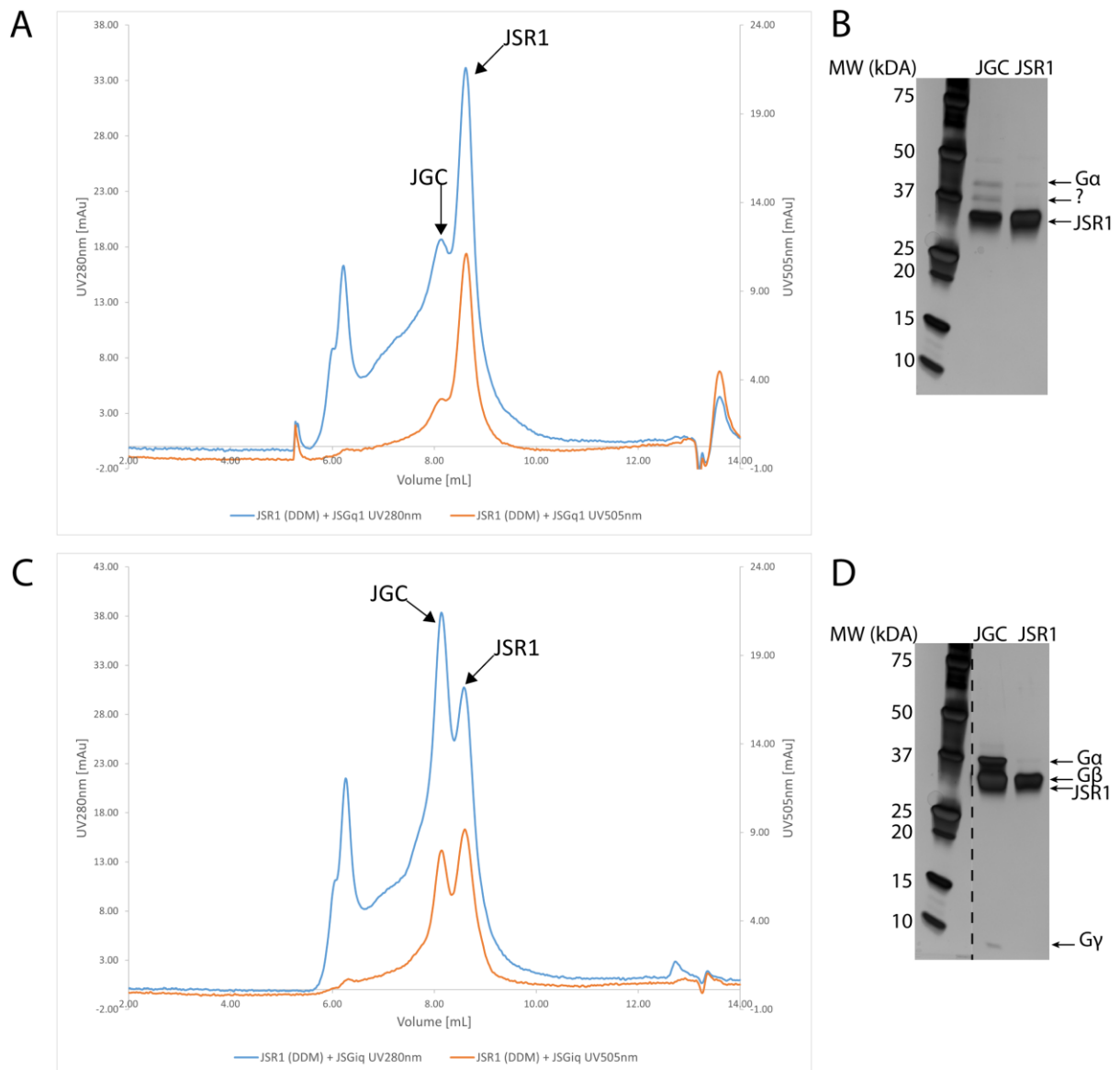
in DDM or in LMNG showed nearly identical profiles (**Figure 42A/B**).

In addition to hGi and hGq, other G proteins from these two subtypes were tested for complex formation, namely hG11 (Gq/11 subtype; **Figure 42D**), bovine Gt (bGt) (Gi/o/t subtype; **Figure 42E**), mini-Go (Gi/o/t subtype; **Figure 42F**), two mini-Gsq variants (**Figure 42G/H**) and a hGiq chimera developed by Dr. Agnieszka Olechwier (PSI) called GP04 (**Figure 42C**). The bGt heterotrimer and the mini-Go proteins were expressed and purified by Dr. Ching-Ju Tsai. GP04 was expressed and purified by Dr. Agnieszka Olechwier. The mini-Gsq proteins were expressed and purified by Dr. Ching-Ju Tsai and Dr. Filip Pamula.

Unfortunately, the JSR1 sample used for these experiments showed large aggregates (**Figure 42C/D/E/F/G/H**, void peak) making the analysis more difficult. Nevertheless, it could clearly be seen that none of these G proteins showed clear complex formation as it was seen with the hGi and hGq (**Figure 42**). Both mini-Gsq, hG11 and mini-Go proteins showed no signs of complex formation because no shift in the peak could be detected (**Figure 42D/F/G/H**). For the GP04 a small shift in the peak was detected in both the UV absorption curves at 280 nm and 505 nm (**Figure 42C**), indicating that a complex can be formed. But there was still a lot of free G protein and JSR1 present, meaning that JSR1 likely does not form a high affinity complex with GP04. The JSR1-bGt complex sample also showed a slight shift of the absorption peak for the 280 nm signal but not as clearly for the 505 nm signal (**Figure 42E**). There might be some complex formation but again it is not as clear as for the hGi and hGq (**Figure 42A/B/E**). Nevertheless, a JSR1-Gt sample was prepared for cryo-EM analysis (see next section).

Overall, making a complex with either the hGi or hGq seemed to be the most promising for structure determination. During my time working on this project, we received the jumping spider visual Gq1 sequence from Prof. Akihisa Terakita. A sequence alignment of the  $\alpha$ 5-helix of the hGi, hGq and JSGq1 showed that there are several differences in the amino acid sequence (**Figure 40A**). As mentioned before, we designed constructs for expression of the full length JSGq1 and a chimeric G protein with the human Gi as template (JSGiq). After purification, these two proteins were tested for activity with the GTPase Glo assay (**Figure 41**) and subsequently for complex formation with activated JSR1 using analytical SEC experiments.

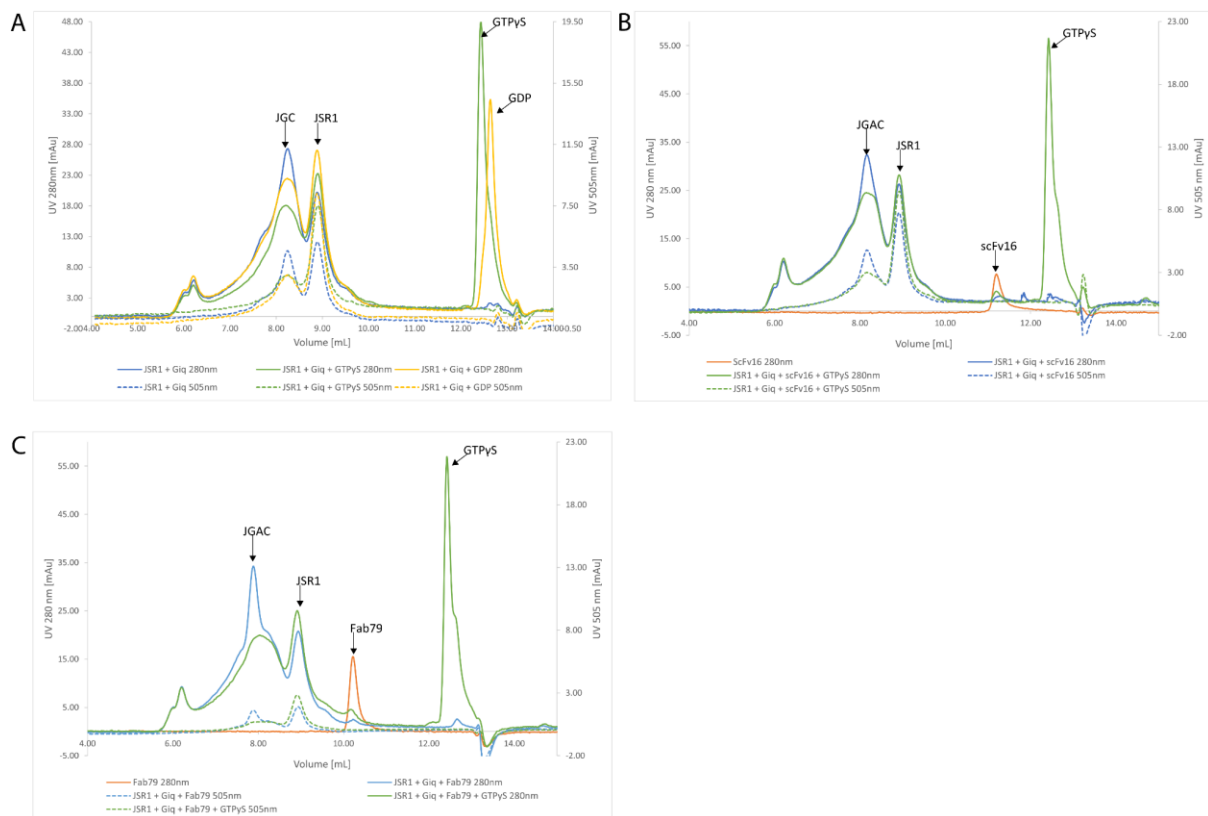
Surprisingly, complex formation with the JSGq1 could not clearly be detected. The SEC profile showed a smaller second peak with lower retention volume but samples taken from this peak did not show all the components of the complex on SDS-PAGE (**Figure 43A/B**). On the other hand, complex formation with the JSGiq chimera could clearly be detected on the SEC profile and was confirmed by SDS-PAGE (**Figure 43C/D**). Because of the silver staining the bands for JSR1 and the G $\beta$  subunit are overlapping but the G $\alpha$  and G $\gamma$  subunits can clearly be seen (**Figure 43D**).



**Figure 43: Analytical SEC for complex formation of JSR1 (ATR611) with the JSGq1 and JSGiq chimera.** Blue lines in the SEC profiles indicate absorption at 280 nm and the orange lines at 505 nm. Both SDS-PAGE were silver stained. **AB:** SEC profile (A) and SDS-PAGE (B) for complex formation with JSR1 and JSGq1. **CD:** SEC profile (C) and SDS-PAGE (D) for complex formation with JSR1 and JSGiq. Abbreviations: JGC=JSR1-G protein complex; JSR=JSR1.

For the JSR1-JSGiq complex, complex dissociation was induced by GDP or GTP $\gamma$ S (non-hydrolysable GTP) and analyzed by analytical SEC in order to validate signaling function of the complex. Moreover, it was tested if the scFv16 or the Fab79 that bind to the hGi protein can stabilize the complex, making it resistant to dissociation by GTP $\gamma$ S. Both the scFv16 and Fab79 have been shown to make the bovine rhodopsin-hGi complex resistant to dissociation (unpublished data from Dr. Ching-Ju Tsai) [119]. Similar to the previous analytical SEC experiments, clear complex formation of JSR1 and JSGiq was detected (**Figure 44A**). Addition of GDP led to a small decrease in the peak signal for the complex, hinting that there is some dissociation (**Figure 44A**). Further decrease of the complex peak signal was detected by addition of GTP $\gamma$ S but there was still a considerable signal, meaning that GTP $\gamma$ S did not lead to the dissociation of all complex molecules, at least not in the experimental conditions used

(Figure 44A). The experiments showed that both the scFv16 and the Fab79 bind to the complex (Figure 44B/C). Addition of the Fab79 lead to a shift in the retention volume from 8.3 ml to 7.9 ml, indicating binding of Fab79 to the complex. The shift for the scFv16 complex was smaller because the scFv16 has a low MW and the peak was shifted only by 0.1 ml from 8.3 ml to 8.2 ml. The JSR1-JSGiq-Fab79 sample showed shoulders of the complex peak in the SEC profile, hinting that it may be a heterogenous population of the complex present. Addition of GTP $\gamma$ S led to complex dissociation in both the JSR1-JSGiq-scFv16 and JSR1-JSGiq79 samples but there was still a considerable amount of complex molecules.



**Figure 44: Analytical SEC experiments with JSR1-JSGiq complex.** Solid lines represent absorption at 280 nm and dashed lines represent absorption at 505 nm. **A:** SEC chromatogram for the JSR1-JSGiq complex (blue) and with addition of GDP (yellow) and GTP $\gamma$ S (green). **B:** SEC chromatogram of the JSR1-JSGiq-scFv16 complex (blue) and with the addition of GTP $\gamma$ S (green). The scFv16 sample is shown in orange. **C:** SEC chromatogram of the JSR1-JSGiq-Fab79 complex (blue) and with the addition of GTP $\gamma$ S (green). The Fab79 sample is shown in orange. Abbreviations: JGAC=JSR1-G protein-antibody complex; JGC=JSR1-G protein complex; JSR=JSR1.

Overall, we could show that the scFv16 and Fab79 bind to the JSR1-JSGiq complex but they do not make the complex completely resistant to dissociation by GTP $\gamma$ S as it was seen for the bovine rhodopsin-hGi complex (unpublished data from our lab; personal communication with Dr. Ching-Ju Tsai).

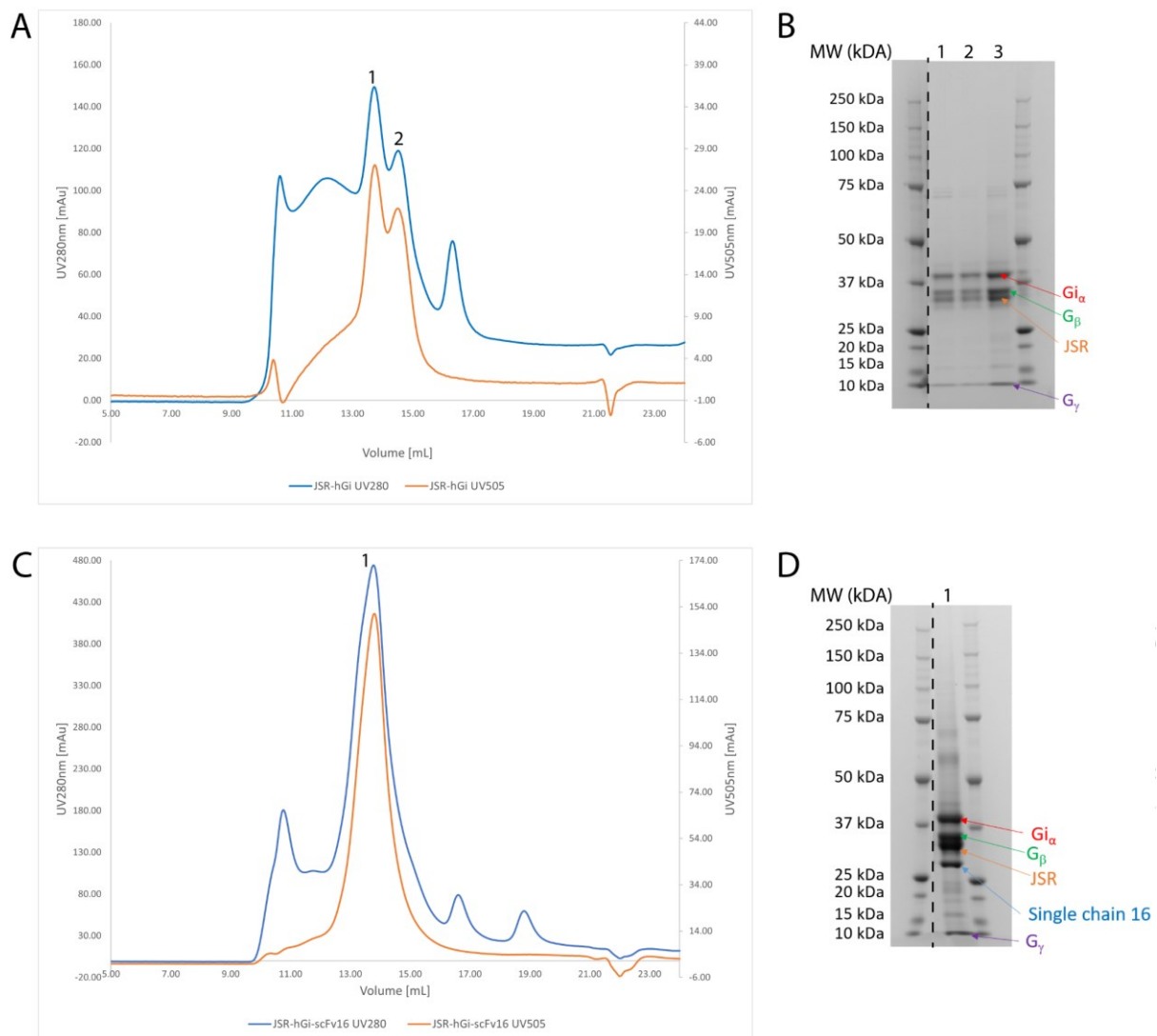
The analytical SEC experiments showed clear complex formation of JSR1 with hGi, hGq and the JSGiq. Because the hGq protein is difficult to work with in terms of expression yield, thermal stability and purification yield, cryo-EM samples were prepared with the hGi and the JSGiq chimera.



#### 4.2.1.6 Preparation of JSR1-G protein complex samples for cryo-EM

The analytical SEC experiments showed which G proteins could be used for complex formation with JSR1 (ATR611) to make a sample suitable for cryo-EM grid preparation and analysis.

The first two samples we prepared were both a JSR1 (ATR611)-hGi complex but once we also included the svFv16 that can potentially help stabilizing the complex. Additionally, the antibody fragment provides an asymmetric feature that can help in aligning the particles from cryo-EM data. G protein was added at a molar excess of 1.25:1 (G protein:JSR1) to JSR1 (ATR611) and incubated for 2h. For the preparation of all the complexes, ~500  $\mu$ g of JSR1 (ATR611) was used. Interestingly, the complex without the scFv16 showed two peaks but on SDS-PAGE both peaks show that all the



**Figure 45: Complex formation and purification of the JSR-hGi (AB) and JSR-hGi-scFv16 (CD) complexes.** Blue lines in the SEC profiles indicate absorption at 280 nm and the orange lines at 505 nm. Both SDS-PAGE were Coomassie stained. **AB:** SEC profile (A) and SDS-PAGE (B) for complex formation with JSR1 and hGi. Lane number three on SDS-PAGE is the combined fractions from peak one and two. **CD:** SEC profile (C) and SDS-PAGE (D) for complex formation with JSR1, hGi and the scFv16.

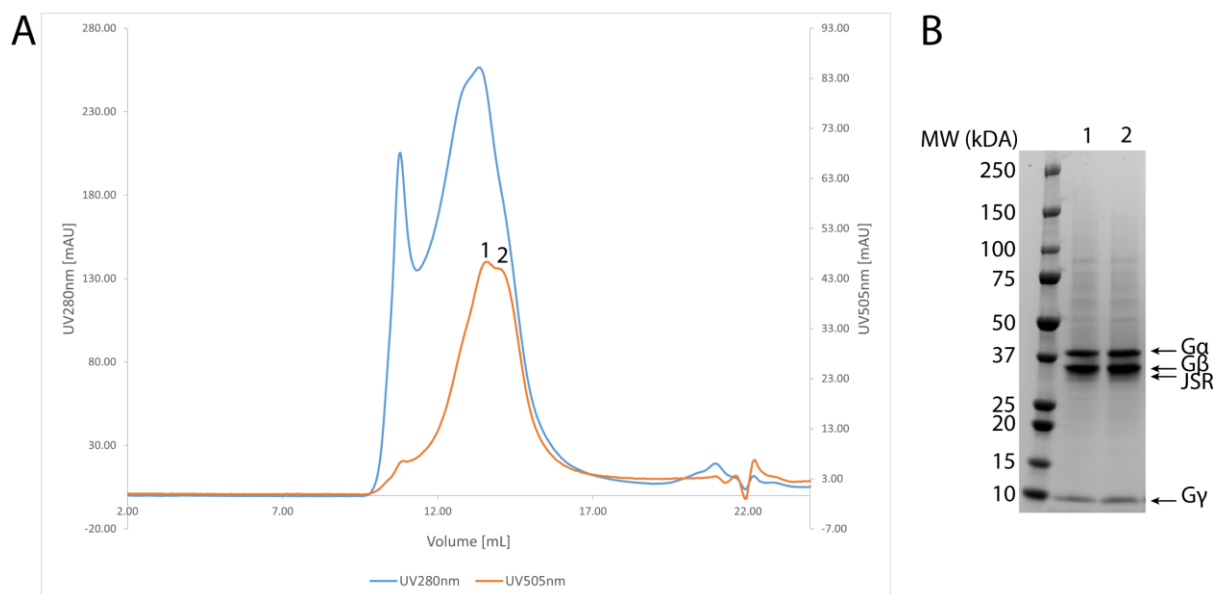
components of the complex are present (**Figure 45A/B**). Therefore, the fractions from these two peaks were combined and concentrated to 4 mg/ml. The concentrated sample was then used to make cryo-



EM grids. Addition of the scFv16 resulted in a single peak, likely creating a more homogenous population compared to the sample without the scFv16 (**Figure 45A/C**). All the components of the complex were clearly visible on SDS-PAGE (**Figure 45D**). The peak fractions were pooled and concentrated to ~3.5 mg/ml for grid preparation.

Although the analytical SEC experiments for the JSR1-Gt complex did not look as good as the ones for the hGi, we decided to prepare the complex and make cryo-EM grids. On SDS-PAGE all of the components of the complex were visible. Fractions were concentrated to ~6 mg/ml and cryo-EM grids were prepared.

Because the cryo-EM analysis of the JSR1-hGi (-scFv16) complexes did not look very promising (see “4.2.2 JSR-Gi complexes cryo-EM data analysis”), we decided to form a complex with another G protein that hopefully forms a more stable complex for structure determination. The hGq also showed clear complex formation in the analytical SEC experiments and could be an option for cryo-EM. However, the JSGi<sub>q</sub> chimera also showed promising results in the analytical SEC experiments and because the chimera was designed based on the JSGq1 sequence (the natural G protein for JSR1) we decided to make a JSR1-JSGi<sub>q</sub> complex sample and prepare cryo-EM grids. The SEC profile showed a void peak with aggregates and then one large peak that actually look like a fusion of two peaks (**Figure 46A**). The



**Figure 46: Complex formation and purification of the JSR-JSGi<sub>q</sub> complex.** Blue lines in the SEC profiles indicate absorption at 280 nm and the orange lines at 505 nm. SDS-PAGE were Coomassie stained. **AB:** SEC profile (A) and SDS-PAGE (B) for complex formation with JSR1 and JSGi<sub>q</sub>.

two peaks are so close together that separating the fractions is not possible. On SDS-PAGE all of the components of the complex are visible over the whole 505 nm peak (**Figure 46B**), therefore the fractions were pooled, concentrated to 8.5 mg/ml and cryo-EM grids were prepared.

#### 4.2.1.7 Summary of JSR1-G protein complex biochemistry

- Purification of JSR1 bound to all-trans 6.11 retinal works well.
- Purification of different G proteins for complex formation with JSR1.
- *In vitro* activity assay showed that all-trans 6.11 retinal is an agonist without illumination.
- *In vitro* activity assays showed that JSR1 can couple and activate hGi, hGq and the JSGiq.
- JSR1 forms a complex, that can be SEC purified and used to make cryo-EM grids, with hGi, hGq and the JSGiq.
- The JSGq1 surprisingly did not form a stable complex with JSR1; invertebrate G proteins might need different buffer conditions than human G proteins.
- Neither hG11, mini-Go or mini-Gsq showed complex formation with JSR1.
- Cryo-EM samples were prepared for the JSR1-hGi, JSR1-hGi-scFv16, JSR1-Gt and JSR1-JSGiq complexes.

## 4.2.2 Cryo-EM grid preparation and data analysis

### 4.2.2.1 Grid preparation and data collection

Cryo-EM grids were prepared for the following samples: JSR1-hGi complex, JSR-hGi-scFv16 complex, JSR1-bGt complex and the JSR1-JSGiq complex. All cryo-EM grids were prepared with a Vitrobot Mark IV at 4°C and 100% humidity. 3 µl of protein sample was applied to the grid for blotting and freezing. The grid type used was mainly Quantifoil R1.2/1.3 Cu 300 or 200 mesh. The blot force was kept constant with a value of 10 and the blot time varied depending on the sample.

The samples for the JSR1-hGi and JSR1-hGi-scFv16 complexes were concentrated to ~4 mg/ml and ~3.5 mg/ml respectively. Several grids were prepared for both samples with blotting times ranging from 2 s to 6 s. Atlas pictures of the grids were acquired on a Titan Krios and datasets on suitable grids were collected at a nominal magnification of 165'000x, corresponding to a pixel size of 0.51 Å.

Grids for the JSR1-bGt complex were prepared with two different concentrations. Once the sample was not concentrated after SEC and grids were prepared with a concentration of 1.25 mg/ml. The blot force was 10 and blotting time was either 2 s or 3 s. The other sample was concentrated after SEC to 6.4 mg/ml and grids were prepared with blot force 10 and blotting time of 3 s or 4s. The atlas pictures of the grids were taken on a Titan Krios microscope. The grids from the unconcentrated sample did not look promising in terms of the ice quality and a dataset was only collected for the concentrated JSR1-bGt complex at a nominal magnification of 165'000x.

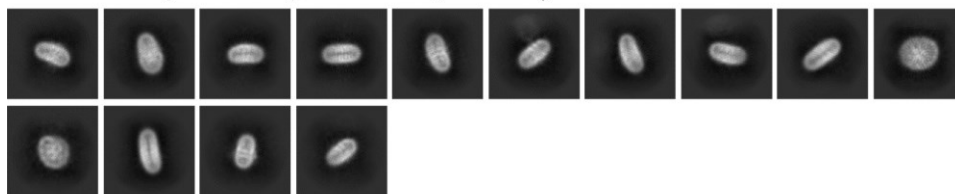
The JSR1-JSGiq complex sample was further concentrated after SEC to 8.5 mg/ml. Grids were prepared with a blot force of 10 and blotting times of 4 s, 6 s or 8 s. Several grids were prepared with the different blotting times and atlas pictures were again taken on a Titan Krios. From two suitable grids, data was collected at a nominal magnification of 165'000x. Because the complex showed a considerable amount of flexibility and the low amount of good particles per micrograph, a large dataset of more than 70'000 micrographs was collected for structure determination (see "4.2.2.4 JSR-JSGiq complex cryo-EM data analysis"). We also tried a different method for preparing cryo-EM grids. We used a machine called Vitrojet (CryoSol) where a drop of the protein solution is put on a pin and the pin "prints" the solution on a grid [142]. This method might reduce the number of particles at the air-water interface which can lead to denaturation of proteins [143]. These grids were screened but not suitable for a large data collection (See next section).

Most datasets were collected together with Dr. Pavel Afanasyev (Cryohub at ScopeM ETHZ), who also taught me how to operate a Titan Krios microscope. Additionally, Pavel also taught me how to process cryo-EM data.

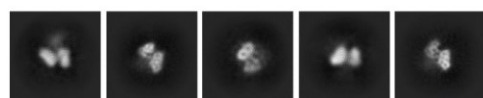
#### 4.2.2.2 Cryo-EM data analysis of the JSR1-bGt complex

For the JSR1-bGt complex only a small dataset was collected since the analytical SEC data did not look as good as for the hGi, hGq and JSGi<sub>q</sub>. Nevertheless, it showed that bGt forms a complex with JSR1 and we wanted to check this by analyzing the sample with cryo-EM. Unfortunately, after particle picking and 2D classification, 2D class averages showed either the micelle or the dissociated bGt heterotrimer (**Figure 47**). The 2D class averages showing the micelles probably have JSR1 in the micelle because there are signs of the TM helices that can be seen (**Figure 47**). This hints that bGt does not form a high affinity complex with JSR1.

2D class averages showing the micelle potentially with JSR1



2D class averages showing dissociated bGt heterotrimer



**Figure 47: Cryo-EM data analysis of the JSR1-bGt complex.** *Top: 2D class averages showing micelles potentially with JSR1 present. Bottom: 2D class averages showing dissociated bGt heterotrimer. Box size is 240x240 pixels with a pixel size of 1.02 Å.*

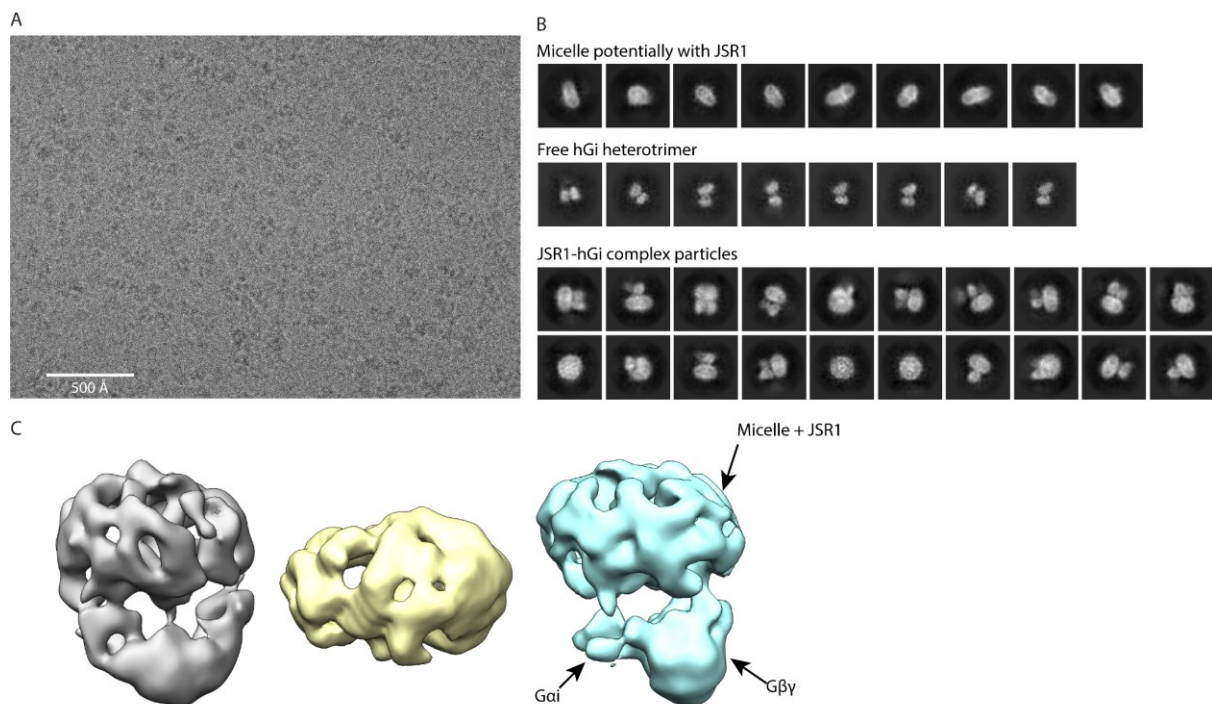
Overall, the JSR1-bGt complex data did not result in any 2D class averages showing the features of the whole complex, meaning that this sample is not stable and falls apart during freezing of the grids. There could be several reasons why the complex dissociated: 1. The JSR1-bGt complex is likely not as stable as the complexes of JSR1 with the hGi, hGq or JSGi<sub>q</sub> as already seen in the analytical SEC experiments; 2. Concentration of the sample after SEC could also lead to dissociation because of the strong centrifugal forces; 3. Protein particles tend to stick to the air-water interface during blotting and freezing which can lead to denaturation and therefore also dissociation of protein complexes [143, 144]. The last two reasons are not only true for the JSR1-bGt complex but also for the other JSR1-G protein complexes (see next two sections).

To conclude, the JSR-bGt complex was not suitable for structure determination because all the complex particles dissociated and no 2D class averages were detected showing the full complex. Therefore, I continued with another G protein to form a better JSR1-G protein complex for structure determination by single particle cryo-EM.

#### 4.2.2.3 Cryo-EM data analysis of the JSR1-hGi complexes

As already mentioned, two JSR1-hGi complex samples were used to make cryo-EM grids. One sample included the scFv16 in addition to the JSR1-hGi complex.

For the JSR1-hGi complex, a suitable grid in terms of ice quality was found and around 11'000 micrographs were collected for analysis. The particle distribution also looked fine, and many particles could be picked from the micrographs (> 2'000'000 particles). Unfortunately, after getting the first 2D class averages, it became clear that the complex dissociation problem seen in the JSR-Gt complex data was also present in the JSR1-hGi data, indicating that both complexes were not very stable (**Figure 47&48**). In contrast to the JSR1-bGt dataset, 2D class averages showing the full JSR1-hGi complex were generated, indicating that this complex is more suitable than the JSR1-bGt complex (**Figure 48B**). However, most 2D class averages of the complex did not show any high frequency information, indicating that the quality is not good enough for high resolution structure determination (**Figure 48B**). Although many particles were picked initially, after 2D classification only ~62'000 particles grouped into 2D class averages that show complex features (~3.1% of initially picked particles).

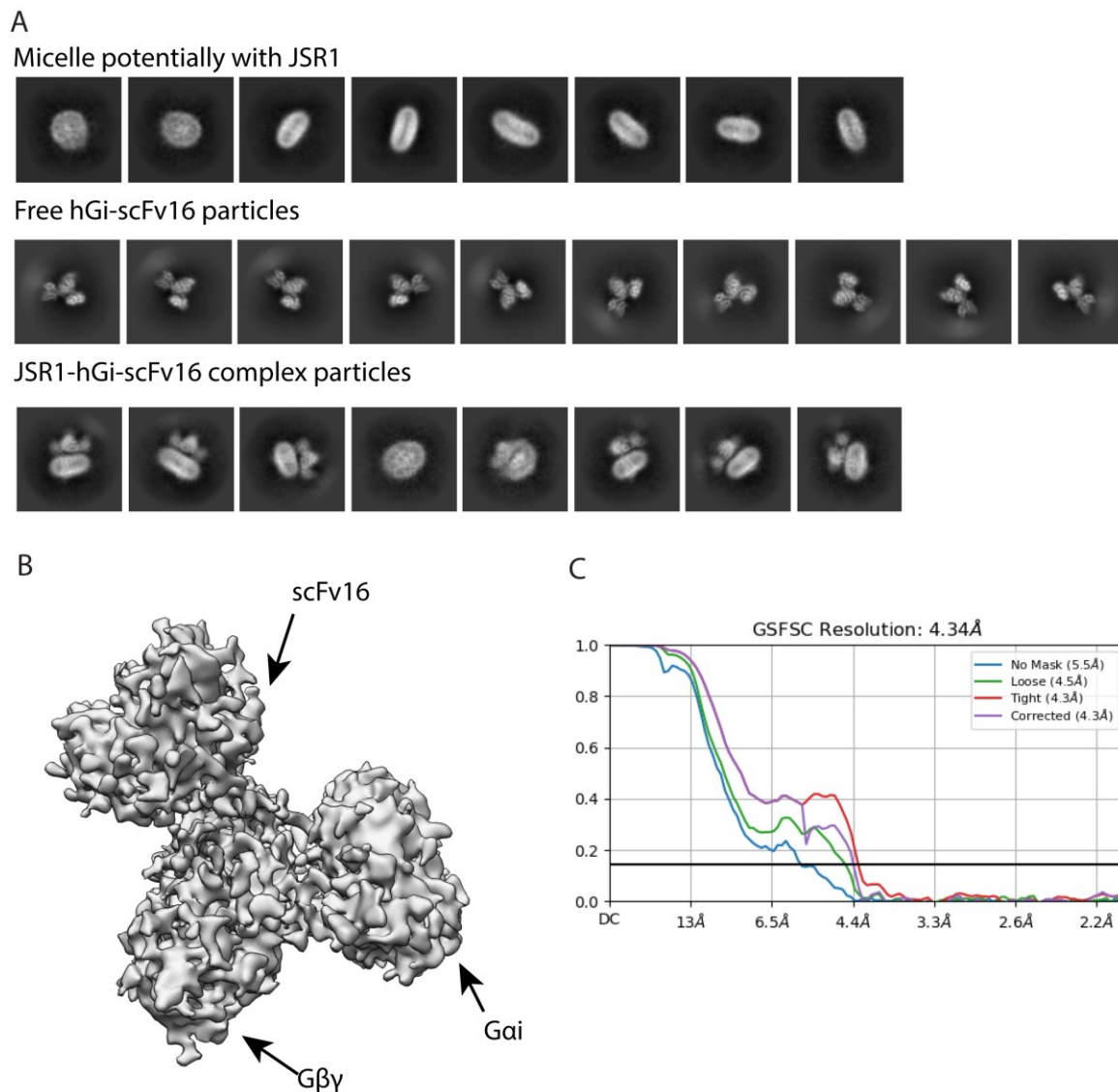


**Figure 48: Cryo-EM analysis of the JSR1-hGi complex.** **A:** Example of an EM micrograph. **B:** 2D class averages showing only the micelle from potentially dissociated complex particles (top); 2D class averages showing free hGi heterotrimer particles (middle); 2D class averages showing the full JSR-hGi complex (bottom). **C:** Ab initio experimental maps generated from the particles in the 2D class averages shown in C at the bottom.

Creating initial models from these ~62'000 particles showed that there are still particles containing components of the dissociated complex (yellow density map in **Figure 48C**). Approximately 53'000 particles grouped into the two density maps that show features of the full complex (**Figure 48C**). Even

so, the number of particles and quality of the data did not allow high resolution structure determination and another sample was prepared for cryo-EM.

Because the JSR1-hGi complex still showed many dissociated complex particles in the cryo-EM dataset, we decided to add the scFv16 to the complex that can potentially help to stabilize the complex. A first dataset was collected and processed. The 2D class averages showed that even with the addition of the scFv16 there were still many complex particles that dissociated, indicating that the scFv16 did not help stabilizing the complex. There were 2D classes of the dissociated hGi-scFv16 sub-complex that showed high frequency information (**Figure 49A**). There were enough particles of this complex to generate an experimental map with 4.34 Å according to the FSC-curve (**Figure 49B/C**) but our focus is actually on the active state of JSR1. Nevertheless, the 2D class averages showing the full complex also showed high frequency information, allowing identification of secondary structure elements in the 2D classes



**Figure 49: Cryo-EM analysis of the JSR1-hGi-scFv16 complex.** A: 2D class averages showing only the micelle from potentially dissociated complex particles (top); 2D class averages showing free hGi-scFv16 particles (middle); 2D class averages showing the full JSR-hGi-scFv16 complex (bottom). Box size is 256x256 pixels with a pixel size of 1.02 Å. B: Experimental 3D map for the hGi-scFv16 sub-complex. C: FSC plot for the experimental map (B).

(**Figure 49A**). However, the number of complex particles was too low to get a high resolution map. Therefore, we decided to collect more data in order to have more complex particles.

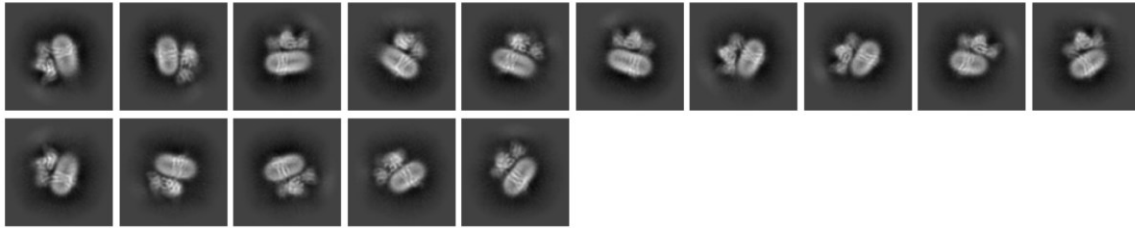
After more data was collected, we could detect good 2D classes showing the full complex and in some 2D class averages we could detect secondary structure information (**Figure 50A**). However, because of the dissociation of many complex particles, the total number of particles after several rounds of 2D classification was still low with ~127'000. Creating three ab-initio maps also showed that some of the particles are not useful for structure determination (light blue map in **Figure 50B**). In addition, in the 2D class averages and the 3D initial maps the micelle size varies (**Figure 50A/B**). This can create problems in aligning the particles because the programs will try to align particles based on the micelle density and not the protein density. After 3D classification only ~40'000 particles were left. 3D refinements without a mask around the protein complex resulted in low resolution maps, presumably also because of the differences in the micelle size. The transmembrane region was not well resolved and not all seven TM helices could be detected. Running 3D refinement with a mask around the protein complex worked better and a map showing all subunits with the TM helices visible was obtained (**Figure 50C**). However, no side chain densities could be seen, and the resolution of the map was only 5.6 Å according to the FSC plot (**Figure 50D**).

Around this time, we received the JSGq1 sequence from Prof. Akihisa Terakita and we designed the JSGiq chimera that showed promising results in the *in-vitro* activity assays and analytical SEC experiments. We therefore decided to prepare a new sample for cryo-EM analysis of the JSR-JSGiq complex, hoping that the JSGiq would form a more stable complex with JSR1 than the hGi.

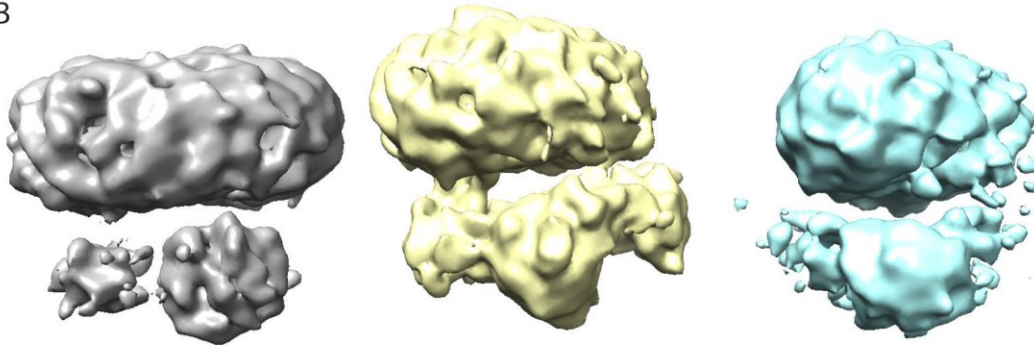


A

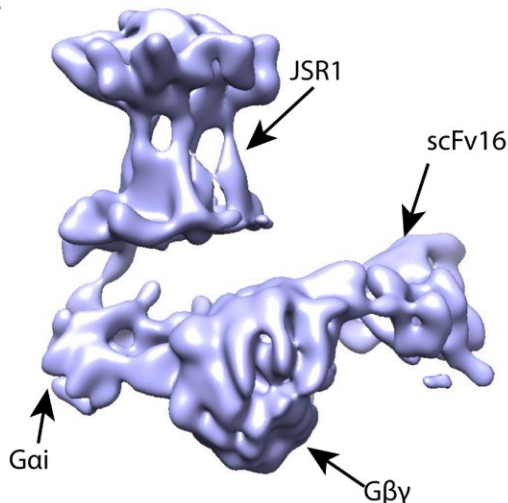
JSR1-hGi-scFv16 complex 2D class averages



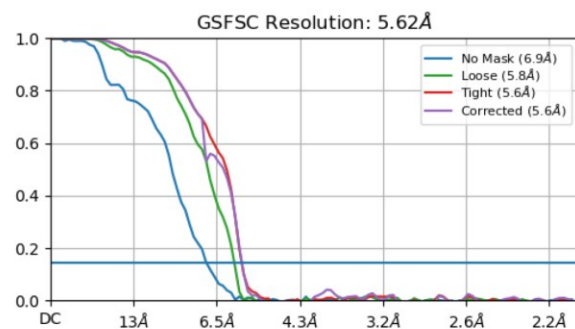
B



C



D



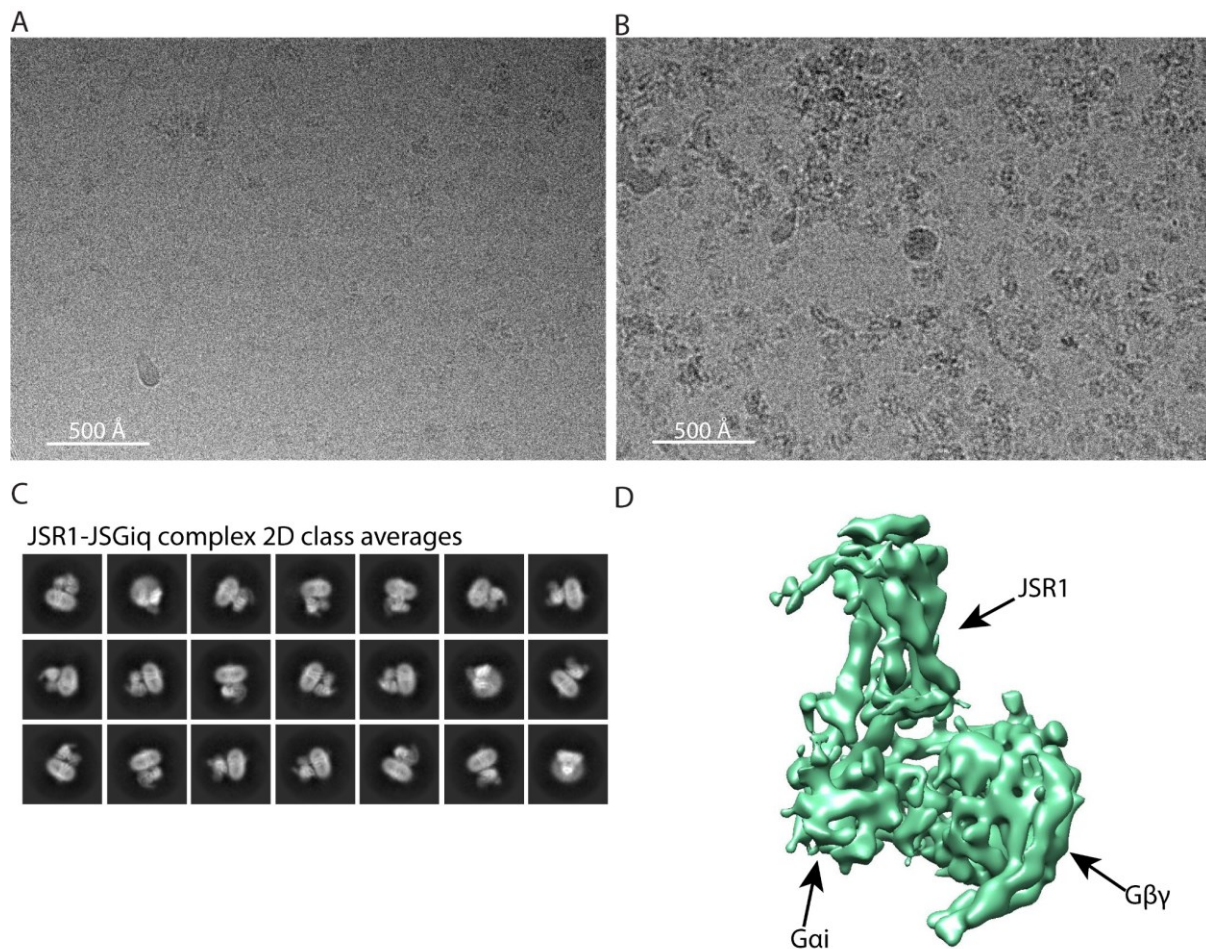
**Figure 50: Cryo-EM analysis of the JSR1-hGi-scFv16 complex.** **A:** 2D class averages showing the JSR1-hGi-scFv16 complex in different orientations. Box size is 256x256 pixels with a pixel size of 1.02 Å. **B:** Ab-initio experimental maps from the particle stack. Each color represents a different experimental map. **C:** 3D refined experimental map using a mask around the protein complex. Subunits are labelled. **D:** FSC plot for the experimental map in (C).

#### 4.2.2.4 Cryo-EM data analysis of the JSR1-JSGiq complex

The JSGiq chimera is based on the hGai and residues, mainly in the  $\alpha 5$  helix, were mutated to match the JSGq1 sequence (**Figure 40**). We prepared several grids, screened them and identified two grids for data collection. We started collecting from very thin ice areas but after some time, we decided to change the ice filter a bit to include areas with slightly thicker ice. This was done because micrographs from very thin ice areas did not show many particles whereas micrographs from areas with slightly thicker ice showed many more particles (**Figure 51A/B**). Nevertheless, from the first dataset good looking 2D class averages were obtained showing secondary structure features (**Figure 51C**). From the



high concentration of the protein sample (~8.5 mg/ml) we would have expected to have more particles



**Figure 51: Cryo-EM analysis of the JSR1-JSGiq complex.** **A:** Example of an EM micrograph from a very thin ice area. **B:** Example of a micrograph from a slightly thicker ice area than (A). **C:** 2D class averages showing the JSR1-JSGiq complex in different orientations. Box size is 240x240 pixels with a pixel size of 1.02 Å. **D:** Experimental map generated from the particles in the 2D class averages shown in C.

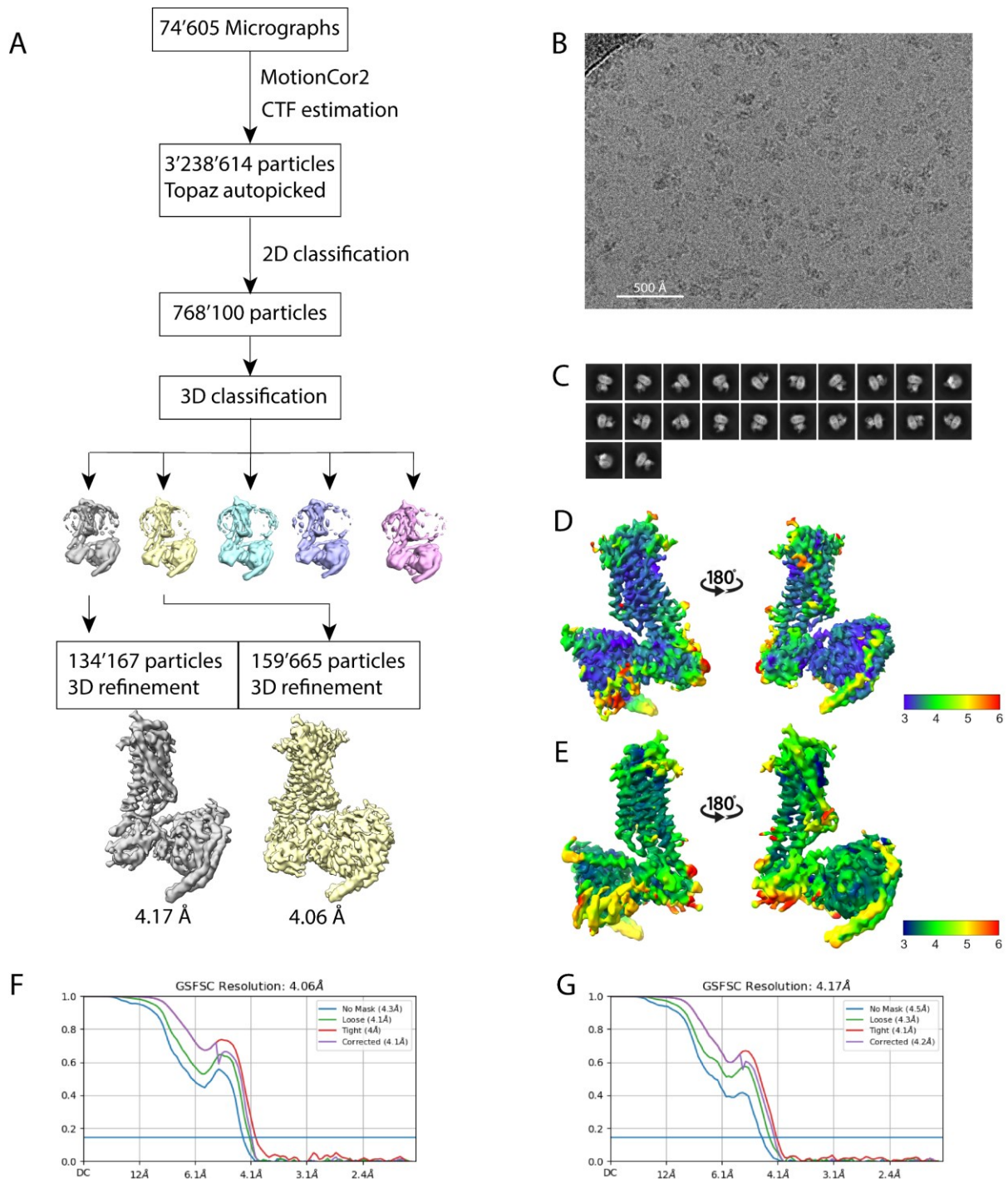
but in the end we only had ~70'000 particles after 2D and 3D classification. Still, the 3D map from these particles looked already better than every map from the JSR-hGi-scFv16 complex (**Figure 51D**). The 3D experimental map looked promising, but the resolution was not sufficient yet to see side chain densities (**Figure 51C**). Additionally, a 3D variability analysis showed that the G protein possesses large flexibility which needs further 3D classification to isolate distinct conformational states.

We collected another large dataset which resulted in a combined dataset of more than 70'000 micrographs. After several rounds of 2D classifications, 768'100 particles grouped into 2D class averages showing the JSR1-JSGiq complex (**Figure 52C**). These particles were used for a 3D classification which allowed the separation of particles into two different conformations of the complex (**Figure 52A**). These two maps were refined to a global resolution of 4.06 Å (JSR1-JSGiq\_1) and 4.17 Å (JSR1-JSGiq\_2) including all subunits of the complex (**Figure 52A/F/G**). The main difference between the two maps is the orientation of the G protein which will be discussed in "4.2.3 JSR1-JSGiq structure analysis". The local resolution maps showed that the TM core and Gβ core of the JSR1-

JSGiq\_1 complex has high resolution with values around 3-3.5 Å whereas for the JSR1-JSGiq\_2 complex it is between 3.5-4 Å (**Figure 52D/E**). In both maps the G $\alpha$  subunit shows the lowest resolution. This is not surprising for the alpha helical (AH) domain of the G $\alpha$  subunit because it is known to be flexible, but also the Ras domain has lower resolution compared to the other subunits (**Figure 52D/E**). This likely comes from the flexibility of the G protein and this flexibility also influences the resolution in the ICL3 and cytoplasmic end of TM6 from JSR1 which shows lower resolution compared to the rest of JSR1 (**Figure 52D/E**). Overall, the JSR1-JSGiq\_1 map showed good side chain densities for most parts of the receptor which allowed modelling of the active state JSR1.

Several approaches were explored to increase the resolution like using different masks for refinement or trying different 3D classification methods. Higher resolution maps were obtained with masks but the maps showed density artifacts along the borders of the masks which is faulty. Visual comparison of these maps with the JSR1-JSGiq\_1 map also showed that there is no improvement in the density of JSR1 which we are mostly interested in. Processing was mainly done with cryosparc and changing the software to relion also did not improve resolution.

In the end, the map of JSR1-JSGiq\_1 was used for model building and the refined model was then used to build the model for the JSR1-JSGiq\_2 complex. Density for the AHD was seen in both complexes but the resolution was not sufficient to confidently build side chains. The main model was built without the AHD (for the publication) but to analyze the AHD position in both cases, models including the AHD were built too. These will however only be analyzed regarding the AHD comparison.



**Figure 52: EM processing of the JSR1-JSGiq complex.** **A:** EM processing workflow for the JSR1-JSGiq complex. **B:** Example of a micrograph taken at a nominal magnification of 165'000. **C:** 2D class averages for the JSR1-Giq complex in different orientations. Box size is 240x240 pixels with a pixel size of 1.02 Å. **D:** 4.06 Å JSR1-JSGiq complex electron potential map colored by local resolution values. **E:** 4.17 Å JSR1-JSGiq complex electron potential map colored by local resolution values. **F:** FSC plot for the JSR1-JSGiq complex at 4.06 Å resolution. **G:** FSC plot for the JSR1-JSGiq complex at 4.17 Å resolution.

#### 4.2.2.5 Summary cryo-EM analysis of JSR1-G protein complexes

- The data of the JSR1-bGt complex showed no 2D class averages of the full complex but all complex particles seem to be dissociated or denatured.
- The JSR1-hGi complex also showed many dissociated complex particles but in contrast to the JSR1-bGt sample, 2D class averages showing the full complex were detected.
- The number and quality of the JSR1-hGi particles was not sufficient to create a high-resolution electron density map.
- Addition of the scFv16 to the JSR1-hGi complex did not help to stabilize the complex.
- Addition of the scFv16 likely helped in the alignment of the particles and allowed detection of secondary structure information in the 2D class averages.
- The number and quality of the JSR1-hGi-scFv16 only resulted in a low-resolution electron density map.
- The JSGi<sub>q</sub> chimera forms a more stable complex with JSR1 than the hGi.
- Two maps of the JSR1-JSGi<sub>q</sub> complex were created at 4.06 Å and 4.17 Å.
- 3D variability analysis shows the flexibility of the G protein, particularly the G $\alpha$  subunit.

### 4.2.3 JSR1-JSGiq structure analysis

In our lab we have now a 4.9 Å structure of JSR1 (all-trans retinal) bound to the hGi solved by Dr. Filip Pamula and two structures of the JSR1 (ATR611)-JSGiq complex. In Filip's case, JSR1 was reconstituted with 9-cis retinal as described by Varma *et al.* [30]. A JSR1-hGi complex with all-trans retinal was prepared that resulted in a 4.9 Å resolution cryo-EM map. Several reasons prevented determination of a higher resolution structure. First, JSR1 bound to the 9-cis retinal needs to be light activated to couple to the G protein and even though the  $\lambda_{\max}$  of the active and inactive states are shifted (9-cis vs all-trans JSR1), they still overlap, leading to a mixed population of active/inactive state JSR1 upon illumination [100]. Second, the low stability of the JSR1-hGi complex resulted in dissociation during freezing for grid preparation, lowering the number of complex particles in the EM data. Lastly, the flexibility of the complex resulted in heterogeneity in the particle set, lowering the overall resolution of the cryo-EM map.

An alternative strategy was therefore adopted for my work in order to determine the structure of the JSR1-G protein complex. Firstly, rather than activating the receptor with light, JSR was reconstituted with ATR6.11, a non-natural retinal analogue that has an agonist activity (see "4.2.1.4 GTPase Glo assay JSR-G proteins). Secondly, the JSR1-hGi structure was used to identify residues in the G protein that interfaced with the receptor. These residues, mainly in the  $\alpha 5$  helix, were substituted to match the amino acids of the jumping spider visual Gq1 G protein (**Figure 40**; see Methods for details). In short, the sequence alignment for the  $\alpha 5$  helix of the human G $\alpha_i$ , human G $\alpha_q$  and jumping spider G $\alpha_{q1}$  revealed several differences (**Figure 40**). The C-terminus of the jumping spider G $\alpha_{q1}$  resembles a combination of the human G $\alpha_i$  and G $\alpha_q$ . We therefore designed a chimeric G protein (JSGiq) based on the hGi and mutated residues to match the jumping spider Gq1 sequence.

The JSR1-JSGiq complex was reconstituted in DDM and used for structure determination (**Figure 46**). Cryo-EM analysis resulted in a map at 4.06 Å global resolution including all domains of the complex (JSR1-JSGiq\_1) (**Figure 52**). This structure will be the main focus in this section because the improvement in resolution compared to JSR1-hGi allowed modelling of side chains into the experimental map. 3D classification and 3D variability analysis revealed flexibility in the particle set; the G $\alpha$  subunit shows a particularly high degree of flexibility that lowers the resolution in this area and therefore also the global resolution (**Figure 52D/E**). In addition, there are larger motions of the whole G protein relative to JSR1. With 3D classification, a second conformation of the complex was isolated and refined to a global resolution of 4.17 Å (JSR1-JSGiq\_2) (**Figure 52**). These EM maps were used for model building of the JSR1-JSGiq complexes. The JSR1-JSGiq\_1 structure will be discussed in detail in the following sections and the JSR1-JSGiq\_2 and Filip's JSR1-hGi will mainly be discussed regarding the G protein binding site (see "4.2.3.4 JSR-G protein interactions").

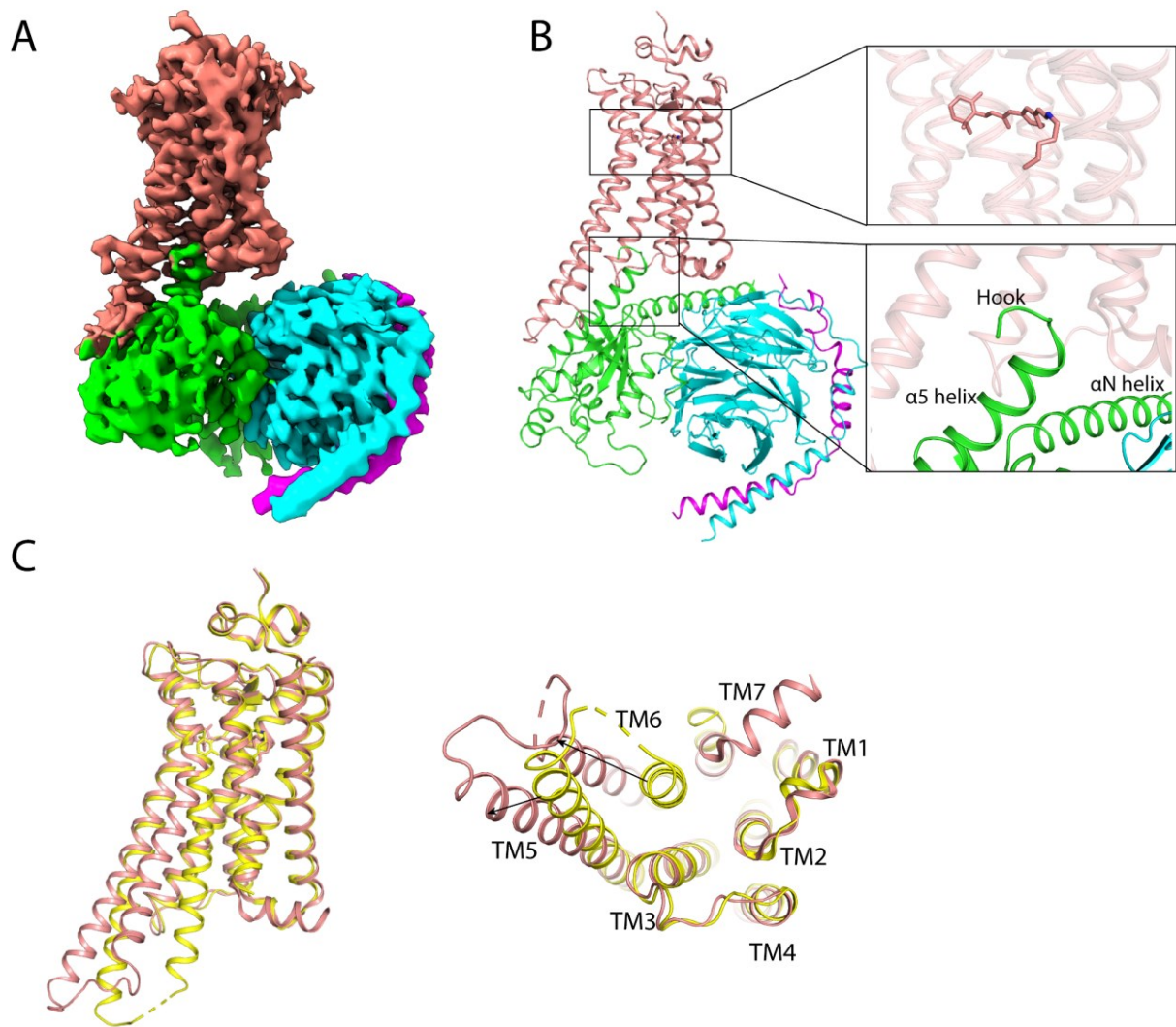
#### 4.2.3.1 JSR1-JSGiq complex overall structure

The structure of the JSR1-JSGiq complex shows the typical features of a class A GPCR-G protein complex. The outward movement of the TM5 and TM6 (relative to the inactive state) opens a cytoplasmic cleft that allows the JSGiq to bind to JSR1 (**Figure 53**). The C-terminus of JSGaiq forms a typical hook-like structure that inserts into the formed binding cleft as observed in all the GPCR-G protein complexes (**Figure 53B**). In contrast to other GPCRs, JSR1 has an elongated TM5 that is involved in interactions with the G protein (see “4.2.3.4 JSR-G protein interactions”). Due to the flexibility of the ICL3 and cytoplasmic end of TM6 the resolution in this area is lower compared to the rest of JSR1 (**Figure 52D**). For a small part at the end of ICL3/start of TM6 the experimental map did not allow confident modelling. The main difference between the two JSR1-JSGiq structures is the orientation of the G protein relative to JSR1 (**Figure 60C/D**) leading to different interactions of the JSGiq protein with JSR1 (discussed in detail later). The map quality allowed modelling for most residues in the G protein (**Figure 53A/B**). The map also showed clear density for the AHD of the JSGaiq subunit but because of the low resolution in this area confident modelling of the side chains of the AHD was not possible. The structures shown here do not include the AHD and the AHD position will only be analysed in 4.2.3.4.

The N-terminus and the extracellular loop 2 (ECL2) of JSR1 form a cap over the orthosteric binding site, separating the ligand binding site from the extracellular milieu (**Figure 53B, Figure 66**). This cap is conserved in all opsins but the structure of the cap can vary between receptors [30, 31, 145, 146], indicating that it might be important for the stability of the retinal-binding pocket and receptor stability overall.

In the next sections we will discuss the JSR1-JSGiq\_1 structure in detail from the ligand binding site to the G protein binding site.

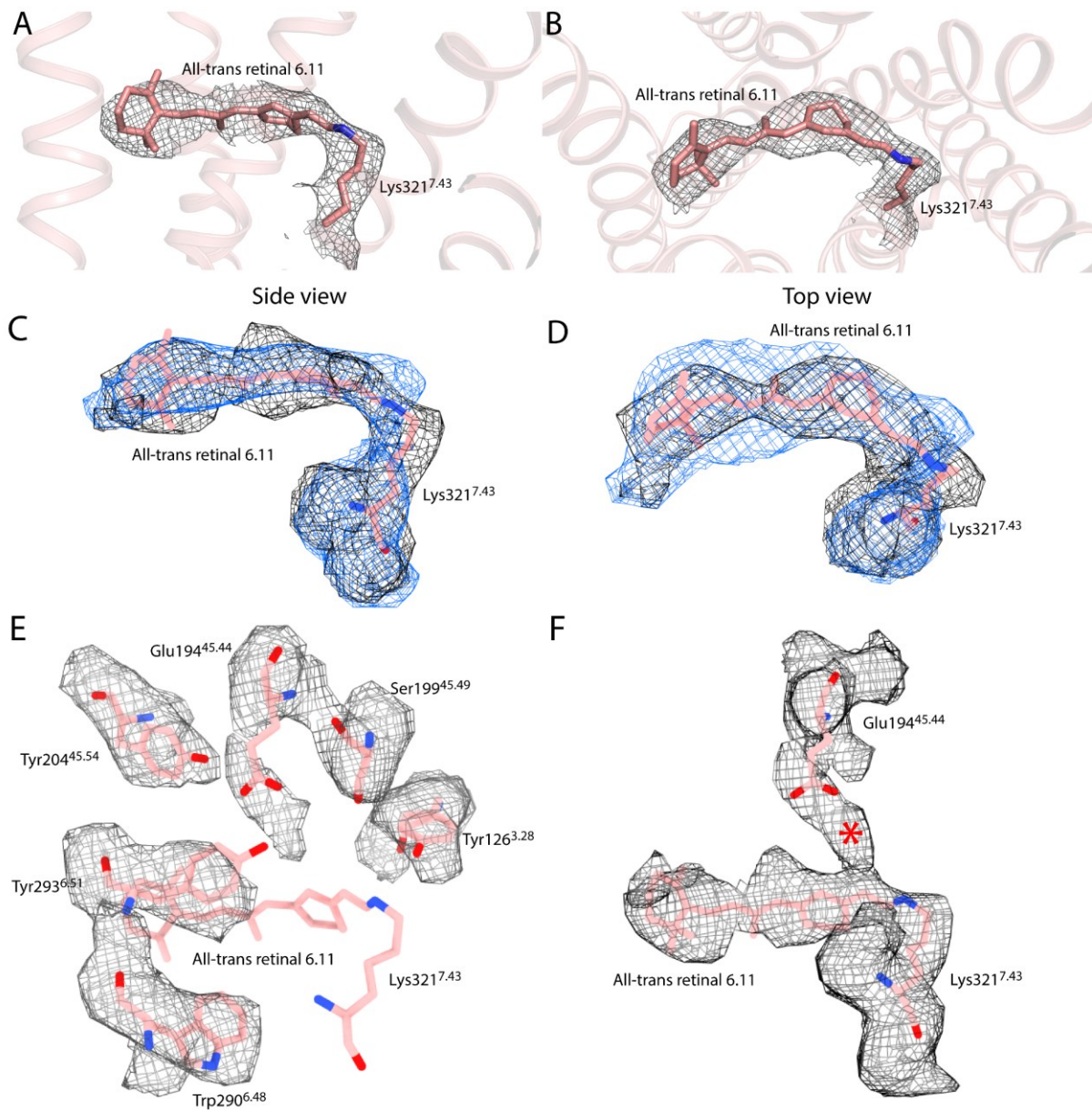




**Figure 53: Overall structure of the JSR1-JSGiq<sub>1</sub> complex.** **A:** Cryo-EM map colored by subunit. JSR1, JSGaiq, G $\beta$  and Gy subunits are shown in salmon, green, cyan and magenta, respectively. **B:** Overall model of the JSR1-JSGiq complex colored in the same way as the cryo-EM map. The retinal binding site and the G protein binding site are shown enlarged in separate windows. **C:** Overlay of the inactive JSR1 (yellow) and the active JSR1 (salmon). Left shows a side view and right shows a view from the cytoplasmic side of the receptor.

#### 4.2.3.2 The retinal binding pocket

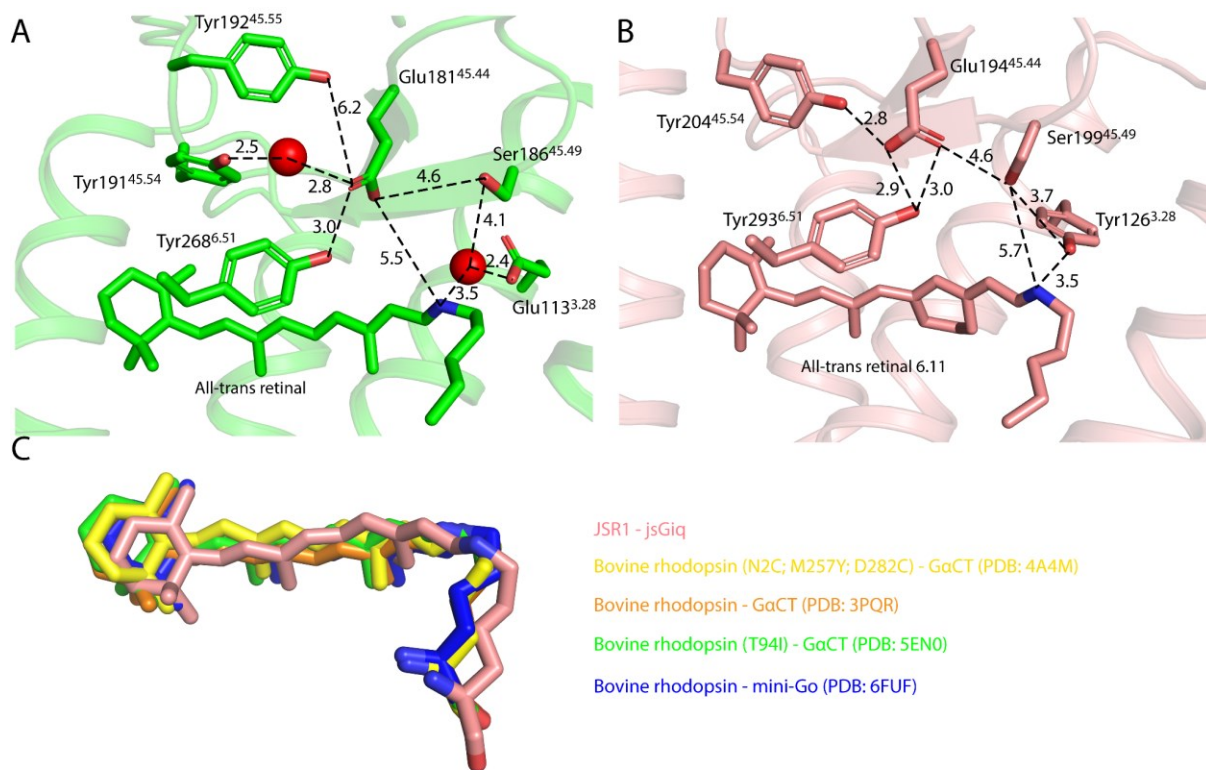
The electron density in the orthosteric binding pocket confirms the presence of a covalently linked retinal to Lys321<sup>7.43</sup> via a protonated Schiff base (**Figure 54A/B; 55B**) (Personal communication with Dr. Matthew Rodrigues). In both cryo-EM maps, the density for the retinal ligand is clearly visible and highly conserved, meaning that the retinal pose is similar between the different conformations of the JSR1-JSGiq complex. Even in the lower resolution JSR1-hGi map, density for the all-trans retinal was visible and overlays well with the ATR6.11 analogue (**Figure 54C/D**). The biggest difference in the JSR1-hGi structure is the lysine position and the Schiff base link that adopts a different conformation than in the JSR1-JSGiq structures (**Figure 56A**). However, the low resolution does not allow unambiguous modelling which might explain the difference in the JSR1-hGi structure.



**Figure 54: All-trans 6.11 retinal with the map and comparison to the all-trans retinal from the JSR1-hGi complex.** **AB:** Visualization of the all-trans 6.11 retinal from the JSR1-Giq complex including the experimental map for the Lys321 and all-trans 6.11 retinal showed as an isomesh. A shows a side view and B shows a top view from the extracellular side. **CD:** Overlay of the all-trans 6.11 retinal with the density map from the JSR1-JSGiq complex (black mesh) and the density map of all-trans retinal from the JSR1-hGi complex (blue mesh). C shows a side view and D shows a top view from the extracellular side. **E:** Residues in the retinal environment shown with the experimental map overlaid as a mesh. **F:** Counterion (Glu194) and retinal with the Schiff base link with the experimental map overlaid. The red asterisk marks a potential water molecule position.

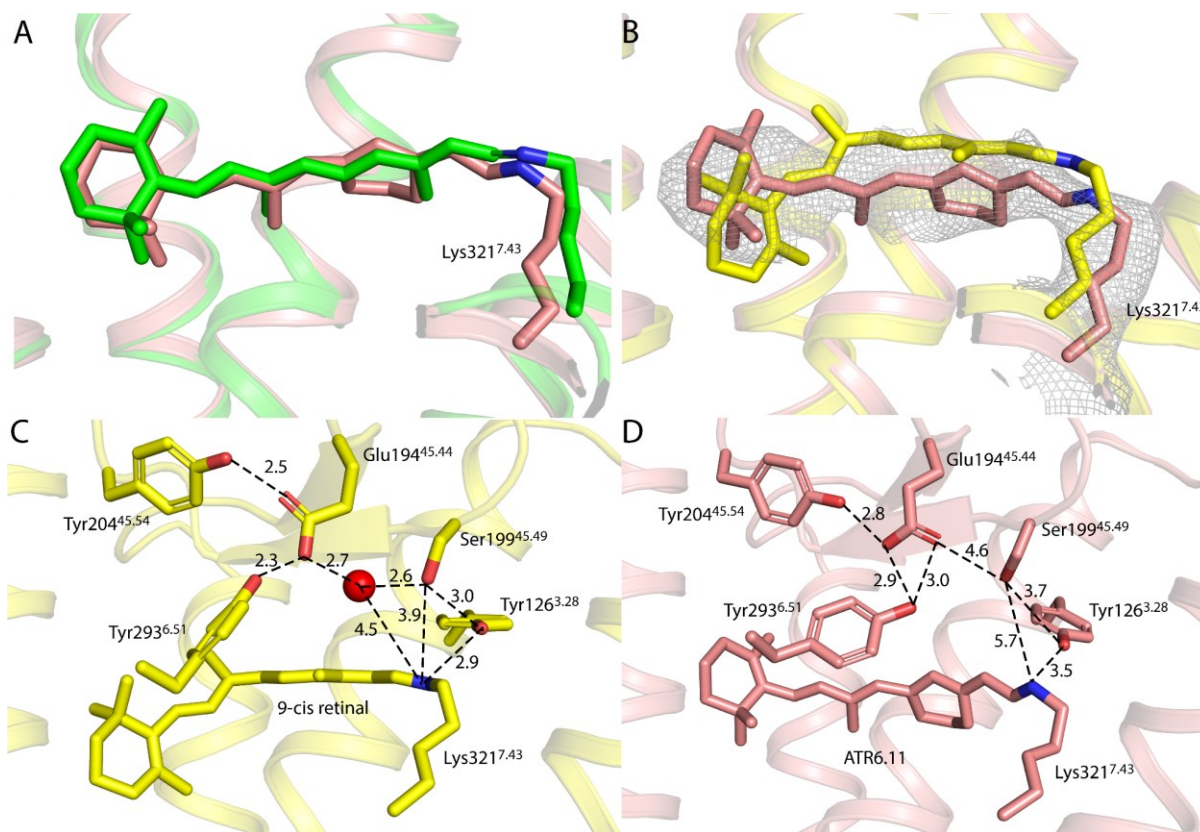
The density for the  $\beta$ -ionone ring of ATR6.11 is not sufficiently well defined to determine its orientation based on the experimental map alone. We have therefore modelled it rotated 180° relative to that of the 9-cis retinal  $\beta$ -ionone ring, which is consistent with the pose observed for all-trans retinal in the active bovine rhodopsin metarhodopsin-II structure [30] [146-149](**Figure 55C; 56B**). The polyene chain shifts to the intracellular side compared to the 9-cis retinal while the  $\beta$ -ionone ring shifts further toward the extracellular side (**Figure 56B**).





**Figure 55: Retinal binding site comparison of bovine metarhodopsin-II and active state JSR1.** **A:** Retinal environment of bovine metarhodopsin II (PDB:5EN0). Important residues are shown as sticks. Carbon atoms are colored green, nitrogen atoms blue and oxygen atoms red. Waters are shown as red spheres. **B:** All-trans retinal 6.11 environment of the active JSR1. Corresponding residues from A are shown as sticks for JSR1 and carbon atoms are shown in salmon, nitrogen atoms in blue and oxygen atoms in red. **C:** Overlay of the all-trans retinal 6.11 from our JSR1 structure and all-trans retinal from four bovine metarhodopsin II structures (PDB: 4A4M (yellow); PDB: 3PQR (orange); PDB: 6FUF (blue); PDB: 5EN0 (green)).

In contrast to the monostable bovine rhodopsin, the Schiff base in JSR1 remains protonated upon photon absorption and isomerisation from cis to trans conformation [91, 98]. The counterion, a negative charge in the ligand binding pocket, is important for the stabilization of the protonated Schiff base link. Monostable rhodopsins have a proximal Glu113<sup>3.28</sup> performing the role of the counterion [87, 145] (**Figure 55A**). In JSR1, as well as other bistable rhodopsins, this position is occupied by a conserved tyrosine (Tyr126<sup>3.28</sup>), and a distal Glu194<sup>45.44</sup> plays the role of the counterion [30, 88, 98] (**Figure 55B; 56C/D**). The crystal structure of the inactive state of JSR1 showed that the counterion is too distant for a direct interaction and is instead linked to the protonated Schiff base via a water-mediated hydrogen bond network (**Figure 56C**) [30]. The retinal binding pocket undergoes subtle changes from the inactive to the active state of JSR1. The polyene chain and the Schiff base move further towards the intracellular side, increasing the distance to the Tyr126<sup>3.28</sup> (2.9 Å to 3.5 Å), Ser199<sup>45.49</sup> (3.9 Å to 5.5 Å) and the distal counterion Glu194<sup>45.44</sup> (6.7 Å to 6.8 Å) (**Figure 56C/D**) [30]. The distal counterion also changes the rotamer in the active state but is still oriented by the “glutamate cage” formed by Tyr204<sup>45.54</sup> and Tyr293<sup>6.51</sup> (**Figure 56C/D**). Densities reconstructed for the side chains mentioned are shown in **figure 54E**.



**Figure 56: Retinal binding site comparison of the inactive and active states of JSR1.** **A:** Comparison of the all-trans retinal (JSR1-hGi complex) and the all-trans retinal 6.11 (JSR1-JSGiq complex) **B:** Comparison of the 9-cis retinal of the inactive JSR1 (6I9K, yellow) and the all-trans retinal 6.11 of the active JSR1 (salmon). Carbon atoms are colored in the respective color of JSR1 and nitrogen atoms are blue. Lysine 321, the Schiff base link and the retinal ligand are shown as sticks. Experimental map is shown for the all-trans retinal 6.11. **C:** 9-cis retinal environment of the inactive JSR1. Residues involved in the Schiff base-counterion link are shown as sticks. Carbon atoms are colored yellow, nitrogen atoms blue and oxygen atoms red. **D:** All-trans retinal 6.11 environment of the active JSR1. Same residues as in B are shown as sticks and carbon atoms are shown in salmon, nitrogen atoms in blue and oxygen atoms in red.

Previous experiments suggested that Glu194 also acts as the counterion in the active state of JSR1 but that Ser199 is not involved in the protonated Schiff base-counterion link [98]. In our structure, the proton of the protonated Schiff base is oriented towards the Glu194 but is not close enough for a direct interaction. Ser199 adopts a different conformation compared to the inactive state and orients away from the protonated Schiff base, implicating that Ser199 is not involved in the protonated Schiff base-counterion link (**Figure 56D**). In the crystal structure of the inactive state there are ordered water molecules around the protonated Schiff base involved in key polar interactions linking to the counterion (**Figure 56C**) [30]. In our structure there is sufficient space in the binding pocket for water molecules to mediate a hydrogen bond network between the protonated Schiff base and the counterion. Interestingly, there is a density in the experimental map that connects the counterion with the protonated Schiff base, possibly coming from 1-2 ordered water molecule(s) (**Figure 54F**). The retinal binding pocket of bovine rhodopsin in the active state also shows that ordered water molecules are present (**Figure 55A**).

The retinal isomerization is the first step in the activation process that triggers conformational changes in the protein, leading to the active state JSR1.

#### 4.2.3.3 Micro-switch domains in JSR1

In GPCRs, there are several conformational changes in conserved micro-domains upon activation of a receptor. In this section we will investigate three major conserved motifs and compare these to the inactive state JSR1 and monostable bovine rhodopsin. These motifs are the C-W-x-P motif in TM6, the P-I-F motif between TM3/5 and 6 and the D/E-R-Y motif in TM3. The model and the density for the side chains mentioned can be seen in **figure 57**.

Although the C<sup>6.47</sup>-W<sup>6.48</sup>-x-P<sup>6.50</sup> motif is highly conserved in class A GPCRs, invertebrate rhodopsins do not possess a cysteine at position 6.47 but either an alanine or serine (**Figure 58G**). Pro<sup>6.50</sup> is conserved in the opsin family (**Figure 58G**) and acts as a hinge in the activation process for the movement of TM6. Interestingly, the proline is followed by a conserved tyrosine at position 6.51 in most opsins and is implicated in the coordination of the distal counterion (**Figure 55A/B; 56C/D; 58G**) [30, 145, 149]. Trp<sup>6.48</sup>, also called the toggle switch, is positioned just below the retinal binding site and is one of the major activation micro-switches. The polyene chain of the all-trans retinal “pushes” on Trp<sup>6.48</sup> resulting in a cascade of structural changes that lead to the outward movement of TM6 (**Figure 57A; 58A**). Trp<sup>6.48</sup> mainly shifts towards the intracellular side but not much laterally like bovine rhodopsin (**Figure 57A; 58A/B**) [30, 145, 149]. This shift initiates the outward movement of TM6 and subsequent conformational changes of the residues in TM6.

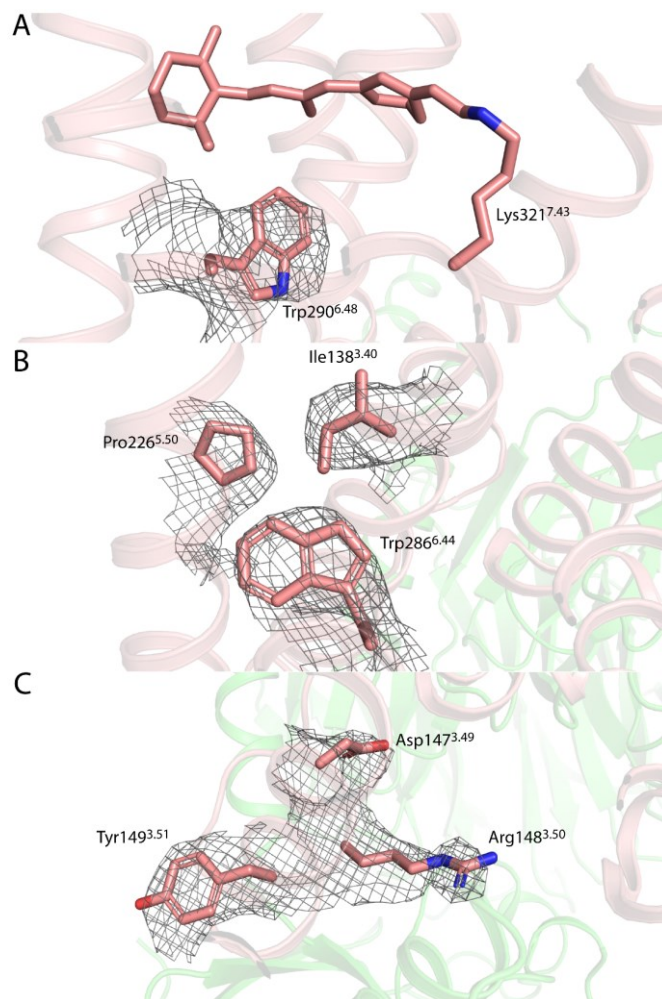
One helix turn below Trp<sup>6.48</sup> is Trp<sup>6.44</sup>, which is part of the P-I-F motif. The P<sup>5.50</sup>-I<sup>3.40</sup>-F<sup>6.44</sup> motif is conserved in class A GPCRs but we found that JSR1 possesses a tryptophan at position 6.44. Sequence alignments show that most opsins have a phenylalanine at position 6.44 (~80%), some have a tyrosine (~19%) and only very few have a tryptophan (~1%) (**Figure 58H**). Curiously, the majority invertebrate rhodopsins have a tryptophan at this position. The position of the tryptophan is highly similar to the position of the phenylalanine in bovine rhodopsin in both the active and inactive states (**Figure 58C/D**)

[30, 145, 149]. In contrast to bovine rhodopsin, the translational shift of proline is smaller and the isoleucine does not undergo a rotamer change but only a small translational shift of the backbone (**Figure 57B; 58C/D**) [30, 145, 149].

The third micro-switch is the conserved D/E<sup>3.49</sup>-R<sup>3.50</sup>-Y<sup>3.51</sup> motif, which is crucial in stabilizing the active state and the interaction with the G protein in the active state (**Figure 58I**). In the inactive state structure of JSR1, the DRY motif is in a “closed” conformation and Asp147<sup>3.49</sup> forms an intra-helical ionic interaction with Arg148<sup>3.50</sup>. Arg148<sup>3.50</sup> forms another inter-helical ionic interaction with Glu272<sup>6.30</sup>. These ionic interactions are characteristic of the inactive state of class A GPCRs, although the inter-helical ionic interaction is less conserved than the intra-helical ionic interaction (**Figure 58E/F**) [18, 19, 22, 30]. In the active state

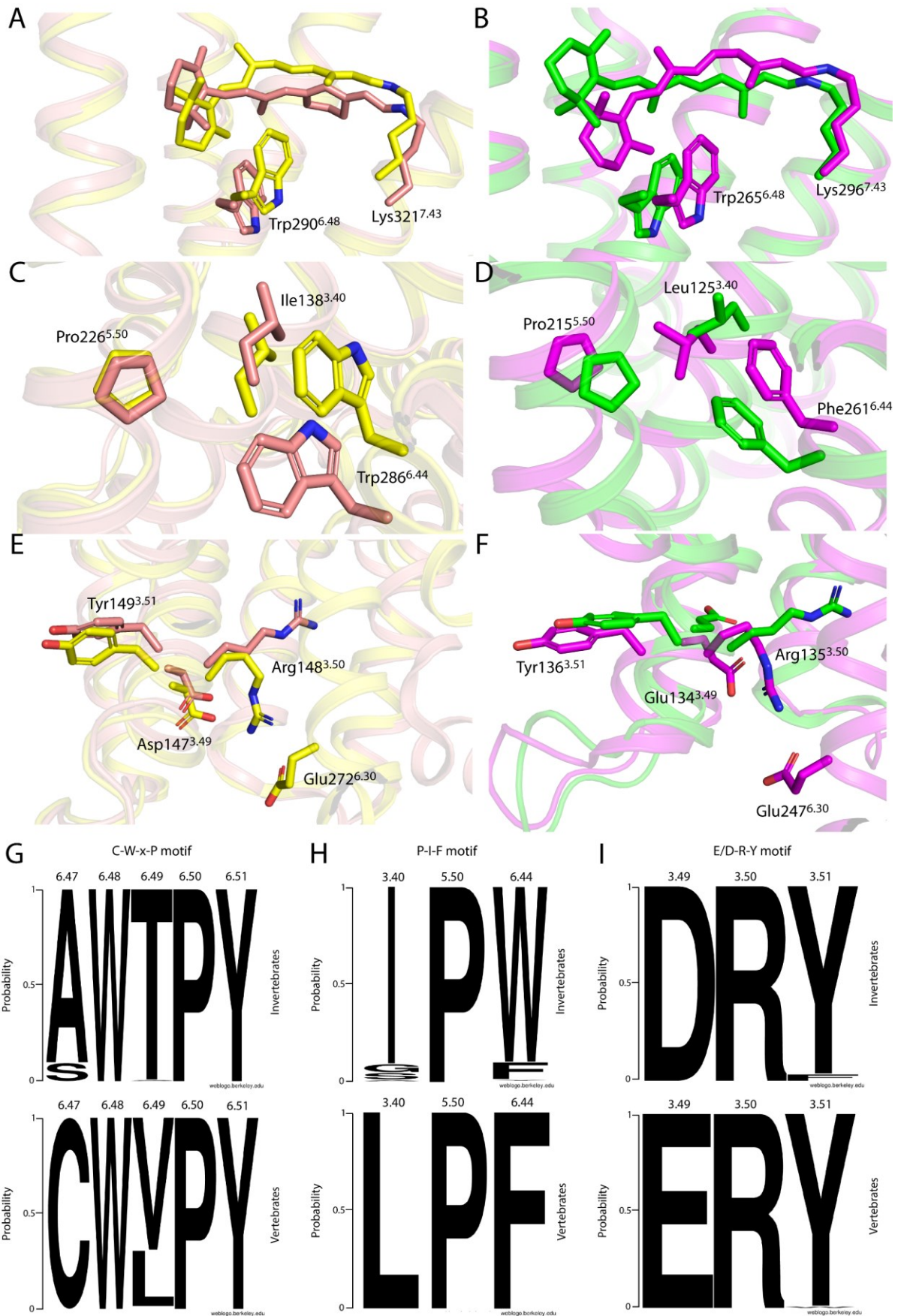
of JSR1, this ionic lock is broken and Arg148<sup>3.50</sup> adopts a different rotamer that interacts with the G protein (**Figure 57C; 58E**). This rotamer change is further stabilized by an interaction with Tyr234<sup>5.58</sup> (**Figure 69; 70**). This is very similar to bovine rhodopsin and the  $\beta$ 2-adrenergic receptor where it was shown that this interaction is critical in forming the active state [28, 29]. In bovine rhodopsin this Tyr points away from the TM core in the inactive state and moves towards the TM core in the active state (**Figure 70**) [28, 29]. On the other hand, in the inactive state of JSR1 or the  $\beta$ 2-adrenergic receptor this Tyr is already in the TM core and it only changes the rotamer in the active state (**Figure 70**) [24, 29, 30]. Additionally, due to the outward movement of TM6, Glu272<sup>6.30</sup> is no longer near Arg148<sup>3.50</sup>, breaking this ionic interaction.

Together, these conformational changes lead to the opening of a cytoplasmic cleft that allows the G protein to bind to the receptor. **Figure 65** shows the activation microswitches from the all-trans 6.11 retinal to the G protein binding site with a comparison to the inactive JSR1.



**Figure 57: Microswitches with experimental map. A:** TRP290 from the C-W-x-P motif shown in stick representation with the experimental map shown as mesh. **B:** P-I-F motif with side chains shown as sticks and the experimental map shown as mesh. **C:** D-R-Y motif with side chains shown as sticks and the experimental map shown as mesh.





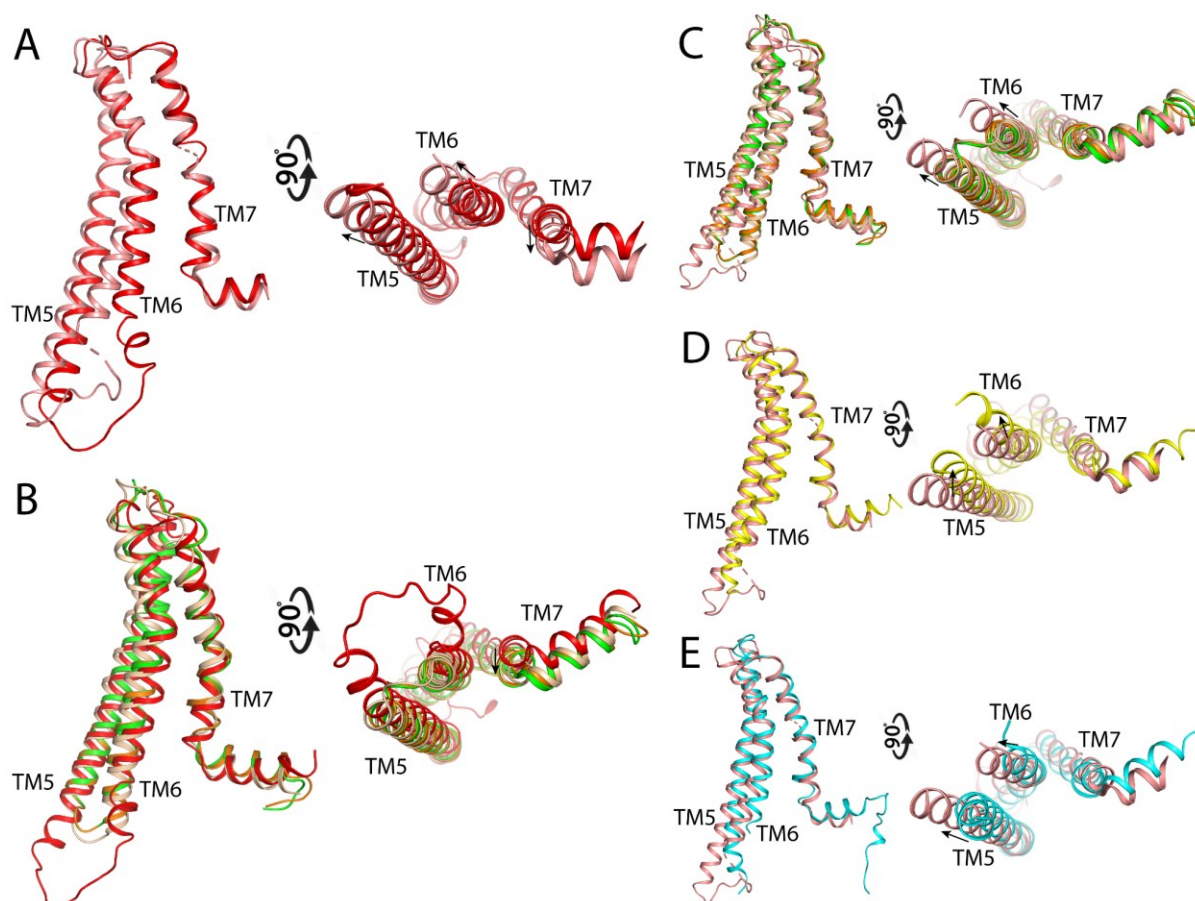
**Figure 58: Micro-switch domains and comparison of the active JSR1 (salmon) to the inactive JSR1 (619K, yellow), inactive bovine rhodopsin (1GZM, magenta) and active bovine rhodopsin (5EN0, green). Carbon atoms are shown in the respective color of the receptor, nitrogen atoms are blue and oxygen atoms red. A: Comparison of the inactive JSR1, showing the 9-cis**

retinal with Lys321<sup>7,43</sup> and the conserved Trp290<sup>6,48</sup> of the C-W-x-P motif as sticks, with the active JSR1 (salmon) showing the same residues as sticks. **B:** Comparison of the inactive bovine rhodopsin with the active metarhodopsin-II showing the retinal, Lys296<sup>7,43</sup> and conserved Trp265<sup>6,48</sup> as sticks. **C:** P-I-F motif (shown as sticks) comparison of the inactive and active states of JSR1. **D:** P-I-F motif (shown as sticks) comparison of the inactive and active states of bovine rhodopsin. **E:** D-R-Y motif (shown as sticks) comparison of the active and inactive state JSR1. Glu272<sup>6,30</sup> is also shown as stick representation. **F:** E-R-Y motif (shown as sticks) comparison of the inactive bovine rhodopsin with metarhodopsin-II. Glu247<sup>6,30</sup> is shown as sticks. **G:** Sequence conservation of the C-W-x-P motif in invertebrate rhodopsins (top) and vertebrate rhodopsins (bottom). **H:** Sequence conservation of the P-I-F motif in invertebrate rhodopsins (top) and vertebrate rhodopsins (bottom). **I:** Sequence conservation of the E/D-R-Y motif in invertebrate rhodopsins (top) and vertebrate rhodopsins (bottom).

#### 4.2.3.4 JSR1-G protein interactions

Opening of the cytoplasmic cleft enables binding of the G protein, which is a pre-requisite for catalysis of nucleotide exchange by the activated receptor. In this section, we will compare the G protein binding sites of the JSR1-hGi complex with the two conformations of the JSR1-JSGiq complex. Due to the lower resolution of the JSR1-hGi complex, we will not compare detailed side chain interactions but rather focus on secondary structure positions.

The biggest movements between inactive and active states of GPCRs are present in the relocation of the TM5/6 (**Figure 53C**) [18, 19]. It was shown before that the distance of relocation of TM6 is dependent on the subtype of G $\alpha$  protein [150, 151]. Indeed, we observe that the relocation of TM6 is larger in the JSGiq bound JSR1 compared to the hGi bound JSR1 (**Figure 59A; 60A**). The movement of TM6 in the JSR1-hGi complex is comparable to other GPCR-Gi/o structures or rhodopsin bound to the C-terminal peptide of transducin [117, 146-148] (**Figure 59B**). Similarly, the displacement of TM6 in the JSR1-JSGiq complex is slightly larger to that of the M1 receptor bound to hG11 but smaller than in the  $\beta$ 2-adrenergic receptor bound to hGs (**Figure 59D/E**) [24, 152].



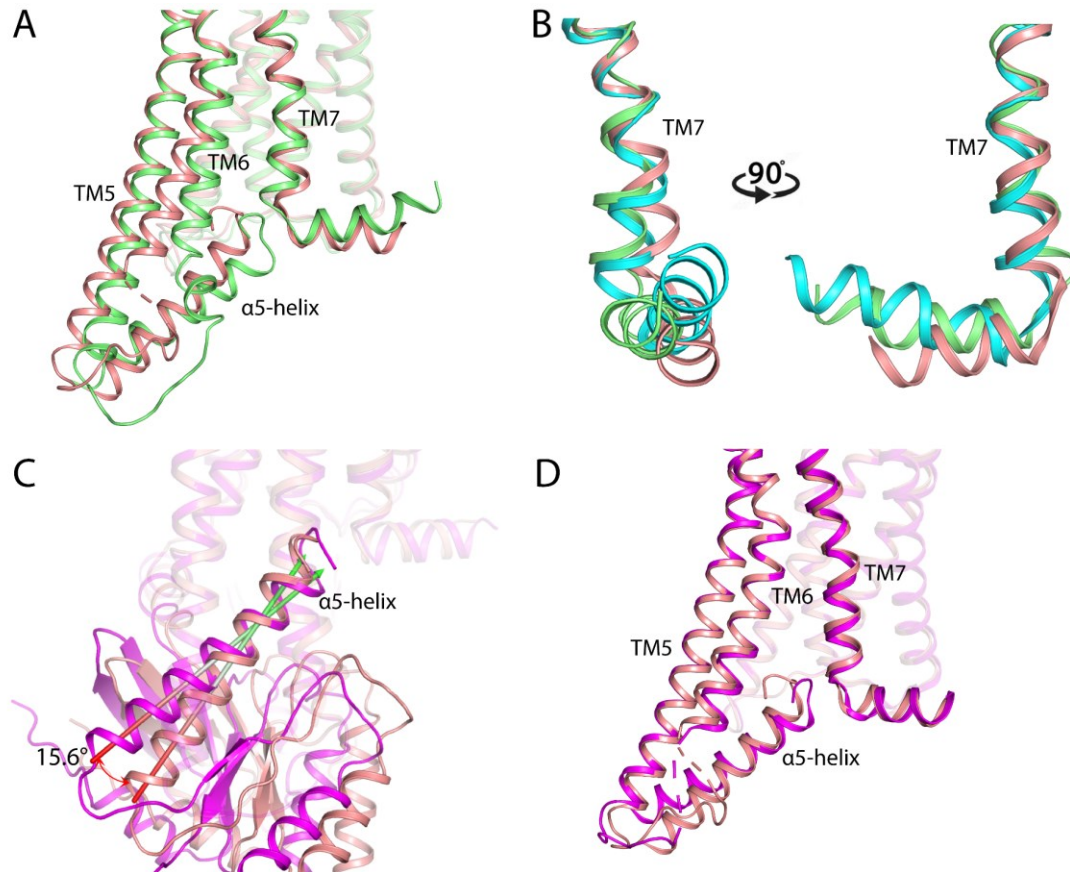
**Figure 59: TM5/6 and TM7 relocation comparison.** **A:** Overlay of TM5/6/7 of the JSR1-JSGiq complex (salmon) and the JSR1-hGi complex (red). **B:** Overlay of TM5/6/7 of the JSR1-hGi complex with bovine metarhodopsin II bound to G $\alpha$ T (PDB: 3PQR (orange); PDB: 4A4M (green)). **C:** Overlay of TM5/6/7 of the JSR1-JSGiq complex with different bovine metarhodopsin II structures (PDB: 3PQR (orange); PDB: 4A4M (green)) and the rhodopsin-hGi complex (PDB:6CMO (wheat)). **D:** Overlay of TM5/6/7 of the JSR1-JSGiq complex (salmon) with the  $\beta$ 2-adrenergic receptor-Gs complex (PDB: 3SN6, yellow). **E:** Overlay of TM5/6/7 of the JSR1-JSGiq complex (salmon) with the M1 receptor-G11 complex (PDB: 6OIJ, cyan).

Although the TM6 movement of the JSR1-hGi complex was similar to other GPCR-Gi/o/t complexes, we found that the conformation of TM7 is slightly different compared to other GPCR-G protein complexes. Interestingly, it is similar to the position of TM7 in the non-canonically activated NTS1R-Gi complex [153], indicating that JSR1 might bind the hGi differently compared to other GPCRs (**Figure 60B**). Additionally, the position of the  $\alpha$ 5 helix of the hG $\alpha$ i subunit is different compared to other GPCR-hGi complexes. The  $\alpha$ 5 helix adopts a shallower binding pose and the hook does not insert as deeply into the binding cleft compared to the JSR1-JSGiq complex (**Figure 60A**). Together, these observations indicate that JSR1 interacts with the hGi in a different manner compared to other GPCRs but still forms a functional signalling complex.

In contrast, our JSR1-JSGiq structure shows all the features of a fully active receptor. Usually, TM7 moves towards the transmembrane core and Tyr331<sup>7.53</sup> follows this movement, packing it against residues at position 3.44 and 3.47, leading to a hydrogen bond network between Tyr331<sup>7.53</sup> and Tyr234<sup>5.58</sup> that stabilizes the active state [26, 27]. TM7 relocates towards the transmembrane core and



the Tyr331<sup>7,53</sup> shifts into the core, packing against Leu141<sup>3,44</sup> and Ile144<sup>3,47</sup>, and is located above the conserved Arg148<sup>3,50</sup>. Arg148<sup>3,50</sup> hydrogen bonds to the backbone oxygen of Cys351 from the hook of the JS*G* $\alpha$ i $q$  which is a conserved interaction in GPCR-Gi complexes (see **Table 5** for detailed interactions).



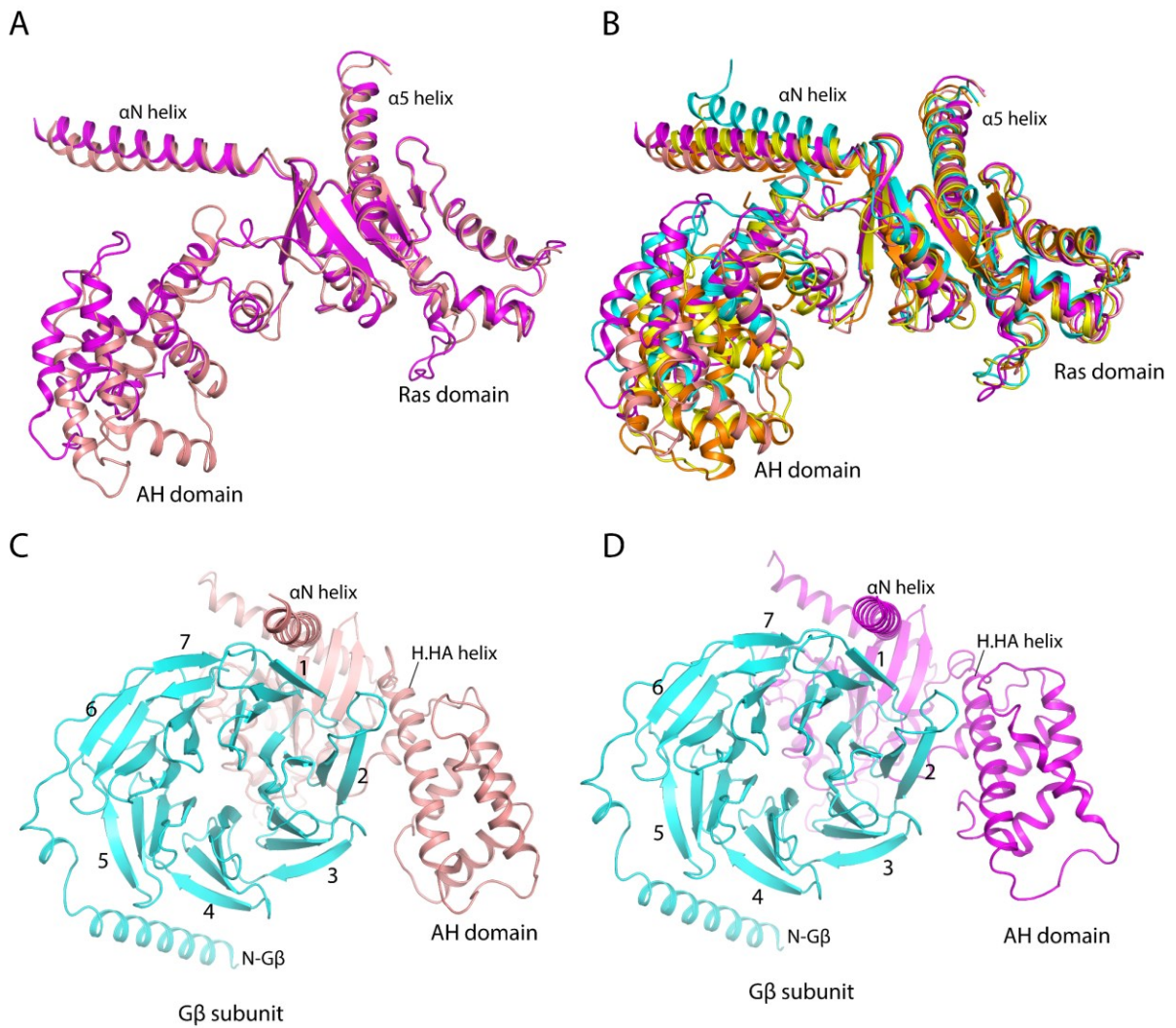
**Figure 60:** G protein binding site analysis of the JSR1-hGi complex and two conformations of the JSR1-JSGiq complex. **A:** G protein binding site comparison of the JSR1-hGi complex (lime) and the JSR1-JSGiq\_1 complex (salmon). From the  $G\alpha$  subunits only the  $\alpha 5$  helix is shown. **B:** Visualization and comparison of the cytoplasmic end of TM7 from the JSR-hGi (lime), JSR-JSGiq\_1 (salmon) and the non-canonical conformation of NTS1R (6OSA, cyan) shown from two different views. **C:** Visualization of the  $\alpha 5$  helix rotation between the two conformations of the JSR1-JSGiq complex. The JSR1-JSGiq\_1 complex is shown in salmon and the second conformation of the JSR1-JSGiq\_2 complex is shown in magenta. The main axis of the  $\alpha 5$  helices are shown as arrows. **D:** Comparison of the cytoplasmic ends of the two JSR1 conformations bound to JSGiq. Colors are the same as in figure 5D. From the  $G\alpha$  subunits only the  $\alpha 5$  helix is shown.

As mentioned before, during our analysis of the JSR1-JSGiq cryo-EM data we found that there are large movements primarily from the G protein relative to the receptor. 3D classification showed that a second conformation of the JSR1-JSGiq complex was present where the  $\alpha 5$  helix is rotated by 15.6° (**Figure 60C**). The whole G protein shows a similar degree of rotation to the  $\alpha 5$  helix. This rotation leads to a shallower binding pose of the  $\alpha 5$  helix of JS*G* $\alpha$ i $q$  where the G protein does not seem to be fully engaged to JSR1 (**Figure 60D**). Moreover, the cytoplasmic end of TM5 shifts slightly more outwards while the cytoplasmic end of TM6 moves further towards the transmembrane core to accommodate the rotated G protein (**Figure 60D**). The density for the hook of  $\alpha 5$  helix is less well resolved in the JSR-JSGiq\_2 map, indicating that this region is more flexible in this conformation. In



the JSR1-hGi complex, the  $\alpha 5$  helix adopts an even shallower binding pose than this alternate conformation of the JSR1-JSGiq complex, further indicating that JSR1 interacts with the hGi in a unique manner (**Figure 60A/D**). In this alternate conformation of the JSR1-JSGiq complex there are fewer interactions between JSR1 and the JSGiq compared to the fully engaged G protein conformation (17 interactions vs. 31 interactions) and nine interactions are conserved in both conformations (**Table 5**). Especially that the conserved interaction of Arg148<sup>3.50</sup> with the backbone oxygen of the Cys351 is missing, indicates that the JSR1-JSGiq<sub>2</sub> complex shows an intermediate state and not the fully engaged G protein state. Otherwise, most of the differences in the interactions are located at the cytoplasmic end of TM5 and ICL3. The density at the end of ICL3 and TM6 is not well resolved in both cryo-EM maps, indicating that this is a highly flexible region. From the density it might even be that the last helical turn of TM6 (inactive JSR1; PDB=6i9k) adopts a flexible structure because a helical structure usually shows strong electron density in the experimental cryo-EM maps. In the JSR1-JSGiq<sub>2</sub> complex two more residues in ICL3 had to be deleted because the density did not allow confident modelling which might further reduce the interactions seen in the model. Together these two conformations could visualize that JSR1 can interact differently with the JSGiq and still form a stable complex. In nature binding of the G protein is likely a dynamic process and our two structures present two snapshots of this process.

In both of the JSR1-JSGiq maps, density for the AHD could be observed, allowing modelling of the main chain but the resolution was not good enough to model side chains. Nevertheless, models with the AHD were created. Aligning both JSGiq structures showed that the AHD is at a similar position in both complexes, indicating that the AHD follows the rotation of the G protein to a similar degree. However, the AHD in the JSR1-JSGiq<sub>1</sub> complex is slightly more closed than in the JSR1-JSGiq<sub>2</sub> complex (**Figure 61A**). In most GPCR-G protein complexes the density for the AHD is not well resolved which does not allow modelling of the domain. There are three GPCR-hGi complex structures where the AHD was modelled without addition of an antibody fragment that stabilizes the AHD [154-156]. Comparison with these structures showed that the AHD is at similar overall positions like we see in our JSR1-JSGiq structures but the exact position of the helices differs a bit. In both of our JSR1-JSGiq structures the H.HA helix from the AHD is in proximity to the side of blade 2 and a loop in blade 3 of the G $\beta$  subunit (**Figure 61C/D**). These regions might interact with each other to stabilize the position of the AHD which allows reconstruction of the density for the AHD.



**Figure 61:** **A:** Structure comparison of the jsGaiq subunit from the JSR1-JSGiq\_1 (salmon) and JSR1-JSGiq\_2 (magenta) complexes. The  $\alpha$ -helical domain (AH domain),  $\alpha$ N helix,  $\alpha$ 5 helix and the Ras domain are labelled. **B:** Structure comparison of the G $\alpha$  subunit from the JSR1-JSGiq\_1 (salmon), JSR1-JSGiq\_2 (magenta), GABA(B) receptor-hGi (PDB: 7EB2 [154]; cyan), NTSR1-hGi (PDB: 7LOS [155]; orange) and the cannabinoid receptor 2-hGi (PDB: 6PT0 [156]; yellow) complex structures. The  $\alpha$ -helical domain (AH domain),  $\alpha$ N helix,  $\alpha$ 5 helix and the Ras domain are labelled. **C:** Structure of the JSGaiq and G $\beta$  subunits from the JSR1-JSGiq\_1 complex. The G $\alpha$  subunit is colored in salmon and the G $\beta$  subunit in cyan. The AH domain,  $\alpha$ N helix and the H.HA helix of the JSGaiq are labelled. The numbers of the G $\beta$  subunit represent the numbering of the  $\beta$  blades according to Wall et al. [157]. **D:** Structure of the JSGaiq and G $\beta$  subunits from the JSR1-JSGiq\_2 complex. The G $\alpha$  subunit is colored in magenta and the G $\beta$  subunit in cyan. The AH domain,  $\alpha$ N helix and the H.HA helix of the JSGaiq are labelled. The numbers of the G $\beta$  subunit represent the numbering of the  $\beta$  blades according to Wall et al. [157].

<b>Residue</b>	R148 (3.50)	V151 (3.53)	I152 (3.54)	M156 (ICL2)	A157 (ICL2)	P160 (ICL2)	I237 (5.61)	H244 (5.68)	L248 (5.72)	Q251 (5.75)	K254 (5.78)	M255 (5.79)	I257 (ICL3)	K258 (ICL3)	L260 (ICL3)	R261 (ICL3)	S262 (ICL3)	N263 (ICL3)	R274 (6.32)	L275 (6.33)	V278 (6.36)	S334 (7.56)	H335 (8.47)	P336 (8.48)	K337 (8.49)	
R31 (G.hns1.02)				Blue	Green																					
I194 (G.S3.01)				Orange																						
E298 (G.H4.05)													Orange	Orange												
A301 (G.H4.08)													Orange													
E308 (G.H4.16)																	Green									
E318 (G.h4s6.12)																Green		Green								
I319 (G.h4s6.13)																	Green									
H322 (G.S6.04)												Green														
C325 (G.s6h5.02)												Green														
Q333 (G.H5.5)										Green	Green															
F334 (G.H5.6)														Green	Green											
C337 (G.H5.9)								Orange		Green																
K340 (G.H5.12)								Blue																		
D341 (G.H5.13)								Blue									Green									
I343 (G.H5.15)				Orange																						
N347 (G.H5.19)			Green																							
L348 (G.H5.20)			Green																							
K349 (G.H5.21)																				Orange					Green	
E350 (G.H5.22)																									Orange	Blue
C351 (G.H5.23)	Green	Green																								
N352 (G.H5.24)	Blue																						Blue	Blue	Green	
L353 (G.H5.25)	Green							Green																		
V354 (G.H5.26)																				Blue			Green			

Table 5: JSR1-JSGiG complexes interaction table. Top row shows the JSR residue number and the B.W. numbering. Left column shows the JSGiG residue number and hGi protein numbering. The distance cut-off for interactions was 4 Å. Green fields indicate interactions present in the JSR1-JSGiG\_1, orange fields interactions present in the JSR1-JSGiG\_2 and blue fields indicate interactions found in both structures.

#### 4.2.3.5 Summary JSR1-JSGiq structure analysis

- The overall structure of the JSR-JSGiq complex shows the typical architecture of GPCR-G protein complexes.
- The densities and models for the all-trans and ATR611 retinal overlay well and are similar to the all-trans retinal from different bovine metarhodopsin II models.
- The orientation of the PSB and a density between the CI and PSB indicate that the PSB-CI link is mediated by water molecules.
- The water mediated PSB-CI link does not involve Ser199 in the active state JSR1 but Tyr293 is likely involved.
- The microswitch domains show the typical conformations seen in other active state GPCRs.
- The G protein binding pose of the JSR1-hGi is different to the JSR-JSGiq and other GPCR-hGi complex models.
- The second conformation of the JSR1-JSGiq complex shows a rotated G protein leading to a shallower binding pose of the  $\alpha 5$  helix compared to the JSR1-JSGiq\_1 complex.

## 5 Discussion

### 5.1 Adhesion GPCR project

The adhesion GPCR field is still relatively young compared to the GPCR field overall. In the last ~10-15 years adhesion GPCRs gained a lot of attention because it was found that these receptors are involved in diverse and important physiological processes. Many studies focused on *in vivo* experiments with model organisms, like mice or zebrafish, or cell signaling assays. *In vitro* assays were focused on the extracellular domains of adhesion GPCRs and their (potential) interaction partners. However, at the start of this project there was no literature about *in vitro* experiments with the 7TM domain of adhesion GPCRs. This could have several reasons. First, the whole adhesion GPCR family lacks high affinity ligands that could facilitate purification. Second, the ligand free state of aGPCRs is likely highly dynamic, making it difficult to study these receptors *in vitro* because of a low stability when extracted from the cell membrane. Lastly, purification from a native source (like it was done for bovine rhodopsin for example [158]) which could facilitate purification is not suitable because of a lack of a native source.

In this project we aimed to recombinantly express aGPCRs in order to characterize them biochemically and structurally. Success in this direction was achieved which allowed the preparation of a sample that could be analyzed by single particle cryo-EM.

#### 5.1.1 AGPCR-G protein complex

##### 5.1.1.1 Biochemistry

The initial constructs were designed for HEK cell expression because we are working with human proteins and our collaborators, Prof. Adriano Aguzzi and his group, showed already that Adgrg6 can be expressed in HEK cells for signaling assays [77]. For our purpose we tried to overexpress the receptors in order to get a high yield. Before I started working on this project, Dr. Gabriella Collu already conducted a first expression test of the full length Adgrg6 in HEK cells. The overexpression of the full length Adgrg6 showed signs of cytotoxicity, likely attributed to the cAMP accumulation from the G<sub>s</sub> protein activation by Adgrg6. For this reason, most of our constructs for HEK cell expression were designed to be inactive by truncating or mutating the Stachel sequence. One Adgrg6 (G6.01) where the GAIN domain is included did not have an altered Stachel sequence but in theory the GAIN domain should protect the Stachel sequence and therefore activation of the receptor. Two Adgrd1 constructs were also potentially active (D1.05/.06) as these were full length receptor isoforms. In our case we did not observe a significantly reduced cell viability with these constructs after 72h expression (**Table 6**). However, from the in-gel fluorescence imaging of SDS-PAGE and FSEC experiments we can

not conclude that the receptors were correctly expressed. The in-gel fluorescence scanning showed a band with the same MW for all the constructs (**Figure 11B**). For Adgrg6 constructs faint bands with different MW were detected (**Figure 11B**). We still do not know what protein showed this fluorescence

Sample	Cell count (cells/mL)	Viability (%)
D1.01	5.41x10 <sup>6</sup>	83.2
D1.02	5.53x10 <sup>6</sup>	77.5
D1.03	4.95x10 <sup>6</sup>	80.6
D1.04	6.16x10 <sup>6</sup>	75.9
D1.05	4.67x10 <sup>6</sup>	83.0
D1.06	4.49x10 <sup>6</sup>	82.6
G6.01	5.61x10 <sup>6</sup>	86.8
G6.02	4.10x10 <sup>6</sup>	83.2
G6.03	6.03x10 <sup>6</sup>	84.3
G6.04	4.77x10 <sup>6</sup>	86.9

**Table 6: HEK cell expression tests for Adgrd1 (D1.XX) and Adgrg6 (G6.XX) constructs.** First column shows the construct number. Second column shows the cell count after 72h expression. Third column shows the cell viability after 72h expression.

signal. The MW is too high that it could be YFP only and attempts to purify the receptor from HEK cell expressions were not successful. Changing the expression host to insect cells was the right decision in hindsight.

Our insect cell constructs contained both active (or potentially active) receptor constructs (G6.05/.06; D1.08) and inactive constructs (G6.07-.09; D1.07/.09-.11) that can be activated by adding chemically synthesized Stachel peptide. For all of these constructs expression could clearly be detected, showing that insect cells are a good expression host for GPCRs, as it was demonstrated before for class B1 GPCRs (closest family to the adhesion GPCRs) [136, 137]. The in-gel fluorescence imaging and FSEC analysis also showed that Adgrg6 is likely less stable than Adgrd1 because in SDS-PAGE

fluorescence signal was detected in the stacking gel, meaning that there are aggregates that can not be dissolved by the SDS buffer. Additionally, the FSEC data showed large peaks in the void volume, indicating aggregation. The same was observed for the Adgrd1 constructs but generally, Adgrd1 constructs showed less signal for the void peak (**Figure 13**).

Further evidence that Adgrg6 is less stable than Adgrd1 was seen in the detergent screening. In most conditions tested the Adgrg6 constructs showed worse ratios (void peak signal to receptor peak signal) compared to Adgrd1 constructs (**Table 3**). Generally, solubilization in a mild detergent like LMNG or LMNG/CHS showed better results as the other detergents tested. However, in almost all the conditions tested, a higher signal was measured for the void peak, indicating that both Adgrg6 and Adgrd1 are unstable in detergent micelles, at least in the apo-state. Nevertheless, the Adgrd1 results showed promise that purification can be achieved by optimizing conditions and addition of a stabilizing interaction partner, in this case a G protein.

Purification trials for the receptors alone were not successful. SDS-PAGE analysis either showed no band for the receptor or only faint bands, but the FSEC data showed that the samples contained aggregated protein because the peak was eluted in the void volume (**Figure 17/18**). The purification trials using TALON resin also showed that there is free his-tagged GFP that binds to the column and the FSEC data showed one main peak for GFP in the elution sample (**Figure 17**). Together, the purification trials showed that the receptors need an interaction partner for stabilization in order to purify the receptors.

A surprising result was seen in the complex formation experiment. In our hands, all the constructs tested needed chemically synthesized Stachel peptide in the buffer in order to activate the receptor and induce binding of the mini-G $\alpha$ s (**Figure 20**). This was especially surprising for the constructs D1.12 and G6.10 that express the CTF of the respective receptor. It was shown before that expression of the CTF leads to an active receptor that shows G protein signaling activity in cell assays [47]. This discrepancy could arise because we did not test this directly with co-expression but expressed the receptor and G protein in different cells that were mixed and solubilized together. The reason is that the expression levels of all the components are higher when expressed separately, resulting in technical advantages like having smaller cell culture volumes and having more protein, meaning more signal that can be detected in FSEC experiments. Furthermore, we included a HA signal sequence that was fused to the N-terminus which might affect the ability of the Stachel peptide to bind and activate the receptor. From the aGPCR-G protein complex structures published, we saw that having a methionine before the Stachel sequence is important for the activation [67]. Regardless, our experiment showed clearly that we can activate the receptors by adding their respective Stachel peptide, leading to complex formation with the G protein.

Co-expression of the receptor with the G protein was only successful for D1.12 (Adgrd1-CTF). D1.12 already showed the highest expression levels when expressed alone (**Figure 19B**) and was the only construct where clear expression was detected when co-expressed with the mini-G $\alpha$ s heterotrimer (**Figure 19C**). The equivalent construct for Adgrg6 (G6.10; Adgrg6-CTF) showed much lower expression levels when expressed alone (**Figure 19B**) and showed no clear bands in the co-expression with the mini-G $\alpha$ s heterotrimer. This shows that there are big differences in terms of expression and biochemical properties between Adgrg6 and Adgrd1, indicating that Adgrd1 is a good target for biochemical and structural studies. Further support that Adgrd1 is a good target for structural studies is that the structure was eventually solved by two different groups [68, 70] whereas we still lack a structure of the Adgrg6-CTF.

The initial purification of the Adgrd1<sup>CTF</sup>-mini-G $\alpha$ s complex worked as all components were detected on SDS-PAGE but the SEC profile did not look good because a large peak in the void volume was detected



and no clear peak for the complex was seen. In total three more purifications were done with variations in the solubilization step (0.5% vs 1% LMNG/CHS) and buffer conditions. Changes in the buffer conditions included: Addition of glycerol in the solubilization, wash and elution buffer of the FLAG column and addition of glyco-diosgenin (GDN) to the buffers. Although these changes resulted in a better SEC profile (**Figure 22A/C**), the yield was significantly lower than in the first purification. This could also be because the expression was not as good as the first time which is based on the intensity of the color from the dsRed protein that is used as an infection marker. For making cryo-EM grids only very little amounts of protein is needed, meaning that several grids could be made from the purifications, but the sample could not be concentrated a lot to increase the particle density. Grids from all the purifications were analyzed by cryo-EM.

Taken together, these experiments showed that aGPCR are highly dynamic, making it difficult to work with these receptors *in vitro*. Without an interaction partner, the receptors are too unstable in detergents to be able to purify these receptors. Especially the lack of high affinity ligands hinders the *in vitro* work with these receptors. The only successful strategy to purify the receptors was to co-express them with the G protein, allowing complex formation in the cell membrane and subsequent extraction and purification of the whole complex for cryo-EM studies.

#### 5.1.1.2 Cryo-EM analysis of the *Adgrd1<sup>CTF</sup>-mini-Gs* complex

Initially, grids were screened manually either at a JEOL JEM 2200FS microscope (with help of Dr. Ching-Ju Tsai) or at a Titan Krios microscope (with help from either Dr. Jacopo Marino or Dr. Pavel Afanasyev). During screening, grids from two purifications did not show any particles in the thin ice layers. In both cases the SEC profile showed a signal and in gel fluorescence imaging and western blot also showed that the components are present, but the yield was very low. For this reason, the samples could only be concentrated to less than 1 mg/ml. Even with this low concentration we would have expected to see particles in the thin ice layers. It could be that either the time between SEC and making the grids was too long, leading to precipitation of the protein, or that concentrating the protein led to dissociation and/or precipitation of the complex.

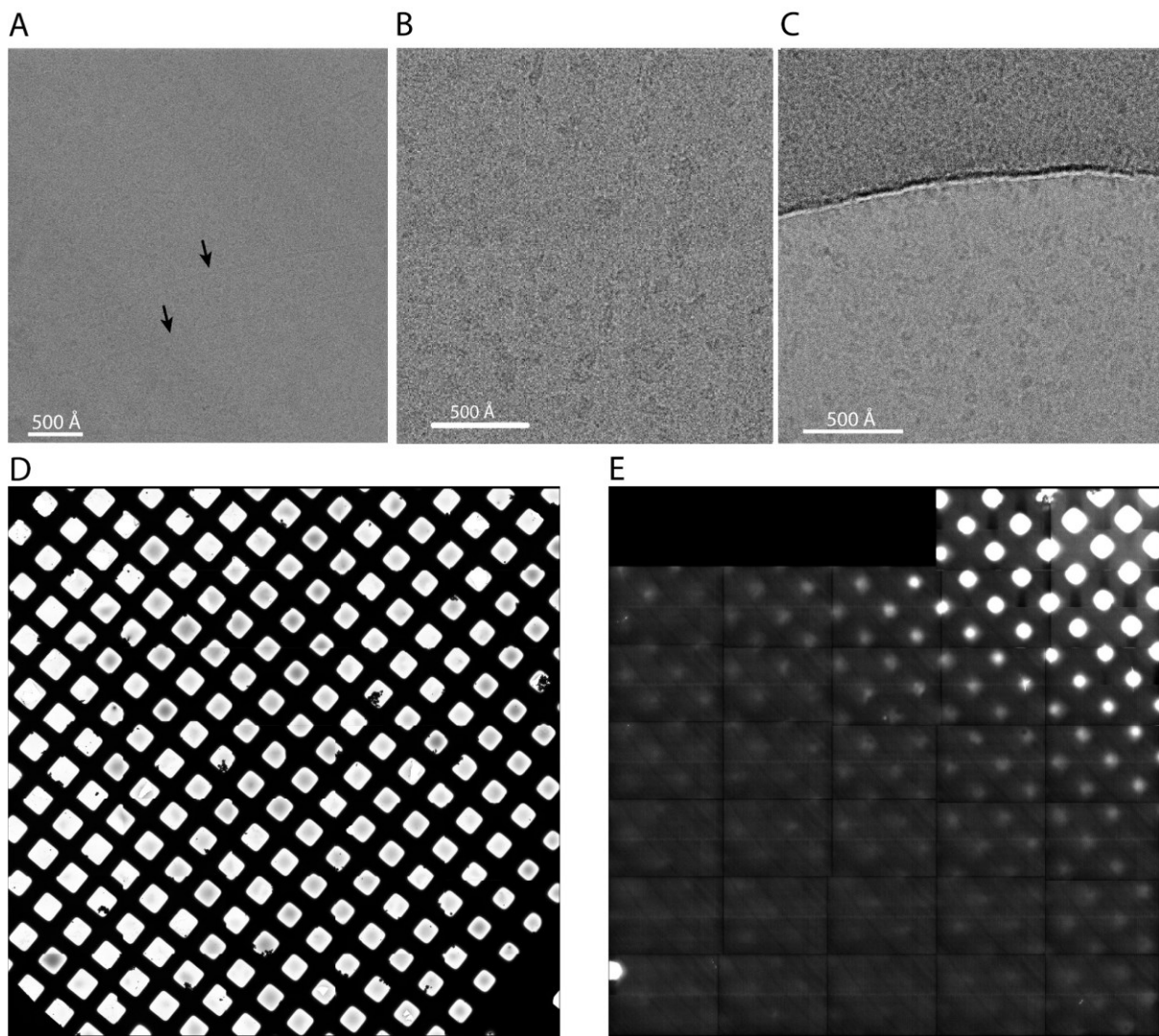
The grids prepared from the first (**Figure 22A/B**; called sample A from here) and fourth (**Figure 22C/D**; called sample B from here) purification both showed particles in the thin ice layers, and we collected a large dataset for both samples. Sample B was concentrated after SEC and sample A was not. In the micrographs from sample B, long fibril like structures were seen that could come from the detergent LMNG. Because sample B was concentrated it is possible that the detergent concentration was also increased, leading to the formation of fibril-like structures of the detergent (**Figure 62A**, black arrows). This indicates that a lower detergent concentration could be used, or the detergent concentration has

to be diluted during the protein concentration process. Grids from sample A generally had thicker ice which can be seen from the quality of the micrograph (**Figure 62B**) compared to a micrograph from sample B where the ice is thinner, and particles can easily be spotted (**Figure 62C**). During the grid preparation and optimization process we noticed that short blotting times (1s or 2s) result in very thick ice layers that can not be used for data collection (**Figure 62E**) and longer blotting times (4s or 5s) give a good ice thickness for data collection (**Figure 62D**).

Even though the SEC and ice thickness of sample A did not look very promising, the cryo-EM data was much better than from sample B. In sample B, a lot of particles were dissociated, as seen in the 2D classes (**Figure 24B**) and complex particles showed less high frequency information (**Figure 24C**). The main difference between sample A and B (besides the buffer composition) is that sample A was directly used for grid preparation whereas sample B had to be concentrated, also increasing the time between the SEC step and grid preparation. It is possible that the complex is not very stable and the concentration step resulted in many dissociated complex particles. Furthermore, we used longer blotting times for making the grids with sample B, which gives us thinner ice layers but the proteins are also more likely to be at the air-water interface which can lead to protein denaturation and complex dissociation [143, 144]. In hindsight, adding additives like n-Octyl- $\beta$ -D-glucoside ( $\beta$ -OG) could have helped making better grids (personal communication with Prof. Skiniotis) and should have been tried.

Unfortunately, neither of the two datasets collected was good enough to solve a high-resolution structure of the complex.

The 2D classes from sample B did not clearly show all of the subunits of the complex and many of the 2D classes were blurry (**Figure 24C**). This translated to the 3D volumes that were created based on these particles. The 3D volumes did not allow identification of all the subunits or did not even show density for the G protein clearly (**Figure 24D**). Furthermore, the particles mostly showed side views and no top views which is needed for a good 3D reconstruction (**Figure 24C**).



**Figure 62: Cryo-EM analysis of the Adgrd1<sup>CTF</sup>-mini-Gs complex.** **A:** Micrograph from sample B taken at a nominal magnification of 40'000x. Black arrows indicate fibril like structures. **B:** Example of a micrograph from sample A, collected at a nominal magnification of 165'000x. **C:** Example of a micrograph from sample B, collected at a nominal magnification of 165'000x. **D:** Atlas of a grid blotted with blot force 10 and blot time of 5s, collected at a nominal magnification of 175x. **E:** Atlas of a grid blotted with blot force 10 and blot time 2s, collected at a nominal magnification of 175x.

The cryo-EM data from sample A, which was collected first, showed more promise as the 2D classes clearly showed signals for all the subunits of the complex (**Figure 23B/C**). The 2D classes also showed high-frequency information and secondary structure details, like the  $\alpha$ N or  $\alpha$ 5 helix, could be recognized from the 2D class averages (**Figure 23B/C**). However, attempts to create a high-resolution 3D volume were not successful. At a low contour level all subunits of the complex, micelle with receptor, G $\alpha$  and G $\beta\gamma$ , could be identified but at a higher contour level it became clear that the transmembrane helices of the receptor are not well resolved (**Figure 23D**). This can have several reasons: 1. The number of complex particles was relatively low with less than 100'000 particles present in the best 2D classes; 2. The particle set likely included heterogeneity, lowering the signal to noise ratio; 3. The ice thickness was relatively thick, further lowering the signal to noise ratio. These points, together with the biochemistry of sample A, lead us to believe that we needed a better

sample and grids for high-resolution structure determination. In the end, it turned out that the samples from subsequent purification resulted in worse cryo-EM data, even though the biochemistry showed improvements. Retrospectively speaking, it might have been a good idea to collect more data from sample A and with sufficient particles it might have been possible to solve a high-resolution cryo-EM map of the complex.

To conclude, the cryo-EM data from sample A looked very promising but the changes implemented for subsequent purifications did not result in a better sample for cryo-EM. Sample purification and grid preparation still needed improvements for high-resolution structure determination but because of the four nature papers that published structures of aGPCRs [67-70], showing the tethered peptide activation mechanism, this project was abandoned in favor of the jumping spider rhodopsin-1 project.

## 5.1.2 Adgrd1-NTF

### 5.1.2.1 Biochemistry and crystallization

The Adgrd1-NTF was expressed as a secreted protein in insect cells. Expression of the protein worked from the first try and pull down assays already showed promising results for purification (**Figure 25**). The best expression conditions were achieved with High Five cells, a VOI of 3% and an expression time of 48h (Yield: 2 mg/L). The purification protocol could be kept simple with one affinity chromatography step and a SEC step, yielding pure protein for subsequent assays (**Figure 26**).

The thermal stability assay showed that the Adgrd1-NTF has a high  $T_m$  with  $\sim 60^\circ\text{C}$  (in SEC buffer). It also showed that freezing and thawing the protein has a negative influence on the protein. Because the melting curve split into two melting curves where two  $T_m$  can be calculated, it is likely that one of the domains of the Adgrd1-NTF has a lower  $T_m$  after a freeze/thaw cycle while the rest of the protein is not affected by this. This result meant that it was the best to use the protein directly after purification. Luckily, the protein is stable enough that it can be stored at  $4^\circ\text{C}$  for at least 4 days without precipitation or changes of the  $T_m$ , further indicating that the protein is very stable. Salt (mostly NaCl) is generally used in protein buffers to increase solubility and disrupt unspecific interactions. The NaCl concentration had the biggest influence on the  $T_m$  of the Adgrd1-NTF with around  $\sim 15^\circ\text{C}$  difference between the sample with no NaCl and the sample with 300 mM NaCl (**Figure 28B/D**).

The crystallization trials were not very successful because only in one condition clear crystal growth could be detected (**Figure 29A/B**). Many of the imaged drops were clear (even with the higher concentrated sample), indicating that the protein concentration can be increased even further. Ultimately, the crystallization trials were abandoned in favor of structure determination of the Adgrd1-G protein complex and the jumping spider rhodopsin-G protein complex.

### 5.1.3 Adgrd1 fusion proteins

#### 5.1.3.1 Adgrd1<sup>CTF</sup>-mini-Gα<sub>s</sub> fusion protein

From the purification trials of Adgrd1 and Adgrg6 alone, it became clear that the receptors need an interaction partner for stabilization in order to purify the receptor. As an alternative strategy to the co-expression with the G proteins, we designed a receptor-G protein fusion protein. In the complex formation experiment we showed that Adgrd1 couples to mini-Gα<sub>s</sub>, so we decided to create a fusion protein with it (**Figure 30**). Expression and purification of the Adgrd1<sup>CTF</sup>-mini-Gα<sub>s</sub> fusion was working from the first experiment, showing that this approach is a valuable alternative to the co-expression approach. Additionally, the protein yield after the purification was largely increased by ~7.5-fold from 200 μg/L (co-expression strategy) to 1.5 mg per Liter of cell culture. This could have several reasons. First, in this approach only one baculovirus is added to infect the insect cells whereas in the co-expression approach several baculoviruses are added, meaning the insect cells have to produce several different proteins, lowering the yield for each individual protein. Second, in the fusion protein approach the ratio of receptor molecules to mini-Gα<sub>s</sub> proteins is 1:1 and the G protein is in proximity to the receptor. This allows every activated receptor to couple to the fused mini-Gα<sub>s</sub>, thereby stabilizing the receptor. On the other hand, the ratio of receptor molecules to G protein molecules in the co-expression approach is likely not 1:1 and the G proteins might not be in proximity to the receptors. This could result in activated receptors that are not coupling to a G protein and are likely to precipitate/aggregate during the purification.

Although the fusion protein can be purified with good yields, coupling the fusion protein to the Gβγ subunits did not show efficient coupling (**Figure 31E**). For cryo-EM studies having a bigger complex is usually beneficial but with today's technological advances the mini-Gα<sub>s</sub> fused to the receptor may be enough to align particles well [159]. Because the structure of Adgrd1 was solved by two groups [68, 70], we did not continue analyzing this sample by cryo-EM. Nevertheless, this work presents a proof-of-concept for the fusion protein approach and shows that a receptor, that can not be purified separately, can be purified with this strategy.

#### 5.1.3.2 Adgrd1<sup>GAIN-CTF</sup>-mini-Gα<sub>s</sub> fusion protein

The other fusion proteins we designed included the GAIN domain of Adgrd1 in addition to the CTF. Additionally, cysteine mutations were introduced that, based on MD simulations, could help immobilize the GAIN domain for structure determination. Immobilization is likely necessary because Barros-Alvarez *et al.* [67] showed that the GAIN domain is very flexible, hindering structure determination. The results showed that the constructs and/or the mutations introduced had a negative effect. Already expression was not clearly seen for all of the tested constructs compared to

the Adgrd1<sup>CTF</sup>-mini-Gα<sub>s</sub> fusion protein (**Figure 31/32**). Although protein was eluted from the FLAG affinity column, the SEC chromatogram showed only a peak in the void volume (**Figure 33**). This indicates that the protein is either aggregated or the cysteine mutations we introduced might lead to the formation of intermolecular disulfide bridges. The second theory is less likely because addition of 100 mM DTT did not change the SEC profile and this concentration of DTT should cleave potential disulfide bridges.

Further evidence that the proteins are aggregated came from the negative staining grids analysis. In micrographs large aggregates were seen in samples with both a high and a low protein concentration, indicating that the protein is not healthy (**Figure 34**). These GAIN-CTF-G protein fusion proteins would need further optimization and/or a redesign of the constructs but because of time reasons, this was not possible anymore in my PhD work.

To summarize, the Adgrd1<sup>CTF</sup>-mini-Gα<sub>s</sub> fusion protein was well expressed and could be purified which shows that this approach can be useful for other receptors as well. On the other hand, the Adgrd1<sup>GAIN-CTF</sup>-mini Gα<sub>s</sub> fusion proteins did not express well and could not be purified. However, for these fusion proteins it is difficult to draw any conclusions because not much work was done to try to optimize expression and purification. In the aGPCR field, the interplay of the GAIN domain with the CTF is still an unsolved problem. Functional and structural characterization of the interactions between the GAIN and CTF domain will remain one of the most interesting topics of aGPCR research.

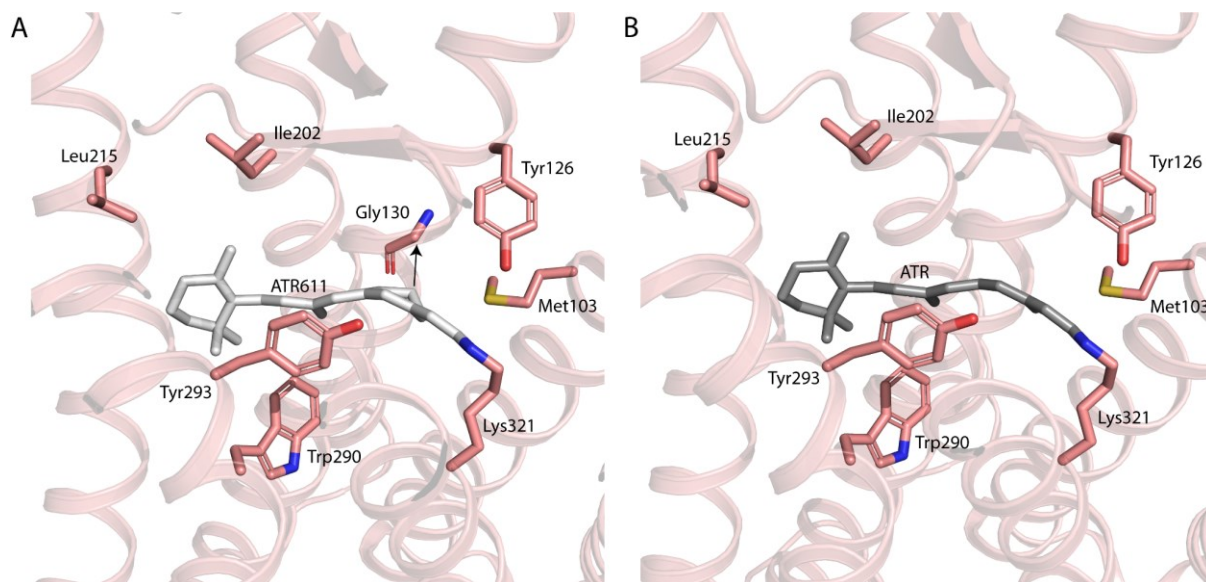


## 5.2 Jumping spider rhodopsin-1 project

### 5.2.1 JSR1-G protein complexes

#### 5.2.1.1 Purifications

To be able to do *in vitro* studies with any protein, one first needs to purify the protein. Before I started working on this project, members of our group (Dr. Niranjana Varma; Dr. Elena Lesca and Dr. Matthew Rodrigues) already established an expression and purification protocol for JSR1. Dr. Niranjana Varma established a purification protocol for JSR1 bound to the inverse agonist 9-cis retinal [30] and Dr. Matthew Rodrigues followed a similar protocol for the purification of JSR1 bound to the non-natural all-trans retinal 6.11 analogue. The purification of JSR1 (ATR611) worked well and pure protein could be obtained from a purification that only involved one 1D4 affinity column (**Figure 35B**). However, the analytical SEC experiments showed that there is aggregation which varies depending on the purification (**Figure 42**). Therefore, the purification of JSR1 might benefit from an additional SEC step to get rid of any aggregated proteins in the 1D4 elution. The absorbance at 505 nm for detecting the retinal is very low in the void peak (**Figure 42**), meaning that the proteins that aggregate are either aggregated JSR1 that can not take up the retinal or other proteins than JSR1. It was surprising that the purification of JSR1 bound to the all-trans retinal 6.11 worked well because previous attempts by several members of our group to purify JSR1 bound to the all-trans retinal were not successful (**Figure 35B**). One hypothesis we had is that the additional carbon atoms from the carbon ring around the double bond at position 11 make van-der-Waals interactions with surrounding residues, leading



**Figure 63: Retinal binding site analysis with active JSR1 and either ATR611 or all-trans retinal.** **A:** Retinal binding site of the JSR1-JSGiq complex model. ATR611 is shown in light grey and JSR1 in salmon. Residues within 4 Å of the ATR611 are shown as sticks and are labelled. The additional interaction of the ATR611 is indicated with an arrow. **B:** Retinal binding site of the JSR1-JSGiq complex model with the all-trans retinal. All-trans retinal is shown in grey and JSR1 in salmon. Residues within 4 Å of the ATR are shown as sticks and are labelled.



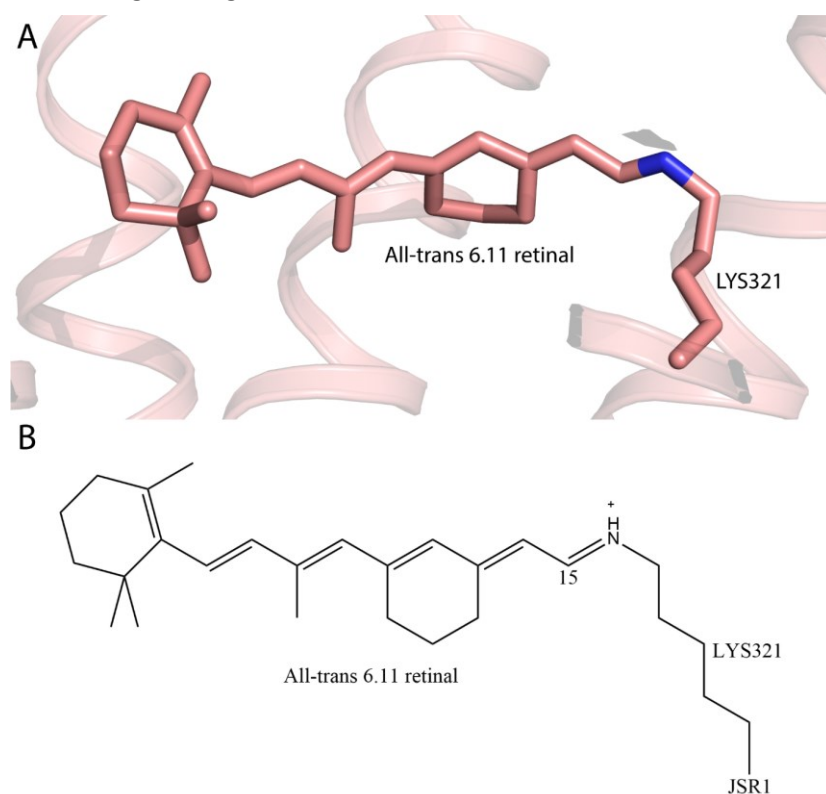
to a higher affinity of the ATR611 compared to the all-trans retinal. Because the resolution of the JSR1-hGi complex is not good enough to see side chain densities, the JSR1-JSGiq\_1 complex model was used for this analysis and for the all-trans retinal the additional carbon atoms from the ring of the ATR611 were deleted (**Figure 63**). The analysis showed that only one additional residue is present within 4 Å of the ATR611 compared to the all-trans retinal (**Figure 63**). This residue is Gly130 which is close to the carbon atom of the ring introduced for the ATR611 (**Figure 63A**). This additional interaction might increase the affinity enough that ATR611 can be incorporated to JSR1.

Expression and purification worked well for all G proteins, but the yields varied greatly. The yields for the purification of hG $\alpha$ i and G $\beta$  $\gamma$  from bovine retinas can not be directly compared to the yields of the hGq, hG11, JSGq1 and JSGiq because the expression method was very different. The hG $\alpha$ i was expressed as a soluble protein in *E.coli* which resulted in a high yield of ~15 mg/L (**Figure 36**). G $\beta$  $\gamma$  was purified from bovine retinas and yielded ~8.7 mg of protein from 450 retinas. When the G $\beta$  $\gamma$  subunits were expressed in insect cells and purified the yield was relatively low with ~100  $\mu$ g/g of cell pellet. Nevertheless, this protocol gave pure protein in the end (**Figure 37**) and therefore it is a valuable alternative to the purification of G $\beta$  $\gamma$  from bovine retinas. This protocol was only followed once and there is probably room for improving the expression and purification conditions in order to achieve a higher yield. Expression and purification of the hGq, hG11, JSGq1 and JSGiq heterotrimers were done with the same protocol and worked in all cases (**Figure 38**). The yields for the hGq, hG11 and JSGq1 were similar with ~500-750  $\mu$ g of purified protein per liter of cell culture but the yield for the JSGiq chimera was much higher with ~2.5 mg/L. This is not surprising based on previous experiences in our lab. The hGi heterotrimer is a well expressing protein with high thermal stability, resulting in high yields. The basis of the JSGiq chimera is the hGi, explaining why the yield is much higher compared to the hGq, hG11 and JSGq proteins. The hGq expression protocol is also special because the chaperon Ric8 is needed for efficient expression (Communication with Jonas Mühle) [160]. Ric8 was not added to the expression of the JSGq1 which might have lowered the expression yield. It was also the first time we worked with an invertebrate G protein, meaning that our protocols are not optimized for invertebrate G proteins but for human G proteins. The JSGq1 might need different conditions in order to achieve a higher yield but this will be subject of future research. Furthermore, the SDS-PAGE analysis of the JSGq1 showed an additional, well visible band between the JSG $\alpha$ q1 and the G $\beta$  that was not seen in the SDS-PAGE of the other G proteins (**Figure 38**). This could be a differentially modified G $\alpha$  or G $\beta$  subunit but further investigation is needed to uncover what protein this band corresponds to.

Overall, all purifications of the different G proteins yielded enough protein to conduct several GTPase Glo assays, analytical SEC experiments and preparation of cryo-EM samples.

### 5.2.1.2 GTPase Glo assay

The GTPase Glo assay was performed for JSR1 bound to one of three ligands (9-cis, 9-cis 6.11 or all-trans retinal 6.11) in combination with either hGi, hGq, JSGq1 or JSGiq heterotrimer (**Figure 39/41**). The assays done with JSR1 bound to either 9-cis or 9-cis 6.11 retinal showed that these two ligands behave similarly. Dark state JSR1 did not increase the GTPase activity of hGi or hGq when bound to 9-cis or 9-cis retinal 6.11, showing that the analogue also acts as an inverse agonist like 9-cis retinal (**Figure 39A/C**). In both cases, illumination of the sample leads to an increase in GTPase activity, indicating that isomerization happens in both ligands, leading to an active JSR1. The GTPase activity also reaches a very similar level, hinting that both ligands activate JSR1 to a similar level (**Figure 39**). This is also true for illuminated JSR1 bound to the ATR611 which shows a very similar GTPase activity in combination with hGi or hGq (**Figure 39**). However, the activity of the dark state JSR1 bound to ATR611 is lower, both in combination with hGi or hGq, indicating that JSR1 does not reach full activity even though the ligand is the same as the illuminated 9-cis 6.11 retinal in theory. There were previous



**Figure 64: All-trans 6.11 retinal conformation.** A: All-trans 6.11 retinal from the JSR1-JSGiq model. Carbon atoms are displayed in salmon and nitrogen in blue. B: Schematic representation of the all-trans 6.11 retinal from the JSR1-JSGiq model.

reports from bovine rhodopsin that showed that the conformation of the Schiff base link changes the activity of the receptor [161]. This study showed that retinal bound in the *syn* conformation acts as a partial agonist and the *anti* conformation acts as a full agonist [161]. Something similar might happen in our case. Our active JSR1 model shows that the retinal is bound in the *anti* conformation (**Figure 64**). The sample that was prepared for

the EM studies was exposed to light, indicating that the illuminated sample has the *anti* conformation. This suggests that in the dark state JSR1 (ATR611) the retinal could have the *syn* conformation that does not act as a full agonist and therefore we see less GTPase activity in the assay (**Figure 39B**). However, we do not have the data at the moment to confirm this. Currently, there are experimental

methods for determining the conformation of the Schiff base link between JSR1 and ATR611 before and after illumination of JSR1. Such experiments are ongoing in our lab.

For the design of the human/jumping spider Giq chimeric heterotrimer we used the hGi as the template and mutated certain residues to match the JSGq1 sequence (**Figure 40**). Surprisingly, we saw that the JSGi<sub>q</sub> chimera had a higher basal activity than the hGi (**Figure 41**). The mutations we introduced are not located in proximity to the GTP binding site but are mainly located in the  $\alpha 5$  helix (**Figure 40B**). We expected that these mutations might influence binding to the receptor and therefore the GTPase activity in combination with JSR1 but not such a dramatic change in the basal activity of the JSGi<sub>q</sub>. It was shown before that a conserved phenylalanine in the  $\alpha 5$  helix acts as a relay to the GDP binding site and this residue can influence the basal activity [162]. In our chimera we do not mutate this phenylalanine (Phe336) but we mutate Glu337 to a cysteine which might influence the Phe336 conformation and therefore the basal activity. We also detected a high basal activity, similar to the JSGi<sub>q</sub>, for the JSGq1 (**Figure 41D**), indicating that jumping spider G proteins might show higher basal activity than human G proteins; at least for the JSGq1. While vertebrate and especially human G proteins have been studied intensely, literature about invertebrate G proteins is very limited.

### 5.2.1.3 Analytical SEC experiments

The analytical SEC experiments were done to test complex formation of JSR1 (ATR611) to different G proteins. As expected, JSR1 (ATR611) formed a complex with hGi that can clearly be detected on SEC (**Figure 42A**). This was already shown by Dr. Filip Pamula who formed a JSR1-hGi complex by activating JSR1 (9-cis) with light. However, the cryo-EM analysis of the JSR1 (ATR611)-hGi complex showed that the complex dissociates during freezing of the cryo-EM grids (**Figure 48**). This was already one of the main problems in Dr. Filip Pamula's work and did not improve with the new ligand. The scFv16 can stabilize GPCR-hGi complexes [119], so naturally the next sample prepared for cryo-EM analysis included the scFv16. However, in our case the scFv16 did not help to stabilize the JSR-hGi complex, seen in the cryo-EM data analysis (**Figure 49**). Therefore, a range of different G proteins from the Gi/o and Gq/11 family were tested for complex formation. Sequences of the  $\alpha 5$  helix of the different G proteins are shown in **Table 7**.

	H5.01	H5.02	H5.03	H5.04	H5.05	H5.06	H5.07	H5.08	H5.09	H5.10	H5.11	H5.12	H5.13	H5.14	H5.15	H5.16	H5.17	H5.18	H5.19	H5.20	H5.21	H5.22	H5.23	H5.24	H5.25	H5.26
Human Gi	T	K	N	V	Q	F	V	F	D	A	V	T	D	V	I	I	K	N	N	L	K	D	C	G	L	F
Mini-Go	T	N	N	I	Q	V	V	F	D	A	V	T	D	I	I	I	A	N	N	L	R	G	C	G	L	Y
Bovine Gt	T	Q	N	V	K	F	V	F	D	A	V	T	D	I	I	I	K	E	N	L	K	D	C	G	L	F
Human Gq	T	E	N	I	R	F	V	F	A	A	V	K	D	T	I	L	Q	L	N	L	K	E	Y	N	L	V
Mini-Gsq	T	E	N	A	R	R	I	F	N	D	C	K	D	I	I	L	Q	M	N	L	R	E	Y	N	L	V
Human G11	T	E	N	I	R	F	V	F	A	A	V	K	D	T	I	L	Q	L	N	L	K	E	Y	N	L	V
Hasarius Gq1 visual	T	E	N	I	R	L	V	F	C	A	V	K	D	T	I	L	Q	N	N	L	K	E	C	N	L	V
Hasarius Gq2	T	E	N	I	R	F	V	F	A	A	V	K	D	T	I	L	Q	L	N	L	K	E	Y	N	L	V

**Table 7: Sequence alignment of the  $\alpha 5$  helix from the G proteins tested for complex formation with JSR1 (ATR611).**

JSR1 has been shown to couple and activate hGq protein in cellular assays [98] and our GTPase Glo assays. As expected, we detected clear complex formation of JSR1 with the hGq on SEC (**Figure 42B**). The hG11 belongs to the same subfamily as hGq with a sequence identity of 90% overall and 100% for the  $\alpha$ 5-helix. To our surprise, we did not detect clear complex formation of JSR1 (ATR611) with the hG11 (**Figure 42D**). This is even more surprising because we previously detected activation of hG11 by JSR1 in the GTPase Glo assay (Personal communication with Jonas Mühle).

Complex formation was detected for JSR1 (ATR611) in combination with the hGi/q chimera GP04 and the bovine Gt (bGt) (**Figure 42C/E**). Overall, the SEC data did not look as promising as for the JSR1-hGi or JSR1-hGq complexes, but we decided to make a JSR1-bGt sample for cryo-EM analysis which turned out to dissociate as well (**Figure 47**).

The human Go belongs to the same subfamily as hGi and the mini-Go has a similar sequence of the  $\alpha$ 5 helix as the hGi (**Table 7**) but there are three residues that are very different (H5.02; H5.17; H5.22). Mini-Go did not show complex formation to JSR1 (ATR611), indicating that these differences do not allow mini-Go to couple to JSR1 (**Figure 42F**). Additionally, other residues in the G protein might be different which does not allow complex formation with JSR1. The overall sequence similarity of hGo and hGi is 73%. Bovine rhodopsin was shown to couple to bGt, hGi and mini-Go which all belong to the same subfamily [117, 146, 148, 149] but for JSR1 we only detect clear coupling to hGi (**Figure 42**). However, the natural G protein for JSR1 is the JSGq1 and not a jumping spider Gi. The sequence of the  $\alpha$ 5-helix from JSGq1 is more similar to the hGq than the hGi but it includes two key residues which match the hGi (H5.18; H5.23 in **Table 7**). This might explain why we also see coupling to the hGi but JSR1 does not form a stable, high affinity complex with the hGi as seen in our cryo-EM data analysis. Both mini-Gsq chimeras did not show complex formation with JSR1 (**Figure 42G/H**). This can be expected because the main part of the chimeric protein is based on the hGs which does not couple to JSR1. The  $\alpha$ 5 helix of the mini-Gsq was mutated at certain positions to match residues of the hGq but when comparing the sequences, one can still detect several differences to the hGq (**Table 7**). Additionally, it was shown that not only the interactions with the  $\alpha$ 5 helix are important for the G protein specificity of a GPCR but also interactions with other parts of the GPCR/G protein are involved [163, 164]. Because the main part is from the hGs, these interactions are likely missing, leading to no complex formation with both mini-Gsq chimeras tested.

As mentioned already, the JSGq1 shows several differences in the sequence of the  $\alpha$ 5 helix compared to the hGi and hGq (**Table 7**). Interestingly, these differences are only seen in JSGq1, which is the visual Gq in jumping spiders, but not in the other JSGq2 which has an identical sequence of the  $\alpha$ 5 helix as the hGq (**Table 7**). Because we never worked with jumping spider G proteins before and were not sure that we can express and purify the protein, we also designed a chimeric G protein where we mutated

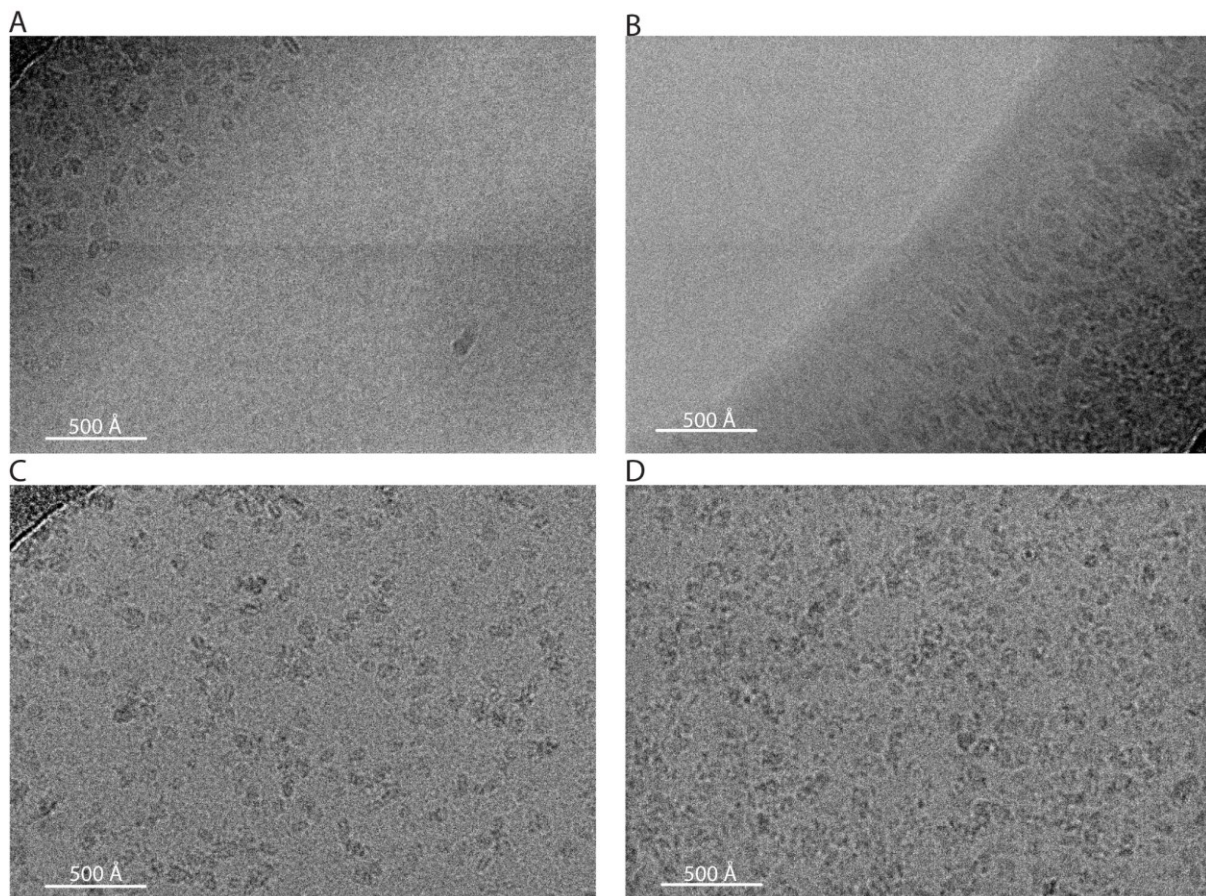
residues of the hGi to match the sequence of the JSGq1 (**Figure 40**). In the end, expression and purification of the JSGq1 was also possible but with lower yields than the JSGiQ chimera. For efficient expression of functional human Gq protein co-expression of a chaperon, Ric-8A, is important [160]. For the JSGq1 we did not co-express this chaperon which could be an additional reason for the lower yield. JSR1 (ATR611) showed clear complex formation with the JSGiQ chimera but surprisingly not with the JSGq1 which was confirmed by SDS-PAGE (**Figure 43**). As mentioned before, we never worked with jumping spider G proteins, or invertebrate G proteins in general, which is also the case for our collaborators Prof. Akihisa Terakita and his group. Therefore, we also lack the experience and vital information about working with jumping spider G proteins *in vitro*. The JSGq1 has a sequence similarity of 50% and 80% compared to the hGi and hGq respectively. It may be that jumping spider G proteins need different buffer conditions than the human G proteins, including also different detergents. One hypothesis from the SEC data was that the G protein is not functional and therefore we do not see any complex formation. However, the GTPase Glo assay data showed that the JSGq1 has intrinsic GTPase activity, showing that the protein is functional. Surprisingly, the basal activity was much higher than the basal activities of the hGi and hGq proteins (**Figure 41**). The reason behind this significant difference is still unclear, but it may be related to the incubation temperature used in the GTPase Glo assay, typically set at room temperature (~22°C). In many cases, certain groups of invertebrates, such as ectothermic species, may exhibit lower body temperatures compared to some vertebrates, potentially indicating adaptations for optimal protein function at lower temperatures in these organisms. Given that our assay's temperature is lower than the human body's usual temperature, it's possible that human G proteins may be less active due to the cooler conditions. In the time frame of my project, it was not possible to further investigate the reason we do not see complex formation with the JSGq1, but this will be subject of future research. Nevertheless, the JSGiQ chimera showed clear complex formation with JSR1 (ATR611) and a sample was prepared for cryo-EM analysis which resulted in the JSR1-JSGiQ structure.

#### 5.2.1.4 Cryo-EM analysis of JSR1-G protein complexes

Grid preparation for cryo-EM single particle analysis is a critical step which determines the quality of the data that can be collected. Ideally, we want to have a thin ice layer with many protein particles in different orientations. However, protein complexes, especially low affinity complexes, can dissociate during grid preparation [143, 144], meaning that a compromise between ice thickness and particle quality has to be made. Initially, our samples had rather low concentrations (~1-~4 mg/ml) and blotting times were between two and six seconds. At the microscope we saw that short blotting times lead to thick ice layers which were not suitable for data collection. Blotting times of four to six seconds



resulted in grids with good ice thickness but also in holes where the ice was too thin in the middle, leading to particle clustering to the side of the hole (**Figure 65A/B**). This leads to a further problem that particles overlap at the edges or even show blurry particles that suffer from beam induced motion (**Figure 65A/B**). Even if we detected particles in these thin ice areas, these particles showed a high degree of dissociation because they might be located at the air-water interface which can lead to denaturation and therefore complex dissociation which was seen for the JSR1-bGt, JSR1-hGi and JSR1-hGi-scFv16 samples (**Figure 47/48/49**). We adjusted the ice filter to also include areas with slightly thicker ice than in **figure 65AB** in order to have complex particles (**Figure 65C**). The ice filter has to be adjusted for each dataset or even changed during the data collection because if we collect micrographs from too thick ice areas, the particles will overlap and the signal to noise ratio will be worse (**Figure 65D**).



**Figure 65:** Examples of micrographs collected at a nominal magnification of 165'000 with a Titan Krios microscope. **A:** Example of a micrograph in a very thin ice area with the JSR1-hGi sample. **B:** Example of a micrograph in a very thin ice area with the JSR1-hGi-scFv16 sample. **C:** Example of a good micrograph with the JSR-hGi-scFV16 sample. **D:** Example of a micrograph in a thick ice area with the JSR-hGi-scFv16 sample.

For the JSR1-JSGiq complex we concentrated the sample much more to ~8.5 mg/ml and blotted the grids for four to eight seconds. This change, together with the improved protein sample, resulted in better grids and a dataset suitable for high resolution structure determination (**Figure 51/52**).

However, we had to collect many micrographs because the ice filter in the first dataset was not optimal and included too many thin ice areas where not many particles were located (**Figure 51A**).

Because many of our datasets showed dissociation of the complex, we also tried a different grid preparation method with the VitroJet. In the end, we could not see any particles on these grids. This could also be because of the sample, which was frozen and thawed for making these grids and might have led to denaturation or aggregation of the proteins.

To summarize, grid preparation is still one of the bottlenecks in structure determination with cryo-EM and one should screen different conditions for grid preparation. Additionally, during data collection it is important to screen areas with different ice thickness in order to optimize the ice filter settings and collect good data for structure determination.

## 5.2.2 JSR1-JSGiq structure

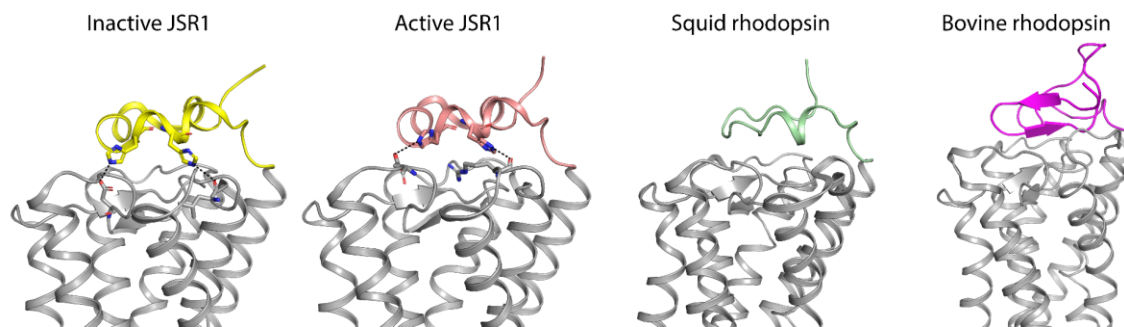
### 5.2.2.1 Active state JSR1

In this project we solved the first active state structure of an invertebrate bistable rhodopsin, namely JSR1. Dr. Filip Pamula already solved a 4.9 Å structure but the resolution did not allow confident modelling of side chains. I managed to improve the resolution in my JSR1 (ATR611)-JSGiq\_1 map to 4 Å. This includes all domains of the complex and the resolution varies between the different subunits of the complex (**Figure 52D**). The map quality for JSR1 was sufficient to be able to model side chains which allows detailed analysis of the structure. The second map (JSR1 (ATR611)-JSGiq\_2) reached a global resolution of 4.17 Å but this structure will only be discussed in regard to the G protein binding site. Here we will discuss the active state JSR1 structure from the JSR1 (ATR611)-JSGiq\_1 map.

### 5.2.2.2 N-terminal extracellular domain

The N-terminal extracellular domain of JSR1 forms a crown-like structure, acting as a lid over the transmembrane bundle and protecting the ligand binding site from the extracellular milieu (**Figure 66**). This fold is conserved in the inactive and active state of JSR1 (**Figure 66**) [30]. In the inactive structure it was seen that this fold is stabilized by two histidine residues in the crown that form hydrogen bonds with residues in the transmembrane bundle [30]. These interactions are also present in the active state JSR1, hinting that these might be important for the fold and/or stability of the N-terminus (**Figure 66**). This cap is also conserved in other monostable or bistable rhodopsins but the structure can vary, hinting at the importance of the cap to protect the retinal binding site from the extracellular environment (**Figure 66**) [30, 31, 145]. Furthermore, this fold might also be important for the receptor stability overall.



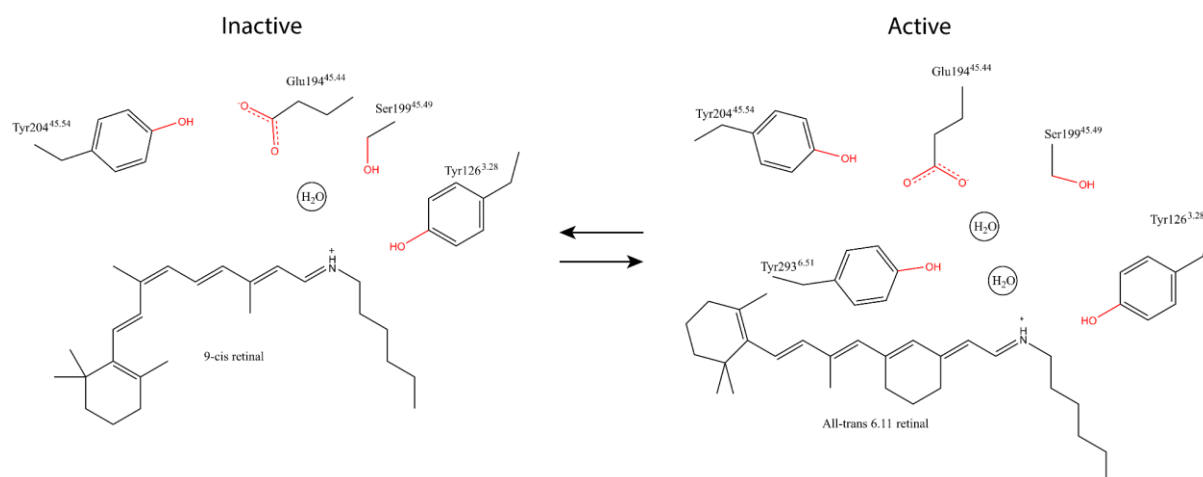


**Figure 66: N-terminal domains of different rhodopsin structures.** The N-terminal domains are labelled in yellow for the inactive JSR1 (PDB: 6I9K [30]), salmon for the active JSR1, palegreen for squid rhodopsin (PDB: 2Z73 [31]) and magenta for bovine rhodopsin (PDB: 1GZM [145]). The rest of the protein is labelled in gray. Both JSR1 structures show the two histidines and the hydrogen bond partners as sticks.

### 5.2.2.3 The retinal binding pocket

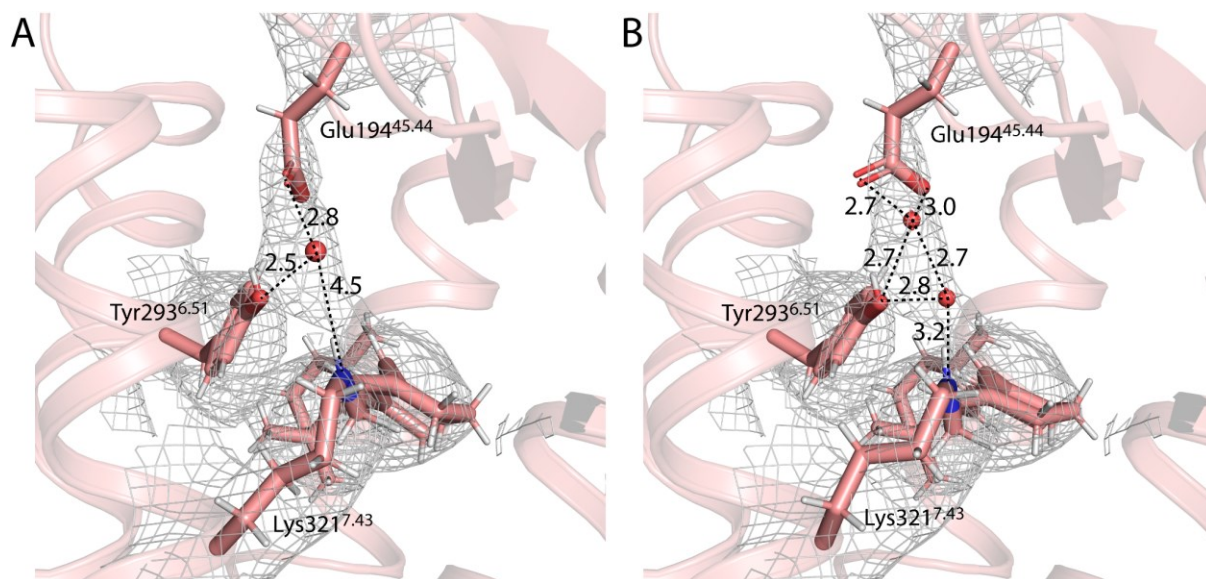
Bistability confers the ability to maintain the Schiff base link to the retinal in both the inactive and active state [91]. Comparing the environment of the retinal ligand between the active and inactive state of JSR1 showed that the 9-cis retinal in the inactive state is tightly packed by surrounding residues (15 residues within 4 Å) [30]. In contrast, the ATR6.11 is only within 4 Å of seven residues. Similarly, 11-cis retinal in bovine rhodopsin is also tightly packed by surrounding residues (15 residues within 4 Å) [145] and the all-trans retinal only has eight residues within 4 Å [146]. This could indicate larger flexibility of the retinal binding pocket and/or water molecules surround the retinal, bridging interactions with side chains in the active state.

In contrast to the monostable bovine rhodopsin, the Schiff base in JSR1 is not deprotonated and hydrolysed upon retinal isomerisation from 11-cis to the all-trans conformation [91, 98]. The proximal counterion, Glu<sup>3.28</sup>, in monostable rhodopsin together with a water molecule are responsible for hydrolysis of the Schiff base [149, 165, 166]. In JSR1 and other bistable opsins this position is occupied by a bulkier Tyr126<sup>3.28</sup>, leaving no space for a water molecule at a similar position as it is seen in the active bovine rhodopsin structure [149]. The presence of a Tyr instead of a Glu might contribute to the thermal stability of the Schiff base in the active state JSR1 and other bistable opsins. Our experimental map for the JSR1-JSGiq complex showed clear density for the Schiff base, demonstrating the thermal stability in the active state (**Figure 54A/B**). In the inactive state of JSR1, the protonated Schiff base is stabilized by the counterion Glu194 via a water mediated hydrogen bond network (**Figure 56C**) [30]. From our data it is likely that in both the inactive and active states of JSR1, water molecules are involved in the Schiff base counterion link [30], highlighting the importance of water molecules in JSR1. The existence of similar water mediated networks in the inactive and active states may facilitate the bistability of JSR1. Based on our data, we propose that the water mediated network changes in the active state compared to the inactive state JSR1. From our JSR1-JSGiq\_1 structure it seems likely that Ser199 is not involved in the active state, but Tyr293 might be involved instead (**Figure 67**).



**Figure 67: Schematic representation of the proposed mechanism of the PSB – Cl link in the inactive state (left) and active state (right).** In the inactive state JSR1 the PSB – Cl link is mediated by a hydrogen bond network where Glu194, Ser199, a water molecule and the PSB are participating. This is based on the crystal structure 6I9K. In the active state we propose that Ser199 is not participating in this network anymore. The PSB – Cl link is mediated by a similar water mediated hydrogen bond network between Glu194, two water molecules, the PSB and possibly the Tyr293.

From our analysis it also seems likely that there are two water molecules present that mediate the PSB – Cl link in the active state (**Figure 67/68**). Two models were generated where hydrogens were added to the model and then either one or two water molecules were added manually in Coot [133]. Real space refinement for the retinal binding pocket was used to optimize the model. The model with one water molecule showed that the distance between the nitrogen of the PSB and the water molecule is too large for a direct interaction (4.5 Å), meaning there is also free space that can be occupied by a second water molecule (**Figure 68**). On the other hand, when we insert two waters to the retinal binding site, the PSB can make a direct interaction with one of the water molecules (3.2 Å) and the second water molecule can interact with the counterion Glu194 which changes the rotamer slightly to accommodate the water molecules (**Figure 68**). This creates a water mediated PSB – Cl link. The Tyr293 is also in hydrogen bonding distance and could participate in this network (**Figure 68**).



**Figure 68: PSB – CI link in the active state JSR1 models based on our JSR1-JSGiq\_1 complex model. A:** Model of the PSB – CI link with an active JSR1 model and one water molecule. Experimental map is shown as mesh around the residues shown as sticks. **B:** Model of the PSB – CI link with an active JSR1 model and two water molecules. Experimental map is shown as mesh around the residues shown as sticks.

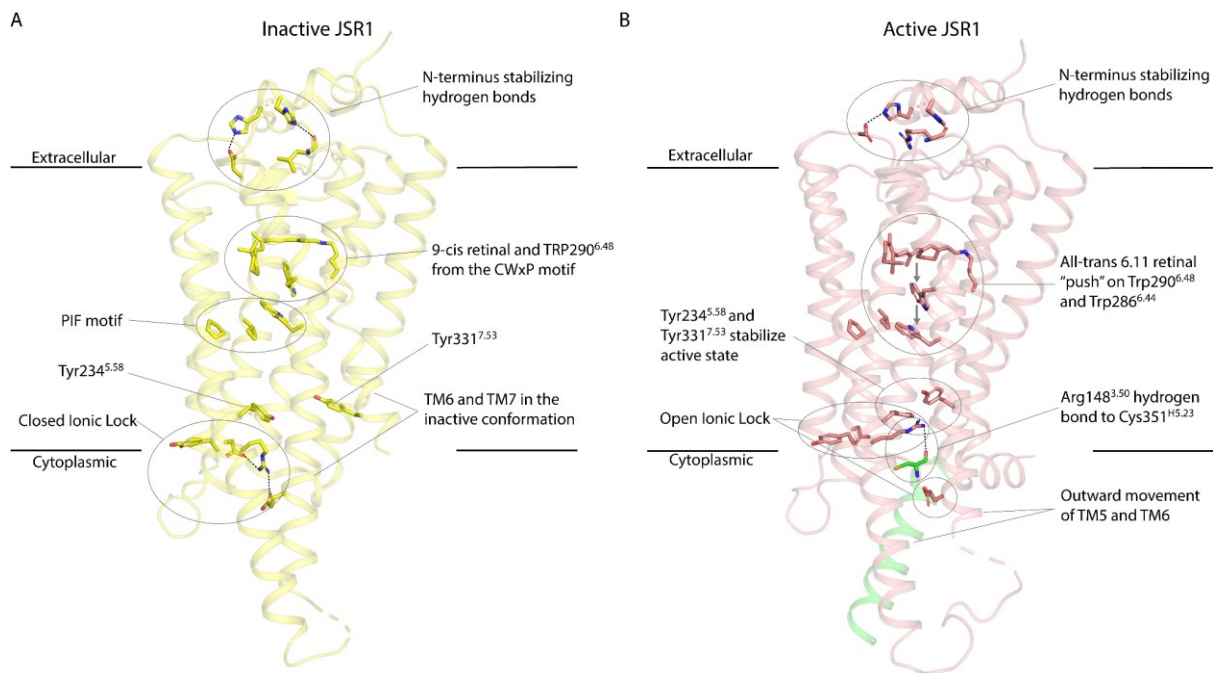
Bistable rhodopsins are interesting receptors for optogenetic applications because they can be activated and inactivated with specific wavelengths of light. JSR1 has been shown to work as an optogenetic tool in zebrafish reticulospinal V2a neurons where activation of JSR1 led to an increase of  $\text{Ca}^{2+}$  levels and evoked a change in swimming behaviour [92]. The main issue with using JSR1 as an optogenetic tool is that the  $\lambda_{\text{max}}$  of the active and inactive states overlap, meaning that there will be a mixture of active and inactive states upon illumination [100]. One JSR1 mutant (JSR1 S199F) has been shown to have a  $\lambda_{\text{max}}$  of 380nm in the inactive state and 540nm in the active state (similar to WT JSR1) [98]. Ser199 is involved in the Schiff base-counterion link in the dark state but not in the active state, explaining the difference of  $\lambda_{\text{max}}$  in the dark state compared to the WT JSR1 [98]. The disruption of the Schiff base-counterion link in the dark state of the S199F mutant leads to a deprotonated Schiff base, explaining the shift of  $\lambda_{\text{max}}$  (Personal communication with Dr. Matthew Rodrigues) [98]. UV-light activation is not very desirable for *in vivo* applications because UV-light does not penetrate tissues well. Our structure will facilitate rational protein engineering of the retinal binding pocket to change  $\lambda_{\text{max}}$  of the active and inactive states. We identified a Tyr293 that does not participate in the PSB – CI link in the dark state but could participate in the illuminated state (**Figure 68**). Mutating this residue could shift  $\lambda_{\text{max}}$  of the illuminated state but not in the dark state. There is an alanine (Ala317) which is close to the water network but is neither important in the dark state or illuminated state for the PSB – CI link. Mutating this residue to a polar residue could change the water network, leading to a change in  $\lambda_{\text{max}}$  of JSR1. In mouse melanopsin (mMeOp) it has been shown that mutating the equivalent residue (Ala333 in mMeOp) to a serine changes the  $\lambda_{\text{max}}$  [167]. A further change was observed when additionally Tyr309 (Tyr293 in JSR1) was mutated to a phenylalanine [167]. Currently, experiments are

ongoing or planned in our lab to mutate several residues in the retinal binding pocket of JSR1 in order to shift  $\lambda_{\max}$ .

#### 5.2.2.4 Micro-switch domains

Although we are working with an invertebrate GPCR, our structure showed that all the major microswitch domains like the C-W-x-P, P-I-F or D/E-R-Y motif follow the typical structural changes seen in vertebrate active class A GPCR structures (see Results “4.2.3.3 Micro-switch domains in JSR1”). This hints that both invertebrate and vertebrate rhodopsins (or vertebrate class A GPCRs) have similar activation mechanisms and over the course of evolution these mechanisms are conserved. However, there are subtle differences in JSR1 compared to monostable bovine rhodopsin or the  $\beta$ 2-adrenergic receptor (class A GPCR).

One difference can be found in the C<sup>6.47</sup>-W<sup>6.48</sup>-x-P<sup>6.50</sup> motif. Our sequence analysis showed that invertebrate rhodopsins possess either an alanine or serine at position 6.47 (**Figure 58G**). On the other hand, vertebrate rhodopsins almost exclusively possess a cysteine at this position (**Figure 58G**). The exact role of the cysteine is not very well known but studies suggested that it participates in the rearrangement of the TM6 and TM7 interface upon receptor activation [21, 168]. One study suggested that the protonation state of the cysteine changes upon activation which is mainly attributed to the movement of TM6 which changes the local environment of the cysteine [168]. JSR1 possess an alanine at position 6.47, meaning that similar functions are not possible. JSR1, and other invertebrate rhodopsins, might have a slightly different mechanism where this position is less critical for receptor function compared to vertebrate rhodopsins, or other class A GPCRs, that possess a cysteine at position 6.47. The tryptophane at position 6.48 and the proline at 6.50 are highly conserved residues (**Figure 58G**). The Pro<sup>6.50</sup> acts as a hinge for the outward movement of TM6 in the activation process and the Trp<sup>6.48</sup>, also referred to as the “toggle-switch”, is often considered as the primary activation switch in class A GPCRs [18, 19, 21, 23]. The retinal isomerization leads to a movement of the polyene chain towards the intracellular side. Trp290<sup>6.48</sup> acts as a lever and the all-trans retinal “pushes” onto Trp290<sup>6.48</sup> and subsequently Trp286<sup>6.44</sup> from the P-I-F motif, leading to the outward movement of TM6 (**Figure 69**). Our JSR1 structure shows that the Trp290<sup>6.48</sup> is likely also the primary activation switch as it was shown for other class A GPCRs. Curiously, all rhodopsin sequences analysed showed that Pro<sup>6.50</sup> is followed by a conserved Tyr<sup>6.51</sup>. This tyrosine has been shown to be important for the coordination of the distal counterion (**Figure 55A/B; 56C/D; 58G**) [30, 31, 145, 146, 149]. Interestingly, this is true for both bistable rhodopsins and monostable rhodopsins although monostable rhodopsins have an additional proximal counterion [88].



**Figure 69: Activation of JSR1 by the agonist all-trans 6.11 retinal.** **A:** Inactive JSR1 (PDB: 6I9K) is shown in yellow. Side chains of the same key residues as in A are shown as sticks. **B:** Active JSR1 is shown in salmon and the C-terminus (residues 334-354) of the Gaiq in green. Side chains of key residues involved in the activation are shown as sticks.

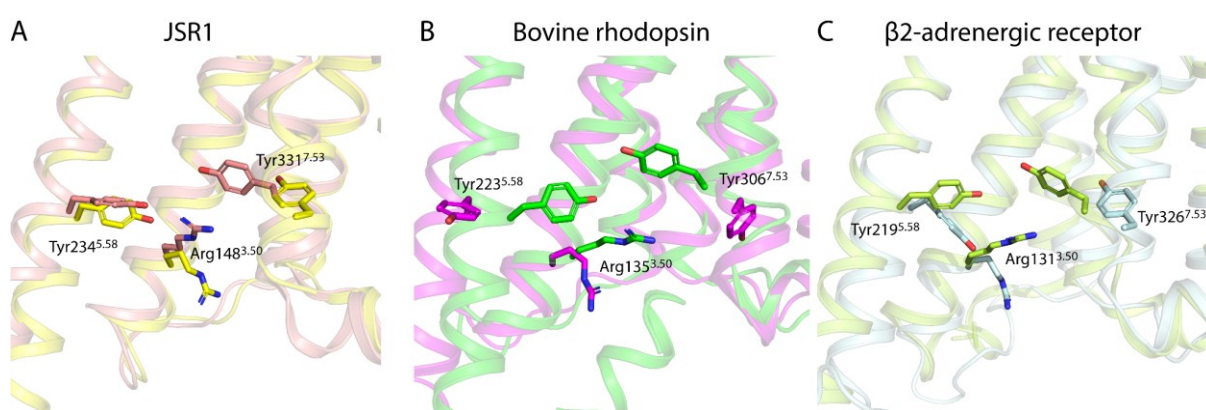
Another difference was in the P<sup>5.50</sup>-I<sup>3.40</sup>-F<sup>6.44</sup> motif where JSR1 possesses a tryptophane instead of a phenylalanine at position 6.44 (**Figure 57B; 58C/H**). Our sequence analysis showed that invertebrate rhodopsins mainly possess a tryptophane at position 6.44 whereas vertebrate rhodopsins have the typical phenylalanine which is conserved in most class A GPCRs (**Figure 58H**). Although tryptophane is a bulkier residue than phenylalanine, the position in the active and inactive states overlaps well and shows a similar movement from the inactive to the active state (**Figure 58C/D**). This hints that the tryptophane has a similar function to the phenylalanine, meaning that it is also involved in the outward movement of TM6 during the activation process. Compared to bovine rhodopsin or the  $\beta$ 2-adrenergic receptor, Pro<sup>5.50</sup> shifts very little in JSR1 (**Figure 58C/D**), hinting that TM3 in JSR1 changes the location very little compared to other receptors.

The D/E<sup>3.49</sup>-R<sup>3.50</sup>-Y<sup>3.51</sup> is another highly conserved motif which is located at the cytoplasmic side of TM3. In the inactive state JSR1 Arg148<sup>3.50</sup> forms an intra-helical ionic interaction with Asp147<sup>3.49</sup> and an inter-helical ionic interaction with Glu272<sup>6.30</sup> [30]. Both ionic interactions are broken in the active JSR1 (**Figure 58E**). Arg148<sup>3.50</sup> forms a hydrogen bond with the backbone oxygen of a cysteine from the JSGaiq (**Figure 69**). The rotamer change of Arg148<sup>3.50</sup> is stabilized by a tyrosine at position 5.58 and the active state is further stabilized by another tyrosine at position 7.53 from the N-P-x-x-Y motif (**Figure 69**) [28, 29]. Curiously, JSR1 shows similar conformational arrangements like the non-photosensitive  $\beta$ 2-adrenergic receptor which is different from bovine rhodopsin (**Figure 70**) [24, 145, 146, 169]. In JSR1 and the  $\beta$ 2-adrenergic receptor, Tyr<sup>5.58</sup> is oriented towards the transmembrane core



in both the inactive and active states and undergoes a rotamer change during activation (**Figure 70A/C**) [24, 30, 169]. On the other hand, in bovine rhodopsin the tyrosine points away from the transmembrane core in the inactive state and upon activation moves towards the transmembrane core (**Figure 70B**) [145, 146]. This is similar to the changes seen for Tyr<sup>7.53</sup>, although this tyrosine undergoes bigger conformational changes in all three receptors due to the movement of TM7 compared to Tyr<sup>5.58</sup> (**Figure 70**).

Together, these observations indicate that the activation mechanism of JSR1 is more similar to non-photosensitive class A GPCRs than vertebrate rhodopsins.



**Figure 70: Structural comparison of residues Arg<sup>3.50</sup>, Tyr<sup>5.58</sup> and Tyr<sup>7.53</sup> between JSR1, bovine rhodopsin and β2-adrenergic receptor.** The before mentioned residues are shown as sticks and the rest of the receptor as cartoon. **A:** Comparison of the inactive (yellow, PDB: 6I9K) and active JSR1 (salmon, JSR1-JSGiq\_1). **B:** Comparison of inactive (magenta, PDB: 1GZM) and active (green, PDB: 4A4M) bovine rhodopsin. **C:** Comparison of the inactive (palecyan, PDB: 2RH1) and active (lime, PDB: 3SN6) β2-adrenergic receptor.

### 5.2.2.5 JSR1-G protein interactions

In total we have three JSR1-G protein complex structures in our lab. The JSR1 (ATR)-hGi structure was solved by Dr. Filip Pamula and two different conformations of the JSR1 (ATR611)-JSGiq complex solved by the author. Our analysis of the JSR1-hGi and JSR1-JSGiq complexes also emphasizes several differences in the G protein engagement. We have not used any antibodies to stabilize these complexes, which might increase the flexibility of the G protein and decrease the resolution, but these motions provide us interesting insights into G protein binding. Initially, we tried to form a JSR1-hGi complex because JSR1 can potentially be used as an optogenetic tool. Compared to other class A GPCR-hGi complexes, our JSR1-hGi complex shows an unusual binding pose of the human G<sub>ai</sub>. The α5 helix adopts a shallower binding pose compared to other GPCR-hGi complexes. Nevertheless, JSR1 is active and in our *in vitro* activity assays we see clear activation of the hGi by JSR1 (**Figure 39**). This different binding pose could arise from the difference of the amino acid sequence of the α5 helix when compared to the visual jumping spider G<sub>q</sub> (**Figure 40A**). JSR1 is able to couple to hGi and activate it, as seen in our *in vitro* activity assay, but it might do it through different interactions than vertebrate GPCRs. Additionally, Tyr331<sup>7.53</sup> does not seem to be in a fully active conformation which might also

come from the unusual movement of TM7 in the JSR1-hGi complex. However, the resolution does not allow confident modelling of side chains so this information has to be considered carefully.

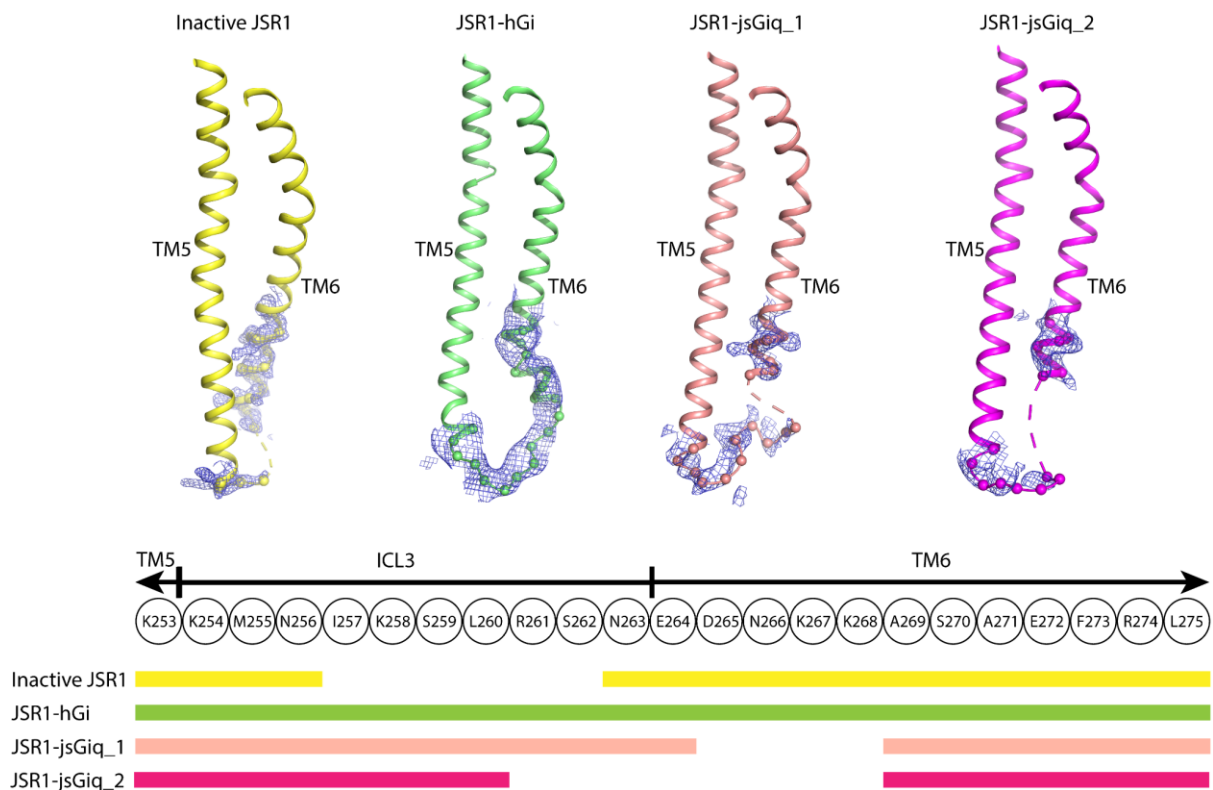
On the other hand, our JSR1-JSGiq\_1 complex shows a fully active JSR1 and a fully engaged JSGiq in one conformation of the complex. The other conformation of the JSR1-JSGiq\_2 complex shows also a fully active JSR1 but a 16° rotated G protein (**Figure 60C/D**). This rotation of the G protein likely results from a dynamic equilibrium between several states of the GPCR/G protein complex, and our structures likely represent the two most populated states. This indicates that binding and activation of the G protein is a highly dynamic process in which the G protein samples different binding poses. The density for the hook is also less resolved in the JSR1-JSGiq\_2 structure compared to the JSR1-JSGiq\_1 structure, indicating that the hook is not fully formed and engaged with JSR1. Overall, there are less JSR1-JSGiq interactions in the JSR1-JSGiq\_2 structure, further indicating that this represents maybe a prebound G protein conformation but not the fully engaged state (**Table 5**). Arg148<sup>3,50</sup> interacts with the backbone oxygen of Cys351 and this interaction is very conserved in GPCR-hGi structures. This interaction is missing in the JSR1-JSGiq\_2 structure, further indicating that it represents an intermediate state and the JSR1-JSGiq\_1 structure represents the fully engaged G protein state. Our work shows that despite differences in the structures of the JSR1-hGi and JSR1-jsGiq complexes, the receptor is capable of catalyzing nucleotide exchange in both G proteins. The implications of these different interactions will be subject of future research of bistable invertebrate rhodopsins.

Another interesting region is the cytoplasmic end of TM5, ICL3 and cytoplasmic end of TM6. We recognized that in all the different structures we have of JSR1, there are missing or not well resolved densities at different positions (**Figure 71**). In the inactive JSR1 structure most of ICL3 is missing [30]. In the JSR1-JSGiq\_1 structure a small part at the cytoplasmic end of TM6 is missing and in the JSR1-JSGiq\_2 structure there are additional missing residues in ICL3 (**Figure 71**). This indicates that this region is more flexible in the JSR1-JSGiq\_2 structure, likely having less interactions with the G protein. Indeed, the additionally missing residues in the JSR1-JSGiq\_2 structure (R261, S262 and N263) make interactions with the JSGiq in the JSR1-JSGiq\_1 structure, likely stabilizing this part (**Table 5**). Interestingly, in the JSR1-hGi structure from Dr. Filip Pamula, clear density for the whole region can be seen (**Figure 71**). It might be that this region makes more interactions with the hGi and therefore it is less flexible, and a clear density can be detected. However, detailed interactions are difficult to analyse because the resolution of the JSR1-hGi complex map does not allow modelling of side chains (**Figure 71**).

Nevertheless, together these observations indicate that flexible regions like the ICL3 can be stabilized by interactions with the G protein as seen in the JSR1-hGi and JSR1-JSGiq\_1 structures. Furthermore, both of our JSR1-JSGiq structures show missing parts at the cytoplasmic end of TM6, indicating that it



might not form an  $\alpha$ -helix in this part but rather a flexible loop-like structure in the active JSR1. The JSR1-hGi structure shows an atypical bend of TM6 exactly where the density disappears in the JSR1-JSGiq maps, meaning that the  $\alpha$ -helical structure might also be disrupted.



**Figure 71: TM5-ICL3-TM6 comparison in JSR1 models.** Top shows models of TM5-ICL3-TM6 for inactive JSR1 (PDB=6i9k), active JSR1 from the JSR1-hGi model, active JSR1 from the JSR1-JSGiq\_1 model and active JSR1 from the JSR1-JSGiq\_2 model. Experimental map is shown for JSR1 and residues 253-275. Bottom shows a schematic representation of the residues 253-275. Each row below shows which residues are present (solid bar) and missing (no bar) from the respective models.

In addition, it was suggested before that monostable rhodopsins can activate G proteins more efficiently due to the displacement of the counterion from Glu181 to Glu113 (bovine rhodopsin numbering) and a larger opening movement of TM6 [88, 170]. In our JSR1-hGi structure we see a similar outward movement of TM6 compared to bovine rhodopsin-Gi/Gt structures [117, 146, 147] and in the JSR1-JSGiq the outward movement of TM6 is even bigger (**Figure 59**). Our structures indicate that the G protein subtype determines the extent of the TM6 movement rather than mono- or bistability. For optogenetic applications, further experiments will be needed to determine the signalling profile of JSR1 with different G protein subtypes and homologues.

This work presents the first bistable invertebrate structure showing the activation mechanism and how bistability is achieved in JSR1. Furthermore, we highlight differences in the G protein interaction of JSR1 when either using a hGi or the JSGiq chimera for complex formation. This provides a framework for future structural experiments of bistable rhodopsins and rational engineering of JSR1 for optogenetic applications.

## 6 Conclusion

In my PhD I was working with GPCRs from two different families, called the adhesion GPCR and opsin family, which perceive very different signals. While opsins perceive photons, aGPCR activity is modulated by the tethered Stachel peptide or other proteins that bind to the receptor. However, in both cases the main goal was to solve the three dimensional structure of the active state receptor coupled to a G protein. The aGPCR project was a new project in the group, meaning that it involved many experiments before a cryo-EM sample could be prepared for analysis. On the other hand, the JSR1 project has been running in our group for years which allowed fast preparation of cryo-EM samples for structure determination.

### 6.1 Adhesion GPCR project

In this work, I contributed to the biochemical characterization of two aGPCRs called Adgrg6 and Adgrd1. Additionally, I could show that Adgrd1 and Adgrg6 can be activated with the respective Stachel peptide and couple to the Gs heterotrimer. Adgrd1 forms a stable complex with the Gs heterotrimer that can be purified and used for cryo-EM studies.

At the start of my PhD, there were no publications where *in vitro* experiments with the aGPCRs 7TM domain were conducted, suggesting that these receptors are difficult to express recombinantly and subsequently to purify. *In vitro* experiments focused on the ECDs of aGPCRs. My aim was to purify an aGPCR in order biochemically and structurally characterize the receptor. Because several studies have performed signaling assays in HEK cells with aGPCRs, the first expression constructs we designed aimed for recombinant expression in HEK cells. However, we could not confirm expression of any of the constructs in HEK cells. This was surprising as other groups, among them our collaborators Prof. Adriano Aguzzi and his group, could express similar constructs in HEK cells for cell signaling assays. In our case we aim to massively overexpress the receptors which might lead to negative effects like wrong folding and/or aggregation of the overexpressed receptors. Our group mainly worked on light sensitive GPCRs that could be purified either in the apo state or bound to an inverse agonist from either a native source (bovine retinas) or from recombinant expression in HEK cells. Therefore, we searched the literature how other groups prepared their GPCR-G protein complex samples for cryo-EM studies. Class B1 GPCRs (closest family to the aGPCR family) were mostly co-expressed in insect cells with the G protein heterotrimer which allows complex formation directly in the cell membrane. The interaction with the G protein stabilizes the active GPCR and allows extraction and purification of the complex. This approach turned out to be very successful and during the time of this project many GPCR-G protein complexes for cryo-EM studies have been prepared in this way. With this information we designed several truncated Adgrg6 and Adgrd1 constructs for insect cell expression similar to the

published class B1 GPCR constructs. I was able to show that both receptors can be expressed in insect cells, but it became clear that they are not stable in the detergent micelle when extracted from the plasma membrane. This means that neither receptor could be purified in the apo state. The lack of high affinity ligands for aGPCRs further hinders the possibilities for purification of aGPCRs, meaning that the only viable option is likely the co-expression with the G protein. In cell signaling assays it was shown that expression of the Adgrg6 and Adgrd1 CTF leads to an active receptor that is capable of signaling through the Gs pathway. I was able to show that both receptors can be activated with the corresponding Stachel peptide, leading to complex formation with the Gs heterotrimer. However, in my case I needed to add chemically synthesized Stachel peptide to activate the receptors and form the complex even though we also expressed the CTF of Adgrg6 and Adgrd1. After the structures of several aGPCRs were published, we compared their constructs to ours and recognized that one publication mentioned that inserting a methionine between the N-terminal HA signal sequence and the CTF is important for proper cleavage of the HA signal peptide and therefore proper activation of the receptor. Nevertheless, complex formation of Adgrg6 and Adgrd1 with the Gs heterotrimer was possible by adding the respective chemically synthesized Stachel peptide. Complex formation was necessary to stabilize the receptor to be able to purify the complex for cryo-EM studies. During the biochemical characterization it became clear that Adgrg6 is a less suitable receptor because of the low expression yields and low stability of the receptor. The co-expression screenings with the Gs heterotrimer showed that Adgrg6 expression is very low compared to Adgrd1 expression. As of the end of my PhD several structures of aGPCRs have been solved but the structure of Adgrg6 is still missing, further hinting that the receptor is not stable enough for structural studies. Therefore, a Adgrd1-Gs heterotrimer complex was purified and used to prepare cryo-EM grids. Purification of the Adgrd1-Gs complex was achieved by co-expressing Adgrd1 and the Gs heterotrimer in insect cells, followed by activation of Adgrd1 and complex formation with the Gs heterotrimer in the cell membrane. Subsequently, the complex was extracted from the cell membrane by detergents and purified to make cryo-EM grids. In addition to the co-expression approach, I showed that an alternative approach, where we created a fusion protein of Adgrd1 with the mini-G $\alpha$ s, can be deployed to purify Adgrd1. Moreover, the yield after purification was significantly higher for the fusion protein, indicating that this can be a valuable approach to purify Adgrd1 or other GPCRs that can not be purified without a binding partner. The SEC profile of the first sample prepared for cryo-EM studies showed a large void peak indicating aggregation of the proteins. Nevertheless, initial cryo-EM data looked promising and 2D class averages showing the full complex were generated. However, the amount and quality of the data did not allow high-resolution 3D reconstruction of the complex. The SEC profile of the next sample looked better with less aggregation but ultimately the cryo-EM data was not better

than the first sample. This was likely not only because of the changes done in the purification protocol

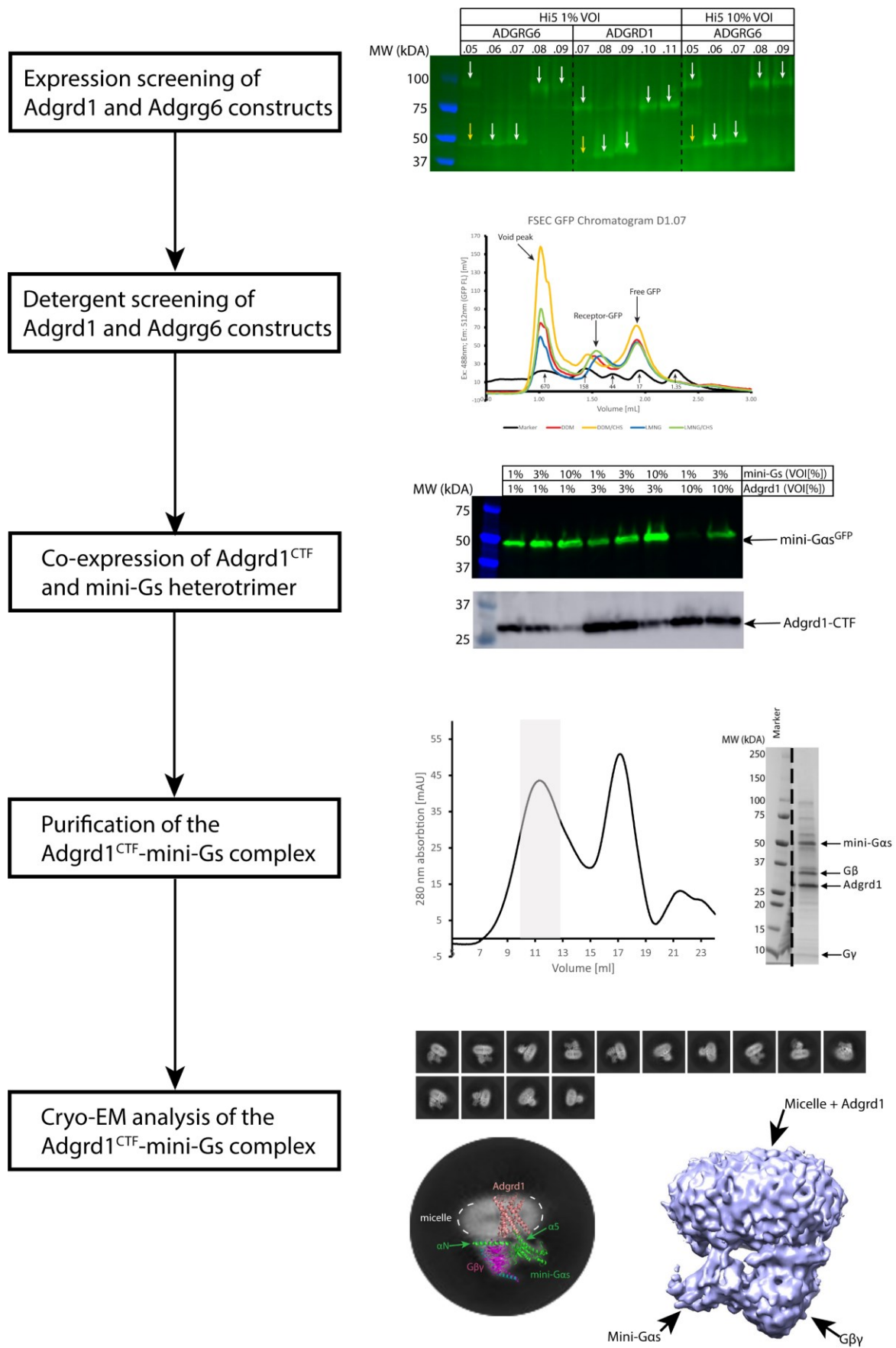


Figure 72: Overview of the structure determination of an aGPCR-G protein complex.

but also because of lower expression levels as the yield was much lower compared to the first purification. During optimization of the purification protocol and grid preparation, four studies were published which solved structures of different aGPCR-G protein complexes. The Adgd1-Gs heterotrimer complex was solved by two groups which ultimately led to the decision that we will not continue with the structure determination of the Adgrd1<sup>CTF</sup>-Gs complex. Nevertheless, the published results agreed well with our data. **Figure 72** shows an overview of the work done for the structure determination of an aGPCR-G protein complex.

All of the published structures focused on the CTF of several aGPCRs. In two publications they tried to determine the structure of an aGPCR with either the GAIN domain or the full-length NTF bound. In one publication they saw density for the NTF but it was too flexible for high-resolution structure determination. In the other publication they created an uncleavable GAIN domain by introducing mutations. Surprisingly, they did not see any density for the GAIN domain and the Stachel peptide showed the same binding pose as the CTF structure. This hints that the GAIN domain is extremely flexible. Together, these studies showed that structure determination of the CTF together with the NTF is difficult. However, the interaction of the NTF with the CTF is one of the big remaining open questions in the aGPCR field. We attempted to address this by creating GAIN-CTF constructs of Adgrd1 where we introduced mutations that could potentially lead to formation of disulfide bridges between the CTF and GAIN domain, thereby restricting the flexibility of the GAIN domain for structure determination. Unfortunately, expression of our constructs was already not very promising as only faint bands with a lot of contaminants or oligomeric structures were detected on SDS-PAGE. A purification trial resulted in a sample where the SEC peak was located in the void volume, indicating aggregation or formation of oligomeric clusters. The negative staining grids from the sample confirmed aggregation of the protein. This could be because the protein was not correctly folded or the cysteine mutations we introduced could have led to inter-molecular disulfide bridges. These results have to be taken with caution as the time was not sufficient to conduct thorough screenings for expression and purification of these constructs.

Overall, more work is needed to be able to solve a structure of an aGPCR including the NTF and CTF. Such a structure would give an insight into possible activation mechanisms of aGPCR and the interplay of the NTF and CTF.

## 6.2 Jumping spider rhodopsin-1 project

The JSR1 project has been running in our group for years. Members of our group biochemically and spectroscopically characterized JSR1 and in addition solved the crystal structure of the inactive state JSR1 bound to the inverse agonist 9-cis retinal. This already gave us interesting insights how JSR1 functions but crucial aspects of the active state JSR1 were still missing. I solved the structure of the

active state JSR1 bound to a chimeric Giq heterotrimer. The resolution of the cryo-EM map allowed modelling of the side chains, giving us insights into the molecular mechanism of JSR1 activation. There are three major take aways from the structure. First, the structure allowed us to make a hypothesis how the PSB is stabilized in the active state. We propose that water molecules stabilized by Tyr293 and Glu194 contribute to the extended hydrogen bond network linking the PSB and the counterion in the active state. Second, the structure gave us insights into the activation mechanism of JSR1. The structural changes seen in the major conserved microswitches are similar to other vertebrate GPCRs, indicating that the activation mechanism is conserved between invertebrates and vertebrates. Third, the structure provided detailed interactions of JSR1 with the G protein. The interaction map of JSR1 with the G protein can guide rational engineering of JSR1 mutants with different G protein specificity. This would result in versatile optogenetic tools.

Initial efforts towards structure determination of the active state JSR1 were already done by a former member of our group, Dr. Filip Pamula. My goal was to improve sample preparation and determine a high-resolution structure of the active state JSR1 bound to a G protein. To achieve this goal we purified JSR1 bound to the non-natural agonist all-trans retinal 6.11 (ATR611). It was rather surprising that JSR1 incorporated the ATR611 because previous attempts to incorporate the all-trans retinal were unsuccessful. An analysis of the structure showed that the carbon ring only leads to one additional interaction with a glycine residue compared to the all-trans retinal. However, this additional interaction is apparently enough that JSR1 can incorporate the ATR611. I showed that this retinal analogue acts as an agonist for JSR1 without illumination and that JSR1 can activate and couple to the human Gi and Gq proteins. Interestingly, the activity assays showed that the activity of JSR1 (ATR611) to induce G protein nucleotide exchange increases after light stimulation. We hypothesis that light stimulation could change the conformation of the Schiff base link which potentially can modulate the activity of JSR1. It was shown before that the conformation of the Schiff base link influences the activity of bovine rhodopsin. In the active structure of JSR1 the Schiff base link shows the *anti* conformation of the Schiff base link, indicating that this is the state after illumination. However, this is still subject to future research in our group. A preliminary HPLC analysis of extracted retinal molecules showed that illumination leads to a change in the *syn-anti* conformation of the Schiff base link, but further experiments are needed to confirm which isoforms are present in the dark and illuminated JSR1 (ATR611). Additionally, the GTPase Glo assay showed that the basal activity of the JSGiq is much higher than the hGi which was surprising. It was shown before for the hGi that a phenylalanine in the  $\alpha 5$  helix acts as a relay between the  $\alpha 5$  helix and GDP binding site. Mutations of this phenylalanine influence the basal activity, showing that mutations distant from the nucleotide binding site can influence the GTPase activity. We do not mutate this residue in our chimera, but we

do mutate a glutamate to a cysteine next to the phenylalanine which might influence the conformation of this phenylalanine, leading to a change in basal activity. Together, the GTPase Glo assay data showed that the ATR611 is an agonist of JSR1 and JSR1 couples and activates hGi, hGq and our JSGi<sub>q</sub> chimera.

Several attempts, from Dr. Filip Pamula and myself, to solve a high resolution structure of the JSR1-hGi complex were not successful because the complex has low stability, resulting in many dissociated complex particles and heterogeneity in the cryo-EM data. I showed that addition of the scFv16 which stabilizes the bovine rhodopsin-hGi complex helps aligning the particles but does not improve the stability of the JSR1-hGi-scFv16 complex. In order to improve complex stability for solving the high-resolution structure, we designed a chimeric human Gi-jumping spider Gq protein (JSGi<sub>q</sub>) where we mutated several residues of the human G $\alpha$ i to match the sequence of the jumping spider visual G $\alpha$ q1. The cryo-EM data of the JSR1-JSGi<sub>q</sub> complex indicated that the complex is more stable as no dissociated complex particles were detected. This sample allowed structure determination of the JSR1-JSGi<sub>q</sub> complex. The resolution was limited by the flexibility of the complex, especially the JSGi<sub>q</sub> subunit shows high flexibility. 3D classification allowed the identification of a second conformation of the JSR1-JSGi<sub>q</sub> complex (JSR1-JSGi<sub>q</sub>\_2) that could be resolved to a similar resolution as the first conformation (JSR1-JSGi<sub>q</sub>\_1). These structures are similar at the extracellular side and retinal binding pocket but differ at the G protein binding site.

The JSR1-JSGi<sub>q</sub>\_1 structure allowed identification of a possible mechanism how bistability is achieved in JSR1. Our hypothesis agrees well with published mutagenesis data that indicates which residues are involved in the protonated Schiff base-counterion link in the inactive and active JSR1. One JSR1 mutant where the serine 199 was mutated to a phenylalanine (JSR1S199F) shows a large shift of the  $\lambda_{\max}$  value for the inactive state but not for the active state. This agrees well with our active JSR1 structure where serine 199 does not seem to be involved in the protonated Schiff base-counterion link. In contrast, the crystal structure of the inactive JSR1 indicated that serine 199 participates in the protonated Schiff base-counterion link, explaining the large shift of the  $\lambda_{\max}$  value in the JSR1S199F mutant. Further mutations to support our hypothesis are currently being investigated. The proximal counterion, Glu<sup>3.28</sup>, in monostable rhodopsin together with a water molecule are responsible for hydrolysis of the Schiff base. We found that bistable invertebrate rhodopsins possess a tyrosine at this position which leaves no space for a water molecule at a similar position like in monostable vertebrate rhodopsins. This might contribute to the stability of the Schiff base in bistable rhodopsins. However, it is likely not the main reason for the bistability because there are bistable opsins that also possess a glutamate at position 3.28 (e.g. Zebrafish parapinopsin). Because we now have an idea which residues are important in stabilizing the protonated Schiff base – counterion link in the inactive and active states,

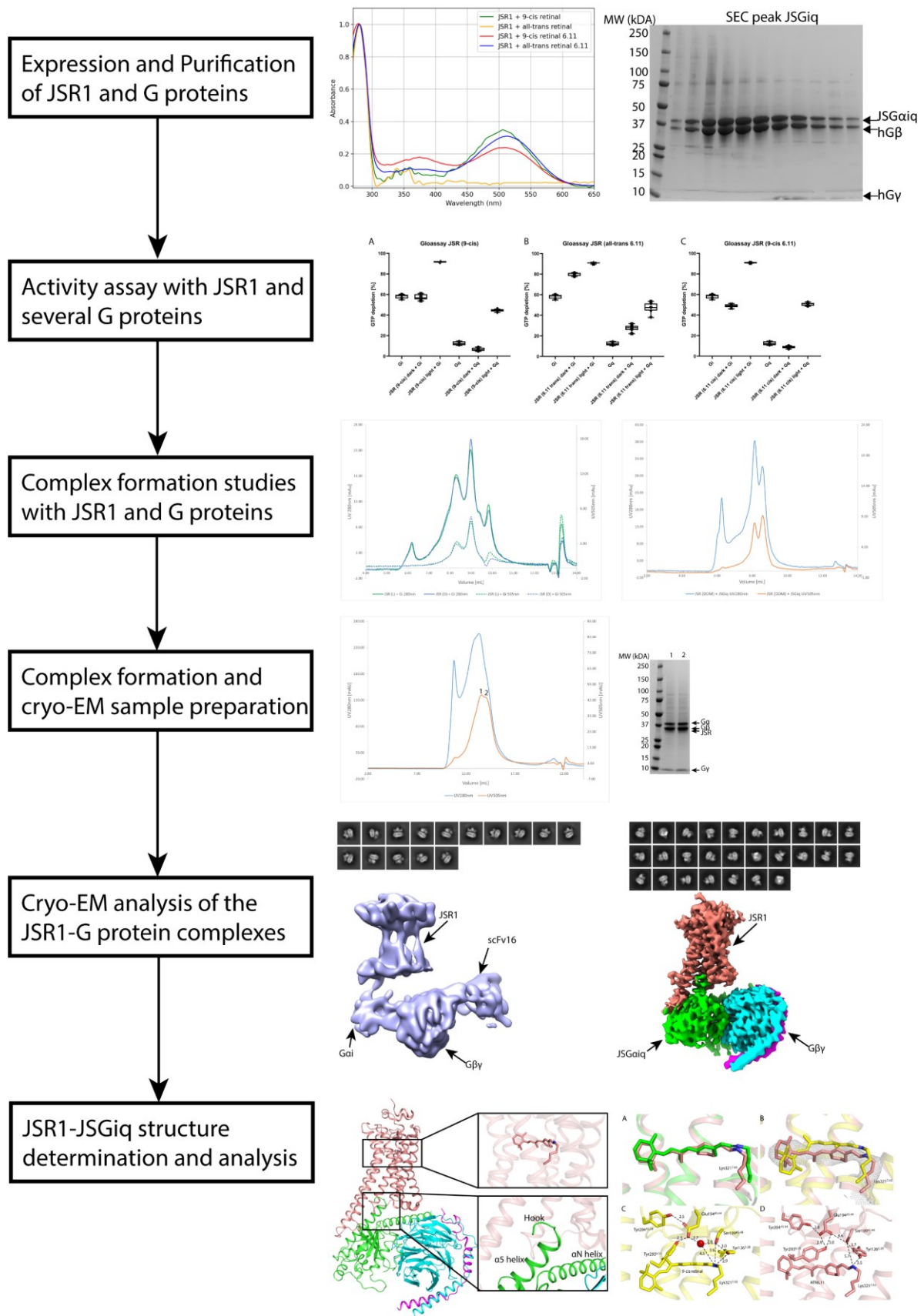


we can rationally engineer JSR1 to shift the  $\lambda_{\max}$  value of the inactive and/or active state JSR1. This will help engineering a JSR1 mutant which is more suitable for optogenetic applications.

In addition, the active state JSR1 structure showed that the major conserved motifs follow the typical structural changes seen in other vertebrate class A GPCRs, indicating that invertebrate GPCRs have a similar activation mechanism. Compared to monostable vertebrate rhodopsin, there is a difference in the sequence of the C-W-x-P motif where invertebrate rhodopsins use A/S-W-x-P instead. Until now the role of this cysteine is not very well known but it is suggested that it participates in the rearrangement of TM6 and TM7. Invertebrate rhodopsins might have a slightly different mechanism for the rearrangement of the TM6 and TM7 interface. Additionally, invertebrate rhodopsins possess a tryptophane instead of a phenylalanine in the P-I-F motif. The conformational change seen in the tryptophane in JSR1 upon activation is very similar to the change seen for the phenylalanine of bovine rhodopsin, indicating a similar mechanism. Curiously, the conformational changes of conserved motifs close to the retinal binding pocket of JSR1 (C-W-x-P and P-I-F motifs) are similar to bovine rhodopsin but the changes seen in motifs close to the G protein binding site (E/D-R-Y and N-P-x-x-Y motifs) are more similar to non-photosensitive class A GPCRs than to the vertebrate monostable rhodopsin. Therefore, JSR1 might be an attractive target to conduct time-resolved experiments to gain more insights into class A GPCR activation.

Furthermore, our JSR1-G protein structures also indicate that JSR1 might interact differently with human G proteins compared to the JSGiq chimera. In the case of the JSR1-hGi structure, JSR1 shows a different binding pose of the hGi protein compared to other GPCR-hGi complex structures. The  $\alpha 5$  helix adopts a shallower binding pose and does not seem to be fully engaged with JSR1. Nevertheless, our *in vitro* activity assay showed clear activation of the hGi by JSR1. Further research is needed to clarify this discrepancy. The lower resolution of the JSR1-hGi structure also does not allow identification of the exact interactions between JSR1 and the hGi. Interestingly, the JSR1-JSGiq\_1 structure shows a fully engaged G protein similar to other GPCR-G protein complexes. This is likely due to the mutations we made to match the sequence of the jumping spider visual Gq1. The JSR1-JSGiq\_2 structure shows a  $\sim 16^\circ$  rotated G protein compared to the JSR1-JSGiq\_1 structure. 3D variability analysis of the cryo-EM data showed that there is motion of the G protein and the two structures we isolated are likely representing two snapshots with stable conformation, respectively the most populated states. It is possible that the G protein shows different binding poses during engagement of the receptor which is visualized in the 3D variability analysis. Most GPCR-G protein complex structures show very similar binding poses of the G protein and these probably represent the most stable poses. However, *in vivo* this is a dynamic process and our data might visualize part of the G protein binding or dissociation process. Because our cryo-EM data has no temporal resolution, we can

not say which of the two conformations would be “first” in a time series. Our JSR1-JSGiq\_2 structure



also shows less interaction between the G protein and JSR1, indicating that it could be an intermediate state. Because we have the nucleotide free G protein it is more likely that the JSR1-JSGiq\_2 complex shows a G protein that starts to dissociate from the receptor.

Additionally, the electron density maps from these structures indicate that the ICL3 is flexible in the inactive state JSR1 but can be stabilized by interactions with the G protein because we can see density for ICL3 in the JSR1-G protein structures. Curiously, in both JSR1-JSGiq maps density for a small part of the intracellular end of TM6 is missing, indicating that it might not form an  $\alpha$ -helical structure but rather a flexible structure.

Furthermore, it was suggested before that monostable rhodopsins are able to activate G proteins more efficiently than bistable rhodopsins. One hypothesis was that a contributing factor is a larger opening of TM6 in monostable rhodopsins. Our JSR1-hGi structure shows a similar outward movement compared to bovine rhodopsin-hGi structures and the JSR1-JSGiq structure shows an even larger outward movement. Based on our data we propose that the extent of the TM6 movement is determined by the G protein subtype rather than mono- or bistability of the receptor. Taken together, our JSR1-G protein structures can help engineering of a JSR1 variant that is less promiscuous or to change the G protein specificity. **Figure 73** shows an overview of the work done for the structure determination of the JSR1-G protein complex.

Together with the engineering of the  $\lambda_{\max}$  values, JSR1 mutants could be developed that are activated by a desired wavelength and signal through a specific G protein. This would yield versatile tools for various optogenetic applications.

My work presents the first structure of the active state of a bistable rhodopsin, providing insights into how bistability is achieved and how bistable invertebrate GPCRs interact with G proteins. My studies provide a basis for future research on bistable rhodopsins and will help develop JSR1 as a selective optogenetic tool.

## 7 Outlook

Structural biology focused on static structures for a long time but in nature these proteins are highly dynamic, meaning it is also important to study the protein dynamics. However, static protein structures provide us with valuable insights into the function. Development of time-resolved methods like time-resolved x-ray crystallography or time-resolved cryo-EM allows researchers to gain valuable information about protein dynamics. Studying proteins with time-resolved methods will be an important task in the future. A combination of time-resolved structural and biophysical methods might shed light on intermediate states of proteins that can not be captured by conventional X-ray crystallography or single particle cryo-EM. With these methods transient events like dissociation of G protein from GPCRs or binding of GRKs to GPCRs could be studied and provide valuable information. Furthermore, computational methods are advancing at a rapid pace which could provide researchers with new software packages. These software packages, especially in combination with machine learning approaches, could help extract more data from already existing datasets or help in the analysis of large datasets which are too difficult for humans to analyze.

Structural studies of aGPCRs are still limited to either the CTF or the NTF but we lack an understanding of the interplay of these two fragments. In my studies I tried to purify a sample where we tried to immobilize the GAIN domain in order to get a structure of the CTF with the GAIN domain. However, these experiments only yielded negative results. In the future it will be important for aGPCR research to develop methods to study the interactions of the NTF with the CTF to gain further insights into the function of aGPCRs. Moreover, some aGPCRs bind large protein ligands that modulate aGPCR activity and gaining structural and biophysical insights to the mechanisms will provide valuable information. In most cases it is not clear yet how binding of a ligand to the NTF will transmit the information to the CTF and thereby regulate the activity of the receptor. Overall, we only just started understanding the atomic details of aGPCR function and it is mostly limited to the activation of aGPCRs by the corresponding Stachel peptide.

AGPCRs are involved in a wide variety of physiological processes, meaning that aGPCRs could be valuable pharmaceutical targets. However, for many aGPCRs the only known ligand is the corresponding tethered Stachel peptide. For many aGPCRs it was shown that they can be activated by adding chemically synthesized Stachel peptide but the affinity is in the micromolar range. The low affinity makes sense because the Stachel peptide is tethered to the receptor but for drug development micromolar affinity is too low. The aGPCR structures bound to the Stachel peptide and mutational studies showed which residues of the Stachel peptide are important for aGPCR activation. This information can be used to design similar peptides with higher affinity for aGPCR activation. There were already first efforts to design such peptides where they also introduced unnatural amino acids

to improve the affinity. For every aGPCR there is at least one agonist known, the corresponding Stachel peptide, but we lack orthosteric antagonists for aGPCRs. To my knowledge there is only one small molecule antagonist known for Adgrg1 and Adgrg5. Identification and development of antagonists for aGPCRs will be an important task for future research. In addition, antagonists can potentially help to determine the structure of inactive state aGPCRs. A few groups also started developing antibodies that can modulate aGPCR activity. Overall, we still lack high affinity ligands for aGPCRs that can potentially be developed as an aGPCR therapeutics.

Bistable rhodopsins are of great interest to researchers because they can potentially be used as optogenetic tools to control signaling activity and the signaling pathway by specific wavelengths of light. First studies already showed that JSR1 can be used as an optogenetic tool to activate the Gq signaling pathway. Although JSR1 can be activated with green light, precise inactivation of the receptor is not possible because the  $\lambda_{\max}$  values for the active and inactive state JSR1 are overlapping. Based on the inactive state and active state structures of JSR1, researchers will be able to rationally engineer JSR1 in order to develop a more selective optogenetic tool. First experiments are already being conducted in our lab to find a JSR1 mutant where the  $\lambda_{\max}$  values are separated. However, separation of the  $\lambda_{\max}$  values is not enough. Ideally, we would like to control JSR1 activity with long wavelength light (red and green light) because it can penetrate tissues better than short wavelength light which is beneficial for *in vivo* applications.

A good optogenetic tool would not only allow precise control of the activity of the receptor but also control of the signaling pathway. JSR1 was shown to signal through the human Gq and Gi pathways, meaning that JSR1 is promiscuous. It was shown before that GPCRs can be altered to signal through a different G protein pathway by exchanging the intracellular parts of the receptor (e.g., ICLs or intracellular ends of TM5/6/7) with the corresponding parts of a receptor that signals through the desired G protein pathway. Similar experiments are ongoing in our lab to create JSR1 mutants that signal through a desired G protein pathway. Together with engineering of the  $\lambda_{\max}$  values JSR1 can potentially be developed into a selective optogenetic tool.

Our JSR1 structures are not only valuable for understanding JSR1 in atomic detail but also for other bistable receptors where JSR1 can be taken as a template to engineer other bistable rhodopsins for optogenetic applications. Furthermore, our data can help researchers determine structures of other bistable rhodopsins.

In addition, JSR1 is a potential target for time-resolved structural methods because the activation can be precisely controlled by light pulses, possibly allowing structure determination of intermediate states which would further improve our understanding of JSR1. This would provide further insights

into the changes in the retinal binding pocket upon isomerization of the retinal. Because we can also inactivate JSR1 with light pulses we could gain insights into the re-isomerization to the 11-cis isomer. Additionally, intermediate state structures will improve our understanding how bistable rhodopsins function on an atomic level. We could also show that the structural changes in conserved microswitch domains during the activation process of JSR1 are more similar to non-photosensitive class A GPCRs meaning that JSR1 could be used to further improve our understanding of class A GPCR activation.

## 8 Appendix

### 8.1 Supplementary Tables

**Supp. Table 1:** Protein sequences for the Adgrd1 constructs in the pcDNA4-TO vector for HEK cell expression. Color code: yellow=signal peptide; grey=linker; light blue=3C protease cleavage site; green=YFP; magenta=His-tag; red=TwinStrep tag

Expression Host	Receptor	Construct name	Expressed protein Sequence
HEK	Adgrd1	D1.01	MEKLLRLCCWYSWLLLFYVNFQVRG AILMQVVPLELARGHQVALSSISYVGCSSLVLC LVATLVTFAVLSSVSTIRNQRYHIHANLSFAVLVAQVLLISFRLEPGTTPCQVMVALLH YFLLSAFAWMLVEGLHLYSMVIKVFGESESKHRYYYGMGWGFLLICIISLFSAMDY GTSNNCWLSLASGAIWAFVAPALFVIVVNIIGILIAVTRVISQISADNYKIHGDPSAFKLT AKAVAVLLPILGTSWVFGVLAVNGCAVVFQYMFATLNSLQGLFIFLHCLLNSEVRAA FKHKTKVWSLTSSSARTSNAKPFHSDLMNGTRPGMASTKLPWDKSSSHAHRVDS AVGSALEVLFGQGG VSKGEELFTGVVPILVELDGDVNGHKFVSVEGEGDATYGKL TLKFICTTGKLPVPWPTLVTTFGYGLQCFARYPDHMKQHDFFKSAMPEGYVQERTIF FKDDGNYKTRAEVKFEGDTLVNRIELKIDFKEDGNILGHKLEYNYNHNVYIMADK QKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHLYSYQSALSADPNE KRDHMLLEFVTAAGITLGMDELYK GSGSHHHHHHHHSAWSHPQFEKGGGGGGG GSGGSAWSHPQFEK*
HEK	Adgrd1	D1.02	MEKLLRLCCWYSWLLLFYVNFQVRG AILMQVVPLELARGHQVALSSISYVGCSSLVLC LVATLVTFAVLSSVSTIRNQRYHIHANLSFAVLVAQVLLISFRLEPGTTPCQVMVALLH YFLLSAFAWMLVEGLHLYSMVIKVFGESESKHRYYYGMGWGFLLICIISLFSAMDY GTSNNCWLSLASGAIWAFVAPALFVIVVNIIGILIAVTRVISQISADNYKIHGDPSAFKLT AKAVAVLLPILGTSWVFGVLAVNGCAVVFQYMFATLNSLQGLFIFLHCLLNSEVRAA FKHKTKVWSLTSSSARTSNAKPFHSDGSALEVLFGQGG VSKGEELFTGVVPILVELD GDVNGHKFVSVEGEGDATYGKLT TKFICTTGKLPVPWPTLVTTFGYGLQCFARYPD HMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKIDFKED GNILGHKLEYNYNHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGD GPVLLPDNHLYSYQSALSADPNEKRDHMLLEFVTAAGITLGMDELYK GSGSHHHH HHHSAWSHPQFEKGGGGGGGGGSAWSHPQFEK*
HEK	Adgrd1	D1.03	MEKLLRLCCWYSWLLLFYVNFQVRG TNFAIAMQVVPLELARGHQVALSSISYVGCSSL SVLCLVATLVTFAVLSSVSTIRNQRYHIHANLSFAVLVAQVLLISFRLEPGTTPCQVMA VLLHYFLLSAFAWMLVEGLHLYSMVIKVFGESESKHRYYYGMGWGFLLICIISLFA MDSYGTSNNCWLSLASGAIWAFVAPALFVIVVNIIGILIAVTRVISQISADNYKIHGDPS AFKLTAKAVAVLLPILGTSWVFGVLAVNGCAVVFQYMFATLNSLQGLFIFLHCLLNS EVRAAFKHKTKVWSLTSSSARTSNAKPFHSDGSALEVLFGQGG VSKGEELFTGVVP LVELDGDVNGHKFVSVEGEGDATYGKLT TKFICTTGKLPVPWPTLVTTFGYGLQCFAR YVPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKID FKEDGNILGHKLEYNYNHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQN TPIGDGPVLLPDNHLYSYQSALSADPNEKRDHMLLEFVTAAGITLGMDELYK GSGS HHHHHHHSAWSHPQFEKGGGGGGGGGSAWSHPQFEK*



HEK	Adgrd1	D1.04	<p>MEKLLRLCCWYSWLLLFYFNQVRG TNFAIAMQVVPLELARGHQVALSSISYVGC SL  SVLCLVATLVTFVAVLSSVSTIRNQRYHIHANLSFAVLVAQVLLISFRLEPGTTPCQVMA  VLLHYFFLSAFAWMLVEGLHLYSMVIKVFGESESKHRYYYGMGWGFPLLIISLSFA  MDSYGTSNNCWLSLASGAIWAFVAPALFVIVVNIIGILIAVTRVISQISADNYKIHGDPS  AFKLTAKAVAVLLPILGTSWVFGVLAVNGCAVVFQYMFATLNSLQGLFIFLHCLLNS  EVRAAFKHKTWVSLTSSSARTSNAKPFHSDLMNGTRPGMASTKLSPWDKSSSHAH  RVDLSAVGSALEVLFGQGGVSKGEELFTGVVPILVELDGDVNGHKFVSVEGEGDA  TYGKLT LKFICTTGKLPVPWPTLVTTFGYGLQCFARYPDHMKQHDFFKSAMPEGYVQ  ERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIM  ADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHLYSYQSALS  KDPNEKRDHMLLEFVTAAGITLGMDELYK GSGSHHHHHHHHSAWSHPQFEKGGGS  GGGGGGSAWSHPQFEK*</p>
HEK	Adgrd1	D1.05	<p>MEKGTLLVSPSQSGPGGDQPLLVKHR EDSAVVLSLIDTIDTVMGHVSSNLHGSTPQ  VTVEGSSAMAEFSVAKILPKTVNSSHYRFPAHGQSQFIQIPHEAFHRHAWSTVVGLLYH  SMHYLLNNIWPAHTKIAEAMHHQDCLLFATSHLISLEVSPPTLSQNLGSGPLITVHLK  HRLTRKQHSEATNSSNRVYVYCAFLDFSSGEGVWSNHGCALTRGNLTYSVCRCTHLT  NFAILMQVVPLELARGHQVALSSISYVGC SL SVLCLVATLVTFVAVLSSVSTIRNQRYHIH  ANLSFAVLVAQVLLISFRLEPGTTPCQVMAVLLHYFFLSAFAWMLVEGLHLYSMVIK  VFGSESESKHRYYYGMGWGFPLLIISLSFAMDSYGTSNNCWLSLASGAIWAFVAPA  LFVIVVNIIGILIAVTRVISQISADNYKIHGDPSAFKLTAKAVAVLLPILGTSWVFGVLAVN  GCAVVFQYMFATLNSLQGLFIFLHCLLNSVRAAFKHKTWVSLTSSSARTSNAKPF  HSDLMNGTRPGMASTKLSPWDKSSSHAHRVDLSAVGSALEVLFGQGGVSKGEELF  TGVPILVELDGDVNGHKFVSVEGEGDATY GKLTLKFICTTGKLPVPWPTLVTTFGY  GLQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNR  IELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD  HYQQNTPIGDGPVLLPDNHLYSYQSALS KDPNEKRDHMLLEFVTAAGITLGMDELY  K GSGSHHHHHHHHSAWSHPQFEKGGGGGGGGSAWSHPQFEK*</p>
HEK	Adgrd1	D1.06	<p>MHRVCFLSFQTRKQHSEATNSSNRVYVYCAFLDFSSGEGVWSNHGCALTRGNLTYS  VCRCTHTNFAILMQVVPLELARGHQVALSSISYVGC SL SVLCLVATLVTFVAVLSSVSTI  RNQRYHIHANLSFAVLVAQVLLISFRLEPGTTPCQVMAVLLHYFFLSAFAWMLVEGL  HLYSMVIKVFGESESESKHRYYYGMGWGFPLLIISLSFAMDSYGTSNNCWLSLASGAI  WAFVAPALFVIVVNIIGILIAVTRVISQISADNYKIHGDPSAFKLTAKAVAVLLPILGTSW  VFGVLAVNGCAVVFQYMFATLNSLQGLFIFLHCLLNSVRAAFKHKTWVSLTSSSA  RTSNAKPFHSDLMNGTRPGMASTKLSPWDKSSSHAHRVDLSAVGSALEVLFGQGG  GVSKGEELFTGVVPILVELDGDVNGHKFVSVEGEGDATY GKLTLKFICTTGKLPVPW  PTLVTTFGYGLQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVK  FEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIE  DGSVQLADHYQQNTPIGDGPVLLPDNHLYSYQSALS KDPNEKRDHMLLEFVTAAG  ITLGMDELYK GSGSHHHHHHHHSAWSHPQFEKGGGGGGGGSAWSHPQFEK*</p>

**Supp. Table 2:** Protein sequences for the Adgrg6 constructs in the pcDNA4-TO vector for HEK cell expression. Color code: yellow=signal peptide; grey=linker; light blue=3C protease cleavage site; green=YFP; magenta=His-tag; red=TwinStrep tag

Expression Host	Receptor	Construct name	Expressed protein Sequence
HEK	Adgrg6	G6.01	<p>MMFRSDRMWSCHWKWKPSLLFLFALYIMCVPHSVWG ANQILNLTADGQNL TSA  NITNIVEQVKRIVNKEENIDITL GSTLMNIFSNLSSSDSDLLESSEALKTIDELAFKIDL N  STSHVNITRNLALSVSLLPGTNAISNFSIGLPSNNESYFQMD FESGQV DPLASVILPP  NLL ENLSPEDSVL VRR A QFTFFNKTGLFQDVG PQRKTLVSYVMACSIGNITI QNLKDP  VQIKIKHTRTQEVHHPICAFWDLNKNK SFGGWNTSGCVAHRSDASET VCLCNHFT  HFGV LMDL PRSASQLDARNTKVLTFISYIGCGISAIFSAATLLTYVA FEKLR RDYPSKIL  MNLSTALLFLNLLFLLDGWITSFNVDGLCIAVAVLLHFFLLATFTWMGLEAIHMYIALV  KVFNTYIRRYILKFCIIGWGLPALVVS VVLASRNNNEVYGKESYGKEKGDEF CWIQDP  VIFYVT CAGYFGVMFFLNIA MFIVVMVQICGRNGKRSNRTLREEVLRNLR SVVSLTFL  LGMTWGF AFAFWG PLNIPFMYLFSIFNSLQGLFIFHFCAMKENVQKQWRQHLCCG  RFRLADNSDWSKTATNIIKSSDNLGKLS SSSIGSNSTYLTSKSKSSSTTYFKRNSHTD  NVSYEHSFNKSGSLRQC FHGQVLVKTGPCGSALEVL FQGGVSKGEELFTGVVPILV  ELDGDVNGHKFSVSGEGEGDATY GKLTLKFICTTGKLPVPWPTLVTTFGYGLQCFARY  PDHMKQHDFFKS AMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFK  EDGNILGHKLEYNYN SHNVYIMADKQKNGIKVNFKIRHNIEDG SVQLADHYQQNTPI  GDGPVLLPDNHYSYQSALS KDPNEKRDMVLEFVTAAGITLGMDELYK GSGSHH  HHHHHSAW SHPQFEKGGGGGGGGGSAW SHPQFEK*</p>
HEK	Adgrg6	G6.02	<p>MMFRSDRMWSCHWKWKPSLLFLFALYIMCVPHSVWG GVLMDL PRSASQLDAR  NTKVLTFISYIGCGISAIFSAATLLTYVA FEKLR RDYPSKILMNLSTALLFLNLLFLD GWI  TSFNVDGLCIAVAVLLHFFLLATFTWMGLEAIHMYIALV KVFNTYIRRYILKFCIIGWGL  PALVVS VVLASRNNNEVYGKESYGKEKGDEF CWIQDPVIFYVT CAGYFGVMFFLNIA  MFIVVMVQICGRNGKRSNRTLREEVLRNLR SVVSLTFLGMTWGF AFAFWG PLNIPF  MYLFSIFNSLQGLFIFHFCAMKENVQKQWRQHLCCGRFRLADNSDWSKTATNIIK  SSDNGSAL EVL FQGGVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATY G  KLT LKFICTTGKLPVPWPTLVTTFGYGLQCFARYPDHMKQHDFFKS AMPEGYVQERT  IFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYN SHNVYIMAD  KQKNGIKVNFKIRHNIEDG SVQLADHYQQNTPIGDGPVLLPDNHYSYQSALS KDPN  EKRDH MVLEFVTAAGITLGMDELYK GSGSHHHHHHSAW SHPQFEKGGGGGGG  GGSGGSAW SHPQFEK*</p>
HEK	Adgrg6	G6.03	<p>MMFRSDRMWSCHWKWKPSLLFLFALYIMCVPHSVWG THAGV LMDL PRSASQL  DARNTKVLTFISYIGCGISAIFSAATLLTYVA FEKLR RDYPSKILMNLSTALLFLNLLFLD  GWITSFNVDGLCIAVAVLLHFFLLATFTWMGLEAIHMYIALV KVFNTYIRRYILKFCIIG  WGLPALVVS VVLASRNNNEVYGKESYGKEKGDEF CWIQDPVIFYVT CAGYFGVMFF  LNIA MFIVVMVQICGRNGKRSNRTLREEVLRNLR SVVSLTFLGMTWGF AFAFWG P  LNIPFMYLFSIFNSLQGLFIFHFCAMKENVQKQWRQHLCCGRFRLADNSDWSKTAT  NIIKSSDNGSAL EVL FQGGVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGD  ATY GKLTLKFICTTGKLPVPWPTLVTTFGYGLQCFARYPDHMKQHDFFKS AMPEGYV  QERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYN SHNVYI</p>

			<p>MADKQKNGIKVNFKIRHNIEDGSQLADHYQQNTPIGDGPVLLPDNHLYSQSALS</p> <p>DPNEKRDHMLLEFVTAAGITLGMDELYKGGSHHHHHHHHSAWSHPQFEKGG</p> <p>SGGGGGGGSAWSHPQFEK*</p>
HEK	Adrg6	G6.04	<p>MMFRSDRMWSCHWKWKPSLLFLFALYIMCVPHSVVWGGVLMDLPRSASQLDAR</p> <p>NTKVLTFISYIGCGSAIFSAATLLTYVAFELRRDYPKILMNLSTALLFNLLFLLDGWI</p> <p>TSFNVDGLCIAVAVLLHFFLATFTWMGLEAIHMYIALVKVFNTYIRRYILKFCIIGWGL</p> <p>PALVSVVLASRNNNEVYGKESYGKEKGFECWIQDPVIFYVTCAGYFGVMFFLNIA</p> <p>MFIVVMVQICGRNGKRSNRTLREEVLRNLRVSVSLTFLGMTWGAFFAWGPLNIPF</p> <p>MYLFSIFNSLQGLFIFHCAMKENVQKQWRQHLCCGRFRLADNSDWSKTATNIKK</p> <p>SSDNLGKLSSSSIGSNSTYLTSKSKSSSTTYFKRNSHTDNVSYEHSFNKSGSLRQCFC</p> <p>QVLVKTGPCGSALEVLVQGPGGVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEG</p> <p>DATYGLTLKFICTTGKLPVPWPTLVTTFGYGLQCFARYPDHMKQHDFFKSAMPEGY</p> <p>VQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVY</p> <p>IMADKQKNGIKVNFKIRHNIEDGSQLADHYQQNTPIGDGPVLLPDNHLYSQSALS</p> <p>KDPNEKRDHMLLEFVTAAGITLGMDELYKGGSHHHHHHHHSAWSHPQFEKGG</p> <p>SGGGGGGGSAWSHPQFEK*</p>

**Supp. Table 3:** Protein sequences for the Adgrd1 constructs in the pAC8REDNK vector for insect cell expression. Color code: yellow=signal peptide; grey=linker; light blue=TEV protease cleavage site; green=GFP; magenta=His-tag; red=FLAG tag; residues with a orange letter are point mutations

Expression Host	Receptor	Construct name	Expressed protein Sequence
Insect cells	Adgrd1	D1.07	<p>MKTIIALS<b>YIFCLVFA</b>DYK<b>DDDD</b>ENLY<b>FQGL</b>TEERKTFQSPGVLSYLQNVLSLSPSKSLS  EQTALNLTKTLKAVGEILLPGWIALSEDSAVVLSLIDTIDTVMGHVSSNLHGSTPQV  TVEGSSAMAEFVAKILPKTVNSSHYRFPAHGQSFQIPHEAFHRHAWSTVVGLLYHS  MHYYLNNIWPAAHTKIAEAMHHQDCLLFATSHLISLEVSPPTLSQNLSGSPLITVHLK  HRLTRKQHSEATNSSNRVYVYCAFLDFSSGEGVWSNHGCALTRGNLTYSVCRCTHLT  NFAILMQVVPLELARGHQVALSSISYVGCSSLVCLVATLVTFVAVLSSVSTIRNQRYHIH  ANLSFAVLVAQVLLISFRLEPGTTPCQVMVAVLLHYFFLSAFWMLVEGLHLYSMVIK  VFGSEDSKHRYYYGMGWGFPLLICIISLFAMDSYGTSNNCWLSLASGAIWAFVAPA  LFVIVVNIGILIAVTRVISQISADNYKIHGDPSAFKLTAKAVAVLLPILGTSWVFGVLAVN  GCAVVFQYMFATLNSLQGLFIFLHCLLNSEVRAAFKHKTKVWLSLSSARTSNAKPF  HSDLMNGTRPGMASTKLSPWDKSSSAHRVDLSAVGSPAGENLY<b>FQGV</b>SKGEELFT  GVVPILVELDGDVNGHKFSVSGEGEGDATY<b>GKLT</b>LFICTTGKLPVPWPTLVTTFSY<b>G</b>  VQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRI  ELKGIDFKEDGNILGHKLEYNYNHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADH  YQNTPIGDGPVLLPDNHYLSTQSALS<b>KDPNEKRDH</b>MVLEFVTAAGITHGMDELY<b>K</b>  GGGGS<b>HHHHHHHH</b>*</p>
Insect cells	Adgrd1	D1.08	<p>MKTIIALS<b>YIFCLVFA</b>DYK<b>DDDD</b>ENLY<b>FQGL</b>TNFAILMQVVPLELARGHQVALSSISYVGC  SSLVCLVATLVTFVAVLSSVSTIRNQRYHIHANLSFAVLVAQVLLISFRLEPGTTPCQV  MAVLLHYFFLSAFWMLVEGLHLYSMVIKVFGESEDSKHRYYYGMGWGFPLLICIISL  FAMDSYGTSNNCWLSLASGAIWAFVAPALFVIVVNIGILIAVTRVISQISADNYKIHGD  PSAFKLTAKAVAVLLPILGTSWVFGVLAVNGCAVVFQYMFATLNSLQGLFIFLHCLLN  SEVRAAFKHKTKVWLSLSSARTSNAKPFHSDLMNGTRPGMASTKLSPWDKSSHS  AHRVDLSAVGSPAGENLY<b>FQGV</b>SKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGD  ATY<b>GKLT</b>LFICTTGKLPVPWPTLVTTFSY<b>G</b>VQCFSRYPDHMKQHDFFKSAMPEGYV  QERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNHNVYI  MADKQKNGIKVNFKIRHNIEDGSVQLADHYQNTPIGDGPVLLPDNHYLSTQSALS<b>K</b>  DPNEKRDH<b>MVLEFVTAAGITHGMDELYK</b>GGGGS<b>HHHHHHHH</b>*</p>
Insect cells	Adgrd1	D1.09	<p>MKTIIALS<b>YIFCLVFA</b>DYK<b>DDDD</b>ENLY<b>FQGGGGSGGG</b>SVPLELARGHQVALSSISYVGC  CSLSVCLVATLVTFVAVLSSVSTIRNQRYHIHANLSFAVLVAQVLLISFRLEPGTTPCQV  MAVLLHYFFLSAFWMLVEGLHLYSMVIKVFGESEDSKHRYYYGMGWGFPLLICIISL  FAMDSYGTSNNCWLSLASGAIWAFVAPALFVIVVNIGILIAVTRVISQISADNYKIHGD  PSAFKLTAKAVAVLLPILGTSWVFGVLAVNGCAVVFQYMFATLNSLQGLFIFLHCLLN  SEVRAAFKHKTKVWLSLSSARTSNAKPFHSDLMNGTRPGMASTKLSPWDKSSHS  AHRVDLSAVGSPAGENLY<b>FQGV</b>SKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGD  ATY<b>GKLT</b>LFICTTGKLPVPWPTLVTTFSY<b>G</b>VQCFSRYPDHMKQHDFFKSAMPEGYV  QERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNHNVYI  MADKQKNGIKVNFKIRHNIEDGSVQLADHYQNTPIGDGPVLLPDNHYLSTQSALS<b>K</b>  DPNEKRDH<b>MVLEFVTAAGITHGMDELYK</b>GGGGS<b>HHHHHHHH</b>*</p>

Insect cells	Adgrd1	D1.10	<p>MKTIIALS<b>YIFCLVFA</b>DYK<b>DDDD</b>ENLY<b>FQGL</b>TEERKTFQSPGVILSYLQNVLSLSPSKLS  EQTALNLTKTFLKAVGEILLPGWIALSEDSAVVLSLIDTIDTVMGHVSSNLHGSTPQV  TVEGSSAMA<del>EF</del>SVAKILPKTVNSSHYRFP<b>AHGQ</b>SFIQIPHEAFHRHAWSTVVGLLYHS  MHYYLNNIWPAHTKIAEAMHHQDCLLFATSHLISLEVSPPTLSQNLSGSPLITVHLK  HRLTRKQHSEATNSSNRV<b>FVYCA</b>FLDFSSGEGVWSNHGCALTRGNLTYSVCRCT<b>RLT</b>  NFAILMQVVPLELARGHQVALSSISYVGC<b>SLV</b>LCLVATLVTF<b>AVL</b>SSVSTIRNQRYHIH  ANLSFAVLVAQVLLISFRLEPGTTPCQVM<b>AVLLHY</b>FFLSAF<b>AWML</b>VEGLHLYSMVIK  VFGSEDSKHRYYYGMGWGF<b>PLLI</b>CIISL<b>SFAM</b>DSYGT<b>SNNC</b>WLSLASGAIWAFVAPA  LFVIVVNIGILIAVTRVISQISADNYKI<b>HGD</b>PSAFKLTAKAVAVLLPILGT<b>SWV</b>FGVLAVN  GCAV<b>VFQY</b>MFATLNSLQGLFIFL<b>FHCL</b>LNSEVRAAFKHKT<b>VW</b>SLTSS<b>SART</b>SN<b>AKPF</b>  HSDLMNGTRPGMA<b>STK</b>LSPWDKSSSAHRVDLSAV<b>GSPAG</b>ENLY<b>FQGV</b>SK<b>GEELF</b>  <b>GVV</b>PILVELDGDVNGHK<b>FVS</b>GEGEGDATY<b>GKLT</b>LK<b>FICTT</b>GKLPVPWPTLVTT<b>FSY</b>  <b>VQC</b>FSRYPDHMKQHDF<b>FKS</b>AMPEGYVQ<b>ERTIFF</b>KDDGNYK<b>TRAE</b>VK<b>FEGD</b>TLVNR<b>I</b>  <b>ELK</b>GIDFKEDGNILGHK<b>LEYN</b>YNSHN<b>VYI</b>MADKQK<b>NGIK</b>VNFKIRHN<b>IEDG</b>SVQLAD<b>H</b>  <b>YQ</b>QNTPIGDGPVLLPDNH<b>YLS</b>TQSALS<b>KDP</b>NEKRDHMV<b>LLE</b>FVTAAGITHGM<b>DELY</b>K  GGGG<b>SHHHHHHHH</b>*</p>
Insect cells	Adgrd1	D1.11	<p>MKTIIALS<b>YIFCLVFA</b>DYK<b>DDDD</b>ENLY<b>FQGL</b>TEERKTFQSPGVILSYLQNVLSLSPSKLS  EQTALNLTKTFLKAVGEILLPGWIALSEDSAVVLSLIDTIDTVMGHVSSNLHGSTPQV  TVEGSSAMA<del>EF</del>SVAKILPKTVNSSHYRFP<b>AHGQ</b>SFIQIPHEAFHRHAWSTVVGLLYHS  MHYYLNNIWPAHTKIAEAMHHQDCLLFATSHLISLEVSPPTLSQNLSGSPLITVHLK  HRLTRKQHSEATNSSNRV<b>FVYCA</b>FLDFSSGEGVWSNHGCALTRGNLTYSVCRCTHL<b>G</b>  GGGGGGG<b>SV</b>PLELARGHQVALSSISYVGC<b>SLV</b>LCLVATLVTF<b>AVL</b>SSVSTIRNQRYHI  HANLSFAVLVAQVLLISFRLEPGTTPCQVM<b>AVLLHY</b>FFLSAF<b>AWML</b>VEGLHLYSMVI  KVFGSEDSKHRYYYGMGWGF<b>PLLI</b>CIISL<b>SFAM</b>DSYGT<b>SNNC</b>WLSLASGAIWAFVAP  ALFVIVVNIGILIAVTRVISQISADNYKI<b>HGD</b>PSAFKLTAKAVAVLLPILGT<b>SWV</b>FGVLAV  NGCAV<b>VFQY</b>MFATLNSLQGLFIFL<b>FHCL</b>LNSEVRAAFKHKT<b>VW</b>SLTSS<b>SART</b>SN<b>AKPF</b>  FHS<b>DL</b>MNGTRPGMA<b>STK</b>LSPWDKSSSAHRVDLSAV<b>GSPAG</b>ENLY<b>FQGV</b>SK<b>GEELF</b>  <b>TGV</b>VPIVELDGDVNGHK<b>FVS</b>GEGEGDATY<b>GKLT</b>LK<b>FICTT</b>GKLPVPWPTLVTT<b>FSY</b>  <b>VQC</b>FSRYPDHMKQHDF<b>FKS</b>AMPEGYVQ<b>ERTIFF</b>KDDGNYK<b>TRAE</b>VK<b>FEGD</b>TLVNR<b>I</b>  <b>ELK</b>GIDFKEDGNILGHK<b>LEYN</b>YNSHN<b>VYI</b>MADKQK<b>NGIK</b>VNFKIRHN<b>IEDG</b>SVQLAD  <b>HY</b>QQNTPIGDGPVLLPDNH<b>YLS</b>TQSALS<b>KDP</b>NEKRDHMV<b>LLE</b>FVTAAGITHGM<b>DEL</b>  <b>YK</b>GGGG<b>SHHHHHHHH</b>*</p>
Insect cells	Adgrd1	D1.12	<p>MKTIIALS<b>YIFCLVFA</b>TNFAILMQVVPLELARGHQVALSSISYVGC<b>SLV</b>LCLVATLVTF<b>AVL</b>  VSSVSTIRNQRYHIHANLSFAVLVAQVLLISFRLEPGTTPCQVM<b>AVLLHY</b>FFLSAF<b>A</b>  WMLVEGLHLYSMVIK<b>V</b>FGSEDSKHRYYYGMGWGF<b>PLLI</b>CIISL<b>SFAM</b>DSYGT<b>SNNC</b>  WLSLASGAIWAFVAPALFVIVVNIGILIAVTRVISQISADNYKI<b>HGD</b>PSAFKLTAKAVAV  LLPILGT<b>SWV</b>FGVLAVNGCAV<b>VFQY</b>MFATLNSLQGLFIFL<b>FHCL</b>LNSEVRAAFKHKT  VW<b>SL</b>TSS<b>SART</b>SN<b>AKPF</b>HSDLMNGTRPGMA<b>STK</b>LSPWDKSSSAHRVDLSAV<b>DYKD</b>  <b>DDDK</b>*</p>
Insect cells	Adgrd1	D1.13	<p>MKTIIALS<b>YIFCLVFA</b>DYK<b>DDDDA</b>ENLY<b>FQGG</b>GGGGGGG<b>SV</b>PLELARGHQVALSSISYV  GC<b>SLV</b>LCLVATLVTF<b>AVL</b>SSVSTIRNQRYHIHANLSFAVLVAQVLLISFRLEPGTTPCQ  VM<b>AVLLHY</b>FFLSAF<b>AWML</b>VEGLHLYSMVIK<b>V</b>FGSEDSKHRYYYGMGWGF<b>PLLI</b>CIISL  <b>SFAM</b>DSYGT<b>SNNC</b>WLSLASGAIWAFVAPALFVIVVNIGILIAVTRVISQISADNYKI<b>HGD</b>  <b>DP</b>SAFKLTAKAVAVLLPILGT<b>SWV</b>FGVLAVNGCAV<b>VFQY</b>MFATLNSLQGLFIFL<b>FHCL</b></p>

			LNSEVRAAFKHKTQVWVSLTSSARTSNAKPFHSDLMNGTRPGMASTKLSPWDKSSH SAHRVDLSAV*
<b>Insect cells</b>	<b>Adgrd1</b>	<b>D1.14</b>	<b>MKTIIALSIFCLVFA</b> <b>DYKDDDDA</b> <b>ENLYFQGL</b> LTEERKTFQSPGVLSYLQNVLSLPSKSL SEQTALNLTkTFLKAVGEILLPGWIALSEDSAVVLSLIDTIDTVMGHVSSNLHGSTPQ VTVEGSSAMAEFVAKILPKTVNSSHYRFP AHGQSFQIPHEAFHRHAWSTVVGLLYH SMHYLLNNIWPAHTKIAEAMHHQDCLLFATSHLISLEVSPPTLSQNLGSPPLITVHLK HRLTRKQHSEATNSSNRV FVYCAFLDFSSGEGVWSNHGCALTRGNLTYSVCRCT <b>ALA</b> NFAILMQVVPLELARGHQAVALSSISYVGC SLSVLCLVATLVTFAVLSSVSTIRNQRYHIH ANLSFAVLVAQVLLISFRLEPGTTPCQVMALLHYFFLSAFAWMLVEGLHLYSMVIK VFGSEDSKHRYYYGMGWGFPLLCIISL FAMD SYGTSNNCWLSLASGAIWAFVAPA LFVIVVNIGILIAVTRVISQISADNYKIHGDPSAFKLTAKAVAVLLPILGTSWVFGVLAVN GCAVVFQYMFATLNSLQGLFIFLFHCLLNSEVRAAFKHKTQVWVSLTSSARTSNAKPF HSDLMNGTRPGMASTKLSPWDKSSHSAHRVDLSAV*
<b>Insect cells</b>	<b>Adgrd1</b>	<b>D1.15</b>	<b>MKTIIALSIFCLVFA</b> <b>DYKDDDDA</b> <b>ENLYFQGL</b> VYSRQDHPGFQVLASASHYWPLENV DGIHELQD TTGDIVEGKVNKGIYLKEEKGVTL LLYGRYNSSCISKPEQCGPEGVTFSSFF WKTQGEQSRPIPSAYGGQVISNGFKVCSGGGRGSVELYTRDNSMTWEASFPPGPY WTHVLF TWKSKEGLKVYVNGTLSTSDPSGKVS RDYGESNVNLVIGSEQDQAKCYEN GAFDEFIWERALTPDEIAMYFTA AIGKHALLSSTLPSLFMTSTASPVMPTDAYHPIITN LTEERKTFQSPGVLSYLQNVLSLPSKSLSEQTALNLTkTFLKAVGEILLPGWIALSED SAVVLSLIDTIDTVMGHVSSNLHGSTPQVTVEGSSAMAEFVAKILPKTVNSSHYRFP AHGQSFQIPHEAFHRHAWSTVVGLLYHSMHYLLNNIWPAHTKIAEAMHHQDCLLF ATSHLISLEVSPPTLSQNLGSPPLITVHLK HRLTRKQHSEATNSSNRV FVYCAFLDFSS GEGVWSNHGCALTRGNLTYSVCRCTHLTNFAILMQVVPLELARGHQAVALSSISYVGC SLSVLCLVATLVTFAVLSSVSTIRNQRYHIHANLSFAVLVAQVLLISFRLEPGTTPCQV MAVLLHYFFLSAFAWMLVEGLHLYSMVIK VFGSEDSKHRYYYGMGWGFPLLCIISL FAMD SYGTSNNCWLSLASGAIWAFVAPALFVIVVNIGILIAVTRVISQISADNYKIHGD PSAFKLTAKAVAVLLPILGTSWVFGVLAVNGCAVVFQYMFATLNSLQGLFIFLFHCLL NSEVRAAFKHKTQVWVSLTSSARTSNAKPFHSDLMNGTRPGMASTKLSPWDKSSH AHRVDLSAV*

**Supp. Table 4:** Protein sequences for the Adgrg6 constructs in the pAC8REDNK vector for insect cell expression. Color code: yellow=signal peptide; grey=linker; light blue=3C protease cleavage site; green=YFP; magenta=His-tag; red=FLAG tag; residues written in orange letters are point mutations

Expression Host	Receptor	Construct name	Expressed protein Sequence
Insect cells	Adgrg6	G6.05	<p>MKTIIALSIFCLVFA<sup>DYKDDDD</sup>ENLYFQGPVDISNCLKEANEVANQILNLTADGQNLTSANITNIVEQVKRIVNKEENIDITLGSTLMNIFSNILSSSDSDLLESSEALKTIDELAFKIDLNSTSHVNITTRNLALSVSLLPGTNAISNFSIGLPSNNESYFQMDFESGQVDPLASVILPPNLLLENLSPEDSVLVRRAQFTFFNKTGLFQDVGPPQRKTLVSYVMACSIGNITIQNLKDPVQIKIKHTRTQEVHHPICAFWDLNKNKSFGGWNTSGCVAHRDSDASETVCLCNHFTHFGLMDLPRASQLDARNTKVLTFISYIGCGISAIFSAATLLTYVAFEKLRRDYPSKILMNLSTALLFLNLLFLLDGWITSFNVDGLCIAVAVLLHFFLLATFTWMGLEAIHMYIALVKVFNTYIRRYILKFCIIGWGLPALVVSVVLASRNNNEVYGKESYGKEKGEFCWIQDPVIFVYTCAGYFGVMFFLNIAMFIVVMVQICGRNGKRSNRTLREEVLRNLRNLSVSLTFLLGMTWGF<sup>FAFFAWGPLNIPFMYLFSIFNSLQGLFIFIHFCAMKENVQKQWRQHLCCGRFRLADNSDWSKTATNIIKSSDNLGKSLSSSSIGSNSTYLTSKSKSSSTTYFKRNSHTDNVSYEHSFNKSGSLRQCFCFHGQVLVKTGPCGSPAG</sup>ENLYFQGVSKGEELFTGVVPILVELDGDVNGHKFSVS<sup>GEGEGDATY</sup>GKLT<sup>LKFI</sup>CTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYN<sup>SHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALS</sup>KDPNEKRDHMV<sup>LLEFVTAAGITHGMD</sup>ELYKGGGSHHHHHHHH*</p>
Insect cells	Adgrg6	G6.06	<p>MKTIIALSIFCLVFA<sup>DYKDDDD</sup>ENLYFQGT<sup>HFGVLM</sup>DLPRASQLDARNTKVLTFISYIGCGISAIFSAATLLTYVAFEKLRRDYPSKILMNLSTALLFLNLLFLLDGWITSFNVDGLCIAVAVLLHFFLLATFTWMGLEAIHMYIALVKVFNTYIRRYILKFCIIGWGLPALVVSVVLASRNNNEVYGKESYGKEKGEFCWIQDPVIFVYTCAGYFGVMFFLNIAMFIVVMVQICGRNGKRSNRTLREEVLRNLRNLSVSLTFLLGMTWGF<sup>FAFFAWGPLNIPFMYLFSIFNSLQGLFIFIHFCAMKENVQKQWRQHLCCGRFRLADNSDWSKTATNIIKSSDNLGKSLSSSSIGSNSTYLTSKSKSSSTTYFKRNSHTDNVSYEHSFNKSGSLRQCFCFHGQVLVKTGPCG</sup>SPAGENLYFQGVSKGEELFTGVVPILVELDGDVNGHKFSVS<sup>GEGEGDATY</sup>GKLT<sup>LKFI</sup>CTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYN<sup>SHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALS</sup>KDPNEKRDHMV<sup>LLEFVTAAGITHGMD</sup>ELYKGGGSHHHHHHHH*</p>
Insect cells	Adgrg6	G6.07	<p>MKTIIALSIFCLVFA<sup>DYKDDDD</sup>ENLYFQGGSGGGSGGGSSASQLDARNTKVLTFISYIGCGISAIFSAATLLTYVAFEKLRRDYPSKILMNLSTALLFLNLLFLLDGWITSFNVDGLCIAVAVLLHFFLLATFTWMGLEAIHMYIALVKVFNTYIRRYILKFCIIGWGLPALVVSVVLASRNNNEVYGKESYGKEKGEFCWIQDPVIFVYTCAGYFGVMFFLNIAMFIVVMVQICGRNGKRSNRTLREEVLRNLRNLSVSLTFLLGMTWGF<sup>FAFFAWGPLNIPFMYLFSIFNSLQGLFIFIHFCAMKENVQKQWRQHLCCGRFRLADNSDWSKTATNIIKSSDNLGKSLSSSSIGSNSTYLTSKSKSSSTTYFKRNSHTDNVSYEHSFNKSGSLRQCFCFHGQVLVKTGPCG</sup>SPAGENLYFQGVSKGEELFTGVVPILVELDGDVNGHKFSVS<sup>GEGEGDATY</sup>GKLT<sup>LKFI</sup>CTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDD</p>



			<p> <b>G</b>NYKTRAEVKFEGD<b>T</b>LVNRIELK<b>G</b>IDFKEDGNILGHKLEYNYN<b>S</b>HN<b>V</b>YIMADKQKNGI  <b>K</b>VNFKIRHN<b>I</b>EDGS<b>V</b>QLADHYQ<b>Q</b>NTPIGDGPVLLPD<b>N</b>HYL<b>S</b>TQSALS<b>K</b>DPNEKR<b>D</b>H  <b>M</b>VLL<b>E</b>FVTAAGITHGM<b>D</b>ELYKGGG<b>S</b>HHHHHH<b>H</b>* </p>
<b>Insect cells</b>	<b>Adgrg6</b>	<b>G6.08</b>	<p> <b>M</b>KTI<b>I</b>ALS<b>I</b>YIFCL<b>V</b>FA<b>D</b>Y<b>K</b>DD<b>D</b>ENLY<b>F</b>Q<b>G</b>PVDISNCLKEANEVANQILNLTADGQNL<b>T</b>  SANITNIVEQVKRIVNKEENIDITLGSTLMNIFSNILSSSDSDLLESSEALKTIDELAFKID  LNSTSHVNITTRNLALS<b>V</b>SS<b>L</b>PGTNAISNFSIGLPSNNESYFQ<b>M</b>DFESGQ<b>V</b>DPLAS<b>V</b>IL  PPNLLENLSPEDSVLVRRAQFTFFNK<b>T</b>GLFQDVGPQRK<b>T</b>LVS<b>V</b>YMACS<b>I</b>GNITIQNLK  DPVQIKIKH<b>T</b>R<b>T</b>QEVHHPICAFWDLNKNK<b>S</b>FGGWNTSGCVAHRSDA<b>S</b>ETVCLCN<b>H</b>  <b>F</b>AHFGVLMDLPR<b>S</b>ASQLDAR<b>N</b>TKV<b>L</b>TFISYIGCGISAIFSAATLLTYVAF<b>E</b>KLR<b>R</b>DYPSKI  LMNLSTALLFLNLLFLLDGWIT<b>S</b>FNVDGLCIAVAVLLHFF<b>L</b>LATFTWMGLEA<b>I</b>HMYIAL  VKVFNTYIRRYILK<b>F</b>CIIGWGLPALV<b>S</b>VVLASRNNNEVY<b>G</b>KESY<b>G</b>KE<b>G</b>DEFCW<b>I</b>QD  PVIFV<b>T</b>CAGYFGVMFFLN<b>I</b>AMFIVMV<b>Q</b>ICGRNGKR<b>S</b>NRTLREEVLRN<b>L</b>RSV<b>S</b>LT<b>F</b>  LLGMTWGF<b>A</b>FA<b>F</b>AWG<b>P</b>LNIPFMYL<b>S</b>IFNS<b>L</b>QGLFIF<b>I</b>FCAMKEN<b>V</b>QKQWR<b>Q</b>HLCC  GRFRLADNSD<b>S</b>KTATN<b>I</b>IKSSD<b>N</b>L<b>G</b>KSLS<b>S</b>SSIGSN<b>S</b>T<b>L</b>TSKS<b>S</b>SSTTYFKR<b>S</b>HT  DNV<b>S</b>YEHSFNKSGSLR<b>Q</b>CFHGQVL<b>V</b>K<b>T</b>G<b>P</b>CGSPAGENLY<b>F</b>Q<b>G</b>V<b>S</b>KGEEL<b>F</b>TGV<b>V</b>PI<b>L</b>  <b>V</b>ELDGDV<b>N</b>GHK<b>F</b>S<b>V</b>SGEGEGDATY<b>G</b>KL<b>L</b>KF<b>I</b>CT<b>T</b>G<b>K</b>L<b>P</b>VP<b>W</b>PT<b>L</b>VT<b>T</b>FSY<b>G</b>V<b>Q</b>CF<b>S</b>R  <b>Y</b>PDH<b>M</b>K<b>Q</b>H<b>D</b>FF<b>K</b>SAM<b>P</b>EGY<b>V</b>Q<b>E</b>R<b>T</b>IF<b>F</b>K<b>D</b>D<b>G</b>NYK<b>T</b>RAEV<b>K</b>FEGD<b>T</b>LVNRIELK<b>G</b>ID<b>F</b>  <b>K</b>EDGNILGHKLEYNYN<b>S</b>HN<b>V</b>YIMADKQKNGIKVNFKIRHN<b>I</b>EDGS<b>V</b>QLADHYQ<b>Q</b>NT  <b>P</b>IGDGPVLLPD<b>N</b>HYL<b>S</b>TQSALS<b>K</b>DPNEKR<b>D</b>H<b>M</b>VLL<b>E</b>FVTAAGITHGM<b>D</b>ELYKGGG<b>S</b>  <b>H</b>HHHHHH<b>H</b>* </p>
<b>Insect cells</b>	<b>Adgrg6</b>	<b>G6.09</b>	<p> <b>M</b>KTI<b>I</b>ALS<b>I</b>YIFCL<b>V</b>FA<b>D</b>Y<b>K</b>DD<b>D</b>ENLY<b>F</b>Q<b>G</b>PVDISNCLKEANEVANQILNLTADGQNL<b>T</b>  SANITNIVEQVKRIVNKEENIDITLGSTLMNIFSNILSSSDSDLLESSEALKTIDELAFKID  LNSTSHVNITTRNLALS<b>V</b>SS<b>L</b>PGTNAISNFSIGLPSNNESYFQ<b>M</b>DFESGQ<b>V</b>DPLAS<b>V</b>IL  PPNLLENLSPEDSVLVRRAQFTFFNK<b>T</b>GLFQDVGPQRK<b>T</b>LVS<b>V</b>YMACS<b>I</b>GNITIQNLK  DPVQIKIKH<b>T</b>R<b>T</b>QEVHHPICAFWDLNKNK<b>S</b>FGGWNTSGCVAHRSDA<b>S</b>ETVCLCN<b>H</b>  <b>F</b>GGG<b>S</b>GGG<b>S</b>GGG<b>S</b>ASQLDAR<b>N</b>TKV<b>L</b>TFISYIGCGISAIFSAATLLTYVAF<b>E</b>KLR<b>R</b>DYPSK  ILMN<b>L</b>STALLFLNLLFLLDGWIT<b>S</b>FNVDGLCIAVAVLLHFF<b>L</b>LATFTWMGLEA<b>I</b>HMYIAL  VKVFNTYIRRYILK<b>F</b>CIIGWGLPALV<b>S</b>VVLASRNNNEVY<b>G</b>KESY<b>G</b>KE<b>G</b>DEFCW<b>I</b>QD  PVIFV<b>T</b>CAGYFGVMFFLN<b>I</b>AMFIVMV<b>Q</b>ICGRNGKR<b>S</b>NRTLREEVLRN<b>L</b>RSV<b>S</b>LT<b>F</b>  LLGMTWGF<b>A</b>FA<b>F</b>AWG<b>P</b>LNIPFMYL<b>S</b>IFNS<b>L</b>QGLFIF<b>I</b>FCAMKEN<b>V</b>QKQWR<b>Q</b>HLCC  GRFRLADNSD<b>S</b>KTATN<b>I</b>IKSSD<b>N</b>L<b>G</b>KSLS<b>S</b>SSIGSN<b>S</b>T<b>L</b>TSKS<b>S</b>SSTTYFKR<b>S</b>HT  DNV<b>S</b>YEHSFNKSGSLR<b>Q</b>CFHGQVL<b>V</b>K<b>T</b>G<b>P</b>CGSPAGENLY<b>F</b>Q<b>G</b>V<b>S</b>KGEEL<b>F</b>TGV<b>V</b>PI<b>L</b>  <b>V</b>ELDGDV<b>N</b>GHK<b>F</b>S<b>V</b>SGEGEGDATY<b>G</b>KL<b>L</b>KF<b>I</b>CT<b>T</b>G<b>K</b>L<b>P</b>VP<b>W</b>PT<b>L</b>VT<b>T</b>FSY<b>G</b>V<b>Q</b>CF<b>S</b>R  <b>Y</b>PDH<b>M</b>K<b>Q</b>H<b>D</b>FF<b>K</b>SAM<b>P</b>EGY<b>V</b>Q<b>E</b>R<b>T</b>IF<b>F</b>K<b>D</b>D<b>G</b>NYK<b>T</b>RAEV<b>K</b>FEGD<b>T</b>LVNRIELK<b>G</b>ID<b>F</b>  <b>K</b>EDGNILGHKLEYNYN<b>S</b>HN<b>V</b>YIMADKQKNGIKVNFKIRHN<b>I</b>EDGS<b>V</b>QLADHYQ<b>Q</b>NT  <b>P</b>IGDGPVLLPD<b>N</b>HYL<b>S</b>TQSALS<b>K</b>DPNEKR<b>D</b>H<b>M</b>VLL<b>E</b>FVTAAGITHGM<b>D</b>ELYKGGG<b>S</b>  <b>H</b>HHHHHH<b>H</b>* </p>
<b>Insect cells</b>	<b>Adgrg6</b>	<b>G6.10</b>	<p> <b>M</b>KTI<b>I</b>ALS<b>I</b>YIFCL<b>V</b>FA<b>T</b>HFGVLMDLPR<b>S</b>ASQLDAR<b>N</b>TKV<b>L</b>TFISYIGCGISAIFSAATLLTY  VAF<b>E</b>KLR<b>R</b>DYPSKI<b>M</b>N<b>L</b>STALLFLNLLFLLDGWIT<b>S</b>FNVDGLCIAVAVLLHFF<b>L</b>LATFT  <b>W</b>MGLEA<b>I</b>HMYIALVKVFNTYIRRYILK<b>F</b>CIIGWGLPALV<b>S</b>VVLASRNNNEVY<b>G</b>KESY  <b>G</b>KE<b>G</b>DEFCW<b>I</b>QDPVIFV<b>T</b>CAGYFGVMFFLN<b>I</b>AMFIVMV<b>Q</b>ICGRNGKR<b>S</b>NRTLRE  <b>E</b>VL<b>R</b>N<b>L</b>RSV<b>S</b>LT<b>F</b>LLGMTWGF<b>A</b>FA<b>F</b>AWG<b>P</b>LNIPFMYL<b>S</b>IFNS<b>L</b>QGLFIF<b>I</b>FCAM<b>K</b>  <b>N</b>VQKQWR<b>Q</b>HLCCGRFRLADNSD<b>S</b>KTATN<b>I</b>IKSSD<b>N</b>L<b>G</b>KSLS<b>S</b>SSIGSN<b>S</b>T<b>L</b>TSK<b>S</b>  <b>K</b>SSSTTYFKR<b>S</b>HTDNV<b>S</b>YEHSFNKSGSLR<b>Q</b>CFHGQVL<b>V</b>K<b>T</b>G<b>P</b>C<b>D</b>Y<b>K</b>DD<b>D</b><b>K</b>* </p>

<p><b>Insect cells</b></p>	<p><b>Adgrg6</b></p>	<p><b>G6.11</b></p>	<p>MKTIIALS<b>YIFCLVFA</b>DYK<b>DDDDA</b>ENLY<b>FQG</b>GGSGGGGGGSASQLDARNTKVLTFIS  YIGCGISAIFSAATLLTYVAFEKLRDYPISKILMNLSTALLFLNLLFLLDGWITSFNVDGL  CIAVAVLLHFFLLATFTWMGLEAIHMYIALVKVFNTYIRRYILKFCIIGWGLPALVVSVV  LASRNNNEVYGKESYGKEKGDEFWCWIDPVIYVTCAGYFGVMFFLNIAMFIVVMV  QICGRNGKRSNRTLREEVLRNLRVSVSLTFLGMTWGF<del>FAFFAWG</del>PLNIPFMYLFSIFN  SLQGLFIFIFHCAMKENVQKQWRQHLCCGRFRLADNSDWSKTATNIIKSSDNLGKS  LSSSIGSNSTYLTSKSKSSSTTYFKRNSHTDNVSYEHSFNKSGSLRQC<b>FHGQVLVKTGP</b>  C*</p>
<p><b>Insect cells</b></p>	<p><b>Adgrg6</b></p>	<p><b>G6.12</b></p>	<p>MKTIIALS<b>YIFCLVFA</b>DYK<b>DDDDA</b>ENLY<b>FQG</b>PVDISNCLKEANEVANQILNLTADGQN  LTSANITNIVEQVKRIVNKEENIDITLGSTLMNIFSNLSSSDSDLESSEALKTIDELAFK  IDLNSTSHVNITRNLALS<del>VSSLLPGTNAISNF</del>SIGLPSNNESYFQMDFESGQVDPLAS  VILPPNLENLSPEDSVLVRRAQFTFFNKTLGFQDVGPQRKTLVSYVMACSIGNITIQN  LKDPVQIKIKHTRTQEVHHPICAFWDLNKNKSF<del>GGWNTSGCVAHR</del>SDASETVCCLC  NA<b>F</b>AHF<del>GV</del>LMDLPRASQLDARNTKVLTFISYIGCGISAIFSAATLLTYVAFEKLRDYP  SKILMNLSTALLFLNLLFLLDGWITSFNVDGLCIAVAVLLHFFLLATFTWMGLEAIHMYI  ALVKVFNTYIRRYILKFCIIGWGLPALVVS<del>VV</del>LASRNNNEVYGKESYGKEKGDEFWCWIDP  DPVIYVTCAGYFGVMFFLNIAMFIVVMVQICGRNGKRSNRTLREEVLRNLRVSVSLT  FLLGMTWGF<del>FAFFAWG</del>PLNIPFMYLFSIFNSLQGLFIFIFHCAMKENVQKQWRQHL  CGRFRLADNSDWSKTATNIIKSSDNLGKSLSSSIGSNSTYLTSKSKSSSTTYFKRNSH  TDNVSYEHSFNKSGSLRQC<b>FHGQVLVKTGPC</b>*</p>

**Supp. Table 5:** Protein sequences for the Adgrd1-mini G $\alpha$ s fusion protein constructs in the pAC8REDNK vector for insect cell expression. Color code: yellow=signal peptide; grey=linker; light blue=3C/TEV protease cleavage site; green=mini-G $\alpha$ s 399; red=FLAG tag; residues written in orange letters are point mutations

Expression Host	Receptor	Construct name	Expressed protein Sequence
Insect cells	Adgrd1 – mini-Gs	D1-mGs.01	MKTIIALSIFCLVFA <sup>DYKDDDDKDYKDDDDK</sup> ENLYFQG <sup>TN</sup> FAILMQVVPLELARGH QVALSSISYVGCSSLVCLVATLVTFAVLSSVSTIRNQRYHIHANLSFAVLVAQVLLISF RLEPGTTPCQVMAVLLHYFFLSAFWMLVEGLHLYSMVIKVFGESESKHRYYYGMG WGFPLLICISLSFAMDSYGTSNNCWLSLASGAIWAFVAPALFVIVNIGILIAVTRVIS QISADNYKIHGDPSAFKLTAKAVAVLLPILGTSWVFGVLAVNGCAVVFQYMFATLNSL QGLFIFLFHCLLNSEVRAAFKHKTQVWSLTSSSARTSNAKPFHSDLMNGTRPGMAST KLSPWDKSSSAHRVDLSAV <sup>DYKDDDDK</sup> GCLGNSKTEDQRNEEKAQREANKKIEKQ LQKDKQVYRATHRLLLLGADNSGKSTIVKQMRILHGGSGGGTSGIFETKFQVDKVF NFMFDVGGQRDERRKWIQCFNDVTAIFVVDSSDYNRLQEALNDFKSIWNNRWL RTISVILFNKQDLLAEKVLGKSKIEDYFPEFARYTTPEDATPEPGEDPRVTRAKYFIRD EFLRISTASGDGRHYCYPHFTCAVDTENARRIFNDCRDIIQRMHLRQYELL*
Insect cells	Adgrd1 – mini-Gs	D1-mGs.02	MKTIIALSIFCLVFA <sup>DYKDDDDKDYKDDDDK</sup> ENLYFQG <sup>TN</sup> FAILMQVVPLELARGH QVALSSISYVGCSSLVCLVATLVTFAVLSSVSTIRNQRYHIHANLSFAVLVAQVLLISF RLEPGTTPCQVMAVLLHYFFLSAFWMLVEGLHLYSMVIKVFGESESKHRYYYGMG WGFPLLICISLSFAMDSYGTSNNCWLSLASGAIWAFVAPALFVIVNIGILIAVTRVIS QISADNYKIHGDPSAFKLTAKAVAVLLPILGTSWVFGVLAVNGCAVVFQYMFATLNSL QGLFIFLFHCLLNSEVRAAFKHKTQVWSLTSSSARTSNAKPFHSDLMNGTRPGMAST KLSPWDKSSSAHRVDLSAV <sup>ENLYFQGG</sup> CLGNSKTEDQRNEEKAQREANKKIEKQLQ KDKQVYRATHRLLLLGADNSGKSTIVKQMRILHGGSGGGTSGIFETKFQVDKVN FMFDVGGQRDERRKWIQCFNDVTAIFVVDSSDYNRLQEALNDFKSIWNNRWLRTI SVILFNKQDLLAEKVLGKSKIEDYFPEFARYTTPEDATPEPGEDPRVTRAKYFIRDEFL RISTASGDGRHYCYPHFTCAVDTENARRIFNDCRDIIQRMHLRQYELL*
Insect cells	Adgrd1 – mini-Gs	D1-mGs.03	MKTIIALSIFCLVFA <sup>DYKDDDDADYKDDDDK</sup> TNFAILMQVVPLELARGHQVALSSISY VGCSSLVCLVATLVTFAVLSSVSTIRNQRYHIHANLSFAVLVAQVLLISFRLEPGTTPC QVMAVLLHYFFLSAFWMLVEGLHLYSMVIKVFGESESKHRYYYGMGWGFPLLICII SLSFAMDSYGTSNNCWLSLASGAIWAFVAPALFVIVNIGILIAVTRVISQISADNYKIH GDPSAFKLTAKAVAVLLPILGTSWVFGVLAVNGCAVVFQYMFATLNSLQGLFIFLFH CLLNSEVRAAFKHKTQVWSLTSSSARTSNAKPFHSDLMNGTRPGMASTKLSPWDKSS HSAHRVDLSAV <sup>ENLYFQGG</sup> CLGNSKTEDQRNEEKAQREANKKIEKQLQKDKQVYRAT HRLLLLGADNSGKSTIVKQMRILHGGSGGGTSGIFETKFQVDKVNFMFDVGGQ RDERRKWIQCFNDVTAIFVVDSSDYNRLQEALNDFKSIWNNRWLRTISVILFNKQD LLAEKVLGKSKIEDYFPEFARYTTPEDATPEPGEDPRVTRAKYFIRDEFLRISTASGDG RHYCYPHFTCAVDTENARRIFNDCRDIIQRMHLRQYELL*
Insect cells	Adgrd1 – mini-Gs	D1-mGs.04	MKTIIALSIFCLVFA <sup>DYKDDDDADYKDDDDKDYKDDDDADYKDDDDK</sup> LEVLFGQPA YHPIITNLTEERKTFQSPGVLSYLQNVLSLPSKSLSEQTALNLTKTKAVGEILLPGW IALSESAVVLSLIDTIDTVMGHVSSNLHGSTPQVTEGSSAMAEFVAKILPKTVNSS HYRFPAGQSFQIPHEAFHRHAWSTVVGLLYHSMHYLNNIIPAHTKIAEAMHH

			<p>QDCLLFATSHLISLEVSPPTLSQNLGSPPLITVHLKHRLTRKQCEATNSSNRVVFYCA  FLDFSSGEGVWSNHGCALTRGNLTYSVCRCTHLTNFAILMQVVPLELARGHQVALSS  ISYVGCSLSVLCLVATLVTFVAVLSSVSTIRNQRYHIHANLSFAVLVAQVLLISFRLEPGTT  PCQVMVALLHYFFLSAFWMLVEGLHLYSMVIKVFGESESKHRYYYGMGWGFPLLI  CIISLAFAMDSYGTSNCWLSLASGAIWAFVAPALFVIVVIGILIAVTRVISQISADNY  KIHGDPSAFKLTAKAVAVLLPILGTSWVFGVLAVNGCAVVFQYMFATLNSLQGLFIFL  FHCLLNSEVRAAFKHKTWVSLTSSARTSNAKPFHSDLMNGTRPGMASTKLSPWD  KSSSAHRVDLSAVGGGGGGGGGGGGGGSEENLYFQGLGNSKTEDQRNEEKAQRE  ANKKIEKQLQKDKQVYRATHRLLLLGADNSGKSTIVKQMRILHGGSGGGSGTSGIFET  KFQVDKVNFMFVGGQRDERRKWIQCFNDVTAIIFVVDSSDYNRLQEALNDFKSI  WNNRWLRTISVILFNKQDLAEKVLGKSKIEDYFPEFARYTTPEDATPEPGEPRVT  RAKYFIRDEFRLISTASGDGRHYCPHFTCAVDTENARRIFNDCRDIQRMHLRQYELL</p> <p>*</p>
<p><b>Insect cells</b></p>	<p><b>Adgrd1 – mini-Gs</b></p>	<p><b>D1- mGs.05</b></p>	<p>MKTIIALSIFCLVFA<sup>Y</sup>DKDDDDADYKDDDDKDYKDDDDADYKDDDDK<sup>LE</sup>VLFQGPA  YHPIITNLTEERKTFQSPGVLSYLQNVLSLPSKSLSEQTALNLTKTFLKAVGEILLPGW  IALSEDSAVVLSLIDTIDTVMGHVSSNLHGSTPQVTVEGSSAMAEFSVAKILPKTVNSS  HYRFPAHGQSFQIPHEAFHRHAWSTVVGLLYHSMHYLNNIIPAHTKIAEAMHH  QDCLLFATSHLISLEVSPPTLSQNLGSPPLITVHLKHRLTRKQCEATNSSNRVVFYCA  FLDFSSGEGVWSNHGCALTRGNLTYSVCRCTHLTNFAILMQVVPLELARGHQVALSS  ISYVGCSLSVLCLVATLVTFVAVLSSVSTIRNQRYHIHANLSFAVLVAQVLLISFRLEPGTT  PCQVMVALLHYFFLSAFWMLVEGLHLYSMVIKVFGESESKHRYYYGMGWGFPLLI  CIISLAFAMDSYGTSNCWLSLASGAIWAFVAPALFVIVVIGILIAVTRVISQISADNY  KIHGDPSAFKLTAKAVAVLLPILGTSWVFGVLAVNGCAVVFQYMFATLNSLQGLFIFL  FHCLLNSEVRAAFKHKTWVSLTSSARTSNAKPFHSDLMNGTRPGMASTKLSPWD  KSSSAHRVDLSAVGGGGGGGGGGGGGGSEENLYFQGLGNSKTEDQRNEEKAQRE  ANKKIEKQLQKDKQVYRATHRLLLLGADNSGKSTIVKQMRILHGGSGGGSGTSGIFET  KFQVDKVNFMFVGGQRDERRKWIQCFNDVTAIIFVVDSSDYNRLQEALNDFKSI  WNNRWLRTISVILFNKQDLAEKVLGKSKIEDYFPEFARYTTPEDATPEPGEPRVT  RAKYFIRDEFRLISTASGDGRHYCPHFTCAVDTENARRIFNDCRDIQRMHLRQYELL</p> <p>*</p>
<p><b>Insect cells</b></p>	<p><b>Adgrd1 – mini-Gs</b></p>	<p><b>D1- mGs.06</b></p>	<p>MKTIIALSIFCLVFA<sup>Y</sup>DKDDDDADYKDDDDKDYKDDDDADYKDDDDK<sup>LE</sup>VLFQGPA  YHPIITNLTEERKTFQSPGVLSYLQNVLSLPSKSLSEQTALNLTKTFLKAVGEILLPGW  IALSEDSAVVLSLIDTIDTVMGHVSSNLHGSTPQVTVEGSSAMAEFSVAKILPKTVNSS  HYRFPAHGQSFQIPHEAFHRHAWSTVVGLLYHSMHYLNNIIPAHTKIAEAMHH  QDCLLFATSHLISLEVSPPTLSQNLGSPPLITVHLKHRLTRKQHSEACNSSNRVVFYCA  FLDFSSGEGVWSNHGCALTRGNLTYSVCRCTHLTNFAILMQVVPLELARGHQVALSS  ISYVGCSLSVLCLVATLVTFVAVLSSVSTIRNQRYHIHANLSFAVLVAQVLLISFRLEPGTT  PCQVMVALLHYFFLSAFWMLVEGLHLYSMVIKVFGESESKHRYYYGMGWGFPLLI  CIISLAFAMDSYGTSNNCCLSLASGAIWAFVAPALFVIVVIGILIAVTRVISQISADNYK  IHGDPSAFKLTAKAVAVLLPILGTSWVFGVLAVNGCAVVFQYMFATLNSLQGLFIFL  HCLLNSEVRAAFKHKTWVSLTSSARTSNAKPFHSDLMNGTRPGMASTKLSPWDK  SSSAHRVDLSAVGGGGGGGGGGGGGGSEENLYFQGLGNSKTEDQRNEEKAQREA  NKKIEKQLQKDKQVYRATHRLLLLGADNSGKSTIVKQMRILHGGSGGGSGTSGIFETK  FQVDKVNFMFVGGQRDERRKWIQCFNDVTAIIFVVDSSDYNRLQEALNDFKSIW</p>

			<p>NNRWLRTISVILFNKQDLLAEKVLGKSKIEDYFPEFARYTTPEDATPEPGEDPRVTR  AKYFIRDEFRLISTASGDGRHYCYPHFTCAVDTENARRIFNDCRDIIRMHLRQYELL*</p>
Insect cells	Adgrd1 – mini-Gs	D1- mGs.07	<p>MKTIIALSIFCLVFA<sup>DYK</sup>DDDDADYKDDDDKDYKDDDDADYKDDDDK<sup>LEVL</sup>FQGPAYHPIITNLTEERKTFQSPGVLSYLQNVLSLPSKSLSEQTALNLTKTKFLKAVGEILLPGW  IALSEDSAVVLSLIDTIDTVMGHVSSNLHGSTPQVTVEGSSAMAEFVAKILPKTVNSS  HYRFPAHGQSFQIPHEAFHRHAWSTVVGLLYHSMHYLNNIWPAAHTKIAEAMHH  QDCLLFATSHLISLEVSPPTLSQNLGSPPLITVHLKHLRTRKQHSEATN<sup>C</sup>SNRVFVYCA  FLDFSSGEGVWSNHGCALTRGNLTYSVCRCTHLTNFAILMQVVPLELARGHQVALSS  ISYVGCSLSVLCLVATLVTFVAVLSSVSTIRNQRYHIHANLSFAVLVAQVLLISFRLEPGTT  PCQVMAVLLHYFFLSAFWMLVEGLHLYSMVIKVFGESESKHRYYYGMGWGFPLLI  CIISLFSAMDYSGTNN<sup>C</sup>SLASGAIWAFVAPALFVIVVNIIGILIAVTRVISQISADNYK  IHGDPSAFKLTAKAVAVLLPILGTSWVFGVLAVNGCAVVFQYMFATLNSLQGLFIFL  HCLLNSEVRAAFKHKT<sup>K</sup>VWVSLTSSSARTSNAKPFHSDLMNGTRPGMASTKLSPWDK  SSHS<sup>A</sup>HRVDLSAVGGGGGGGGGGGG<sup>SE</sup>ENLYFQGLGNSKTEDQRNEEKAQREA  NKKIEKQLQKDKQVYRATHRLLLLGADNSGKSTIVKQMRILHGSGSGGGTSGIFETK  FQVDKVNFMFDVGGQRDERRKWIQCFNDVTAIIFVVDSSDYNRLQEALNDFKSIW  NNRWLRTISVILFNKQDLLAEKVLGKSKIEDYFPEFARYTTPEDATPEPGEDPRVTR  AKYFIRDEFRLISTASGDGRHYCYPHFTCAVDTENARRIFNDCRDIIRMHLRQYELL*</p>
Insect cells	Adgrd1 – mini-Gs	D1- mGs.08	<p>MKTIIALSIFCLVFA<sup>DYK</sup>DDDDADYKDDDDKDYKDDDDADYKDDDDK<sup>LEVL</sup>FQGPAYHPIITNLTEERKTFQSPGVLSYLQNVLSLPSKSLSEQTALNLTKTKFLKAVGEILLPGW  IALSEDSAVVLSLIDTIDTVMGHVSSNLHGSTPQVTVEGSSAMAEFVAKILPKTVNSS  HYRFPAHGQSFQIPHEAFHRHAWSTVVGLLYHSMHYLNNIWPAAHTKIAEAMHH  QDCLLFATSHLISLEVSPPTLSQNLGSPPLITVHLKHLRTRKQHSEATN<sup>S</sup>SNRVFVYCA  FLDFSSGEGVWSNHGCALTRGNLTYSVCRCTHLTNFAILMQVVPLEL<sup>C</sup>GHQVALSS  ISYVGCSLSVLCLVATLVTFVAVLSSVSTIRNQRYHIHANLSFAVLVAQVLLISFRLEPGTT  PCQVMAVLLHYFFLSAFWMLVEGLHLYSMVIKVFGESESKHRYYYGMGWGFPLLI  CIISLFSAMDYSGTNN<sup>C</sup>WLSLASGAIWAFVAPALFVIVVNIIGILIAVTRVISQISADNY  KIHGDPSAFKLTAKAVAVLLPILGTSWVFGVLAVNGCAVVFQYMFATLNSLQGLFIFL  FHCLLNSEVRAAFKHKT<sup>K</sup>VWVSLTSSSARTSNAKPFHSDLMNGTRPGMASTKLSPWD  KSSH<sup>A</sup>HRVDLSAVGGGGGGGGGGGG<sup>SE</sup>ENLYFQGLGNSKTEDQRNEEKAQRE  ANKKIEKQLQKDKQVYRATHRLLLLGADNSGKSTIVKQMRILHGSGSGGGTSGIFET  KFQVDKVNFMFDVGGQRDERRKWIQCFNDVTAIIFVVDSSDYNRLQEALNDFKSI  WNNRWLRTISVILFNKQDLLAEKVLGKSKIEDYFPEFARYTTPEDATPEPGEDPRVT  RAKYFIRDEFRLISTASGDGRHYCYPHFTCAVDTENARRIFNDCRDIIRMHLRQYELL  *</p>
Insect cells	Adgrd1 – mini-Gs	D1- mGs.09	<p>MKTIIALSIFCLVFA<sup>DYK</sup>DDDDADYKDDDDKDYKDDDDADYKDDDDK<sup>LEVL</sup>FQGPAYHPIITNLTEERKTFQSPGVLSYLQNVLSLPSKSLSEQTALNLTKTKFLKAVGEILLPGW  IALSEDSAVVLSLIDTIDTVMGHVSSNLHGSTPQVTVEGSSAMAEFVAKILPKTVNSS  HYRFPAHGQSFQIPHEAFHRHAWSTVVGLLYHSMHYLNNIWPAAHTKIAEAMHH  QDCLLFATSHLISLEVSPPTLSQNLGSPPLITVHLKHLRTRKQHSEATN<sup>S</sup>SNRVFVYCA  FLDFSSGEGVWSNHGCALTRGN<sup>C</sup>TYSVCRCTHLTNFAILMQVVPLELARGHQVALSS  ISYVGCSLSVLCLVATLVTFVAVLSSVSTIRNQRYHIHANLSFAVLVAQVLLISFRLEPGTT  PCQVMAVLLHYFFLSAFWMLVEGLHLYSMVIKVFGESESKHRYYYGMGWGFPLLI  CIISLFSAMDYSGTNN<sup>C</sup>WLSGAIWAFVAPALFVIVVNIIGILIAVTRVISQISADNY</p>

			<p>KIHGDPSAFKLTAKAVAVLLPILGTSWVFGVLAVNGCAVVFQYMFATLNSLQGLFIFL  FHCLLNSEVRAAFKHKTQVWSLTSSSARTSNAKPFHSDLMNGTRPGMASTKLSPWD  KSSSAHRVDLSAVGGGGSGGGGGGGGGGSE<sup>N</sup>LYFQ<sup>G</sup>CLGNSKTEDQRNEEKAQRE  ANKKIEKQLQKDKQVYRATHRLLLLGADNSGKSTIVKQMRILHGGSGGGSGGTS<sup>G</sup>IFET  KFQVDKVNFMFDVGGQRDERRKWIQCFNDVTAIIFVVDSSDYNRLQEALNDFKSI  WNNRWLRTISVILFLNKQDLLAEKVLGKSKIEDYFPEFARYTTPEDATPEPGEDPRVT  RAKYFIRDEFRLISTASGDGRHYCYPHFTCAVDTENARRIFNDCRDIIQRMHLRQYELL</p> <p>*</p>
<p><b>Insect cells</b></p>	<p><b>Adgrd1 – mini-Gs</b></p>	<p><b>D1- mGs.10</b></p>	<p>MKTIIALS<sup>Y</sup>IFCLVFA<sup>D</sup>YK<sup>D</sup>DDDD<sup>A</sup>DYK<sup>D</sup>DDDD<sup>K</sup>DYK<sup>D</sup>DDDD<sup>A</sup>DYK<sup>D</sup>DDDD<sup>K</sup>LE<sup>V</sup>LFQ<sup>G</sup>PA  YHPIITNLTEERKTFQSPGVLSYLQNVLSLPSKSLSEQTALNLTKTKFLKAVGEILLPGW  IALSEDSAVVLSLIDTIDTVMGHVSSNLHGSTPQVTEGSSAMA<sup>E</sup>FSVAKILPKTVN<sup>S</sup>  HYRFPAHGQSFIQIPHEAFHRHAWSTVVGLLYHSMHYLNNI<sup>W</sup>PAHTKIAEAMHH  QDCLLFATSHLISLEVSPPTLSQNL<sup>S</sup>GSPLITVHLKHRLTRKQ<sup>H</sup>SEATN<sup>S</sup>SNRVFVYCA  FLDFSSGEGVWSNHGCALTRGN<sup>C</sup>TYSVCRCTHLTN<sup>F</sup>FAILMQV<sup>V</sup>PLELAC<sup>G</sup>HQVALSS  ISYVGC<sup>S</sup>LSVLCLVATLVTF<sup>A</sup>VLSSVSTIRNQR<sup>Y</sup>HIHANLSFAVLVAQVLLIS<sup>F</sup>RLEPGTT  PCQVMAVLLHYFFLSAFAWMLVEGLHLYSMVIKVF<sup>G</sup>SEDSKHRY<sup>Y</sup>YGMGWGFPLLI  CIISL<sup>S</sup>FAMDSYGT<sup>S</sup>NNCWL<sup>S</sup>LSL<sup>C</sup>SGAIWAFVAPALFVIV<sup>N</sup>IGILIAVTRVISQISADNY  KIHGDPSAFKLTAKAVAVLLPILGTSWVFGVLAVNGCAVVFQYMFATLNSLQGLFIFL  FHCLLNSEVRAAFKHKTQVWSLTSSSARTSNAKPFHSDLMNGTRPGMASTKLSPWD  KSSSAHRVDLSAVGGGGSGGGGGGGGGGSE<sup>N</sup>LYFQ<sup>G</sup>CLGNSKTEDQRNEEKAQRE  ANKKIEKQLQKDKQVYRATHRLLLLGADNSGKSTIVKQMRILHGGSGGGSGGTS<sup>G</sup>IFET  KFQVDKVNFMFDVGGQRDERRKWIQCFNDVTAIIFVVDSSDYNRLQEALNDFKSI  WNNRWLRTISVILFLNKQDLLAEKVLGKSKIEDYFPEFARYTTPEDATPEPGEDPRVT  RAKYFIRDEFRLISTASGDGRHYCYPHFTCAVDTENARRIFNDCRDIIQRMHLRQYELL</p> <p>*</p>

**Supp. Table 6: Cryo-EM data collection and processing details of Adgrd1-Gs complex**

	<b>Adgrd1-Gs complex</b>	
	<b>Sample A</b>	<b>Sample B</b>
<b>Direct Electron Detector</b>	<b>K3 + GIF</b>	<b>K3 + GIF</b>
<b>Number of movies</b>	<b>22'504</b>	<b>31'141</b>
<b>Acceleration Voltage [kV]</b>	<b>300</b>	<b>300</b>
<b>Total Exposure [e<sup>-</sup>/Å<sup>2</sup>]</b>	<b>55</b>	<b>80</b>
<b>Number of Frames per Movie</b>	<b>40</b>	<b>40</b>
<b>Defocus Range [μm]</b>	<b>-0.8 to -2.0</b>	<b>-1.2 to -3.0</b>
<b>Automated EM Data Acquisition Software</b>	<b>EPU</b>	<b>EPU</b>
<b>Pixel Size [Å/px]</b>	<b>0.51</b>	<b>0.51</b>
<b>Symmetry</b>	<b>C1</b>	<b>C1</b>
<b>Total Number of picked Particles</b>	<b>1'055'084</b>	<b>1'941'387</b>
<b>Number of Particles after 2D classification</b>	<b>95'413</b>	<b>33'278</b>
<b>Number of Particles in final Map</b>	<b>-</b>	<b>-</b>
<b>Final Map Resolution [Å]</b>	<b>-</b>	<b>-</b>
<b>FSC threshold</b>	<b>-</b>	<b>-</b>
<b>Map Sharpening B-factor [Å<sup>2</sup>]</b>	<b>-</b>	<b>-</b>



**Supp. Table 7: Cryo-EM data collection and processing details of the JSR1-Gt complex**

	<b>JSR1-Gt</b>
<b>Direct Electron Detector</b>	<b>K3 + GIF</b>
<b>Number of movies</b>	<b>11'408</b>
<b>Acceleration Voltage [kV]</b>	<b>300</b>
<b>Total Exposure [<math>e^-/\text{\AA}^2</math>]</b>	<b>50</b>
<b>Number of Frames per Movie</b>	<b>40</b>
<b>Defocus Range [<math>\mu\text{m}</math>]</b>	<b>-1.0 to -2.2</b>
<b>Automated EM Data Acquisition Software</b>	<b>EPU</b>
<b>Pixel Size [<math>\text{\AA}/\text{px}</math>]</b>	<b>0.51</b>
<b>Symmetry</b>	<b>C1</b>
<b>Total Number of picked Particles</b>	<b>2'387'106</b>
<b>Number of Particles after 2D classification</b>	<b>No particles for the full complex</b>
<b>Number of Particles in final Map</b>	<b>-</b>
<b>Final Map Resolution [<math>\text{\AA}</math>]</b>	<b>-</b>
<b>FSC threshold</b>	<b>-</b>
<b>Map Sharpening B-factor [<math>\text{\AA}^2</math>]</b>	<b>-</b>

**Supp. Table 8: Cryo-EM data collection and processing details of JSR1-hGi complexes**

	<b>JSR1-hGi</b>	<b>JSR1-hGi-scFv16</b>
<b>Direct Electron Detector</b>	<b>K3 + GIF</b>	<b>K3 + GIF</b>
<b>Number of movies</b>	<b>11'801</b>	<b>28'128</b>
<b>Acceleration Voltage [kV]</b>	<b>300</b>	<b>300</b>
<b>Total Exposure [<math>e^-/\text{\AA}^2</math>]</b>	<b>70</b>	<b>54</b>
<b>Number of Frames per Movie</b>	<b>40</b>	<b>40</b>
<b>Defocus Range [<math>\mu\text{m}</math>]</b>	<b>-1.0 to -2.4</b>	<b>-1.0 to -2.4</b>
<b>Automated EM Data Acquisition Software</b>	<b>EPU</b>	<b>EPU</b>
<b>Pixel Size [<math>\text{\AA}/\text{px}</math>]</b>	<b>0.51</b>	<b>0.51</b>
<b>Symmetry</b>	<b>C1</b>	<b>C1</b>
<b>Total Number of picked Particles</b>	<b>2'023'050</b>	<b>681'002</b>
<b>Number of Particles after 2D classification</b>	<b>62'660</b>	<b>127'009</b>
<b>Number of Particles in final Map</b>	<b>-</b>	<b>39'274</b>
<b>Final Map Resolution [<math>\text{\AA}</math>]</b>	<b>-</b>	<b>5.62</b>
<b>FSC threshold</b>	<b>-</b>	<b>0.143</b>
<b>Map Sharpening B-factor [<math>\text{\AA}^2</math>]</b>	<b>-</b>	<b>-420</b>

**Supp. Table 9: Cryo-EM data collection and processing details of JSR1-JSGiq**

	JSR1-JSGiq_1	JSR1-JSGiq_2
Direct Electron Detector	K3 + GIF	K3 + GIF
Number of movies	74'605	74'605
Acceleration Voltage [kV]	300	300
Total Exposure [ $e^-/\text{\AA}^2$ ]	70	70
Number of Frames per Movie	40	40
Defocus Range [ $\mu\text{m}$ ]	-1.0 to -2.4	-1.0 to -2.4
Automated EM Data Acquisition Software	EPU	EPU
Pixel Size [ $\text{\AA}/\text{px}$ ]	0.51	0.51
Symmetry	C1	C1
Total Number of picked Particles	2'964'739	2'964'739
Number of Particles after 2D classification	768'100	768'100
Number of Particles in final Map	159'665	134'167
Final Map Resolution [ $\text{\AA}$ ]	4.06	4.17
FSC threshold	0.143	0.143
Map Sharpening B-factor [ $\text{\AA}^2$ ]	-162	-206

Supp. Table 10: Refinement and validation statistics of the JSR1-JSGiq complex models.

	JSR1-JSGiq_1	JSR1-JSGiq_2
Initial models used	6I9K (JSR1), 6CRK (Gi)	6I9K (JSR1), 6CRK (Gi)
Model resolution [Å]	4.0	4.1
FSC threshold	0.143	0.143
Non-hydrogen atoms	7482	7489
Protein residues	946	948
Ligand	ATR611	ATR611
R.m.s.d. Bond length [Å]	0.004	0.004
R.m.s.d. Bond angles [°]	0.718	0.756
MolProbity score	2.10	2.27
Clashscore	14.16	19.96
Rotamer outliers [%]	1.47	1.72
<i>Ramachandran plot</i>		
Favored [%]	95.49	95.73
Allowed [%]	4.40	4.17
Outliers [%]	0.11	0.11

**Supp. Table 11: List of software packages used**

<b>Software</b>	<b>Function</b>	<b>Provider</b>
<b>PyMOL[134]</b>	<b>Three-dimensional molecular viewer</b>	<b>Schrödinger</b>
<b>Coot: CCP4 software suite [133]</b>	<b>Display and modification of three-dimensional models</b>	<b>CCP4 software suite</b>
<b>PHENIX [132]</b>	<b>Phasing, map modifications and structure refinement</b>	<b>PHENIX</b>
<b>Biorender.com</b>	<b>Graphics program</b>	<b>BioRender</b>
<b>Adobe Illustrator</b>	<b>Graphics program</b>	<b>Adobe</b>
<b>CryoSPARC [125]</b>	<b>Cryo-EM data processing</b>	<b>Structura Biotechnology</b>
<b>Relion [128, 171]</b>	<b>Cryo-EM data processing</b>	<b>MRC LMB / Scheres lab</b>
<b>CisTEM [172]</b>	<b>Cryo-EM data processing</b>	<b>UMASS/Grigorieff lab</b>
<b>GraphPad PRISM</b>	<b>Statistics program</b>	<b>GraphPad Software</b>
<b>Chimera/ChimeraX [131, 173]</b>	<b>Three-dimensional molecular viewer and editor</b>	<b>RBVI</b>
<b>ChemDraw</b>	<b>Graphics program for chemical structures</b>	<b>Revvity Signals</b>

## 8.2 Presentations at scientific events

### 8.2.1 Poster presentations

- Synapsis Forum 2020 and 2022
- 4<sup>th</sup> ERNEST conference 2020
- 4<sup>th</sup> GPCRnet International Conference 2022
- International Retinal Protein Conference 2022
- 13<sup>th</sup> D-BIOL Symposium ETH Zurich

### 8.2.2 Scientific talk

- Synapsis Forum 2021

## 8.3 List of publications

- Thomas Gruhl, Tobias Weinert, Matthew J. Rodrigues, Christopher J. Milne, Giorgia Ortolani, Karol Nass, Eriko Nango, Saumik Sen, Philip J. M. Johnson, Claudio Cirelli, Antonia Furrer, Sandra Mous, Petr Skopintsev, Daniel James, Florian Dworkowski, Petra Båth, Demet Kekilli, Dmitry Ozerov, Rie Tanaka, Hannah Glover, Camila Bacellar, Steffen Brünle, Cecilia M. Casadei, Azeglio D. Diethelm, Dardan Gashi, Guillaume Gotthard, Ramon Guixà-González, Yasumasa Joti, Victoria Kabanova, Gregor Knopp, Elena Lesca, Piyee Ma, Isabelle Martiel, Jonas Mühle, Shigeki Owada, Filip Pamula, Daniel Sarabi, **Oliver Tejero**, Ching-Ju Tsai, Niranjan Varma, Anna Wach, Sébastien Boutet, Kensuke Tono, Przemyslaw Nogly, Xavier Deupi, So Iwata, Richard Neutze, Jörg Standfuss, Gebhard Schertler & Valerie Panneels. Ultrafast structural changes direct the first molecular events of vision. *Nature* **615**, 939–944 (2023). <https://doi.org/10.1038/s41586-023-05863-6>
- **Oliver Tejero**, Filip Pamula, Mitsumasa Koyanagi, Takashi Nagata, Pavel Afanasyev, Ishita Das, Xavier Deupi, Mordechai Sheves, Akihisa Terakita, Gebhard F.X. Schertler, Matthew J. Rodrigues, Ching-Ju Tsai. Active state structures of a bistable visual opsin bound to G proteins. *bioRxiv* 2024.04.09.588704; <https://doi.org/10.1101/2024.04.09.588704>
- Matthew J. Rodrigues, **Oliver Tejero**, Jonas Mühle, Filip K. Pamula, Ishita Das, Ching-Ju Tsai, Akihisa Terakita, Mordechai Sheves, Gebhard F.X. Schertler. Activating an invertebrate bistable opsin with the all-trans 6.11 retinal analogue. *bioRxiv* 2024.04.08.588240; doi: <https://doi.org/10.1101/2024.04.08.588240>

## 9 Bibliography

1. Schonenbach, N.S., S. Hussain, and M.A. O'Malley, *Structure and function of G protein-coupled receptor oligomers: implications for drug discovery*. Wiley Interdiscip Rev Nanomed Nanobiotechnol, 2015. **7**(3): p. 408-27.
2. Latek, D., et al., *G protein-coupled receptors--recent advances*. Acta Biochim Pol, 2012. **59**(4): p. 515-29.
3. Lefkowitz, R.J., *The superfamily of heptahelical receptors*. Nat Cell Biol, 2000. **2**(7): p. E133-6.
4. Sriram, K. and P.A. Insel, *G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs?* Mol Pharmacol, 2018. **93**(4): p. 251-258.
5. Lappano, R. and M. Maggiolini, *G protein-coupled receptors: novel targets for drug discovery in cancer*. Nat Rev Drug Discov, 2011. **10**(1): p. 47-60.
6. Fang, Y., T. Kenakin, and C. Liu, *Editorial: Orphan GPCRs As Emerging Drug Targets*. Front Pharmacol, 2015. **6**: p. 295.
7. Lander, E.S., et al., *Initial sequencing and analysis of the human genome*. Nature, 2001. **409**(6822): p. 860-921.
8. Baldwin, J.M., *Structure and function of receptors coupled to G proteins*. Curr Opin Cell Biol, 1994. **6**(2): p. 180-90.
9. Attwood, T.K. and J.B. Findlay, *Fingerprinting G-protein-coupled receptors*. Protein Eng, 1994. **7**(2): p. 195-203.
10. Kolakowski, L.F., Jr., *GCRDb: a G-protein-coupled receptor database*. Recept Channels, 1994. **2**(1): p. 1-7.
11. Hu, G.-M., T.-L. Mai, and C.-M. Chen, *Visualizing the GPCR Network: Classification and Evolution*. Scientific Reports, 2017. **7**(1): p. 15495.
12. Civelli, O., et al., *G protein-coupled receptor deorphanizations*. Annu Rev Pharmacol Toxicol, 2013. **53**: p. 127-46.
13. Harmar, A.J., *Family-B G-protein-coupled receptors*. Genome Biol, 2001. **2**(12): p. Reviews3013.
14. Fredriksson, R., et al., *The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints*. Mol Pharmacol, 2003. **63**(6): p. 1256-72.
15. Schiöth, H.B. and R. Fredriksson, *The GRAFS classification system of G-protein coupled receptors in comparative perspective*. Gen Comp Endocrinol, 2005. **142**(1-2): p. 94-101.
16. Dorsam, R.T. and J.S. Gutkind, *G-protein-coupled receptors and cancer*. Nat Rev Cancer, 2007. **7**(2): p. 79-94.
17. Zheng, C., J. Tholen, and V.V. Gurevich, *Critical role of the finger loop in arrestin binding to the receptors*. PLoS One, 2019. **14**(3): p. e0213792.
18. Venkatakrisnan, A.J., et al., *Molecular signatures of G-protein-coupled receptors*. Nature, 2013. **494**(7436): p. 185-194.
19. Rosenbaum, D.M., S.G.F. Rasmussen, and B.K. Kobilka, *The structure and function of G-protein-coupled receptors*. Nature, 2009. **459**(7245): p. 356-363.
20. Barak, L.S., et al., *The conserved seven-transmembrane sequence NP(X)2,3Y of the G-protein-coupled receptor superfamily regulates multiple properties of the beta 2-adrenergic receptor*. Biochemistry, 1995. **34**(47): p. 15407-14.
21. Olivella, M., G. Caltabiano, and A. Cordero, *The role of Cysteine 6.47 in class A GPCRs*. BMC Struct Biol, 2013. **13**: p. 3.
22. Vogel, R., et al., *Functional role of the "ionic lock"--an interhelical hydrogen-bond network in family A heptahelical receptors*. J Mol Biol, 2008. **380**(4): p. 648-55.
23. Trzaskowski, B., et al., *Action of molecular switches in GPCRs--theoretical and experimental studies*. Curr Med Chem, 2012. **19**(8): p. 1090-109.



24. Rasmussen, S.G.F., et al., *Crystal structure of the  $\beta_2$  adrenergic receptor–Gs protein complex*. *Nature*, 2011. **477**(7366): p. 549-555.
25. Ballesteros, J.A. and H. Weinstein, [19] *Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors*, in *Methods in Neurosciences*, S.C. Sealfon, Editor. 1995, Academic Press. p. 366-428.
26. Weis, W.I. and B.K. Kobilka, *The Molecular Basis of G Protein-Coupled Receptor Activation*. *Annu Rev Biochem*, 2018. **87**: p. 897-919.
27. Zhou, Q., et al., *Common activation mechanism of class A GPCRs*. *eLife*, 2019. **8**: p. e50279.
28. Goncalves, J.A., et al., *Highly conserved tyrosine stabilizes the active state of rhodopsin*. *Proceedings of the National Academy of Sciences*, 2010. **107**(46): p. 19861-19866.
29. Valentin-Hansen, L., et al., *The Arginine of the DRY Motif in Transmembrane Segment III Functions as a Balancing Micro-switch in the Activation of the  $\beta_2$ -Adrenergic Receptor*. *Journal of Biological Chemistry*, 2012. **287**(38): p. 31973-31982.
30. Varma, N., et al., *Crystal structure of jumping spider rhodopsin-1 as a light sensitive GPCR*. *Proceedings of the National Academy of Sciences*, 2019. **116**(29): p. 14547-14556.
31. Murakami, M. and T. Kouyama, *Crystal structure of squid rhodopsin*. *Nature*, 2008. **453**(7193): p. 363-367.
32. Bjarnadóttir, T.K., et al., *The human and mouse repertoire of the adhesion family of G-protein-coupled receptors*. *Genomics*, 2004. **84**(1): p. 23-33.
33. Baud, V., et al., *EMR1, an unusual member in the family of hormone receptors with seven transmembrane segments*. *Genomics*, 1995. **26**(2): p. 334-344.
34. Gray, J.X., et al., *CD97 is a processed, seven-transmembrane, heterodimeric receptor associated with inflammation*. *J Immunol*, 1996. **157**(12): p. 5438-47.
35. Kwakkenbos, M.J., et al., *The EGF-TM7 family: a postgenomic view*. *Immunogenetics*, 2004. **55**(10): p. 655-666.
36. Krishnan, A., et al., *The origin of GPCRs: identification of mammalian like Rhodopsin, Adhesion, Glutamate and Frizzled GPCRs in fungi*. *PLoS One*, 2012. **7**(1): p. e29817.
37. Srivastava, M., et al., *The Amphimedon queenslandica genome and the evolution of animal complexity*. *Nature*, 2010. **466**(7307): p. 720-6.
38. Putnam, N.H., et al., *Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization*. *Science*, 2007. **317**(5834): p. 86-94.
39. Nordström, K.J., et al., *Independent HHsearch, Needleman--Wunsch-based, and motif analyses reveal the overall hierarchy for most of the G protein-coupled receptor families*. *Mol Biol Evol*, 2011. **28**(9): p. 2471-80.
40. Nordström, K.J., R. Fredriksson, and H.B. Schiöth, *The amphioxus (Branchiostoma floridae) genome contains a highly diversified set of G protein-coupled receptors*. *BMC Evol Biol*, 2008. **8**: p. 9.
41. Kamesh, N., G.K. Aradhyam, and N. Manoj, *The repertoire of G protein-coupled receptors in the sea squirt Ciona intestinalis*. *BMC Evol Biol*, 2008. **8**: p. 129.
42. Fredriksson, R. and H.B. Schiöth, *The repertoire of G-protein-coupled receptors in fully sequenced genomes*. *Mol Pharmacol*, 2005. **67**(5): p. 1414-25.
43. Nordström, K.J., et al., *The Secretin GPCRs descended from the family of Adhesion GPCRs*. *Mol Biol Evol*, 2009. **26**(1): p. 71-84.
44. Lin, H.H., et al., *Autocatalytic cleavage of the EMR2 receptor occurs at a conserved G protein-coupled receptor proteolytic site motif*. *J Biol Chem*, 2004. **279**(30): p. 31823-32.
45. Araç, D., et al., *A novel evolutionarily conserved domain of cell-adhesion GPCRs mediates autoproteolysis*. *Embo j*, 2012. **31**(6): p. 1364-78.
46. Krasnoperov, V., et al., *Post-translational proteolytic processing of the calcium-independent receptor of alpha-latrotoxin (CIRL), a natural chimera of the cell adhesion protein and the G*

- protein-coupled receptor. Role of the G protein-coupled receptor proteolysis site (GPS) motif.* J Biol Chem, 2002. **277**(48): p. 46518-26.
47. Liebscher, I., et al., *A tethered agonist within the ectodomain activates the adhesion G protein-coupled receptors GPR126 and GPR133.* Cell Rep, 2014. **9**(6): p. 2018-26.
  48. Krishnan, A., et al., *Classification, Nomenclature, and Structural Aspects of Adhesion GPCRs.* Handb Exp Pharmacol, 2016. **234**: p. 15-41.
  49. Prömel, S., et al., *Characterization and functional study of a cluster of four highly conserved orphan adhesion-GPCR in mouse.* Developmental Dynamics, 2012. **241**(10): p. 1591-1602.
  50. Formstone, C.J., et al., *Basal enrichment within neuroepithelia suggests novel function(s) for Celsr1 protein.* Mol Cell Neurosci, 2010. **44**(3): p. 210-22.
  51. Langenhan, T., G. Aust, and J. Hamann, *Sticky signaling--adhesion class G protein-coupled receptors take the stage.* Sci Signal, 2013. **6**(276): p. re3.
  52. Gupte, J., et al., *Signaling property study of adhesion G-protein-coupled receptors.* FEBS Lett, 2012. **586**(8): p. 1214-9.
  53. Iguchi, T., et al., *Orphan G protein-coupled receptor GPR56 regulates neural progenitor cell migration via a G alpha 12/13 and Rho pathway.* J Biol Chem, 2008. **283**(21): p. 14469-78.
  54. Paavola, K.J., et al., *The N terminus of the adhesion G protein-coupled receptor GPR56 controls receptor signaling activity.* J Biol Chem, 2011. **286**(33): p. 28914-28921.
  55. Mogha, A., et al., *Gpr126 functions in Schwann cells to control differentiation and myelination via G-protein activation.* J Neurosci, 2013. **33**(46): p. 17976-85.
  56. Okajima, D., G. Kudo, and H. Yokota, *Brain-specific angiogenesis inhibitor 2 (BAI2) may be activated by proteolytic processing.* J Recept Signal Transduct Res, 2010. **30**(3): p. 143-53.
  57. Yang, L., et al., *GPR56 Regulates VEGF production and angiogenesis during melanoma progression.* Cancer Res, 2011. **71**(16): p. 5558-68.
  58. Ward, Y., et al., *LPA receptor heterodimerizes with CD97 to amplify LPA-initiated RHO-dependent signaling and invasion in prostate cancer cells.* Cancer Res, 2011. **71**(23): p. 7301-11.
  59. Stoveken, H.M., et al., *Adhesion G protein-coupled receptors are activated by exposure of a cryptic tethered agonist.* Proc Natl Acad Sci U S A, 2015. **112**(19): p. 6194-9.
  60. Demberg, L.M., et al., *Identification of the tethered peptide agonist of the adhesion G protein-coupled receptor GPR64/ADGRG2.* Biochem Biophys Res Commun, 2015. **464**(3): p. 743-7.
  61. Wilde, C., et al., *The constitutive activity of the adhesion GPCR GPR114/ADGRG5 is mediated by its tethered agonist.* Faseb j, 2016. **30**(2): p. 666-73.
  62. Müller, A., et al., *Oriented Cell Division in the C. elegans Embryo Is Coordinated by G-Protein Signaling Dependent on the Adhesion GPCR LAT-1.* PLoS Genet, 2015. **11**(10): p. e1005624.
  63. Langenhan, T., *Adhesion G protein-coupled receptors-Candidate metabotropic mechanosensors and novel drug targets.* Basic Clin Pharmacol Toxicol, 2020. **126 Suppl 6**: p. 5-16.
  64. Kishore, A. and R.A. Hall, *Versatile Signaling Activity of Adhesion GPCRs.* Handb Exp Pharmacol, 2016. **234**: p. 127-146.
  65. Vizurraga, A., et al., *Mechanisms of adhesion G protein-coupled receptor activation.* J Biol Chem, 2020. **295**(41): p. 14065-14083.
  66. Liebscher, I. and T. Schöneberg, *Tethered Agonism: A Common Activation Mechanism of Adhesion GPCRs.* Handb Exp Pharmacol, 2016. **234**: p. 111-125.
  67. Barros-Álvarez, X., et al., *The tethered peptide activation mechanism of adhesion GPCRs.* Nature, 2022. **604**(7907): p. 757-762.
  68. Ping, Y.-Q., et al., *Structural basis for the tethered peptide activation of adhesion GPCRs.* Nature, 2022. **604**(7907): p. 763-770.
  69. Xiao, P., et al., *Tethered peptide activation mechanism of the adhesion GPCRs ADGRG2 and ADGRG4.* Nature, 2022. **604**(7907): p. 771-778.

70. Qu, X., et al., *Structural basis of tethered agonism of the adhesion GPCRs ADGRD1 and ADGRF1*. *Nature*, 2022. **604**(7907): p. 779-785.
71. Monk, K.R., et al., *A G protein-coupled receptor is essential for Schwann cells to initiate myelination*. *Science*, 2009. **325**(5946): p. 1402-5.
72. Monk, K.R., et al., *Gpr126 is essential for peripheral nerve development and myelination in mammals*. *Development*, 2011. **138**(13): p. 2673-80.
73. Scherer, S.S., et al., *Axons regulate Schwann cell expression of the POU transcription factor SCIP*. *J Neurosci*, 1994. **14**(4): p. 1930-42.
74. Petersen, S.C., et al., *The adhesion GPCR GPR126 has distinct, domain-dependent functions in Schwann cell development mediated by interaction with laminin-211*. *Neuron*, 2015. **85**(4): p. 755-69.
75. Paavola, K.J., et al., *Type IV collagen is an activating ligand for the adhesion G protein-coupled receptor GPR126*. *Sci Signal*, 2014. **7**(338): p. ra76.
76. Sigoillot, S.M., et al., *Adhesion GPCRs as Novel Actors in Neural and Glial Cell Functions: From Synaptogenesis to Myelination*. *Handb Exp Pharmacol*, 2016. **234**: p. 275-298.
77. Küffer, A., et al., *The prion protein is an agonistic ligand of the G protein-coupled receptor Adgrg6*. *Nature*, 2016. **536**(7617): p. 464-468.
78. Bayin, N.S., et al., *GPR133 (ADGRD1), an adhesion G-protein-coupled receptor, is necessary for glioblastoma growth*. *Oncogenesis*, 2016. **5**(10): p. e263.
79. Frenster, J.D., et al., *Functional impact of intramolecular cleavage and dissociation of adhesion G protein-coupled receptor GPR133 (ADGRD1) on canonical signaling*. *J Biol Chem*, 2021. **296**: p. 100798.
80. Frenster, J.D., et al., *Expression profiling of the adhesion G protein-coupled receptor GPR133 (ADGRD1) in glioma subtypes*. *Neurooncol Adv*, 2020. **2**(1): p. vdaa053.
81. Frenster, J.D., et al., *PTK7 is a positive allosteric modulator of GPR133 signaling in glioblastoma*. *Cell Rep*, 2023. **42**(7): p. 112679.
82. Stephan, G., et al., *Activation of the adhesion G protein-coupled receptor GPR133 by antibodies targeting its N-terminus*. *J Biol Chem*, 2022. **298**(6): p. 101949.
83. Terakita, A., E. Kawano-Yamashita, and M. Koyanagi, *Evolution and diversity of opsins*. *Wiley Interdisciplinary Reviews: Membrane Transport and Signaling*, 2012. **1**(1): p. 104-111.
84. Koyanagi, M. and A. Terakita, *Diversity of animal opsin-based pigments and their optogenetic potential*. *Biochim Biophys Acta*, 2014. **1837**(5): p. 710-6.
85. Sekharan, S. and K. Morokuma, *Why 11-cis-retinal? Why not 7-cis-, 9-cis-, or 13-cis-retinal in the eye?* *J Am Chem Soc*, 2011. **133**(47): p. 19052-5.
86. Ernst, O.P., et al., *Microbial and animal rhodopsins: structures, functions, and molecular mechanisms*. *Chem Rev*, 2014. **114**(1): p. 126-63.
87. Palczewski, K., et al., *Crystal structure of rhodopsin: A G protein-coupled receptor*. *Science*, 2000. **289**(5480): p. 739-45.
88. Terakita, A., et al., *Counterion displacement in the molecular evolution of the rhodopsin family*. *Nat Struct Mol Biol*, 2004. **11**(3): p. 284-9.
89. Sakmar, T.P., R.R. Franke, and H.G. Khorana, *Glutamic acid-113 serves as the retinylidene Schiff base counterion in bovine rhodopsin*. *Proc Natl Acad Sci U S A*, 1989. **86**(21): p. 8309-13.
90. Sakmar, T.P., R.R. Franke, and H.G. Khorana, *The role of the retinylidene Schiff base counterion in rhodopsin in determining wavelength absorbance and Schiff base pKa*. *Proc Natl Acad Sci U S A*, 1991. **88**(8): p. 3079-83.
91. Koyanagi, M. and A. Terakita, *Gq-coupled rhodopsin subfamily composed of invertebrate visual pigment and melanopsin*. *Photochem Photobiol*, 2008. **84**(4): p. 1024-30.
92. Hagio, H., et al., *Optogenetic manipulation of Gq- and Gi/o-coupled receptor signaling in neurons and heart muscle cells*. *eLife*, 2023. **12**: p. e83974.

93. Rodgers, J., et al., *Using a bistable animal opsin for switchable and scalable optogenetic inhibition of neurons*. EMBO Rep, 2021. **22**(5): p. e51866.
94. Rost, B.R., et al., *Optogenetics at the presynapse*. Nat Neurosci, 2022. **25**(8): p. 984-998.
95. Wagdi, A., et al., *Selective optogenetic control of G(q) signaling using human Neuropsin*. Nat Commun, 2022. **13**(1): p. 1765.
96. Nagata, T., et al., *Depth Perception from Image Defocus in a Jumping Spider*. Science, 2012. **335**(6067): p. 469-471.
97. Koyanagi, M., et al., *Molecular evolution of arthropod color vision deduced from multiple opsin genes of jumping spiders*. J Mol Evol, 2008. **66**(2): p. 130-7.
98. Nagata, T., et al., *The counterion–retinylidene Schiff base interaction of an invertebrate rhodopsin rearranges upon light activation*. Communications Biology, 2019. **2**(1): p. 180.
99. Hardie, R.C. and P. Raghu, *Visual transduction in Drosophila*. Nature, 2001. **413**(6852): p. 186-93.
100. Ehrenberg, D., et al., *The Two-Photon Reversible Reaction of the Bistable Jumping Spider Rhodopsin-1*. Biophysical Journal, 2019. **116**(7): p. 1248-1258.
101. Milligan, G. and E. Kostenis, *Heterotrimeric G-proteins: a short history*. Br J Pharmacol, 2006. **147 Suppl 1**(Suppl 1): p. S46-55.
102. Huang, S.K., et al., *Mapping the conformational landscape of the stimulatory heterotrimeric G protein*. Nature Structural & Molecular Biology, 2023. **30**(4): p. 502-511.
103. Morris, A.J. and C.C. Malbon, *Physiological regulation of G protein-linked signaling*. Physiol Rev, 1999. **79**(4): p. 1373-430.
104. Ross, E.M. and T.M. Wilkie, *GTPase-activating proteins for heterotrimeric G proteins: regulators of G protein signaling (RGS) and RGS-like proteins*. Annu Rev Biochem, 2000. **69**: p. 795-827.
105. Sato, M., et al., *Accessory proteins for G proteins: partners in signaling*. Annu Rev Pharmacol Toxicol, 2006. **46**: p. 151-87.
106. Umemori, H., et al., *Activation of the G protein Gq/11 through tyrosine phosphorylation of the alpha subunit*. Science, 1997. **276**(5320): p. 1878-81.
107. Gu, J.L., et al., *Regulation of hematopoietic-specific G-protein Galpha15 and Galpha16 by protein kinase C*. J Cell Biochem, 2003. **88**(6): p. 1101-11.
108. Manganello, J.M., et al., *Protein kinase A-mediated phosphorylation of the Galpha13 switch I region alters the Galphabeta13-G protein-coupled receptor complex and inhibits Rho activation*. J Biol Chem, 2003. **278**(1): p. 124-30.
109. Chakravorty, D. and S.M. Assmann, *G protein subunit phosphorylation as a regulatory mechanism in heterotrimeric G protein signaling in mammals, yeast, and plants*. Biochem J, 2018. **475**(21): p. 3331-3357.
110. Jennings, B.C. and M.E. Linder, *Chapter 200 - Regulation of G Proteins by Covalent Modification*, in *Handbook of Cell Signaling (Second Edition)*, R.A. Bradshaw and E.A. Dennis, Editors. 2010, Academic Press: San Diego. p. 1629-1633.
111. Dupré, D.J., et al., *The role of Gbetagamma subunits in the organization, assembly, and function of GPCR signaling complexes*. Annu Rev Pharmacol Toxicol, 2009. **49**: p. 31-56.
112. Tang, W.J. and A.G. Gilman, *Type-specific regulation of adenylyl cyclase by G protein beta gamma subunits*. Science, 1991. **254**(5037): p. 1500-3.
113. Zhang, S., et al., *A C-terminal mutant of the G protein beta subunit deficient in the activation of phospholipase C-beta*. J Biol Chem, 1996. **271**(33): p. 20208-12.
114. Logothetis, D.E., et al., *The beta gamma subunits of GTP-binding proteins activate the muscarinic K<sup>+</sup> channel in heart*. Nature, 1987. **325**(6102): p. 321-6.
115. Zamponi, G.W., et al., *Crosstalk between G proteins and protein kinase C mediated by the calcium channel alpha1 subunit*. Nature, 1997. **385**(6615): p. 442-6.

116. Yoon, E.J., et al., *Gbetagamma interferes with Ca<sup>2+</sup>-dependent binding of synaptotagmin to the soluble N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) complex*. *Mol Pharmacol*, 2007. **72**(5): p. 1210-9.
117. Tsai, C.-J., et al., *Cryo-EM structure of the rhodopsin-Gai-βγ complex reveals binding of the rhodopsin C-terminal tail to the gβ subunit*. *eLife*, 2019. **8**: p. e46041.
118. Zhang, Y., et al., *Cryo-EM structure of the activated GLP-1 receptor in complex with a G protein*. *Nature*, 2017. **546**(7657): p. 248-253.
119. Maeda, S., et al., *Development of an antibody fragment that stabilizes GPCR/G-protein complexes*. *Nature Communications*, 2018. **9**(1): p. 3712.
120. Carpenter, B. and C.G. Tate, *Engineering a minimal G protein to facilitate crystallisation of G protein-coupled receptors in their active conformation*. *Protein Eng Des Sel*, 2016. **29**(12): p. 583-594.
121. Carpenter, B. and C.G. Tate, *Expression and Purification of Mini G Proteins from Escherichia coli*. *Bio Protoc*, 2017. **7**(8).
122. Maeda, S., et al., *Crystallization scale preparation of a stable GPCR signaling complex between constitutively active rhodopsin and G-protein*. *PLoS One*, 2014. **9**(6): p. e98714.
123. Wessling-Resnick, M. and G.L. Johnson, *Transducin interactions with rhodopsin. Evidence for positive cooperative behavior*. *J Biol Chem*, 1987. **262**(26): p. 12444-7.
124. Abdulrahman, W., et al., *A set of baculovirus transfer vectors for screening of affinity tags and parallel expression strategies*. *Anal Biochem*, 2009. **385**(2): p. 383-5.
125. Punjani, A., et al., *cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination*. *Nat Methods*, 2017. **14**(3): p. 290-296.
126. Bepler, T., et al., *Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs*. *Nat Methods*, 2019. **16**(11): p. 1153-1160.
127. Albeck, A., et al., *Role of retinal isomerizations and rotations in the photocycle of bacteriorhodopsin*. *Journal of the American Chemical Society*, 1986. **108**(15): p. 4614-4618.
128. Kimanius, D., et al., *New tools for automated cryo-EM single-particle analysis in RELION-4.0*. *Biochem J*, 2021. **478**(24): p. 4169-4185.
129. Zheng, S.Q., et al., *MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy*. *Nat Methods*, 2017. **14**(4): p. 331-332.
130. Punjani, A. and D.J. Fleet, *3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM*. *J Struct Biol*, 2021. **213**(2): p. 107702.
131. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis*. *J Comput Chem*, 2004. **25**(13): p. 1605-12.
132. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution*. *Acta Crystallogr D Biol Crystallogr*, 2010. **66**(Pt 2): p. 213-21.
133. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics*. *Acta Crystallogr D Biol Crystallogr*, 2004. **60**(Pt 12 Pt 1): p. 2126-32.
134. Schrödinger, L.L.C. and D. Warren, *PyMOL*.
135. Grubbs, F.E., *Procedures for Detecting Outlying Observations in Samples*. *Technometrics*, 1969. **11**(1): p. 1-21.
136. Liang, Y.-L., et al., *Phase-plate cryo-EM structure of a class B GPCR–G-protein complex*. *Nature*, 2017. **546**(7656): p. 118-123.
137. Liang, Y.-L., et al., *Cryo-EM structure of the active, Gs-protein complexed, human CGRP receptor*. *Nature*, 2018. **561**(7724): p. 492-497.
138. Bohnekamp, J. and T. Schöneberg, *Cell adhesion receptor GPR133 couples to Gs protein*. *J Biol Chem*, 2011. **286**(49): p. 41912-41916.
139. Ping, Y.-Q., et al., *Structures of the glucocorticoid-bound adhesion receptor GPR97–Go complex*. *Nature*, 2021. **589**(7843): p. 620-626.
140. Leon, K., et al., *Structural basis for adhesion G protein-coupled receptor Gpr126 function*. *Nature Communications*, 2020. **11**(1): p. 194.

141. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold*. Nature, 2021. **596**(7873): p. 583-589.
142. Ravelli, R.B.G., et al., *Cryo-EM structures from sub-nl volumes using pin-printing and jet vitrification*. Nature Communications, 2020. **11**(1): p. 2563.
143. D'Imprima, E., et al., *Protein denaturation at the air-water interface and how to prevent it*. Elife, 2019. **8**.
144. Noble, A.J., et al., *Reducing effects of particle adsorption to the air-water interface in cryo-EM*. Nat Methods, 2018. **15**(10): p. 793-795.
145. Li, J., et al., *Structure of bovine rhodopsin in a trigonal crystal form*. J Mol Biol, 2004. **343**(5): p. 1409-38.
146. Deupi, X., et al., *Stabilized G protein binding site in the structure of constitutively active metarhodopsin-II*. Proceedings of the National Academy of Sciences, 2012. **109**(1): p. 119-124.
147. Choe, H.-W., et al., *Crystal structure of metarhodopsin II*. Nature, 2011. **471**(7340): p. 651-655.
148. Tsai, C.-J., et al., *Crystal structure of rhodopsin in complex with a mini-G<sub>o</sub> sheds light on the principles of G protein selectivity*. Science Advances, 2018. **4**(9): p. eaat7052.
149. Singhal, A., et al., *Structural role of the T94I rhodopsin mutation in congenital stationary night blindness*. EMBO reports, 2016. **17**(10): p. 1431-1440.
150. Rose, A.S., et al., *Position of Transmembrane Helix 6 Determines Receptor G Protein Coupling Specificity*. Journal of the American Chemical Society, 2014. **136**(32): p. 11244-11247.
151. Nygaard, R., et al., *The dynamic process of  $\beta(2)$ -adrenergic receptor activation*. Cell, 2013. **152**(3): p. 532-42.
152. Maeda, S., et al., *Structures of the M1 and M2 muscarinic acetylcholine receptor/G-protein complexes*. Science, 2019. **364**(6440): p. 552-557.
153. Kato, H.E., et al., *Conformational transitions of a neurotensin receptor 1–Gi1 complex*. Nature, 2019. **572**(7767): p. 80-85.
154. Shen, C., et al., *Structural basis of GABAB receptor–Gi protein coupling*. Nature, 2021. **594**(7864): p. 594-598.
155. Zhang, M., et al., *Cryo-EM structure of an activated GPCR–G protein complex in lipid nanodiscs*. Nature Structural & Molecular Biology, 2021. **28**(3): p. 258-267.
156. Xing, C., et al., *Cryo-EM Structure of the Human Cannabinoid Receptor CB2-G(i) Signaling Complex*. Cell, 2020. **180**(4): p. 645-654.e13.
157. Wall, M.A., et al., *The structure of the G protein heterotrimer Gi alpha 1 beta 1 gamma 2*. Cell, 1995. **83**(6): p. 1047-58.
158. Okada, T., K. Takeda, and T. Kouyama, *Highly selective separation of rhodopsin from bovine rod outer segment membranes using combination of divalent cation and alkyl(thio)glucoside*. Photochem Photobiol, 1998. **67**(5): p. 495-9.
159. Robertson, M.J., et al., *Structure determination of inactive-state GPCRs with a universal nanobody*. Nature Structural & Molecular Biology, 2022. **29**(12): p. 1188-1195.
160. Seven, A.B., et al., *Structures of G $\alpha$  Proteins in Complex with Their Chaperone Reveal Quality Control Mechanisms*. Cell Rep, 2020. **30**(11): p. 3699-3709.e6.
161. Mahalingam, M. and R. Vogel, *The all-trans-15-syn-retinal chromophore of metarhodopsin III is a partial agonist and not an inverse agonist*. Biochemistry, 2006. **45**(51): p. 15624-32.
162. Kaya, A.I., et al., *A conserved phenylalanine as a relay between the  $\alpha 5$  helix and the GDP binding region of heterotrimeric Gi protein  $\alpha$  subunit*. J Biol Chem, 2014. **289**(35): p. 24475-87.
163. Flock, T., et al., *Selectivity determinants of GPCR–G-protein binding*. Nature, 2017. **545**(7654): p. 317-322.
164. Inoue, A., et al., *Illuminating G-Protein-Coupling Selectivity of GPCRs*. Cell, 2019. **177**(7): p. 1933-1947.e25.

165. Jastrzebska, B., K. Palczewski, and M. Golczak, *Role of bulk water in hydrolysis of the rhodopsin chromophore*. J Biol Chem, 2011. **286**(21): p. 18930-7.
166. Janz, J.M. and D.L. Farrens, *Role of the retinal hydrogen bond network in rhodopsin Schiff base stability and hydrolysis*. J Biol Chem, 2004. **279**(53): p. 55886-94.
167. Wijayaratna, D., et al., *Engineered blue-shifted melanopsins for subcellular optogenetics*. bioRxiv, 2023: p. 2023.10.07.561352.
168. Zhang, X.C., Y. Zhou, and C. Cao, *Proton transfer during class-A GPCR activation: do the CWxP motif and the membrane potential act in concert?* Biophysics Reports, 2018. **4**(3): p. 115-122.
169. Cherezov, V., et al., *High-Resolution Crystal Structure of an Engineered Human  $\beta_2$ -Adrenergic G Protein-Coupled Receptor*. Science, 2007. **318**(5854): p. 1258-1265.
170. Tsukamoto, H., et al., *The Magnitude of the Light-induced Conformational Change in Different Rhodopsins Correlates with Their Ability to Activate G Proteins\**. Journal of Biological Chemistry, 2009. **284**(31): p. 20676-20683.
171. Scheres, S.H.W., *RELION: Implementation of a Bayesian approach to cryo-EM structure determination*. Journal of Structural Biology, 2012. **180**(3): p. 519-530.
172. Grant, T., A. Rohou, and N. Grigorieff, *cisTEM, user-friendly software for single-particle image processing*. Elife, 2018. **7**.
173. Goddard, T.D., et al., *UCSF ChimeraX: Meeting modern challenges in visualization and analysis*. Protein Sci, 2018. **27**(1): p. 14-25.



## 10 Annexed Documents

### 10.1 Curriculum Vitae

#### **Curriculum vitae**

---



#### **Personal Details:**

Name: Tejero  
First Name: Oliver Miguel  
Address: Weinberstrasse 3D, 5417 Untersiggenthal  
Telephone: +41 79 660 88 21  
E-Mail: oliver.tejero@hotmail.com  
Date of Birth: 28.01.1995  
Nationality: Swiss  
LinkedIn: [www.linkedin.com/in/oliver-tejero-642a0118a](http://www.linkedin.com/in/oliver-tejero-642a0118a)  
ORCID ID: 0000-0002-2905-7006

#### **Education:**

02/2020-04/2024 Doctor of Sciences (PhD) of ETH Zurich  
PhD Thesis in the group of Prof. Gebhard Schertler at  
ETH Zurich/Paul Scherrer Institute  
Title of the PhD thesis: "Biochemical and structural  
characterization of GPCR-G protein complexes"

09/2017-11/2019 Master's degree in biochemistry at University of Zurich;  
Master thesis in the group of Prof. Dr. Martin Jinek  
Title of the thesis: "Characterization of Type III CRISPR-  
associated Protein Csm6"

09/2014-06/2017 Bachelor's degree in biochemistry at University of Zurich

08/2010-06/2014 Kantonsschule Baden

## **Research and Working Experience:**

02/2020-04/2024	PhD Thesis at Paul Scherrer Institute, Villigen PSI and Eidgenössische Technische Hochschule Zurich (ETHZ)
10/2018-07/2019	Master thesis at University of Zurich
10/2017-12/2018	FIFA World Football Museum Zurich Security

## **Personal skills:**

### **Language skills:**

German:	Mother tongue
English:	fluent
French:	intermediate

### **Scientific skills:**

#### Protein expression and purification:

Expression hosts: bacterial cells, insect cells, mammalian cells; native source (bovine retinas)

Protein types: Soluble (secreted and intracellular) and membrane proteins

Purification methods: Affinity chromatography, Size exclusion chromatography, Ion exchange chromatography

#### Biochemical and biophysical characterization of proteins

Thermal stability assays

Activity assays

Binding assays

UV-Vis spectroscopy

#### Molecular Biology

PCR

DNA extraction and purification

Site directed mutagenesis

#### X-ray crystallography

Automated crystallization screens

Manual crystallization

Sample preparation

Data collection at SLS (PSI) and SwissFEL (PSI)

#### Single particle analysis by cryo-EM

Sample and cryo-EM grid preparation

Titan Krios operator license at ScopeM ETHZ

Data processing (Cryosparc, Relion, Chimera, Coot, Phenix)

#### Scientific workshops

EMBO workshop on Molecular Neurobiology 2021

Instruct course on Single Particle Analysis by Cryo-EM 2021

### **Transferable skills:**

#### Software skills:

Microsoft office applications (Word, Excel, PowerPoint, Outlook)

Adobe Illustrator

Graphpad Prism  
PyMol  
Basic knowledge with Command-Line Interface of windows and linux  
Project management  
Scientific presentation  
Poster presentation (Synapsis Forum 2020/2022, 4<sup>th</sup> ERNEST conference 2020, International Retinal Protein Conference 2022, 4<sup>th</sup> GPCRnet International Conference 2022)  
Scientific talk at Synapsis Forum 2021  
Interdisciplinary Collaboration  
Supervision of students  
ETH students from "Blockkurse" "Membrane Protein Biology" and "Multigene Expression in Mammalian Cells"

## Publications

Carmela Garcia-Doval, Frank Schwede, Christian Berk, Jakob T. Rostøl, Ole Niewoehner, **Oliver Tejero**, Jonathan Hall, Luciano A. Marraffini & Martin Jinek. Activation and self-inactivation mechanisms of the cyclic oligoadenylate-dependent CRISPR ribonuclease Csm6. *Nat Commun* **11**, 1596 (2020). <https://doi.org/10.1038/s41467-020-15334-5>

Thomas Gruhl, Tobias Weinert, Matthew J. Rodrigues, Christopher J. Milne, Giorgia Ortolani, Karol Nass, Eriko Nango, Saumik Sen, Philip J. M. Johnson, Claudio Cirelli, Antonia Furrer, Sandra Mous, Petr Skopintsev, Daniel James, Florian Dworkowski, Petra Bãth, Demet Kekilli, Dmitry Ozerov, Rie Tanaka, Hannah Glover, Camila Bacellar, Steffen Brünle, Cecilia M. Casadei, Azeglio D. Diethelm, Dardan Gashi, Guillaume Gotthard, Ramon Guixà-González, Yasumasa Joti, Victoria Kabanova, Gregor Knopp, Elena Lesca, Pikyee Ma, Isabelle Martiel, Jonas Mühle, Shigeki Owada, Filip Pamula, Daniel Sarabi, **Oliver Tejero**, Ching-Ju Tsai, Niranjana Varma, Anna Wach, Sébastien Boutet, Kensuke Tono, Przemyslaw Nogly, Xavier Deupi, So Iwata, Richard Neutze, Jörg Standfuss, Gebhard Schertler & Valerie Panneels. Ultrafast structural changes direct the first molecular events of vision. *Nature* **615**, 939–944 (2023). <https://doi.org/10.1038/s41586-023-05863-6>

Three manuscripts are currently in preparation/under review (one first author, two co-author). Please check ORCID page for updates.

## References:

- Prof. em. Gebhard F.X. Schertler, Department of Biology and Chemistry, Paul Scherrer Institute, +41 56 310 42 65, Email: [gebhard.schertler@psi.ch](mailto:gebhard.schertler@psi.ch)
- Dr. Ching-Ju Tsai, Department of Biology and Chemistry, Paul Scherrer Institute, +41 56 310 54 84, Email: [ching-ju.tsai@psi.ch](mailto:ching-ju.tsai@psi.ch)
- Prof. Dr. Martin Jinek, Department of Biochemistry, University of Zurich; Tel.: +41 44 635 5572; Email: [jinek@bioc.uzh.ch](mailto:jinek@bioc.uzh.ch)