universität wien

MASTERARBEIT

„Expression and Refolding of human GPCRs for NMR Studies”

verfasst von

Daniel Mayer BSc

angestrebter akademischer Grad

Master of Science (MSc)

Wien, 2013

Studienkennzahl lt. Studienblatt:	A 066 862
Studienrichtung lt. Studienblatt:	Masterstudium Chemie
Betreut von:	Prof. Dr. Gebhard F. X. Schertler

Acknowledgement

I am truly indebted and thankful to my parents, Anton and Ulrike Mayer, who made it possible to get a good education for my future life. Mum and Dad did not only provide financial support but also a very nice act of friendship. They also were always stand right next to me and helped me in difficult situations of my life. Thanks, without you both it would not be possible what I have done.

It is also a great pleasure to thank all my friends who I will not like to miss in my life. We had a really nice time together and I hope it will stay how it was, also if we are going in different directions: Sophie, Julia, Lisa, Alexander, Markus, Steffi, Chris, Chris & Chris, Mathias, Thomas, Patrick,….

I am sincerely and heartily grateful to my supervisor, Dmitry Veprintsev, for his great support during my master's thesis. He was the person who made this project possible. I would also like to thank Gebhard Schertler and all my colleges from the Laboratory of Biomolecular Research at the Paul Scherrer Institut who have been a great help during my thesis.

Daniel

Declaration

I declare that the work presented here is, to the best of my knowledge and belief, original and the result of my own investigations, except as acknowledged, and has not been submitted, either in part or as a whole, for a degree at this or any other university.

Formulations and ideas taken from other sources are cited as such. This work has not been published.

Date, 2013.08.01

Daniel Mayer BSc.

Abstract

G-protein-coupled receptors (GPCRs) are membrane proteins which sense a variety of extracellular stimuli and transmit the signal into the cell. They are responsible for the controlling of many important functions in the organism, and their malfunction is associated with a number of diseases such as including cancer, cardiac dysfunction, central nervous system disorders, obesity, inflammation and pain. They are targeted by approximately a third of currently used drugs. In order to understand how GPCRs function at molecular level, we need to know their structure and dynamics. NMR is most powerful and established technique to study dynamics of proteins. To do NMR measurements, proteins need to be labelled with NMR-active isotopes. *E. coli* is the most widely and versatile expression system to produce labelled proteins. I have expressed the human GPCRs wild-type version of vasopressin V2 receptor (V2R) and a stabilised version of β1-adrenergic receptor (beta1AR-FB1) in *E. coli* in inclusion bodies in milligram amounts. A purification protocol was established which delivered high yields and purity of unfolded GPCRs. In order to develop a refolding protocol, I have designed a small-scale screen and identified initial refolding conditions. I characterised the refolded samples by stability measurements, radio-ligand binding assay, mass spectrometry, size-exclusion chromatography and analytical ultracentrifugation. The results suggest that the refolding strategy is a promising approach, and with additional optimisation will result in samples suitable for NMR studies.

Aim of the Project

GPCRs are a class of membrane proteins involved in cellular signalling. They are responsible for signal transduction which is taking place across the cell membranes (Millar & Newton, 2010). The stimulation targets are ranging from small molecules such as photons, ions, neurotransmitter over hormones to big proteins (Kristiansen, 2004). Especially the small molecules which can be easily transported by our vascular system make them an interesting target for pharmaceutical industry. Today several high resolution crystal structures of GPCRs are known, mostly from the Class-A family of *rhodopsin-like* GPCRs (Venkatakrishnan et al., 2013). Recently, a better understanding how GPCRs might work was delivered by the first x-ray structure of an activated GPCR with its ligand (Rosenbaum, Rasmussen, & Kobilka, 2009). Even though X-ray crystallography provided us with huge sets of information about GPCRs, NMR would improve our understanding for dynamics of GPCRs which is not possible with crystallography. This technique allows us to study protein-ligand, protein-protein interaction and conformational changes of those highly flexible targets in real-time. If we want to study GPCRs by NMR it is necessary to produce isotopically-labelled recombinant protein. This NMR-active isotope enrichment is limited to single amino acid labelling in mammalian and insect expression systems which are well-established for GPCRs (Saxena, Dutta, Klein-Seetharaman, & Schwalbe, 2012). Study of larger proteins requires deuteration which is not possible in mammalian or insect cells. However, deuteration is feasible in *E. coli*. An overexpression of GPCRs in *E. coli* usually leads to aggregation of the protein and the formation of inclusion bodies (Banères, Popot, & Mouillac, 2011). One of the exception is the neurotensin receptor which was functionally expressed in the inner membrane of *E. coli* (White, Trinh, Shiloach, & Grisshammer, 2004). It was shown that refolding of GPCRs is possible from inclusion bodies. Recently also the first NMR spectrum of the refolded chemokine receptor CXCR1 was published. The amino acids of the wild-type chemokine receptor were fully assigned in this publication (Park, Das, et al., 2012). The host organism *E. coli* is usually the expression system of choice for NMR studies because a variety of labelling strategies are possible. In addition, it is the most cost effective expression system.

The aim of the project was to establish an expression and purification protocol for GPCRs in *E. coli*. The additional aim was to develop a refolding screen for GPCRs. The refolding screen should allow probing of several different refolding conditions in parallel. The success of refolding was evaluated by several biophysical assays. The final aim was to record initial NMR spectra of refolded proteins. Two GPCRs from the Class-A family were used as target molecules for this thesis. The first GPCR is a stabilised version of the β1-Adrenergic Receptor (beta1AR-FB1) and the second target is the wild-type version of human Vasopressin V2 receptor (V2R).

Table of contents

Abbreviations

α5I	alpha-5 Integrin
Beta1AR	β1-adrenergic receptor
$BINDPO_4$	phosphate binding buffer for IMAC purification
BL21(DE3)	competent E. coli cells
βME	β-Mercaptoethanol
CPM	thiol-specific fluorochrome IMAC purification
cfu	colony formation unit
c.o.	codon optimized
DBIS	dilution buffer for inclusion bodies purification
E. coli	Escherichia coli
EDTA	ethylendiaminetetraacetic acid
EL	extracellular loop
ElutionIMAC	elution buffer for IMAC purification
$ElutionPO_4$	phosphate elution buffer for
Equ buffer	equilibration buffer for IMAC purification
FB1	stabilised by Florian Brückner
GPCR	G-protein-coupled receptor
H8	His-tag with eight histidines
IB	inclusion body
IB_A	Inclusion body wash buffer A
IB_S	Inclusion body solubilisation buffer
IMAC	immobilized metal affinity chromatography
IMP	integral membrane protein
IL	intracellular loop
LB medium	Lysogeny Broth medium
M9 medium	minimal growth medium
MB	magnetic bead
NTA	nitrilotriacetic acid
pET-21a(+)	expression vector
PIC	Protease Inhibitor Cocktail
Rosetta2(DE3)	competent E. coli cells
RLB	radio-ligand binding
RT	room temperature
s.c.	cleavage site
SDS	sodium dodecyl sulphate
SLS	N-lauroyl sarcosine
TB medium	Terrific broth medium
TBS	Tris buffered saline
TBST	TBS + Tween 20
TED	tris(carboxymethyl)-ethylenediamine
TM	transmembrane helical segment
V2R	vasopressin V2 receptor
WashIMAC	wash buffer for IMAC purification

1 Introduction

1.1 G-protein-coupled receptors

G-protein-coupled receptors (GPCRs) are the largest family of eukaryotic membrane-bound proteins; the human genome encodes more than 800 different GPCRs. They are responsible for the initiation of a wide range of cellular responses by transmitting signals from the extracellular side to the inside of the cell (Grisshammer, 2013; Hiller, Kühhorn, & Gmeiner, 2013). GPCRs are involved in a lot of diseases including cancer, cardiac dysfunction, diabetes mellitus, central nervous system disorders, obesity, inflammation and pain (Grisshammer, 2013). GPCRs relay a variety of different signals initiated by hormones and neurotransmitters. Ligand binding to one receptor may induce different responses depending on the type of ligand. They are in charge of different functions like vision, olfaction and taste that make life worth living (Rosenbaum et al., 2009). These integral membrane proteins (IMPs) are subdivided in different classes by either sequence identity, ligand binding, signalling, pharmacologic response or a combination of those (Lagerström & Schiöth, 2008). In 1994 Kolakowski introduced a classification system that is based on structural similarity to prototypical GPCRs. It is built up of six classes: Class A (*Rhodopsin-like*), Class B (*Secretin receptor family*), Class C (*Metabotropic Glutamate/Pheromone*), Class D (*Fungal pheromone*) Class E (*Cyclic AMP Receptors*) and Class F (*Frizzled/Smoothened*) (Kolakowski, 1994). An additional nomenclature system has been introduced only five years later by Bockaert and Pin which classifies GPCRs into five families on the basis of structural and ligand-binding criteria. The biggest and most diverse family is represented by the rhodopsin family, which is characterised by shared structural features and similarity of the activation mechanisms (Bockaert & Pin, 1999).

All GPCRs share one important structural feature: they all have seven membrane-spanning α-helical segments that are separated by alternating intracellular and extracellular loop regions of varying length (Figure 1). Almost 30% of all available drugs on the market target GPCRs and are responsible for approximately half of the revenue of the pharmaceutical industry. GPCRs are involved in heart, muscle and nervous system function which makes them pharmacologically important targets. Because of this, it is very important to get more GPCR structures by x-ray crystallography as well as structural and dynamic information by NMR (Grisshammer, 2013). NMR studies may provide a better understanding how signalling of these highly flexible molecules works.

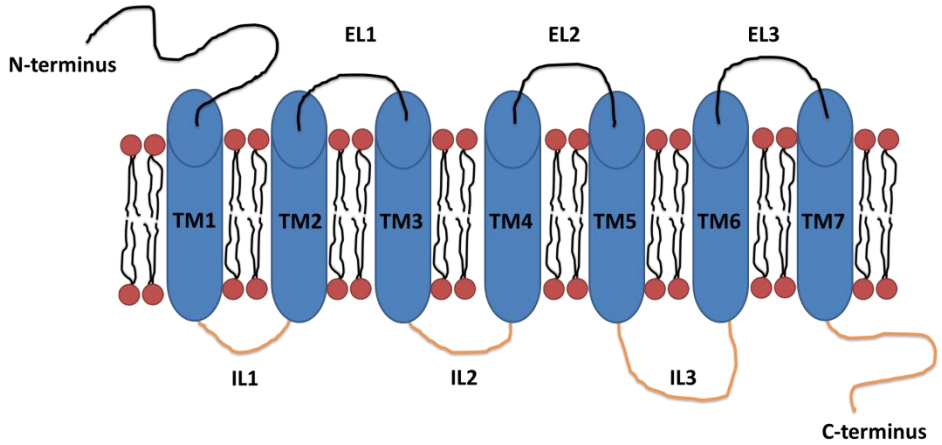

Figure 1: GPCRs are IMPs and are characterised by a seven trans-membrane α-helices (TM1-7) of a single peptide chain which is incorporated in a lipid bilayer membrane. The seven α-helices are connected by alternating extracellular (EL1-3)and intracellular loops (IL1-3). The N-terminus is located at the extracellular side and the C-terminus at the intracellular side.

1.1.1Signalling & dynamics of GPCRs

A significant part of signalling in vertebrate physiology is based on signal transduction by GPCRs. This signal transduction can be mediated by a variety of different compounds such as neurotransmitters, hormones, ions, photons and other kinds of stimuli (Kristiansen, 2004). GPCRs can also be seen as information gate of the cell, being very specific for each targeting compound. Rosenbaum, *et al*. described how the selectivity of the signal of one receptor (β2-adrenergic receptor)results in a variety of responses depending on which drug is binding to the binding pocket (=orthosteric site) (Rosenbaum *et al*., 2009). The early idea of how GPCRs work was rather simple and differs a lot from what can be shown nowadays. It was thought that GPCRs have only two different states; inactive, where no signal is transmitted and active, where transmitting of the signal takes place.

Across all subtypes, GPCRs share a high sequence homology in their orthosteric site. It has been shown that although this homology in their orthosteric site, receptors can activate very specific signalling pathways depending on which ligand has bound. This ligand depended signalling is called biased agonism. It is unlikely that targeting of the sides alone will be the way to go for future drug discovery. (Lane, Abdul-Ridha, & Canals, 2013). In case of the signalling pathway involving heterotrimeric G-proteins, which is the best understood pathway nowadays, the variety of cellular effects is usually explained with the multiple different G-protein combinations, which end up in discrete signalling pathways. The principle behind signal transduction is rather simple: a ligand of the GPCR is binding on the extracellular side of GPCR-complex with heterotrimeric G-protein (G_{α}–subunit with $G_{\beta\gamma}$-subunit). As response of this interaction a conformational change of the GPCR is taking place, which leads to the exchange of the GDP to GTP at the G_{α}-subunit. The result of this exchange is the dissociation of the heterotrimeric G-protein. The free G_{α}-subunit is able to trigger several different signalling cascades like activation of different kinases. Also 37 other regulators of G-proteins signalling (RGS) are known to regulate G-protein signalling pathways.

There are two additional types of proteins which are responsible for the up- and down-regulation of GPCRs: GPCR kinases (GRKs) and β-arrestins (1 & 2). These two groups of proteins are also able to bind to GPCRs. The GRKs is responsible for the phosphorylation of the GPCR at the cytosolic side which can results in the binding of the β-arrestins. This can lead to a termination or reduction of the signalling through the G-protein signalling pathway because the G-protein is blocked from future interaction with the GPCR (Zhang & Eggert, 2013). One of the main interests of our department is to reach a better understanding for *rhodopsin-like* GPCRs.

1.1.2 Targets

This work is based on two GPCRs from the subfamily Class-A (*rhodopsin-like* GPCRs). The wild type of human vasopressin V2 receptor (V2R) and a stabilised version of human β1-adrenergic receptor, which is abbreviated by "beta1AR-FB1" (engineered by Florian Brückner, unpublished), have been used.

Human vasopressin V2 receptor:

The human vasopressin V2 receptor is targeted by the pituitary hormone vasopressin (AVP) which is responsible for the regulation of the water balance in the tissue of vertebrates. Several diseases are correlated with disturbance in the vasopressin action such as congestive heart failure, kidney, liver and brain disease. It is poorly understood how the regulation of these states is taking place (Armstrong, Seeber, Ayoub, Feldman, & Pfleger, 2013). At the moment, there is no crystal or NMR structure of the full length receptor.

Human β1-adrenergic receptor

The human β1-adrenergic receptor belongs to the adrenoceptors and plays a critical role in maintaining the cardiac homeostasis. The β2-adrenergic receptor is involved in the same regulation process and is responsible for the down-regulation of the β1-adrenergic receptor trans-signalling. This protects the cell from dying (Ahmet *et al*., 2004). The structure of the turkey β1-adrenergic receptor is already known. The functional expression and crystallisation of this GPCR required a highly engineered construct (Warne *et al*., 2008, 2011).The construct design of the turkey version was transferred onto the human β1-adrenergic receptor by Florian Brückner to increase his chances to crystallise the human β1-adrenergic receptor. The construct has several optimisations starting from truncated N- and C-termini, stabilising mutations and the replacement of the IL3 by T4 lysozyme.

1.1.3 Available structures

The Protein Data Bank (PDB), depository of all solved protein structures, contains several thousands of structures of soluble proteins but only 280 structures of integral membrane proteins (IMPs) (Banères *et al*., 2011). The first high resolution crystal structure of a GPCR has been solved in the year 2000. The

crystal structure of bovine rhodopsin (2.8 angstroms) was the first structure to show the interaction of a ligand with a GPCR and provided information on how molecular GPCR activation might work (Palczewski *et al*., 2000). Afterwards several structures of aminergic GPCRs, notably avian β1-adrenergic receptor, human β2-adrenergic receptor, human muscarinic acetylcholine 2 receptor have been solved (Figure 2) (Venkatakrishnan *et al*., 2013). Recently, the first x-ray structures of ligand activated GPCRs were published (Rosenbaum *et al*., 2009). In 2012 a NMR spectrum of a full-length GPCR was recorded that allowed a full sequence alignment. The NMR spectrum was done with an unmodified GPCR in phospholipid bilayers which is unprecedented in the field. The phospholipid bilayer simulates a nature-like environment (Park, Das, *et al*., 2012). The structure of this CXCR1 is mostly comparable with what we already known from modelling and other crystal structures of GPCRs. One significant difference is still given: The structure shows a kink in helix six which has not been found in any GPCR structure before. Is this kink now an artefact which is coming from the refolding process or will we find this structural feature in all future NMR structures? It could be possible that this structure characteristic occurs in NMR studies of functional expressed GPCRs from insect and mammalian cell too.

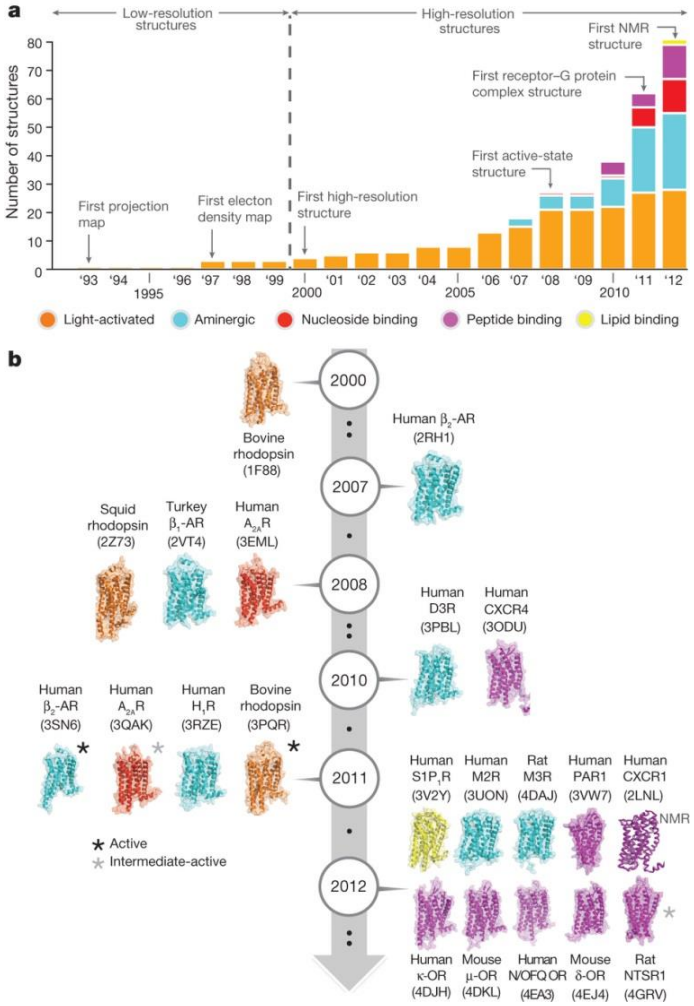

Figure 2: The illustration **A** shows how the numbers of structures of GPCRs have be increased over the last few years. Figure **B** shows which new crystal structures were published recently. Active structures are marked with a black star and intermediate-active with a grey star.

1.2 Nuclear magnetic resonance

Nuclear magnetic resonance (NMR) measurements make use of the phenomenon that nuclei with uneven-numbered spin absorb and re-emit electromagnetic radiation into a magnetic field. This emitted electromagnetic radiation induces a current in the magnet which can be later on translated to a signal using Fourier-transformation. Every isotope which has this uneven-spin (odd number or protons and/or of neutrons) absorbs its specific resonance frequency which allows to distinguish between the different NMR-active isotopes (^{1}H, ^{13}C, ^{15}N, ^{19}F,…). The specific resonance frequency depends on the strength of the magnetic

field and the magnetic properties of each isotope. One important factor is that the resonance frequency is directly proportional to the applied magnetic field.

Measurement of a NMR signal requires three major steps: The nucleuses with uneven spin have to align to the constant magnetic field $H_{0,}$ afterwards perturbation of this alignment has to happen by shooting a pulse sequence into the sample. After perturbation, the realignment is taking place. This leads to an induced current in the static magnetic field H_0 which finally gives information on environment of each disturbed nucleus. As the signal of each nucleus is depending on its environment, structural changes or even the whole structure can be interpreted from these induced currents (Cavanagh *et al*., 2010). Nowadays nuclear magnetic resonance is a well-established technique which is used for structure determination of small molecules up to big proteins and protein complexes, but it was not always like that.

1.2.1Protein NMR

Proteins larger than 30 kDa posed a problem for NMR studies until 1997. This is due to the fact that the relaxation times are much longer for those large macromolecules compared to smaller molecules. NMR is a useful technique to study large molecules. In 1997, a technique which is called transverse relaxation-optimised spectroscopy (TROSY) was developed. The technique allows us to deal with the transverse relaxation in multidimensional NMR studies. The principle behind this technique is the constructive use of interference between anisotropy interaction and dipole-dipole coupling (Pervushin, Riek, Wider, & Wüthrich, 1997).

Especially in case of GPCRs it could deliver additional information which might not be achievable by x-ray crystallography:

- To have a dynamic view on GPCRs during their conformational changes, this would be induced by addition of diverse ligands (agonist, antagonist and biased ligand) to the sample. It could be possible to distinguish between multiple active/inactive and partially active states.
- To discover the ligand entrance channel and the ligand binding site/sites.
- To obtain the dynamic changes and flexibility of the GPCR during the binding process
- To follow the full conformational change of the GPCR between the ground and fully activated state, this might give information on the signalling process and a better understanding how biased signalling can be so specific.
- To study interaction of the heterotrimeric G-protein- and arrestin-complex with the GPCR and get an impression how the forwarding of the signal of different ligands is taking place.
- To analyse the stability of G-protein or arrestin GPCR complexes depending on single amino acid mutations in one of the involved proteins. This is already possible with other techniques, but it would provide additional information on different conformational changes which may obtain a better understanding why this single amino acid mutation has this kind of influence.

To study those interactions and conformational changes it is necessary to do an isotopic labelling of the proteins.

1.2.2Production of isotope labelled proteins

A study of proteins by NMR is only possible with a NMR-active isotope enrichment as mentioned above. Possibilities are single amino acid labelling or ^{15}N-, ^{13}C-, ^{2}H-labelling or a combination of those. Depending on which expression system has been chosen different labelling techniques are possible or not. The expression of GPCRs in mammalian and insect cells is a well-establish technique in most labs, who are working on these targets. IMPs can be functionally expressed in these two systems. The yields are also in an acceptable range (between 100 µg/l to 2 mg/l) for NMR studies for well-expressing constructs. However, the labelling in those expression systems is limited to single amino acid labelling, as the costs of ^{15}N- and ^{13}C-labelled amino acids for uniform labelling are too high. The second limitation is the possible level of deuteration which is reachable in those expression systems. It is only possible to reach up to 30% deuteration in insect and mammalian cells, because the cells are dying if higher levels of deuterium oxide are used in the culture medium. The level of deuteration close to 100% is necessary for significant improvement of NMR properties of large macromolecules.

Yeast represents one of those host organisms which allows to produce recombinant protein with huge diversity of NMR-active labels at reachable costs. It has been already shown in literature that functional folded GPCRs can be expressed in the yeast *Pichia pastoris* (Krettler, Reinhart, & Bevans, 2013). *Pichia pastoris* is capable of glycosylation which can be an advantage or disadvantage. Many GPCRs are naturally glycosylated, but non-uniform glycosylation may also pose a problem for NMR. In some cases, the glycosylation plays a significant role in the functioning of the receptor. This is important for studies with the protein in its natural environment. However, yeast has the tendency hyperglycosylate recombinant proteins, this means that glycosylation occurs at sites which are not naturally glycosylated. It has to be mentioned that yeast is not able to do all types of glycosylation which are appearing in human cells or other higher eukaryotes.

The most commonly used expression host for recombinant production of labelled protein for NMR studies is *E. coli*. It allows the expression of proteins in high yields and the labelling is cheap compared to other expression systems such as insect or mammalian cells. This expression host also allows almost 100% deuteration of the recombinant protein which is very important if high molecular weight proteins are to be studied by NMR (Saxena *et al*., 2012). There are two strategies for expression of GPCRs in *E. coli*: Expression in the inner membrane with fusion proteins which yield in low levels of functional folded GPCRs (Ma, Kubicek, & Labahn, 2013; Weiss & Grisshammer, 2002; White *et al*., 2004). The other possibility is expression of unfolded protein in inclusion bodies (IBs) (Michalke *et al*., 2010). The expression of recombinant protein in IBs usually gives high expression yields. High purity can be achieved with simple purification protocols, because IBs can be easily separated from other components by a few centrifugation steps. It is worth mentioning that in some cases the overexpressed protein can occupy up to 80% of the total proteins in IBs (Figure 3). Sometimes, the recombinant production of the protein is only possible in form of IBs, due to the fact that a few proteins cannot be correctly folded in absence of their original folding machinery. An additional problem is that a few proteins are toxic to their host organism which can be overcome by aggregation of the protein in IBs. In literature, several promot-

ers are described which allow expression of unfolded protein aggregated in IBs at high yields. One of the reported targeting partners for the production of unfolded *rhodopsin-like* GPCRs is the α5-Integrin (*α5I*) fragment at the N-terminus of the expression construct which usually delivers high yield of unfolded protein (Banères *et al*., 2011). The omnipresent problem of this strategy - refolding of GPCRs may not be as easy as for other proteins - has been discussed above.

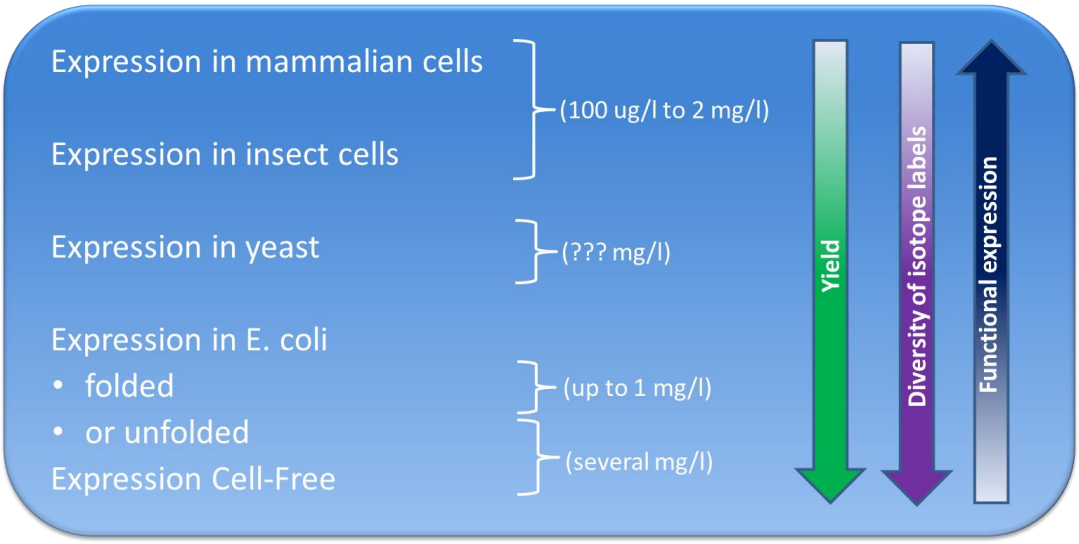

Figure 3: This figure should summarise what expression system can be used for the production of GPCR. It illustrate additionally that the functionality of the expressed protein is usually depending on which system is used for recombinant production. The tendency shows that a choice have to be done between high expression or high yield. In rare possibility you are reaching both.

1.3 Refolding of proteins

Overexpression of heterologous proteins is a common strategy which often leads to the production of dense, spherical, amorphous, insoluble aggregates which are called inclusion bodies. It is usually the case that biologically active protein cannot be isolated from IBs, but there may be correctly folded protein between the aggregate layers of an inclusion body. If very high expression is achieved, the expressed proteins account for 40 – 50 % of the total protein mass of the cell. The amount of the produced IBs in the host organism is depending on the growth condition and the protein which is overexpressed (Kumar, 2011). In order to increase the yield of proteins in IBs, proteins are often expressed as fusion proteins. The most common fusion proteins for functional overexpression of integral membrane proteins in *E. coli* are β-galactosidase (β-gal, 114 kDa) and maltose-binding protein (MBP, 43 kDa) at the N-terminus. Thioredoxin (TRX, 10 kDa) and green fluorescent protein (GFP, 27 kDa) are fusion partners which can lead to functional expression if they are located at the C-terminus of a GPCR fusion construct. Other fusion partners such as α5 integrin fragment (*α5I*, 31 kDa) at the N-terminus and schistosemal glutathione D-transferase (GST, 25 kDa) at the C-terminus have shown to promote the expression of unfolded protein which leads to an accumulation in inclusion bodies (Banères *et al*., 2011).

Harsh denaturation conditions like 8 M urea, 6 M guanidine-hydrochloride or harsh detergents like sodium dodecyl sulphate (SDS) are used to solubilise the protein aggregations. These harsh solubilisation

conditions lead to total unfolding of the proteins if native-like secondary structures already existed in the IBs. This has been shown for several examples already. Several hypotheses on refolding of proteins exist. One hypothesis says that already existing native-like secondary structure elements should be preserved, because they are reducing the interaction between different proteins chains. This reduction in interactions may lead to a higher success rate of refolding due to the fact that less protein is aggregating. Milder chemicals such as n-propanol have to be used to preserve secondary structure elements (Singh *et al*., 2012). This was shown for globular proteins; however, intermolecular interactions are tighter in IMPs. This can lead to accumulation of large hydrophobic interaction regions which complicates efficient refolding. Due to this IMPs usually need harsher detergents for their solubilisation like SDS or SLS (Huang, Bayley, Liao, London, & Khorana, 1981).

Several different refolding strategies are used today which all have certain advantages and disadvantages:

- refolding by dilution
- refolding by dialysis or diafiltration
- refolding using chromatographic techniques (size-exclusion chromatography, ion-exchange chromatography, hydrophobic interaction chromatography, reverse-phase chromatography, liposome chromatography and affinity chromatography)
- refolding using zeolites
- refolding using high hydrostatic pressure

The reason why so many techniques exist is that every protein behaves differently. It is not possible to predict which method should be chosen. Also a lot of different parameters have to be optimized for a successful refolding such as pH, concentration of protein, ionic strength, surfactant, temperature, refolding time, additives, reducing/oxidizing conditions, ligand, cofactors, substrate analoges (Qoronfleh, Hesterberg, & Seefeldt, 2007). At the end, many combinations have to be tried to find the optimal refolding condition resulting in high yield of refolded protein (Kumar, 2011). While refolding of soluble proteins is a well-established technique in a lot of labs, refolding of IMPs remains to be a challenge. Additional problems common to IMPs are discussed below.

1.3.1Refolding of membrane proteins

Due to large hydrophobic regions on the surface of IMPs their refolding is more complicated than the refolding of soluble proteins. These hydrophobic areas have a high tendency to form aggregates in a hydrophilic environment where the refolding is usually taking place. The problem of aggregation is that aggregated proteins cannot be refolded, which decreases the efficiency of refolding process. Nature evolved special machineries which support native folding and the incorporation of these proteins into lipid bilayer membranes. Until today it is not possible to express and reconstitute these machineries *in vitro*, so we have to use different strategies for refolding of IMPs (Figure 4). It has been shown in literature that refolding in detergents (Michalke *et al*., 2010; Park, Casagrande, *et al*., 2012), amphipoles (Bazzacco *et al*., 2012) and lipids/bicells/vesicles (Park, Casagrande, *et al*., 2012) is possible. The effi-

ciency of each folding depends on the competition between 3D-structure formation and aggregation. An important input factor for successful refolding is to employ a surfactant or surfactant mixture that has the ability to stabilise the native 3D-structure of a functional folded IMP (Hans Kiefer, 2003).

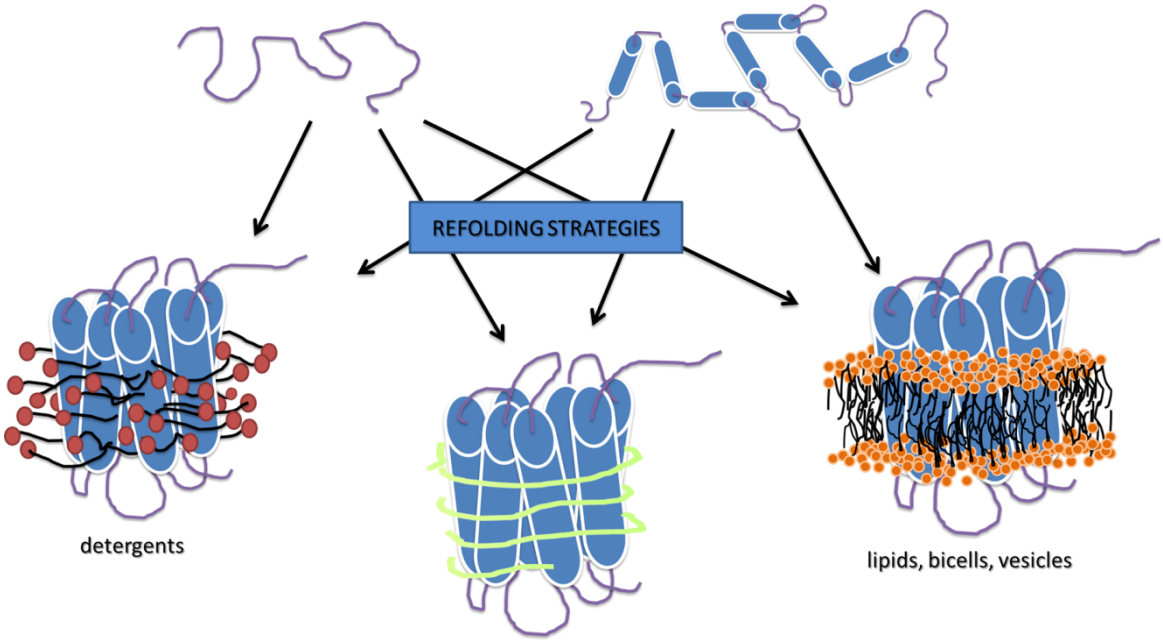

Figure 4: Refolding strategies for IMPs shown by a model of a GPCR: How the unfolded version of a GPCR looks like is not clear at the moment. It could be completely unfolded or have some or all of its secondary structure elements performed after solubilisation.

This surfactant or surfactant mixture has to promote intermolecular interaction of the helices in case of GPCRs. The additional factor is that the surfactants should not only be able to provide assistance for folding, they should also have a stabilisation effect on the folded protein. Favourable folding conditions have been identified already for several IMPs, anyways, the conditions under which the functionality of the refolded protein has been tested has sometimes to be regarded with suspicion. Successful refolding reports are published on functionally folded human leukotriene B4 receptor 1 (BLT1) in 30% lauryl-N,N-dimethylamine-N-oxide (LDAO) (Baneres *et al*., 2003), up to 4% functional folded human leukotriene B4 receptor2 (BLT2) receptor in a mixture of n-dodecyl phosphocholine (DPC) and hexadecyl-β-d-maltoside (HMD) (Arcemisbéhère *et al*., 2010), refolding of OR5 receptor with digitonin and SLS which has been exchanged against phosphatidylcholine (PC), phosphatidylglycerol and dodecyl-β-d-maltoside (H Kiefer *et al*., 1996), refolding of CXCR1 chemokine interleukin-8 receptor by exchanging detergent SDS against DDM and Cymal-6 (Park, Casagrande, *et al*., 2012) or in phospholipid bilayers (Park, Das, *et al*., 2012).

Furthermore, a new technique for refolding of IMPs is the reconstitution using amphipoles. It has been shown to be successful for five different receptors (four published and one unpublished) which had been refolded before using different conditions: BLT1 & 2, serotonin receptor 5-HT_{4a}, cannabinoid receptor CB1 and ghrelin receptor GHSR-1a (unpublished). The amphipol strategy provided a refolding efficiency between 40 to 70% for the tested GPCRs. This refolding efficiency is always higher than the one reported before for the other strategy. (Popot *et al.*, 2011).

All refolding strategies have the problem that membrane proteins have the tendency to form aggregates in a hydrophilic environment. On-column refolding seemed to be a useful technique to minimise aggregation. Successful refolding was done for BLT1 and 5-HT4 receptor using this technique (Baneres *et al*., 2003; Banères *et al*., 2005).

1.3.2On-column refolding

On-column refolding is one of the newer refolding techniques and is based on the specific binding of the protein to a column matrix. Different strategies and principles of how on-column refolding can be done have been published. A very sophisticated innovation of on-column refolding is a resin with an immobilised antibody which acts like a chaperone during the refolding process. The antibody-antigen interaction can promote refolding and also lead to a significant increase in refolding efficiency. It is shown in literature that higher refolding efficiencies can be achieved using this strategy compared to other approaches. The main problem of antibody-assisted refolding is that conditions have to be found where the protein does not aggregate and the antibody does not unfold (Katoh, Kumada, & Maeshima, 2005). A more common strategy of on-column refolding is immobilized metal-ion affinity chromatography (Dong, Chen, & Sun, 2009; Glynou, Ioannou, & Christopoulos, 2003). The big advantage of this strategy is that it can reduce inter-molecular interactions of proteins which generally have the tendency to aggregate, because they are immobilised with one side of the protein on the resin surface. This strategy also allows one-step purification and refolding of the protein. The solubilized protein binds to the column matrix and the denaturant concentration can be reduced by different wash steps. The reducing of the denaturing condition can be as needed; slowly or fast. It is also possible to change the redox system very easily during the refolding process. This may be very important for GPCRs as successful refolding also depends on the correct formation of disulphide bridges. Another advantage is that the refolding is done with highly purified protein which inhibits additional interaction with other proteins during refolding. This may have a significant increase on the refolding yield. The protein does not need to be diluted to low concentrations, which can save a lot of time and resources. The other advantage is that less refolding buffer is needed which can be an important cost factor if expensive detergent has to be used. The elution of the protein can be done after refolding using either imidazole, EDTA or a reduced pH (Dong *et al*., 2009; Kumar, 2011). It might be important that the resin surface can also have an influence on the refolding efficiency which was shown for the on-column refolding of a C-terminal fragment of human alpha-fetoprotein (Sharapova *et al*., 2011). In some cases it might be better to use a refolding technique which was shown to be efficient for most proteins refolded during the last decades.

1.3.3Refolding by dilution

Refolding by dilution is the most used and oldest way how refolding of proteins can be done. The folding process is not controlled by a refolding machinery as in an intact cell. The result of this uncontrolled process is usually a mixture of folded and unfolded protein. The principle of the technique is fast dilution, which decreases the denaturant concentration and forces the polypeptide chain to fold. These procedure leads to a very fast bypassing of the intermediate state compare to other refolding strategies like dialysis. The protein has less time to form inter-molecular interactions and is forced to form a higher percentage of intra-molecular interactions which might lead to a higher refolding efficiency especially for proteins that have a tendency to form aggregates(Kumar, 2011; Tsumoto, Ejima, Kumagai, & Arakawa, 2003; Vallejo & Rinas, 2004).

There are several possibilities for dilution refolding:

a) Solubilized protein can be added to a tank containing refolding buffer. The concentration of the protein increases over the time. The addition of the protein has to be stopped at the point where the denaturant concentrations are too high to allow any further increase in yield of folded protein.
b) Another possibility is the pulsed dilution method, where small amounts of solubilized protein are added over a longer period of time. It has been documented that this refolding strategy can result in yields comparable to those achieved by dilution refolding *method a)*. The theory behind this is that if a small amount of protein refolds to its native conformation, it is out of the game and has a lower tendency to form aggregates.
c) Dilution refolding can also be done the other way around. Refolding buffer is constantly added to solubilise protein in this case. The denaturing conditions are decreased over time which allows the protein to go through all its intermediate states during the refolding process.
d) Refolding by dilution can be also achieved by mixing solubilized protein and refolding buffer at a constant rate, which prevents increasing the concentration of misfolded proteins and may result in a higher yield of refolded protein (Kumar, 2011).

2 Methods

2.1 Expression & purification

2.1.1Constructs

The general construct design was kindly provided by Mouillac's group in Montpellier. They shared with us a variation of construct A (Figure 5) which had a 6xHis-tag instead of a 8xHis-tag. The His-tag was replaced by Dalibor Milic to improve purification purity. The constructs contained an alpha-5 integrin domain (*α5I*) which promotes the expression of *rhodopsin-like* GPCRs in *E. coli* in inclusion bodies (Banères *et al*., 2011). Between the cleavage site/sites was a poly-glycine linker included. Afterwards the linker a cleavage site/sites and particular GPCR sequence was arranged in the sequence. The end of the constructs contains an additional linker with a C-terminus 6xHis-tag for IMAC purification. The whole sequence was inserted in a pET-21a(+) vector which was resistance to ampicillin.

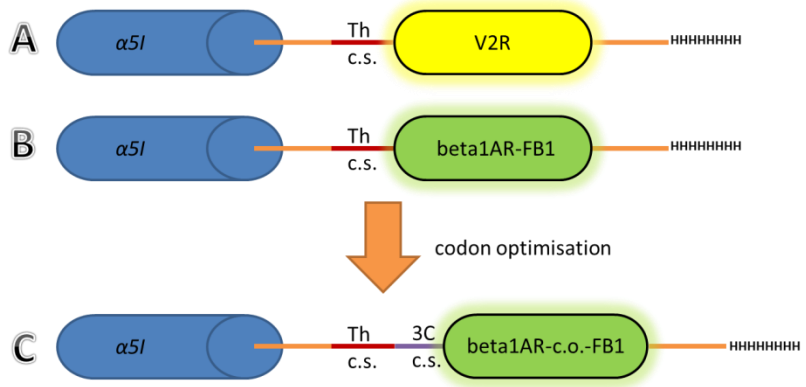

Figure 5: All used construct contained a thrombin cleavage site after a poly-glycine linker. Construct C had an additional 3C-cleavage site after the thrombin cleavage site located. The beta1AR-c.o.-FB1 stands for codon optimised version for the expression in *E. coli* and was ordered from Genewiz Inc. (USA). The construct are called in this thesis: A = *α5I*-V2R-H8, B = *α5I*-beta1AR-FB1-H8 and C = *α5I*-beta1AR-c.o.-FB1-H8. (Th = thrombin, 3C = 3C Protease, c.s. = cleavage site, c.o. = codon optimised)

2.1.2Sequencing

Sequencing was done by GATC Biotech. They use the new single molecule real-time method (SMART^{TM}) for sequencing. This technique has been developed by Pacific Bioscience and allows parallelized single molecule DNA sequencing. The method uses a single DNA polymerase enzyme which is immobilized on a solid support. The polymerase is thereby located on the bottom of tiny well where only "one" single molecule of DNA fits in. A mixture of the four DNA bases is added where each DNA base carries a different fluorescent dye. If the polymerase incorporates a single nucleotide to copy the DNA template, it results in the release of the fluorescent dye. The occurring fluorescence signal is measured in

real-time with the new zero-mode waveguide technology. The colour code allows determination of the DNA sequence.

The beta1AR-FB1 sequence was cloned into the pET-21a(+) vector of the *α5I*-V2R-H8 construct by ligation-independent cloning. The sequencing of the new plasmid was tried at the coding sequence to ensure that no part of the construct had been lost during the PCR reaction. New sequencing primers (Table 1) were designed and ordered from Microsynth AG in Switzerland. For the primer design three important rules were considered: the use of a GC-clamp, length of primers between 20-28 bases and a melting temperature of approximately 60°C.

Table 1: Sequencing primers for *α5I*-beta1AR-FB1 construct in the expression vector pET-21a(+)

Primer name/direction	Sequence	T_m [°C]
DM1/forward	TGATCTGATTGTGGGGTCCTTTGG	61
DM2/reverse	GTCTGGGAGTCAAAGCTATCG	56

2.1.3 Transformation

Transformation is a technique which is used in molecular biology to introduce genetic information into a host cell. A plasmid encoding the gene of interest and an antibiotic resistance is introduced into the cells by heat shocking chemically competent cells. Cells which took up the plasmid can be selected by growth on selective media containing antibiotics.

A 200 µl aliquot of chemically competent *E. coli* cells (BL21(DE3) for construct V2R-H8, beta1AR-FB1-c.o.-H8 and Rosetta2(DE3) for construct beta1AR-FB1-H8) from -80°C stock was thawed on ice. 100 µl of this cell suspension were transferred to a new sterile 1.5 ml eppendorf tube and 1µl of plasmid DNA which had a concentration of 10-80 ng/µl was added. After mixing the cells, they were incubated for 10 to 30 min on ice. A heat shock was done in a heat block for 45 s at 42°C. Afterwards the cell suspension was put on ice immediately for 2 min. 700 µl of preheated SOC media (approximately 37°C) was added and the eppendorf tube was stored in an incubator at 37°C for 45 min. After revitalisation phase 100 µl of this cell suspension was pipetted on a 9.5 cm agar plate which contained the necessary antibiotic resistance (in case of BL21(ED3) 100 µg/ml ampicillin and for Rotetta2(ED3) 100 µg/ml ampicillin and 30 µg/ml chloramphenicol). The plating of the cells was done by adding sterile glass beads and shaking the closed agar plate a few times. The glass beads were discarded and the plate was incubated at 37°C overnight or at RT over the weekend.

2.1.4 Plasmid miniprep

Minipreps are a small scale method for extraction and purification of plasmid DNA from bacterial cultures. The plasmid of interest is transformed into competent cells, a culture is grown and the plasmids are extracted from the bacterial cells. The yield of minipreps depends on the copy number of the plasmid, and the type of competent cells used.

For plasmid DNA preparation XL1-Blue cells ($\geq 2.4*10^7$ cfu/µg) were used. A transformation was done as described above. A single colony was picked and grown in 4 ml LB with 100 µg/ml ampicillin at 37°C overnight. The cells were harvested by centrifugation and the protocol for plasmid preparation (GenElute™ Plasmid Miniprep Kit) from Sigma-Aldrich was followed. The protocol is available on the webpage of Sigma-Aldrich (http://www.sigmaaldrich.com/life-science/molecular-biology/dna-and-rna-purification/plasmid-miniprep-kit.html).

2.1.5 Media

The most common used growth media for bacteria cultures are Lysogeny Broth medium (LB medium), Terrific broth medium (TB medium) and 2x Yeast extract and Tryptone (2xYT broth medium). They are usually delivered as dry powder where water has to be added and an autoclaving step has to be done before use. Two different types of minimal media (M9 media) were used for the expression of ^{15}N enriched V2R-H8 in bacterial cells. The two different media vary in their complexity. In this thesis, they will be called low-complexity M9 medium and high-complexity M9 medium. For the preparation of the media phosphate buffer and sodium chloride were autoclaved in2 l baffled flasks. All other ingredients were filtered through a 0.22 µm syringe filter. The buffer was preheated in presence of antibiotic to approximately 37°C for several hours or overnight. The metals and vitamins were added before the culture was started.

Recipe for low-complexity ^{15}N-label M9 medium:

50	mM	$Na_2HPO_4*7H_2O$
25	mM	KH_2PO_4
10	mM	NaCl
2.5	g/L	D-Glucose-Anhydrogen (Grebu)
1	g/l	$^{15}NH_4Cl$
2	mM	$MgSO_4$
50	µM	$CaCl_2$
50	µM	$FeSO_4*7H_2O$
110	µM	$ZnCl_2$
0.2	%	Yeast Extract
50	µM	Thiamin-hydrochlorid
41	µM	D-Biotin
20	µM	L-Methionine

Recipe for high-complexity ^{15}N-label M9 medium:

50	mM	$Na_2HPO_4*7H_2O$
25	mM	KH_2PO_4
10	mM	NaCl
2.5	g/l	D-Glucose-Anhydrogen (Grebu)
1	g/l	$^{15}NH_4Cl$
2	mM	$MgSO_4$
50	µM	$FeSO_4$
2.50	µM	$CaCl_2*2H_20$
2.07	µM	H_3BO_3
150	nM	$CoCl_2*6H_2O$
50	nM	$CuCl_2*2H_2O$
5	µM	$ZnCl_2$
5	µM	$Na_2MoO_4*2H_2O$
400	nM	$MnCl_2*4H_2O$
200	nM	Thiamin-hydrochlorid
409	nM	D-Biotin
716	nM	Choline chloride
227	nM	Folic acid
819	nM	Niacinaminde
456	nM	D-Pantothenic acid
27	nM	Riboflavin

2.1.6 Overexpression of proteins in inclusion bodies

Expression of unlabelled protein

All expression optimisations and protein productions were done in 2 l flasks. At the beginning of the practical work also normal 2 l Erlenmeyer flask had been used, but later during the practical work only 2 l baffled flask were used, because of better oxygenation.

The cultures were started either with a cell suspension directly from an agar plate or with a starting culture which had been grown in LB medium at 37 or 25°C overnight. For starting the culture directly from an agar plate, 5 ml of preheated medium was used to get a cell suspension directly on the agar plate using a 1 ml Gilson pipet for suspending of the cells. This cell suspension was usually enough to start four to six flasks containing 500 ml of medium. For the first expression trials the medium was inoculated at room temperature and only heated for bacterial growth. Later only pre-warmed medium was used for inoculation. As culture medium mostly TB medium was used, if that was not available LB medium was used. The culture was usually grown to an OD_{600} of 0.6 to 0.8 before the culture was induced with IPTG. The final IPTG concentration in the medium was 1 mM. Two different antibiotics were used to culture the cells. At the beginning 100 µg/ml ampicillin had been used which was changed to 50 µg/ml carbenicillin during the practical work of the thesis. The expression of non-codon optimized construct of beta1AR-FB1 was done using 100 µg/ml ampicillin and 30 µg/ml chloramphenicol, because the Rosetta2(DE3) cells have an additional plasmid inside which allows them to do the expression of rare codons. The cells were harvested by centrifugation after 4 h of expression. The harvesting was performed on two different ways:

a) The cells suspension was transferred to a bottle which had a plastic bag inside and afterwards the cells were spun down with 6000 g at 4°C for 30 min (SORVAL® RC3B Plus). The supernatant was discarded afterwards by decanting. The plastic bag containing the cell pellet was frozen in liquid nitrogen and stored at -80°C.
b) The cell suspension was transferred to a 1 l plastic bottle without plastic bag was centrifuged at 4°C for 30 min with 6000 g. The supernatant was discarded and the cell pellet was resuspended in 25 ml Lysis-Buffer. The whole procedure was done on ice. The suspensions were transferred to 50 ml falcon tubes and were frozen in liquid nitrogen. The cells were stored at -80°C.

Expression of ^{15}N-labelled protein

Two different M9 media were used for production of labelled protein. The recipes are listed under *2.1.5 Media*. Different protocols were used for the two media:

a. Expression in low-complexity M9 medium:

The steps for low-complexity M9 medium were the same as for the expression of unlabelled protein, only M9 medium was used instead of TB medium. The medium was supplemented with 100 µg/ml ampicillin. The medium was preheated and inoculated with a starter culture grown at25°C overnight at 180 rpm. The culture was induced with 1 mM IPTG final when it reached an OD of 0.6 to 0.8. The harvesting was done as in *Variation b).*

b. Expression in high-complexity M9 medium:

Prior to induction the cells were grown in high-complexity medium, transferred to M9 minimal medium and inducted with IPTG. Sivashanmugam *et al.* used this two-step protocol to obtain higher expression yield of labelled proteins (Sivashanmugam *et al*., 2009).

The growing of the culture in TB medium was done as described above for the production of unlabelled protein. The pre-culture was inoculated with a starter culture grown at 25°C overnight. The culture was started in 2 l baffled flask with 500 ml TB medium which had been preheated in presence of antibiotic to 37°C. First, 100 µg/ml ampicillin was used as antibiotic, this was changed to 50 µg/ml carbenicillin during the optimisation trials. The cultures were grown in 2 l baffled flasks to a density of 2 to 2.5 (OD_{600}). The culture was transferred to 1 l plastic bottles and centrifuged at 6000 g for 22 min at 20°C). The cell pellets were resuspended in 25 ml pre-warmed M9-Buffer (50 mM NaH_2PO_4/25 mM Na_2HPO_4, 10 mM NaCl, pH 7.0-7.1) and the suspension was transferred to 2 l baffled flask with preheated 475 ml of high-complexity M9 medium. The M9 medium contained either 100 µg/ml ampicillin or 50 µg/ml carbenicillin. Before inducing the culture with 1 mM IPTG the cell were grown in M9 medium for 1 h. The harvesting was done as usual by *Variation b)* after 4 h of expression.

2.1.7Purification

In order to facilitate purification most proteins are expressed as fusions containing an affinity tag at their N- or C-terminal end. Several different affinity tags and their corresponding resins are available for purification. His-tag purification is the most common strategy nowadays and it was also used for this project. Several His-tag resins with different properties are available for protein purification. Two different resins (Table 2) have been used during the practical work: Ni-NTA Sepharose 6 Fast Flow (GE Healthcare Life Science) and silica based Ni-TED resin (Macherey-Nagel).

Table 2: Comparison of the properties of two different His-tag purification resins, which have been used during purification optimisation trials. The listed resins have different chelating groups: nitrilotriacetic acid (NTA) and tris(carboxymethyl)ethylene diamine) (TED) (data of the resins are provided by GE Healthcare Life Science and Macherey-Nagel).

	Ni-NTA Sepharose 6 Fast Flow	**silica based Ni-TED**
technology	Affinity chromatography (IMAC, immobilized metal-ion affinity chromatography)	
material/backbone	cross-linked 6% agarose beads immobilized nickel(II)-ions	macroporous silica with immobilized nickel(II)-ions
useable for	FPLC and gravity flow columns	
procedure	interaction of His-tag of recombinant expressed protein with nickel(II)-ion complex	
free coordination centres	2	1
purified protein/g swallowed resin	40-50 mg/ml GFP	5 mg/ml GFP
unspecific binding	is possible	low

The cell pellets stored at -80°C were thawed in the falcon tubes and transferred to a beaker. The thawed pellets of 3 l high-density expression (grown to OD_{600} of approximately 2) or 6 l low-density expression (grown to OD_{600} of approximately 0.7) was filled up to approximately 200 ml with l Lysis-Buffer. One spatula tip of bovine DNAse (Sigma-Aldrich) was added and this mixture was stirred until a homogenous suspension was reached. The cracking of the cells was performed by a French Press system (1000-1500 bar, 5 to 6 cycles, N_2 pressure of 4 to 5 bar, 4°C) of the company AVESTIN (model: EmulsiFlex-C3). The cracked cell suspension was split to five centrifuge tubes (volume approximately 32 ml). The suspension was centrifuged in an A8.24 rotor at 29.000 g at 4°C for 30 min (HiCEN XL from Herolab). The supernatant was discarded afterwards and the pellet was resuspended in IB_A buffer. Afterwards, the suspension was spun down again with 29.000 g at 4°C for 30 min. After this centrifugation step the supernatant was discarded. Approximately 20 ml of IB_S buffer ($IB_{S\text{-}Guanidinium}$ or $IB_{S\text{-}10\%SLS}$) was added to each centrifuge tube. The pellets were resuspended again. Afterwards the resuspending step they were put on a roller mixer for the solubilisation step at 4°C overnight. The solubilised material was transferred to ultracentrifuge tubes and spun down with 117.000 g at 4°C for 1 h (Ti-70). Afterwards the supernatant was carefully decanted into a 100 ml Duran bottle. The mostly clear pellets were discarded. For purification, either Ni-NTA Sepharose 6 Fast Flow resin or silica based Ni-TED resin was used. Depending on the purification resin different buffers were used:

Buffers for recombinant protein purification by Ni-NTA Sepharose 6 Fast Flow resin:

Lysis-Buffer	50 mM Tris*HCl pH7.8, 300 mM NaCl, 1 mM EDTA, 5 mM $MgCl_2$
IB_A	50 mM Tris*HCl pH 7.8, 300 mM NaCl, 1 mM EDTA, 1 M Urea and 0.05% βME
$IB_{S\text{-}Guanidinium}$	50 mM Tris*HCl pH 7.8, 500 mM NaCl, 6 M Guanidinium*HCl, 2%(v/v) SLS, 10% Glycerol, 0.05% βME + 1 PIC tablet per 50 mL
$IB_{S\text{-}10\%SLS}$	50 mM Tris*HCl pH 7.8, 500 mM NaCl, 10%(w/w) SLS, 10% Glycerol, 0.05% βME + 1 PIC tablet (Complete™ Protease Inhibitor Cocktail Tablet, EDTA-free) per 50 mL

Buffers for recombinant protein purification by silica based Ni-NTA resin:

Lysis-Buffer	50 mM NaH_2PO_4 pH 7.3, 300 mM NaCl, 1 mM EDTA, 5 mM $MgCl_2$
IB_A	50 mM NaH_2PO_4 pH 7.3, 300 mM NaCl, 1 mM EDTA, 1 M Urea and 0.05% βME
$IB_{S\text{-}10\%SLS}$	50 mM NaH_2PO_4 pH 7.3, 500 mM NaCl, 10%(w/w) SLS, 10% Glycerol, 0.05% βME + 1 PIC tablet (Complete™ Protease Inhibitor Cocktail Tablet, EDTA-free) per 50 mL

Purification with Ni-NTA columns:

The supernatant was filtered through a 0.22 µm syringe filter before it was loaded on the column. Up to four syringe filters were used to filter 100 ml of solubilisation suspension. The filtered supernatant was diluted 1:1 with DBIS buffer to decrease the viscosity (in case of using $IB_{S\text{-}Guanidinium}$ solubilisation buffer also to decrease additionally the concentration of Guanidinium) before doing an automatic purification (Table 3) with the ÄKTAxpress FPLC system (GE Healthcare).

Table 3: Parameters for ÄKTAxpress FPLC system

Steps	Volumes / Buffer	Flow [ml/min]
Equilibration	15 col. vol. of Equ buffer	3.0
Loading	0.22 µm filtered and diluted supernatant	2.0
Washing	15 col. vol. of $IMAC_{Wash}$ buffer	2.0
Elution	4 col. vol. of ElutionIMAC buffer	0.1

Buffers:

DBIS	40 mM Tris*HCl pH 7.8, 1 M NaCl, 2% SLS, 10% Glycerol, 0.05% βME
Equ	20 mM Tris*HCl pH 7.8, 500 mM NaCl, 1% SLS, 5% Glycerol, 0.05% βME
$IMAC_{Wash}$	20 mM Tris*HCl pH 7.8, 300 mM NaCl, 0.1% SLS, 40 mM Imidazole, 0.05% βME
$IMAC_{Elution}$	20 mM Tris*HCl pH 7.8, 150 mM NaCl, 0.1% SLS, 1 M Imidazole, 0.05% βME

Purification with silica based Ni-TED resin:

For the purification with silica based Ni-TED resin a normal gravity flow column (Glass Econo-Column (BioRad) for chromatography) was used instead of a FPLC system and packed Ni-NTA columns, because only bulk resin was available. The supernatant was not filtered or diluted in this purification protocol, because the hydrostatic pressure was high enough to produce an adequate flow through the resin. There was not the problem of blocking any filter. The whole purification was done on the bench at RT.

The silica based Ni-TED resin (1.5 to 3 g dry resin) was pre-soaked in $BINDPO_4$ buffer and used for packing of the column. After sedimentation of the resin the supernatant was carefully added to the column After protein binding the column was washed with six to ten column volumes of $BINDPO_4$ buffer. The elution was done with $ElutionPO_4$ buffer. Fraction sizes of 2 ml at the beginning to 4 ml at the end were collected.

Buffers:

$BINDPO_4$	50 mM NaH_2PO_4/Na_2HPO_4 pH 7.3, 500 mM NaCl, 0.1%(w/w) SLS and 0.05% βME
$ElutionPO_4$	50 mM NaH_2PO_4/Na_2HPO_4 pH 7.3, 500 mM NaCl, 0.1%(w/w) SLS, 500 mM Imidazole and 0.05% βME

2.1.8Thrombin cleavage

Thrombin recognises and cleaves LeuValProArg↓GlySer. Additionally it may recognise a few other sites where it cleaved unspecifically.

The thrombin cleavage was done in batch mode. First, the protein was dialysed in a 10 kDa cut-off Slide-A-Lyser Dialysis cassette (Thermo Scientific) using a sample to buffer ratio of approximately 1 to 200. After overnight dialysis the amount of sample was evaluated by measuring the UV absorption (NanoDrop 1000, Thermo Scientific). The cleavage was done directly in the dialysis cassette for at least

4 h. During the project thrombin from two different sources was used: human thrombin with an activity of ≥1,000 NIH units/mg (Sigma-Aldrich was replaced by bovine thrombin with an activity of ≥40-300 NIH units/mg (Sigma-Aldrich). The cleavage conditions used for the different constructs are listed below:

Cleaving conditions for α5I-beta1AR-FB1-H8 and α5I-beta1AR-c.o.-FB1-H8

Two different cleavage buffers were used depending on the purification protocol:

$IMAC_{Wash}$ buffer for Ni-NTA agarose resin or $BINDPO_4$ buffer (with 1.0% (v/v) instead of 0.1% (v/v) SLS) for silica based Ni-TED

Amount of bovine thrombin: 25 U per mg of construct

Cleaving conditions for α5I-V2R-H8:

Two different cleavage buffers were used depending on the purification protocol:

$IMAC_{Wash}$ buffer for Ni-NTA agarose resin or $BINDPO_4$ buffer (with 1.0% (v/v) instead of 0.1% (v/v) SLS) for silica based Ni-TED

Amount of bovine thrombin: 1 U per mg of construct

2.1.9Purification of cleaved GPCRs

The separation of the GPCR and *α5I*-domainwas done as described in *2.1.8 Purification*. After the separation, the elution fractions which contained the GPCR were pooled and concentrated using spin concentrators (Vivaspin® 20, 10 kDa cut-off). The concentrated protein was dialysed against $IMAC_{Wash}$ buffer or $BINDPO_4$ buffer at a 1:200 protein to buffer ratio. After dialysis the sample was concentrated (Vivaspin® 20, 10 kDa cut-off) to approximately 9 mg/ml for the refolding experiments.

2.1.10 SDS-Page

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-Page) is a common method for separation of proteins by molecular weight on a discontinuous polyacrylamide gel by electricity as driving force. The proteins are mixed with an SDS loading buffer which results in a covering of all proteins with dodecyl sulphate. Due to this covering all proteins are negatively charged. and move independent of their previous charge in the direction of the anode. The migration rate of larger proteins is slower than that for smaller proteins and leads to separation by molecular weight. For this technique its common to use a stacking gel today which leads to a focusing of the proteins before the separation starts. This result in sharper lines on the gel. Different staining methods are available for protein visualisation: tracking dye, loading aids, coomassie brilliant blue or silver staining.

Protein samples of different purifications, separations, cleavage experiments and refolding experiments were analysed using NuPage® Novex® 4-12% Bis-Tris Gels (1.0 mm, 15 well) form Life Technologies™,

because this gel showed the best separation and resolution for membrane proteins. The first position of the gel was generally loaded with a marker, either Precision Plus Protein™ Unstained Standards (250–10 kDa, for normal InstantBlue® staining) or Precision Plus Protein™ All Blue Standard (250-10 kDa for Western Blot), both were from BIO-RAD. Before the samples were loaded into the gel pockets they were mixed with a self-made 5x loading SDS-loading buffer (recipe: 2.25 ml of 1 m Tris*HCl pH 6.8, 5 ml glycerol, 0.5 g SDS, 5 mg Bromophenol Blue, 2.5 ml of 1 M DDT) in a ratio of five to one. 10 µl of this mixture was loaded into one gel pocket. The gel was run in MOPS buffer with an applied electric current of 200 V for 45 min. Afterwards the gel was transferred to a clean plastic box and 10 to 15 ml of InstantBlue®, a fast version of coomassie staining reagent, was added. The gel was incubated on an orbital shaker for 15 to 45 min and washed afterwards with water. An extensive destaining was not necessary for that type of staining. One hour of shaking in 200 ml fresh water lead to a complete destaining. The gel was scanned using a common desktop scanner, which is available in every office.

2.1.11 Western Blot

Western blot is a technique which allows to detect specific sequences/structures of proteins transferred to a membrane from a SDS-Page. It is used to identify the protein/proteins of interest. Usually two different antibodies are used for the detection in order to minimize false positive signals. The first antibody binds to a specific peptide sequence (or amino acid structure) such as the His-tag. The second antibody (anti-antibody) is an antibody which binds specifically to the first antibody and it is also carrying an enzyme or radioactive label. The enzymatic reaction or radioactive labelling allows a sensitively detection of proteins.

The SDS-Page (NuPage® Novex® 4-12% Bis-Tris Gels (1.0 mm, 15 well) form Life Technologies™) was run with the same condition as described under point *2.1.10 SDS-Page*, but no staining was done. Hoefer Semi Phorwas used for transfer of the proteins from a SDS-Page to a PVDF membrane. The layers were built up as follows: A thick filter paper soaked in TBS buffer, was put directly on the anode of the blotting machine. PVDF membrane (Millipore) was placed on top. The next layer was the unstained SDS gel and on that a second thick filter paper soaked in TBS buffer was placed on it. Approximately 10 ml of TBS buffer was added on top of all these layers to ensure that it will not dry out during overnight blotting. A roller was used to press everything together and to get rid of air bubbles which might be between the layers. Finally the cathode part of the blotting machine was used to close the gadget. Some weight was added on top of the machine. The blotting was done by applying current of 10 mA overnight. After blotting the PVDF membrane was stored in TBS buffer.

Several washing/binding steps (Table 4) were done before quantification with an imager was possible:

Buffers:

TBS	10 mM Tris*HCl pH 7.5, 150 mM NaCl, filtered through 0.22 µm
TBST	10 mM Tris*HCl pH 7.5, 500 mM NaCl, 0.05% (v/v) Tween 20, 0.2% (v/v) Triton X-100, filtered through 0.22 µm
Blocking-Buffer	3% (w/w) BSA dissolved in TBS Buffer

Primary Antibody	Blocking-Buffer + 1:2000 Anti-Penta-His (Mouse, Qiagen #34660)
Secondary Antibody	10% (w/w) milk powder dissolved in TBS buffer + 1:4000 Anti-Mouse-AP (Southern Biotech #1031-04)

Table 4: Protocol of washing/binding steps for Western Blot

Step	solution	volume/ml	time/min
1	TBS buffer	20	5
2	Blocking buffer	10	45
3	TBST buffer	20	5
4	TBST buffer	20	5
5	TBS buffer	20	5
6	Primary Antibody	10	45
7	TBST buffer	20	5
8	TBST buffer	20	5
9	TBS buffer	20	5
10	Secondary Antibody	10	45
11	TBST buffer	20	5
12	TBST buffer	20	5
13	TBST buffer	20	5
14	TBST buffer	20	5
		Total time	290

After washing/binding steps the membrane was carefully drained from excess buffer and covered with 3 ml of Lumi-Phos^{TM} WB Chemiluminescent Substrate at RT for 3 min. The excess substrate was carefully drained before detection with a CCD Imager (GE Healthcare ECL Plus Western Blotting Detection System) was done (Table 5). It was necessary to cool down the camera to -25°C, which lead to a more sensitive measurement.

Table 5: Settings for CCD Imager of GE Healthcare ECL Plus Western Blotting Detection System for cemiluminescence and visual detection

Settings	Chemiluminescence	Visual Detection
Chemi display	On	off
White light	Off	on
PMT	Medium/medium	Medium/medium
Time	12 s	120 ms

2.2 Refolding & characterisation

2.2.1Refolding on magnetic beads

In our lab we have a machine called Kingfisher mL Magnetic Practical Processors (Thermo Fisher Scientific) which is a small scale purification system and would allow to test 15 small refolding conditions in parallel.

Beads:

BcMag® His Magnetic beads from Bioclon	→ Ni beads I
EcoMag™ His-Ni-Magnetic particles from Bioclon	→ Ni beads II
EcoMag™ His-Co-Magnetic particles from Bioclon	→ Co beads

Buffers:

$MB_{Loading}$	20 mM Tris*HCl pH 7.8, 300 mM NaCl and 0.1% SLS
MB_{Wash}	20 mM HEPES pH 7.5, 200 mM NaCl and 0.125% DDM
$MB_{Elution}$	20 mM HEPES pH 7.5, 200 mM NaCl, 0.125% DDM and 500 mM Imidazole

Coating solutions:

BSA coating solution	1 g in 10 ml $MB_{Loading}$ buffer
Peptone coating solution	2 g in 10 ml $MB_{Loading}$ buffer
Polyethylenimine coating solution	1% (w/w) in $MB_{Loading}$ buffer

2.2.2Refolding on Ni-NTA resin

Resin:

Ni-NTA Sepharose 6 Fast Flow resin from GE Healthcare

Buffer:

Running buffer	20 mM NaH_2PO_4 pH 7.3, 150 mM NaCl, 0,05% DDM

2.2.3Refolding by dilution

In small scale refolding tests 1 µl or less of solubilized GPCR stock solution (approximately 9 mg/ml) was pipetted to the inner wall of an 1.5 ml eppendorf tube and a pre-mixed refolding buffer was added quickly. The refolding mixture was resuspended and stored at 20°C for 1 h 20 min. Oxidized glutathione was added to change the redox system from a more reducing environment (1 to 0.1 red./ox.) to an equal ratio of red./ox. glutathione (except in condition C where it was the other way around → it was started from a more oxidised environment). 1 h later the eppendorf tube was transferred to 4°C and ox. glutathione (red. glutathione in condition C) was added a second time. Afterwards an additional incubation time of at least 2 h was done to allow the GPCR to go through its refolding process.

The refolding experiments were done in refolding buffers which contained 20 mM HEPES pH 7.3 and 150 mM NaCl. The additional additives are listed in Table 6. The concentration of the V2R-H8 stock was 9 mg/ml to provide a high enough dilution ratio which should guarantee dilution of SLS (N-lauroyl-sarcosine) to a minimal concentration at which it is not interfering with the refolding process anymore.

Table 6: Refolding condition which were tested by the dilution refolding strategy

Buffer composition	Condition A	Condition B	Condition C	Condition D	Condition E
Red./ox. Glutathione (stating ratio)	2 mM/0.2 mM	2 mM/0.2 mM	2 mM/0.2 mM	0.2 mM/2 mM	2 mM/0.2 mM
DDM	4 mM	2 mM	----	4 mM	4 mM
Cymal-6	----	2 mM	4 mM	----	----
Tolvaptan	----	----	----	----	10 µM

2.2.4 CPM-Assay

One of the newer biophysical characterization assays for membrane proteins is the microscale fluorescent stability screen based on the reaction of a thiol-specific fluorochrome N-[4-(7-diethylamino-4methyl-3-coumarinyl)-phenyl]-maleimide dye (CPM, extinction/emission maxima of ~384/470 nm) (Figure 6) with free cysteines of the protein. This dye allows monitoring of the unfolding process of cysteine-rich proteins while the temperature is increased continuously. During the unfolding of the protein more and more non-accessible cysteines from inside of the protein become accessible to the dye. The unbound dye does not give any measurable fluorescence signal. After the reaction a dramatic increase of the signal can be detected at a wavelength of 470 nm (Alexandrov, Mileni, Chien, Hanson, & Stevens, 2008).

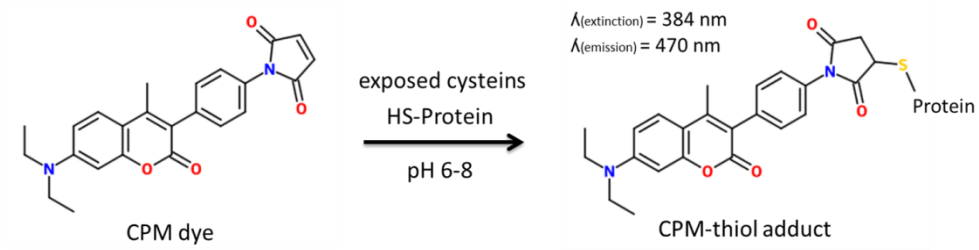

Figure 6: Reaction equation of CPM dye (N-[4-(7-diethylamino-4methyl-3-coumarinyl)-phenyl]-maleimide) with a free cysteine of a protein.

Buffers for setup of the CPM-Assay:

$CPM_{unfolding}$	20 mM HEPES pH 7.5, 100 mM NaCl, 1% (w/w) SLS
$CPM_{unfolding\_elution}$	20 mM HEPES pH 7.5, 100 mM NaCl, 1% (w/w) SLS, 10 mM EDTA
CPM_{DDM}	20 mM HEPES pH 7.5, 200 mM NaCl, 0.125% DDM
$CPM_{Elution}$	20 mM HEPES pH 7.5, 200 mM NaCl, 0.125% DDM, 500 mM Imidazole

2.2.5Radio-ligand binding

Radio-ligand binding assays are very common for characterisation of GPCRs. Thex give information on the functional folding of the receptors, because only a native folded receptors should be able to bind its ligands specifically. The radio-ligand binding is a perfect tool to test on the one side if native folding was reconstituted during the refolding process and on the other side it allows a very precise quantification of the refolding efficiency. A filter assay based on intact cells and membrane fractions of V2R was already available. This assay was not useable for measuring solubilised material, because solubilised material will not be hold back by a pore size of 0.22 µm. The ligands which were used for this experiment, the hot ligand, [H3] labelled vasopressin peptide (Vasopressin (8-L-Arginine), [Phenylalanyl-3,4,5-3H(N)]),cold vasopressin peptide and the cold ligand Tolvaptan.
Due to that I tried to establish a radio-ligand binding assay for solubilised material by size-exclusion chromatography and a silica based Ni-TED separation.

Size-exclusion based radio-ligand binding assay

A 96-well filter plate was prepared with 1 cm height size-exclusion column (Sephadex G-25 Fine from GE Healthcare) inside a few wells. 20 µl of a testing mixture (1 mg/ml BSA with 1 µM fluorescein) was loaded on each small size exclusion column. The elution of BSA and fluorescein was done by stepwise adding of running buffer (20 mM HEPES pH 7.3, 150 mM NaCl, 0.05% DDM, filtered through 0.22 µm filter). Each fraction was collected in a new well. The concentration of BSA was measured by UV absorption at 280 nm (NanoDrop 1000 from Thermo Scientific). The fluorescein concentration was determined by fluorescence absorption at 515 nm (NanoDrop 3300 Fluorospectrometer).
The elution of hot vasopressin ligand was done too, exactly how it was done for the mixture of BSA and fluorescein. The readout of eluted peptide was measured by the TopCount NXT™ from Packard.

Silica based Ni-TED radio-ligand binding assay:

The protein with the bound ligand should be separated from the unbound ligand by a simple column wash step after binding. The whole procedure should take place in a 96-well format(filter plates Millipore Multiscreen®-HV, 0.45 µm hydrophilic low protein binding); beginning with incubation of the receptor with hot ligand, over a wash step, to the final quantification by plate reader. The used assay does not allow an absolute quantification, because the bound ligand may dissociate from receptor during the wash step. This process of dissociation can be minimized by a fast wash step and a ligand which has a slow off rate from the receptor. The used hot ligand vasopressin has a very high affinity of approximately 1 nM.

This assay was performed in two different ways:

a) Variation A:

   This experiment was working with an excess of protein to hot ligand during the batch binding. It was done to see if there is any refolding at all.

The protein refolding solution (1 µM) direct from the refolding procedure was incubated with an approximately 11 times lower concentration of hot vasopressin (88.2 nM). For this 50 µl of refolding solution was mixed with 50 µl of hot ligand solution on ice. After 1 h of incubation the 100 µl of protein/hot ligand solution was pipetted onto the pre-soaked resin (20 ± 0.2 mg of dry resin) in a well of the 96-well filter plate. Afterwards, the plate was spun by 300 g at 4°C for 1 min. The flow-through was loaded twice again after spinning it through the resin. Afterwards, 200 µl of RLB_{Wash} buffer was added to each well and spun down by 250 g at 4°C for 3 min. This step was repeated once. Then, 100 µl of $RLB_{Elution}$ buffer was added to the resin with the bound protein. The resin was resuspended twice at 4°C during a 10 min incubation time. The elution of the receptor with ligand was done by a 1 min centrifugation step into a white 96-well plate (Pico Plate™-96, white microplate) using 1000 g at 4°C. 100 µl of scintillation liquid (TRI-CARB® Liquid Scintillation from Packard) was added to the elution buffer. After 1 h of equilibration time the plate was read by TopCount NXT™ from Packard. Counting was done for 1 min per well. The samples were normalized by a positive control which contained hot ligand in a known amount and scintillation liquid.

b) Variation B:

This experiment was performed with the same protein-ligand ratio but the hot ligand was supplemented by cold ligand to reach almost equal concentration of protein and ligand. The reason for this is that the experiment wanted to be done also direct from the refolding buffer without any additional manipulation steps. Hot ligand in the amount of protein is not affordable, due to this it was supplemented by cold vasopressin peptide.

The ligand mixture contained 88.2 nM of hot ligand and 1 µM of cold ligand. The binding affinity of both ligands to the V2R should be almost the same, because the tritium label of the hot vasopressin peptide will not have a big influence on the binding process. 50 µl of refolding solution were mixed with 50 µl of hot and cold ligand stock. After 1 h of incubation 50 µl of this mixture were pipetted into a well which contained soaked silica based Ni-TED resin. The loaded sample was spun through the resin by 250 g at 4°C for 30 s. The flow-through was loaded twice again, after spinning it through the resin. The resin was washed three times with 200 µl of RLB_{Wash} buffer which was spun through by 250 at 4°C for 4 min directly after pipetting. The resin was incubated with 100 µl of $RLB_{Elution}$ buffer for 15 min at 4°C. The eluted protein was spun down by 1000 g at 4°C for 1 min. The flow-through was mixed with 200 µl of scintillation liquid. After 1 h of equilibration time the read out was done with TopCount NXT™ from Packard.

This procedure was done in parallel for refolding condition A, B, E unfolded receptor (stored in 0,1% SLS buffer) and for an additional negative control which contained resin only.

Buffers:

RLB_{Wash}	20 mM Na_2HPO_4 pH 7.3, 150 mM NaCl and 0.05% DDM (filtered through 0.22 µm)
$RLB_{Elution}$	20 mM Na_2HPO_4 pH 7.3, 150 mM NaCl, 0.05% DDM and 500 mM Imidazole (filtered through 0.22 µm)

2.2.6Analytical ultracentrifugation

Analytical ultracentrifugation (AUC) is a method which allows to study molecular weight of protein in solution under physiological conditions. Two different techniques are used today which are based on different principles: sedimentation velocity run and sedimentation equilibrium run.
The sedimentation velocity run is a hydrodynamic method. The measuring takes place at high speeds where only sedimentation contributes to the result. The sedimentation is depending on shape and buoyant mass of the macromolecule. This methods offers to estimate the molecular mass and sedimentation coefficient precisely if the right parameters are chosen. The other method is sedimentation equilibrium run which is a thermodynamic method. The experiment is done at much lower g-forces compare to the sedimentation velocity run. This allows the macromolecule to reach an equilibrium between sedimentation and diffusion. The result is only dependent on buoyant mass and not on the shape because there is no net transport of molecules. This technique delivers information about the oligomerisation states of a macromolecules (Howlett, Minton, & Rivas, 2006).

Both experiments were performed with the Proteome Lab XL-I ultracentrifuge from Beckman Coulter. The parameters for both measurements are listed in Table 7.The analysis of the sedimentation velocity experiment was done with the SETFIT software (developed by Peter Schuck Phd) to determine the sedimentation coefficient. The data analysis for the sedimentation equilibrium run was done by Ultraspin software (Dr. Dmitry Veprintsev).

Buffer:

20 mM NaH_2PO_4 pH 7.3, 150 mM NaCl, 0.05% DDM

Table 7: The used parameters of the equilibration velocity and sedimentation velocity run.

	Equilibration velocity run	Sedimentation velocity run
Measuring cells	Six-sector cell	Two-sector cell
Speed	18.000 and 24.000 rpm rotor An-60 Ti	42 000 rpm rotor An-60 Ti
Time	measured in 6 h intervals	continuously
Temperature	22°C	22°C

2.2.7Liquid chromatography-mass spectrometry

The measurement of the accurate mass of the intact GPCRs was performed by two different protocols:

a) Variation A:

10 to 50 µl GPCR sample with a concentration of approximately 0.5 mg/ml was injected i into the LC-MS system. The parameters for the liquid chromatography-mass spectrometry (LC-MS) run are listed in Table 8. The measurement was done with refolded sample in the detergent DDM and unfolded sample in the detergent SLS.

b) Variation B:

In Variation B a precipitation protocol was used to get rid of the most detergent which should lead to a better ionisation of the protein during the ionisation step.

The whole precipitation procedure was done with glass ware which is very important to prevent contamination of any polymer from plastic ware. 10 times of Solution A was added to the sample which was stored in a buffer which contained 0.1% (w/w) SLS. The solution was mixed and spun down by 4000 g at 4°C for 30 min and afterwards stored at -20°C for nearly one day. The sample was spun down again and he liquid was carefully removed with a Pasteur pipet. The same volume which was added of solution A, was added of solution B. The vial was vortexed and stored at -20°C for 1 h. The last step of replacing the solution, vortexing and storing was repeated twice before the acetone was evaporated at room temperature. The close to dry pellet was dissolved in $MS_{Solubilisation}$ buffer before it was measured by LC-MS. The parameters for the LC-MS run are listed in Table 8.

The precipitation protocol for analysing integral membrane protein was adapted from protocol developed by Sasa Bjelic (personal communication). DM, DDM and Fos-Choline-12 are the recommended detergents for this protocol. I used the detergent SLS and solubilisation buffer. At this point thanks to Alain Blanc for helping me with the MS measurements.

Table 8: Parameters of the LC-MS measurements of Variation A and B.

	Varaiation A	**Variation B**
Type	TOF MS (LCT Premier – KD048)	
Ionisation mode	positive mode/electro spray	
Column	Zorbax C3 (3 µm 2x150 mm,)	XTerra RP 8 Guard Column (125Å, 5 µm, 3 x 20 mm)
Column temperature	50°C	60°C
Solvents	A: ACN + 0.1% (v/v) HCOOH B: H2O + 0.1% (v/v) HCOOH C: isopropanol	A: ACN + 0.1% (v/v) HCOOH B: H_2O + 0.1% (v/v) HCOOH C: isopropanol
Gradient	0 – 5 min……. 5% A/90% B/5% C 5 – 10 min……....5% A/90% B/5% C 10 – 15 min…...….15% A/80% B/5% C 15 – 20 min……....25% A/70% B/5% C 20 – 25 min……....35% A/60% B/5% C 25– 30 min……...90% A/5% B/5% C	0 – 5 min……. 5% A/90% B/5% C 5 – 10 min……....5% A/90% B/5% C 10 – 15 min…...….15% A/80% B/5% C 15 – 20 min……....25% A/70% B/5% C 20 – 25 min……....35% A/60% B/5% C 25– 30 min……...90% A/5% B/5% C
Injection volume	10 – 50 µl	10 µl

Solution:

Solution A	10% (w/w) Trichloroacetic acid in Acetone
Solution B	Acetone
$MS_{Solubilisation}$	100 mM Tris*HCl pH 6-7, 7 M Guandidinium-hydrochlorid, 1 mM EDTA

2.3 NMR Experiments

Sample preparation:

The refolding was done as described under point *2.2.3 Refolding by dilution*. 50 ml falcon tubes were used instead of 1.5 ml eppendorf tubes. The next day the protein was concentrated using either a 50 kDa or 100 kDa Vivaspin® 20 centrifugal spin concentrator. Centrifugation steps were done using the table top centrifuge (Eppendorf centrifuge 5810R, A-4-62 rotor) 4.000 g at 4°C for 5 min in an Eppendorf centrifuge 5810R equipped with A-4-62 rotor. The sample was resuspended every 5 min. This procedure was followed until a volume of approximately 500 µl was reached. This protein solution containing folded and unfolded material was split into two aliquots which were both run over a size-exclusion column.

The size-exclusion was performed on a Superdex 200 16/60 column with a flow of 0.5 ml/min in the cold room. As running buffer 20 mM NaH_2PO_4 pH 7.3, 150 mM NaCl and 0.05% DDM was used on an ÄKTAprime system. The elution fractions of the monodisperse peak were pooled and concentrated again with 50 kDa or 100 kDa Vivaspin® 20 concentrators to approximately 500 µl. The sample was stored at 4°C until the NMR spectra were recorded.

Measurement

Shortly before the measurement 5% D_2O was added to the sample. The measurement was performed in a 5 mm NMR tube at 20°C. All NMR experiments were performed on Bruker Avance 600 and 700 spectrometers equipped with triply tuneable cryogenic probeheads and Z-axis pulsed field gradients. A Watergate was applied.

3 Results

3.1 Expression and Purification in TB and M9 Media

Expression of beta1AR-FB1:

The expression of construct *α5I*-beta1AR-FB1 was done during the start of the thesis in Rosetta2(DE3) cells, because the expression level were much lower or not given in the other cell lines. The screening of expression in different cell lines was done by Antoine Gautier whose project I continued.

The reason for low expression levels was that the sequence of the beta1AR was not codon optimized for *E. coli*. Due to this it was important to use Rosetta2(DE3) cells, but the expression resulted there in low protein yields. The highest yield was of the uncut construct of approximately 0.5 mg form 1 l of bacteria culture expressed in TB medium. Unfortunately, this result was not reproducible, despite using several optimisation strategies which were successfully used for the expression of *α5I*-V2R-H8 (see below).

The use of a codon optimised version of the *α5I*-beta1AR-FB1-H8 construct and the optimised expression-purification protocol for the fusion construct *α5I*-V2R-H8 leaded to a significant increase of the expression levels. It was possible to achieve a yield of 2.3 mg/l of the uncut construct. Several changes may have contributed to this almost five-fold increase of the yield per litre *E. coli* culture: the codon optimization, the change of the antibiotic, the expression in BL21(DE3) cells and the new purification protocol with silica based Ni-TED resin. If the results of expression levels between the two studied GPCRs are comparable, it can be concluded that the codon optimisation and the changed antibiotic had the biggest influence on the yield. The codon optimisation brought the advance that it was possible to do the expression of the *α5I*-beta1AR-FB1-H8 in BL21(DE3) cells without two antibiotics. It was already possible to reach good yields with the *α5I*-V2R-H8 construct in those cells. The antibiotic change from 100 µg/ml ampicillin to 50 µg/ml carbenicillin had an big improvement on the expression level of V2R. The expression of the codon optimised *α5I*-beta1AR-FB1-H8 construct was not tested by using the antibiotic ampicillin in combination with BL21(DE3) cells which makes it hard to compare these optimisations. How big the influence of the changed purification protocol for the final yield was not evaluated, but it made the purification less time and resources consuming (see below).

During the construct design of *α5I*-beta1AR-FB1-H8 it was assumed that the second cleavage site in the α-helical part of the GPCR sequence will not be accessible for the protease. Because of alpha helical segments which can already exist in the unfolded state of a GPCR (Banères *et al*., 2011).Thrombin cleavage experiments showed that the presumption was wrong. Figure 7 illustrate that unspecific cleavage was taking place by doing the thrombin cleavage in 0.1% SLS buffer which was used for the purification of the protein. After modification of the cleavage condition was it possible to reduce the unwanted cleavage in the alpha helical segment almost to zero. It was necessary to use higher amounts of thrombin to yield up with a close to 100% cleavage of the fusion construct under modified conditions. The reason for this could be either that the higher SLS concentration had an influence on the activity of the enzyme or the detergent is annealing now on the cleavage sites which was less the case before. The formation of the

detergent micelle around the helical segment of the unfolded protein may explain why the cleavage in the helical part was not taking place anymore. The cleavage site in the helical part is more hydrophobic because of the amino acid sequence, which may lead to a tighter binding of the single detergent molecules in that area. The poly-glycine linker next to the original cleavage site is hydrophilic and flexible, therefore potentially reducing binding of detergent molecules in that part of the construct.

As already mentioned above it was necessary to use higher concentration of thrombin during the cleavage. Due to this it was tried to replace the more expensive human thrombin with an activity of ≥1,000 NIH units/mg (Sigma-Aldrich) with a cheaper thrombin version from bovine sources. The bovine thrombin which had activity of ≥40-300 NIH units/mg (Sigma-Aldrich) was working in my hand as well as the more expensive human one.

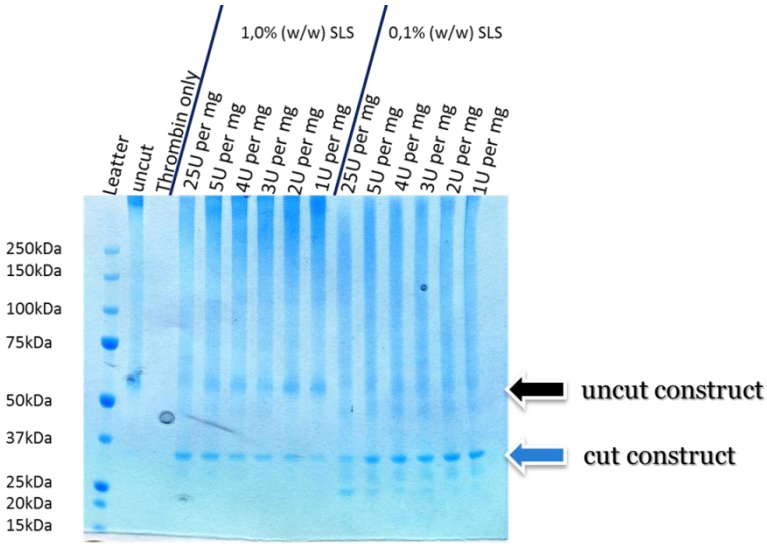

Figure 7: Thrombin cleavage of the construct beta1AR-FB1-H8: The black arrow shows the uncut construct which was running at a molecular mass of approximately 57 kDa. The blue arrow shows the cleaved GPCR and the *α5I*-domain which were both running at the same height under this condition. Unspecific cleavage was present at thrombin concentrations between 1 to 25 U per mg of uncut construct using 0.1% SLS (w/w) in the cleavage buffer. Higher concentration of 1.0% (w/w) SLS lead to a reduction of unspecific cleavage to almost "zero", but higher concentration of thrombin per mg of uncut construct were necessary to reach a close to 100% cleavage.

The cleavage with the 3C Protease was tried only once with different ratios of construct to protease in 0.1% SLS buffer. The efficiency of this protease compare to thrombin was very low, despite using a very high one to one molar ratio of protease to construct. The first reason might be the buffer redox conditions were suboptimal, especially because the sample was stored over one week before the cleavage experiment was done. This problem could be easily solved by adding additional reducing agent which is planned for the next experiment. The second reason may be that SLS inhibits the activity of 3C Protease, but no data to support or disprove the assumption were found in literature.

Expression of unlabelled V2R-H8

The expression of *α5I*-V2R-H8 was done during the practical work in BL21(DE3) cells only, but it was tried by Dalibor Milic in several other cell lines. The expression yield of those was the same or lower compare to BL21(DE3) cells.

The expression level of the fusion construct *α5I*-V2R-H8 was optimised during several optimisation trials. Two different inoculation strategies were tested where the strategy of using a starting culture which was grown overnight delivered better results compare to a directly inoculation from an agar plate. The starting culture which was grown at 25°C lead to a higher expression yield as the starting culture at 37°C. The OD_{600} of those was only slightly differing what should not have a dramatic influence on the expressed protein. Therefore, a starting culture grown at 25°C was used for all future experiments. One of the key points that was identified for delivering good expression levels was to prewarm culture media. The preheating was done usually in an incubator in presence of antibiotics overnight. The improve was that the growth rate of the cells was much better and with a more reliable duplication of the biomass over a longer time period. The final expression yield after optimisation was 2.0 mg/l cut construct after purification. The purification was done using 5 ml Ni-NTA columns (5ml HisTrap FF crude column from GE Healthcare) which is important to mention because the purification was not as sufficient/efficient as it could be. During the setup trials of the refolding screens on magnetic beads and Ni-NTA resin was recognised that unspecific binding of the protein to the agarose surface was taking place. This unspecific binding probably also occurred during the purification with the Ni-NTA column. It may also explain for the reduced binding capacity of a recharged column. These 2 mg/l purified protein were reached by repeating the purification three to four times, using the flow-through from the previous purification. The column were regenerated after each purification. The binding capacity of the Ni-NTA resin was significantly lower than anticipated. In our experience the Ni-NTA resin used has the possibility to bind 8 times more protein per ml of pre-soaked resin compare to the silica based version. The interesting thing is that the efficiency was almost the other way around. The silica based Ni-TED resin was possible to purify approximately 3 to 4 times more protein compare to the Ni-NTA columns. The problems of unspecific binding which might be the reason for the observed phenomena are documented below point *3.2.1 Refolding on magnetic beads*. For the His-tag purification was it important to change from Tris*HCl to phosphate buffers, because the tertiary nitrogen of Tris*HCl is able to block the last left free coordination site of immobilized Ni-complex on the silica based resin surface. This can lead to a significant reduction of the binding capacity of the resin.

The influence of the solubilisation buffer can be excluded because only during the start of the thesis a different one was used. The $IB_{S\_Guanidinium}$ buffer was changed to a solubilisation buffer that contained only SLS ($IB_{S\text{-}10\%SLS}$). The solubilisation buffer was changed because the high concentration of guanidinium-hydrochloride is on the one side the tendency to crystallise which can become a problem using a FPLC system. On the other side a theory exist that if the solubilisation of IBs is too harsh a refolding might not be possible anymore (Singh *et al.*, 2012). The solubilisation of inclusion bodies was as efficient for the

10% (w/w) SLS solubilisation buffer as for the one which contained guanidinium-hydrochloride and SLS. Due to this $IB_{S\text{-}10\%SLS}$ was used subsequently.

The expression of unlabelled *α5I*-V2R-H8 construct was tried with the antibiotic carbenicillin too. The yield was approximately twice as high as for the expression with ampicillin. The expression of unlabelled protein was done only once but it yield up with 4.2 mg purified V2R-H8 from 1 l of bacteria culture grown in TB medium. This antibiotic allowed to grow the cells to higher biomass (OD_{600} of 2.0 to 2.5), most likely because it allowed to keep the plasmid in the cells for longer. The influence of the silica based resin was not evaluated, but it made the purification less time and resources consuming.

The thrombin cleavage of the *α5I*-V2R-H8 construct was done in 0.1% SLS purification buffers and showed specific and efficient cleavage. It was observed that for a successful thrombin cleavage a dialysis was necessary to reduce the concentration of Imidazole before the thrombin was added. If the thrombin was added direct into the elution buffer of the IMAC purification, a very low efficiency was observed also after one day of cleavage. The efficiency did not improved by using a longer time period. The combination of SLS detergent and high concentration of Imidazole might lead to a non-reversible inhibition of thrombin. Only one protein band and its oligomer bands were visual on the gel and on the Western Blot (Figure 8). The Western Blot gave the additional information that the right constructs were expressed for both GPCRs. This can be concluded from the Western Blot result, because an anti His-tag primary antibody was used and the His-tags were located at the C-termini of the construct sequences.

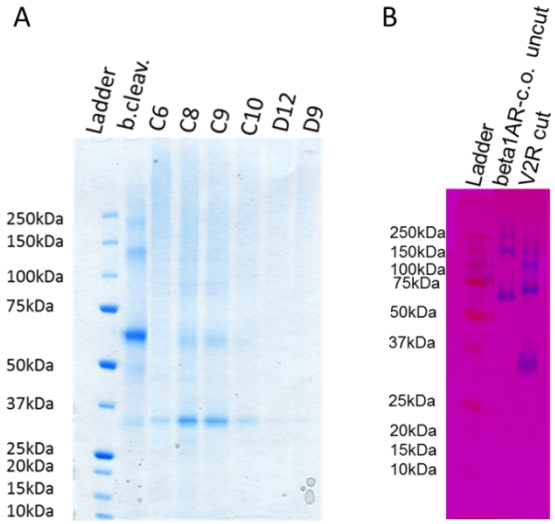

Figure 8: A shows the elution fraction of the separation of cut *α5I*-V2R-H8 construct and the uncut construct including the *α5I*-domain on position 2**. B:** Western Blot of *α5I*-beta1AR-c.o.-FB1-H8 uncut construct and V2R-H8: It was an anti His-tag Western Blot and shows that the constructs both constructs were successfully expressed in inclusion bodies. The *α5I*-beta1AR-c.o.-FB1-H8 construct was running very close to its predicted mass of 67 kDa. In case of V2R-H8 the protein was running at lower molecular weight than it should be. It was running at approximately 31 - 32 kDa instead of 43 kDa, but membrane proteins are usually running at lower molecular mass

Expression of ^{15}N- labelled V2R-H8

The yields of the expressed purified V2R-H8 were looking promising, so I tried to express uniformly labelled ^{15}N-V2R-H8 by two different expression strategies. First, I used less complex M9 medium (Variation A) where the growing of the culture and the expression was done in the same medium with 100 µg/ml ampicillin. This expressions resulted in no yield of the construct at all. Second, I tried to do an expression with a protocol which promised to have a dramatic increase on the expression of labelled proteins (Sivashanmugam *et al*., 2009). The idea is to grow the bacteria culture first in high dense medium and transferring those cells afterwards to M9 medium for the expression. There was no expression too. Changing the antibiotic from 100 µg/ml ampicillin to 50 µg/ml carbenicillin resulted in expression of the protein. The final yield after purification was approximately 0.5 mg/l of the purified GPCR. The purification was done thereby with silica based Ni-TED resin.

3.2 Refolding & Characterisation

3.2.1Refolding on magnetic beads

Magnetic beads are one of the newer techniques in molecular biology to do a fast purification of recombinant proteins. The principle behind this technique is to have a magnetic carrier which has for example an immobilized affinity tag for His-tag purification on its surface. The idea was to use the small scale purification system for setting up a refolding screen by magnetic His-tag beads in combination with this fast purification machine, because it would allow to do 15 refolding experiments in parallel. Refolding on beads was reported before (Sharapova *et al*., 2011; Wang, Wang, & Geng, 2009).

The principle of a refolding screen using magnetic beads is shown in Figure 9.

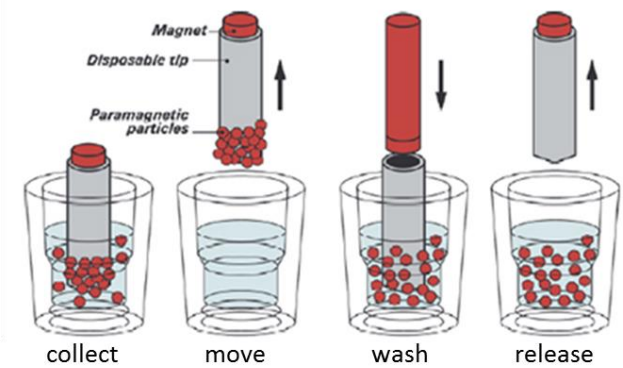

Figure 9: Principle of magnetic bead refolding: Binding of the protein in the first well of five (in case of the used machine, which had a total well number of 5) to the magnetic beads. The added disposable tip was used for mixing and transferring the bead. The collect the beads with the immobilized protein were released from the tip by releasing the magnetic insert. Afterwards, the disposable tip was used alone to mix everything by the machine. This steps of collecting and releasing were done four times until well four was reached. For the elution steps the beads were collected again and transferred to the fifth well where imidazole was used to disturb the interactions between the protein and the Ni-complex. This leaded to elution of the "refolded" GPCR in the fifth well. The beads were collected a last time before they were released in well number two (the graphic is from the homepage of Thermo Fisher).

For the setup of the refolding screen uncut construct of *α5I*-V2R-H8 was used during the beginning. 4 µg of unfolded construct in 100 µl of $MB_{Loading}$ buffer were pipetted into the loading well (1st well) and 10 µl of magnetic bead suspension was added afterwards. 1 ml of MB_{wash} buffer were provided in well two to four, which should simulated the refolding buffers that wanted to be used during a refolding screen. The elution well (5th well) contained 50 µl of $MB_{Elution}$ buffer. After preparation the machine was started (speed: medium, 1st well: 1 min mixing, 1 min collecting, 2nd-5th well: releasing of the beads, resuspending for 5 min and collecting for 1 min, the beads were collected again in the 5th well and released finally in the 2nd well). The concentration of protein in each well was quantified by SDS-Page and tryptophan-fluorescence measurements. This procedure was tried several times, but there was close to zero elution taking place. Also little protein was detectable in the loading well and washing wells. The reason was that unspecific binding of protein to the bead surface was taking place by all tested magnetic beads. This was evaluated by mixing the used beads with 5x $SDS_{Loading}$ buffer and those stripping solution on a SDS-Page (Figure 10). As a control, I tried to do this experiment with $G_{\alpha i}$. The result was that unspecific bind of the soluble protein to the bead surface was also taking place.

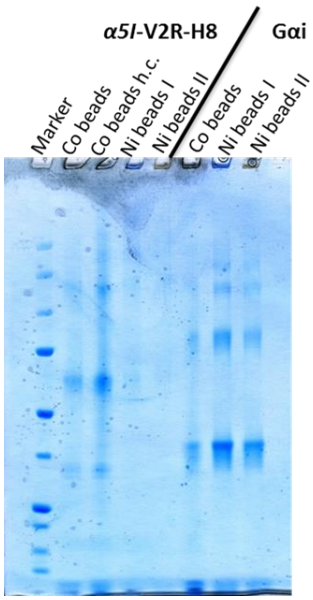

Figure 10: SDS-Page stained with InstantBlue® shows the proteins, which were stripped from the surface of the magnetic beads by 5x $SDS_{Loading}$-Buffer. It visualizes the unspecific binding of the proteins *α5I*-V2R-H8 and $G_{\alpha i}$ to the agerose surface of the different beads. In case of *α5I*-V2R-H8 using Ni beads I and II the stripping could not be shown, but it was taking place. Additionally, the tryptophan fluorescence measurements gave nearly the same result for all tested beads.

After several failed trials at various protein concentrations, it was tried to coat the beads to inhibit or at least to reduce unspecific binding. The covering of the beads was tried with BSA, peptone, polyethylenimine and uncut *α5I*-V2R-H8 construct. The beads were incubated with a solution of those covering agents for approximately 35 min during several times of resuspending them in their coating buffers. Before the beads were used for testing their efficiency, they were washed two times with $MB_{Loading}$ buffer. The same procedure as for uncoated beads was used to test efficiency of coated beads. The coating of the beads was so efficient that no uncut construct of *α5I*-V2R-H8 was binding any more to the beads, which was evaluated by running a SDS-Page of loading/washing/eluting wells (Figure 11). Maybe it is worth to mention why it was tried to do also the covering with *α5I*-V2R-H8 itself. There are two reasons: First, we saw from the previous results that this fusion protein is interacting very tight with the bead surface which would be good property for covering compound. Second, the other more important reason was that the quantification of refolding wanted to be tested by a CPM-assay. This assay would not give clear results if other proteins are in the testing solution too.

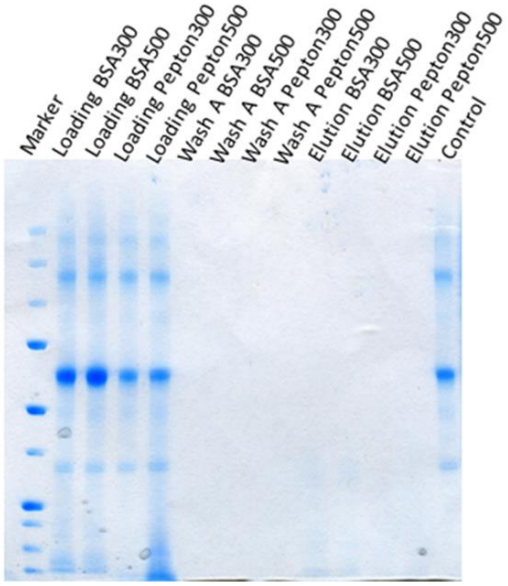

Figure 11: SDS-Page stained with InstantBlue® shows that the coating of the beads with different volumes (BSA300 = 300 µl of 1 g/10 ml, BSA500 = 500 µl of 1 g/10 ml,…) of coating solution did not lead to a better elution of the protein in the fifth well. The coating procedure was so efficient that interaction of the protein with the bead surface was removed to zero in case of BSA coating. For peptone coating could it be that it was less efficient compare to the BSA coating, because less protein is visual on the gel. But there is no protein for both coating solutions visual in the elution well. Additionally, the interaction of the His-tag with the immobilized Ni-NTA complex on the magnetic bead surface were interrupted that no binding at all was taking place anymore. For this experiment *α5I*-V2R-H8 was

It was not known how big the influence of the *α5I*-domain might be for this nonspecific binding, so the experiments was tried also with unfolded V2R-H8, but it leaded to the same result which were showed in Figure 12 with the coating trails. A coating of the beads with V2R-H8 and polyethylenimine was also insufficient and yield up with close to zero elution of protein in the elution well like in all other tested cases before. It was tried to use higher amounts of protein during the binding step, but it was still the same. Protein disappeared from the binding well and almost all of it was sticking unspecific to the bead surface. A similar problem was encountered during refolding of a recombinantly expressed C-terminal fragment of human alpha-fetoprotein on IMAC resin. At the beginning, normal Ni-NTA agarose resin was used and this resulted in close to 0% refolding. However, the use of silica based IMAC beads resulted in refolding efficiency of approximately 60% (Sharapova *et al*., 2011). Most likely, the reduction of the nonspecific binding resulted in successful refolding of their protein. While Sharapova *et al*. did not tried this strategy with a GPCR, but the properties of their protein seem to be very close to a GPCR. The protein is highly hydrophobic and has multiple disulphide bridges.

The strategy of doing refolding on beads has the advantage that the single protein molecules are not able to interact with each other which might yield up with higher yield because aggregation products are reduced. But if the His-tagged protein is interacting with the Ni-beads so tight that the protein cannot be eluted with 500 mM Imidazole, there must be additional interaction between the protein on the bead sur-

face present. A refolding may not be possible if some part of the bound proteins are interacting with the bead surface. I would argue that If nonspecific binding is present, is refolding is impossible. Steps must be taken to find resin with low non-specific binding for this approach to work.

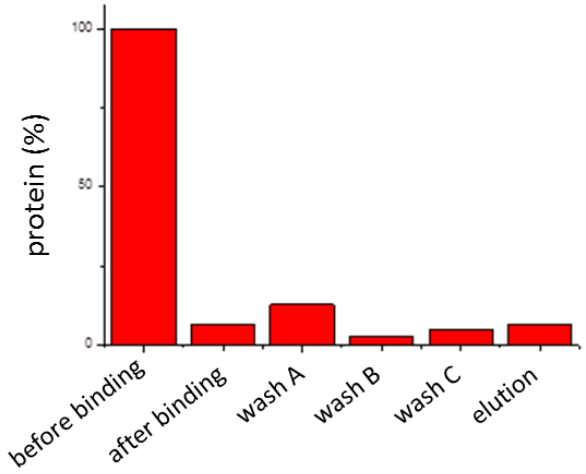

Figure 12: Tryptophan fluorescence measurements were used to measure the percentage of unspecific binging to the bead surface. In this experiment higher amount of V2R-H8 were used to show how high the unspecific binding was. 10 µl of Ni-NTA magnetic beads I were used for this experiment and the concentration of V2R-H8 during the binding process was 0.44 mg/ml (Total volume was 100 µl during the binding process → 44 µg of protein)

3.2.2Refolding on Ni-NTA resin

The idea was to setup a refolding screen on a 96-well format using filter plates and Ni-NTA Agarose 6 Fast Flow resin from GE Healthcare. The unfolded protein should be immobilized on the metal affinity resin and washed by different refolding buffer/buffers to generate an on-column refolding. This technique would allow as well as the refolding on magnetic beads a high throughput screen of several refolding condition in parallel. The elution would be possible in cysteine free buffer, which is necessary if a quantification of refolding efficiency wanted to be tested using the CPM-Assay. It is very important for this assay that the free cysteines of protein are the only ones, which are able to react with the dye, in the buffer.

It was possible to do a purification of the GPCRs with Ni-NTA resin in both cases (*α5I*-beta1AR-FB1-H8 and *α5I*-V2R-H8). The setup of a refolding screen on a 96 - well format using the same resin was less successful. The same problem was obtained like it was already showed under point *3.2.1 Refolding on magnetic beads*. The unspecific binding (Figure 13) of the protein to the resin surface was again the problem why it was not possible to establish an on-column refolding screen. Both types of His-Tag affinity material had an agarose surface. It was not possible to find out if the both surfaces are identical, but it does not matter. It can be concluded that neither magnetic beads with immobilizes Ni-NTA complex or Ni-NTA resin are useable to do a refolding of GPCRs under the tested conditions.

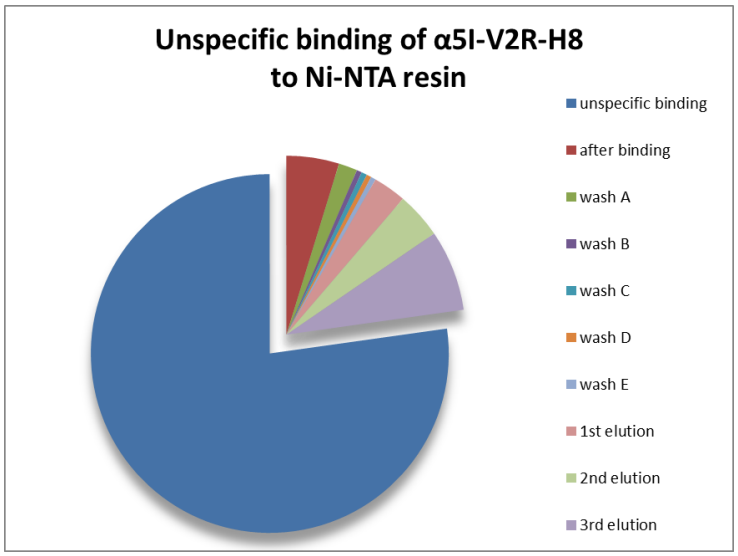

Figure 13: The quantification of protein in each fraction was done by tryptophan fluorescence measurements. 100% is representing thereby the used start-material. This figure shows what was generally observed: most of the protein (more than 75%) was binding unspecific to the surface of the resin (loading volume: 50 µl, washing volumes: 1000 µl, elution volumes: 100 – 170 µl).

3.2.3Refolding by dilution

Refolding by dilution is an alternative to refolding on beads, and eliminated the problem of non-specific interactions with beads. On the other hand, it may be complicated by protein aggregation. I performed fast dilution by adding refolding buffer in a quick way to denatured material. Detergent refolding was done by exchanging harsh detergent (SLS) against milder detergents such as DDM, DM or Cymal-6. A dilution ratio of more than 200 times was chosen to make guaranty that the denaturation conditions were very low during the refolding process. In literature, it is described that refolding by dilution should take place at low concentrations of protein (0.02 – 0.05 mg/ml) to minimise protein aggregation. A final protein concentration of 0.04 mg/ml was chosen to produce sufficient material (4 µg in 100 µl) to quantify of the refolding using CPM-Assay.

The refolding of solubilised GPCR by dilution refolding was tried only with V2R-H8, because good expression levels of beta1AR-FB1 were reached at the end of the practical work. The characterisation was done by CPM-Assay, radio-ligand binding, which was based on a new invited silica based affinity separation protocol and AUC. An Additional characterisation was done by several size-exclusion chromatography runs performed on a Superdex 200 16/30 column. The detection of the protein was done by UV absorption at 280 nm. The size-exclusion was made on an Ettan LC system from GE Healthcare. A running buffer was used which had almost the same components as the buffer which was used for NMR measurements. The refolding procedure had an influence on the protein. The unfolded sample (in

0.1% (w/w) SLS), which was injected in the same amount as the other from the refolding condition E, showed no specific elution peak. But the protein from the refolding condition E did. In case of the unfolded one it looks that some aggregation were eluting with the dead volume. The sample after the refolding process was eluting with a singular monodisperse peak at approximately 14-15 ml (Figure 14).

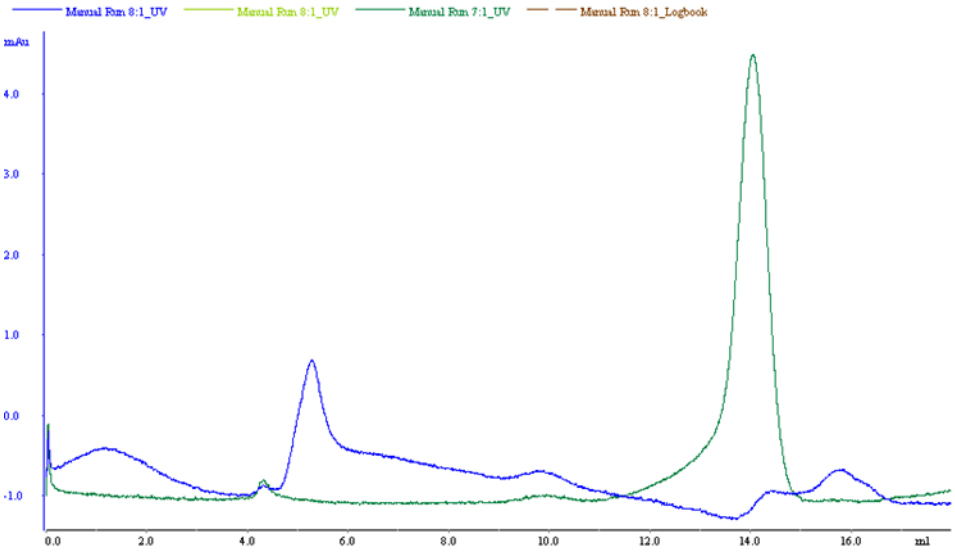

Figure 14: Overlay of: size exclusion chromatogram of unfolded V2R-H8 in blue and V2R-H8 handled by refolding condition E in green. The size-exclusion was performed on a Superdex 200 16/30 column, flow 0.5 ml/min, buffer: 20 mM HEPES pH 7.3, 150 mM NaCl, 0.05% (w/w) DDM, filtered through 0.22 µm, degased.

According to the calibration of the column, this peak position correspond to a molecular weight of less than43 kDa. The expected molecular weight of protein inclusive micelle should be approximately 80 kDa. At the beginning, it was thought that the GPCR might be interacting with the size-exclusion material and that leaded to the later elution time. Later the results of AUC and LC-MS give the evidence that an addition cleavage was taking place in the GPCR sequence during the thrombin cleavage (see *3.2.6 Analytical ultracentrifugation* and *3.2.7 Liquid chromatography mass spectrometry*). It still looked like that maybe a refolding of the truncated protein (TM 3 – 7) was possible because a symmetric size-exclusion peak is an indicator for integrity of GPCRs (Grisshammer, 2009).

Different refolding conditions were compared by their size-exclusion too. The overlay of those shows a correlation of their height to the refolding efficiencies which were measured with the radio-ligand binding assay (Figure 15). The best refolding efficiency is again refolding condition E followed by condition A which contained mostly the same components in the refolding buffer, only Tolvaptan was missing. Tolvaptan was used because it was assumed that the ligand of the vasopressin V2R receptor may assist during the refolding process. If we compare the result of refolding condition A and E, it can be concluded that Tolvaptan was able to increase the refolding efficiency.

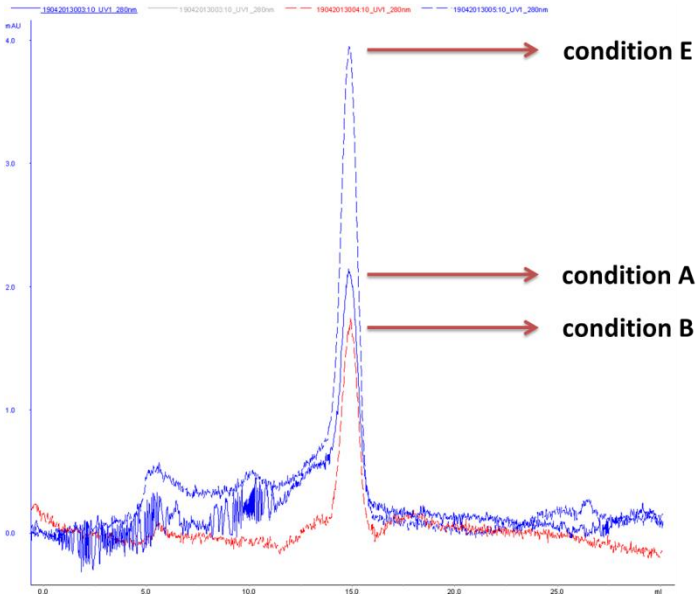

Figure 15: Overlay of the size-exclusion chromatograms of the best three tested refolding conditions. All three refolding conditions were showing a monodisperse peak and same elution time. They differed by their peak heights which were correlating to the radio-ligand binding results of these refolding conditions.

3.2.4CPM-Assay

The CPM-Assay allows measuring the thermal transition or transitions of proteins, which can be used as characterisation method for the refolded material if purified protein from natural sources is available. It is also a fantastic tool to do fast optimization of buffer conditions for your protein by analyse temperature shifts which may induced by different ligands or detergent/lipid combinations. One negative aspect is that purified material is need to make safe assumptions, but this in most cases no problem, because only few µg of protein are necessary for this assay.

The CPM-Assay was used in this thesis as a fast measuring technique for the quantification/qualification of the different refolding condition of the dilution refolding screen: first, is there any refolding taken place and second, how efficient was the refolding in percentage to the used material. For the setup and evaluation of this measurement Opsin was taken as reference material to evaluate how much protein is necessary to be able to measure a refolding in presence of unfolded material which is able to react with the dye. Additionally, it was the aim to find out how high is the corresponding fluorescence intensity of one dye molecule is after the reaction with a free cysteine from the inner core of a GPCR. In other words, how much protein had to be put into the refolding screen that quantification with CPM-Assay is possible.

For this evaluation several different concentrations of Opsin (0.05 to 5 µg/100 µl) were measured in different buffers, in absence and in presence of high percentage of unfolded material (10% of folded Opsin to 90% of unfolded V2R-H8). For the measurement, 90 µl of the test solution was mixed with 10 µl of a 1:40 dilution of a 3 mg/ml CPM dye stock (in DMSO). The dilution of the dye was always done with measuring buffer of the actual experiment. The whole sample preparation was done on ice to be sure that

the protein was not already denaturing during the preparation. The mixed reaction solutions were divided in 25 µl aliquots and pipetted into measuring tubes of the 72-Well Rotor of Rotor-Gene Q from Qiagen. The program to measure the transition states was started at 20°C and a temperature ramping protocol of 5°C per min with a constant gradient up to 90°C was applied. The evaluation was done by the Rotor-Gene Q Software from Qiagen.

For testing of the refolding efficiencies of dilution refolding experiments was an additional purification step necessary, because the refolding buffers contained reduced glutathione. The reduced glutathione reacts with the CPM dye and makes a direct testing from the refolding buffer impossible. The separation of protein from its refolding buffer was done with silica based Ni-TED resin which was used already for the purification of the protein. 20 ± 0.2 mg of dry silica based Ni-TED resin were weighted into a 1.5 ml eppendorf tube. 200 µl of CPM_{DDM} buffer were added to pre-soak the resin for 10 min. Afterwards, the buffer was replaced by 100 µl of the refolding solution. The resin was incubated for approximately 10 min on ice. During this incubation time the beads were resuspended several times. The eppendorf tube was spun down and the refolding solution was discarded. The resin was washed twice with CPM_{DDM} buffer to get rid of almost all refolding additives. The elution was done with 100 µl of $CPM_{Elution}$ buffer for 5 min by batch mode incubation. The CPM-Assay was done as already described above.

Optimization of the CPM-Assay:

The evaluation of the CPM-Assay was done with Opsin and solubilized *α5I*-V2R-H8 construct from the purification step. Opsin gave a clear transition in absence and presence of unfolded *α5I*-V2R-H8 construct. The fluorescence signal at the start was in presence of unfolded material higher compare to the sample without unfolded protein. The higher signal depended on two reasons: First, more material was in each reaction, because it was additionally added unfolded material (→ approximately five fold more). Second, the unfolded material had probably all cysteine exposed to the dye. It was observed that the delta between the two fluorescence signals of 0.5 µg Opsin/100 µl was decreasing a bit compare to the delta of the sample without unfolded material. In case of the higher concentration of Opsin only a shift of the whole signal was observed which could belong to the higher SLS concentration in the reaction volume. The Transition of 1.25 µg/100 µl of Opsin + 90% unfolded *α5I*-V2R-H8 showed a shifted transition state to a slightly lower T_m which might be correlated with higher concentration of SLS in the sample. The 2.5 fold higher concentration of SLS was coming from the adding of the unfolded *α5I*-V2R-H8 which had been purified and stored in this buffer (Figure 16). A transition temperature of 51 to 52°C was measured for Opsin what had been measured from a lab member already before (Table 9).

More different concentrations of Opsin were tested to see how low the concentration could be to get still a significant fluorescence signal which can be used for the quantification of the refolding efficiencies. The optimum concentration was approximately 1 µg/100 µl which lead to a strong enough signal. It was con-

cluded from these results that for a refolding screen would it be the best if 4 µg of protein were used in 100 µl. This would allow to do a detection of a refolding efficiency down to approximately 1-5%.

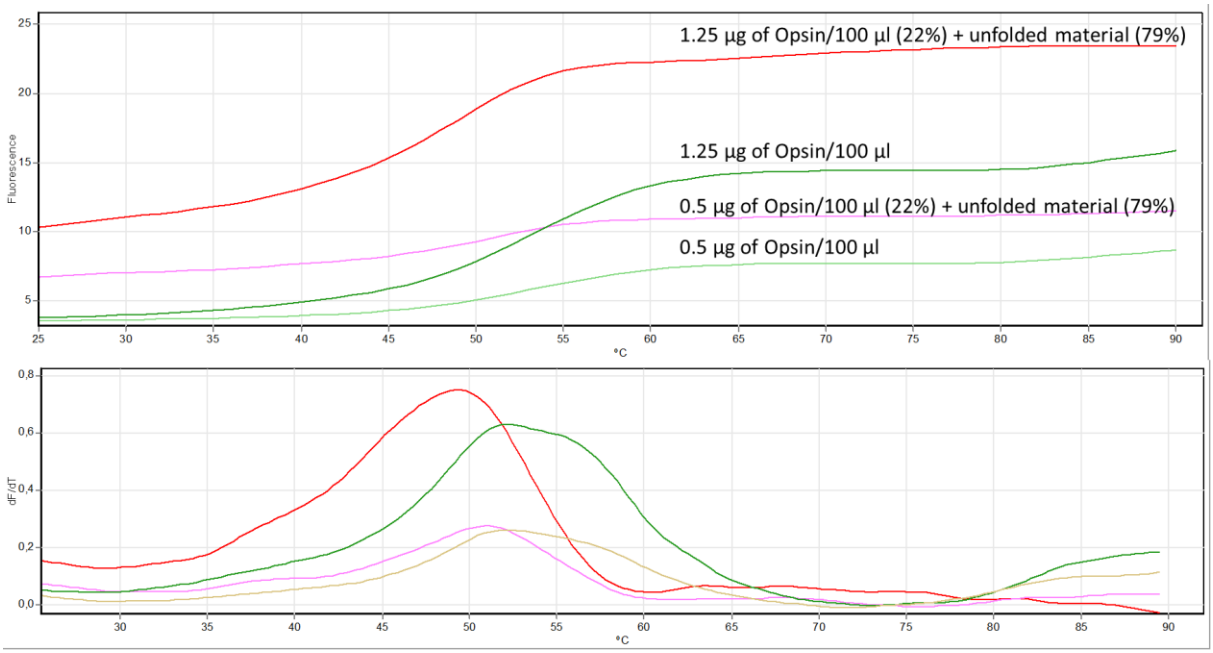

Figure 16: Dataset from CPM-Assay of Opsin without and with unfolded *α5I*-V2R-H8. The upper graph shows the fluorescence signal against the temperature ramping. The lower graph illustrate the first derivation of the fluorescence signal against the temperature.

Table 9: T_m of Opsin with and without of unfolded material (= *α5I*-V2R-H8). It was observed that SLS leads to a destabilization of Opsin already at low concentrations. The T_m is an average out of 4 measurements of each sample.

Experiment	T_m [°C]
1.25 µg of Opsin/100 µl (22%) + unfolded material (79%)	49.3 ± 0.0
0.5 µg of Opsin/100 µl (22%) + unfolded material (79%)	51.0 ± 0.0
1.25 µg of Opsin/100 µl	52.2 ± 0.1
0.5 µg of Opsin/100 µl	52.2 ± 0.1

Characterisation of refolded material:

The approximately refolding efficiencies were calculated from the delta fluorescence signal before and after the transition. The calculation contained information about the amount of protein and also the different number of cysteines which were able to react with the dye during the unfolding. For the identification of the cysteines which can react with the dye a computer model of both GPCRs was used. Opsin has three cysteines which are located in the core of the protein and V2R has two cysteines. The calculated refolding efficiency of refolding condition A was approximately 7%. The high fluorescence signal at the beginning was probably correlating with the amount of unfolded protein in the sample. Because the separation protocol for the protein did not distinguished between folded and unfolded receptor. The fluorescence signal was starting at a high level and was decreasing with increasing temperature the over time. Thermo

quenching might be responsible for that decrease of signal intensity. The refolding efficiency of refolding condition B was approximately the half compare to condition A.

It was concluded that an accurate calculation of refolding efficiency is not really possible by this assay, because the separation process with the silica beads has an influence on the signal intensity (Figure 17). The binding of the protein was not always the same also if approximately the amount of resin was used in all experiments. The CPM-Assay is not the best way to do efficiency calculations of refolding conditions, but it is a good method to distinguish between is there any refolding or no refolding.

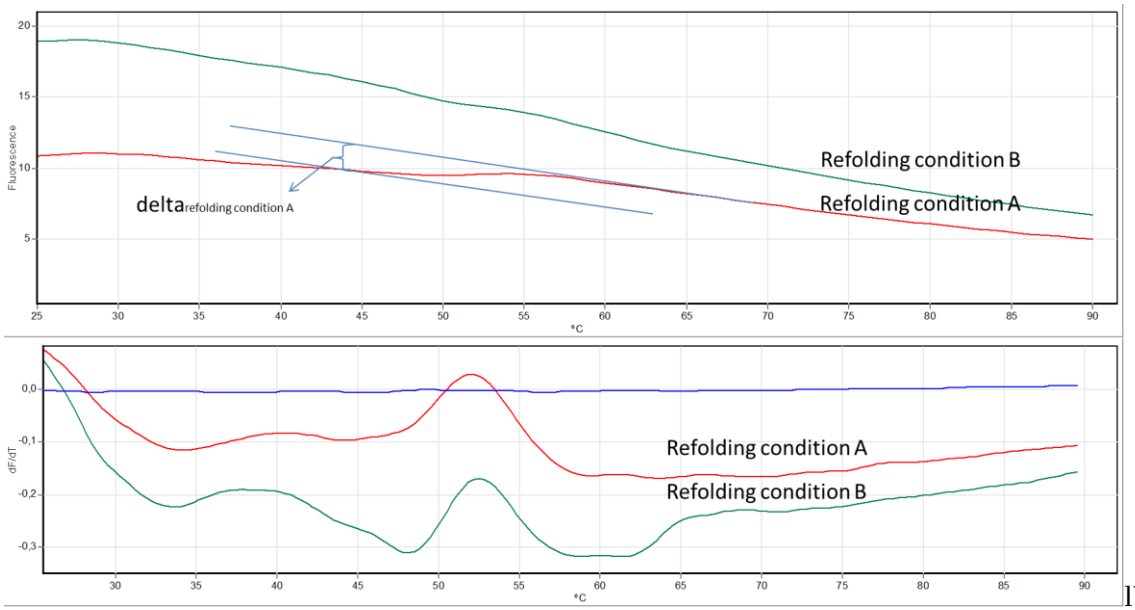

Figure 17: The melting curves of the refolding conditions A (DDM) and B (DDM + Cymal-6) after the separation of protein from refolding buffers using silica based Ni-TED beads. The approximately refolding efficiencies were calculated using the delta between before and after transition like it is shown in the upper graph. The delta fluorescence signals of the refolding conditions were compared with the delta fluorescence signals of Opsin.

3.2.5Radio-ligand binding

Size-exclusion based radio-ligand binding assay

At the beginning, it was tired to establish a radio-ligand binding assay based on size-exclusion separation of bound and unbound hot ligand on a 96-well filter plate. A similar protocol was already used already for the radio-ligand binding assay of β1-adrenergic receptor. The separation is thereby done by size-exclusion spin columns which needs a lot of handling and is unpractical for a high-throughput radio-ligand binding assay. A radio-ligand binding assay for V2R was also well-established in our lab, but it was only usable for whole cells or membrane fraction because it is a filter based assay. Due to this it was important to have a new assay for solubilised/refolded protein too.

The evaluation of the resolution was done by using BSA and fluorescein. A mixture of these two compounds was loaded on the small 1 cm size-exclusion column in the 96-well filter plate and the elution was

done by stepwise adding of 100 µl elution buffer. I observed a baseline separation for these two substances, where BSA should represent the GPCR and fluorescein the small vasopressin peptide. The separation seemed to work very well, but a high background was measured during the radio-ligand binding assay. The background of 1000 counts was to high that it could belong to unspecific binding only. Due to that it was recorded a elution profile of the hot vasopressin too. The elution profile of this small peptide was more comparable to the elution profile of a big protein such as BSA and V2R than to a small molecular compound (Figure 18). The reason for the bad resolution might be that the vasopressin peptide incorporate into the empty detergent micelles which were eluting with the protein together. On this account was it not possible to establish a radio-ligand binding assay based on a size-exclusion separation by the tested conditions.

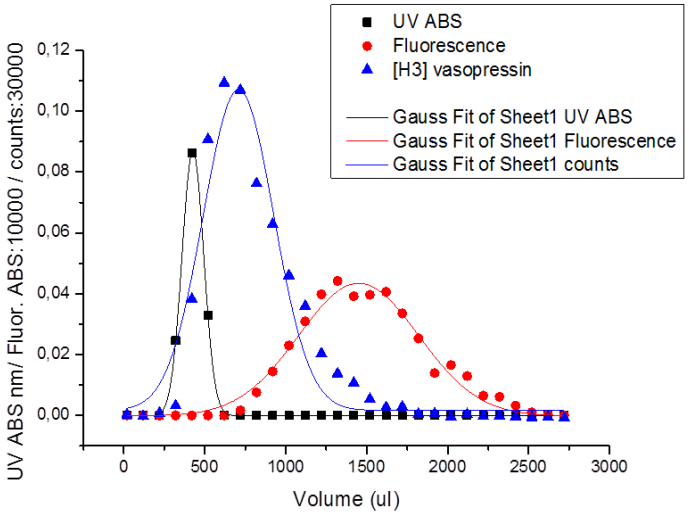

Figure 18: Size-exclusion profile of the separation of BSA (UV signal at 280 nm) and Fluorescein (fluorescence signal at 515 nm) overlaid with the elution profile of [H3] labeled vasopressin peptide (measured by a scintillation counter). The fluorescence signal was divided by 10.000 and the counts by 30.000 to get everything on the same scale. The protein BSA which should represent the GPCR was eluting in the fractions 3 to 5 (between 220 to 520 µl using a step elution of 100 µl of buffer each time). The Fluorescein which should represent the vasopressin peptide was eluting with a much broader peak (started at 620 µl and ended at 2720 µl).

Silica based Ni-TED radio-ligand binding assay:

The silica based Ni-TED radio-ligand binding assay was performed on two different ways like it is described in the method part (*2.2.5 Radio-ligand binding*). The first, Variation A (Figure 19) was done by an excess of protein compare to hot ligand. 11 times more protein was used during 1 h of batch binding. This experiment was done to see if there is any binding of hot ligand to the protein. It was not important to get an exactly refolding efficiency in percentage from this experiment, because it wanted to be evaluated if there is any refolding taking place. The counts from unfolded compare to refolded material were significantly different. This result gave an evidence that refolding condition A assisted the reconstitution of V2R. A very low intensity was measured for the unfolded sample which showed very low unspecific binding. This assay was not designed to calculate the refolding efficiency, but it allows an estimation of it. The reference signal was used as reference for a 100% binding of the used ligand during the incubation. A refolding efficiency of ≥ 6% was calculated for refolding condition A.

The experiment was done above the K_D, so that most of the ligand should be bound to the refolded protein if it was present in solution. The washing step might have resulted in a reduced signal because the washing process on the silica resin took almost 20 min. This extended washing can reduce observed signal due to dissociation of the ligand.

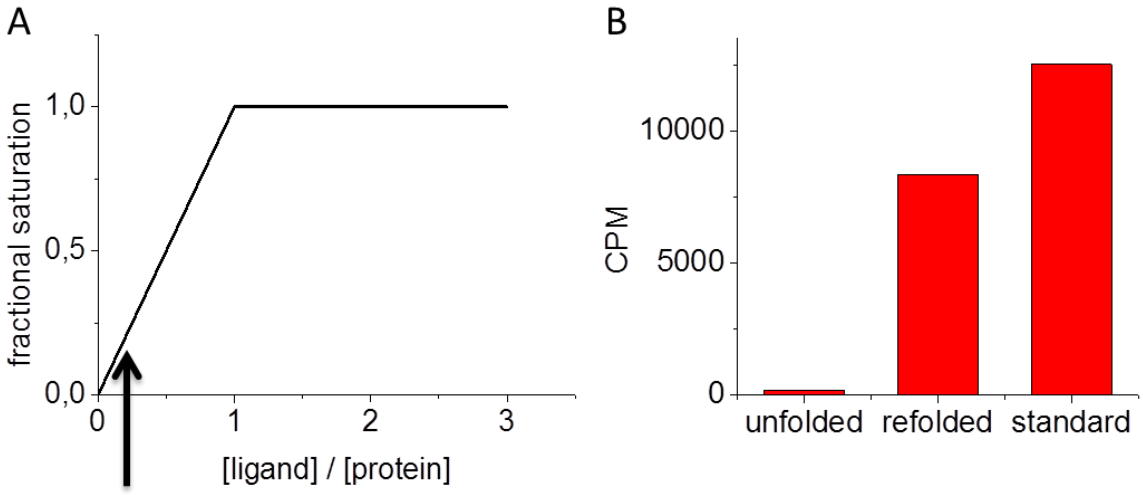

Figure 19: Variation A of silica based Ni-TED radio-ligand binding assay. **Diagram A** should visualise where the concentration of hot ligand was located during the batch binding. **Diagram B** shows the counts which were measured with TopCount NXT™. The counts for unfolded protein were very low compare the counts which were measured for the refolded sample. The refolding was done for this experiment with refolding condition A.

The second radio ligand binding assay based on silica resin is described as Variation B in the method part. The measurement wanted to be done without an additional dilution step from the refolding solution. A hot ligand concentration was not affordable because of the costs of hot vasopressin ligand. Due to that the hot ligand was supplemented by cold ligand to get an equal concentration of vasopressin peptide to protein during the binding process (Figure 20). It was assumed that the tritium label of vasopressin peptide had no influence on the binding process.

Unspecific binding was tested with unfolded protein like in Variation A. The refolding condition E had additionally 10 µM of Tolvaptan inside the refolding solution. The binding affinity of Tolvaptan to V2R is almost the same as for vasopressin peptide. Due to this a competition was taking place between vasopressin peptide and Tolvaptan. For the calculation of the refolding efficiency the concentration of all ligands were pooled. This is the reason why the counts of the refolding condition E (Table 10) were mostly the same but the calculated refolding condition differed by 60%. The refolding efficiency could be also higher, because of the washing steps (see above). The result should be taken critically because this experiment was done only once and has to be repeated more times.

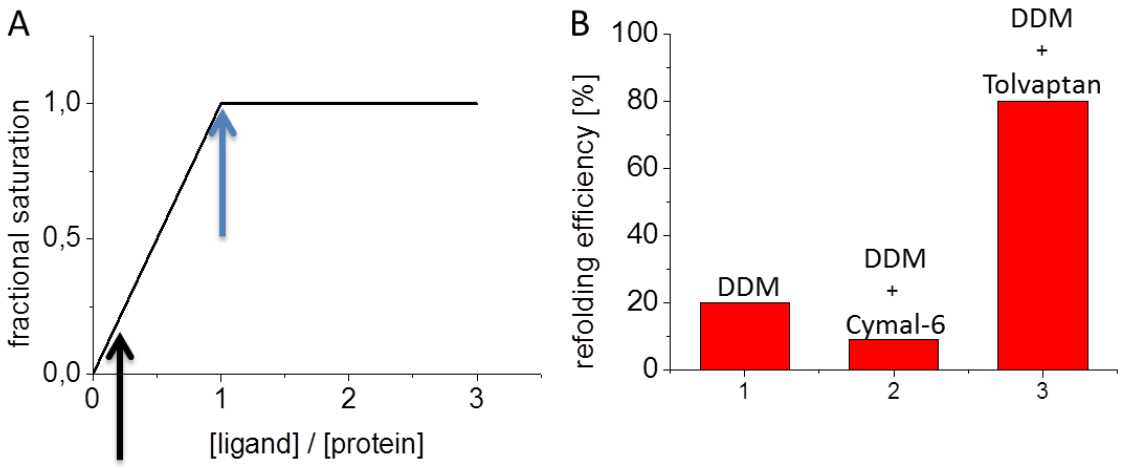

Figure 20: Diagram A: The graph visualise that the binding was taking place at 1 to 1 ratio of protein to ligand for refolding condition A (DDM) and B (DDM + Cymal-6). The refolding condition E (DDM + Tolvaptan) had an large excess of Tolvaptan (10 µM) inside the binding solution. The black arrow shows at which ratio radio-ligand binding Variation A was working and the blue arrow visualises where the binding for refolding condition A and B was located. **Diagram B** shows the refolding efficiency of three different refolding conditions: A, B and E. In case of refolding condition E a refolding efficiency of 80% was calculated.

Table 10: Measuring results of Variation B experiment and calculated refolding efficiencies

	counts	delta (x-resin)	refolding efficiency [%]
refolding condition A	445	250	21
refolding condition B	308	113	9
refolding condition E	291	96	80
unfolded protein	141	-54	
resin	195	0	
positive control	6816	6621	

Due to the result of the AUC and LC-MS measurements (see below) it seemed that the refolding experiments were done with an truncated vasopressin V2 receptor. If the accurate mass of the MS-measurement is the true mass it was worked with a receptor where the first two helices were missing. The experiment of the refolding efficiency was done only once but Variation A showed a significant increase in the counts after refolding. It could be that the truncated GPCR was still able to bind its ligand. However, the positive signal of the radio-ligand binding looked promising for future experiments. Due to that the refolding experiment should be repeated at soon as the full-length V2R has be purified. The new silica based radio-ligand binding assay which was developed during this thesis seemed to work very well, but the assay needs more evaluations.

3.2.6Analytical ultracentrifugation

The analytical ultracentrifugation (AUC) measurements were performed only once with a refolded unlabelled V2R-H8 sample which was prepared on the same way as the NMR samples. The sedimentation velocity run was used to measure to oligomerisation state of the refolded receptor. A monodisperse peak was obtained which had a sedimentation coefficient of 3 S. The concentration of the sample was approximately 20 µM which was five times lower compare to the NMR measurement. The determination of the molecular weight of the receptor-detergent micelle was done by a sedimentation equilibrium run. The measured molecular weight was 27 ± 1 kDa (Figure 21). This result was supporting the result of the LC-MS experiment that a unwanted cleavage of the vasopressin V2 receptor was taking place in the GPCR sequence during the thrombin cleavage. The measured molecular weight did not explained why the detergent micelle which should be there to keep the protein in solution did not contributed to the obtained molecular weight, as it has a higher density compare to water. At the moment we have no explanation for that observed phenomena that the detergent micelle did not contributed to the molecular weight of the receptor. Nevertheless, both experiments have to be repeated with full length receptor.

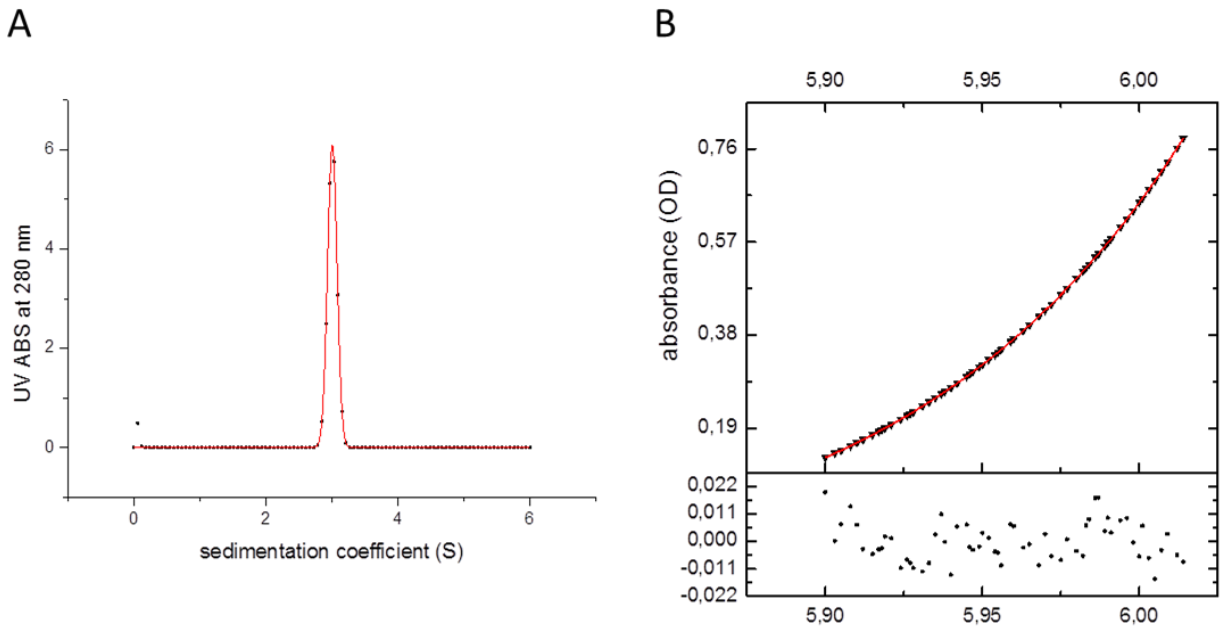

Figure 21: Diagram A shows the results of the sedimentation run. A sedimentation coefficient of approximately 3S was measured. The fitting of the data set was done by a GaussAmp fit. Diagram B visualize the results from the equilibrium velocity run. The lower of this two diagram shows the randomly distributed errors.

3.2.7 Liquid chromatography mass spectrometry

The mass spectrometry of membrane proteins is complicated by the presence of detergent micelle which suppresses ionisation. A measurement of the protein mass using Variation A was not possible. The detergent micelle around the protein was probably the reason for it. The second described protocol Variation B was more useful. The new precipitation protocol allowed to do LC-MS of intact IMPs. The results indicated that an unwanted cleavage of the *α5I*-V2R-H8 construct in the GPCR was probably taking place. The thrombin was cleaving most likely in a second position which was approximately 100 amino acids from the N-terminus of V2R-H8 located (Figure 22). It was not possible to identify the exact position yet. The cleavage site must be located from the measured mass in the first extracellular loop of the GPCR, which mean that 2 helices were missing. This result emphasises the importance of mass spec analysis of membrane proteins and GPCRs in particular due to their often unusual mobility on SDS-Page. The protocol established in this work will help to accurately measure molecular weight of GPCRs and improve quality control of recombinant protein production.

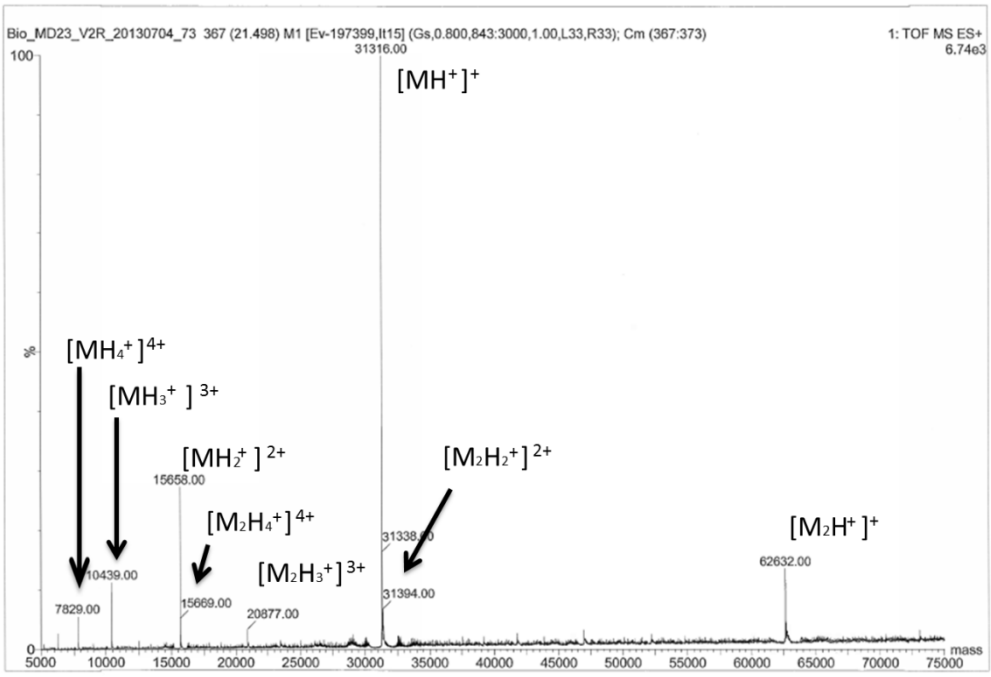

Figure 22: The MS measurement was done in the positive mode and the spectrum showed a predominating protein with an mass of 31.316 Da. Additional, to the predominating species $[MH^+]^+$ several other species were present → multiple charged monomers and dimers of the protein. The mass of the detected protein was approximately 12 kDa less to what it should be (43 kDa). From this data it was concluded that a second cleavage site exists in the construct.

3.3 NMR Experiment

A 2D NOESY experiment was done for unlabelled and ^{15}N-labelled V2R-H8 (Figure 23). The experiments for the unlabelled receptor were performed on a 700 MHz and for the ^{15}N-labelled V2R-H8 on a 600 MHz instrument. High concentration of DDM made the measuring of the unlabelled sample very difficult, because of a low signal to noise ratio. A Watergate element, which is usually used to suppress the solvent signal, was modified to suppress detergent signals. The Watergate element supressed the solvent that it was not flipped back by the first shaped pulse and also supressed the DDM signals because they were not refocused by the second selective pulse and therefore their phase were scrambled along the z-axis by the gradient pair G1 G2.

The modified Watergate made it at least possible to get some information from the amid-signal area where it was impossible to get any information without using the modified Watergate. The results from the unlabelled V2R-H8 sample were not good enough to make an assumption about if the GPCR is correct fooled or not. From that experiment we were able to learn at least something about handling a NMR sample, how the preparation of the sample should be done and that DDM might not be the best detergent to do NMR studies with it.

A ^{15}N-labelled V2R sample was tried to measure too, but the result was disappointing. No increases of the amide-signals were observed compare to the unlabelled sample. It seemed that the labelling strategy with the high-complexity M9 medium and the pre-growing in high dense medium, was not really working. The procedure with the labelled protein was tried only once and has to be repeated again to be sure that this protocol is not working. It could also be that the expression of unlabelled protein was taken place until every easy available nutrient were used up which was still left from the growing in high density media (TB medium). But the cells were grown in the M9 media for 1 h before the induction was done with 1 mM IPTG. This should be enough time to consume the nutrition which were still left in the cells from the media exchange.

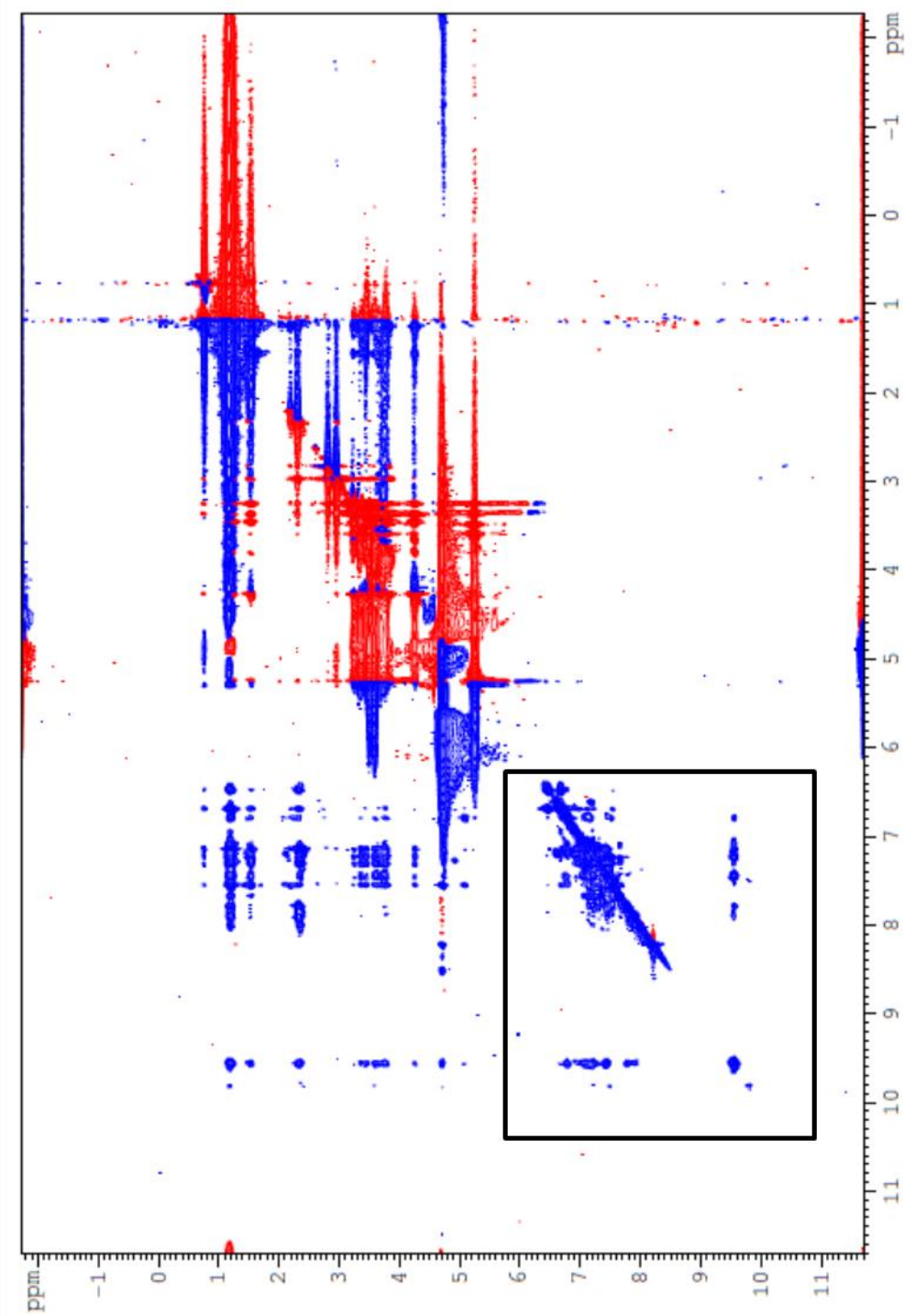

Figure 23: 2D NOSEY spectra of unlabeled V2R-H8 in 20 mM NaH_2PO_4 pH 7.3, 150 mM NaCl, 0.05% (w/w) DDM at 20°C. The 2D NOESY was measured on a 700 MHz spectrometer with 64 scans per increment and maximum evolution times of 21 ms and 100 ms in t_1 and t_2 respectively for a total time of 9 h. The 1D was measured with 64 scans and an evolution time of 180 ms. The framed rectangle should show where we suppose to see long range coupling of amides which gives information about the folding of the protein. From this spectrum it is hard to say if there is any refolding. The signal of unlabeled protein is to week to give a safe assumption.

4 Conclusion

In this work I was able to express both GPCR constructs in inclusion bodies to produce milligram amounts of recombinant protein. The developed purification protocol delivers additionally good yields of those in a high purity which is important for a refolding screen. While the refolding screen either by the magnetic bead or on-column refolding strategy did not work due to non-specific binding of GPCR to the bead surface, the refolding screen by dilution delivered more promising results. The possibly refolded protein V2R showed a monodisperse peak and was also able to bind its ligand vasopressin. The efficiency of refolding was improved in the presence of V2R ligand Tolvaptan. The refolding of beta1AR-FB1 was not tried because of luck of time. Towards the end of the thesis I developed a sample preparation protocol which allows LC-MS analysis of GPCRs. I identified that V2R which was used for the refolding had cryptic thrombin cleavage site approximately between second and third transmembrane helixes. Immediate efforts should be directed to produce full length receptor and its refolding screen performed using methods established during this work. Overall, the results obtained in the thesis show that the *E. coli* is a very promising host for expression of GPCRs for NMR studies, and that the refolding strategy should be pursued in the future.

5 Outlook

The future steps for this project should be to design a new construct of *α5I*-V2R-H8. This means to inhibit the unwanted thrombin cleavage by specific mutations or to replace the thrombin cleavage site with a more specific protease cleavage site. In my opinion it would be better to replace the cleavage site, as it would avoid mutations in the V2R and be applicable to other receptors. In case of construct *α5I*-beta1AR-c.o.-FB1-H8 it has to be find out why the 3C Protease cleavage was not working very efficient. This will include optimisation of buffer condition and testing if the detergent SLS is responsible for the inhibition. Alternative detergents may need to be used for the cleavage step. The refolding screen for both receptors should be performed again using the conditions and evaluation methods which I have developed in my thesis.

For the characterisation of the refolded receptors it is important to do a more precisely quantification of the new radio-ligand binding assay, in comparison to the receptor expressed in native conformation. Additionally, it is important to complement the ligand binding with another functional assay such as the G-protein activation assay. The next step is to find NMR conditions which produce good NMR spectra and at the same time keep the receptor in its native conformation. Finally, we are aiming to measure NMR spectra of the receptor in order to study how its conformation and dynamics is affect by the binding of ligands with different pharmacological properties. The thorough understanding of how GPCR structure is affected by activity modulating ligands will have a great impact on our understanding of their physiological and pharmacological properties and may contribute to the future drug design efforts.

6 Abstract (German)

G-Protein gekoppelte Rezeptoren (GPCRs) gehören zu den Membranproteinen und sind zu einem erheblichen Anteil an den Signalübertragungen, vom extrazellulären in den intrazellulären Raum, beteiligt. Sie sind an vielen wichtigen regulatorischen Prozessen in der Zelle involviert. Des Weiteren werden ihre Fehlfunktionen mit einer großen Anzahl an Krankheiten in den Zusammenhang gebracht, wie zum Beispiel: Entstehung vom Krebs, kardiale Dysfunktion, Störungen des zentralen Nervensystems, Übergewicht, Entzündungen und Schmerz. Etwa ein Drittel aller derzeit auf dem Markt verfügbaren Medikamente wirken auf diese Rezeptoren. Um ein besseres Verständnis über ihre Funktion auf molekularer Ebene zu erhalten, ist es sehr wichtig, ihre Struktur und Dynamik zu kennen. Derzeit stellt NMR die beste und am weitesten entwickelte Technik dar um die Dynamik von Proteinen zu erforschen. Für NMR Untersuchungen ist es jedoch notwendig isotopenmarkierte Proteine zu expremieren, was jedoch oft ein Problem darstellt. *E. coli* ist das am besten etablierte Expressionssystem und ermöglicht eine Vielzahl an unterschiedlichen Isotopenmarkierungen. In dieser Arbeit gelang es zwei humane GPCRs erfolgreich in Milligrammmengen zu expremieren. Die verwendeten GPCRs sind zum Einen der Wildtyp des menschlichen Vasopressin V2 Rezeptors (V2R) und zum Anderen eine stabilisierte Version des menschlichen β1-Adrenergic Rezeptors (beta1AR-FB1). Des Weiteren war es möglich ein Protokoll für die Aufreinigung rekombinant hergestellter Proteine zu etablieren, welches ausgezeichnete Ausbeuten und eine hohe Reinheit lieferte. In weiterer Folge wurde ein Rückfaltungsprotokoll entwickelt, das einen small-scale Screen verschiedenster Rückfaltungsbedingungen erlaubt. Die verschiedenen Rückfaltungsprodukte wurden mit folgenden Methoden charakterisiert: Temperaturstabilitätsassay, Radio-Ligand Binding Test, Massenspektrometrie, Gel-Permeations-Chromatographie und Analytische Ultrazentrifuge. Die Ergebnisse deuten darauf hin, dass die Rückfaltungsstrategie ein vielversprechender Ansatz für das weitere Forschungsprojekt darstellt. Mit ein paar wenigen weiteren Optimierungen lassen sich höchstwahrscheinlich vielversprechenden Proben für zukünftige NMR-Studien produzieren.

7 CV

Education & Qualifications

01/2013 – 07/2013	Master. Thesis at Paul Scherrer Institut, Switzerland Group of Prof. Gebhard F. X. Schertler
10/2011 - present	Master student at Department of Chemistry, University of Vienna Major in Organic, Analytic and Food Chemistry Date of final exam, fixed on 27th of August 2013
10/2009 - 02/2010	Three month internship at Paul Scherrer Institut, Switzerland Group of Prof. Gebhard R. X. Schertler
10/2008 - 09/2011	BSc Department of Chemistry, University of Vienna

Research and Laboratory Experience

01/2013 – present	**Master's Thesis** **Prof. Gebhard F. X. Schertler,** **Laboratory of Biomolecular Research, Paul Scherrer Institut (PSI), Switzerland** Supervisor: Dr Dmitry Veprintsev Project: Expression and Refolding of human GPCRs for NMR studies
07/2012 – 09/2012	**Internship** **Prof. Gebhard F. X. Schertler,** **Laboratory of Biomolecular Research, Paul Scherrer Institut (PSI), Switzerland** Supervisor: Dr Dmitry Veprintsev Project: G-protein selectivity of G-protein-coupled receptors
10/2011 - 03/2012	**Bachelor's Thesis** **Prof. Doris Marko, Department of Food Technology and Toxicology at University of Vienna** Supervisor: Dr. Gudrun Palke Project: The influence of the anthocydien Delphinidine on the human grow factors EGF and VEGF
Education Priorities	Organic and inorganic synthesis Methods: cell culture, protein expression & purification, ELISA, PCR, mutagenesis, Comet-Assay, Western Blot, ELISA, Nanodisc-Assay, HPLC, HPLC-MS, GC, GC-MS, MALDI-MS, LC-MS, AAS, EES, XRF, IR, NMR, DC, diverse methods of titration
Grants	Erasmus Internship fellowship (for three month at PSI, summer 2012) Vorarlberg fellowship (for master thesis at PSI, January to June 2013)

Publications

Sun D[1], Ostermaier MK[1], Heydenreich FM[1] ([1] equal contribution), Mayer D, Jaussi R, Standfuss J, Veprintsev DB. AlaScan-automatic primer design for scanning mutagenesis. Submitted.

8 References

Ahmet, I., Krawczyk, M., Heller, P., Moon, C., Lakatta, E. G., & Talan, M. I. (2004). Beneficial effects of chronic pharmacological manipulation of beta-adrenoreceptor subtype signaling in rodent dilated ischemic cardiomyopathy. *Circulation*, *110*(9), 1083–90.

Alexandrov, A. I., Mileni, M., Chien, E. Y. T., Hanson, M. a, & Stevens, R. C. (2008). Microscale fluorescent thermal stability assay for membrane proteins. *Structure (London, England : 1993)*, *16*(3), 351–9.

Arcemisbéhère, L., Sen, T., Boudier, L., Balestre, M.-N., Gaibelet, G., Detouillon, E., Orcel, H., et al. (2010). Leukotriene BLT2 receptor monomers activate the G(i2) GTP-binding protein more efficiently than dimers. *The Journal of biological chemistry*, *285*(9), 6337–47.

Armstrong, S. P., Seeber, R. M., Ayoub, M. A., Feldman, B. J., & Pfleger, K. D. G. (2013). Characterization of Three Vasopressin Receptor 2 Variants: An Apparent Polymorphism (V266A) and Two Loss-of-Function Mutations (R181C and M311V). *PloS one*, *8*(6), e65885.

Baneres, J.-L., Martin, A., Hullot, P., Girard, J.-P., Rossi, J.-C., & Parello, J. (2003). Structure-based analysis of GPCR function: conformational adaptation of both agonist and receptor upon leukotriene B4 binding to recombinant BLT1. *Journal of molecular biology*, *329*(4), 801–14.

Banères, J.-L., Mesnier, D., Martin, A., Joubert, L., Dumuis, A., & Bockaert, J. (2005). Molecular characterization of a purified 5-HT4 receptor: a structural basis for drug efficacy. *The Journal of biological chemistry*, *280*(21), 20253–60.

Banères, J.-L., Popot, J.-L., & Mouillac, B. (2011). New advances in production and functional folding of G-protein-coupled receptors. *Trends in biotechnology*, *29*(7), 314–22.

Bazzacco, P., Billon-Denis, E., Sharma, K. S., Catoire, L. J., Mary, S., Le Bon, C., Point, E., et al. (2012). Nonionic Homopolymeric Amphipols: Application to Membrane Protein Folding, Cell-Free Synthesis, and Solution Nuclear Magnetic Resonance. *Biochemistry*, *51*(7), 1416–1430.

Bockaert, J., & Pin, J. P. (1999). Molecular tinkering of G protein-coupled receptors: an evolutionary success. *The EMBO journal*, *18*(7), 1723–9.

Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III, Skelton, N. J., & Rance, M. (2010). *Protein NMR Spectroscopy: Principles and Practice (Google eBook)* (p. 912). Academic Press.

Dong, X.-Y., Chen, L.-J., & Sun, Y. (2009). Refolding and purification of histidine-tagged protein by artificial chaperone-assisted metal affinity chromatography. *Journal of chromatography. A*, *1216*(27), 5207–13.

Glynou, K., Ioannou, P. C., & Christopoulos, T. K. (2003). One-step purification and refolding of recombinant photoprotein aequorin by immobilized metal-ion affinity chromatography. *Protein expression and purification*, *27*(2), 384–90.

Grisshammer, R. (2009). *Purification of recombinant G-protein-coupled receptors. Methods in enzymology* (1st ed., Vol. 463, pp. 631–45).

Grisshammer, R. (2013). Why we need many more G protein-coupled receptor structures. *Expert review of proteomics*, *10*(1), 1–3.

Hiller, C., Kühhorn, J., & Gmeiner, P. (2013). Class A G protein-coupled receptor (GPCR) dimers and bivalent ligands. *Journal of medicinal chemistry*.

Howlett, G. J., Minton, A. P., & Rivas, G. (2006). Analytical ultracentrifugation for the study of protein association and assembly. *Current opinion in chemical biology*, *10*(5), 430–6.

Huang, K. S., Bayley, H., Liao, M. J., London, E., & Khorana, H. G. (1981). Refolding of an integral membrane protein. Denaturation, renaturation, and reconstitution of intact bacteriorhodopsin and two proteolytic fragments. *The Journal of biological chemistry*, *256*(8), 3802–9.

Katoh, S., Kumada, Y., & Maeshima, N. (2005). Template Refolding by Use of an Antibody-Coupled Affinity Column. *Chemical Engineering & Technology*, *28*(11), 1394–1397.

Kiefer, H, Krieger, J., Olszewski, J. D., Von Heijne, G., Prestwich, G. D., & Breer, H. (1996). Expression of an olfactory receptor in Escherichia coli: purification, reconstitution, and ligand binding. *Biochemistry*, *35*(50),

Kiefer, Hans. (2003). In vitro folding of alpha-helical membrane proteins. *Biochimica et biophysica acta*, *1610*(1), 57–62.

Kolakowski, L. F. (1994). GCRDb: a G-protein-coupled receptor database. *Receptors & channels*, *2*(1), 1–7.

Krettler, C., Reinhart, C., & Bevans, C. G. (2013). Expression of GPCRs in Pichia pastoris for structural studies. *Methods in enzymology*, *520*, 1–29.

Kristiansen, K. (2004). Molecular mechanisms of ligand binding, signaling, and regulation within the superfamily of G-protein-coupled receptors: molecular modeling and mutagenesis approaches to receptor structure and function. *Pharmacology & therapeutics*, *103*(1), 21–80.

Kumar, D. (2011). *Protein Refolding / Renaturation*. *Comprehensive Biotechnology* (Second Edi., Vol. 1, pp. 765–784).

Lagerström, M. C., & Schiöth, H. B. (2008). Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nature reviews. Drug discovery*, *7*(4), 339–57.

Lane, J. R., Abdul-Ridha, A., & Canals, M. (2013). Regulation of G protein-coupled receptors by allosteric ligands. *ACS chemical neuroscience*, *4*(4), 527–34.

Ma, Y., Kubicek, J., & Labahn, J. (2013). Expression and purification of functional human mu opioid receptor from E.coli. *PloS one*, *8*(2), e56500.

Michalke, K., Huyghe, C., Lichière, J., Gravière, M.-E., Siponen, M., Sciara, G., Lepaul, I., et al. (2010). Mammalian G protein-coupled receptor expression in Escherichia coli: II. Refolding and biophysical characterization of mouse cannabinoid receptor 1 and human parathyroid hormone receptor 1. *Analytical biochemistry*, *401*(1), 74–80.

Millar, R. P., & Newton, C. L. (2010). The year in G protein-coupled receptor research. *Molecular endocrinology (Baltimore, Md.)*, *24*(1), 261–74.

Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., et al. (2000). Crystal structure of rhodopsin: A G protein-coupled receptor. *Science (New York, N.Y.)*, *289*(5480), 739–45.

Park, S. H., Casagrande, F., Chu, M., Maier, K., Kiefer, H., & Opella, S. J. (2012). Optimization of purification and refolding of the human chemokine receptor CXCR1 improves the stability of proteoliposomes for structure determination. *Biochimica et biophysica acta*, *1818*(3), 584–91.

Park, S. H., Das, B. B., Casagrande, F., Tian, Y., Nothnagel, H. J., Chu, M., Kiefer, H., et al. (2012). Structure of the chemokine receptor CXCR1 in phospholipid bilayers. *Nature*, *491*(7426), 779–83.

Pervushin, K., Riek, R., Wider, G., & Wüthrich, K. (1997). Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(23), 12366–71.

Popot, J.-L., Althoff, T., Bagnard, D., Banères, J.-L., Bazzacco, P., Billon-Denis, E., Catoire, L. J., et al. (2011). Amphipols from A to Z. *Annual review of biophysics*, *40*, 379–408.

Qoronfleh, M. W., Hesterberg, L. K., & Seefeldt, M. B. (2007). Confronting high-throughput protein refolding using high pressure and solution screens. *Protein expression and purification*, *55*(2), 209–24.

Rosenbaum, D. M., Rasmussen, S. G. F., & Kobilka, B. K. (2009). The structure and function of G-protein-coupled receptors. *Nature*, *459*(7245), 356–63.

Saxena, K., Dutta, A., Klein-Seetharaman, J., & Schwalbe, H. (2012). Isotope labeling in insect cells. *Methods in molecular biology (Clifton, N.J.)*, *831*, 37–54.

Sharapova, O. a, Yurkova, M. S., Laurinavichyute, D. K., Andronova, S. M., Fedorov, A. N., Severin, S. E., & Severin, E. S. (2011). Efficient refolding of a hydrophobic protein with multiple S-S bonds by on-resin immobilized metal affinity chromatography. *Journal of chromatography. A*, *1218*(31), 5115–9.

Singh, S. M., Sharma, A., Upadhyay, A. K., Singh, A., Garg, L. C., & Panda, A. K. (2012). Solubilization of inclusion body proteins using n-propanol and its refolding into bioactive form. *Protein expression and purification*, *81*(1), 75–82.

Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J., & Li, Q. (2009). Practical protocols for production of very high yields of recombinant proteins using Escherichia coli. *Protein science : a publication of the Protein Society*, *18*(5), 936–48.

Tsumoto, K., Ejima, D., Kumagai, I., & Arakawa, T. (2003). Practical considerations in refolding proteins from inclusion bodies. *Protein Expression and Purification*, *28*(1), 1–8.

Vallejo, L. F., & Rinas, U. (2004). Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microbial cell factories*, *3*(1), 11.

Venkatakrishnan, A. J., Deupi, X., Lebon, G., Tate, C. G., Schertler, G. F., & Babu, M. M. (2013). Molecular signatures of G-protein-coupled receptors. *Nature*, *494*(7436), 185–94.

Wang, C., Wang, L., & Geng, X. (2009). Optimization of refolding with simultaneous purification of recombinant human granulocyte colony-stimulating factor from Escherichia coli by immobilized metal ion affinity chromatography. *Biochemical Engineering Journal*, *43*(2), 197–202.

Warne, T., Moukhametzianov, R., Baker, J. G., Nehmé, R., Edwards, P. C., Leslie, A. G. W., Schertler, G. F. X., et al. (2011). The structural basis for agonist and partial agonist action on a β(1)-adrenergic receptor. *Nature*, *469*(7329), 241–4.

Warne, T., Serrano-Vega, M. J., Baker, J. G., Moukhametzianov, R., Edwards, P. C., Henderson, R., Leslie, A. G. W., et al. (2008). Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature*, *454*(7203), 486–91.

Weiss, H. M., & Grisshammer, R. (2002). Purification and characterization of the human adenosine A(2a) receptor functionally expressed in Escherichia coli. *European journal of biochemistry / FEBS*, *269*(1), 82–92.

White, J. F., Trinh, L. B., Shiloach, J., & Grisshammer, R. (2004). Automated large-scale purification of a G protein-coupled receptor for neurotensin. *FEBS letters*, *564*(3), 289–93.

Zhang, X., & Eggert, U. S. (2013). Non-traditional roles of G protein-coupled receptors in basic cell biology. *Molecular bioSystems*, *9*(4), 586–95.