DISS. ETH NO. 24544

Engineering, Large Scale Expression, and NMR Spectroscopy of the Human Beta1-Adrenergic Receptor

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

Jakob Aastrup Jørgensen

Cand. Scient. Biochemistry, University of Copenhagen

born on 03.12.1981

citizen of Denmark

accepted on the recommendation of
Prof. Dr. Gerhard Wider
Prof. Dr. Gebhard Schertler
Dr. Dmitry Veprintsev

2017

Engineering, Large Scale Expression and NMR-spectroscopy of the Human Beta1-Adrenergic Receptor

Jakob Aastrup Jørgensen

"...det menneske, der slider sig op for guld og ære, eller hvad det nu er, af andre grunde end lidenskab og indre trang, han er og bliver dog en nar."

"...the man who, solely from regard to the opinion of others, and without any wish or necessity of his own, toils after gold, honour, or any other phantom, is no better than a fool."

"...ein Mensch, der um anderer willen, ohne daß es seine eigene Leidenschaft, sein eigenes Bedürfnis ist, sich um Geld oder Ehre oder sonst was abarbeitet, ist immer ein Tor."

Die Leiden des jungen Werther, Den 20. Julius.

Johann Wolfgang von Goethe

Contents

Summary

G-Protein coupled receptors (GPCRs) are membrane proteins, conveying an extracellular stimulus over the plasma membrane with an intracellular response as a consequence. These receptors are found in all eukaryotic organisms, and particularly in higher eukaryotes they play crucial roles, for example, regulating heart rate, transmitting signals in neural synapses in the brain, or enabling vision. Being the targets of an estimated 30% of all pharmaceutical drugs, their importance can hardly be overstated.

Only recently, 3-dimensional structures have become available, facilitating structure-based drug design. To a large extent, however, information on conformational changes when a GPCR is activated or antagonized is lacking. In addition, dynamical characterization is also absent. An ideal tool to study this, conformational dynamics of proteins, is nuclear magnetic resonance (NMR) spectroscopy.

Here, an engineered and thermo-stabilized version of the human beta1-adrenergic receptor has been optimized for structural studies, improving the yield of the receptor in eukaryotic expression systems, while retaining thermo-stability, making NMR studies more feasible in terms of labor and finances. This was done through testing an array of modifications on the protein level. In addition, expression conditions using the baculovirus-insect cell system have been optimized to yield protein of higher quality, an improvement which is most likely generally applicable to membrane protein over-expression.

Likewise, an optimized procedure for fast purification of receptor from large-scale expression was set up. The purified protein exhibited biophysical characteristics, resembling those of the avian model GPCR, the turkey beta1-adrenergic receptor, which is known to readily crystallize and is amenable to NMR studies.

Finally, NMR experiments were carried out with the human beta1-adrenergic receptor, resulting in 2-dimensional [^{1}H,^{15}N]-TROSY spectra, of very similar appearance to those previously obtained for the turkey beta1-adrenergic receptor. Most interestingly, a spectrum with the clinically important beta-blocker S-propranolol, which has also been described as a biased agonist, was obtained, which showed that this ligand induces a unique state of the receptor – an effect not seen for the previously tested antagonist on the avian homolog.

This thesis demonstrates that NMR studies of an extremely challenging GPCR, are indeed possible, but also clearly shows that it is a formidable task that is still in need of additional optimization before the full potential of NMR spectroscopy on this type of protein can be unleashed.

Zusammenfassung

G-Protein-gekoppelte Rezeptoren (GPCRs) sind membranständige Proteine, die einen extrazellulären Stimulus durch die Plasma Membran in ein intrazelluläres Signal überträgt. Diese Rezeptoren findet man in allen Eukaryonten, und sie spielen insbesondere in höher entwickelten Eukaryonten eine entscheidende Rolle in der Regulierung der Herzfrequenz, bei der Übertragung von Signalen an Neuronalen Synapsen des Gehirns, oder der Wahrnehmung von Licht im Auge. Ihre Wichtigkeit kann schwer unterschätzt werden, wenn man bedenkt, dass sie das Ziel von ungefähr 30 Prozent aller Medikamente sind.

Erst kürzlich wurde die 3-dimensionale Struktur der GPCRs publiziert, was die Struktur-basierte Entwicklung von Medikamenten befördert hat. In weiten Teilen fehlen jedoch noch Informationen zur Konformationsänderung bei Aktivierung oder Antagonisierung der GPCRs. Zusätzlich kennt man die dynamischen Karakterzüge nicht. Ein herausragendes Werkzeug zum Studium konformationaler Dynamiken ist die Kernspinresonanz (NMR) Spektroskopie.

In dieser Doktorarbeit wurde eine thermo-stabile Version des menschlichen Beta1-Adrenergen Rezeptor für die Strukturanalyse optimiert, durch Verbesserung der Ausbeute in einem eukaryontischen Expressionssystems, während die Thermo-stabilität beibehalten wurde, was zukünftig den Arbeitsaufwand und die Kosten von NMR studien reduziert. Hierzu, testeten wir eine Reihe von Modifikationen auf Protein Ebene. Zusätzlich, konnten die Expressionsbedingungen durch den Gebrauch des Baculovirus-Insekten Zell Systems so optimiert werden, dass die Ausbeute hochwertiger Proteine erhöht werden konnte, was von allgemeiner Bedeutung für die Überexpression von Membran Proteinen ist.

Damit einher ging die Optimisierung der Aufreinigung des Rezeptors aus Expressionskulturen in grossem Umfang. Das aufgereinigte Protein zeigte biophysikalische Eigenschaften, die jenen von Vogel GPCR Modellen entsprichen - der Truthahn Beta1-Adrenerge Rezeptor ist für seine einfache Kristallisierbarkeit bekannt und damit besonders gut für NMR Studien geeignet.

Schliesslich wurden NMR Experimente mit dem menschlichen Beta1-Adrenergen Rezeptor durchgeführt, bei denen die 2-dimensionalen [^{1}H,^{15}N]-TROSY Spectren eine grosse Ähnlichkeit mit dem Truthahn Beta1-Adrenergen Rezeptor aufwiesen. Besonders spannend war hierbei, dass das Spektrum welches mit einem klinisch relevanten Beta-Blocker S-Propranolol, welcher auch als funktinal selektiver Agonist bekannt ist, gemessen wurde, zeigte, dass dieser Ligand einen einzigartigen Zustand des Rezeptors auslöste, den man nicht mit Antagonisten im vorher getesteten Vogel Homolog ausmachen konnte.

Diese Doktorarbeit demonstriert, dass NMR Spektroskopie von GPCRs extrem kompliziert, aber möglich ist. Das anspruchsvolle Ziel zu erreichen, bedarf jedoch weiterer Optimisierung, bevor das volle Potentzial der NMR Spektroskopie von Membran Proteinen entfesselt werden kann.

Abbreviations

[3H]DHA	Tritiated DiHydroAlprenolol
AcMNPV	*Autographa california* Multiple Nucleo Polyhedrosis Virus
ALAC	Alprenolol Ligand Affiinity Chromatography
CPM	7-Diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin
DDM	DoDecylMaltoside
DM	DecylMaltoside
DMEM	Dulbecco's Modified Eagle's Medium
ER	Endo-plasmic Reticulum
P.ph	polyhedrin promoter of AcMNPV
P.39K	39K promoter of AcMNPV
EGFP	Enhanced Green Fluorescent Protein
EYFP	Enhanced Yellow Fluorescent Protein
FBS	Fetal Bovine Serum
FSEC	Fluorescence Size Exclusion Chromatography
$gp67^{SS}$	gp67 Signal sequence
$gp67L^{SS}$	gp67 Signal sequence, long version
GPCR	G-Protein Coupled Receptor
hADRB1	human beta1-adrenergic receptor
hADRB2	human beta2-adrenergic receptor
HEK293	Human Embryonic Kidney cells clone 293
High five	*Tricuplusia ni* insect cell line
HPLC	High Pressure Liquid Chromatography
IMAC	Immobilized Metal Affinity Chromatography
n. a.	Not applicable
NMR	Nuclear Magnetic Resonance
p0	Initial baculovirus stock from transfection

p1	First passage of a baculovirus stock
p2	Second passage of a baculovirus stock
PBS	Phosphate Buffered Saline
PEI	Poly Ethylene Imine
Sf9	*Spodoptera frugiperda* insect cell line 9
Sf21	*Spodoptera frugiperda* insect cell line 21
SDS	Sodium Dodecyl Sulfate
Sol	Solubilized
T4L	T4 Lysozyme
tADRB1	turkey beta1-adrenergic receptor
TBS	Tris Buffered Saline
TM	TransMembrane
Tot	Total
TROSY	Transverse Relaxation Optimized SpectroscopY
WGA	Wheat Germ Agglutinin
WT	Wild-Type

1. Introduction

7TM and G-Protein Coupled Receptors

G-Protein Coupled Receptors (GPCRs) constitute one of the largest families of protein encoded in the human genome (Lander et al. 2001), and are integral membrane proteins. This family is again part of an even bigger family, the 7TM super family of proteins that are present in all three domains of life: Eubacteria, Archaea and Eukarya. The 7TM family shares a common structural architecture, giving it its name, which is made up seven trans-membrane alpha-helical segments arranged in bundle roughly resembling a cylinder. Members of this family include: Bacteriorhodopsin, a light driven proton pump for which the prototype is found in the extremophile *Halobacterium salinarium* from the domain Archaea; Proteorhodopsin, also a light driven proton pump found in a marine *γ-proteobacterium,* but from the domain Eubacteria (Beja et al. 2000); Bovine rhodopsin, from *Bos Taurus* from the domain Eukarya, which is one of the light activated receptors enabling vision in the kingdom Animalia. It should be noted that there is doubt as to whether the prokaryotic versions really are related to the GPCRs through a common ancestor, or the similarity is a result of convergent evolution (Beja et al. 2000).

Rhodopsin can be seen as the prototypical and probably one of the most well studied G-protein coupled receptors. It is found in abundance in the retina of the visual organs of animals. However, Rhodopsin is but one of many GPCRs in humans and other animals. In one study more than 800 GPCRs in humans were identified, including the olfactory receptors (i.e. related to the sense of smell), and more than 300 receptors as non-olfactory GPCRs, capable of detecting a wide variety of stimuli including amino-acids, small molecule hormones, lipids, peptides and proteins as their ligands and even, as mentioned earlier, light (Fredriksson et al. 2003, Lagerstrom and Schioth 2008). These stimuli are generally referred to as 1st messengers.

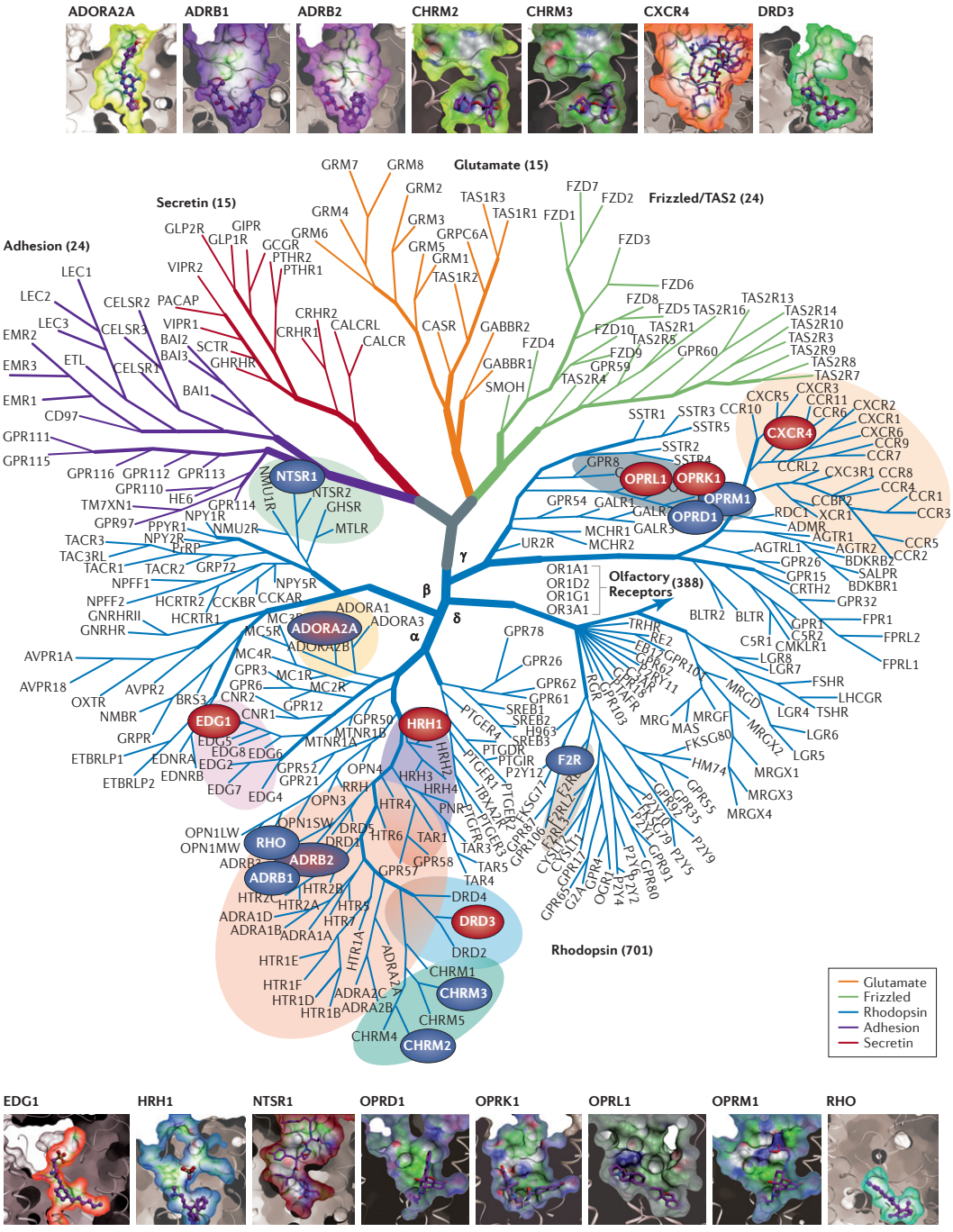


Figure 1. Overview of the human G-Protein Coupled Receptors

Phylogenetic tree of the human G-Protein Coupled Receptors including available structures as of 2013. Binding ligand pocket is shown for select GPCR with available structural information. Figure adapted from Stevens et al. (2013) (previously adapting from Fredriksson et al. 2003)

In addition to rhodopsin, these non-olfactory GPCRs include for example the adrenergic receptors, the opioid receptors, and the majority of the serotonin receptors, all of which have important physiological functions. In

Table 1. Overview of the different GPCR families.

Class	GRAFS Family	Note	Human Receptors
A	Rhodopsin	Includes ∼400 olfactory receptors	282 (+ 400 olfactory)
B	Secretin	Also referred to as B1	15
	Adhesion	Also referred to as B2	33
C		Includes Taste type 1 receptors	22
F	Frizzled		11
O	Taste 2	Re-classified as own family, although originally grouped with Frizzled	24

Table adapted from (Isberg et al. 2015)

turn, these are important pharmacological drug targets, ranging from cardiac hypertension over analgesia to treating mental health conditions such as depression. Thus, the GPCR family is a class of proteins that play key physiological roles. Therefore the understanding of these receptors has a great impact on the knowledge of multicellular organisms, but equally so on society and its economy. It has been estimated that 30%-40% of all drugs are targeting these receptors (Stevens et al. 2013). An overview of the human GPCRs can be seen in the phylogenetic tree in Figure 1.

Based on sequence similarity, the GPCR superfamily has been divided into groups in different ways. One type is the GRAFS division and naming: Glutamate-like (G), Rhodopsin-like (R), Adhesion-like (A), Frizzled/ Smoothened (F) and Secretin-Like (S) (Fredriksson et al. 2003). Another similar division uses the letters A through F (Lagerstrom and Schioth 2008). A comparison between the two groupings can be seen in Table 1. It is noted that not all classes are represented in the human genome.

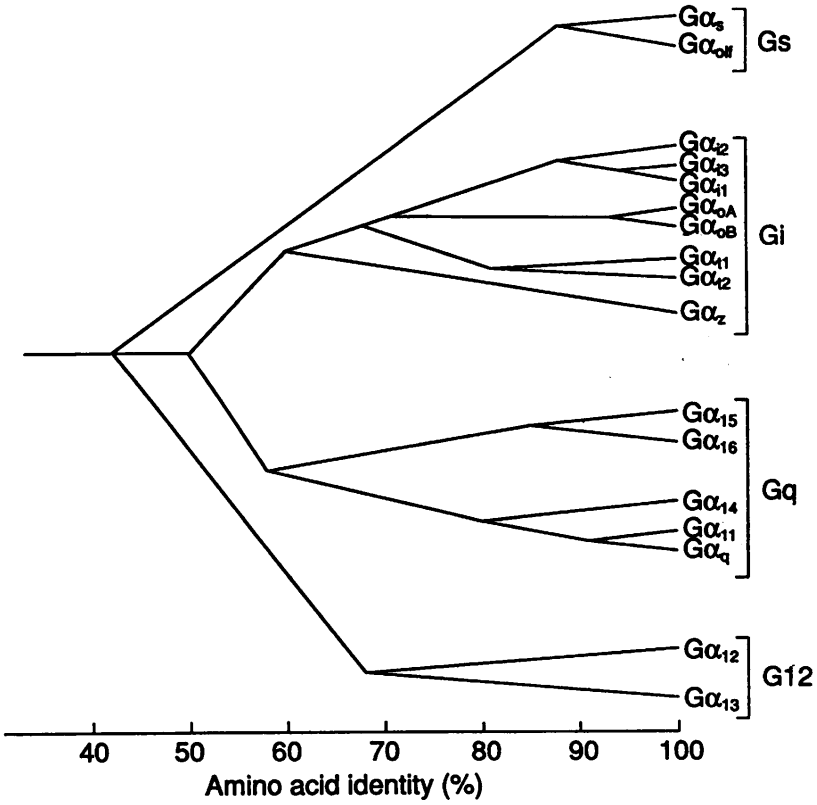


Figure 2. Diversity among mammalian G-proteins.

Phylogenetic tree of mammalian G-proteins, depicting the four classes of G-protein . Figure adapted from (Simon et al. 1991)

G-protein coupling and downstream signaling initiated by GPCRs

As the name suggests, the GPCRs are capable of interacting with a cytosolic but membrane tethered complex called a G-protein: a hetero-trimeric protein consisting of an alpha-, beta-, and gamma-subunit. The alpha subunit binds GDP or GTP, hence the name derived from the nucleotide guanosine, commonly abbreviated to G. A GPCR can be activated by a ligand leading to binding to a given G-protein, which upon binding will exchange a bound GDP for a GTP, in turn leading to downstream signaling events in the signaling cascade and eventually a final cellular response. Replacement of GDP by GTP makes the protein complex active, but eventually the active G-protein will hydrolyze GTP to GDP, and render itself inactive again.

Although there are several different versions of each of the G-protein subunits, and these are present in different combinations *in vivo*, there are vastly many more GPCRs than simply the individual subunits of the G-proteins. As such, the set of G-proteins can be seen as a generic set of signaling proteins serving as signaling mediators expressed in different ways in various tissues with the GPCRs acting as their tissue specific extension reaching from the cytosol across the plasma membrane to the extra cellular environment (Wettschureck and Offermanns 2005). An overview of the mammalian G-protein classes is seen in Figure 2.

The Beta1-Adrenergic Receptors

The adrenergic receptors belong to the Rhodopsin-like family of GPCRs (class A) – the most populous group of GPCRs. Named for their binding and of activation by adrenaline and the closely related nor-adrenaline, they are actually not each other's most closely related relatives sequence-wise (Fredriksson et al. 2003). The beta–adrenergic receptors in particular can be considered prototypical GPCRs, in that they were discover and studied before the concept of a G-protein coupled receptor was coined. In fact the beta-adrenergic receptors were referred to as being adenylyl-cyclase coupled (see for example Mukherjee et al. 1975).

The mammalian beta1-adrenergic receptors have been shown to be predominantly expressed in heart, lung, cerebral cortex and pineal gland tissues. (Frielle et al. 1987). The natural ligands of the receptor are the catechol-amines adrenaline and noradrenaline. Upon release of these compounds into the blood stream, amongst other responses, the heart rate increases, which the beta1-adrenergic receptor is responsible for, and at the molecular level the receptor signals primarily via the Gs G-protein which in turn activates adenylyl cyclase (reviewed by Brueckner et al. 2013).

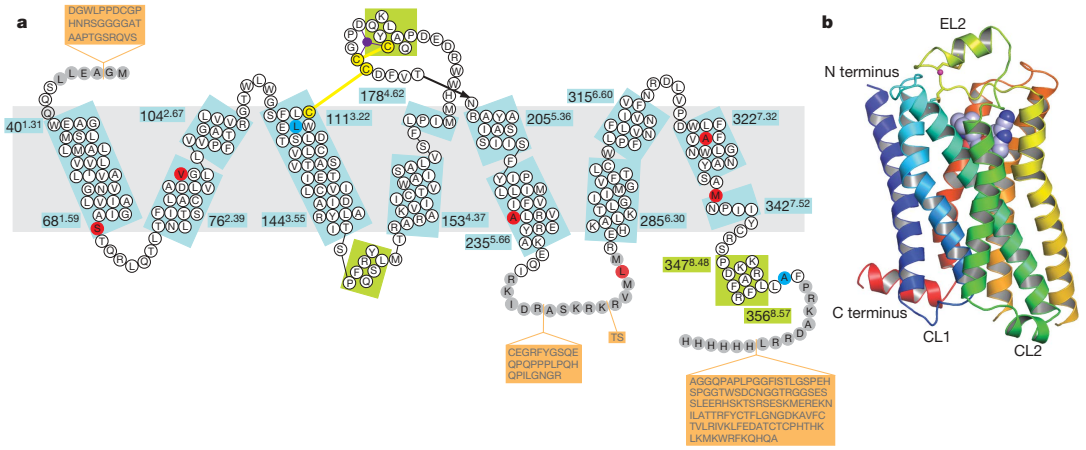


Figure 3. Structure of the turkey beta1-adrenergic receptor.

Outline of the structure of the turkey beta1-adrenergic receptor – the second non-rhodopsin GPCR to have its structure solved (Warne et al. 2008). (**a**) Snake plot showing the primary amino-acid sequence and topology in the plasma membrane(grey bar) including the various modification of this receptor that made crystallization possible. (**b**) 3D cartoon of the crystal structure in complex with the antagonist cyanopindolol. Figure adapted from Warne et al. (2008).

Structural biology of GPCRs and the missing dynamical link

As mentioned above, the adrenergic receptors are involved in key physiological processes, which apply to many GPCRs, and therefore it is easy to realize that these receptor also constitute an important drug development. In fact it has been estimated that 30-40% of all drugs target these receptors (Stevens et al. 2013). The increasing number of GPCR structures provides an understanding of how existing drugs interact with their target receptor, and likewise, these structures facilitate the identification of new drugs (Cooke et al. 2015).

The turkey beta1-adrenergic receptor was in fact the second non-rhodopsin GPCR to have its structure solved by X-ray crystallography (Warne et al. 2008). The structure of this receptor can be seen in Figure 3. While such a structure presents invaluable information, it only offers a snapshot of the ligand binding process. The missing link is the dynamical aspects of ligand binding, receptor activation and interaction with the downstream signaling partners. Such information can be provided by Nuclear

Magnetic Resonance spectroscopy. Using this technique it is possible to gain knowledge about how the receptor moves and responds to ligands, as a population of individual molecules, rather than a frozen state observed in a crystal.

Organization of the experimental section of this thesis

The experimental section of this thesis is organized in four separate chapters, 2 through 5, according to the individual subject of each chapter: Engineering the ultra-stabilized human beta1-adrenergic receptor (Chapter 2); Engineering of the Baculovirus-insect cells expression system (Chapter 3); Establishment of a new purification method of the human beta1-adrenergic receptor (Chapter 4); NMR studies of the human beta1-adrenergic receptor. The organization is seen as a diagram in Figure 4.

The sequence of these chapters have been presented in a way such that a natural progression from the starting point, a GPCR with an very low yield, towards, the end point of performing NMR experiments is apparent. However, it is especially stressed that much of the work from chapters 2 through 4, was essentially done in parallel. As such findings from chapter 4 are for example utilized in chapter 2. On the other hand, the order of chapter 2 and chapter 3 might be switched altogether, from an experimental and intellectual viewpoint, but is kept in the chosen way since a general introduction to the receptor and the particular way it has been engineered is given in this chapter. Likewise, an optimized version of the receptor (identified in Chapter 2) has been used in chapter 3, making this order of chapters more natural. In this context it should also be mentioned that initial NMR experiments were done with constructs made without some of the knowledge obtained in Chapter 3. However, in the final NMR experiment done with a very challenging version of the receptor, all the knowledge has been utilized, and in this way this particular protein concludes all the work done here.

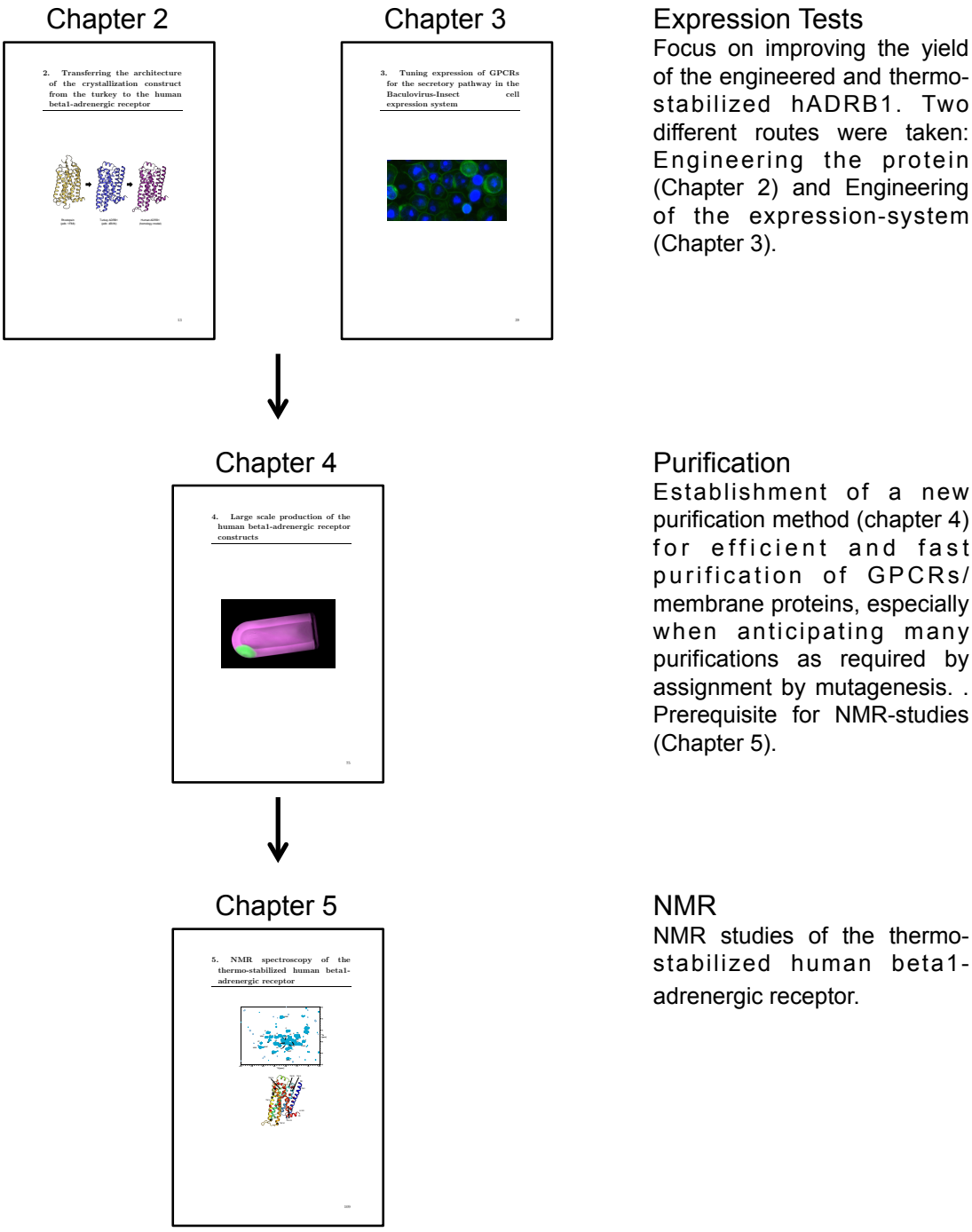
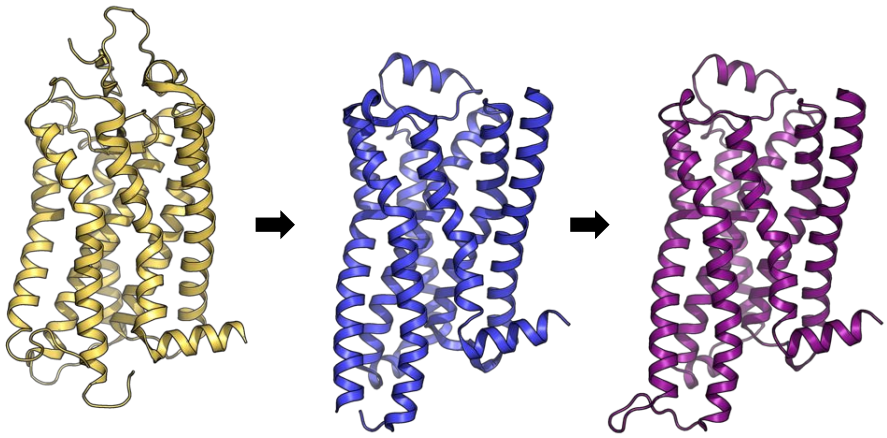


Figure 4. Overview of the organization of the thesis.

A graphical overview of how this thesis has been organized is seen on the right hand side, while a description of each section is found on the left hand side.

2. Transferring the architecture of the crystallization construct from the turkey to the human Beta1-Adrenergic Receptor



Rhodopsin
(pdb: 1F88)

Turkey ADRB1
(pdb: 4BVN)

Human ADRB1
(homology model)

Chapter 2 frontpage. "Rhodopsinization" of adrenergic receptors.

The illustration is showing how the turkey ADRB1 roughly resembles the structure of rhodopsin, and how the anticipated structure of the human ADRB1 (homology model) again resembles the former two. However, the key to the "rhodopsinization" lies in what is not seen with the naked eye, as described in the chapter: improved biophysical characteristics and removal of flexible region of the receptors.

Introduction

G-Protein Coupled Receptors (GPCRs), and in general Membrane proteins, constitute only a small fraction of all available protein structures obtained by X-ray crystallography and Nuclear Magnetic Resonance (NMR) spectroscopy. Therefore, the field of GPCR structural biology remains an area in which large improvements are yet to be made.

The first GPCR to be crystallized and subsequently having its X-ray structure solved was the bovine rhodopsin purified from native sources – cow and bull retinas (Palczewski et al. 2000). The ample availability in and of the retinas and easy purification from Rod outer segments made it the obvious first target for pursuing crystallographic studies of a GPCR.

It would be another 7 years until a thermo-stabilized mutant of rhodopsin had its structure solved. Here the receptor was produced and purified from Chinese hamster ovary cells, making it the first record of a recombinantly expressed GPCR to be crystallized (Standfuss et al. 2007).

Almost immediately after the engineered rhodopsin structure came out, the structure of the human beta2-adrenergic Receptor (hADRB2) was published (Cherezov et al. 2007, Rasmussen et al. 2007). The year after, the structure of the turkey beta1-adrenergic receptor (tADRB1) was solved (Warne et al. 2008).

Facing similar problems, the successful crystallizations of the hADRB2 and the tADRB1 were achieved via different strategies. The inherent instability of membrane proteins, regarding the protein overall integrity, once extracted from its natural environment, the plasma membrane, was a primary obstacle, necessitating use of very mild detergent, which could bury hydrophilic sites, which are the most probable candidates for forming crystal contacts. However conformational heterogeneity of the active protein was also a concern (reviewed by Bill et al. 2011).

In the case of the hADRB2, the problems were dealt with in two different but qualitatively similar ways, one of which was to isolate the receptor, form a complex with an antibody Fab fragment, and transfer the complex to DMPC-CHAPSO bicelles for conventional crystallization. Here the antibody fragment was intended to and did provide additional crystal contacts, while also making the receptor more conformationally homogenous (Rasmussen et al. 2007).

The other direction that was taken was to exchange the third intracellular loop with T4 lysozyme (T4L), a protein known to readily crystalize, and as such serve as a crystallization facilitator, with the T4L protruding from the receptor, while using lipidic cupic phase for crystallization (LCP) (Cherezov et al. 2007, Rosenbaum et al. 2007). This method gave the best crystals, resulting in a larger part of the receptor being resolved crystallographically and to higher resolution.

The two chosen ways share a common feature, which is to keep the receptor in a relatively mild environment, and artificially provide the means for efficiently forming crystal contacts, here via either Fab fragment or the T4L insertion. As such, by means of including crystallization partner (the aforementioned T4L or Fab) the entity to be crystallized had an expanded size with more hydrophilic surface available for forming crystal contacts.

The approach taken for the tADRB1 receptor went in the opposite direction. Originally chosen for its high functional expression, the receptor had been extensively modified mainly involving deletions and truncations of regions that were predicted to be very flexible and therefore potentially hampering successful crystallization (Warne et al. 2003).

A realization was here again that mild detergent is not optimal for crystallization, due to the large micelle size compared to the size of the GPCR. Thus, identifying mutations that improve thermo-stability was a road to making a receptor, which would be able to tolerate harsher detergent. Such

mutations were found in by means of an alanine scanning of nearly the entire receptor, which identified several stabilizing mutations, by screening for thermo-stability. In this set of mutations, six were chosen, and the combination, called m23, gave a much more thermo- and detergent-stable receptor, with a decreased affinity for agonists, suggesting a potential conformational stabilization (Serrano-Vega et al. 2008). In addition, three more stabilizing mutations were identified, by performing a leucine scanning on the same m23 construct, which identified two more thermo-stabilizing mutations, with an even more stable construct as a consequence (Miller and Tate 2011).

In this way, instead of keeping the membrane mimetic (e.g micelle, bicelle) at a relatively large size in order for it to be as nondenaturing as possible, while making the object to be crystallized bigger by introducing the aforementioned crystallization partner for improved crystallizability, the membrane mimetic, here a micelle, could now be made smaller owing to the increased detergent- and thermo-stability. This would mean that more of the extracellular parts of the GPCR would be exposed making them available for forming crystal contacts. Consequently, the structure of the tADRB1 (engineered as briefly described in paragraph immediately below) was solved from crystals conventionally grown using the relatively harsh and short-chained detergent octyl-thio-glucoside (Warne et al. 2008).

As previously stated (Serrano-Vega et al. 2008), the thermo-stability of the tADRB1 approached that of rhodopsin, while in addition it had been modified in a way that made it very similar regarding overall shape – flexible parts were systematically removed (Warne et al. 2009). As such the construct is mimicking the properties of wild-type bovine rhodopsin, so far the only GPCR that readily crystallizes without modifications, including unaltered N- and C- termini, and all extra intra- and extra-cellular loops intact.

While the tADRB1 gave the second structure of a non-Rhodopsin GPCR, it is still questionable whether this structure provides the appropriate information to correctly explain the pharmacology of the beta1-adrenergic receptors in mammals, and particularly in Homo sapiens. In fact, almost since the discovery of the ligand binding entities in a select adrenergic system, the beta-adrenergic receptors, it has been known that it should be approached with caution, as stated by Minneman et al. (1980). This is so, due to the observed differences in the pharmacology between the turkey ADRB1 and the two studied mammalian beta-adrenergic receptors, an ADRB1 and an ADRB2, including different selectivity for a variety of beta-blockers. The tADRB1 is found in between the two afore mentioned mammalian receptors pharmacologically, but closer to the mammalian ADRB1, making it referred to as the turkey ADRB1 (tADRB1). It should be noted that the sequence homology between tADRB1 and the two human adrenergic receptors, hADRB1 and hADRB2, is 81 % and 59%, respectively.

Therefore, as necessitated by the different pharmacological profiles of the mammalian and the avian beta1-adrenergic receptor, and due to the relevance of the human beta1-adrenergic receptor in basic physiology research and drug development structural studies of the hADRB1 were initiated. In the study at hand, the overall construct architecture is being transferred from tADRB1 crystallization construct to the hADRB1 to obtain an analogous construct with similar properties regarding detergent and thermo-stability and overall shape.

Results

Construct design: The tADRB1.44-m23 crystallization construct as a template

As mentioned above the tADRB1 was modified extensively. This included a systematic series of N-terminal truncations, C-terminal truncations and deletions in the third intracellular loop. The criteria for selecting particular modifications were expression level and importantly, stability in detergent in purified form (Warne et al. 2003, Warne et al. 2009). The construct used here is the tADRB1.44-m23 (also referred to as the β1AR44-m23 elsewhere).

The N-terminus of the tADRB1 was truncated in three steps, and it was found that truncating down to A33$^{1.16}$ would give the most stable receptor, while going further would decrease stability in detergent. The super script is denoting the position of the GPCR amino-acid position using the Ballesteros-Weinstein convention (Ballesteros and Weinstein 1995), and the numbering has been extended to include residue beyond those in the trans-membrane segment for simplicity. Likewise the C-terminus had also been truncated such that 10 residues would remain in the C-terminal end of L311$^{8.58}$.

The third modification was finding an as stable and as small as possible version of the Intra Cellular Loop (ICL) 3, due to its relatively large size compared to that of rhodopsin, which was known to crystallize readily. Again an optimal modification was found having a certain size, while at least in some cases, reducing this further would lead to decreased stability (Warne et al. 2009). The position for the 28 amino-acid long deletion in the tADRB1.44-m23 construct was between C244$^{5.75}$ and R271$^{6.12}$, both residues included.

Using sequence alignment, and the significant degree of identity and homology between the turkey and the human ADRB1, it was reasonable to believe that the hADRB1 could be modified in a similar way to the above mentioned. Hence, the overall architecture of the tADRB1 construct was

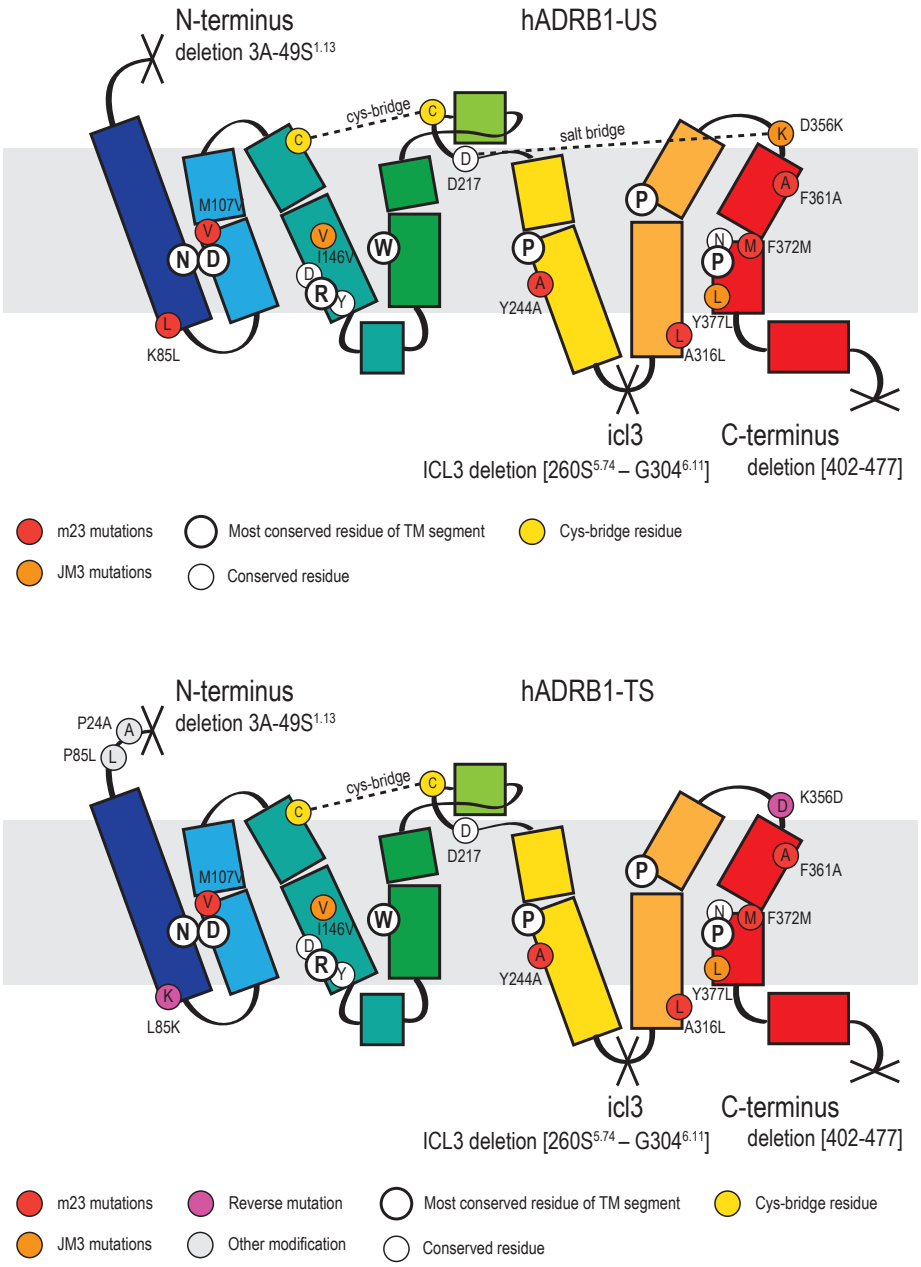


Figure 5. Overview of constructs

Cartoon showing the minimized and thermo-stabilized constructs of the hADRB1 receptor, hADRB1-US (**top panel**) and hADRB1-TS (**bottom panel**). Grey rectangles represent the cell membrane, while colored rectangles represent the trans-membrane segments. Positions of deletions/truncations are shown in the primary amino-acid sequence, while superscripts indicate Ballesteros-Weinstein number. Key residues are indicated in figures. An additional disulfide bridge in extra-cellular loop 3 is not shown in either figure.

transferred directly to the hADRB1. This means that the truncations and the deletions were made at the exact same distance in amino-acid residues from a

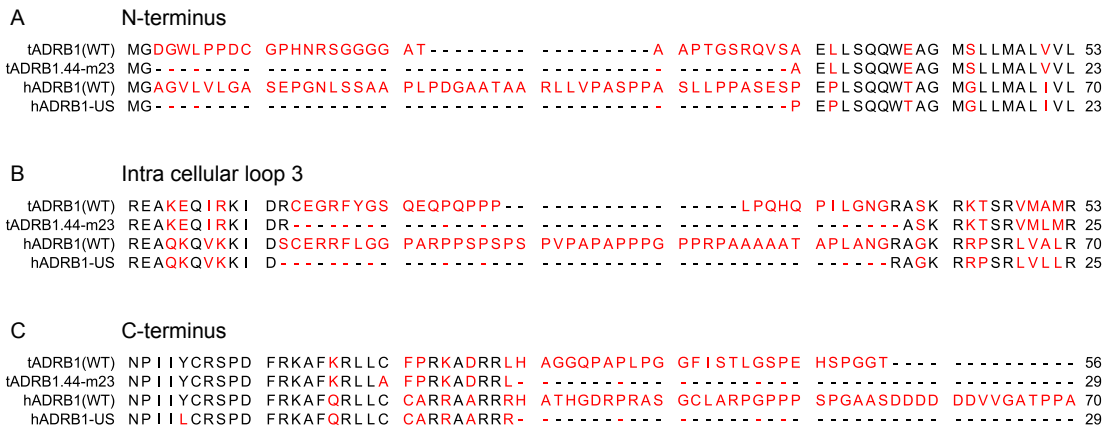


Figure 6. Transferred truncations and deletions from the tADRB1.44-m23 construct to the hADRB1-US construct.

Alignment of key regions in the tADRB1 and hADRB1 wild-type and crystallization constructs (tADRB1.44-m23 and hADRB1-US). In the alignment, gaps are then representing deleted residues, in either the tADRB1.44-m23 and the hADRB1-US, when compared to their respective wild-type receptors. (**A**) N-terminus, (**B**) the intracellular loop 3, and (**C**) C-terminus. Identical residues shown in black while non-identical residues are shown in red. Note that the entire C-terminus has not been shown in panel (**C**) of the two wild-type receptors.

relevant highly conserved region or residue, for example x.50 (the most conserved residue) in the relevant helix or helices, here using the Ballesteros-Weinstein to denote that conserved residue. The relevant regions, N-terminus, ICL3, and C-terminus, and truncations and deletions, can be seen Figure 5, while a detailed view of these regions is shown in Figure 6.

As previously shown, the m23 mutations in the tADRB1 (R68S, M90V, Y227A, A282L, F327A, and F338M; position of residue in wild-type tADRB1 sequence (Serrano-Vega et al. 2008)) could be transferred to the same positions in the hADRB1, however with a single change in the individual amino-acid identities (K85L, M107V, Y244A, A316L, F361A and F372M; position of residue in wild-type hADRB1), with and increase in thermo-stability as a consequence (Serrano-Vega and Tate 2009). With the additions of the three extra stabilizing mutations found later (I129V, D322K, Y343L; position of residue in wild-type tADRB1 sequence (Miller-Gallacher et al. 2014), n.b. a fourth mutation (D200E) was originally identified ((Miller and

Tate 2011)) but was not used), the tADRB1 was now referred to as being ultra-stabilized (Miller and Tate 2011). All nine mutations were added to the modified construct of the hADRB1. This construct was called hADRB1-US, with US denoting the ultra-stabilizing mutations, but also the truncations and deletions. The thermo-stabilizing mutations can also be seen in Figure 5.

hADRB1-US gives a highly thermo-stable and homogenous receptor preparation

The hADRB1-US construct exhibited a high thermo-stability as measured by thermal denaturation in CPM assay developed by Alexandrov et al. (2008), while analytical size exclusion chromatography also gave a sharp peak, much like what can be obtained for the tADRB1.44-m23 crystallization construct. However, a big problem was the very low yield of the receptor when expressed in the insect cell line Sf9 from the species *Spodoptera frugiperda,* using the baculovirus system as for the expression of the tADRB1.44-m23. Here the receptor was purified by nickel affinity purification followed by Alprenolol Ligand Affinity Chromatography, with a final yield in the order of approximately 75ug receptor per liter of insect cell culture (data not shown, Florian Brueckner, personal communication).

The low yield seriously hindered convenient production for crystallizations, but importantly, an even lower yield would be expected for NMR experiments, due to lower yield in specialized and less rich medium, as required for the necessary isotope labelling. Therefore, an increase in the yield of the receptor was of uttermost importance.

Improving expression by diverse mutations

It has previously been observed that the stabilizing mutations can cause a decrease in the functional yield of the hADRB1 receptor (Serrano-Vega and Tate 2009). Therefore, a simple screen was setup to determine whether any of the nine stabilizing mutations were causing a significant drop in the yield. At the same time, the thermo-stability was measured with scope of finding one or

more mutations that could be reversed in order to obtain a higher yield, while retaining a high thermo-stability and receptor integrity

The region immediately next to the membrane can play a critical role when membrane proteins are co-translationally incorporated into the membrane at the ER (Endoplasmic Reticulum). Especially charged residue are known to have an impact on the incorporation of the membrane-spanning segments into the membrane, especially negatively charged residues on the extracellular-side of trans-membrane helix 1 (Beltzer et al. 1991, Wallin and von Heijne 1995). By inspection of the N-terminal end of the construct it is seen that the tADRB1.44-M23 has an additional negatively charged residue, compared to the hADRB1-US construct, while also several prolines are present potentially making incorporation difficult due to their limited degree of freedom. On the other hand a charged residue, a glutamate, is present N-terminally of the point of truncation, distanced by only a few residues.

Therefore, several constructs were made, in which the very N-term in the hADRB1-US was exchanged with that of the tADRB1.44-m23 in a stepwise manner for every differing residue, such that more and more of the tADRB1.44-m23 N-term was exchanged, and also a single point mutation, T11E, was included. In addition, the N-terminus was slightly extended, in single amino acid steps, to include the aforementioned glutamate.

To easily perform these experiment, the strategy developed by Kawate and Gouaux (2006) was used, in that an Enhanced Green Fluorescent Protein (EGFP) fused to the C-terminus of ADRB1-US by a flexible linker. The EGFP was followed by a purification tag (initially the 1D4-tag, but in later constructs, the Twin-Strep was added in addition to the 1D4-tag), which would allow for detection of the receptor in an unpurified state, as well as purified in small amounts, due to the high sensitivity when using EGFP as a probe. For easy expression, the constructs were transiently transfected into HEK293 cells, assuming that these effects on the yield would be expression

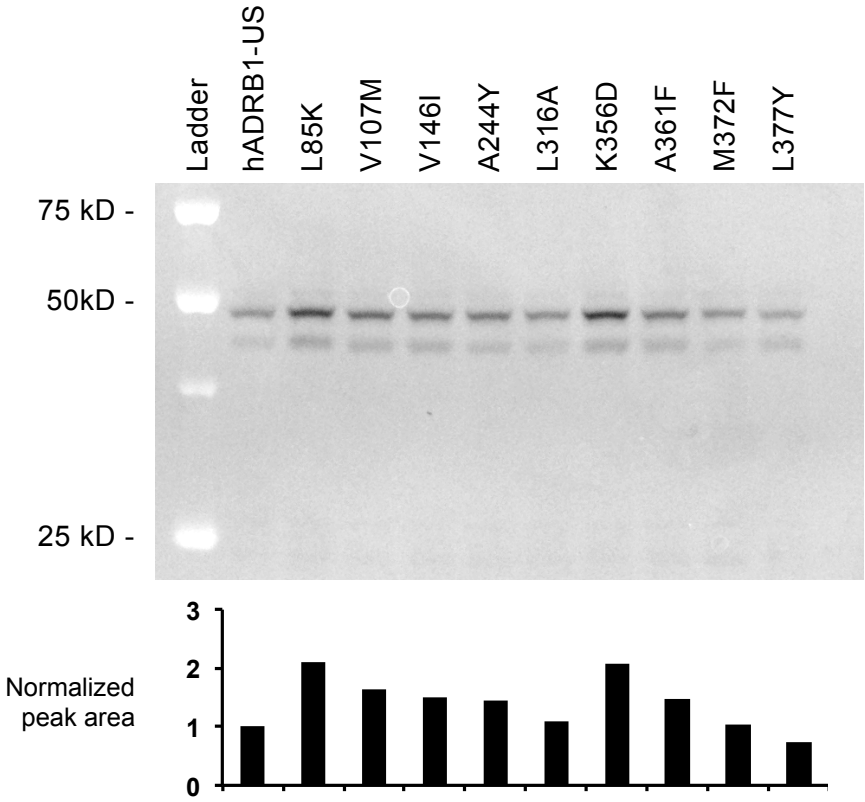


Figure 7. Screen by reversion of transferred stabilizing mutations for improved yield of the hADRB1-US construct.

In-gel fluorescence of reversed stabilizing mutations in the hADRB1-US-EGFP fusion constructs. Mutants expressed in HEK 293 cells by transient transfection, solubilized in 1%DDM and cleared by ultra-centrifugation. Individual reversion mutants are indicated above each lane, and the residue number corresponds to the numbering in the wild-type hADRB1 receptor. Below gel are the peak areas found as normalized to the peak area of hADRB1-US.

system independent, when considering insect cells and a mammalian cell type, due to the high similarity of the two expression systems, namely that both are cells derived from higher eukaryotic organisms, and thus have very similar translation machinery.

Initially all 18 constructs were evaluated by in-gel fluorescence of solubilized receptor in cleared lysate. The main band is migrating at around 50kD, a little shorter than expected which is common for protein with a high fraction of hydrophobic alpha helices. Sometimes a slightly faster migrating band is seen, which disappears upon purification. Likewise, free EGFP can

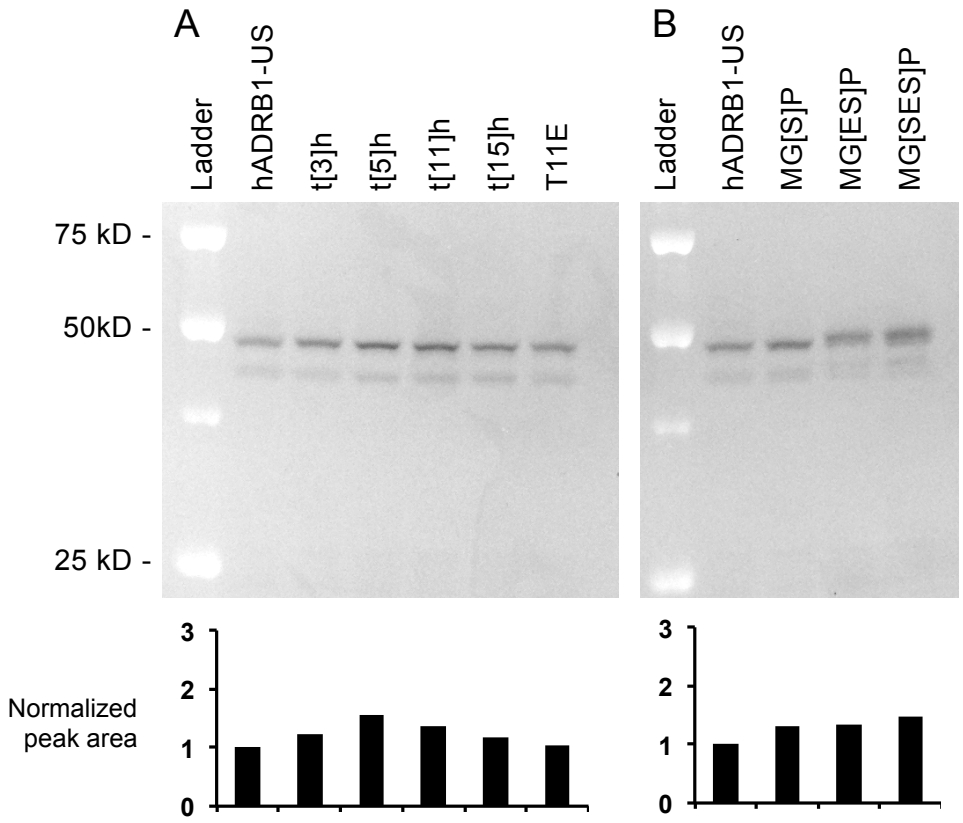


Figure 8. Screen of N-terminal modifications of improved yield in the hADRB1-US construct.

In-gel fluorescence of N-terminal modifications of the hADRB1-US-EGFP fusion constructs. Mutants expressed in HEK 293 cells by transient transfection, solubilized in 1%DDM and cleared by ultra-centrifugation. (**A**) Exchange of N-term with that of tADRB1.44-m23. Number in bracket indicates the last residue of the tADRB1.44-m23 sequence, T11E is a single point mutation. (**B**) N-terminal extension of hADRB1-US construct. Letter(s) in individual brackets indicate what residue(s) was/were inserted in this position. Below gel are the peak areas found as normalized to the peak area of hADRB1-US.

also sometimes be seen in the gels, which is interpreted as it being a product of digestion by endogenous proteases.

It is clearly seen in Figure 7, showing the screen of reversion mutations, that two of them stand out regarding increased yield. The first is L85K, where K85L is a mutation belonging to the original m23 mutations. The second is K356D, the reversion of D356K, which is expected to form the introduced salt bridge together with D217.

In the experiments exploring the modifications of the N-terminus, several potential candidates for improving the yield were identified. Figure 8 (A)

shows that the two first steps exchanging the N-terminus with that of the turkey construct (see legend of Figure 8 for details of the constructs and corresponding names), lead to a slightly increased amount of solubilized receptor, whereas exchanging it further from there leads to a drop in the yield almost back to the original level. The best modification was the t5h, which is a double mutant (P3A, P5L). The point mutant, T11E, did not lead to any increased yield. Extending the N-terminus slightly did lead to an increased yield, but especially for the highest yielding extensions, the ones including the glutamate residue, a splitting of the main band was observed, as evident from Figure 8 (B).

All constructs were tested for their thermo-stability and it was found that only the K356D mutation showed a drop in thermo-stability, supporting that this particular mutation, D356K in the hADRB1-US, could be forming a salt bridge. However, the drop in thermo-stability was only around 2 degrees as measured by the CPM assay. In summary, this meant that four modifications had been identified, all of which improved the yield but exhibited no, or only a very modest, change in thermo-stability. However, due to the band splitting of the construct MG[SES]P, possibly indicating some inhomogeneity, this particular modification was not used for further studies. It has to be mentioned however, that upon purification, the band-splitting also disappeared here, hinting at a possible effect caused by cellular lipids or other factors that are removed during purification.

Improved yield and retained thermo-stability in combinations of modifications

Since the improvement of the yield was in the order of 1.5 to 2-fold in the best cases for each individual modification, an increase was still needed to make structural biology experiments feasible. Therefore the two highest yielding reverse mutations, along with the best N-terminal modification, the t5h, were combined in all possible ways to explore whether the yield could be

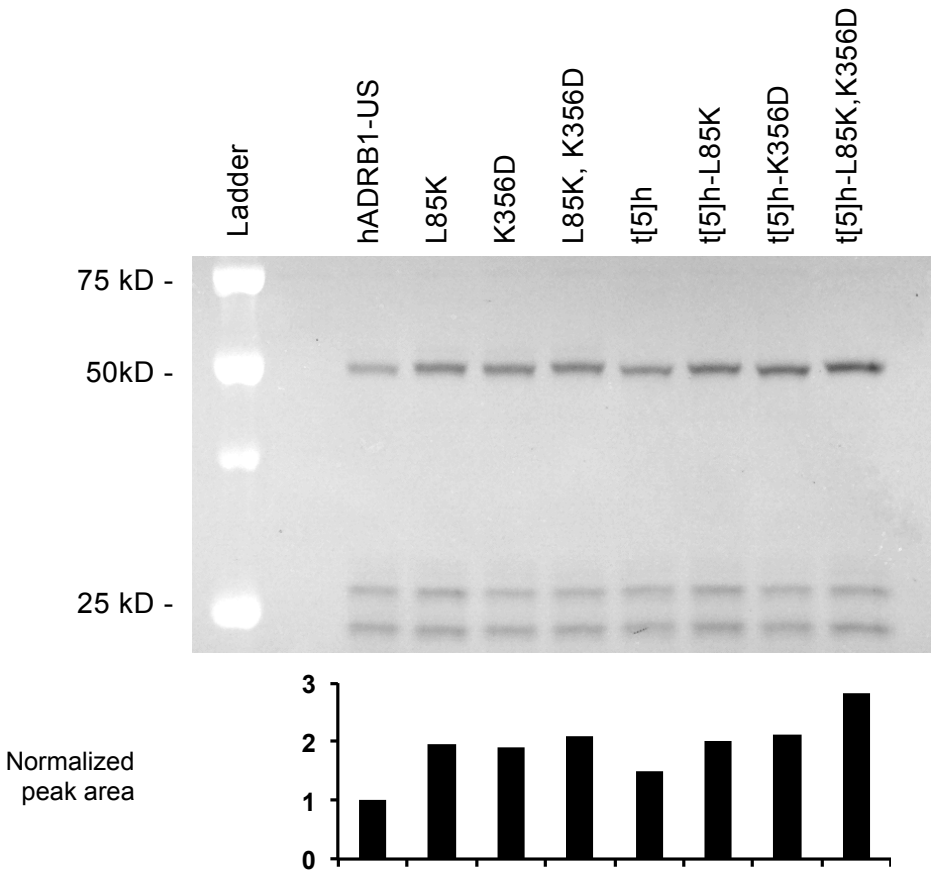


Figure 9. Effect of combining three modifications of the hADRB1-US construct.

In-gel fluorescence of combinations of the three modifications of the hADRB1-US-EGFP fusion constructs including the point mutations L38K, K264 and the t[5]h modification. Mutants expressed in HEK 293 cells by transient transfection, solubilized in 1%DDM and cleared by ultra-centrifugation. Each particular combination is indicated above each lane. Below gel are the peak areas found as normalized to the peak area of hADRB1-US.

improved further. As evident from Figure 9, a further increase in the yield is seen in nearly all cases, especially when combining all three modifications. It is noted, though, that the combination of modifications was not absolutely additive in terms of yield.

Importantly, the melting temperature did not decrease further when several modifications were combined. Figure 10 shows melting temperatures Tm of all combinations, including individual modification. The melting temperatures were measured for the apo-form, that is the ligand free form, as well as for an agonist (isoproterenol) and several antagonists (atenonol, alprenolol, cyanopindolol) (chemical structures of ligands can be seen in Tabel

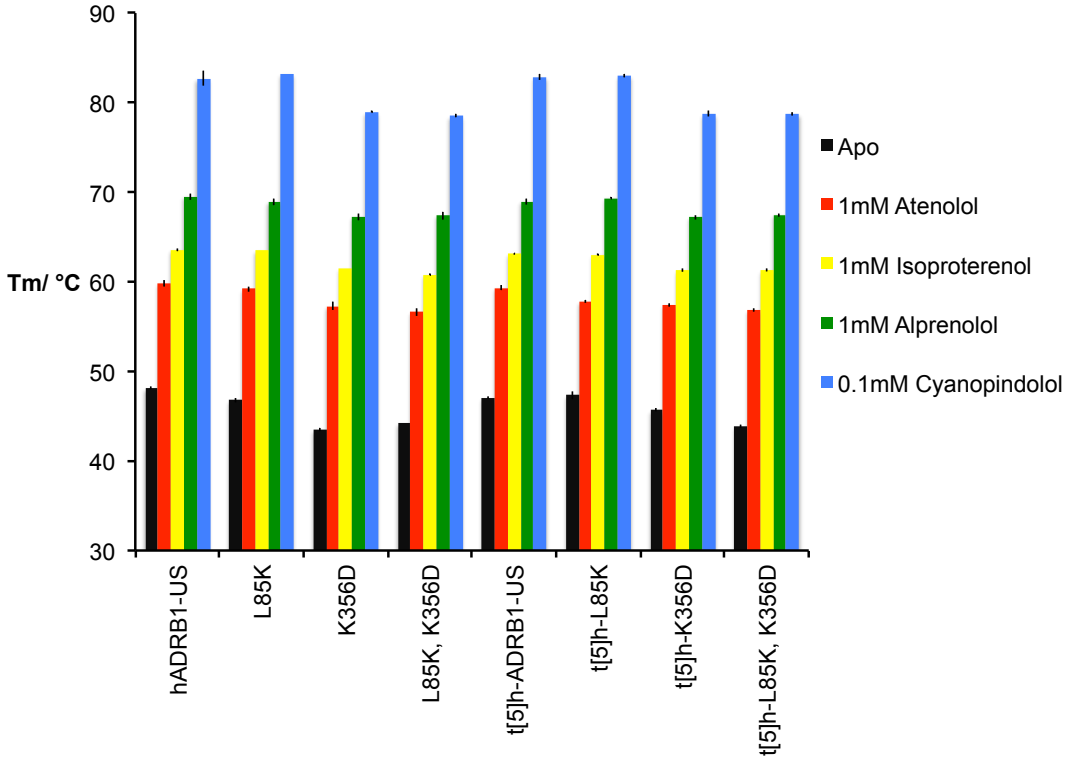


Figure 10. Thermo-stability of combinations of modifications measured with various ligands

Melting temperatures of receptors with combinations of reversion mutants without ligand or in the presence of ligand as indicated. Melting temperatures were determined by thermal denaturation of the purified receptor in DDM, in the presence of the CPM dye. The CPM dye becomes fluorescent upon binding to cysteine thiol group, that are exposed upon denaturation, but otherwise buried, and therefore inaccessible to the CPM dye, when the receptor is in its folded form. As such the degree of unfolding is detected by the relative degree of fluorescence. The assay was performed using a thermal cycler with fluorescence detection.

A 1). In all cases, whenever the K264D reversion mutation is present, it is accompanied by a very consistent drop in the Tm. It is also noted that as expected, the higher the affinity of the ligand, the higher degree of stabilization (Baker 2010).

To test whether the homogeneity of the receptor was retained the receptor mutants were analyzed by fluorescence size exclusion chromatography (FSEC), detecting EGFP fluorescence from the EGFP fused to the receptor upon separation by the size exclusion column. Figure 11 shows FSEC profiles

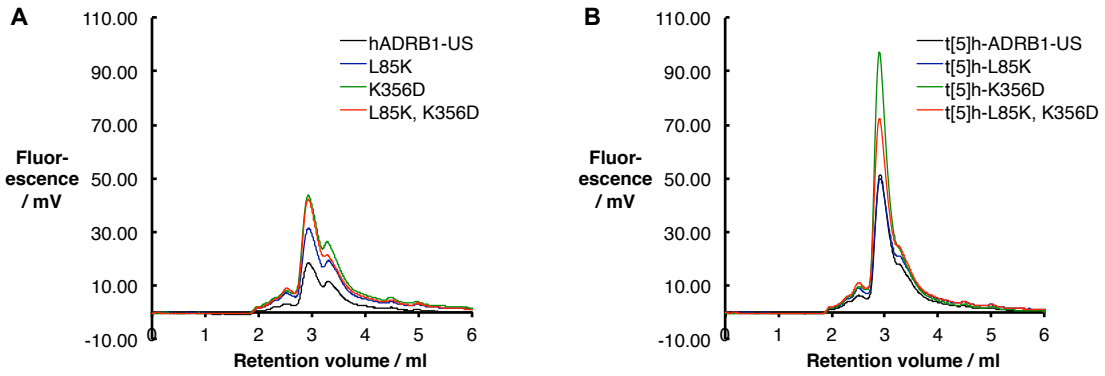


Figure 11. FSEC profiles of combinations of modifications.

Fluorescence-Size Exclusion Chromatography (FSEC) profiles of combination mutants. HEK293 cells with overexpressed receptor mutants, produced by transient transfection, were solubilized in 1% DDM and lysate was cleared by ultracentrifugation (150'000 g for 20 min). FSEC runs performed at 4 degrees Celsius on a TOSOH Super SW3000 column, for which the column void is 2 ml and the total column volume is approximately 4.5 ml. A calibration of the column is found in Figure A 1. Individual mutations and combinations as indicated in panels.

of all constructs run in cleared lysate after solubilization in 1% DDM. A sharp main peak is observed around 2.8 ml of retention volume for all combinations corresponding to size of 110 kD which roughly equals the weight of the receptor, EGFP and a DDM micelle combined. A minor peak is seen at around 3.4 ml, which corresponds to free EGFP. In all cases, only very little signal is detected for lower retention volumes, suggesting that the receptors are almost exclusively present in a monomeric form. Therefore, the main difference between the peaks is their individual sizes. One can observe that the hierarchy roughly corresponds to that observed in the in-gel fluorescence experiment in terms of intensity, with the original construct at the bottom and the combinations at the top.

Discussion

In order to do structural studies of the human beta1-adrenergic receptor (hADRB1), a crystallization type construct has been made using a similar construct for a closely related avian receptor, the turkey ADRB1 as a mold (Warne et al. 2003, Warne et al. 2008, Warne et al. 2009), which was the second non-rhodopsin GPCR to be crystallized. The human construct called hADRB1-US, was truncated in its N-and C-termini as well as its intra cellular loop 3, to remove flexible regions. A total of 9 nine thermo-stabilizing mutations were transferred to the human construct, originally identified in the tADRB1 (Serrano-Vega et al. 2008, Serrano-Vega and Tate 2009, Miller and Tate 2011, Miller-Gallacher et al. 2014), making the hADRB1 thermo-and detergent stable. However, two of the nine stabilizing mutations were found to have a positive impact on the yield when reversed to their original identity, while also removing two prolines from the N-terminus had a positive impact on the yield. All three modifications were combined to obtain a receptor construct with an only 2 degrees lower melting temperature, as determined by the CPM assay, but with a significantly higher yield.

An interesting observation is that several of the transferred stabilizing mutations can be removed without changing the melting temperature drastically. This means that one or more of the mutations are redundant, and as such can be avoided. A directly following question is then if even more of the mutations can be removed without a further decrease in the thermo-stability. Of course, an exhaustive analysis to find a minimal set of thermo-stabilizing mutations, while retaining same thermo-stability, would be virtually impossible due to sheer number of combinations available. An alternative approach for thermo-stabilization has been developed using directed evolution, and might more easily identify a minimal set of mutations (Sarkar et al. 2008).

The selection criteria applied here was very simple: only the mutations giving a significant increase in the yield were of interest. In this particular case only two fulfilled this requirement, limiting the number of combinations to 1, not considering any other modifications. Here it was also observed that these two stabilizing mutations could be removed without a synergistic decrease on the melting temperature. This particular finding, however, is in hindsight not a surprise. The two mutations reside in different types of regions in the receptor. Although both are polar residues, both in contact with the solvent, L85K is positioned at end of TM1, while K356D is positioned well into ECL3. The different character of these positions makes it unlikely that they would be correlated in the thermo-stabilization.

The high similarities in terms of sequence homology between the two receptors, hADRB1 and tADRB1, but their relatively distinct physical properties, pharmacological profiles, and difference in their abilities to be over-expressed, offers a window to how these individual characteristics can be explained structurally, since the explanation lies in the set of residues that differ among the two receptors. It is especially advantageous here since the regions that differ amount to only around 20%.

Materials and Methods

Plasmids

All plasmids for the initial expression screen of single modifications of the hADRB1-US receptor were derived from the pACMVtetO plasmid, in which an EGFP with a 1D4-tag at its C-terminus preceded by a 15 amino-acids linker, is fused C-terminally of the receptor ORFs. All plasmids used in the second screen of combinations of modifications were derived from the plasmid pEGFP-C1, to which the whole expression cassette of the construct hADRB1-US-EGFP 1D4 had been transferred. This plasmids was modified further by exchanging the EGFP-1D4 with a C-terminally Twin-Strep tagged EGFP(gBlock, purchased from IDT). All modification, combinations of modifications, in the hADRB1-US coding DNA region were done by PCR or overlap extension PCR, and the resulting products were sub-cloned into the appropriate vector background.

Mammalian cell culture and transient transfection

All expression tests were done using HEK 293 cells. Cells were grown in DMEM(high glucose , Bioconcept) supplemented with 10% FBS and 1% Penicillin/streptomycin (Bioconcept). Cells were grown in 15cm dishes to an approximately 80-90% confluence at the point of transfection. Transfection was done using linear PEI at a ratio of 3µg PEI per 1µg DNA. 40µg of DNA was used per 15cm dish. PEI was added to 5 ml serum free DMEM per dish, to which DNA was added. After 10 min, the mixture was added to the cells. The cells were harvested after 72 hours with a rubber police man, washed in phosphate buffered saline, and stored at -20 degrees Celsius.

Solubilization and In-Gel fluorescence

The thawed cells with over expressed receptors, were re-suspended in buffer (150 mM NaCl, 50 mM Hepes) with protease inhibitors (Roche complete EDTA free), were homogenized with a2 9 gauge needle and a syringe,

and briefly sonicated at low amplitude. All steps were carried out either on ice or at 4 degrees Celsius. The cell suspension was added DoDecylMaltoside to a final concentration of 1%, and the solubilization took place for 1 hour. After solubilization the DDM insoluble material was pelleted by ultracentrifugation at 150'000 x g for 20 min. The supernatant was separated by SDS-PAGE(12% Trupage, Sigma), and imaged in a Amersham Imager AI600(GE Lifesciences).

Small scale purifications

The solubilized fractions described above were used for purification by means of the 1D4 tag or where possible the Twin-Strep. For the 1D4 purification the solubilized fraction was incubated with purification resin on which anti-1D4 antibody has been conjugated (gift from Chayne Piscitelli). The binding was done in batch mode in an Eppendorf tube, and was left overnight to bind. The following day, the resin was gently spun down, and washed 3 times in wash buffer (150 mM NaCl, 50 mM Hepes pH7.5, 0.03 % DDM), and eluted off the resin by incubation in wash added 1D4 peptide to a concentration of 80µM. The elution step was carried out over night.

For the purifications using the Twin-Strep tag, the solubilized fractions were applied onto a 3ml gravity flow columns with 0.2 ml Strep-Tactin sepharose, which had been equilibrated in wash buffer (150 mM NaCl, 50 mM Hepes pH7.5, 0.03 % DDM). The columns were washed with 6 ml wash buffer, and the proteins were eluted off resin with 0.6 ml wash buffer with 5 mM desthiobiotin.

Thermo-stability assay

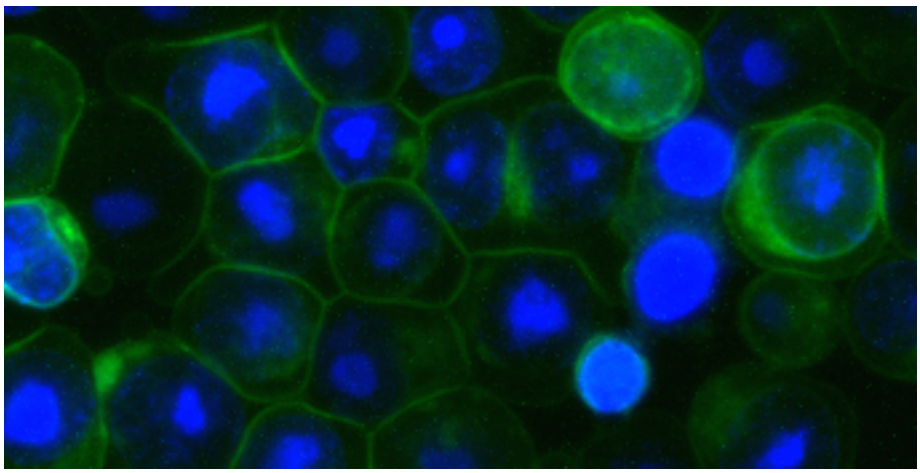
Thermal stability was determined by means of the CPM assay (Alexandrov et al. 2008). Briefly, the purified proteins described above, were thermally denatured by increasing the temperature from 30 to 90 degrees Celsius at 4 degrees per minute. This is done in the presence of the CPM dye, which becomes fluorescent when its maleimide moiety reacts with cysteine Sulhydryl group. In this way chemically, inaccessible cysteines, become

exposed when the receptor is denatured, and react with the CPM dye. During the denaturation the fluorescence is detected. All runs were performed in a Qiagen Rotorgene QPCR machine. The denaturation profiles were analyzed by the software accompanying the apparatus, and the melting temperature was defined as the infliction point of the denaturation curve, at which the deriviative of the fluorescence as a function of temperature assumed its maximum value. Noise of the fluorescence raw data was removed using the software by applying a sliding average over the temperature, termed a heavy filter digital filter. All species were measure as an average of 4. Approximately 0.5-4 ug of protein was used for each sample, using a concentration 0.2µg/ml of the CPM Dye.

Fluorescence Size Exclusion Chromatography

DDM solubilized samples describe above were filtered using 0.22 µm table-top micro centrifuge filter (Millipore) in order to remove large particles that could damage chromatography column. Upon filtration samples were added NaCl to a final concentration of 500 mM. Samples were then analyzed by being run on a Tosoh SuperSW3000 analytical gel filtration column using an Äkta Ettan HPLC system (GE) equipped with an auto-sampler and fluorescence detector set to excite EGFP(485nM), with emission detected at 520 nm. The mobile phase was 50 mM Hepes pH 7.5, 500 mM NaCl, 0.05% DDM. The column void, V0, is 2ml, while the total column volume, Vt, is 4.5 ml as estimated void peak and small molecule peak using a Biorad protein standard. The run was continued for 6.5 ml for each sample. All runs were performed at 4C.

3. Tuning expression of GPCRs for the secretory pathway in the Baculovirus-Insect cell expression system



Chapter 3 frontpage. Fluorescent image of insect cells expressing EGFP-tagged GPCR.

Cells were imaged using a fluorescent microscope (FLoid). Nuclei were stained with DAPI. EGFP and DAPI was visualized with the EGFP and DAPI setting, respectively.

Introduction

Structural studies of G-Proteins Coupled Receptors (GPCRs) and membrane proteins in general, are an eminent task. Hurdles include lack of long-term stability when extracted by detergent from the cell membrane, and difficulties forming crystals for X-ray crystallography due to the sometimes-limited hydrophilic surface area, which are the main points for crystal contact formation. Last but not least is the yield of membrane proteins when recombinantly overexpressed, which is almost always much lower than that obtained of cytosolic proteins.

Sufficient stability upon extraction from the plasma membrane can be achieved in a number of ways including protein engineering like scanning mutagenesis (Zhou and Bowie 2000, Serrano-Vega et al. 2008), use of stabilizing ligands, or the by choosing a homolog of the target protein with better characteristics. Likewise, crystallizability, can be improved in much the same way, and in addition new crystallization techniques have emerged which include lipidic cubic phase, which is a type of crystallization method mainly applied to integral membrane proteins (Caffrey 2015).

In spite of having these relatively new tools, obtaining sufficient amounts of a given membrane protein is one of the biggest obstacles in the field. Historically, overexpression of GPCRs was initially tested in the most common expression system for overexpression of proteins, *E. Coli*, with initial and successful attempt on the rat Neurotensin receptor (se for example Tucker and Grisshammer 1996). However, due to low yields, the most successful expression systems are Eukaryotic, especially from higher eukaryotes, which include mammalian cells and insect cells, the rat Neurotensin structure was eventually solved using a eukaryotic expression hosts (White et al. 2012, Xiao et al. 2013). Curiously, the Neurotensin receptor was indeed produced in E. Coli to yield an X-ray structure (Egloff et al. 2014). This involved a newly

developed method for directed evolution of GPCRs, toward detergent-stable mutants of high expression yield in this organism (Sarkar et al. 2008).

The model GPCR, the turkey beta1-adrenergic receptor (tADRB1), was originally chosen for structural studies due to its relatively high yield when overexpressed and high thermo-stability (Warne et al. 2003, Warne et al. 2009). However, early studies showed that this particular receptor exhibited a distinct pharmacology from that of the mammalian counterparts (Minneman et al. 1980). Thus, using the turkey in place of a mammalian beta1-adrenergic receptor is clearly illustrating the limitations of this approach. However, abolishing the "easier homolog"-strategy presents one with the obstacle of challenging production and handling conditions in order to obtain high quality protein for structural studies. Therefore, rather than engineering only the protein, another option is to optimize the expression system.

Amongst the eukaryotic expression systems, the baculovirus-insect cell expression system is one of the most used and successful systems. It is routinely used for expression of membrane proteins and is, in fact, by far the most frequently used expression system for GPCRs for which an X-ray crystallographic structure has been determined (Lv et al. 2016). However, there have been reports of large amounts of misfolded protein especially for secreted or membrane proteins (Tate and Blakely 1994, Tate et al. 2003).

The co-translational folding of membrane protein differs from that of cytosolic proteins in that it engages a fundamentally different and elaborate mechanism. The pathway consists of the Sec-translocon, a protein-conducting channel, requiring several co-factors, on the rough ER membrane in eukaryotes (Zimmermann et al. 2011). This pathway is also responsible for the co-translational folding of secreted proteins and concomitant translocation of these to the ER. Thus, the folding pathways of secreted and membrane proteins are referred to as being convergent in eukaryotes.

Previous studies have shown that including an N-terminal signal sequence (also referred to as signal peptide) can have a modest to large effect on the expression level of membrane proteins including GPCRs (Guan et al. 1992, Korepanova et al. 2009). One advantage with the signal sequence is that it is always cleaved off in its native form, thus when fused to a recombinant target, it is normally cleaved off, leaving only a minimal alteration of the protein construct of interest.

In this study the effect of both different promoters, and the AcMNPV gp67 signal sequence have been explored in the overexpression of several related GPCRs using engineered crystallization type constructs, specifically belonging to beta-adrenergic receptors.

Results

Experimental strategy: Different amount of functional receptor of two closely related GPCR homologs

Previous studies have shown that the 6 mutations, the m23 mutation, that were found to stabilize the turkey ADRB1 (Serrano-Vega et al. 2008) could be transferred to both the human ADRB1 and human ADRB2, with the former exhibiting the highest degree of increased stability, however with the latter initially being significantly more stable than the former (Serrano-Vega and Tate 2009). In addition 3 mutations were identified that increased the thermo-stability of the turkey ADRB1 even further, to give what is referred to as an ultra-stabilized receptor (Miller and Tate 2011). It has to be mentioned that this receptor, an engineered form of the turkey ADRB1, has been truncated in the N- and C-termini, and has a portion of its intracellular loop 3 deleted, in order to minimize flexibility to improve crystalizability (Warne et al. 2003, Warne et al. 2009).

It has been found that all these mutations can be combined together with the truncations and the deletion mentioned above, to obtain an analogous construct of the ultra-stabilized turkey ADRB1, for the human ADRB1, which is called the ultra-stabilized human ADRB1, referred to from now on as hADRB1-US. However, due to very low yield, the receptor was re-engineered, including removal of several of the thermo-stabilizing mutations, which had previously been reported to have a negative impact on the yield (Serrano-Vega and Tate 2009). In that way a receptor with an improved yield was obtained, which only had a slightly lower thermo-stability (Chapter 2; Joergensen et al., manuscript in preparation). This receptor, had 2 of the 9 thermo-stabilizing mutations removed, while two prolines at position 3 and 5, in the construct, were mutated to alanine and leucine, respectively, in an otherwise identical construct. This receptor is referred to as the thermo-stabilized human ADRB1, abbreviated hADRB1-TS. The particular turkey

ADRB1 construct used here and in the just mentioned manuscript is the tADRB1.44-m23 construct also referred to as beta1-AR44-m23 (Warne et al. 2012).

Initial expression of the hADRB1-US showed that the yield was only a fraction of the yield of the analogous tADRB1.44-m23 construct. In addition, it was noted that only a fraction of the hADRB1-US could be solubilized in DoDecylMaltoside (DDM), while the degree of solubilization appeared slightly higher for the tADRB1.44-m23 although notably still exhibiting a large amount of DDM (using 1% DDM W/V) insoluble material (data not shown). The inability of a large fraction of the produced material to be solubilized was a curious and contradictory finding, since it meant that part of the receptor population was not compatible with DDM, while the other part was indeed compatible.

In membrane protein work a screen is often performed to identify a suitable detergent for solubilization (Newby et al. 2009), that is, the process of solubilizing both proteins and membrane lipids, essentially separating these constituents from one another, if enough detergent is used, while ideally not destroying the integrity of the functional protein. Likewise, one might also screen detergents for further processing. The ability to solubilize membrane proteins and lipids differs among detergents. On one end of the spectrum is Sodium-Dodecyl-Sulfate (SDS) solubilizing almost anything, however, almost always with denaturation as a consequence and is thus termed harsh. At the other end detergents such as DDM solubilize while typically retaining functionality of the protein, and are thus termed mild (reviewed by le Maire et al. 2000).

Due to the high sequence similarity and thus their inferred structural similarity, the general ability of the two receptors, tADRB1.44-m23 and hADRB1-US, to be solubilized by DDM, in their functional form, would be expected to be very similar. However, one must keep in mind that if a given

protein does not tolerate the detergent, here DDM, it might unfold which in turn most likely would lead to aggregation and not staying in solution. As such, screening for detergents with the ability to solubilize more of the receptor, and hence probably being harsher than DDM, did not seem attractive, nor did it seem reasonable.

Alternatively, a scenario where, in addition to the soluble receptor, an insoluble form of the receptor is also present upon overexpression was considered a possibility to investigate further. One hypothesis, which is consistent with this scenario, could be the simple situation where accumulation of misfolded receptor is observed during over-expression. This could also be consistent with the large fraction of insoluble material when using DDM to solubilize. In other words, it is hypothesized that the insoluble receptor is misfolded receptor. While this study was being carried out another group published a method exactly along those lines, using solubilization in DDM or Digitonin as the means to assess whether or not a membrane protein was correctly folded (Thomas and Tate 2014).

Since co-translation folding of a membrane protein is a complex multistep process requiring many factors, a possible explanation for the accumulation of misfolded protein could be that the pathway is being saturated to a point where it cannot produce the receptor in its correctly folded form. It is known that the standard promoter in the baculovirus expression system, the polyhedrin promoter, is extremely strong (Matsuura et al. 1987), and that saturation of a biosynthetic pathway is not uncommon in engineered systems for high expression (Grimm et al. 2006). In this way, testing whether or not this could be the case seemed reasonable.

Relieving the stress on the SEC-translocon could potentially be achieved by reducing the amount of transcribed mRNA for the particular constructs being translated through this folding pathway. Thus, a search for promoters of different strengths was initiated. A recently published study by Lin and Jarvis

(2013), examined several baculovirus promoters of ACMNPV, and found that a delayed early promoter the 39K promoter (P.39K) gave the highest amount of the functional form of a secreted protein. Therefore, a screen was set up to test different promoters effect on the protein yield and ability to be solubilized by DDM.

In addition, all constructs were tested with or without the AcMNPV gp67 signal sequence ($gp67^{SS}$), in order to assess the effect of this on constructs with this particular architecture. The $gp67^{SS}$ was chosen due to its previous use on GPCR and membrane proteins (Korepanova et al. 2009, Kofuku et al. 2012). Additionally, it is one of the most well characterized signal sequences. As such, the $gp67^{SS}$, is considered part of the expression system, rather than being part of the protein itself.

Construct design

Initial experiments were carried out using the MultiBac (Berger et al. 2004) derivative EmBacY (Bieniossek et al. 2008, Trowitzsch et al. 2010), which in addition to its ability to express a given target from a certain location in the viral genome, also co-expresses EYFP under control of the late p10 promoter from elsewhere in the genome. This makes it possible to follow the viral infection by monitoring EYFP, either by routine fluorescence microscopy or by other means of assessment of fluorescence. However, due to the desire of using EGFP as a reporter for the receptor in downstream experiments, co-expression of EYFP was not optimal due to spectral overlap of these two fluorescent proteins. Therefore, a new plasmid was designed on the basis of pFL to be used with the Multibac baculo viral backbone. Here mCherry was put under control of the p10 promoter (P.p10) in pFL, while the protein of interest was being expressed under the polyhedrin promoter (P.ph). In addition, two other plasmids were made in which the P.ph was exchanged with the Immediate Early 1 promoter (P.IE), and the 39K promoter (P.39K) (Guarino and Summers 1986, Guarino and Smith 1990), respectively. In this

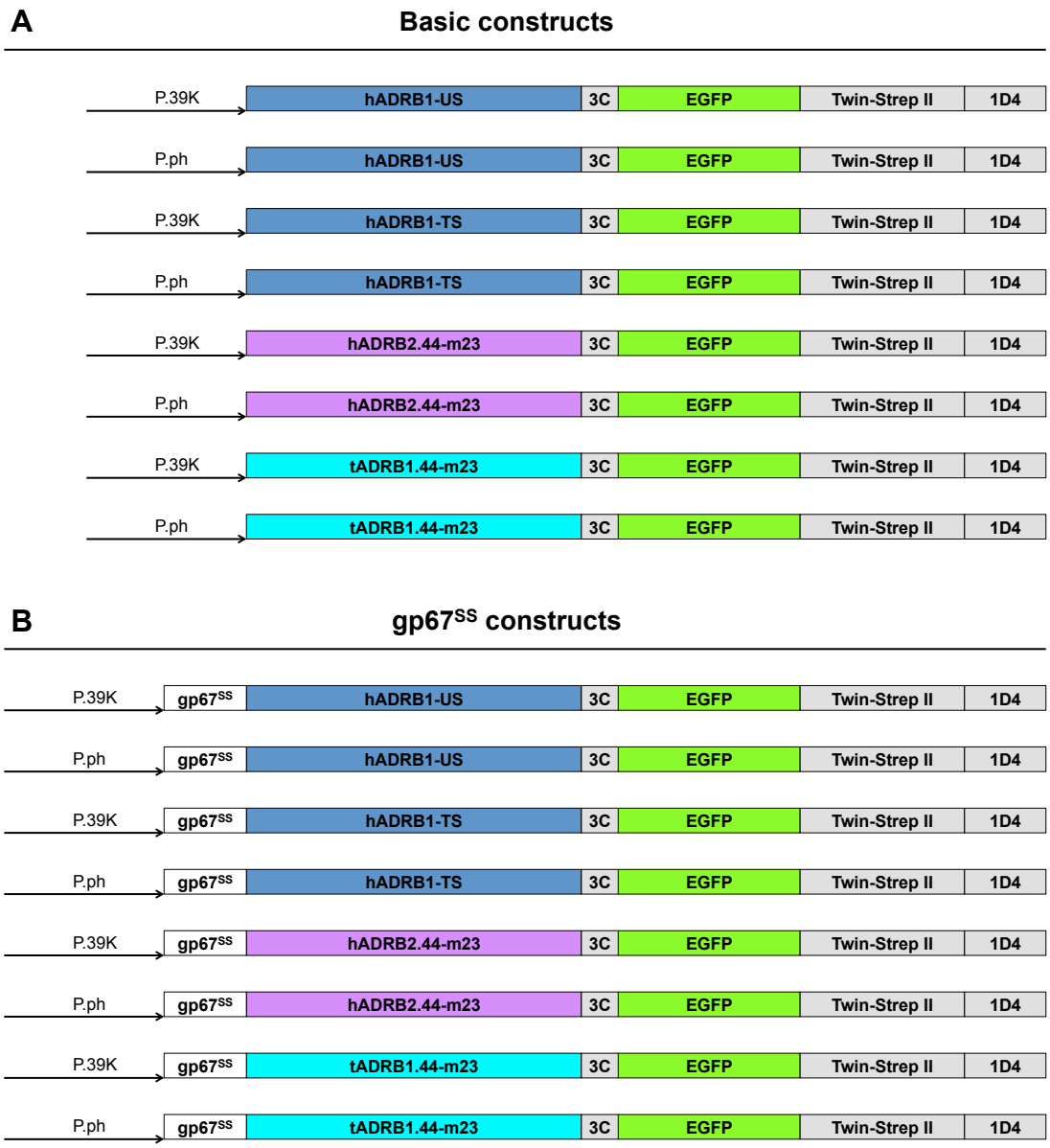


Figure 12. Overview of constructs used in this study.

Promoter and protein coding region have been included for all constructs. Constructs are not drawn to scale. The 39K promoter (**P.39K**) or the polyhedrin promoter (**P.ph**) are indicated as arrows in leftmost part of each construct. (**A**) shows basic construct, while (**B**) shows construct with the gp67 signal sequence (**gp67^{SS}**) .

way, infection could be observed through mCherry fluorescence, while expression of the GPCR could be observed through EGFP fluorescence, with the EGFP being present at the C-terminal end of the GPCR. An overview of the plasmid architecture can be seen in Figure 12. The receptors are joined to the C-terminal EGFP by a flexible linker containing a 3C protease cleavage

site, while the EGFP itself has a purification tag attached to its C-terminus – a Twin-Strep II, and a 1D4 tag.

Higher degree of solubilization in 1% DDM when using the 39K promoter

An initial screen testing all three promoters was done for all constructs without signal sequences, and it quickly became clear that the expression level of the P.IE was not of any immediate relevance for over expression due to the extremely low yield of protein for all constructs (data not shown). Therefore, P.IE was not considered further.

As seen in Figure 13, both a varying degree of total expressed polypeptide is seen when comparing different constructs, while also a varying amount of DDM solubilized receptor is observed. A main band is seen at around 50kD, which correspond to the monomeric only partially SDS-denatured polypeptide, leading a migration at a lower-than the expected mass of around 67kD for all constructs. Also, a weaker upper band is visible, commonly referred to as part of an aggregation ladder (Prive 2007).

Not surprisingly, the tADRB1.44-m23 constructs are giving the highest yield on all counts. It is clear that while P.ph supports a higher total amount of receptor being solubilized in DDM, the P.39K expression is associated with a larger fraction of total material being solubilized by DDM.

Importantly, an extremely small degree of hADRB2.44-m23 receptor could be solubilized, indicating that something was fundamentally wrong with this particular construct. It is clearly seen that this construct leads to two clearly separated main bands. An explanation for that may be found in the fact that a glycosylation site has been left intact in this construct. Although a surprise, and of no apparent interest for any structural studies, this particular construct could offer an example of a negative control for this assay. Therefore, it was kept for further analysis along with the rest of the constructs. Indeed similar results were obtained from a stable HEK 293 cell-line expressing the hADRB2.44-m23 (data not shown). It is of interest that the general construct

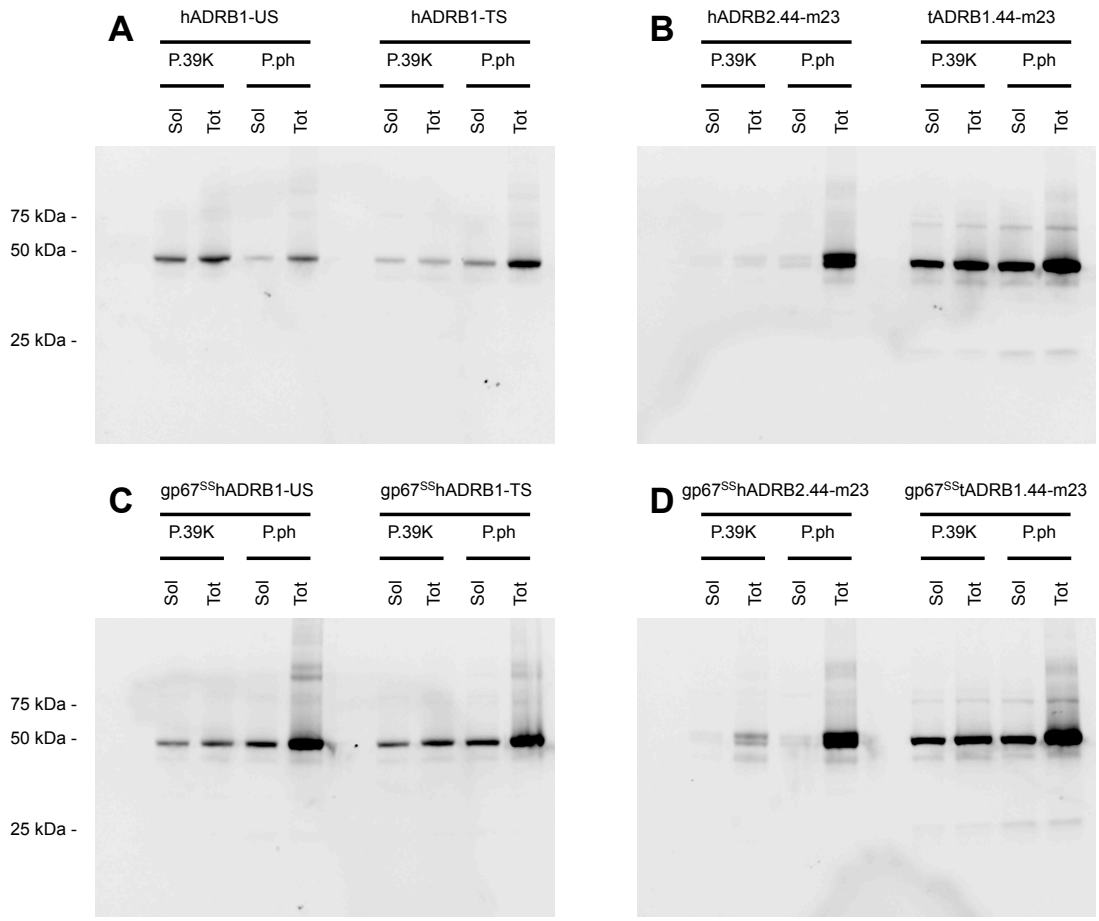


Figure 13. Higher degree of solubilized receptor when using the 39K promoter compared to the polyhedrin promoter.

In-gel fluorescence of various GPCR constructs expressed in Sf9 cells using recombinant baculoviruses, solubilized in 1% DoDecyl-Maltoside (DDM). Constructs, promoter and fraction is indicated above each lane. (**A**, **B**) basic constructs; (**C**, **D**) constructs with gp67 signal sequence. Solubilized material (**Sol**) is the cells solubilized in DDM, after ultracentrifugation (150'000g), while Total (**Tot**) is the exact same material before ultracentrifugation. Samples were separated by SDS-PAGE, and proteins were imaged using the filter setting for EGFP fluorescence

architecture is not transferable from the tADRB1.44-m23 to the human beta2-adrenergic receptor, since the thermo-stabilizing mutations could indeed be transferred (Serrano-Vega and Tate 2009).

It is particularly noteworthy that including the gp67 signal sequence, mainly has a large effect on overall yield and the solubilized material for the two constructs hADRB1-US and hADRB1-TS, while no significant change is seen for the tADRB1.44-m23 construct. Including a signal sequence in the

constructs does not appear to improve the degree of solubilized to total material.

Less aggregation-prone material when using the 39K promoter revealed by FSEC

In order to assess the state the different receptors when solubilized by DDM, but expressed using different promoters, the cleared lysate was analyzed by Fluorescence Size Exclusion Chromatography (FSEC). This allowed for an estimation of the degree of aggregated receptor in the fraction of over expressed receptor that could be solubilized. The result is shown Figure 14. Firstly, it is noted that the tADRB1.44-m23 construct exhibits a very sharp gel filtration profile in all cases, with a retention volume of 2.8ml of the peak corresponding to a molecular weight of around 120 kDa. A subtle difference, however, is seen in the region of higher molecular weight elution volumes and the column void volume of 2 ml, with higher amount being present here under expression using the polyhedrin promoter.

Another extreme are the profiles of the hADRB2.44-m23 constructs, which are markedly different from the rest of the constructs, and appears to be mainly aggregated protein with no apparent peak at 2.8ml. This reaffirms the conclusion stated above, that the hADRB2.44-m23 construct is probably mainly present as misfolded and aggregation prone polypeptide.

The most interesting cases are the human ADRB1 constructs. All constructs have a main elution peak at 2.8 ml, which is similar in appearance to the tADRB1.44-m23 elution peak, characteristic of GPCR in a DDM micelle. However, all human ADRB1 constructs have a large amount of material between the column void and the peak at 2.8 indicating for these receptors a considerable amount is still misfolded and aggregation prone. In all cases a clear and pronounced difference in the profiles is seen when comparing the constructs expressed under the P.39K to the P.ph, with the former leading to less aggregation prone material.

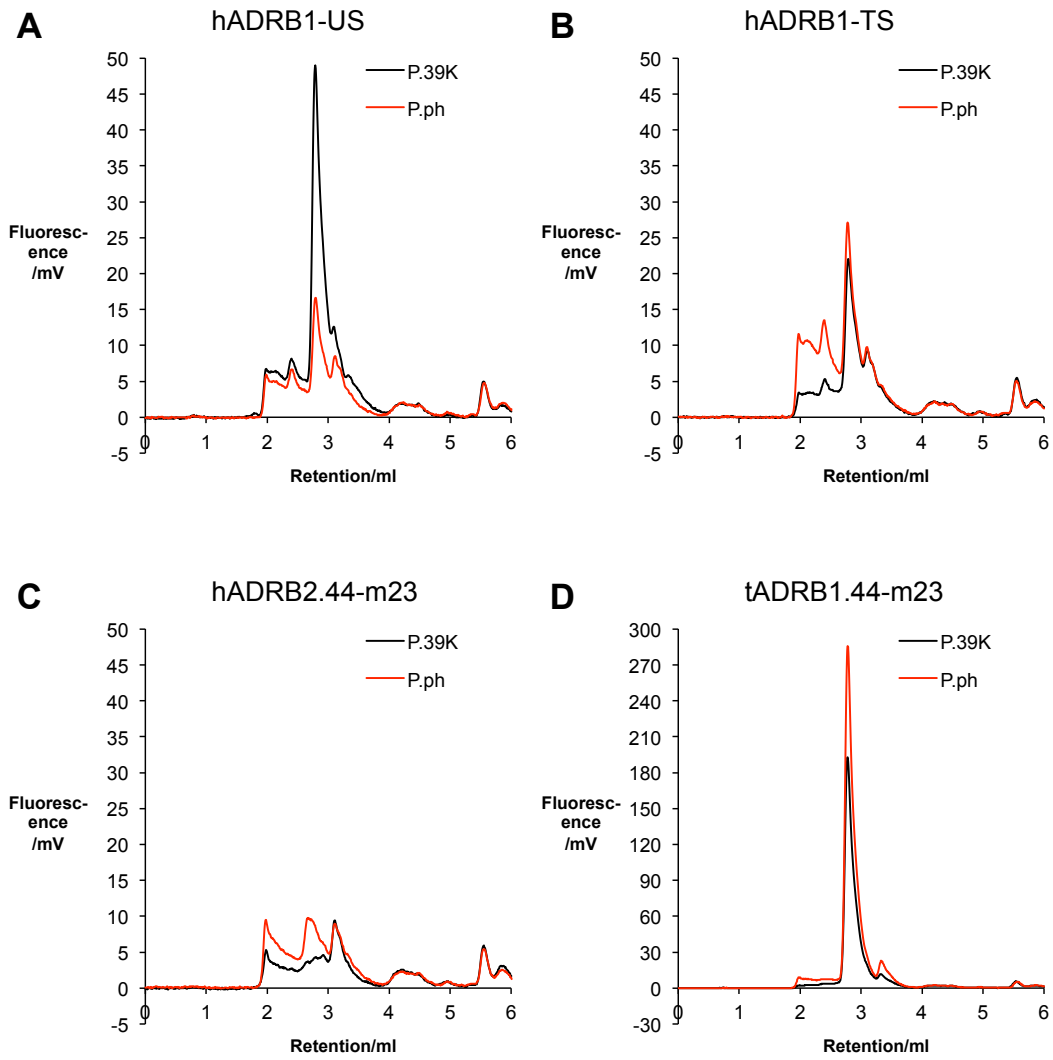


Figure 14. Less aggregation prone receptor in solubilized fraction when using 39K promoter compared to polyhedrin promoter.

Fluorescence–Size Exclusion Chromatography of DDM solubilized fraction of constructs as indicated and treated as described in legend of Figure 13. The run was performed using a TOSOH Super SW3000 column (Figure A 1), which has a void volume of 2ml and a total volume of 4,5. Excitation was set to 485 nm, and emission was set to 520 nm, for detection of the C-terminal EGFP present in all constructs.

Interestingly, again, the gp67 signal peptide did not alter the qualitative behavior of any construct (Figure 15), it had an effect on the overall amount of receptor solubilized by DDM, and hence the peak height in the FSEC profiles, as observed in the previous section.. No clear effect was observed for the turkey construct.

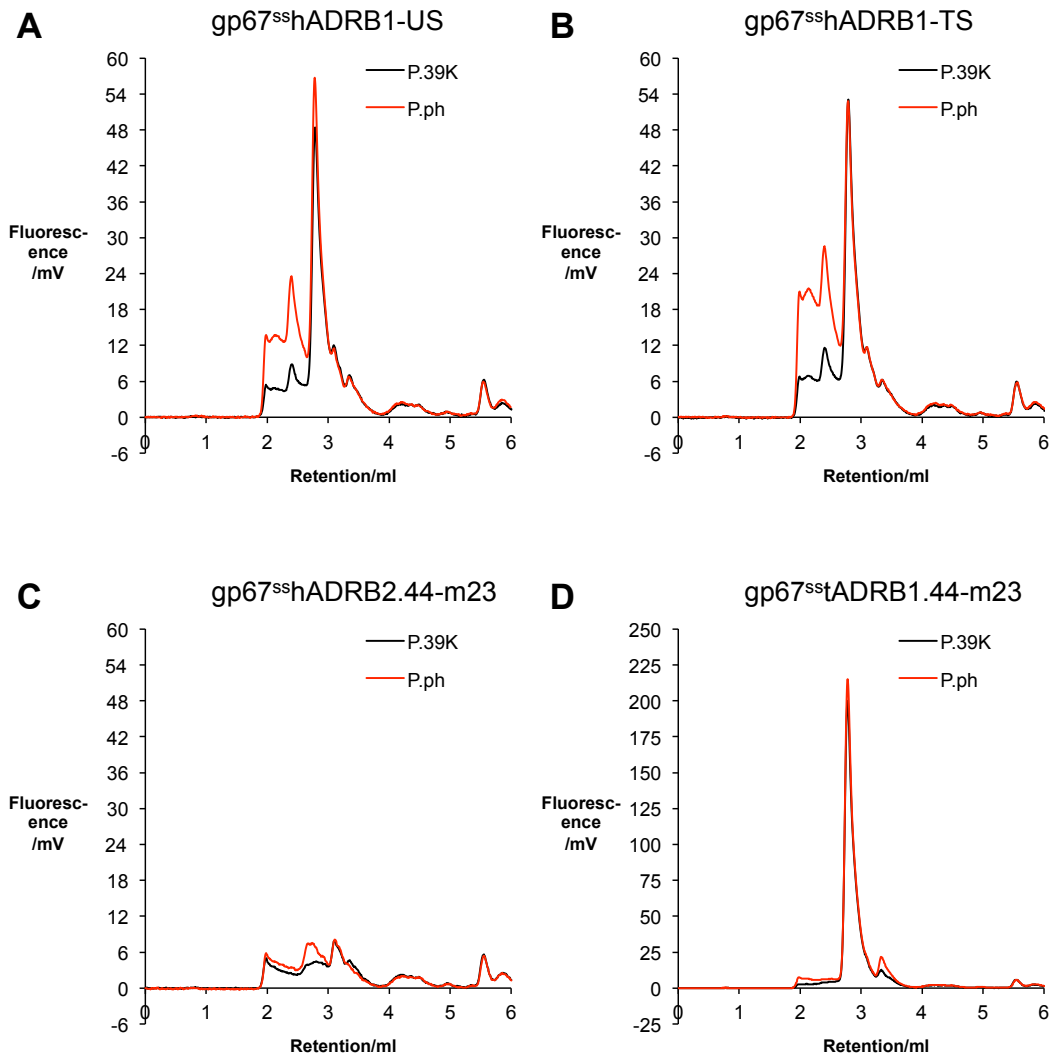


Figure 15. Less aggregation prone receptor in solubilized fraction when using 39K promoter compared to polyhedrin promoter in constructs with gp67^{SS} signal sequence.

Fluorescence–Size Exclusion Chromatography of DDM solubilized fraction of constructs as indicated and treated as described in legend of Figure 13. The run was performed using a TOSOH Super SW3000 column (Figure A 1), which has a void volume of 2ml and a total volume of 4,5. Excitation was set to 485 nm, and emission was set to 520 nm, for detection of the C-terminal EGFP present in all constructs.

To a varying degree a lower molecular weight peak is observed which is most likely coming from free EGFP, which could come from both solubilizable and in-solubilizable receptor from proteolytic cleavage both before and after solubilization.

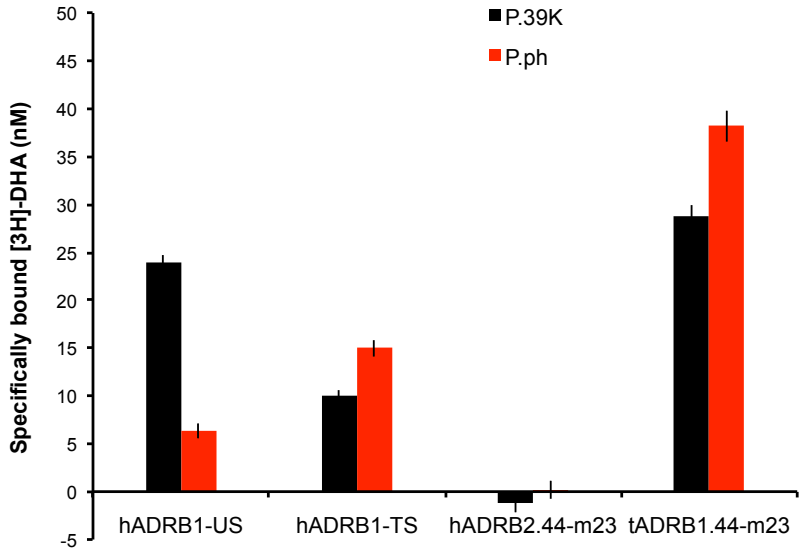


Figure 16. Higher functional yield from the polyhedrin promoter in membranes compared to the 39K promoter

Saturation single point radioligand binding of constructs as indicated on diagram. Tritiated DiHydroAlprenolol ([3H]DHA) (100nM) and crude membranes were incubated for 1 hour, and unbound [3H]DHA was separate from the membrane bound fraction on a glass fiber filter. Each sample corresponds to approximately 100'000 baculovirus infected cells.

The gp67 signal sequence has the largest effect on activity measured by saturation radio ligand binding

Having characterized the DDM solubilized over-expressed material, one obvious experiment was to quantify the amount of functional receptor expressed under the various conditions mentioned above. The absolute gold standard in quantifying functional GPCR, is the binding of radio-actively labelled ligand to the receptor. Therefore, the functional receptor yield was quantified as such using single point binding assay under saturating conditions. In Figure 16 and Figure 17 it is clearly seen that, as expected, all hADRB2.44-m23 constructs show no significant binding, supporting the findings from the in-gel fluorescence and the FSEC experiments and cementing that the construct is indeed only resulting in a misfolded receptor. It is worthwhile

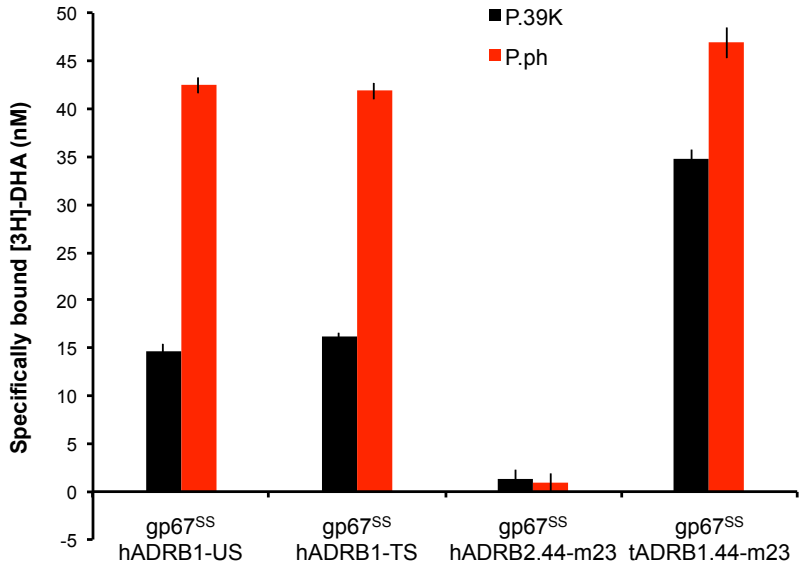


Figure 17. Dramatic increase in yield of hADRB1 construct when using gp67 signal sequence.

Saturation single point radioligand binding of constructs as indicated on diagram. Tritiated DiHydroAlprenolol ([3H]DHA) (100nM) and crude membranes were incubated for 1 hour, and unbound [3H]DHA was separate from the membrane bound fraction on a glass fiber filter. Each sample corresponds to approximately 100'000 baculovirus infected cells.

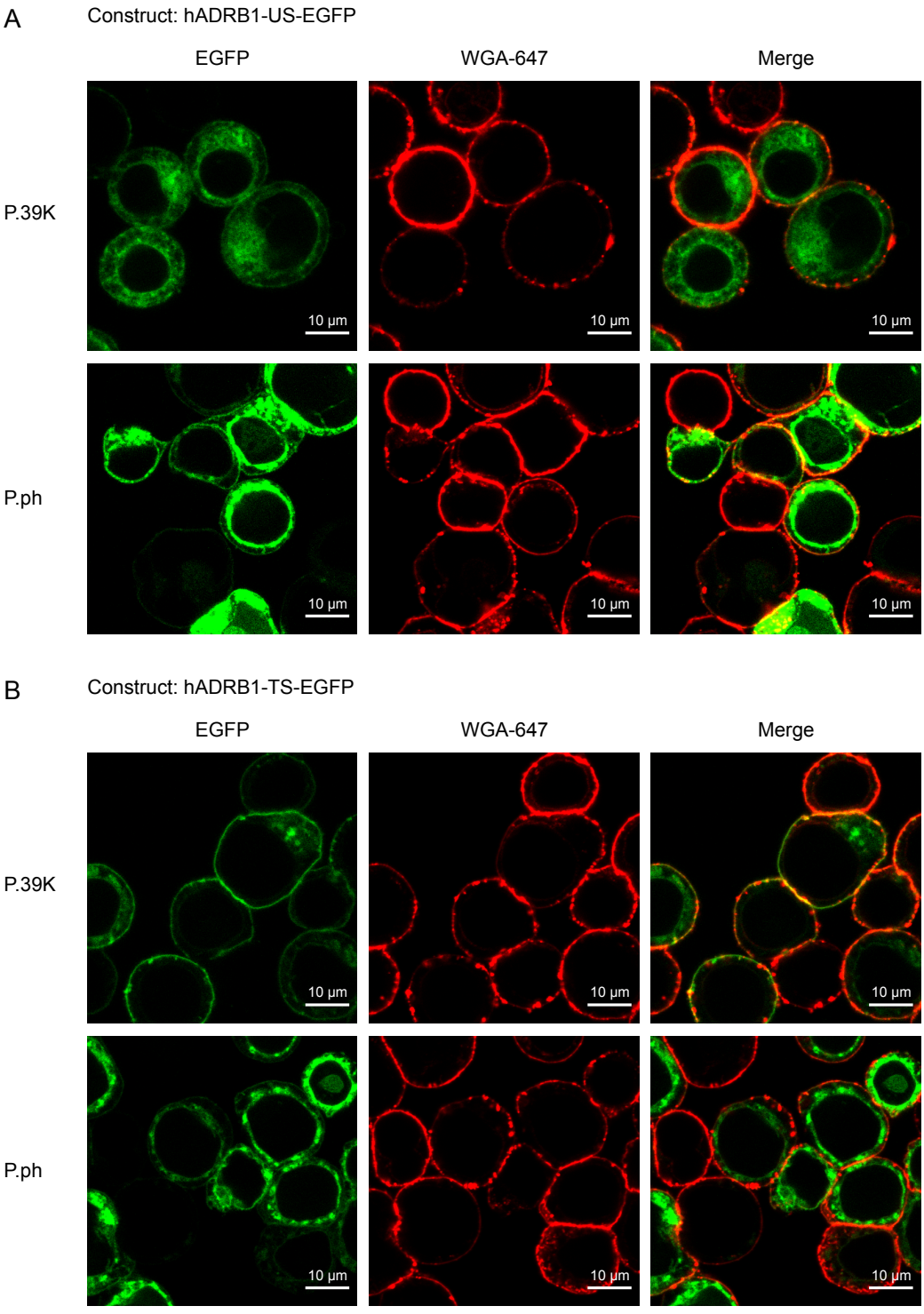
noting that the hADRB2 constructs clearly demonstrate that a C-terminal EGFP remains fluorescent even when the preceding membrane protein is misfolded. However, it cannot be excluded that even more misfolded receptor is present in which the C-terminal EGFP is also misfolded.

On the other hand all other constructs exhibit the ability to specifically bind radio-labelled ligand, however with large discrepancies between the turkey and human counterparts (Figure 16, Figure 17), agreeing with the previous results. It is interesting that the inclusion of the gp67 signal sequence has such a profound effect (approximately 5 fold increase) on the yield especially in the human constructs, while the effect is relatively modest (around 10% increase) in the turkey construct. This could suggest that the human ADRB1 receptor TM1, which serves as a signal anchor sequence for GPCRs without a native signal peptide (Spiess 1995), engages the SEC translocon less efficiently than the turkey counterpart, and explain the low yield of the hADRB1 constructs without the signal sequence ($gp67^{SS}$).

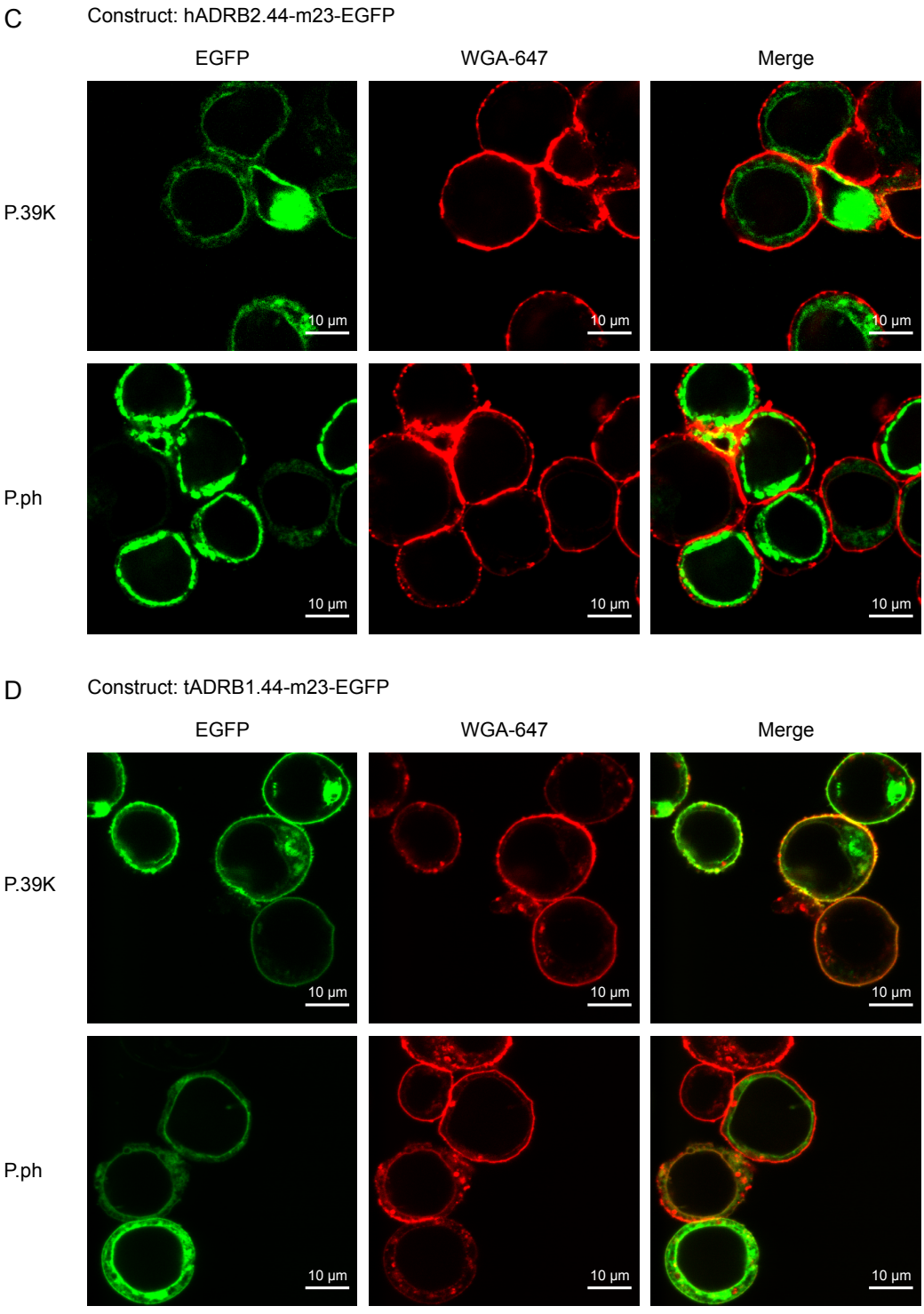
Preferential surface expression with the 39K promoter

To assess the effect of overexpressing the proteins using the two different promoters on the cellular level, each construct's effect regarding cellular distribution was analyzed by laser scanning confocal microscopy. Especially the ability of the cells to make the receptors go to the cell surface is seen as an indication of a correctly folded receptor. Therefore, the major objective was here to elucidate whether the two different promoters exhibited different phenotypes in terms of cell surface expression of the GPCRs. Initial experiments were carried out in the traditional way with relevant cells grown on cover-slips, and eventually being infected with virus. However, it quickly became clear that this way was not reflecting the situation with cells grown in suspension as judged by wide field fluorescence microscopy. Therefore, cells were grown in suspension, fixed in suspension, and deposited on a slide under a cover slip.

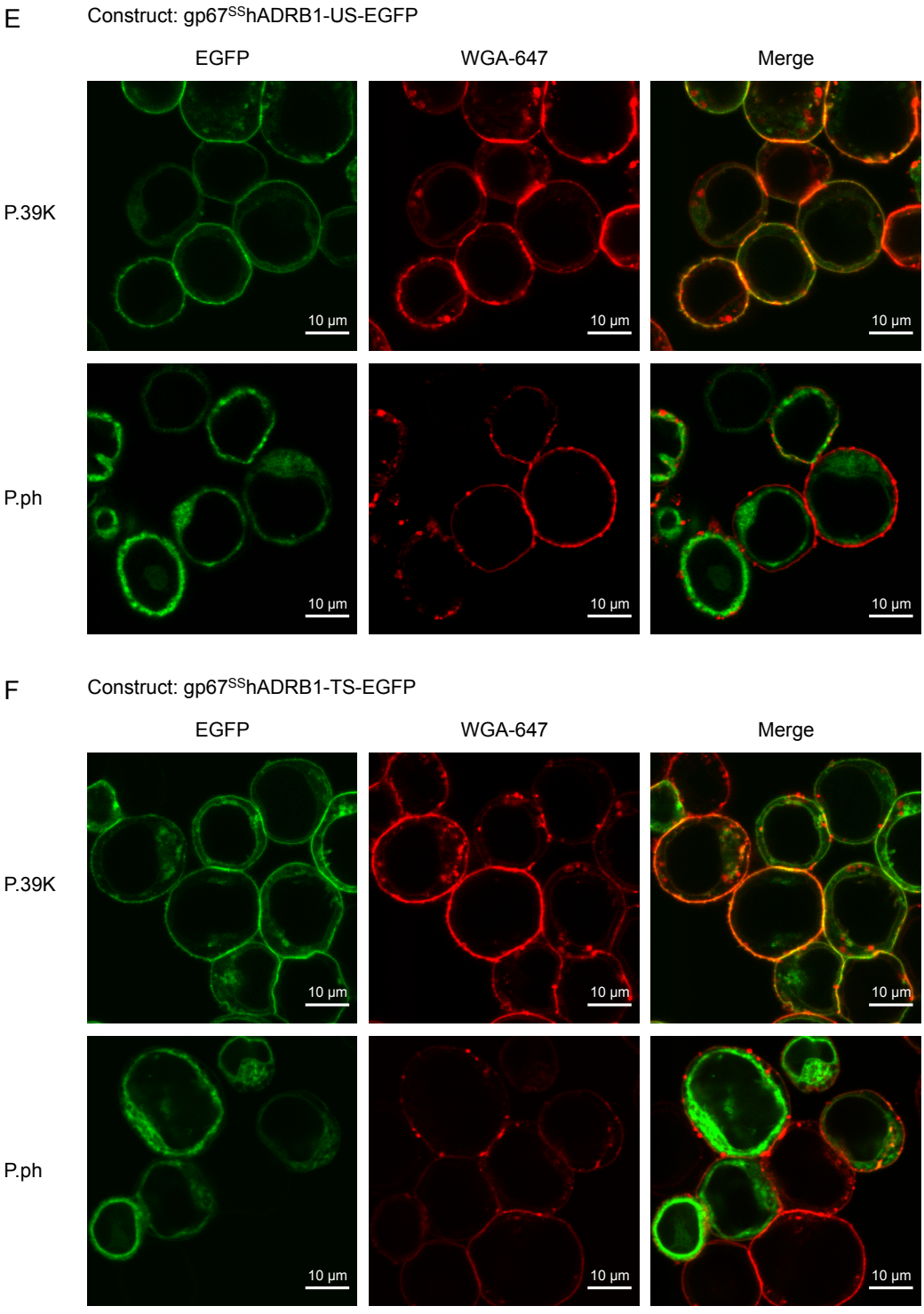
Indeed, there was in fact a remarkable difference between the two promoters (Figure 18). While the P.ph led to accumulation of protein inside the cell, the P.39K was able to produce protein that almost exclusively went to the cell membrane, as defined by the lectin Wheat Germ Agglutinin, here conjugated with Alexa-fluor-647. The difference is most pronounced for the hADRB1 constructs, including those with the gp67 signal sequence (Figure 18 A, B, E, F). For the P.39K, as mentioned, there is some intracellular protein, but it is expected since the production is a continuously ongoing process, both trafficking receptor to the surface and degrading incorrectly folded protein, all taking place inside the cell. It is very interesting to note that the presumably completely misfolded hADRB2 constructs, with and without signal sequence, shows no apparent presence of receptor molecules on the cell surface, for both P.39K and P.ph (Figure 18 C, G). This is an intriguing fact since this is also similar in appearance of the hADRB1 construct expressed using P.ph.



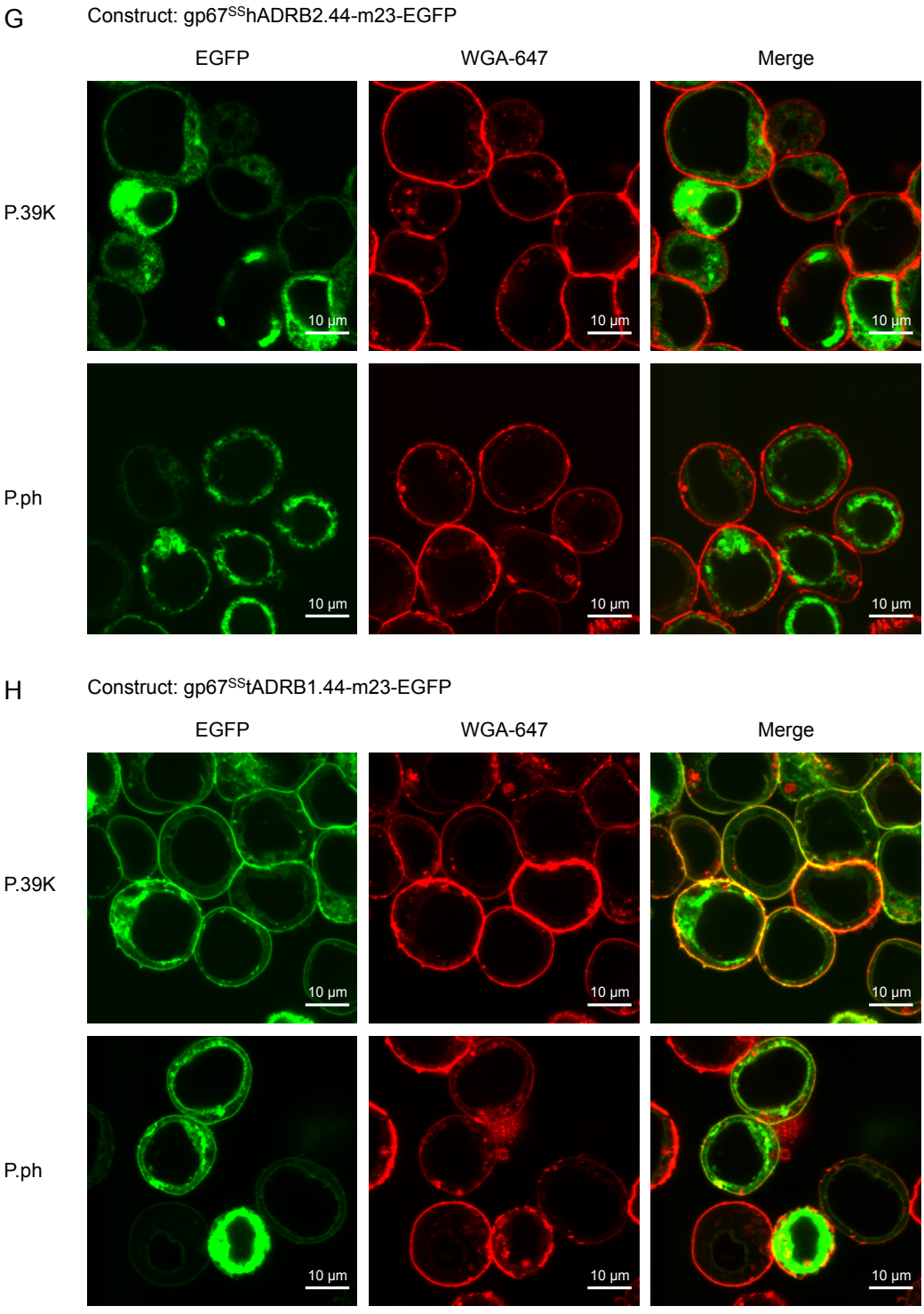
Taken together, the sub-cellular localization pattern of functional receptor is almost indistinguishable from that of the known non-functionally expressed protein, hADRB2.44-m23, when using the P.ph as promoter. On the other hand, the distribution of protein when using the P.39K as promoter is



completely different when comparing a functional protein (e.g. gp67^{SS}-hADRB1-TS), to a non-functional one (e.g. gp67^{SS}-hADRB2.44-m23), with the former being nicely located on the cell surface, and the latter being caught inside the cell.



Generally, in the set of constructs used, it can be stated that a non-functional receptor implies no surface expression. On other hand, surface expression implies a functional protein. However, it is not necessarily true that no surface expression implies non-functional protein (easily observed for the



gp67^{SS}hADRB1 derived constructs; Figure 17, Figure 18 E, F). But, if one considers only expression using the 39K promoter, it can be stated that no surface expression implies no functional protein.

Figure 18. Preferential cell surface expression of functional receptor using the 39K promoter.

Confocal micrograph of Sf9 cells infected with recombinant baculoviruses expressing GPCR constructs using the 39K promoter (**P.39K**) or the polyhedrin Promoter (**P.ph**). The cellular distribution of the various GPCR constructs is elucidated under these conditions. The construct is indicated in upper left corner of each panel, A to H. Localization of receptors was done by visualizing the C-terminal EGFP (**EGFP**) on these. The cell membrane has been stained with Wheat Germ Agglutinin-Alexafluor-647 (**WGA-647**) in order to define it in the micrographs. Sf9 were fixed with 3.7% formaldehyde in PBS.

Discussion

In the study at hand, a variety of complementary methods have been applied to analyze the effect of using different promoter in GPCR crystallization type constructs. This particular approach has been applied not only to identify ways to improve conditions for production of GPCRs and membrane protein, but also to map where in the process of over production the hurdles lie. The main techniques include SDS-PAGE with subsequent In-gel fluorescence of solubilized material and importantly also DDM insoluble material, which cannot be quantified by FSEC, which on the other hand is able to assess the state of the solubilized material in much greater detail separation by SDS-PAGE. Neither, however, are able to quantify the amount of functional receptor, which can then be done by the gold standard of functionality, radio-ligand binding. Finally, microscopy has shown that there is a clear difference in the phenotype of the intact cells when producing the receptors using different promoters. All techniques are commonly used individually, but in combination a greater resolution of the potential problem in production is achieved. In this way, it has been demonstrated by relatively simple means, that an expression system can be rationally engineered, taking a step towards the ultimate goal of producing more functional membrane protein per liter cell culture.

It has here been exploited that a C-terminal EGFP remains fluorescent even when the preceding membrane protein is incorrectly folded as previously found (Thomas and Tate 2014). The example found in this manuscript, the hADRB2.m23 constructs, also constitute perhaps the clearest case of this phenomenon, since these are probably almost completely misfolded. This is different from expression in *E. Coli* in which EGFP (or GFP) to some extent can act as a folding reporter (Drew et al. 2001, Bill et al. 2011).

Using the persistent fluorescence of EGFP even when the preceding membrane protein is misfolded in eukaryotes, together with monitoring both

DDM soluble and insoluble fraction, DDM resistant protein was observed as being very pronounced when using the polyhedrin promoter, while the cells were still able to cope with the misfolded protein in most cases, even the hADRB2 constructs, when using the weaker promoter P.39K. This is similar to and consistent with the results obtained for mutants of rhodopsin with increased tendency to fold incorrectly, found in the genetic disease retinitis pigmentosa. Here inhibition of the proteasome led to accumulation of DDM resistant species of rhodopsin, concluded as being misfolded protein (Saliba et al. 2002). Thomas and Tate (2014) referred to this as differential solubility in DDM and SDS, using DDM as the folding-reporter. However, it has to be applied with some care, as observed here in this manuscript, since even the hADRB2.m23 constructs could be observed as solubilized by DDM, albeit weakly, but not much less than the active hADRB1 counterparts which were found to be active, by specific ligand binding.

In another study a large number of yeast membrane proteins were screened for expression and quality in yeast, and it was found that efficiency of a given detergent, including DDM, was not a good predictor of monodispersity in FSEC experiments (Newstead et al. 2007). On the other hand a different idea can be formed here: The higher degree of solubilzation that can be achieved by DDM the better is the FSEC profile for a given protein, here varying expression conditions and not the detergents. This is what has been observed here, for all pairs of a given protein expressed either using P.39K or P.ph, and it seems likely that this would also apply to other membrane-proteins

One of the most striking and clear observations from the work presented here is the fact that a membrane protein can indeed and indisputably be expressed on the cell surface of insect cells, here Sf9. While there has been speculation as to why Sf9 cells produce membrane protein primarily located intra-cellularly, it is now clearly documented that in the context of

baculovirus mediated expression it is due to the choice of a late strong promoter such as the polyhedrin promoter, since choosing the 39K promoter results in membrane protein located on the cell surface. Whether or not this improvement is due to less mRNA, earlier synthesis in terms of viral lifecycle or a combination of the two, remains to be shown.

The afore-mentioned consistency with the previous findings for rhodopsin also extends to the cellular localization of the recombinantly expressed protein, where using the P.ph to drive expression leads to accumulation of protein inside the cell, rather than being on the cell surface. This was also observed for the Neurotensin receptor when produced in Sf9 cells using a baculovirus with the P.ph promoter, where about 2/3 of the active protein remained inside the cells (Xiao et al. 2013).

As observed, expression using the polyhedrin promoter, still gives the largest amount of active protein in membranes, but not the highest quality as judged by FSEC. The high similarity of the main FSEC peak of the hADRB1 expressed under P.39K and P.ph, respectively, in the construct with gp67 signal sequence, but with the latter leading to higher amount of active receptor in the crude membrane, is quite puzzling. This needs further investigation, but is beyond the scope of this study. It could be that part of receptor expressed using P.ph, is in an active ligand binding state, but resistant to solubilization, perhaps caused by some form of aggregation rendering it DDM resistant, perhaps even aggregating with other proteins.

Although the goal of ultimately improving membrane protein yield has not been reached, it has here been shown that microscopy can potentially serve as an extremely fast screening methodology for weeding out non-functional constructs when using the 39K promoter. As described earlier, surface expression and functionality of the protein is correlated in almost the simplest possible way when using this promoter: Surface expression is necessary and sufficient for functional receptor in this set of constructs. Since

surface expression was indeed achieved for functional constructs, one can ignore considering other factors that could lead to hampered surface expression, such as a non-functional or destroyed ER export motif of which there are several potential possibilities (Wu 2013). This is, of course, a potential issue that could lead to falsely weeded out constructs.

It must be noted that for the constructs used here, glycosylation sites are not present in the hADRB1 and the tADRB1 derived constructs, isolating these receptors from effects arising from such modifications, whereas a site is present in the hADRB2 derived constructs, as mentioned earlier. As such the receptors considered here, constitute a simplified system, which is therefore an ideal model system for optimizing production of GPCRs or membrane proteins, purely regarding the expression system and not the proteins. It has to be noted, that for this purpose, the tADRB1 constructs are less suitable for this task due to the easy of production of these, which can mask improvement as demonstrated by the FSEC profiles almost identical appearances under either P39K or P.ph.

It will be interesting to see how expression of recombinant membrane proteins can be improved even further, using the knowledge obtained here. Obvious routes to take include engineering of the cells, with emphasis on improving their ability to handle larger amounts of protein going through the SEC –translocon. This could potentially be achieved by using CRISPR technology to tune alter the secretory pathway. However, it is clear that no simply fix is available here for this complex system.

Materials and Methods

Plasmids

All plasmid used in this study were derived from the plasmid pFL, which is derived from the commercial pFastbac-dual (Invitrogen). The open reading frame encoding mCherry was cloned into pFL in the expression cassette using the p10 promoter in this plasmid. The resulting plasmid was called pFLmCh.ph. The DNA encoding the hr5-IE1 promoter was sub-cloned by PCR from the plasmid pIEX-1 in to pFLmCh.ph, in which it exchanged the polyhedrin promoter, whereas the 39K promoter was isolated by PCR directly from the recombinant baculovirus genome, also exchanging the polyhedrin promoter, resulting in the plasmids, pFlmCh.IE and pFLmCh.39K, respectively. By sub-cloning mediated by restriction enzymes, all GPCR – EGFP fusion protein open reading frames, were cut and pasted into pFLmCh.IE, pFLmCh.39K and pFLmCh.ph.

Baculovirus production and expression experiments

Sf9 cells were cultured in SF4 Insect Cell Express medium (Bioconcept, CH). The Recombinant baculoviruses were produced according to the Bac-to-Bac (Invitrogen) manufacturer's instructions, however, using the Multibac Bacmid. Recombinant Bacmid was isolated using Qiagen miniprep kit. Recombinant baculovirus Bacmid was transfected into Sf9 cells using ExtremeGene HP(Roche) transfection reagent according to manufacturers instructions. This was done in adherent cells and this would become the p0 stock. The p0 stock was then used to infect approximately 50 ml of Sf9 cells at a cell density around 1.5 mill/ml. 2 days post infection the culture supernatant was harvested to obtain the p1 stock. Finally the p1 stock was amplified in 250 ml of Sf9 cells at a cell density around 1.5 mill/ml to obtain the p2 stock, which was used in in subsequent experiments. For the expression experiments the virus was titered by qPCR according to (Hitchman et al.

2007), and an amount corresponding to an M.O.I. of 5 was added in each condition (M.O.I., multiplicity of infection, here referring to viral genomes per cell).

For each construct, the cells were grown in 50 or 500 ml vessel at a density of approximately 2 mill/ml, as adjusted by counting using a neubauer chamber. 72 hours post infection, the cells were harvested, aliquoted and stored at -20 C, for further analysis.

Solubilization, in-gel fluorescence

Cell pellets were thawed on ice and ice-cold solubilization buffer (containing 50 mM Hepes pH 7.5, 150 mM NaCl, including Roche complete EDTA free protease inhibitor) was added. DDM was added and the total volume was adjusted such that the total volume was a 10% percent of the culturing volume containing 1% DDM. One sample was taken out immediately and added SDS gel sample buffer which would be the cellular fraction named "Total". The samples were then left solubilizing for 1 hour at 4C using a rotating device for constant mixing. Upon completion of solubilization samples were ultra-centrifuged at 150'000 x g for 20 min at 4C, in order to sediment DDM insoluble material. The supernatant was then either stored at 4C for further analysis or a sample was taken out and immediately added SDS sample buffer to obtain the solubilized fraction named "solubilized".

The soluble fraction and the total fraction were separated by SDS-PAGE with 12 % Trupage gels and TEA-Tricine buffer system (Sigma-Aldrich), and the gel resolved species were imaged using EGFP fluorescence in an Amersham imager AI600(GE), with the light and filter setting suitable for EGFP. All gels were imaged at the same time. Virtually identical results were obtained when scrambling gels, indicating that the excitation light intensity was sufficiently homogenous in terms of intensity at the whole imaged area.

Fluorescence-size exclusion chromatography

DDM solubilized samples described immediately above were filtered using 0.22 µm table top micro centrifuge filter (Millipore) in order to remove large particles that could damage chromatography column. Upon filtration samples were added NaCl to a final concentration of 500 mM. Samples were then analyzed by being run on a Tosoh SuperSW3000 analytical gel filtration column using an Äkta Ettan HPLC system (GE) equipped with an auto-sampler and fluorescence detector set to excite EGFP(485nM), with emission detected at 520nm. The mobile phase was 50 mM Hepes pH 7.5, 500 mM NaCl, 0.05% DDM. The column void, V0, is 2ml, while the total column volume, Vt, is 4.5 ml as estimated void peak and small molecule peak using a Biorad protein standard. The run was continued for 6.5 ml for each sample. All runs were performed at 4C.

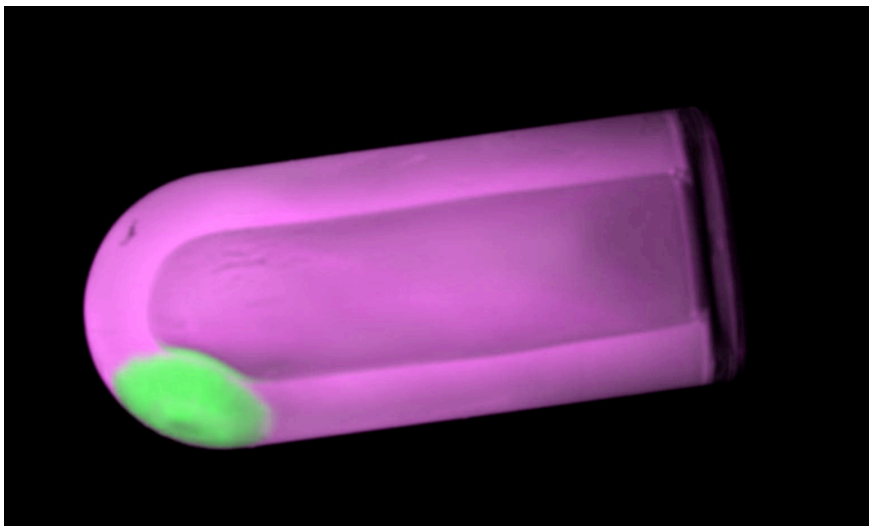
Radioligand bindings assays

Cells were thawed and resuspended in Phosphate-buffered Saline supplemented with Roche complete EDTA free protease inhibitor, and homogenized using a 29G needle with a syringe. For each reaction an amount corresponding to 100'000 cells were used. The crude cell membranes were incubated for one hour with the radioactively labelled beta blocker, [3H]-DiHydroAlprenolol([3H]-DHA). At the end of the incubation, the membrane suspension was applied on a Whatman Grade GF/B Glass Microfiber Filters(GE 1821-025), using a vacuum manifold(Millipore, XX2702550 | 1225). After application, the membranes were washed with 15ml of ice cold PBS, briefly dried on the filter, and transferred to a scintillation vial and added 10 ml of scintillation fluid (ScintLogic U, SG-BXX-01). The samples were incubated at least 16 hours, vortexed, and transferred to a scintillation counter for quantification.

Widefield and confocal microscopy

All samples were harvested from 1 ml culture and washed 2 twice in PBS at pH6.2, by centrifugation at 400g for 3 minutes and re-suspension in said buffer. For wide field fluorescence microscopy the cells were stained with Hoechst, and the cells were finally re-suspended in 10% of the culture volume to achieve a higher density. A sample was deposited on a microscopy slide under a cover slip. This was done in a pairwise manner and cells were imaged immediately. For confocal microscopy, the washed cells were fixed in 3.7 % formaldehyde in PBS pH 6.2. After fixation, the cells were washed in PBS pH 6.2, and stained with wheat germ agglutinin conjugated to Alexafluor-647. The cells were then washed in PBS pH 6.2. Immediately after staining the cells were applied to a microscopy slide under a cover slip, which was fastened and sealed with nail polish. Confocal images were obtained on a Leica SP5 confocal microscope, using a 488nm laser to excite EGFP, with the photomultiplier tube set to 505-530nm, while Alexa-fluor 647 was excited using a 633nm laser, with the photomultiplier tube set to 650-700nm.

4. Large scale production of the human Beta1-Adrenergic Receptor constructs



Chapter 4 front page. Ultracentrifuge tube showing the first step of a receptor purification.

Fluorescence image of a ultracentrifuge tube with the pellet of EGFP-tagged (green) receptor being seperated from the mCherry (pink) containing supernatant.

Introduction

Originally the turkey beta1-adrenergic receptor was affinity purified in a two step manner starting with a nickel IMAC enrichment step followed by alprenolol ligand affinity chromatography (ALAC) (Warne et al. 2003). Therefore, it was natural to adopt this approach.

The original goal of this project was to characterize the human beta1-adrenergic receptor by NMR as carried out on the turkey beta1-adrenergic receptor (Isogai et al. 2016). In this study, the valines in the protein were selectively labelled, and the individual peaks in the 2D [^{1}H, ^{15}N] correlation spectrum, were assigned by mutating each of them into a non-valine residue, either alanine or isoleucine. This approach would entail making viruses, producing the receptor and purifying it, for at least a significant fraction of the 32 valines found in the hADRB1-TS construct, in order to obtain enough assigned peaks. Depending on the similarity of the spectra obtained for the turkey and the human receptors, one possibility could be to assign similar and interesting residues first, and let the peaks with more doubtful identity wait until later, if assigning them at all. In this way one could minimize the work, and not waste time and resources. In any case, at least 10 to 15 peaks would need to be assigned. Keeping in mind that the yield of the hADRB1-TS is still significantly lower than that of the tADRB1.44-m23, requiring around 6-8 liters of cell culture to obtain 1mg of labelled protein, this receptor is very challenging to work with in NMR studies. Therefore, doing the required scaling up would severely heighten the workload at all steps ranging from expression of the receptor to the final purification step.

A popular purification strategy is the one developed in Raymond Stevens' lab using the 6xHis-tag (see for example Wu et al. 2012). This purification method includes several lysis steps with ultracentrifugation, which is followed by several washes of membranes in a buffer containing 1M NaCl, to wash away membrane bound protein, again using ultracentrifugation. Finally, after

solubilization, the receptor is isolated on Talon IMAC Resin, the cobalt counterpart of nickel-based IMAC.

It is obvious that such a purification method is time and labor consuming, to the point where the feasibility of a project with a low yield protein must be questioned. However, an alternative approach had also been carried out successfully for the purification of recombinantly produced rhodopsin, using an antibody for the bovine rhodopsin C-terminus termed 1D4 with the sequence TETSQVAPA (Mackenzie et al. 1984, Reeves et al. 2002). Here, the homogenized cell pellets are solubilized directly after removal of unsolubilized material by ultracentrifugation (Standfuss et al. 2007). Therefore, on the positive side are an extremely simple purification scheme and minimal labor, due to the high specificity of the 1D4 antibody. There are several drawbacks of this method, including relatively long incubation of solubilized material with the antibody-resin conjugate and costly production of this resin.

An attractive alternative to the above-mentioned strategies was the slightly more recently developed Strep-Tactin purification system. Here an engineered form of streptavidin binds an artificial tag termed strep-tag II with the sequence WSHPQFEK. The system has undergone several iterations of evolution of both the tag and streptavidin, and started out with the Strep-tag (WRHPQFGG) (Schmidt et al. 1996, Voss and Skerra 1997, reviewed by Skerra and Schmidt 1999). Due to the artificial nature of this tag, which was evolved to bind streptavidin through phage display of a random peptide library, as mediated by random DNA library, the streptavidin or Strep-Tactin exhibits an extremely low nonspecific binding when the tag is eluted off either of them. (Schmidt and Skerra 2007). This is of course different from IMAC, which has a much higher nonspecific background depending on the yield of the protein being purified and the amount of resin used, and might carry over small amounts of impurities, for example proteases, which should always be minimized, to ensure sample stability which is extremely important for the

inherently long NMR experiments. In addition the IMAC requires elution with several hundred milli-molar imidazole, which can act as a denaturant.

The Strep II tag, therefore, offers several advantages over the other tags mentioned above, and purification with this tag has therefore been tested and optimized for purification of a GPCR of relatively low yield.

Results

Development of purification using Strep tag II and Strep-Tactin

The first expression and purification experiments using the strep II-tag were done using three different constructs expressed in Hi5 cells using the Baculovirus system. The hADRB1-US receptor construct was tested along with two versions of this receptor, in which a short and a long form the AcMNPV gp67 signal sequence, respectively, was included. The short form was called gp67^{SS}, while the longer form was called gp67L^{SS}. The resulting constructs were then hADRB1-US, gp67^{SS}ADRB1-US, gp67L^{SS}ADRB1-US. The latter two were included since signal sequences, and in particular the one derived from gp67, have previously been shown to improve expression of membrane proteins in insect cells (Guan et al. 1992, Korepanova et al. 2009, Kofuku et al. 2012). In this setup, the viral backbone utilized was the Multibac derivative, EmBacY (Berger et al. 2004, Trowitzsch et al. 2010), a version that has an EYFP expression cassette integrated in the genome. The turkey ADRB1.44-m23-6His was also included as a reference, since it constitutes an example of a well-expressed GPCR. The initial constructs can be seen in Figure 19.

Different amounts of viral suspensions were tested for expression, and cells were harvested at 48 hours post infection. The DDM solubilized samples were analyzed by western blotting with a Strep-Tactin-AP (alkaline-phosphatase-conjugate) for detection of strep II tagged protein. The resulting blot showed no signal. However, the marker was detected (data not shown).

The conditions using the highest titer were quantified using single point radio ligand binding. This would show whether there was active receptor present in the cells. As seen in Figure 20, there is clearly active receptor present in all cases. It is immediately demonstrated here that the yield of the ADRB1-US is markedly lower (4-5 times) than that of the tADRB1.44-m23 (as previously found, Florian Brueckner, personal communication). Hence, it

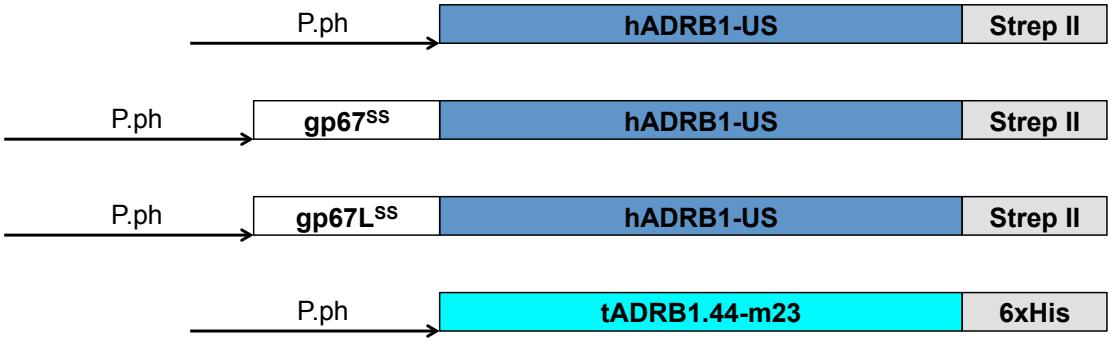


Figure 19. Overview of constructs used in the test of the single Strep tag II.

Promoter, signal sequence, GPCR coding region and Strep tag II is indicated in the relevant section of the expression cassette for each construct. Constructs are not drawn to scale.

could immediately be out-ruled that the lack of signal in the western blot, occurred due to lack of receptor expression.

It is of course very interesting to note that the inclusion of a signal peptide has a profound effect on the expression level, increasing the amount by a factor of more than 2. This is, in itself, a major improvement of the characteristics of the engineered hADRB1-US receptor.

Knowing that all viral constructs did indeed produce functional receptors, a western blot was re-done. In order to ensure the missing signals from the hADRB1-US constructs were not caused by inefficient western transfer, which can be difficult with membrane proteins, a GPCR with a dual Strep II tag (known as the Twin-strep tag) positioned the C-terminal end of a C-terminal EGFP relative to the receptor was included as a known control. The blot is seen in Figure 21 (A). The marker and the Twin-Strep tagged GPCR (control) are clearly producing visible band. The control GPCR sample is seen as multiple bands with the two most prominent bands at apparent masses of around 50kD and 25kD, corresponding to the full length construct and a cleaved off EGFP still having a strep II tag (dual). This confirms that the western transfer did work properly, but in order to rule out faulty western transfer completely, caused by a receptor specific effect, another blot was done using an antibody specific for the hADRB1 receptor. Figure 21 (B) clearly

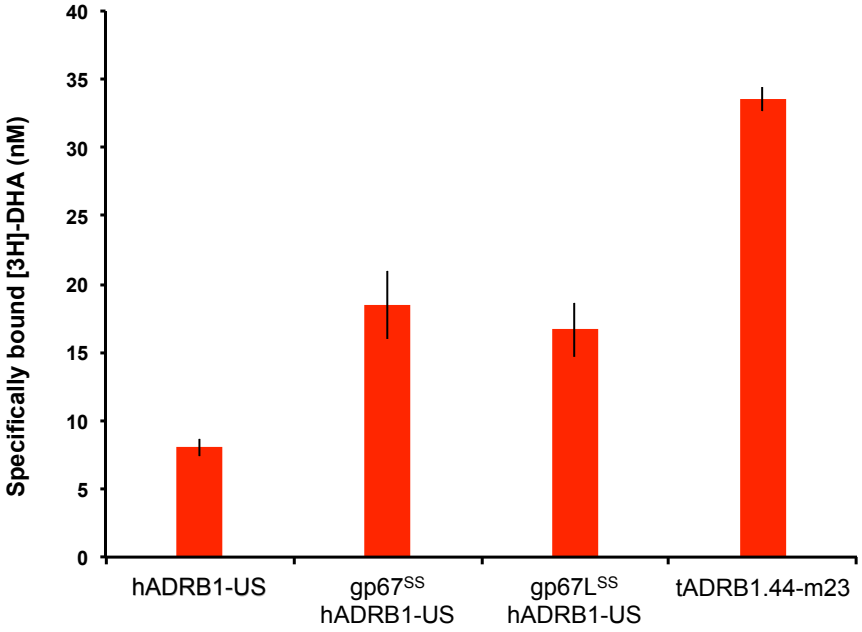


Figure 20. Expression of various ADRB1 constructs from human and turkey in High Five cells using recombinant baculoviruses.

Single point saturation binding in crude membranes from High Five cells containing overexpressed from approximately 100'000 cells per sample. Membranes were incubated with 100 nM [3H]-DiHydroAlprenolol ([3H]-DHA), and isolated on a glass fiber filter. Error bar shows standard deviation (n = 3).

shows that a band is present at the correct size, validating that the western transfer of these construct was working at least with some efficiency. The immediate conclusion following these observations was that either the Strep II tag is not detectable by the Strep-Tactin-AP conjugate, or the Strep II tag is simply cleaved off the receptor by endogenous proteases of the baculovirus infected cells.

A dot blot was made with selected samples of solubilized and total material, to isolate the western detection from potential losses of protein during SDS-PAGE and western transfer. Figure 21 (C) shows the dot blot. Again, the marker and the GPCR-EGFP-Twin-Strep construct clearly appear as visible dots, while the two solubilized samples show no apparent presence. However, the total material from the gp67^{SS}hADRB1-US-StrepII does show a faint dot. This indicates that the receptor is present in a Strep-Tactin-AP

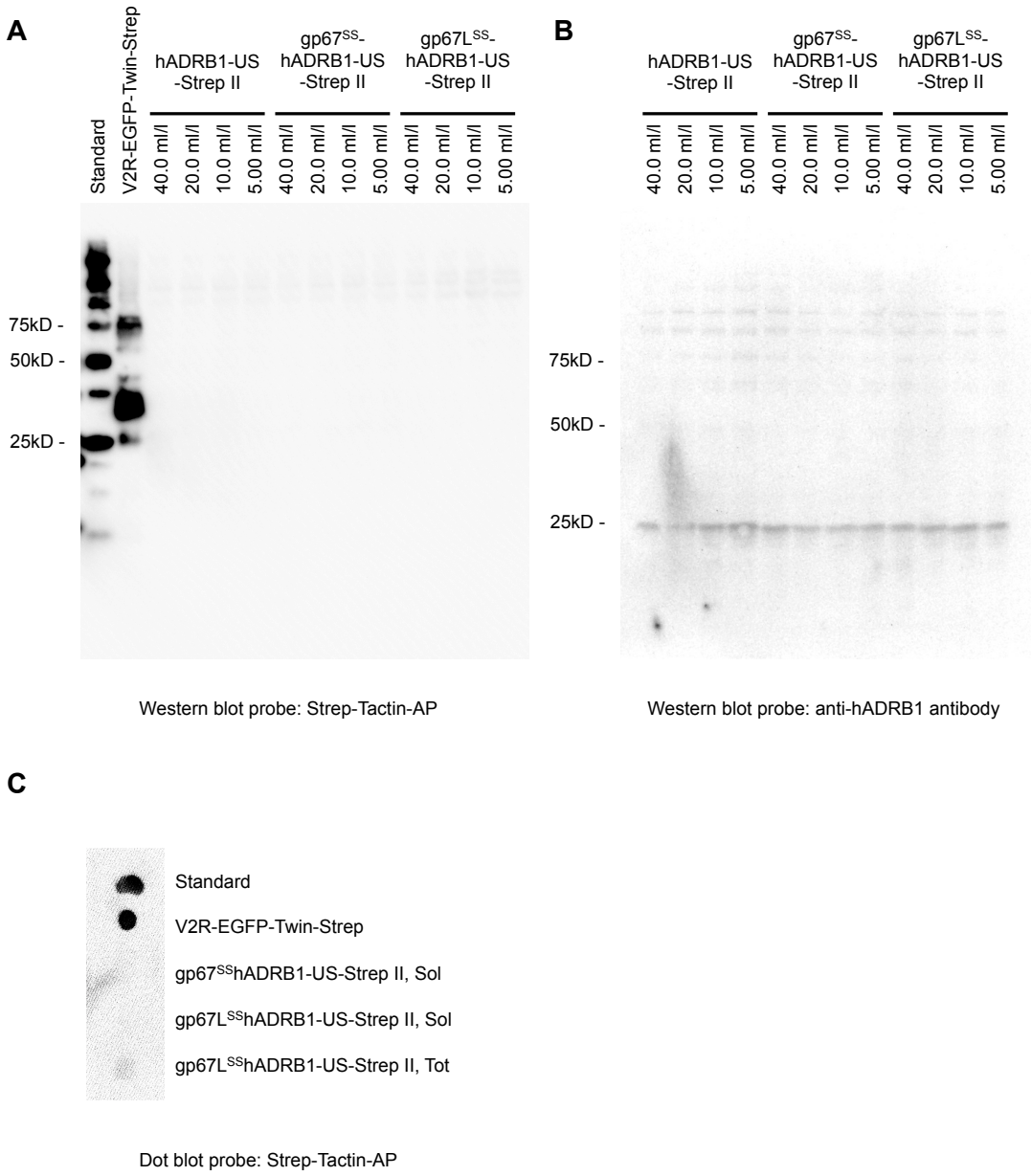


Figure 21. Blots of various hADRB1-US constructs shows presence of these.

Different hADRB1-US derived constructs on different types of blots.
(**A**) Western blot probed with Strep-Tactin-AP, a Twin-Strep (V2R receptor fused to Twin Strep tagged EGFP). (**B**) Western blot probed with rabbit anti-human ADRB1 antibody (immunizing peptide CTVWAISALVSFLPILMHWWRAESDEARRCYNDPKCCDFVTNRA-YAIASS). (**A**, **B**) Constructs are indicated above blot. Different amounts of virus per liter Sf9 cell culture was used and is indicated above each lane. (**C**) Dot blot of select samples (all 20 ml viral suspension/l cell culture) from panel (A, B) as indicated, solubilized fraction (**Sol**), total cellular material (**Tot**).

detectable form. Therefore, a remaining question is whether the Strep II tag is mainly cleaved off the receptor or the lack of detection is simply due to inefficient recognition under the western detection conditions.

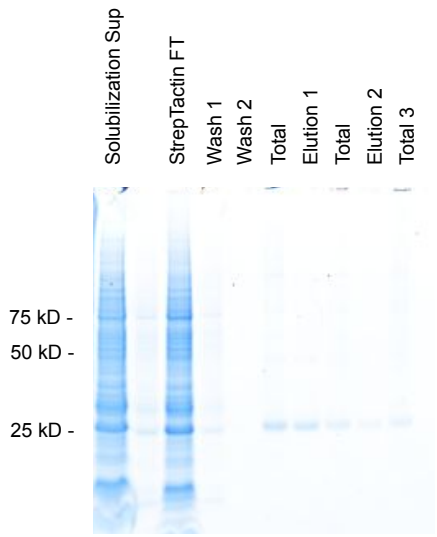


Figure 22. Purification of a single strep-tagged construct: gp67^{SS}hADRB1-TS.

SDS-PAGE stanied with Coomassie showing purification steps: 1% DDM was used for solubilization of Sf9 membranes with over expressed gp67^{SS}hADRB1-TS-strep. 2ml Strep-Tactin Sepharose was used in a gravity flow column. Fraction are as follows: solubilized membranes post ultracentrifugation (**Solubilization Sup**); Flow through of Solubilization Sup post application onto Strep-Tactin Resin (**Strep-Tactin FT**); first wash of Strep-Tactin Resin with 10 ml wash buffer (**Wash 1**); Second wash of Strep-Tactin Resin with 10 ml wash buffer (**Wash 2**); Resin + 2 ml elution buffer (wash buffer + 5 mM desthiobiotin) (**Total 1**); 2 ml elution buffer (**Elution 1**); Resin + 2 ml elution buffer post Elution 1 (**Total 2**); 2 ml elution buffer alone post elution 1 (Elution 2); Resin + 2 ml elution buffer post Elution 2 (**Total 3**).

It is known that the lack of detection can occur, but that the protein can still be purified using the strep tag II and Strep-Tactin (IBA biosciences, personal communication). Therefore, a medium scale purification was attempted. Figure 22 shows a test purification performed with an open column and 2ml Strep-Tactin-Sepharose, which should have a capacity for more than 10 mg of receptor. A Coomassie stained band can be seen in the first 2 ml elution fraction, while in the second 2ml elution fraction only a very faint band can be seen. It is also noted that a significant fraction appears to bind un-specifically on the resin along with other protein from the DMM solubilized membranes, since it is not eluted off, but still appears to be enriched on the resin. It is not a surprise that aggregation prone hADRB1-US-strep is still

enriched on the resin since it has a binding advantage over all other proteins due to it being Strep II tagged. In conclusion, this demonstrates that the construct can actually be purified to very high purity with this single strep II tag.

Development of purification using double Strep tag II (Twin-Strep) and Strep-Tactin using ADRB1-EGFP fusion proteins

The findings immediately above hinted at using a double strep Strep II (known commercially as the Twin-Strep) tag could be a way of recovering more protein. Especially, the fact that a Twin-Strep tagged protein was detected, could be an indication that to low affinity is an issue along with the know fast kinetics of binding and dissociation of the Strep II tag with Strep-Tactin (Schmidt and Skerra 2007).

In order to facilitate fast and efficient quantification of the protein at all stages of production and purification, a version of the hADRB1-TS with a C-terminal EGFP was used. Likewise, a similar construct was used for the tADRB1.44-m23, which was used for establishing part of the purification due to its high yield and, hence, its ease of production. DDM was chosen as the default detergent for solubilization and purification, since that is one of the most commonly used mild detergent, with a proven ability to provide an environment in which membrane proteins an in particular GPCRs retain their functionality (Prive 2007).

Initial test purifications were carried out using the previously described method of directly solubilizing a homogenized cell pellets with over expressed tADRB1.44-m23-EGFP-TS receptor. A surprising result was obtained as seen in Figure 23 (A). The resulting purification led to an enrichment of the low molecular band corresponding to a cleaved off EGFP, with almost no intact construct appearing. It is noted that the flow through after loading onto the column appears to have lost the free EGFP, while also the intensity of the intact receptor band has been reduced. The directly solubilized receptor

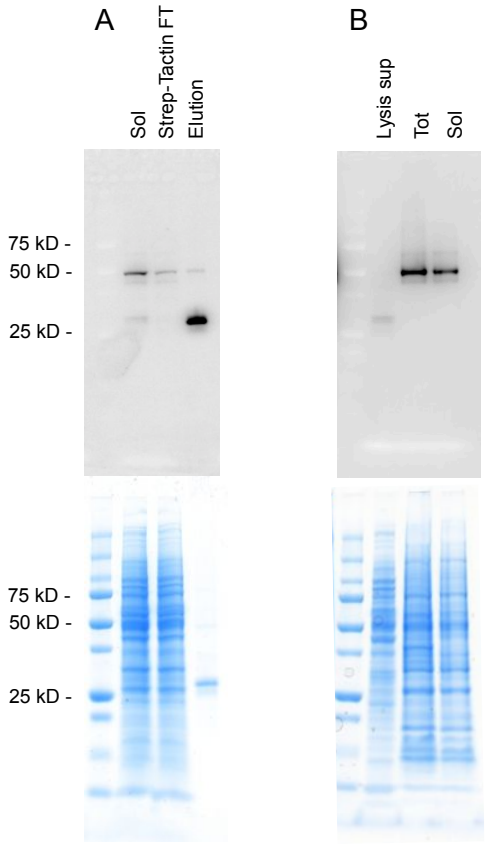


Figure 23. Enrichment of free EGFP, in initial Twin-Strep purification of tADRB1.44-m23-EGFP-Twin-Strep and optimization of this.

SDS-PAGE of purification steps visualized by in-gel fluorescence of EGFP (top part), and the corresponding coomassie stained gels (lower part). Sf9 cells containing overexpressed tADRB1.44-m23-EGFP-Twin-Strep by means infection with recombinant baculovirus were treated as follows:
(**A**) Direct solubilization of Sf9 cells containing over expressed tADRB1.44-m23-EGFP-Twin-Strep. Lanes are showing: Supernatant of DDM solubilized cells post ultracentrifugation (**Sol**); Strep-Tactin column flow through of the Sol sup (**Strep-Tactin FT**); Strep-Tactin elution with 5 mM desthiobiotin (**Elution**).
(**B**) Removal of free EGFP by introduction of a lysis step. Lanes contain the following: supernatant post ultracentrifugation of homogenized sf9 cells in hypotonic lysis buffer (**Lysis sup**); Total material after homogenization of cell membranes and addition of DDM (**Tot**); Sol tot fraction post ultracentrifugation (**Sol**).

appears to have undergone some proteolytic cleavage, most likely in the flexible linker between the receptor itself and the C-terminal EGFP-TS. However, there is clearly several fold more intact construct, compared to the free EGFP band in the solubilized fraction. The purification is an example of a particularly extreme case, and a less extreme enrichment has also been observed (data not shown).

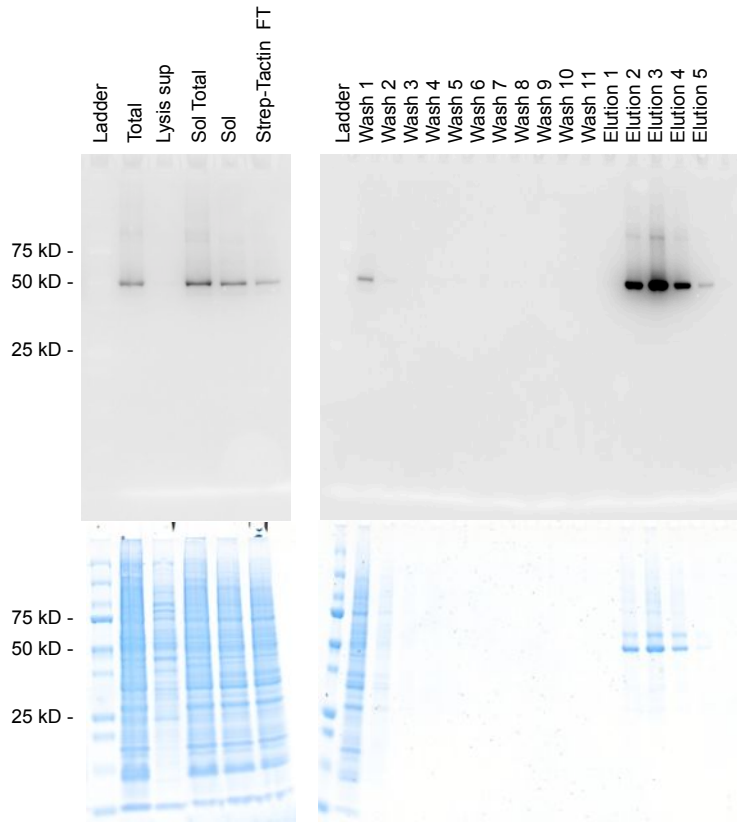


Figure 24. Strep-Tactin purification of hADRB1-TS-EGFP-Twin-Strep produced in Sf9 cells infected with recombinant baculovirus.

SDS-PAGE of purification steps visualized by in-gel fluorescence of EGFP (top part), and the corresponding coomassie stained gels (lower part). Sf9 cells containing overexpressed hADRB1.44-m23-EGFP-Twin-Strep by means infection with recombinant baculovirus were treated as indicated above each lane: Sf9 cells with over expressed receptor homogenized in hypotonic lysis buffer (**Total**); The supernatant of the lysed cells (**Lysis sup**); membranes solubilized in DDM (**Sol Total**); Supernatant of solubilized receptor/membrane post ultracentrifugation (**Sol**); Flow-through of sol sup fraction loaded onto Strep-Tactin column (**Strep-Tactin FT**); washing unbound material off the Strep-Tactin column in 2ml fraction (**Wash** + number); Elution of receptor using 5 mM desthiobiotin in 2ml fractions (**Elution** + number).

The purification result is somewhat puzzling, but a potential explanation could be the fast binding kinetics of the Strep II combined with steric hindrance of binding of the intact construct. This could be so since the N- and C-terminus of the EGFP are not far from each other. In the intact construct, a receptor in a DDM micelle, could block or slow down binding of the Twin-Strep to the Strep-Tactin molecules, on the resin. This would then give the

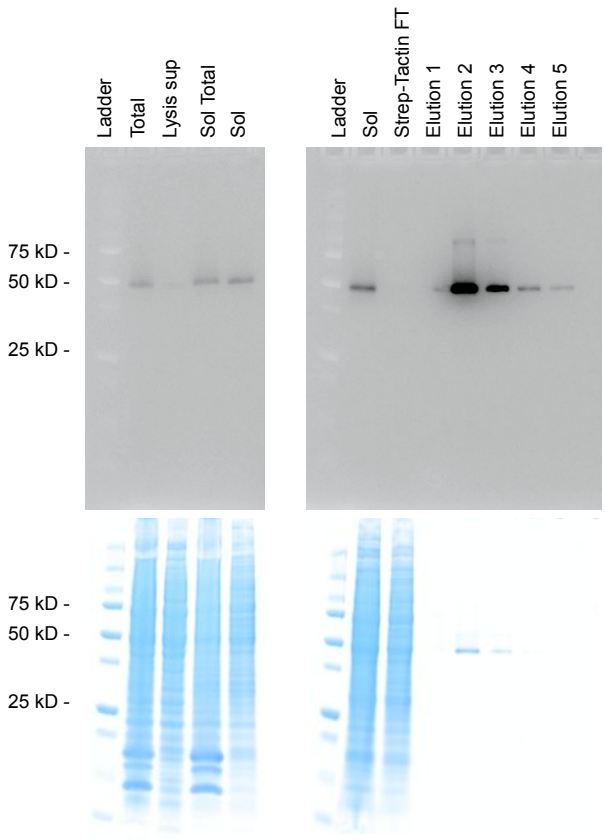


Figure 25. Strep-Tactin purification of ADRB1-TS-EGFP-Twin-Strep from HEK293 GnTi⁻ cells.

SDS-PAGE of purification steps visualized by in-gel fluorescence of EGFP (top part), and the corresponding coomassie stained gels (lower part). Sf9 cells containing overexpressed hADRB1.44-m23-EGFP-Twin-Strep by means infection with recombinant baculovirus were treated as indicated above each lane: Sf9 cells with over expressed receptor homogenized in hypotonic lysis buffer (**Total**); The supernatant of the lysed cells (**Lysis sup**); membranes solubilized in DDM (**Sol Total**); Supernatant of solubilized receptor/membrane post ultracentrifugation (**Sol**); Flow-through of sol sup fraction loaded onto Strep-Tactin column (**Strep-Tactin FT**); Elution of receptor using 5 mM desthiobiotin in 2ml fractions (**Elution** + number).

free EGFP an advantage, still having an intact Twin-Strep, but in a maximally available form.

In order to avoid this effect, the simplest solution was to introduce a lysis step in hypotonic buffer, and pellet the crude membranes by ultracentrifugation while discarding the supernatant containing the free EGFP. The result is seen in Figure 23 (B). It is clearly seen that such a step is efficiently removing the free EGFP to a point where it is not detectable. This,

of course, slows down the purification. However, a single lysis step is still manageable to perform even with a large amount of cells/cell pellets.

Using this strategy the purification of the receptor could now be done and can be seen in Figure 24, showing an examples of a purification of the hADRB1-EGFP-TS construct produced using a recombinant baculovirus and Sf9 insect cells. It is seen that some protein still remains in the Strep-Tactin column flow through.

A similar purification was performed using a stable HEK293 GnTi- cell line expressing exactly the same construct, the hADRB1-TS-EGFP-Twin-Strep (Figure 25). In this case the protein was efficiently capture from the solution, showing no apparent presence in the flow through.

This discrepancy in the ability of the Strep-Tactin column to capture all the protein and leave little to nothing in the flow through was observed to vary from purification to purification. Factors that could influence this include proteolysis of the tag, detergent used and saturation of the column. However, the hADRB1-TS-EGFP-Twin-Strep, does not produce enough receptor, as estimated by A280 absorption, in either system to saturate a column, which has a capacity of approximately 5mg, of the receptor per ml of resin.

Purification and characterization of hADRB1-TS-Twin-Strep construct

Since the constructs used in the previous section all have the EGFP fused to the respective C-termini, further studies including NMR and crystallization would require cleaving off this EGFP. A cleavage step would incur further protein losses, and require a lengthy incubation, since such a step would most likely have to be carried out at 4 degrees Celsius. Therefore, testing the purification of receptor constructs without an EGFP but with a Twin-Strep-tag was performed.

Also, assume the hypothesis mentioned previously is correct: A GPCR in a detergent micelle can interfere with the availability of the Twin-Strep, due to tight spatial constraints especially regarding the Strep tag II closest to the

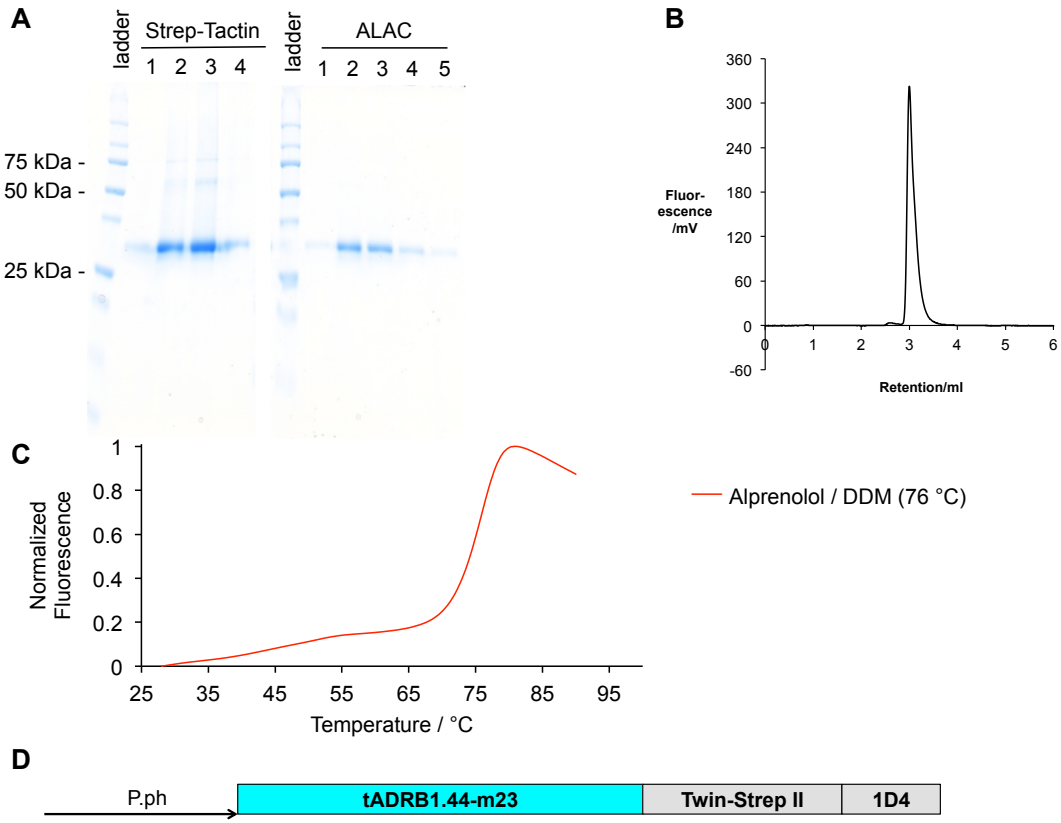


Figure 26. Purification and characterization of the turkey ADRB1.44-m23-TS-Twin-Strep-1D4 construct.

Sf9 cells were infected with recombinant baculovirus using the polyhedrin promoter to drive expression the Sf9 infected with recombinant baculovirus using the polyhedrin promoter tADRB1.44-m23-Twin-Strep-1D4. Membranes containing the receptor were solubilized in DDM. (**A**) Purification steps are indicated above lanes including Strep-Tactin and ALAC; gels stained with coomassie. (**B**) FSEC analysis of ALAC purified receptor, using fluorescence of tryptophan (excitation 280 nm; emission 350 nm). (**C**) Melting temperature with ligand as indicated, determined by thermal denaturation by means of the CPM assays. (**D**) Overview of the tADRB1.44 expression cassette.

EGFP, and thus explain the preferential binding of EGFP-Twin-Strep to the Strep-Tactin Resin, as seen in Figure 23. Having a free Twin-Strep placed directly on receptor C-terminal end, would allow for efficient binding of the Twin-Strep to the Strep-Tactin Resin, which was observed for the proteolytically cleaved off EGFP. An immediate concern was that the micelle could still block or interfere with the binding, which is why a 7 amino-acids flexible linker (GSGGGSA) was introduced between the receptors C-terminal ends and the first Strep tag II in the Twin-Strep. In addition the 1D4 tag was

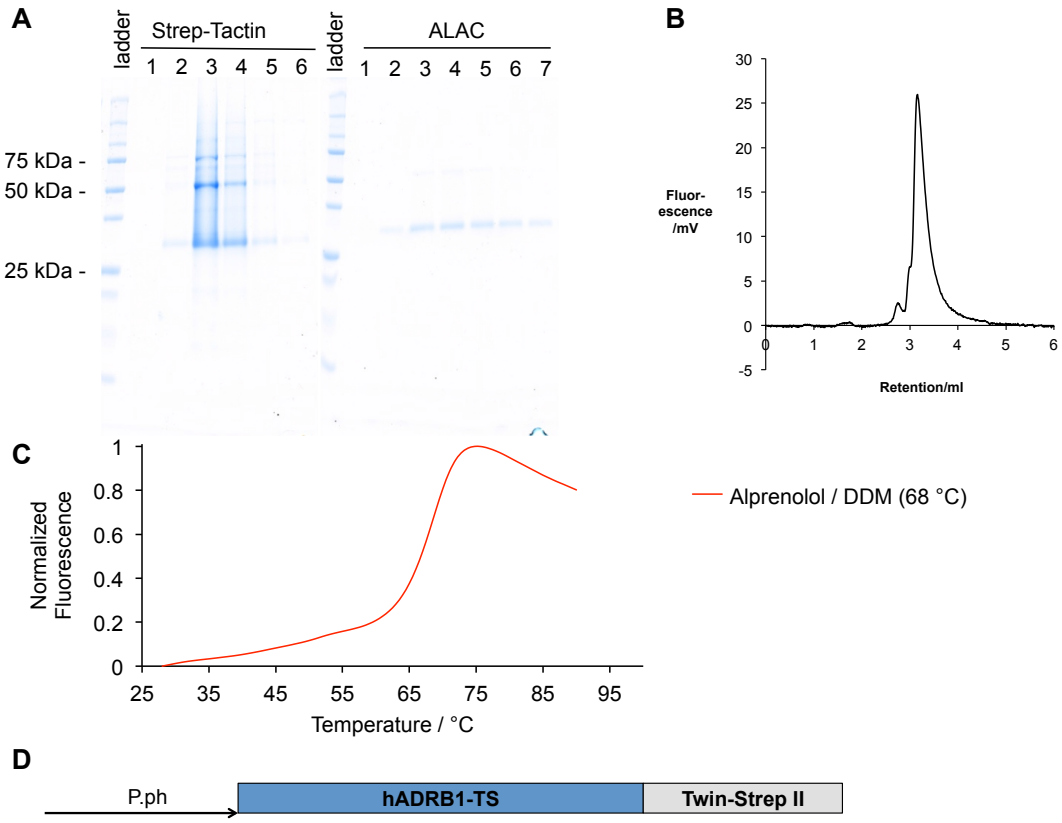


Figure 27. Purification and characterization of the hADRB1-TS-Twin-Strep construct.

Sf9 cells were infected with recombinant baculovirus using the polyhedrin promoter to drive expression the Sf9 infected with recombinant baculovirus using the polyhedrin promoter hADRB1-TS-Twin-Strep. Membranes containing the receptor were solubilized in DDM. (**A**) Purification steps are indicated above lanes including Strep-Tactin and ALAC; gels stained with coomassie. (**B**) FSEC analysis of ALAC purified receptor, using fluorescence of tryptophan (excitation 280 nm; emission 350 nm). (**C**) Melting temperature with ligand as indicated, determined by thermal denaturation by means of the CPM assays. (**D**) Overview of the hADRB1-TS expression cassette.

included at the C-terminal end of the Twin-Strep-tag, as a back-up purification-tag.

Due to the length of the Twin-Strep tag, 26 residues, including two strep II tags and a 10 residue flexible glycine serine linker, a potential problem could arise regarding crystallization, since this is a considerable number of flexible residues. However, this is less of a problem in the NMR studies, especially due to the selective labelling planned for the hADRB1-TS construct.

Purification using this tag, the Twin-Strep tag, was immediately successful and examples of purifications can be seen in Figure 26 (A), Figure

27 (A), Figure 28(A), showing purifications of the tADRB1.44-m23-Twin-Strep-1D4, hADRB1-TS-Twin-Strep, and typical yields would be in the range of 3 mg and 1mg per liter insect cell culture at 2mill cell/ml, respectively.

Alprenolol Ligand Affinity Chromatography and characterization of the hADRB1-TS-Twin-Strep receptor construct

A second purification step was included to ensure that only active receptor was present in the preparation. The chosen step was identical to the one chosen used to purify the tADRB1 crystallization constructs (Warne et al. 2003, Warne et al. 2009) and originally developed by Caron et al. (1979) – Alprenolol Ligand Affinity Chromatography (ALAC). The method was instrumental in the identification of the genes of the beta-adrenergic receptors (Dixon et al. 1986). The principle is based on the antagonist alprenolol bound by a flexible linker to a sepharose matrix. Upon binding of the receptor to the affinity matrix the receptor can be eluted off this matrix with either alprenolol or another ligand for the adrenergic receptors. Purifications are seen in in Figure 26 (A), Figure 27 (A), right most side, of the tADRB1.44-m23, the hADRB1-TS constructs, respectively. The receptors were eluted off the ALAC resin with 0.1 mM alprenolol, with typical yields for the tADRB1.44-m23, the hADRB1-TS of around 1mg, 0.25 mg, respectively, per liter cell culture at 2mill cell/ml.

The hADRB1-TS was analyzed by FSEC. Figure 26 (B), Figure 27 (B), is showing the size exclusion peaks of very similar appearances, of the tADRB1.44-m23 and the hADRB1-TS, both with sharp mono-disperse peaks and little-to–almost no aggregated or higher molecular mass material.

The receptors were characterized by fluorescence detected thermal denaturation (CPM assay). Comparing the hADRB1-TS (Figure 27 (C)) construct to the tADRB1.44-m23 construct (Figure 26)), it is seen that the melting temperatures are 68 °C and 76 °C, respectively. Even though, the melting temperature of the hADRB1-TS is not quite as high as that of the

tADRB1.44-m23 construct, the former clearly approaches the latter in terms of stability.

Thus, it is clear that the hADRB1-TS construct resembles the tADRB1.44-m23 construct biophysically on several counts.

Purification and characterization of the a double tyrosine mutant, hADRB1-TS (A244Y, L377Y) – the hADRB1-TS-YY-Twin-Strep

In the study of the tADRB1 receptor it was necessary to revert two particular thermo-stabilizing mutations in order to achieve an appreciable effect on the receptor in response to the agonist isoprenaline (Isoproterenol) and the G-protein mimic Nb80 (Isogai et al. 2016). The mutations in question are the Y227A, and the Y343L in the particular tADRB1 construct, which corresponds to Y244A and Y377L in the hADRB1 receptor. The equivalent residues in the hADRB2 have both been shown to be involved in its activation (reviewed by Rosenbaum et al. 2009, Goncalves et al. 2010, Rasmussen et al. 2011). Therefore, in order to obtain a receptor that should, in theory, be able to enter an active state, these two residues were reversed in the hADRB1-TS construct to obtain the construct hADRB1-TS-YY. The construct was added the N-terminal Signal sequence from the AcMNPV gp67 protein, since that was shown to increase expression significantly (Chapter 3), while also having the C-terminal Twin-Strep.

The Strep-Tactin and ALAC purification steps can be seen in Figure 28 (A), again showing a highly pure receptor preparation in each step, with typical yields of 0.7 mg and 0.15 mg per liter Sf9 insect cell culture, respectively.

It is noted that, as previously found for a thermo-stabilized tADBR1 construct (Isogai et al. 2016) with the equivalent tyrosines re-introduced, the melting temperature of this construct was lessened notably dropping from from 68 °C in the hADRB1-TS construct (Figure 27 (C)) to 61 °C for the hADRB1-TS-YY construct also purified in DDM (Figure 28 (C)). On the

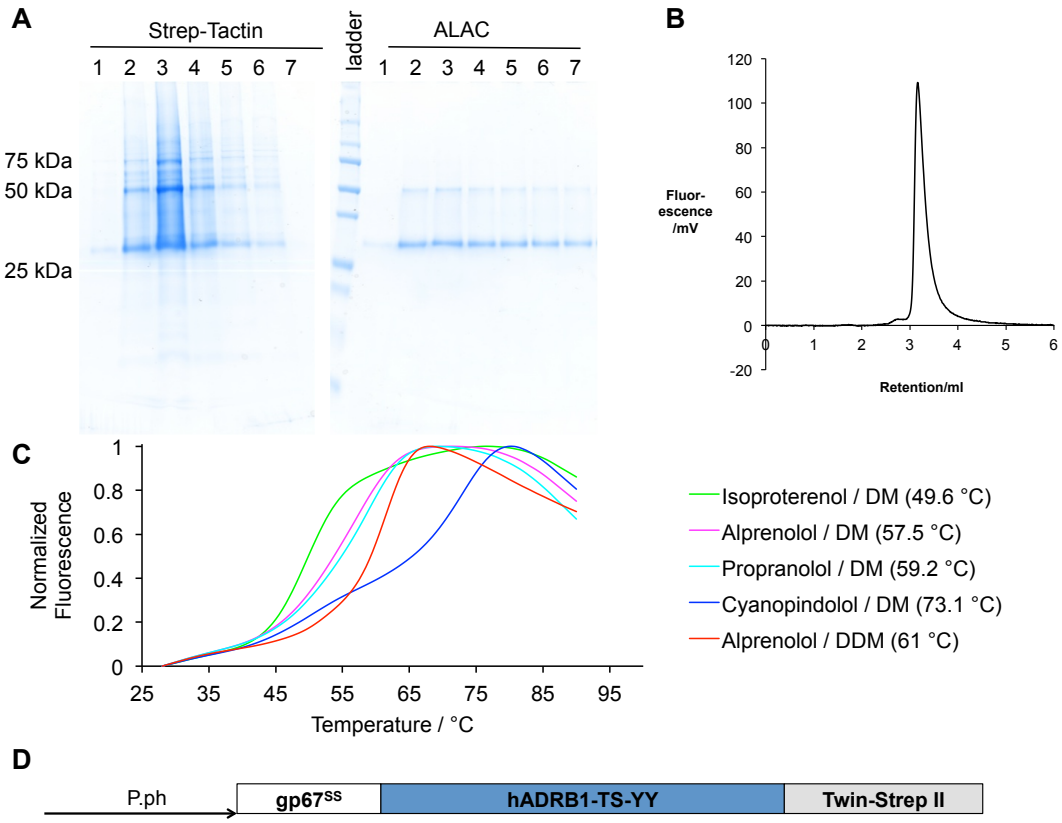


Figure 28. Purification and characterization of the hADRB1-TS-YY-Twin-Strep construct.

Sf9 cells were infected with recombinant baculovirus using the polyhedrin promoter to drive expression the Sf9 infected with recombinant baculovirus using the polyhedrin promoter gp67^{SS}hADRB1-TS-YY-Twin-Strep. Membranes containing the receptor were solubilized in DDM. (**A**) Purification steps are indicated above lanes including Strep-Tactin and ALAC; gels stained with coomassie. (**B**) FSEC analysis of ALAC purified receptor, using fluorescence of tryptophan (excitation 280 nm; emission 350 nm). (**C**) Melting temperature with ligand as indicated, determined by thermal denaturation by means of the CPM assays. (**D**) Overview of the gp67^{SS}hADRB1-TS-YY expression cassette.

other hand the appearance of the FSEC peak (Figure 28 (B)), which is very similar to that of the hADRB1-TS (Figure 27(B)) and the tADRB1.44-m23 (Figure 26(B)), indicates that the integrity of the receptor is still maintained, with an unchanged size compared to that of the hADRB1-TS.

Therefore it has been shown that this mutant, which can in principle be fully activated by an agonist and a G-protein, can still be purified to obtained a highly pure and homogenous receptor preparation, however, as expected, with a reduced thermo-stability. As such this receptor is most likely a very

promising crystallization target to be crystallized with a G-protein or another binding partner.

Discussion

Here it has been shown that a GPCR can be efficiently purified using the Twin-Strep tag giving highly pure protein, while using the single Strep only resulted in a low yield but still highly pure protein. At the time this purification method was tested and developed (2014 - early half of 2015) no record of a GPCR purification using the Twin-Strep tag existed. However, a purification of the cannabinoid receptor using the Twin-Strep was recently published (Yeliseev et al. 2016).

It was here observed that a single Strep-tag was not easily detected. This is most likely an observation that primarily applies to situations where the protein is of extremely low yield such as GPCRs and other hard to express targets. However, it has not been proved completely that weak binding and fast binding kinetics is the cause for the lack of signal when using the strep-tag II alone. It is interesting to note that Grisshammer observed a similar effect with an early generation Strep-tag (Tucker and Grisshammer 1996). To prove that the strep II-tag cannot be efficiently detected would require a simple comparison of the same receptor – one with a single Strep-tag II, and one with the Twin-Strep-tag, which would be easy to perform, in a situation where one has both receptors either purified or with amounts equilibrated by for example radio ligand binding. However, the observation that detection of the single strep II tag ranges from impossible to extremely weak, supports the observation that only slight amounts of single strep II tagged protein could be purified, and that in turn, this could be due to a not high enough affinity of this tag in the single version.

The established purification using the Twin-Strep tag, offers an extremely fast method for obtaining highly pure protein in a very short time, but with relatively little labor. This is a very big advantage exactly in relation to studies such NMR studies in which peak assignment by mutagenesis is to

be carried out, especially when the yield of the protein is low, implying that large amount of bio material must be processed to obtain enough protein.

Upon purification using the alprenolol ligand affinity chromatography (ALAC) it has here been shown that an extremely homogenous preparation of receptor can be obtained, which is most likely very suitable for further structural characterization including crystallography and NMR spectroscopy. This is so since the hADRB1-TS characteristics, although not as stable as the tADRB1.44-m23 construct, do approximate the latter which is known to be readily crystallize and most importantly being amenable for solution NMR studies, although this was done with a more thermo-stable construct – the turkey thermo-stabilized ADRB1 (TS-ADRB1) (Isogai et al. 2016).

Materials and Methods

Plasmids

The plasmid used in the initial studies for purifications using the single Strep tag II were derived from the plasmid commercial pFastbac (Invitrogen). The hADRB1-US ORF was sub-cloned into this backbone pFastbac either with out a signal sequence, or with the gp67 signal peptide in the long or the short form. The Strep tag II was added by including the encoded sequence by PCR. The tADRB1.44-m23-6His ORF was similarly sub-cloned into pFastbac. All plasmids having expression cassettes with GPCRs with C-terminal EGFP fused to them were produced as described earlier.

All mammalian plasmids were derived from pACMVtetO in which the relevant ORF was sub-cloned into. The pACMVtetO plasmid has an Tetracycline inducible CMV fusion promoter, in which to Tet-repressor binding elements are present immediately down stream of the CMV promoter, thus making it repressed in the presence of the Tet-repressor.

Insect cell culture, recombinant baculovirus production and protein production

Sf9 and Hi5 cells were cultured in SF4 Insect Cell Express medium (Bioconcept, CH). The Recombinant baculoviruses were produced according to the Bac-to-Bac (Invitrogen) manufacturer's instructions, however, using either the EmBacY Bacmid for all viruses expressing GPCR with the single Strep tag II or 6xHis tag, an in all other cases the Multibac Bacmid was used. Recombinant Bacmid was isolated using Qiagen miniprep kit. Recombinant baculovirus Bacmid was transfected into Sf9 cells using ExtremeGene HP (Roche) transfection reagent according to manufacturers instructions. This was done in adherent cells and this would become the p0 stock. The p0 stock was then used to infect approximately 50 ml of Sf9 cells at a cell density around 1.5 mill/ml. 2 days post infection the culture supernatant was

harvested to obtain the p1 stock. Finally the p1 stock was amplified in 250 ml of Sf9 cells at a cell density around 1.5 mill/ml to obtain the p2 stock, which was used in the subsequent experiments For small scale experiments the cells were grown in 50ml vessel, while large scale experiments were ranging from 1 to 20 liters. Both small scale and large-scale cultures were infected at a density of approximately 2 mill/ml, as adjusted by counting using a neubauer chamber. 72 hours post infection, the cells were harvested, aliquoted and stored at -20 or -80 degrees Celsius, for further analysis. In experiments in which the receptor was labelled with 15N-valines, Sf9 cells were pelleted at 600 g for 20 min, and resuspended in customized medium lacking valine and yeast extract, in which 15N-valines had been added to a concentration of 120 mg/l. Everything else was done exactly as described for unlabelled receptor production.

Mammalian cell culture

HEK 293 GnTi- stably expressing the Tet-repressor were grown in DMEM (high glucose, Biocencept) supplemented with 10% bovine serum. Stable cell lines expressing the hADRB1-TS-EGFP-Twin-Strep construct, were obtained by transfecting pACMV derived plasmid in to the HEK 293 GnTi- cells, by means of poly-ethylene-imine (PEI) at 3µg PEI to 1 µg DNA, approximately 40 µg were used per 15 cm dish of cells. After 2 days, medium was exchanged with the same fresh medium but with 1mg/ml G418 added as a selection reagent, since a G418 resistance gene is present in the pACMV vector backbone. After 3 to 4 weeks a stable cell line would be obtained expressing the receptor as a polyclonal cell population.

Large scale HEK 293 GnTi protein production,

HEK 293 cell lines described above were grown initially grown in 15 cm dishes as described above. For large scale cultures, several 15 dishes with relevant cells were trypsinized, and the detached cells were transferred spun down at 400 g, for 3 min. Cells were then resuspended in PEM medium

(Invitrogen) supplemented with 10% FBS at a concentration of 0.5 million cells/ml. Cells were allowed to grow while scaling up, and keeping the cell density between 0.5 to 2 million cells/ml. At a suitable volume, the cells were induced with Tetracycline at a final concentration of 4µg/ml, and the expression was allowed to run for 3 days, after which the cells were harvested by centrifugation and stored at -20 or -80 degrees Celsius until further processing.

SDS-PAGE, blotting and In-gel fluorescence

Samples as indicated in figures were separated by SDS-PAGE with 12 % Trupage gels and TEA-Tricine buffer system (Sigma-Aldrich), and the gel resolved species were analyzed further. For western blotting, the gel content was blotted onto a PVDF membrane presoaked in methanol, and washed in Tris Buffered Saline (TBS) (20 mM Tris pH 7.5, 150 mM NaCl) (Immobilon-P, Millipore) using a semi dry blotting system and a semi-dry transfer buffer (14.4g/l glycine, 3g/l Tris, 1g/l SDS, 250ml/l Methanol), using 25 mA per gel for 4 hours. For the dot blotting, samples in SDS sample buffer were simply dotted on a piece of membrane, which had been activated as described immediately above. The dot blot samples were allowed to dry on the membrane before proceeding. For all blots, the membranes were washed in TBS, and blocked by 3% Bovine Serum Albumin in TBS over night. Bands were visualized by incubation with either Strep-Tactin-AP for detection of Strep tag II, or a primary antibody, followed by incubation with secondary anti-body, and membranes were washed 3 times in TBS. For the detection of the human beta1-adrenergic receptor the Rabbit anti-human adrenergic receptor antibody was used (SAB2100064Sigma), while a goat anti-Rabbit conjugate to Alkaline Phosphatase was used as secondary antibody. The detection was carried out using alkaline phosphate substrate solution (Lumiphos, thermofisher, 116002525). For in-Gel fluorescence were imaged using EGFP fluorescence in an Amersham imager AI600(GE), with the light

and filter setting suitable for EGFP.

Cell preparation and Solubilization

All homogenization steps were performed using an Ika Turrax Homogenizer. Cell pellets with overexpressed receptor were thawed and re-suspended in lysis buffer (50 mM Hepes pH 7.5, 10 mM NaCl) with protease inhibitors (Roche Complete, EDTA free) and homogenized. Homogenized material was pelleted by ultracentrifugation at 100'000 x g for 30 min. The pellet was homogenized in solubilization buffer (50 mM Hepes pH 7.5, 1000mM NaCl) with protease inhibitors. The homogenized material was then added DDM to a concentration of 1%, and the solubilization was carried for 1 hour at 4 degrees Celsius. After solubilization, the unsolubilized material was removed by ultra centrifugation at 180'000g for 1 hour.

Strep-Tactin purification

The solubilized material described directly above was loaded at around onto a Strep-Tactin column of either 1ml or 5ml (StrepTrap, GE), which had been equilibrated in 350 wash buffer (350 mM NaCl, 50 mM HEPES pH 7.5, and either 0.1% Decyl-Maltoside or 0.03% DoDecyl-Maltoside). The flow rate was set to 1ml/min for the 1ml column, and 5ml/min for 5ml column. After loading, unbound material was washed away with 5 column volumes, and the protein was eluted in wash buffer of choice in which 5 mM desthiobiotin had been added. All runs were performed on a Biorad NGC 100 chromatography system.

Alprenolol Ligand Affinity Chromatography

The Alprenolol Ligand Affinity Chromatography (ALAC) Resin was produced according to Caron et al. (1979). Materials Sepharose 4B(GE), 1,4-butanediol-di-glycidyl-ether(Sigma), Sodium Thio-Sulfate(Sigma), Potassium Peroxysulfate (Sigma), SodiumBoroHydride (Sigma), SodiumCarbonate (Sigma), Sodium hydroxide (Sigma), DiMethylSulfoxide, Alprenolol (Frontier

Scientific). The purification was performed in the following way: All runs were performed on either an Äkta Express or a Biorad NGC 100. An XK16/25 column, in which 3ml or 5ml of ALAC resin had been packed, was equilibrated with wash buffer (350 mM NaCl, 50 mM HEPES pH 7.5, and either 0.1% Decyl-Maltoside or 0.03% DoDecyl-Maltoside). The eluted protein from the Strep-Tactin column was pooled and loaded onto the ALAC column. Unbound protein was washed away with 5 column volumes of the chosen wash buffer. Finally elution was done using the chosen wash buffer, in which Alprenolol (100 µM) or (RS)-Atenolol(1 mM) (Sigma) was added.

Radioligand binding

Cells were thawed and resuspended in Phosphate-buffered Saline supplemented with Roche complete EDTA free protease inhibitor, and homogenized using a 29G needle with a syringes. For each reaction an amount corresponding to 100'000 cells were used. The crude cell membranes were incubated for one hour with the radioactively labelled beta blocker, [3H]-DiHydroAlprenolol ([3H]-DHA). At the end of the incubation, the membrane suspension was applied on a Whatman Grade GF/B Glass Microfiber Filters (GE 1821-025), using a vacuum manifold (Millipore, XX2702550 | 1225). After application, the membranes were washed with 15ml of ice cold PBS, briefly dried on the filter, and transferred to a scintillation vial and added 10 ml of scintillation fluid (ScintLogic U, SG-BXX-01). The samples were incubated at least 16 hours, vortexed and transferred to a scintillation counter for quantification.

Fluorescence Size-Exclusion Chromatography

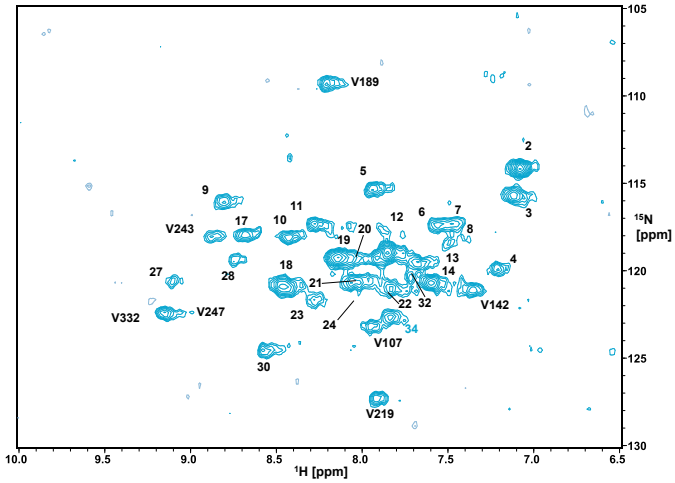
Purified analyzed by being run on a Tosoh SuperSW3000 analytical gel filtration column using an Äkta Ettan HPLC system (GE) equipped with an auto-sampler and fluorescence detector set to excite Tryptophan residues (280nm), with emission detected at 350nm. The flow was set to 0.2 ml/minute. The mobile phase was 50 mM Hepes pH 7.5, 500 mM NaCl, 0.05% DDM. The

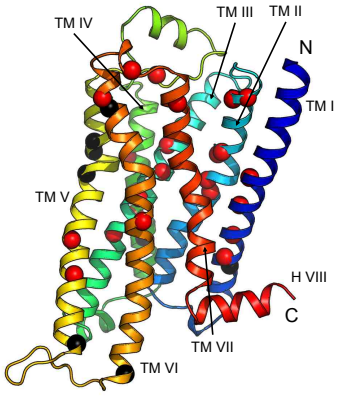
column void, V0, is 2ml, while the total column volume, Vt, is 4.5 ml as estimated by observing void peak and small molecule peak using a Biorad protein standard. The run was continued for 6.5 ml for each sample. All runs were performed at 4C.

Thermo-stability assay

Thermal stability was determined by means of the CPM assay (Alexandrov et al. 2008). Briefly, the purified proteins described above, were thermally denatured by increasing the temperature from 30 to 90 degrees Celsius at 4 degrees per minute. This is done in the presence of the CPM dye, which becomes fluorescent when its maleimide moiety reacts with cysteine Sulhydryl group. In this way chemically, inaccessible cysteines, become exposed when the receptor is denatured, and react with the CPM dye. During the denaturation the fluorescence is detected (excitation 365nm, emission 460 nm). All runs were performed in a Qiagen Rotorgene QPCR machine. The denaturation profiles were analyzed by the software accompanying the apparatus, and the melting temperature was defined as the point infliction point of the denaturation curve, at which the derivative of the fluorescence as a function of temperature assumed its maximum value. Noise of the fluorescence raw data was removed using the software by applying a sliding average over the temperature, termed a heavy digital filter. All species were measure as an average of 4. Approximately 4 µg of protein was used for each sample, using a concentration 0.2µg/ml of the CPM Dye.

5. NMR spectroscopy of the thermo-stabilized human Beta1-Adrenergic Receptor





Chapter 5 frontpage. 2D [^{1}H, ^{15}N]-TROSY spectrum of the hADRB1-TS-YY construct (top) and model of the hADRB1-TS, showing position of nitrogen valines as colored spheres (bottom).

See legends of Figure A 10 and Figure 29 for details.

Introduction

Experimental strategy: Selective labelling of an amino-acid subset to study conformational states

The solution NMR spectroscopy of a G-Protein Coupled Receptor was first carried out in Khorana's lab on the model GPCRs Rhodopsin (Klein-Seetharaman et al. 2002). Importantly, it was recombinantly expressed and selectively labelled. That is, only a subset of the amino acids in the protein, is labelled with residues that in turn harbors the chosen NMR active isotope. As an example consider ^{15}N-valines, in which the nitrogen that forms the peptide bond is NMR active, the only nitrogen atom in this particular amino-acid. As such, in choosing ^{15}N-valines, one gets a NMR-probe positioned in only the peptide back-bone.

The parts of the receptor of the most immediate interest are the trans-membrane helices. Firstly, this is due to the ligand binding pocket of the adrenergic receptors, which is well inside the helical bundle. Secondly, the trans-membrane segments are really the only means for transmitting the signal from one side of the plasma membrane to the other. Therefore, since the helices are mainly composed of residues with hydrophobic side chains, such as L-valine, L-leucine or L-isoleucine, such a residue was an attractive choice for selective labelling. In choosing a specific amino-acid it is important to make sure that the amino acid of choice is present in the region of the protein to be studied. Previous studies have been carried using L-valine to study the conformational response of the ultra-stabilized turkey ADRB1 receptor (turkey TS-ADRB1) (Isogai et al. 2016). An advantage of choosing L-valine is the fact that this amino-acid, among others, is not subject to metabolic scrambling (i.e. conversion of a particular amino-acid or metabolite to another amino-acid or metabolite), since this is an essential amino-acid (Skora et al. 2015). Therefore L-valine was chosen for the selective labelling of the hADRB1-TS receptor and related constructs. A model of the receptor can

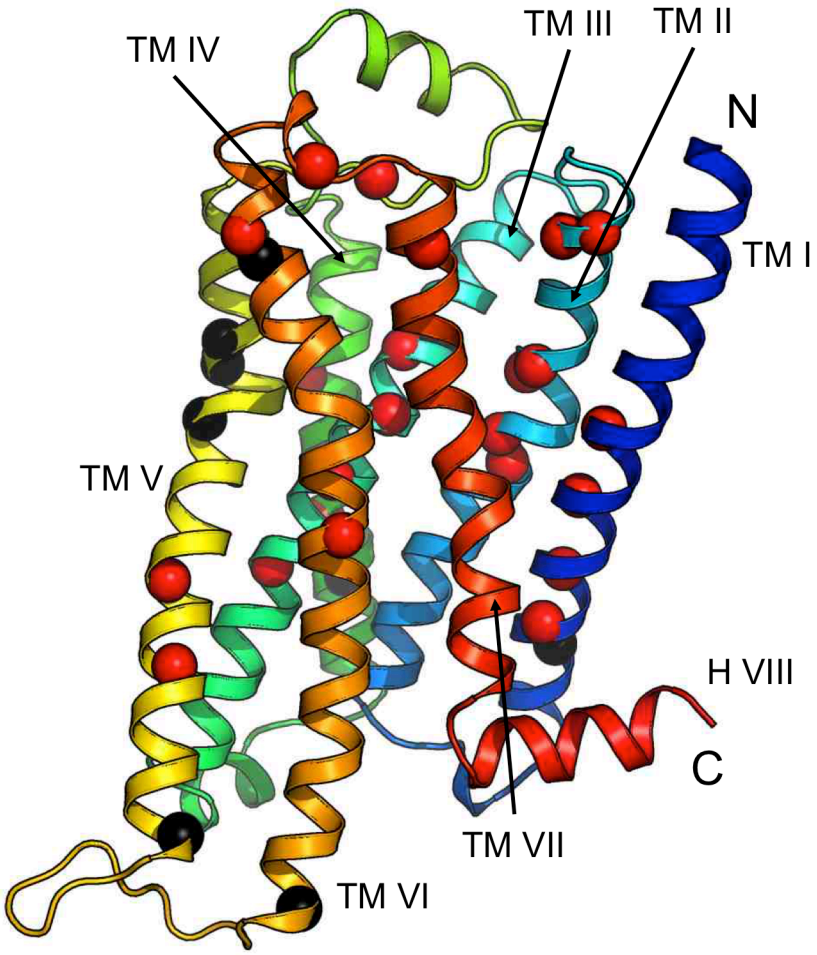


Figure 29. Model of the human ADRB1-TS with nitrogens of valines as spheres.

Model of the human ADRB1-TS construct shown as ribbon, with N and C, indicating N- and C-termini, respectively. Each transmembrane segment is indicated followed by TM followed by a roman number indicating the respective number of this segment, and the amphipathic helix is indicated with and H followed the roman number VIII. Nitrogen atom of valine residues which have been used in all experiments with selective labelling are shown as red ball if valine is conserved in terms of position in the primary sequence, and its predicted position in the 3D structure, and black balls if the valine residue is unique, as compared to the turkey TS-ADRB1. Model produced using Modeller with the ultra-stabilized tADRB1 structure as template (pdb: 4BVN).

be seen in Figure 29, in which all the position of the Nitrogen atoms of the valines have been visualized as balls (see legend of Figure 29 for explanation of coloring). It is clearly seen that the residue are well distributed in the receptor, except for TM7. The approach applied here was to use specialized

medium for the baculovirus infected insect cells. In this particular case this would entail using medium without any yeast extract, due to the presence of the chosen amino acid in the yeast extract (Gossert et al. 2011).

A limitation in the NMR studies, posed by the size of a GPCR-detergent complex, roughly with a mass of 100 kD, is the significant line broadening caused by the comparatively longer correlation time, or phrased differently, slower molecular tumbling. The transverse magnetization, which is responsible for the NMR signal, is relaxing with increased speed as mass of the complex increases. This, in turn leads to broader lines, compared to smaller molecules or molecular complexes. Magnetic dipole-dipole interactions and chemical shift anisotropy (CSA), are both leading to fluctuating magnetic fields due to the molecular tumbling, and in turn this leads to the transverse relaxation of the magnetization. The relaxation caused by the chemical shift anisotropy component and the magnetic dipolar interaction, the spin-spin interaction, can be significantly reduced by using a special type of 2D [^{1}H, ^{15}N] correlation spectroscopy, called Transverse Relaxation Optimized Spectroscopy (TROSY). In the TROSY experiment, the interference between the two aforementioned components is exploited to obtain slower relaxation for some components, hence leading to narrower spectral lines (Pervushin et al. 1997). This type of 2D spectroscopy has been exclusively applied here in the study of the hADRB1-derived constructs

This construct, hADRB1-TS construct contains a total of 32 valine residues, while an additional valine is present in the 1D4 tag, making construct with this tag contain 33 valines.

Results

NMR spectroscopy of the hADRB1-TS receptor: Well-folded and stable receptor.

The first successfully acquired spectrum of the hADRB1-TS protein construct was performed with the specific version hADRB1-TS-Twin-Strep1D4 in complex with the beta-blocker alprenolol. Subsequently, a spectrum was obtained with beta-blocker of highest known affinity, cyanopindolol. An overlay, of the 2D [1H, 15N]-TROSY spectra with these respective ligands can be seen in Figure 30. Firstly, it is noted that the general appearance of the spectra is characteristic of a folded protein, due to several well-defined peaks dispersed over, and occupying, a great proportion of the shown region, compared to the random coil region. The random coil chemical shift values of the peptide bond amide [^{1}H, ^{15}N] pair in a valine residue is (8.03ppm; 119.2 ppm) (Wishart et al. 1995), while very similar values are observed for other types of amide protons. The fact, that the collection of peaks is dispersed over a large region, rather than a set of similar peaks with little dispersion, means that each valine residue is present in a unique and relatively well-defined chemical environment. This is different from an unfolded and potentially aggregated protein, for which it is unlikely that each individual residue will be present in a well-defined state. Rather, an unfolded protein will be a population of species that are very heterogeneous. Therefore, it is reasonable to believe that the protein is present in its functional folded form. It is note worthy the general appearance is very similar to that of the turkey TS-ADRB1 (Isogai et al. 2016), although one very obvious difference is the signal-to-noise (S/N) ratio which is much better in the case of the turkey TS-ADRB1. This difference in S/N is mainly rooted in much lower yield of the human receptor and the chosen amount of starting material for the sample.

32 individual peaks can be identified with some potentially overlapping peaks especially in the central part of the spectrum around (8 ppm ;120 ppm).

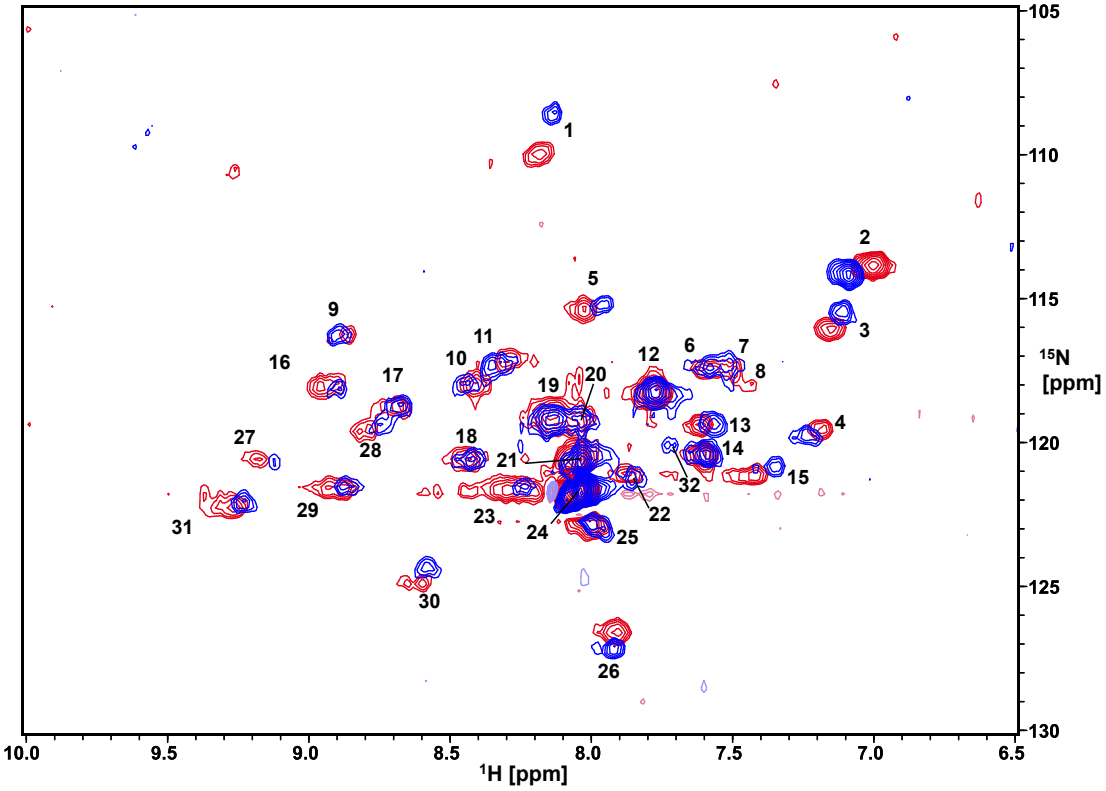


Figure 30. Well-dispersed peaks in [^{1}H, ^{15}N]-TROSY spectrum of hADRB1-TS-Twin-Strep1D4 with the antagonists alprenolol and cyanopindolol.

The [^{1}H, ^{15}N]-TROSY spectra were recorded on a Bruker Avance 900 MHz spectrometer. Sample temperature was 30 degrees Celsius. The receptor was produced in Sf9 cells, purified in Decyl-Maltoside micelles and concentrated to a final concentration of around 100µM with 1 mM Alprenolol (**red contours**) or 0.1 mM cyanopindolol (**blue contours**), in a buffer cotaining 50 mM HEPES (pH 7.5) and 200 mM NaCl, 95%/5% H_2O/D_2O. Peaks were picked manually using the program CARA (Computer Aided Resonance assignment, cara.nmr.ch/). The procedure for picking peaks were carried out by counting from upper most peak from left to right, sloping downwards.

The peaks were manually picked with the program CARA (Computer Aided Resonance Assignment, http://cara.nmr.ch/) and have labelled with a unique number as seen in Figure 30. For all subsequently acquired spectra, the uniquely numbered peak positions were reused by overlaying spectra, and adapting the peak positions, where peaks could be assigned to each other beyond doubt, and otherwise would be picked and receive a new unique number. For example, peak 1 is the same peak in all spectra.

It is noted that particularly peak 19 and peak 25, both relatively strong peaks, might be two or more coinciding peaks. Some weaker peaks appear as

what could be perceived as doublets, for example peak 29 and 31. On the other hand, only a very weak peak is seen for the peak 8. However as evident from several other spectra, this peak persists, while also appearing strongly in other spectra. Taken together, the number of identified peaks corresponds to the number of expected peaks, although, a possibility of more than the expected number of peaks is possibly indicating indicating alternate conformations(s).

One major peak stands out by dwarfing the remaining peak in terms of intensity. This peak, peak 24, is due to an additional valine, which is present in the purification tag, in the 1D4 part of this tag (experimental confirmation discussed later). The entire tag is expected to be entirely disordered, and is present at the C-term of the receptor, allowing it to tumble faster, than the rest of the receptor in the DM micelle. This could explain the larger intensity, since the transverse magnetization causing the signal in the NMR experiments would relax more slowly due to the shorter correlation time, and in turn lead to more signal due less signal loss, as compared to the more slowly tumbling valines in the receptor itself.

Several other prominent peaks stand out from the general population of peaks from the labelled valines in the molecule, and are easily recognized since the peaks lie along the periphery of the cluster of peaks. Especially, the two peaks around 8ppm but each at the extreme upfield or downfield region in the ^{15}N dimension, peak 1 and 26, respectively, stand out. It is clear that these two peaks probably come from the equivalent residue observed in the turkey TS-ADRB1 construct, which result in peaks in the similar spectral regions (Isogai et al. 2016). Similarly, the two peaks, peak 2 and 3, at the extreme end of the upfield region in the hydrogen dimension stand out. One peak of similar and strong appearance is present in the turkey TS-ADRB1.

Tentative assignment by similarity to spectra of the thermo-stabilized turkey ADRB1 spectra.

As mentioned earlier the hADRB1-TS construct, not including purification tags, contain a total of 32 valines. Comparing this construct to the turkey TS-ADRB1 construct, the hADRB1-TS construct has 23 valines that are shared with the turkey TS-ADRB1, in terms of position in the primary structure but also in term of anticipated structural position, and 9 unique valines. Likewise, the turkey TS-ADRB1 has 5 unique valines, and hence a totaling 28 valines (position of valines in this receptor can be seen in Figure A 2, with a side by side comparison to the hADRB1-TS in Figure A 3. In the study by Isogai et al. (2016), a total number of 21 residue, were assigned of which 3 are in unique positions, leaving a maximum of 18 assigned peaks for which the assignment can potentially be transferred. This number of peaks will further be reduced by the difference in the chemical environment of each conserved valine, hence causing a shift in the individual chemical shift value, due to the difference between the hADRB1-TS and the turkey TS-ADRB1 in the primary sequence. The sequence differences might in turn, also contribute indirectly to chemical shift changes through structural differences, in an otherwise conserved region in the vicinity of any of the given valines.

Overlaying the spectra from the two receptors, hADRB1-TS and the turkey TS-ADRB1, with the ligands alprenol and cyanopindol, respectively, reveals similar spectral appearance (Figure 31, Figure 32), however with several potentially corresponding peaks appearing in overlapping regions. Firstly, the aforementioned peaks, 1 and 26 (Figure 30), although not overlapping, are found in spectral region that are only occupied by these respective peaks. This makes it likely that an equivalent valine in the hADRB1-TS and the turkey TS-ADRB1, when in similar chemical environments, will result in peaks with similar chemical shift values. The same argument is applied to other similar peaks although in more crowded regions

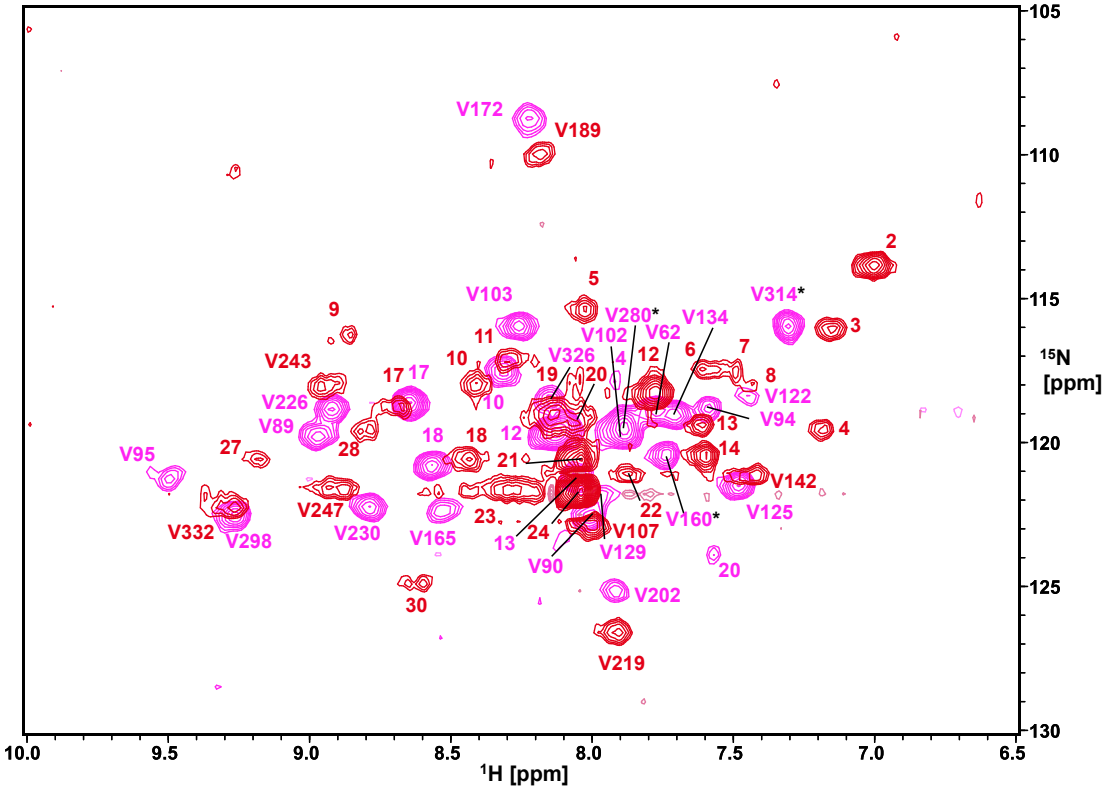


Figure 31. Overlay of [^{1}H, ^{15}N]-TROSY spectra of the hADRB1-TS-Twin-Strep-1D4 receptor, and the turkey TS-ADRB1 receptor, both with the antagonist alprenolol.

Assignments were transferred by similarity of the spectra from turkey (**pink contours**) and the human counter part (**red contours**). Assigned residues in the turkey TS-ADRB1 are indicated by V followed by the position of the valine residue in the primary sequence, while unassigned peaks kept the original number (n.b. No V in front of unassigned peaks), from peak picked manually with CARA. Transferred assignment of valine peak in the spectra of the hADRB1-TS are indicated by the letter V followed by position in primary sequence (n. b. Due to differences in length of N-terminus and intra-cullar loop of wild type receptors, equivalent residues in these will not have the same number in the primary sequence). Asterisk indicates unique valine, only found in the turkey TS-ADRB1 in that particular position. All picked peaks of the hADRB1-TS-Twin-Strep1D4 are summarized in Table 2, while peaks of the turkey TS-ADRB1 are found in Tabel A 4, including assignment (Isogai et al. 2016).

of the spectra, a stricter requirement must be met regarding overlap, due to possible ambiguity. Such peaks include 15, 16, 25, 29 and 31 (Figure 30). The picked peaks along with transfer of assignments for both the ligand alprenolol and cyanopindolol, are found in Table 2. The unusual chemical shift of Val189, and its similarity to the turkey counter part is interesting, and in the study of the TS-ADRB1 it was speculated that this unique chemical shift was due to this particular valines (4.56) position in a helix proline kink, in the position of

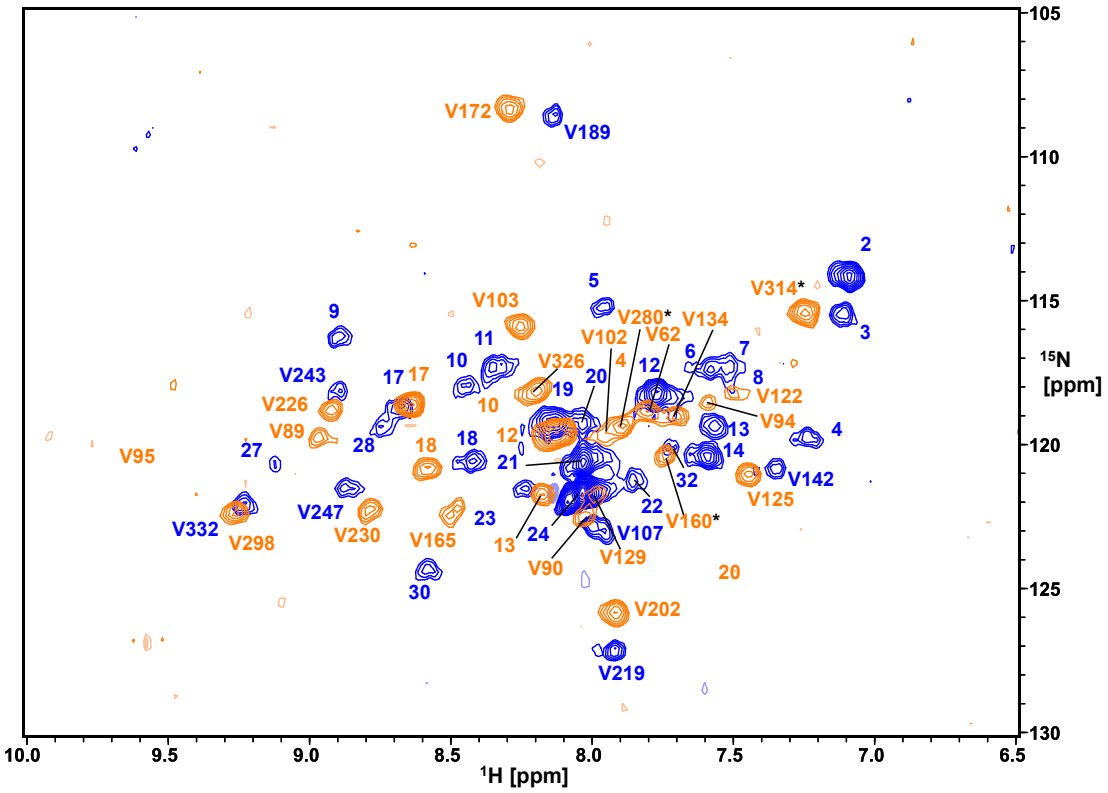


Figure 32. Overlay of [^{1}H, ^{15}N]-TROSY spectra of the hADRB1-TS-Twin-Strep-1D4 receptor, and the turkey TS-ADRB1 receptor, both with the antagonist cyanopindolol.

Assignments were transferred by similarity of spectra from turkey (**orange contours**) to the human counter part (**blue contours**). For details see legend of Figure 31. Asterisk indicates unique valine, only found in the turkey TS-ADRB1 in that particular position. All picked peaks of the hADRB1-TS-Twin-Strep1D4 are summarized in Table 2, while peaks of the turkey TS-ADRB1 are summarized in Tabel A 4, including assignment (Isogai et al. 2016).

the broken hydrogen bond caused by this proline, and that the carbonyl group of the valine would be participating in a hydrogen bond network mediated by water (Isogai et al. 2016).

The chemical shift is due to shielding of the static B_0 field. This is caused by induced currents in the electron density in a molecule, while also the size of the density is important, which can for example be skewed by presence of electronegative atoms. The induced field counteracts the B_0 field, which is called shielding, and this effect is proportional to the B_0 field. In particular, the local electron density surrounding a given atomic nuclei, is relevant since the magnetic field is strongly distance dependent. As such,

Table 2. Valine peaks in the hADRB1-TS-Twin-Strep1D4 construct and their tentative assignment.

Chemical shift values for peaks picked with CARA, from the spectra shown in Figure 30, and the respective transfer of assignments from Figure 31 and Figure 32.

	Alprenolol		Cyanopindolol		Position of residue[1]		
Peak #	^{1}H δ (ppm)	^{15}N δ (ppm)	^{1}H δ (ppm)	^{15}N δ (ppm)	Human ADRB1-TS[2]	Turkey TS-ADRB1[3]	Comment[4]
1	8.18	110.00	8.12	108.59	189 (4.56)	172 (4.56)	
2	7.00	113.90	7.09	114.22			
3	7.14	116.13	7.11	115.43			
4	7.18	119.59	7.23	119.76			
5	8.03	115.42	7.96	115.19			
6	7.59	117.48	7.57	117.34			
7	7.49	117.58	7.51	117.28			
8	7.42	118.17	7.51	117.75			
9	8.84	116.28	8.90	116.25			
10	8.41	118.05	8.44	117.96			
11	8.29	117.21	8.35	117.34			
12	7.77	118.27	7.77	118.27			
13	7.60	119.37	7.57	119.37			
14	7.59	120.45	7.59	120.45			
15	7.42	121.09	7.35	120.85	142 (3.36)	125 (3.36)	
16	8.95	118.05	8.89	118.17	243 (5.57)	226 (5.57)	
17	8.68	118.71	8.68	118.71			
18	8.44	120.62	8.44	120.62			
19	8.15	119.00	8.13	119.26			Multi
20	8.04	119.13	8.04	119.13			
21	8.03	120.48	8.03	120.48			
22	7.87	121.09	7.85	121.35			
23	8.26	121.71	8.24	121.53			
24	8.04	121.66	8.04	121.66			1D4
25	8.00	122.92	8.00	122.92	107 (2.53)	90 (2.53)	Stab,Multi
26	7.89	126.44	7.92	127.15	219 (ECL2)	202 (ECL2)	
27	9.18	120.59	9.12	120.73			
28	8.78	119.61	8.73	119.40			
29	8.87	121.61	8.87	121.61	247 (5.61)	230 (5.61)	
30	8.59	124.92	8.58	124.35			
31	9.26	122.26	9.23	122.09	332 (6.43)	298 (6.43)	
32	n. a.	n. a.	7.73	120.11			

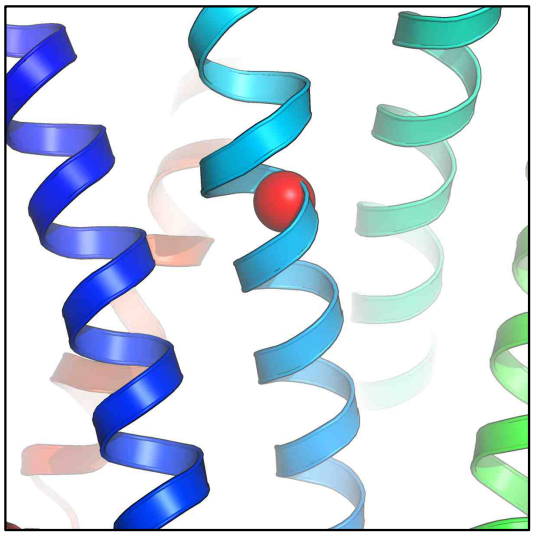
[1]Position in wild-type receptors (Ballesteros-Weinstein number in parenthesis).
[2]Tentative assignment using similarity of spectra and structural information from hADRB1-TS model and the structure of the ultra-stabilized tADRB1 (PDB code 4BVN).
[3]Assignment of peaks obtained from (Isogai et al. 2016)
[4]Comment: 1D4 (valine in 1D4-tag (TETSQVAPA)); Stab (stabilizing mutation); Multi (Multiple coinciding peaks a possibility).

nuclei in identical or similar local and conformational chemical environments

are expected to have identical or similar chemical shifts. For defining this local environment a sphere of 5Å radius has been chosen since that radius is commonly used as the most important region when predicting chemical shifts (He et al. 2009).

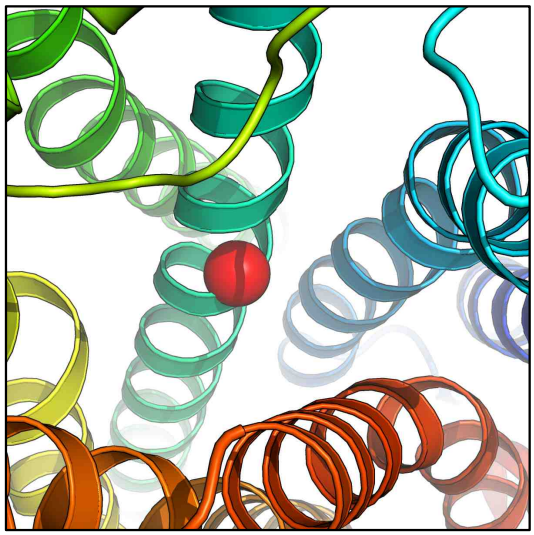
Upon transferring the peak assignment, the immediate chemical environment surrounding each of the valines, for which a peak was tentatively assigned, was examined structurally within a radius of 5Å of the nitrogen of the relevant valine, in the model of the hADRB1-TS. This was done to validate that the local chemical environment is conserved in this region. Figure 33 is showing each of the valines to which a peak has been tentatively assigned. The convention was used, that only residues in hADRB1-TS model, that are different from residues in the equivalent position of the turkey TS-ADRB1 structure are depicted with a side chain in black, while all other residue are rendered without side-chains. In addition, all residues within the 5Å radius of the relevant valine, in both the hADRB1-TS model and the structure of the turkey TS-ADRB1, are shown as a sequence alignment. It is noted that the peptide backbone of model of the human, which was made using the ultra-stabilized tADRB1 (pdb 4BVN) as a template, is almost identical to the template, and therefore no further treatment of the model was done.

It is clearly seen in Figure 33 that for Val107, Val142, Val189, Val219 and Val 332, environment is conserved. On the other hand, Val243 and Val247, do have substitutions. As for Val243, L238C and I241A (substitution from turkey to human) are present, but the side-chains are both more than 5Å away, while each of these residues are two or more peptide bonds away from the relevant Val243. Likewise, the Y248F in the environment of Val247, is positioned distantly from Val247, making it likely that the missing hydroxyl-group will not have a large effect.



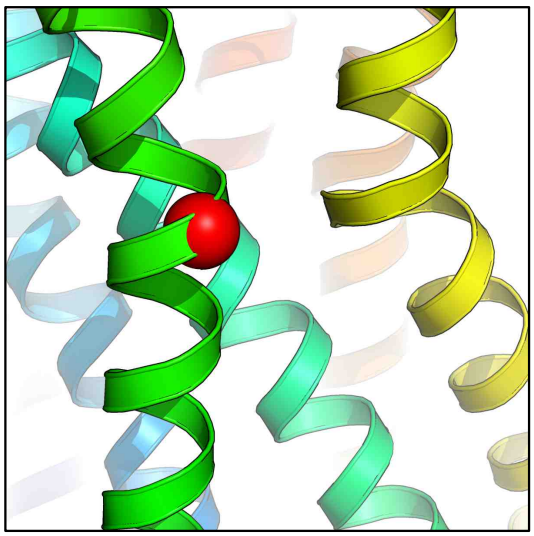
Val107 (2.53)

```
Residues within 5Å radius
#          103        141
human      ADLVVGLLV C
turkey     ADLVVGLLV C
Identity   ********* *
```



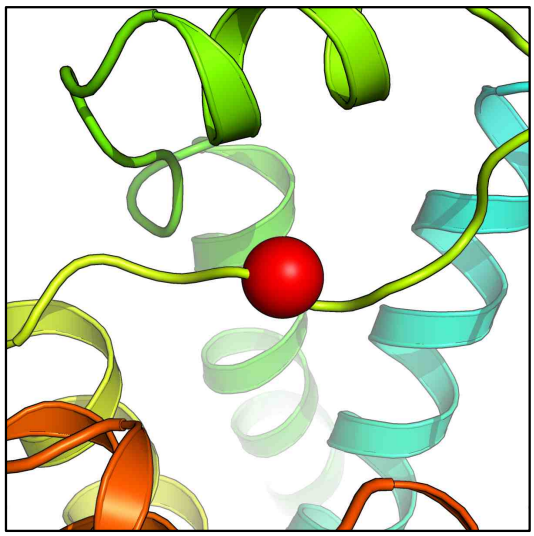
Val142 (3.36)

```
Residues within 5Å radius
#          138
human      DVLCVTAS
turkey     DVLCVTAS
Identity   ********
```



Val189 (4.56)

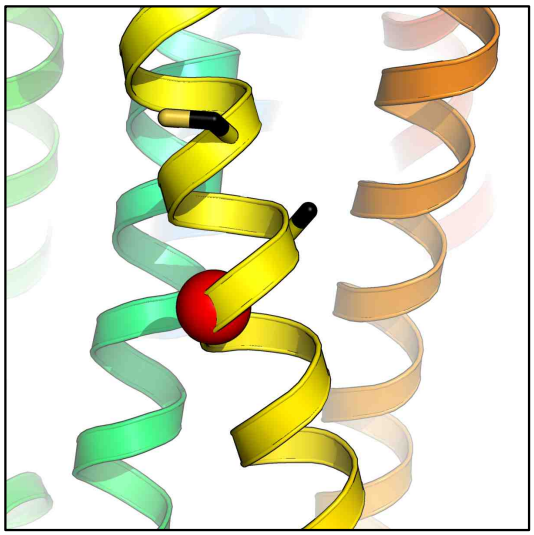
```
Residues within 5Å radius
#          185
human      ISALVSF
turkey     ISALVSF
Identity   *******
```



Val219 (ECL2)

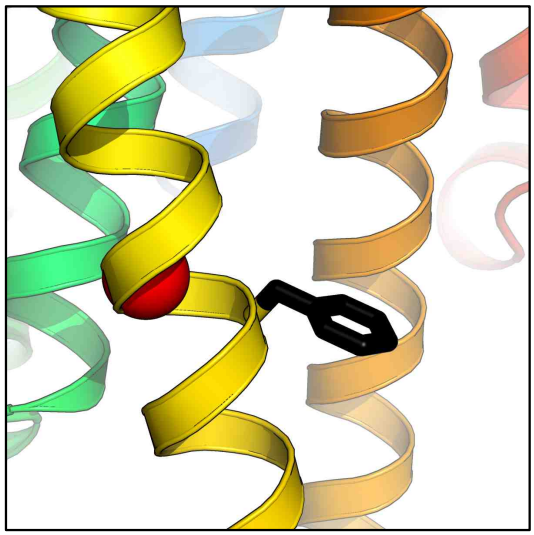
```
Residues within 5Å radius
#          199 210 217
human      W   Y   DFVT
turkey     W   Y   DFVT
Identity   *   *   ****
```

Taken together all the local chemical environments are fairly similar, supporting the tentative assignments of the peaks to these residues.



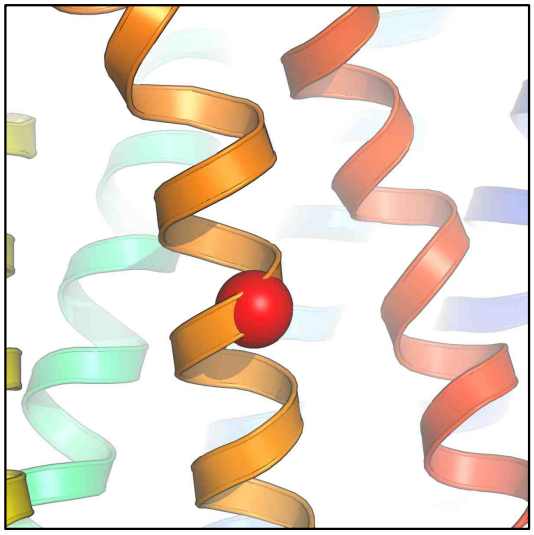
Val243 (5.57)

```
Residues within 5Å radius
#          238
human      CIMAFVALR
turkey     LIMIFVALR
Identity    ** *****
```



Val247 (5.61)

```
Residues within 5Å radius
#          157 243
human      Y   VALRVFRE
turkey     Y   VALRVYRE
Identity   *   ***** **
```



Val332 (6.43)

```
Residues within 5Å radius
#          327
human      GIIMGVFTL
turkey     GIIMGVFTL
Identity   *********
```

Figure 33. Evalution of the chemical environment of each assigned peak.

Cartoon depicting the structural context of each of the tentatively assigned valines, showing the peptide backbone in the hADRB1-TS model, with the same coloring scheme as used in Figure 29. The relevant residue is indicated above each panel with Ballesteros-Weinstein number given in parenthesis. ^{15}N atom of the relevant valine residue (in backbone peptide bond) indicated as a red ball. Sidechain of hADRB1-TS are shown (**black** and hetero-atom standard coloring) for non-identical residues between the hADRB1-TS and turkey TS-ADRB1 within a radius of 5Å of the relevant Nitrogen atom, with all residue within this radius shown as an alignment (numbering refers to position in the human primary sequence; identical residue are indicated by an asterisk below alignment).

The hADRB1-TS-Twin-strep construct

As speculated earlier the comparably immense amplitude of peak 24 might be due to it coming from a presumed fast tumbling valine, which is present in the 1D4 tag of the hADRB1-TS-Twin-Strep1D4. Therefore, with no need for the 1D4 tag, this was removed completely, in order to validate whether or not peak 24 is due to this valine in the 1D4 tag. The resulting construct is the hADRB1-TS-Twin-Strep. The spectrum of this construct can be seen in Figure 34, in which it has been overlaid onto the previously obtained spectrum of the hADRB1-TS-Twin-Strep1D4, both with the ligand alprenolol. Indeed, peak 24 has disappeared from the spectrum, thereby assigning this peak to the valine in the purification 1D4 tag. A similar observation was made for a selectively labelled lysine in the rhodopsin C-terminal tail, which is flexible and unstructured (Klein-Seetharaman et al. 2002).

A good overlap is seen among all residues, excluding peak 24. However, there are some apparent differences. Firstly, it has to be noted that in the case of the hADRB1-TS-Twin-Strep construct, the spectra were obtained using 100 mM NaCl in the buffer instead of 200 mM, which might have account for some of the differences in the peaks positions. The NaCl concentration was reduced to make the experimental conditions more similar to those used in the previous study of the turkey TS-ADRB1 (Isogai et al. 2016), while also lower ionic strength gives rise to less thermal noise in the spectral acquisition, and

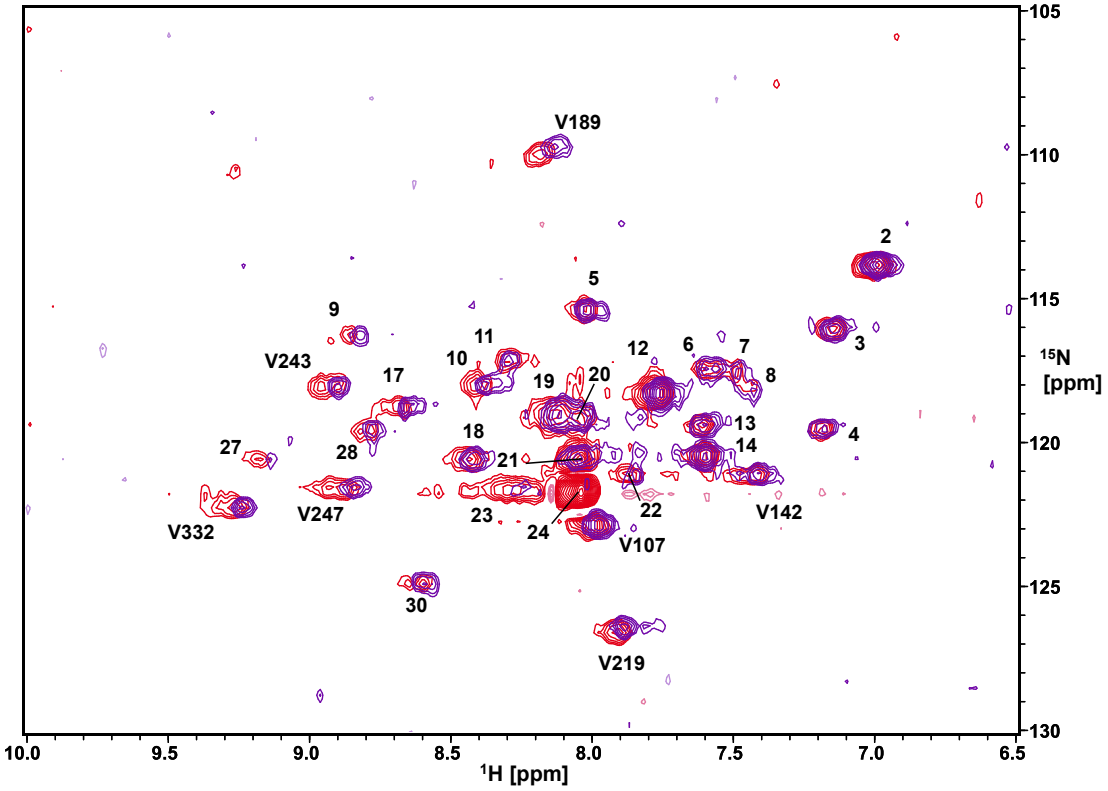


Figure 34. [^{1}H, ^{15}N]-TROSY spectrum of the hADRB1-TS-Twin-Strep receptor and overlaid [^{1}H, ^{15}N]-TROSY spectrum the hADRB1-TS-Twin-Strep1D4 receptor, both with the antagonist alprenolol.

The hADRB1-TS-Twin-Strep was produced, and the spectrum was obtained as described in legend of Figure 30. except for the final buffer which contained 50 mM HEPES (pH 7.5), 100 mM NaCl and 1 mM alprenolol. The [^{1}H, ^{15}N]-TROSY spectrum of the hADRB1-TS-Twin-Strep (**purple contours**) is overlaid onto the spectrum of the hADRB1-TS-Twin-Strep1D4 (**red contours**; from Figure 30). Picked peaks, including tentatively assigned peaks, were transferred directly by similarity of corresponding chemical shifts.

thus a higher S/N ratio. It is also noted that the peak number 8 is appearing stronger in this case.

In principle the peak position should be the same for these two constructs since the removal of a flexible tag should not have any effect on the receptor itself, and hence the chemical shifts of the valines inside the receptor. However, it does illustrate that minor changes should be interpreted with care, especially when there are differences in the experimental conditions (here the measuring buffer).

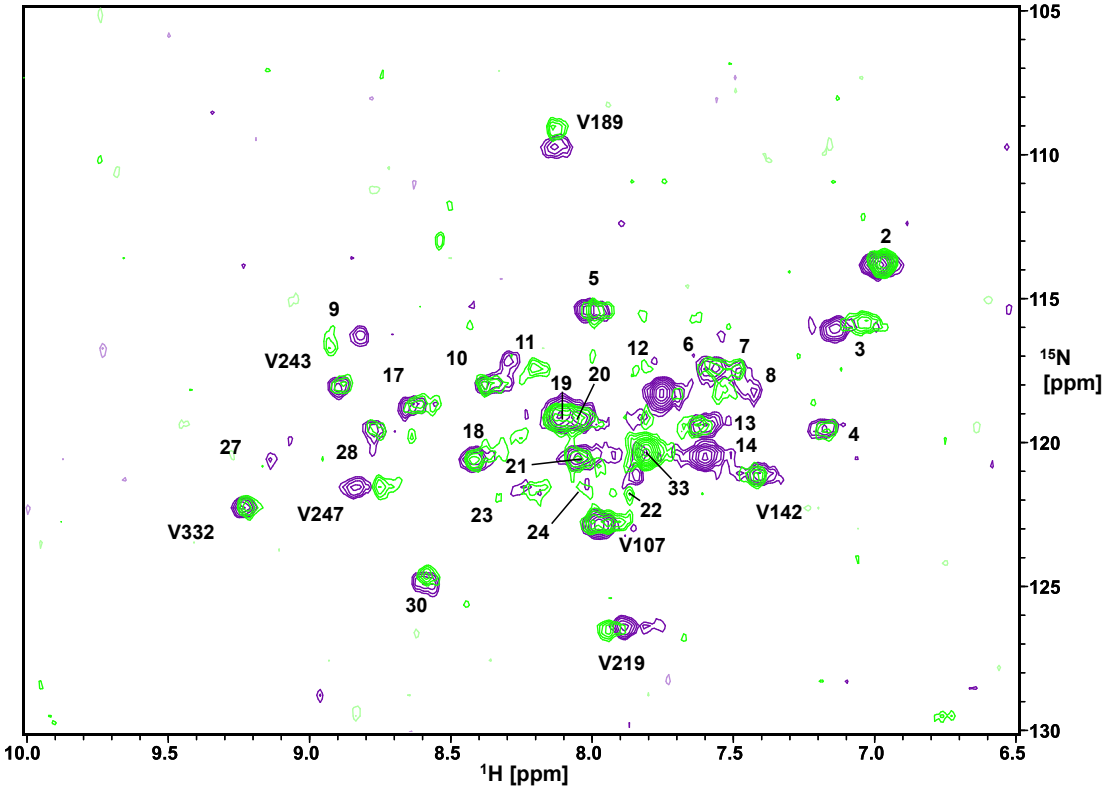


Figure 35. Overlay of [^{1}H, ^{15}N]-TROSY spectra of the hADRB1-TS-Twin-Strep receptor with the antagonists alprenolol and S-propranolol.

[1H, 15N]-TROSY spectrum of the hADRB1-TS-Twin-Strep with 1 mM S-propranolol (**green contours** or 1 mM alprenolol (**purple contours**, from Figure 34). Details of production are given in legend of Figure 34. Peaks of spectrum with S-propranolol were picked and labelled by similarity to the peak positions obtained with alprenolol, or in the case of a dissimilar peak in terms of position, it was given a new number. The chemical shift are summarized in Tabel A 2.

Keeping the aforementioned slightly different experimental conditions in mind, consider then the overlay of the spectra of the hADRB1-TS-Twin-Strep with alprenolol, and the previously by NMR uncharacterized beta-blocker S-propranolol, which was also not characterized in the study of the turkey TS-ADRB1 (Isogai et al. 2016). The spectral overlay is seen in Figure 35. Firstly, it noted that there are some dramatic differences between the to spectra which are not seen between the spectra with the ligands alprenolol and cyanopindolol (discussed below). Secondly, several similarities are observed, including near perfect overlap of for example Val 107, Val332, peak 4, peak 17, peak 30, some of which appeared to be slightly shifted in the comparison of

the two different hADRB1-TS constructs both measured in complex with alprenolol (Figure 34). This observation could be seen as a sign that the differences in these spectra are a consequence of the slightly different buffers, since the two different ligands for which spectra of the hADRB1-TS-Twin-Strep construct were obtained, produced peaks that did not shift between these spectra. As mentioned, there are especially three dramatic changes in the spectrum of the receptor with S-propranolol, this includes the disappearance of the two unassigned peak 12 and peak 14, two of the most prominent peaks, while a entirely new peak has shown up – peak 33, and appears to be a single peak. The peaks of the hADRB1-TS-Twin-Strep are summarized in Tabel A 2. Valine chemical shifts in the hADRB1-TS-Twin-Strep construct.

The double mutant (A244Y, L377Y) of hADRB1-TS, hADRB1-TS-YY gives a well-dispersed spectrum.

Initial attempts at obtaining spectra of the hADRB1-TS-YY-Twin-Strep construct, in complex with the agonist isoproterenol and a engineered minimal Gs protein variant (Carpenter and Tate 2016), failed due to sample degradation. However, it was still possible to obtain a spectrum of this receptor mutant with cyanopindolol. An overlay of this spectrum with the spectrum previously recorded with cyanopindolol of the construct hADRB1-TS-Twin-Strep1D4 is shown Figure 36. It is clearly seen here that several peaks are overlapping with a high degree, reaffirming that the double tyrosine mutant responds in a similar fashion to cyanopindolol, when compared to hADRB1-TS. Generally, several of the peaks appear to be somewhat broader, compared to the peaks of hADRB1-TS with cyanopindolol. This could indicate either that a broader population of states is present for this mutant, or that some chemical exchange is broadening these peaks.

It does not escape notice that several peak positions are not conserved in this mutant. It is interesting that cyanopindolol is not able to force the

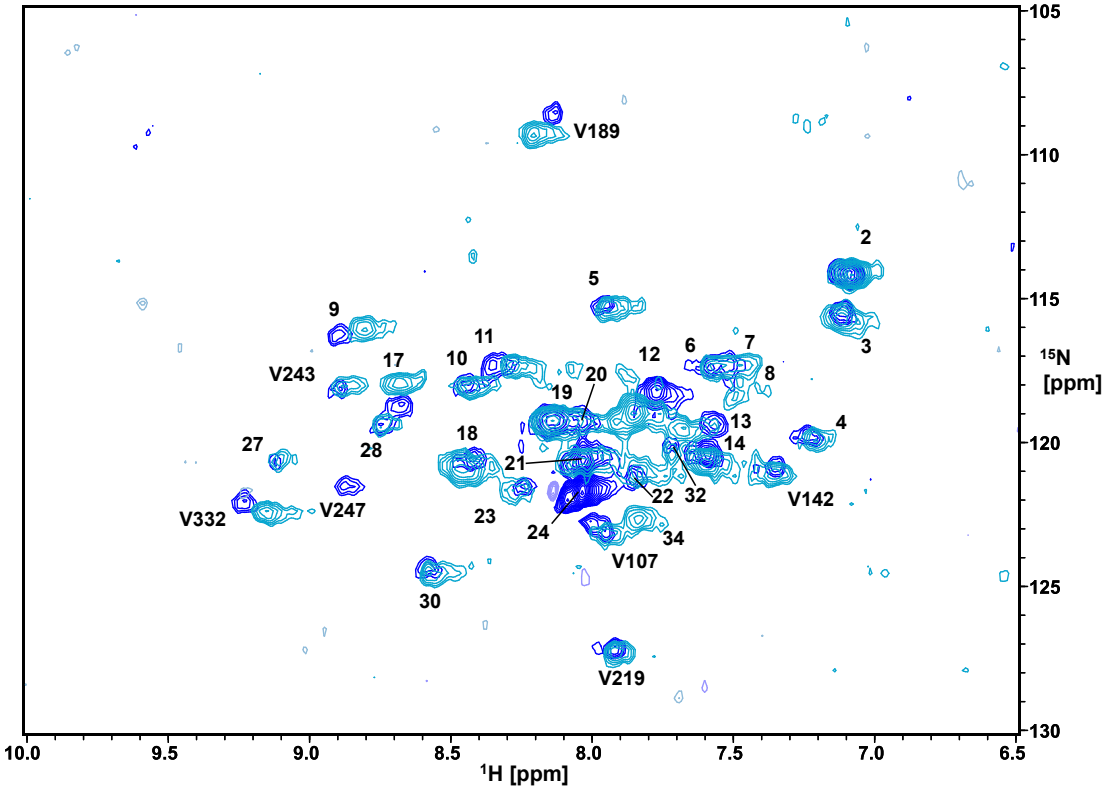


Figure 36. Overlay of [^{1}H, ^{15}N]-TROSY spectra of the gp67^{SS}hADRB1-TS-YY-Twin-Strep receptor and the hADRB1-TS-Twin-Strep1D4 receptor, both with the antagonist cyanopindolol.

[1H, 15N]-TROSY spectrum of the gp67^{SS}hADRB1-TS-YY-Twin-Strep (**light blue contours**) with 0.1 mM cyanopindolol – the double tyrosine mutant (A244Y,L377Y) of the hADRB1-TS with an added N-terminal signal sequence(gp67^{SS}). The receptor was produced exactly as described in legend of Figure 30 The spectrum of the gp67^{SS}hADRB1-TS-YY-Twin-Strep is overlaid the spectrum of the hADRB1-TS-Twin-Strep1D4 with (**blue contours**, from Figure 30). Peaks in the spectrum gp67^{SS}hADRB1-TS-YY-Twin-Strep were picked and labelled by similarity to the spectrum obtained of hADRB1-TS-Twin-Strep1D4 with cyanopindolol, or in the case of a dissimilar peak in terms of position, it was given a new number. The chemical shifts are summarized in Tabel A 3.

receptor into the same state as the hADRB1-TS. This indicates that at least for peak 17 and peak 34, a conformational change could have happened by reintroducing the two tyrosines, alternatively shift could be caused directly by the aromatic group of the tyrosines affecting the chemical shift of valines close by.

The receptor is here regarded as susceptible to NMR studies, but optimization needs to be carried out in order to make studies with the less

stabilizing agonists possible, which is the purpose for working with this mutant, along with having a more wild-type-like receptor.

Discussion

First and foremost it has here been shown that it is possible to obtain 2D spectra of the hADRB1-TS construct and the derived hADRB1-TS-YY construct. The spectra are all characterized by well-dispersed peaks occupying the amide region, but different from that of random coil chemical shift, thus indicating a well-folded protein. The typical measurement was of the length of 24-48 hours, depending on the amount of sample measured (around 1 mg), indicating that it is still a (spectrometer-) time consuming endeavor to study a GPCR by NMR, due to the generally low yield of these proteins. It is interesting to note how the different ligands tested here induce different changes in the obtained NMR spectra, and that similar effect were observed in the study by Isogai et al. (2016).

Taking advantage of the high degree of homology between the human receptor and the turkey counter part, a tentative assignment of 7 peaks to valines in the hADRB1-TS was possible, out of the 18 common peak for which an assignment was available. Similarity of peak positions between the hADRB1-TS and the turkey TS-ADRB1 was the principal method, and peak assigned were transferred wherever little or no ambiguity was observed. Subsequently the local chemical environment was analyzed for similarity or differences, in order to make sure that this was conserved around the assigned valines. An overlay of the tentatively assigned spectra of the hADRB1-TS in complex with Alprenolol and Cyanopindolol, respectively, is seen in Figure 37.

Using the assigned spectra of the turkey TS-ADRB1 as a reference (Figure 38), (Isogai et al. 2016) it is clearly seen that only a few peaks are changing position in the spectra of that receptor in complex with alprenolol and cyanopindolol, respectively. Comparing the effect observed for the turkey TS-ADRB1 to the hADRB1-TS, one must keep in mind that the human receptor has five additional valines, while only 23 valines positions in the primary sequence are shared in the two receptors. Therefore a direct

comparison can only be made between the 7 common and assigned valines. Using both WT-numbering and Ballesteros-Weinstein numbers (in parenthesis) where applicable, these residues are Val107 (2.53), Val142 (3.36), Val189 (4.56), Val219 (ECL2), Val243 (5.57), Val247 (5.61) and Val332 (6.43). Like in the turkey TS-ADRB1 (Figure 38), concerning the hADRB1-TS (Figure 37) the two Val 189 (4.56) and Val219 (ECL2) both move in response to change of the ligand, while an effect is less obvious for the remaining tentatively assigned peaks. This is however consistent with the findings for the TS-ADRB1, in which Val 202 (ECL2) (Val209 in hADRB1), which is situated at the extra cellular part close to the ligand binding pocket, reports on the presence of ligand, and does not appear in the apo-state of the receptor. Likewise, the Val243 (5.47) also stays at a constant position for both alprenolol and cyanopindolol, which was also observed for Val226, the equivalent valine in the turkey TS-ADRB1 (Isogai et al. 2016).

In this way, it is clear that the alprenolol and cyanopindolol induced states appears to be similar, as judged by the Valine probes. However, peak 2, 3, and 4 in the hADRB1-TS spectra have not been taken into account, but are additional information for only the human receptor. It is interesting, that these, for the human receptor unique peaks, are some of the most responsive peaks regarding use of different ligands, and thus, using valines in the NMR experiments, the hADRB1-TS gives potentially more information, compared to the turkey, regarding conformational changes in response to different ligands. Having said that, the spectra obtained with these ligands are still fairly similar, and thus the receptor must be occupying similar states when using alprenolol and cyanopindolol.

As mentioned earlier the receptor in complex with the ligand S-propranolol gives a spectrum in which prominent peaks are missing (peak 12 and 14) but also has a unique peak only observed with this ligand so far. When comparing this spectrum to the spectra obtained with alprenolol one

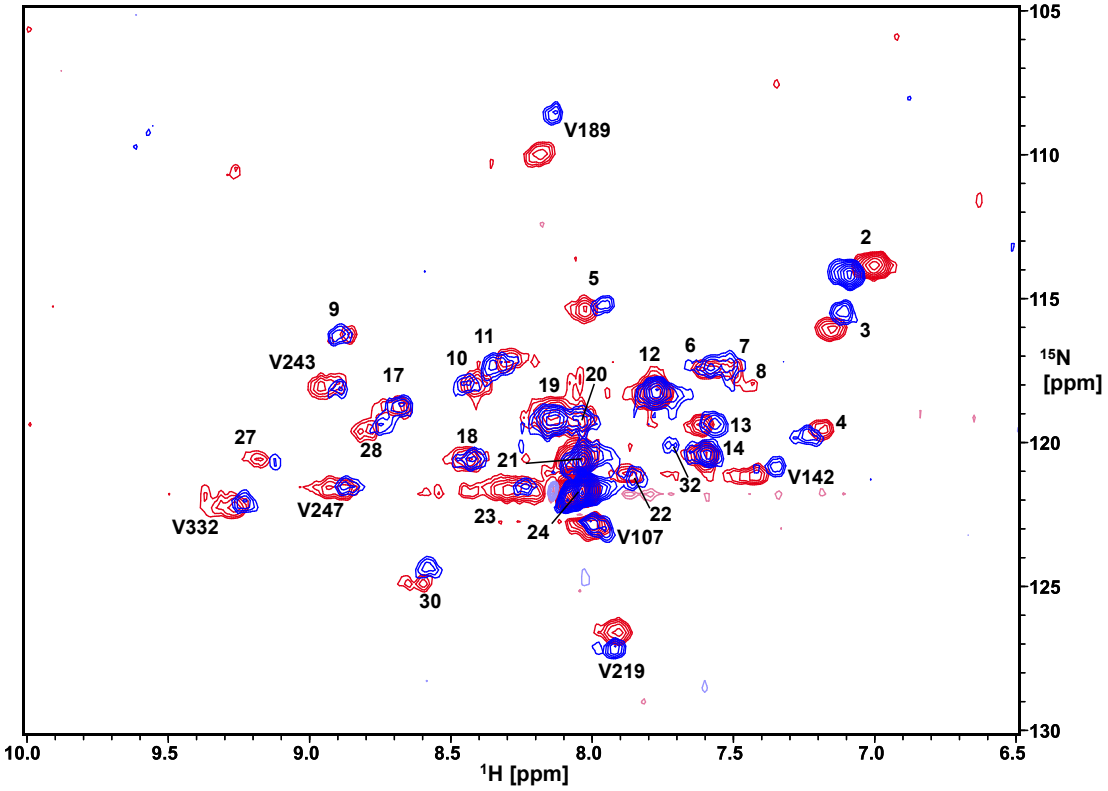


Figure 37. Overlay of tentatively assigned [^{1}H, ^{15}N]-TROSY spectra of the hADRB1-TS-Twin-Strep1D4 receptor, with the ligands alprenolol and cyanopindolol, respectively.

Spectra from Figure 31, Figure 32, with tentatively assigned peaks indicated in the overlay, and summarized in Table 2. Spectra obtained with either 1mM alprenolol (**red contours**) or 0.1 mM cyanopindolol (**blue contours**). Transferred assignment of valine peak indicated by the letter V followed by position in primary sequence.

notices that peak 2 is position in the same place for both ligand, while only peak 3 moves, unlike when comparing spectra obtained with cyanopindolol, in which both peak 2 and peak 3 move, relative to the spectrum with alprenolol. One might therefore speculate that the S-propranolol induced state of the hADRB1-TS is more similar to the alprenolol induced state in the region of the receptor on which these peaks are reporting, but still with some similarities to the cyanopindolol induced state. S-propranolol and alprenolol are both described in the literature as being inverse agonist in the Gs mediated signaling pathway, while being able to activate the MAPK pathway, and could thus be considered biased agonist, for the human ADRB1

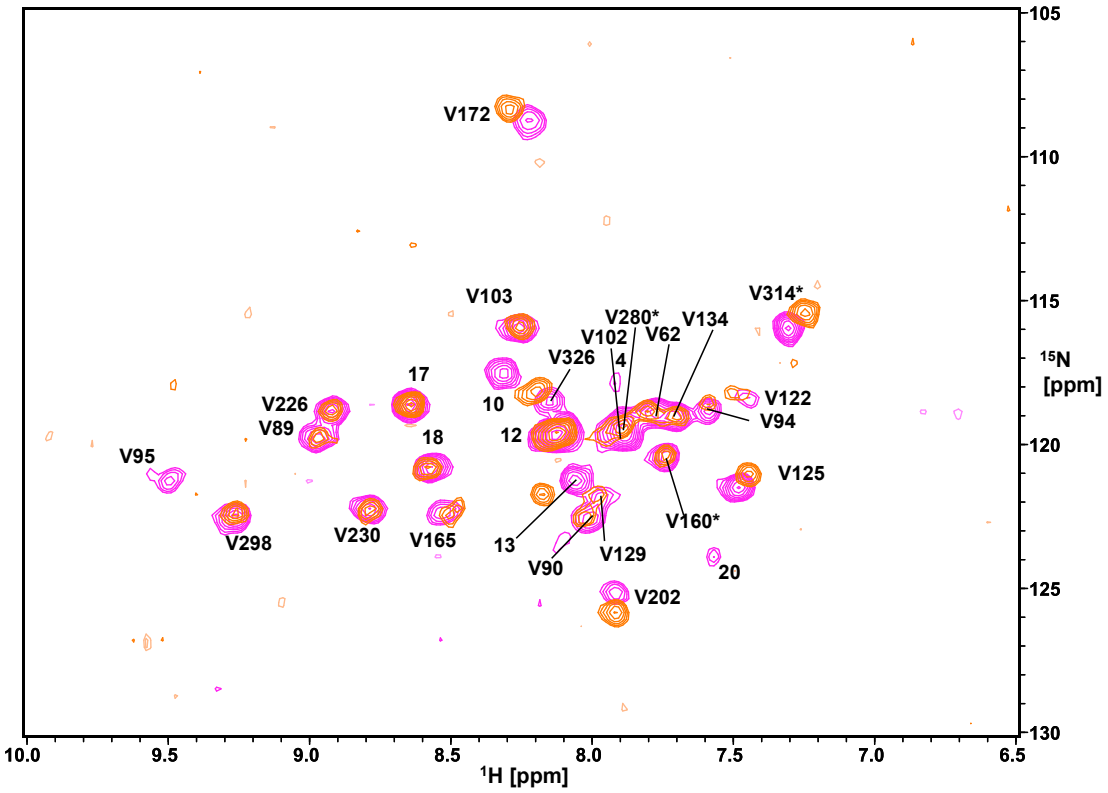


Figure 38. Overlay of [^{1}H, ^{15}N]-TROSY spectra of the turkey TS-ADRB1 receptor, with the antagonists alprenolol and cyanopindolol.

Spectra from Figure 31, Figure 32, with assigned peaks indicated in the overlay, and summarized in Table 2. Spectra obtained with either alprenolol (**pink contours**) or 0.1 mM cyanopindolol (**orange contours**). Assignment of valine peak indicated by the letter V followed by position in primary sequence, while unassigned peaks kept their unique index. Peaks are summarized in Tabel A 4

(Galandrin et al. 2008, Kim et al. 2008). On the other hand cyanopindolol is considered true inverse agonist (Azzi et al. 2001). Therefore peak 2 and peak 3 might be reporting on conformational changes that are related to whether or not ligands are biased.

An interesting fact is that at the proline kink of helix 5, three unique valines are found here (Val230 (5.44), Val231 (5.45) and Val235 (5.49)), compared to the turkey receptor. It is a possibility that these valines gives rise to peak 2, peak 3 and peak 4 that seems to be peaks from unique valines in the hADRB1-TS. The relevant valines are seen in Figure 39(A), while spectral region showing peak 2, 3 and 4 is seen in Figure 39 (B). Val226 (5.57) and

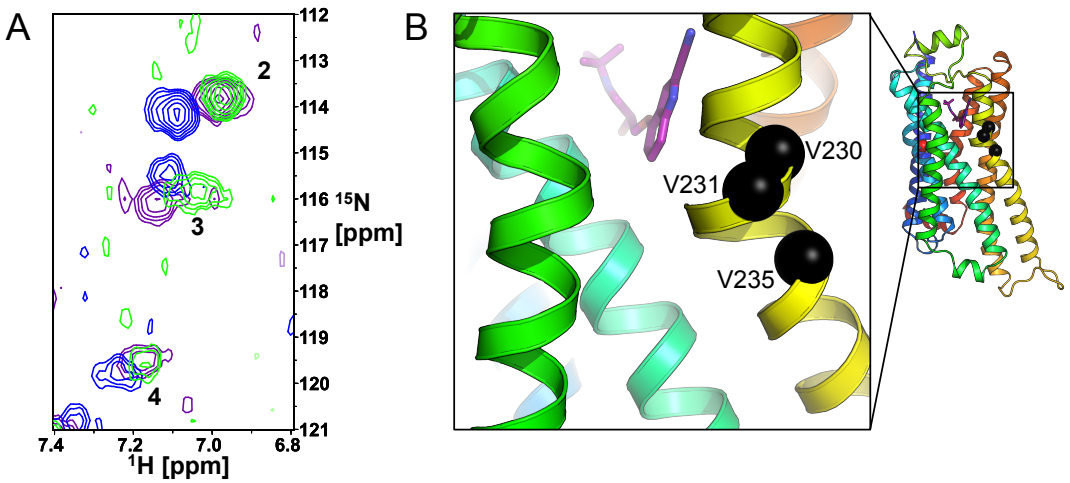


Figure 39. Zoom of peak 2,3, and 4, in the [^{1}H, ^{15}N]-TROSY spectra of the human ADRB1-TS construct with the antagonists alprenolol, S-propranolol and cyanopindolol.

Potential relation ship between unique peaks and the unque valines in the proline kink of TM5, in the hADRB1-TS constructs. (**A**) Region of peak 2, 3 and 4 in overlay from Figure 30, Figure 35, showing the hADRB1-TS in complex with the ligands alprenolol (**purple contours**), S-propranolol, (**green contours**) and cyanopindolol (**blue contours**). (**B**) View of proline kink of TM5, showing 3 of the 9 valine residues only found in the hADRB1-TS construct compared to the turkey TS-ADRB1 in terms of position in the primary structure. Structure depicted with cyanopindolol for which coordinates/pose were taken from the structure of the ultra-stabilized turkey ADRB1 (pdb: 4BVN), after model of hADRB1-TS was aligned onto this.

Val230 (5.61) in turkey TS-ADRB1 were previously found to report on activation, especially distinguishing agonist and antagonists in terms of peak position (Isogai et al. 2016), underscoring the importance of helix 5 in the process of activation by agonists. In hADRB1, Val230 (5.44), Val 231 (5.45) and Val235 (5.49), might also be important positions for activation. This could be supported by the direct link between the kink and Val219 (ECL2), consisting only of 2 alpha-helical turns and a few stretched out residues in the ECL2, which as mentioned earlier, only is visible in the presence of any of tested ligands, but not without ligand in the apo-form (Isogai et al. 2016). It would therefore be interesting to have these three valines assigned, but also observe the response of these to ligands, especially since states induced by for example the biased agonist are hard to study by crystallography due to structures of not highest possible resolution, distortion of the receptor by

crystal packing, and finally the effect caused by stabilizing mutations (Warne et al. 2012).

It has been shown, that an NMR study like the one performed on the turkey TS-ADRB1-TS can be done with the human beta1-adrenergic receptor using the construct applied here. One drawback of the approach used by Isogai et al. (2016), is the amount of labor required for such a study. If that was to be carried out with the hADRB1-TS construct one might estimate that all hands on steps would approximately 6-10 times as labor consuming, a similar increase would be expected regarding the material costs, and some steps could be done in the same time, while other might take longer.

Another potentially more attractive step could be full labeling, including ^{13}C and ^{15}N of all carbon and nitrogen atoms in the receptor and including deuteration. This has recently been done using the turkey beta1-adrenergic receptor (Opitz et al. 2015). An even more promising result has been obtained with a thermo-stabilized neurotensin-receptor using engineered nano-discs for improved stability (Nasr et al. 2017). Having a fully labelled and deuterated, stable receptor would perhaps in the future allow for extraction of information from every single residue, provided the peak was observable, and a sequential assignment was possible. Importantly, information on intra-molecular distances, through acquisition of NOEs, would be the ultimate way to study conformational changes, and likewise performing relaxation measurements to obtain information on the protein dynamics, in the study of a GPCR in response to various ligands.

Materials and Methods

Sample preparation

The ^{15}N-valine labelled samples were exclusively using DecylMaltoside as the membrane mimetic, to which the sample had been exchanged during the Strep-Tactin and the alprenolol ligand affinity Chromatography purification steps. Using an Amicon ultrafiltration device (15ml, 50 kD cut-off) the sample was concentrated to slightly more than 250 µl at a concentration of around 100 µM as estimated by 280 nm absorption. Initial NMR measurements were performed using the measurement buffer containing 250 µM NaCl, 50 mM Hepes pH 7.5, 0.1% DM, while the NaCl concentration was reduced to 100 mM, in an otherwise identical buffer, as indicated in relevant figure legends. All ligands were present at 1 mM, except cyanopindolol, which was used at 0.1 mM, with 1 % DMSO used a co-solvent. All samples contained around 5% D_2O. In experiments where the ligands were exchanged for another ligand of higher affinity, the sample was diluted in the measurement buffer containing the new ligand to 15 ml volume and left for around and hour before being re-concentrated, this step was repeated 4 times.

NMR spectroscopy of the hADRB1 receptor constructs

All measurements were recorded on a Bruker Avance 900 MHz spectrometer at 303 K. The pulse sequences for the [^{1}H, ^{15}N]-TROSY was obtained from Stephan Grzesiek as applied in (Isogai et al. 2016), with an INEPT delay set to 1.5 ms in order to reduce coherence loss, due to fast relaxation, instead of the usual $1/(4^1J([^1H, ^{15}N]))$ = 2.7 ms, where $^1J([^1H, ^{15}N])$ is 92Hz. The acquisition time would be in the range of 24 to 72 hours, although in some instance longer times were needed. All spectra were processed with exponential multiplication in the direct dimension with a line broadening of 20 Hz, while a quadratic sine with a shift of 2 was used for the indirect dimension.

6. Concluding Remarks

In this work a optimized construct for the human beta1-adrenergic receptor that was transferred from an analogous receptor construct of the turkey beta1-adrenergic receptor has been the subject of investigation. The construct was thermo-stabilized by means of mutations previously found in the turkey beta1-adrenergic receptor. Due to extremely low yield of purified receptor different routes were taken to improve the yield of the physiologically important receptor.

Firstly, different modifications of this construct were tested, including removal of the thermo-stabilizing mutations and different modification of the N-terminus. A construct exhibiting higher expression but only slightly reduced thermo-stability was identified. In this way it was shown that at least one of the stabilizing mutations were redundant in terms of stabilization, and could therefore be removed.

Secondly, rather than engineering the protein it-self, the expression system was engineered. Here an earlier and weaker promoter, the 39K promoter, than the conventionally used polyhedrin promoter was used. It was found that using this promoter, preferential surface expression could be obtained for functional proteins, while non-functional protein were exclusively found inside the cell. On the other hand, when using the polyhedrin promoter, constructs giving functional protein, exhibited a phenotype in which the majority of the protein was found intra-cellularly. Using the 39K promoter was also associated with a higher degree of a functional receptor being solubilized by the mild detergent DoDecylMaltoside (DDM). Importantly, the gp67 signal sequence improved the functional yield vastly in the minimized human beta1-adrenergic receptor constructs mentioned above, and it will be interesting to see if this also applies to other receptors.

Thirdly, a new purification method using the Twin-Strep tag and its binding partner Strep-Tactin was established. Convenient and fast purification was achieved, yielding extremely pure protein after a single purification step.

Lastly, it was possible to acquire 2D correlation [^{1}H, ^{15}N]-TROSY spectra of the human beta1-adrenergic receptor using the hADRB1-TS and the hADRB1-TS-YY constructs. The spectra obtained appeared to be those of a correctly folded receptor, showing dispersed resonances in the amide region of the spectrum, and most importantly, the spectra of the human ADRB1-TS resembled the spectra previously published spectra of the turkey TS-ADRB1 (Isogai et al. 2016). The well-folded-ness of the receptor is supported by the fact that the receptor was ligand affinity purified where only the ligand binding fraction of the protein is isolated. This also supports this construct as being an extremely promising candidate for crystallization and X-ray structural studies.

The receptor responded differently to different antagonists, and especially the antagonist S-propranolol induced a unique, and previously uncharacterized state of the receptor, since it was not used in the studies of the turkey TS-ADRB1 (Isogai et al. 2016). In addition, the 2D [^{1}H, ^{15}N]-TROSY spectra of the hADRB1-TS constructs had prominent peaks in unique regions as compared to the turkey TS-ADRB1. These peaks all responded differently, in terms of chemical shift changes, to the different ligands that were tested. Likewise, hADRB1 has three unique valines at the proline kink of TM5, and therefore, it is speculated that the valines might gives rise to the just mentioned unique peaks. It will be especially interesting to further characterize these peaks, and validate whether the relationship between the peaks and these valines holds true, since the particular residue might be important in the process of activating the receptor.

Acknowledgements

I would like to thank my "Doctor Vater" Prof. Gerhard Wider for his continued support, help and advice during this project. I will always remain impressed by his knowledge of the beautiful field of NMR, technically and theoretically, but also in other fields – it is always nice to discuss science with someone who expresses a true interest in whatever one is trying to show, communicate or debate.

I would also like to thank Prof. Gebhard Schertler and Dr. Dmitry Veprintsev. I am especially grateful to them for making this project possible, for providing plentiful valuable advice, and their interest in G-protein coupled receptors.

Florian Brückner deserves special mention for starting part of the project and for being such a good scientific and personal mentor. I still admire his way of working precisely and systematically. Our collaboration really reinforced for me, how good teamwork can be, and it will be something that I will always remember. Likewise, I would like to thank Xavi Deupi, for fun times in and outside the lab, his scientific help, and his always friendly personality. Similarly, I am grateful to Chayne Pisictelli for his help in the lab and always being happy to give friendly advice.

There are several colleagues I would like to thank as well including: Dawei Sun, we had many good scientific chats and many fun times in the lab; Maysam Mansouri, for his friendly spirit, help, advice and our endless scientific discussions; Xiaoqin He, for all the fun times we had discussing science and other things; Maria Mitsi, for our many discussions and fun in the lab; Ye Xie, for her helpfulness, humor and fun in the lab; Christian Peterhans, for his scientific insightfulness, for introducing me to Switzerland and the fun times; Fred Damberger, for his help at the NMR spectrometers and his general interest in many things – it was always fun to talk; Gregory Menchon,

for our friendship and scientific discussions; Dennis Diener, it was fantastic to work in the lab with you, discussing jazz and science – your knowledge will never cease to impress me.

My oldest friend Michael has always helped me, and I will always be grateful to him for his friendship. Johan also deserves my thanks for being the true and good friend that he is (almost forgot to mention that he too is a jazz man). Leon, one of my newest friends, whom it has been great to get to know, with his enthusiastic personality – I am grateful for his help. I am of course also grateful to David and Sarah, for introducing me to Switzerland and their friendship and for our shared interest in jazz. Similarly, I am very thankful to Ulrich, and I had a great time with you and your people.

Then I would like to thank my family for their support. They have really been missed during this time; my siblings and my parents. My dad deserves special thanks for introducing me to science and technology throughout my entire life, but especially as a kid and young adult – It is funny to think of that in the rearview mirror, I should maybe have studied electrical engineering (the first thing I wanted to study), which would have been advantageous for the NMR. On the other hand, had I studied that I would probably never really have encountered NMR... My mom deserves my gratitude for always listening and encouraging me and especially for being who she is. My siblings both know that I really miss them and that I always enjoy being with them.

Lastly and most importantly, I would like to thank my girlfriend, Katya, for her love and support, and her kindness. For always laughing at my jokes – that is a unique talent she possesses. I love you.

References

Alexandrov, A. I., et al. (2008). "Microscale fluorescent thermal stability assay for membrane proteins." Structure **16**(3): 351-359.

Azzi, M., et al. (2001). "Allosteric effects of G protein overexpression on the binding of beta-adrenergic ligands with distinct inverse efficacies." Mol Pharmacol **60**(5): 999-1007.

Baker, J. G. (2010). "A full pharmacological analysis of the three turkey beta-adrenoceptors and comparison with the human beta-adrenoceptors." PLoS One **5**(11): e15487.

Ballesteros, J. A. and H. Weinstein (1995). "Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors." Methods in neurosciences **25**: 366-428.

Beja, O., et al. (2000). "Bacterial rhodopsin: Evidence for a new type of phototrophy in the sea." Science **289**(5486): 1902-1906.

Beltzer, J. P., et al. (1991). "Charged residues are major determinants of the transmembrane orientation of a signal-anchor sequence." J Biol Chem **266**(2): 973-978.

Berger, I., et al. (2004). "Baculovirus expression system for heterologous multiprotein complexes." Nat Biotechnol **22**(12): 1583-1587.

Bieniossek, C., et al. (2008). "MultiBac: multigene baculovirus-based eukaryotic protein complex production." Curr Protoc Protein Sci **Chapter 5**: Unit 5 20.

Bill, R. M., et al. (2011). "Overcoming barriers to membrane protein structure determination." Nat Biotechnol **29**(4): 335-340.

Brueckner, F., et al. (2013). "Structure of beta-adrenergic receptors." Methods Enzymol **520**: 117-151.

Caffrey, M. (2015). "A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes." Acta Crystallogr F Struct Biol Commun **71**(Pt 1): 3-18.

Caron, M. G., et al. (1979). "Affinity chromatography of the beta-adrenergic receptor." J Biol Chem **254**(8): 2923-2927.

Carpenter, B. and C. G. Tate (2016). "Engineering a minimal G protein to facilitate crystallisation of G protein-coupled receptors in their active conformation." Protein Eng Des Sel **29**(12): 583-594.

Cherezov, V., et al. (2007). "High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor." Science **318**(5854): 1258-1265.

Cooke, R. M., et al. (2015). "Structures of G protein-coupled receptors reveal new opportunities for drug discovery." Drug Discov Today **20**(11): 1355-1364.

Dixon, R. A., et al. (1986). "Cloning of the gene and cDNA for mammalian beta-adrenergic receptor and homology with rhodopsin." Nature **321**(6065): 75-79.

Drew, D. E., et al. (2001). "Green fluorescent protein as an indicator to monitor membrane protein overexpression in Escherichia coli." FEBS Lett **507**(2): 220-224.

Egloff, P., et al. (2014). "Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in Escherichia coli." Proc Natl Acad Sci U S A **111**(6): E655-662.

Fredriksson, R., et al. (2003). "The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints." Mol Pharmacol **63**(6): 1256-1272.

Frielle, T., et al. (1987). "Cloning of the cDNA for the human beta 1-adrenergic receptor." Proc Natl Acad Sci U S A **84**(22): 7920-7924.

Galandrin, S., et al. (2008). "Conformational rearrangements and signaling cascades involved in ligand-biased mitogen-activated protein kinase signaling through the beta1-adrenergic receptor." Mol Pharmacol **74**(1): 162-172.

Goncalves, J. A., et al. (2010). "Highly conserved tyrosine stabilizes the active state of rhodopsin." Proc Natl Acad Sci U S A **107**(46): 19861-19866.

Gossert, A. D., et al. (2011). "A simple protocol for amino acid type selective isotope labeling in insect cells with improved yields and high reproducibility." J Biomol NMR **51**(4): 449-456.

Grimm, D., et al. (2006). "Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways." Nature **441**(7092): 537-541.

Guan, X. M., et al. (1992). "Enhancement of membrane insertion and function in a type IIIb membrane protein following introduction of a cleavable signal peptide." J Biol Chem **267**(31): 21995-21998.

Guarino, L. A. and M. W. Smith (1990). "Nucleotide sequence and characterization of the 39K gene region of Autographa californica nuclear polyhedrosis virus." Virology **179**(1): 1-8.

Guarino, L. A. and M. D. Summers (1986). "Functional mapping of a trans-activating gene required for expression of a baculovirus delayed-early gene." J Virol **57**(2): 563-571.

He, X., et al. (2009). "Protein NMR chemical shift calculations based on the automated fragmentation QM/MM approach." J Phys Chem B **113**(30): 10380-10388.

Hitchman, R. B., et al. (2007). "Quantitative real-time PCR for rapid and accurate titration of recombinant baculovirus particles." Biotechnology and bioengineering **96**(4): 810-814.

Isberg, V., et al. (2015). "Generic GPCR residue numbers - aligning topology maps while minding the gaps." Trends in Pharmacological Sciences **36**(1): 22-31.

Isogai, S., et al. (2016). "Backbone NMR reveals allosteric signal transduction networks in the beta1-adrenergic receptor." Nature **530**(7589): 237-241.

Kawate, T. and E. Gouaux (2006). "Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins." Structure **14**(4): 673-681.

Kim, I. M., et al. (2008). "Beta-blockers alprenolol and carvedilol stimulate beta-arrestin-mediated EGFR transactivation." Proc Natl Acad Sci U S A **105**(38): 14555-14560.

Klein-Seetharaman, J., et al. (2002). "Solution NMR spectroscopy of [alpha -15N]lysine-labeled rhodopsin: The single peak observed in both conventional and TROSY-type HSQC spectra is ascribed to Lys-339 in the carboxyl-terminal peptide sequence." Proc Natl Acad Sci U S A **99**(6): 3452-3457.

Kofuku, Y., et al. (2012). "Efficacy of the beta(2)-adrenergic receptor is determined by conformational equilibrium in the transmembrane region." Nat Commun **3**: 1045.

Korepanova, A., et al. (2009). "Expression and purification of human TRPV1 in baculovirus-infected insect cells for structural studies." Protein Expr Purif **65**(1): 38-50.

Lagerstrom, M. C. and H. B. Schioth (2008). "Structural diversity of G protein-coupled receptors and significance for drug discovery." Nat Rev Drug Discov **7**(4): 339-357.

Lander, E. S., et al. (2001). "Initial sequencing and analysis of the human genome." Nature **409**(6822): 860-921.

le Maire, M., et al. (2000). "Interaction of membrane proteins and lipids with solubilizing detergents." Biochim Biophys Acta **1508**(1-2): 86-111.

Lin, C. H. and D. L. Jarvis (2013). "Utility of temporally distinct baculovirus promoters for constitutive and baculovirus-inducible transgene expression in transformed insect cells." J Biotechnol **165**(1): 11-17.

Lv, X., et al. (2016). "In vitro expression and analysis of the 826 human G protein-coupled receptors." Protein Cell **7**(5): 325-337.

Mackenzie, D., et al. (1984). "Localization of Binding-Sites for Carboxyl Terminal Specific Anti-Rhodopsin Monoclonal-Antibodies Using Synthetic Peptides." Biochemistry **23**(26): 6544-6549.

Matsuura, Y., et al. (1987). "Baculovirus expression vectors: the requirements for high level expression of proteins, including glycoproteins." J Gen Virol **68 (Pt 5)**: 1233-1250.

Miller, J. L. and C. G. Tate (2011). "Engineering an ultra-thermostable beta(1)-adrenoceptor." J Mol Biol **413**(3): 628-638.

Miller-Gallacher, J. L., et al. (2014). "The 2.1 A resolution structure of cyanopindolol-bound beta1-adrenoceptor identifies an intramembrane Na+ ion that stabilises the ligand-free receptor." PLoS One **9**(3): e92727.

Minneman, K. P., et al. (1980). "A comparison of the beta-adrenergic receptor of the turkey erythrocyte with mammalian beta1 and beta2 receptors." Mol Pharmacol **17**(1): 1-7.

Mukherjee, C., et al. (1975). "Identification of adenylate cyclase-coupled beta-adrenergic receptors in frog erythrocytes with (minus)-[3-H] alprenolol." J Biol Chem **250**(13): 4869-4876.

Nasr, M. L., et al. (2017). "Covalently circularized nanodiscs for studying membrane proteins and viral entry." Nat Methods **14**(1): 49-52.

Newby, Z. E., et al. (2009). "A general protocol for the crystallization of membrane proteins for X-ray structural investigation." Nature Protocols **4**(5): 619-637.

Newstead, S., et al. (2007). "High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in Saccharomyces cerevisiae." Proc Natl Acad Sci U S A **104**(35): 13936-13941.

Opitz, C., et al. (2015). "An economic approach to efficient isotope labeling in insect cells using homemade 15N-, 13C- and 2H-labeled yeast extracts." J Biomol NMR **62**(3): 373-385.

Palczewski, K., et al. (2000). "Crystal structure of rhodopsin: A G protein-coupled receptor." Science **289**(5480): 739-745.

Pervushin, K., et al. (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." Proc Natl Acad Sci U S A **94**(23): 12366-12371.

Prive, G. G. (2007). "Detergents for the stabilization and crystallization of membrane proteins." Methods **41**(4): 388-397.

Rasmussen, S. G., et al. (2007). "Crystal structure of the human beta2 adrenergic G-protein-coupled receptor." Nature **450**(7168): 383-387.

Rasmussen, S. G., et al. (2011). "Crystal structure of the beta2 adrenergic receptor-Gs protein complex." Nature **477**(7366): 549-555.

Reeves, P. J., et al. (2002). "Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line." Proc Natl Acad Sci U S A **99**(21): 13419-13424.

Rosenbaum, D. M., et al. (2007). "GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function." Science **318**(5854): 1266-1273.

Rosenbaum, D. M., et al. (2009). "The structure and function of G-protein-coupled receptors." Nature **459**(7245): 356-363.

Saliba, R. S., et al. (2002). "The cellular fate of mutant rhodopsin: quality control, degradation and aggresome formation." J Cell Sci **115**(Pt 14): 2907-2918.

Sarkar, C. A., et al. (2008). "Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity." Proc Natl Acad Sci U S A **105**(39): 14808-14813.

Schmidt, T. G. M., et al. (1996). "Molecular interaction between the Strep-tag affinity peptide and its cognate target, streptavidin." Journal of Molecular Biology **255**(5): 753-766.

Schmidt, T. G. M. and A. Skerra (2007). "The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins." Nature Protocols **2**(6): 1528-1535.

Serrano-Vega, M. J., et al. (2008). "Conformational thermostabilization of the beta1-adrenergic receptor in a detergent-resistant form." Proc Natl Acad Sci U S A **105**(3): 877-882.

Serrano-Vega, M. J. and C. G. Tate (2009). "Transferability of thermostabilizing mutations between beta-adrenergic receptors." Mol Membr Biol **26**(8): 385-396.

Simon, M. I., et al. (1991). "Diversity of G-Proteins in Signal Transduction." Science **252**(5007): 802-808.

Skerra, A. and T. G. Schmidt (1999). "Applications of a peptide ligand for streptavidin: the Strep-tag." Biomol Eng **16**(1-4): 79-86.

Skora, L., et al. (2015). "Isotope labeling of proteins in insect cells." Methods Enzymol **565**: 245-288.

Spiess, M. (1995). "Heads or Tails - What Determines the Orientation of Proteins in the Membrane." Febs Letters **369**(1): 76-79.

Standfuss, J., et al. (2007). "Crystal structure of a thermally stable rhodopsin mutant." J Mol Biol **372**(5): 1179-1188.

Stevens, R. C., et al. (2013). "The GPCR Network: a large-scale collaboration to determine human GPCR structure and function." Nat Rev Drug Discov **12**(1): 25-34.

Tate, C. G. and R. D. Blakely (1994). "The Effect of N-Linked Glycosylation on Activity of the Na+-Dependent and Cl--Dependent Serotonin Transporter Expressed Using Recombinant Baculovirus in Insect Cells." Journal of Biological Chemistry **269**(42): 26303-26310.

Tate, C. G., et al. (2003). "Comparison of seven different heterologous protein expression systems for the production of the serotonin transporter." Biochim Biophys Acta **1610**(1): 141-153.

Thomas, J. and C. G. Tate (2014). "Quality control in eukaryotic membrane protein overproduction." J Mol Biol **426**(24): 4139-4154.

Trowitzsch, S., et al. (2010). "New baculovirus expression tools for recombinant protein complex production." J Struct Biol **172**(1): 45-54.

Tucker, J. and R. Grisshammer (1996). "Purification of a rat neurotensin receptor expressed in Escherichia coli." Biochem J **317 (Pt 3)**: 891-899.

Voss, S. and A. Skerra (1997). "Mutagenesis of a flexible loop in streptavidin leads to higher affinity for the Strep-tag II peptide and improved performance in recombinant protein purification." Protein Eng **10**(8): 975-982.

Wallin, E. and G. von Heijne (1995). "Properties of N-terminal tails in G-protein coupled receptors: a statistical study." Protein Eng **8**(7): 693-698.

Warne, T., et al. (2003). "Expression and purification of truncated, non-glycosylated turkey beta-adrenergic receptors for crystallization." Biochim Biophys Acta **1610**(1): 133-140.

Warne, T., et al. (2012). "Crystal structures of a stabilized beta1-adrenoceptor bound to the biased agonists bucindolol and carvedilol." Structure **20**(5): 841-849.

Warne, T., et al. (2008). "Structure of a beta1-adrenergic G-protein-coupled receptor." Nature **454**(7203): 486-491.

Warne, T., et al. (2009). "Development and crystallization of a minimal thermostabilised G protein-coupled receptor." Protein Expr Purif **65**(2): 204-213.

Wettschureck, N. and S. Offermanns (2005). "Mammalian G proteins and their cell type specific functions." Physiological Reviews **85**(4): 1159-1204.

White, J. F., et al. (2012). "Structure of the agonist-bound neurotensin receptor." Nature **490**(7421): 508-513.

Wishart, D. S., et al. (1995). "1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects." J Biomol NMR **5**(1): 67-81.

Wu, G. (2013). "Identification of endoplasmic reticulum export motifs for G protein-coupled receptors." Methods Enzymol **521**: 189-202.

Wu, H., et al. (2012). "Structure of the human kappa-opioid receptor in complex with JDTic." Nature **485**(7398): 327-332.

Xiao, S., et al. (2013). "Transient and stable expression of the neurotensin receptor NTS1: a comparison of the baculovirus-insect cell and the T-REx-293 expression systems." PLoS One **8**(5): e63679.

Yeliseev, A., et al. (2016). "Application of Strep-Tactin XT for affinity purification of Twin-Strep-tagged CB2, a G protein-coupled cannabinoid receptor." Protein Expr Purif **131**: 109-118.

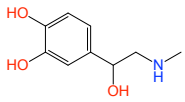
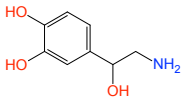
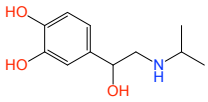
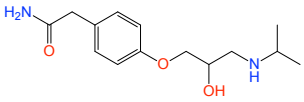
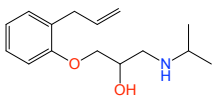
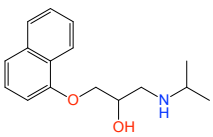
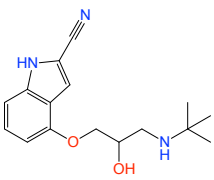
Zhou, Y. and J. U. Bowie (2000). "Building a thermostable membrane protein." J Biol Chem **275**(10): 6975-6979.

Zimmermann, R., et al. (2011). "Protein translocation across the ER membrane." Biochim Biophys Acta **1808**(3): 912-924.

Appendix

A1. Structures of ligands

Tabel A 1. Chemical structures of natural ligands of the human beta1-adrenergic receptor, and all ligands used in this study.

Structure	Common name(s)[1]	Comment[1]
	Adrenaline Epinephrine	Agonist, naturally occuring It stimulates both the alpha- and beta-adrenergic systems Hormone
	Noradrenaline Norepinephrine	Agonist, naturally occuring Hormone Neurotransmitter
	Isoproterenol Isoprenaline	Agonist, artificial
	Atenolol Tenormin	Antagonist, artificial Beta-blocker
	Alprenolol Gubernal	Antagonist, artificial Beta-blocker
	Propranolol Obzidan	Antagonist, artificial Beta-blocker
	Cyanopindolol	Antagonist, artificial Beta-blocker

[1]Information obtained from pubchem (pubchem.ncbi.nlm.nih.gov/)

A2. Tosoh Super SW3000 column

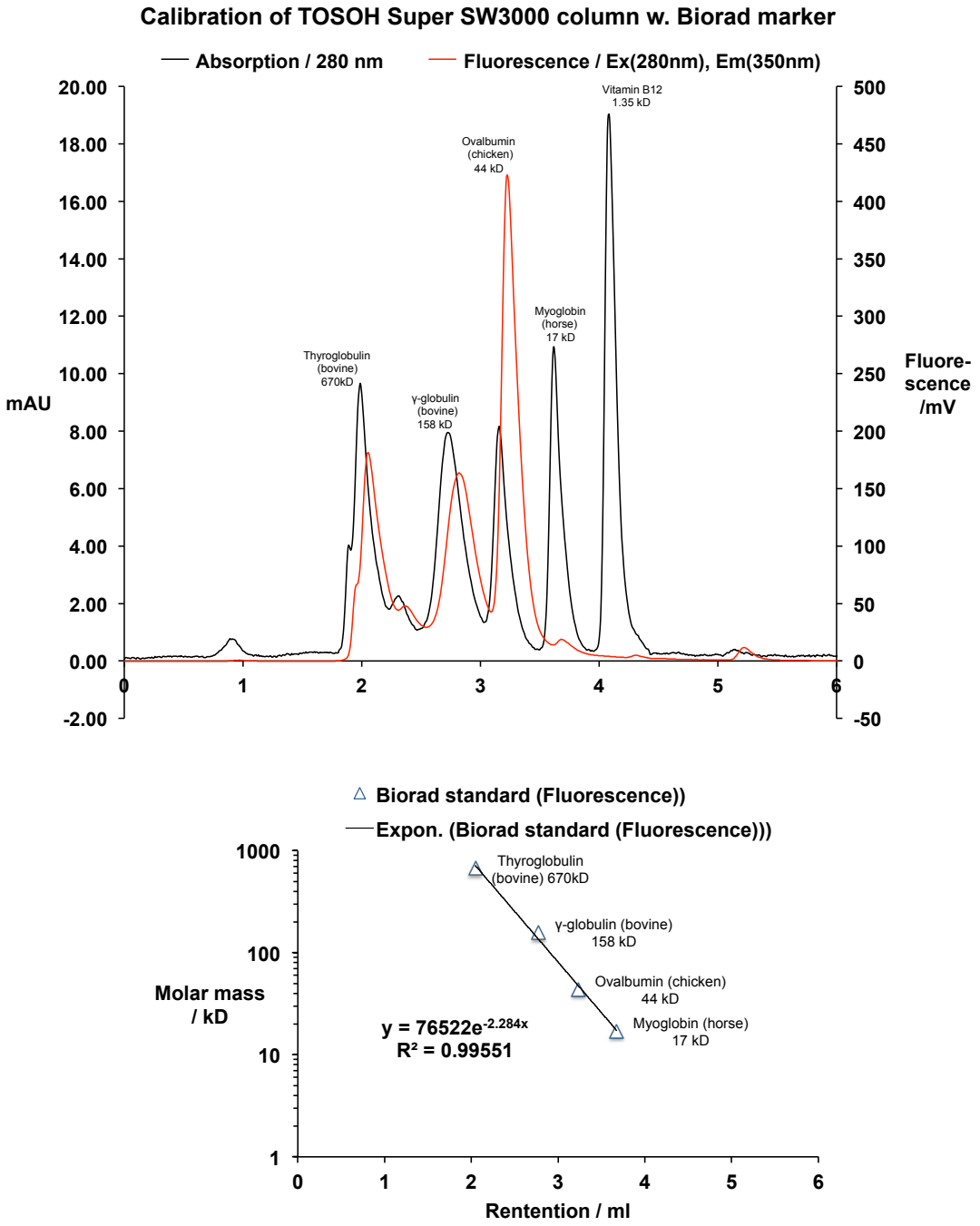


Figure A 1. Calibration of the TOSOH Super SW3000 analytical size exclusion column.

The Biorad standard was run on the column using an ÄKTA ETTAN HPLC system, with UV and fluorescence detection. UV peaks values (ml) of each of the marker constituents were corrected for the mean additional delay of the respective fluorescence peaks, and plotted against the corresponding molar masses (kD).

A3. NMR chemical shift tables

Tabel A 2. Valine chemical shifts in the hADRB1-TS-Twin-Strep construct.
Chemical shift values and tentative assignment for peaks picked with CARA, in the [^{1}H, ^{15}N]-TROSY spectra of the hADRB1-TS-Twin-Strep construct (Figure 34, Figure 35).

	Alprenolol		S-Propranolol		Position of residue[1]	
Peak #	^{1}H δ (ppm)	^{15}N δ (ppm)	^{1}H δ (ppm)	^{15}N δ (ppm)	Human ADRB1-TS[2]	Comment[3]
1	8.13	109.74	8.14	109.09	189 (4.56)	
2	6.98	113.86	6.97	113.77		
3	7.13	116.06	7.04	115.84		
4	7.18	119.59	7.17	119.60		
5	8.03	115.42	7.99	115.45		
6	7.59	117.48	7.58	117.33		
7	7.49	117.58	7.48	117.45		
8	7.42	118.17	7.54	118.19		
9	8.82	116.31	8.92	116.66		
10	8.38	118.02	8.37	118.02		
11	8.29	117.21	8.20	117.37		
12	7.77	118.27	n. a.[4]	n. a.		
13	7.60	119.37	7.63	119.43		
14	7.59	120.45	n. a.	n. a.		
15	7.42	121.09	7.42	121.23	142 (3.36)	
16	8.91	118.15	8.88	117.95	243 (5.57)	
17	8.65	118.76	8.62	118.73		
18	8.41	120.57	8.42	120.63		
19	8.15	119.00	8.13	119.17		Multi
20	8.04	119.13	8.04	119.13		
21	8.03	120.48	8.05	120.62		
22	7.85	121.24	7.87	121.09		
23	8.26	121.71	8.20	121.69		
24	n. a.	n. a.	n. a.	n. a.		1D4
25	7.98	122.87	8.00	122.73	107 (2.53)	Stab
26	7.89	126.44	7.94	126.57	219 (ECL2)	
27	9.14	120.57	n. a.	n. a.		
28	8.78	119.61	8.76	119.56		
29	8.83	121.60	8.74	121.51	247 (5.61)	
30	8.59	124.92	8.59	124.65		
31	9.24	122.27	9.22	122.23	332 (6.43)	
32	n. a.	n. a.	7.73	120.11		
33	n. a.	n. a.	7.81	120.35		

[1]Position in wild-type receptors (Ballesteros-Weinstein number in parenthesis).
[2]Tentative Assignment transferred from spectra of the hADRB1-Twin-Strep-1D4.
[3]Comment: 1D4 (Valine in 1D4-tag (TETSQVAPA)); Stab (stabilizing mutation); Multi (Multiple coinciding peaks a possibility).
[4] n. a.: not applicable due to missing peak.

Tabel A 3. Valine chemical shifts in the hADRB1-TS-YY-Twin-Strep construct.
Chemical shift values and tentative assignment for peaks picked with CARA, in the [^{1}H, ^{15}N]-TROSY spectrum of the gp67^{SS}hADRB1-TS-YY-Twin-Strep construct (Figure 36).

	Cyanopindolol		Position of residue[1]		
Peak #	^{1}H δ (ppm)	^{15}N δ (ppm)	Human ADRB1-TS[2]	Turkey TS-ADRB1[3]	Comment[4]
1	8.21	109.33	189 (4.56)	172 (4.56)	
2	7.09	114.22			
3	7.11	115.70			
4	7.21	119.92			
5	7.94	115.36			
6	7.57	117.34			
7	7.47	117.30			
8	7.49	118.40			
9	8.80	116.06			
10	8.42	118.11			
11	8.28	117.35			
12	7.85	118.93			
13	7.68	119.55			
14	7.59	120.73			
15	7.36	121.18	142 (3.36)	125 (3.36)	
16	8.89	118.17	243 (5.57)	226 (5.57)	
17	8.68	117.96			
18	8.45	120.89			
19	8.13	119.26			Multi
20	8.07	119.32			
21	8.05	120.73			
22	7.84	121.02			
23	8.27	121.78			
24	n. a.[4]	n. a.			1D4
25	7.95	123.21	107 (2.53)	90 (2.53)	Stab
26	7.91	127.29	219 (ECL2)	202 (ECL2)	
27	9.10	120.67			
28	8.73	119.40			
29	n. a.	n. a.	247 (5.61)	230 (5.61)	
30	8.58	124.61			
31	9.16	122.41	332 (6.43)	298 (6.43)	
32	7.72	120.43			
33	n. a.	n. a.			
34	7.83	122.66			

[1]Position in wild-type receptors (Ballesteros-Weinstein number in parenthesis).
[2]Tentative Assignment transferred from spectra of the hADRB1-Twin-Strep-1D4.
[3]Comment: 1D4 (Valine in 1D4-tag (TETSQVAPA)); Stab (stabilizing mutation);
[4] n. a.: not applicable due to missing peak

Tabel A 4. Valine chemical shifts in the turkey TS-ADRB1 construct (Isogai et al. 2016)

	Alprenolol		Cyanopindolol		Position of residue[1]	
Peak #	^{1}H δ (ppm)	^{15}N δ (ppm)	^{1}H δ (ppm)	^{15}N δ (ppm)	Turkey TS-ADRB1[1]	Comment[2]
1	8.21	108.74	8.29	108.37	Val172 (4.56)	
2	7.30	115.94	7.24	115.45	Val314 (6.59)	Unique
3	8.26	116.06	8.26	115.93	Val103 (2.65)	Tentative
4	7.92	117.80	n. a.[4]	n. a.		
5	7.89	119.29	7.89	119.29	Val280 (6.25)	Unique
6	7.77	118.92	7.80	118.85	Val62 (1.53)	
7	7.69	119.07	7.69	119.07	Val134 (3.45)	
8	7.58	118.96	7.59	118.56	Val94 (2.56)	
9	7.45	118.53	7.50	118.20	Val122 (3.33)	Tentative
10	8.33	117.53	n. a.	n. a.		
11	8.14	118.61	8.20	118.18	Val326 (7.36)	
12	8.12	119.85	8.13	119.64		
13	8.05	121.24	8.17	121.74		
14	7.74	120.47	7.74	120.47	Val160 (4.44)	Unique
15	7.49	121.55	7.44	121.05	Val125 (3.36)	
16	8.92	118.76	8.92	118.76	Val226 (5.57)	
17	8.64	118.73	8.64	118.61		
18	8.56	120.82	8.56	120.82		
19	8.00	122.67	8.03	122.51	Val90 (2.53)	Stabilize
20	7.58	123.99	n. a.	n. a.		
21	9.51	121.32	n. a.	n. a.	Val95 (2.57)	Tentative
22	9.26	122.60	9.26	122.39	Val298 (6.43)	
23	8.96	119.85	8.96	119.85	Val89 (2.52)	Tentative
24	8.78	122.21	8.78	122.21	Val230 (5.61)	
25	8.52	122.48	8.48	122.32	Val165 (4.49)	Tentative
26	7.92	125.23	7.91	125.88	Val202 (ECL2)	
27	7.95	121.86	7.99	121.74	Val129 (3.40)	
28	7.88	119.65	7.95	119.66	Val102 (2.64)	

[1]Position in wild-type receptor (Ballesteros-Weinstein number in parenthesis).

[3]Comment: Tentative (peak assigned tentatively, otherwise by mutagenesis); Unique (Unique valine in the turkey TS-ADRB1 as compared to the hADRB1-TS) (Isogai et al. 2016)

[4] n. a.: not applicable due to missing peak

A4. Structure of the ultra-stabilized turkey Beta1-Adrenergic Receptor

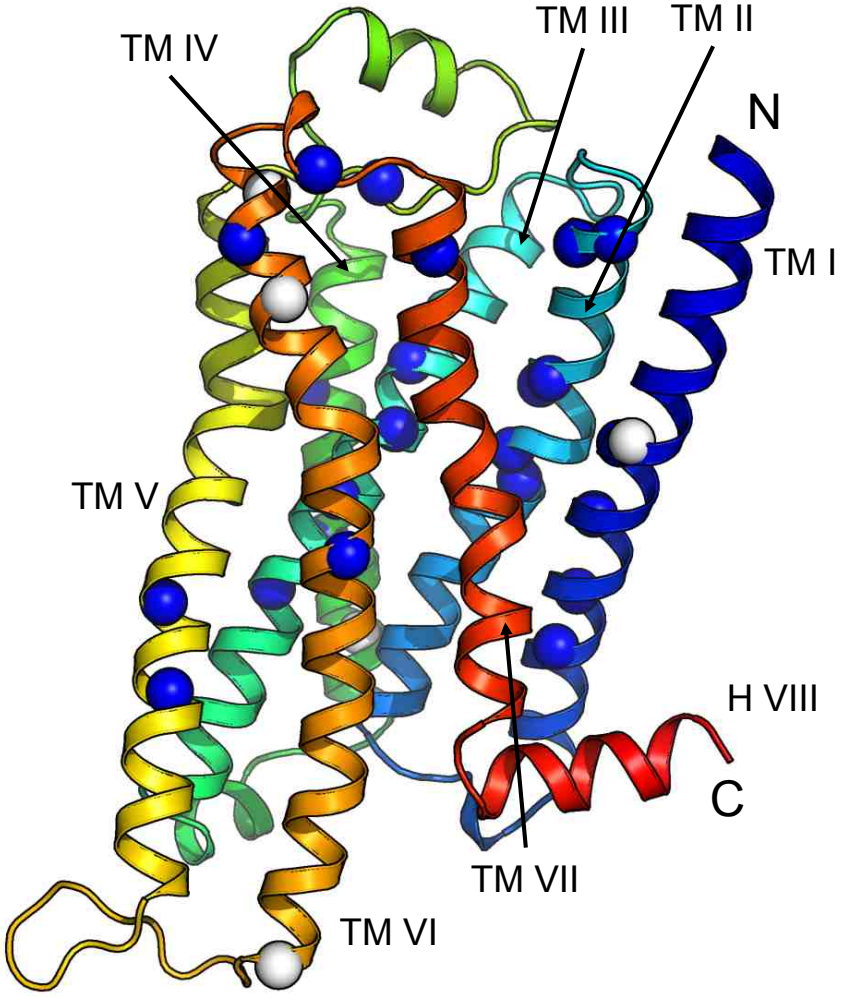


Figure A 2. Structure of the ultra-stabilized-turkey ADRB1 with nitrogens of valines as spheres.

Structure of the ultrastabilized turkey ADRB1 construct (4bvn) shown as ribbon diagram, with N and C, indicating N- and C-terminus, respectively. Each transmembrane segment is indicated followed by TM followed by a roman number indicating the respective number of this segment, and the amphipathic helix is indicated with and H followed the roman number VIII. Nitrogen atom of valine residues, which have been used in all experiments (Isogai et al. 2016) with selective labeling are shown as blue balls if valine is conserved in terms of position in the 3D-structure and primary structure, and white balls if the valine residue is unique, as compared to the human hADRB1-TS.

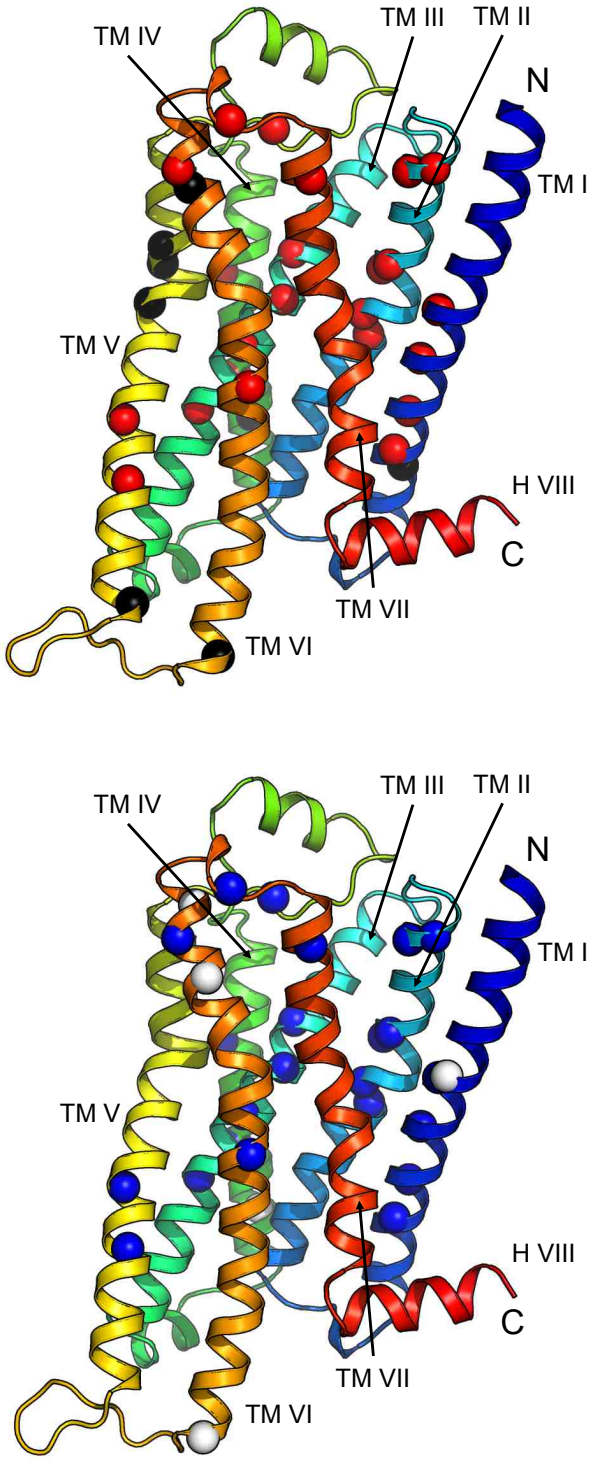


Figure A 3. Comparison of valine position in the hADRB1-TS and the turkey TS-ADRB1.

Side by side comparison of the hADRB1-TS model (**top**), and the ultra-stabilized turkey ADRB1 (4bvn) structure (**bottom**), shown in Figure 29, Figure A 2, respectively. See legend of respective figures for further details.

A5. Individual NMR spectra

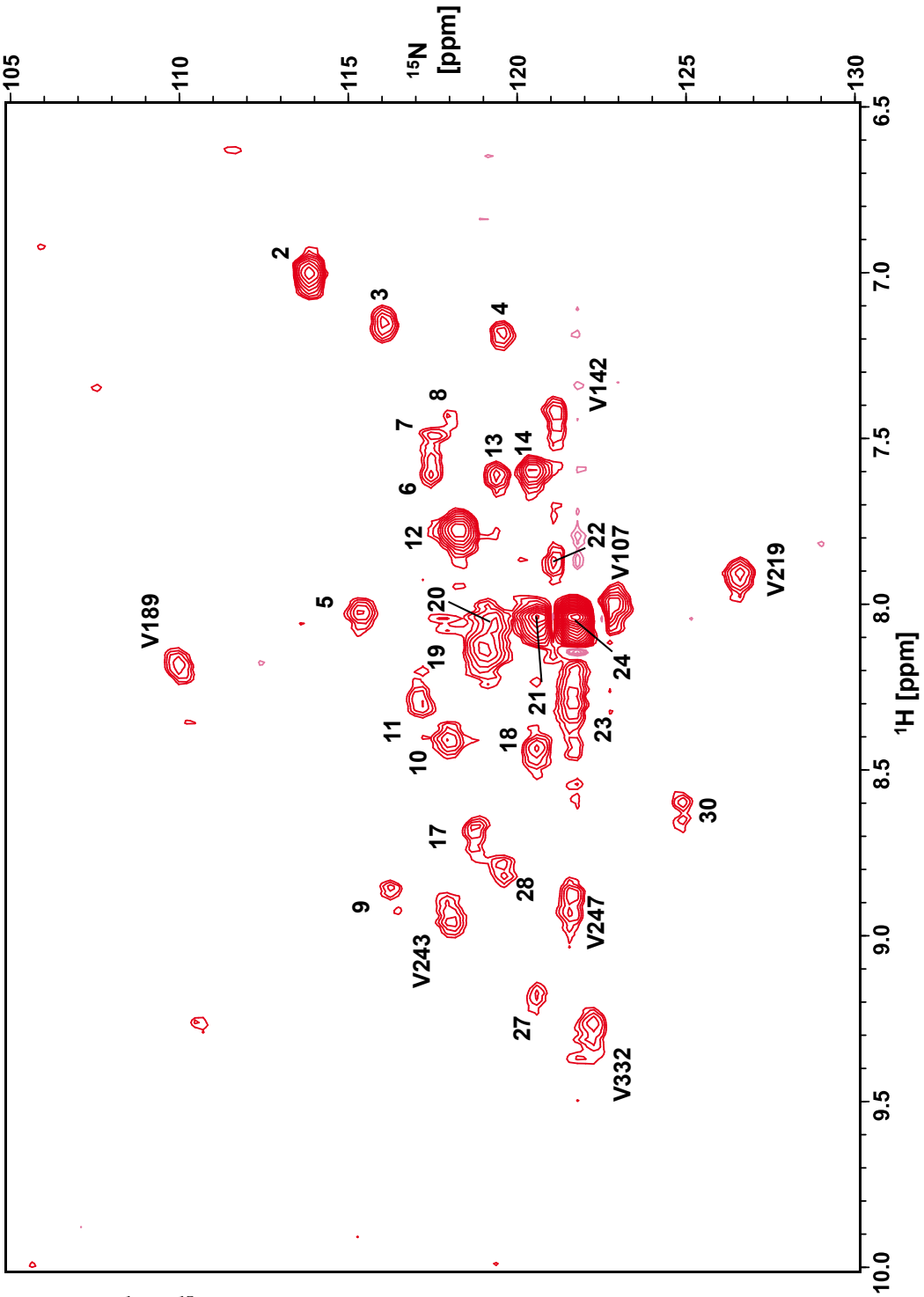


Figure A 4. [^{1}H, ^{15}N]-TROSY spectrum of hADRB1-TS-Twin-Strep1D4 with the antagonist alprenolol.

Experimental details described in legend of Figure 30.

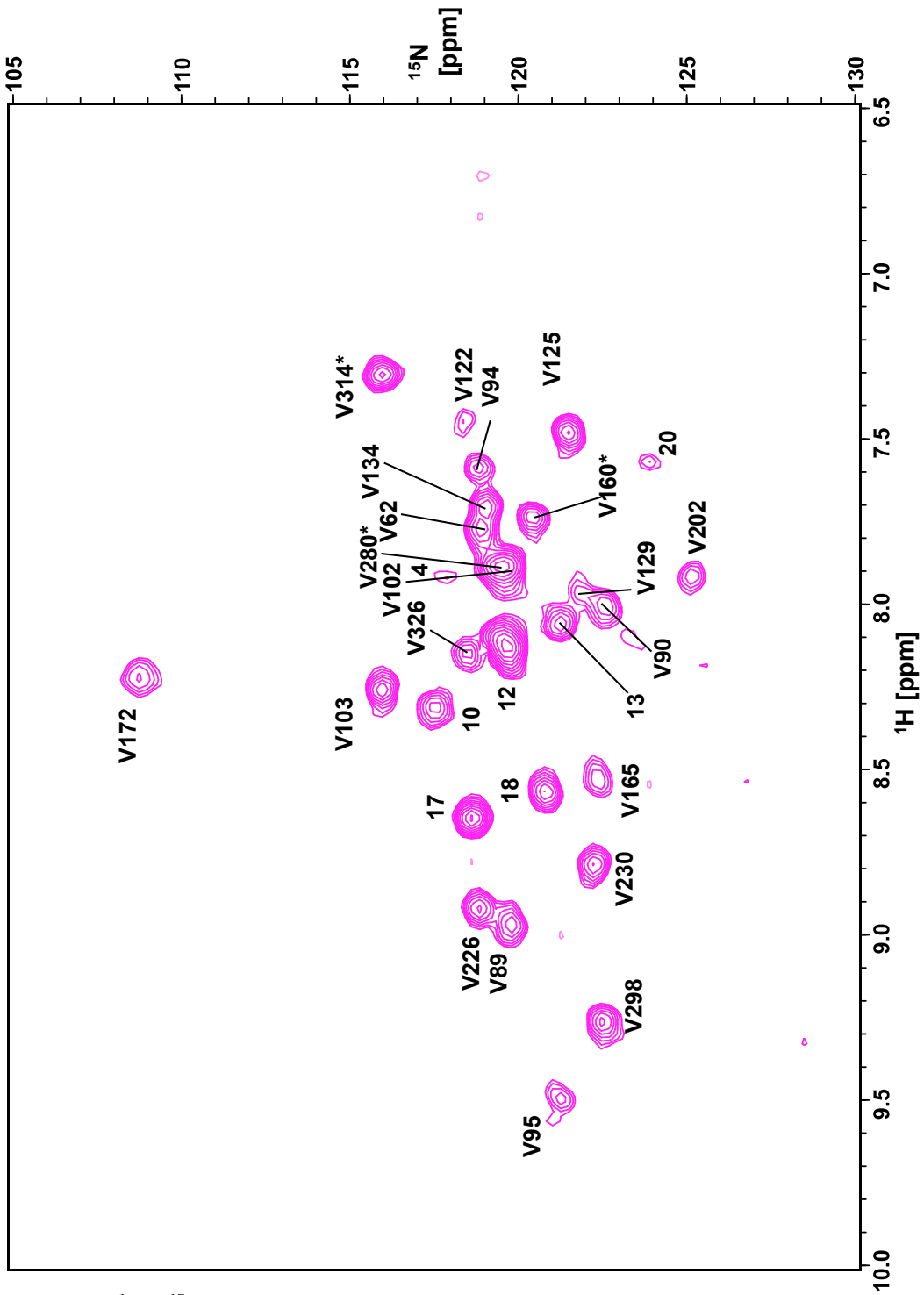


Figure A 5. [^{1}H, ^{15}N]-TROSY spectrum of turkey TS-ADRB1 with the antagonist alprenolol.

Experimental details described in and spectrum obtained from (Isogai et al. 2016).

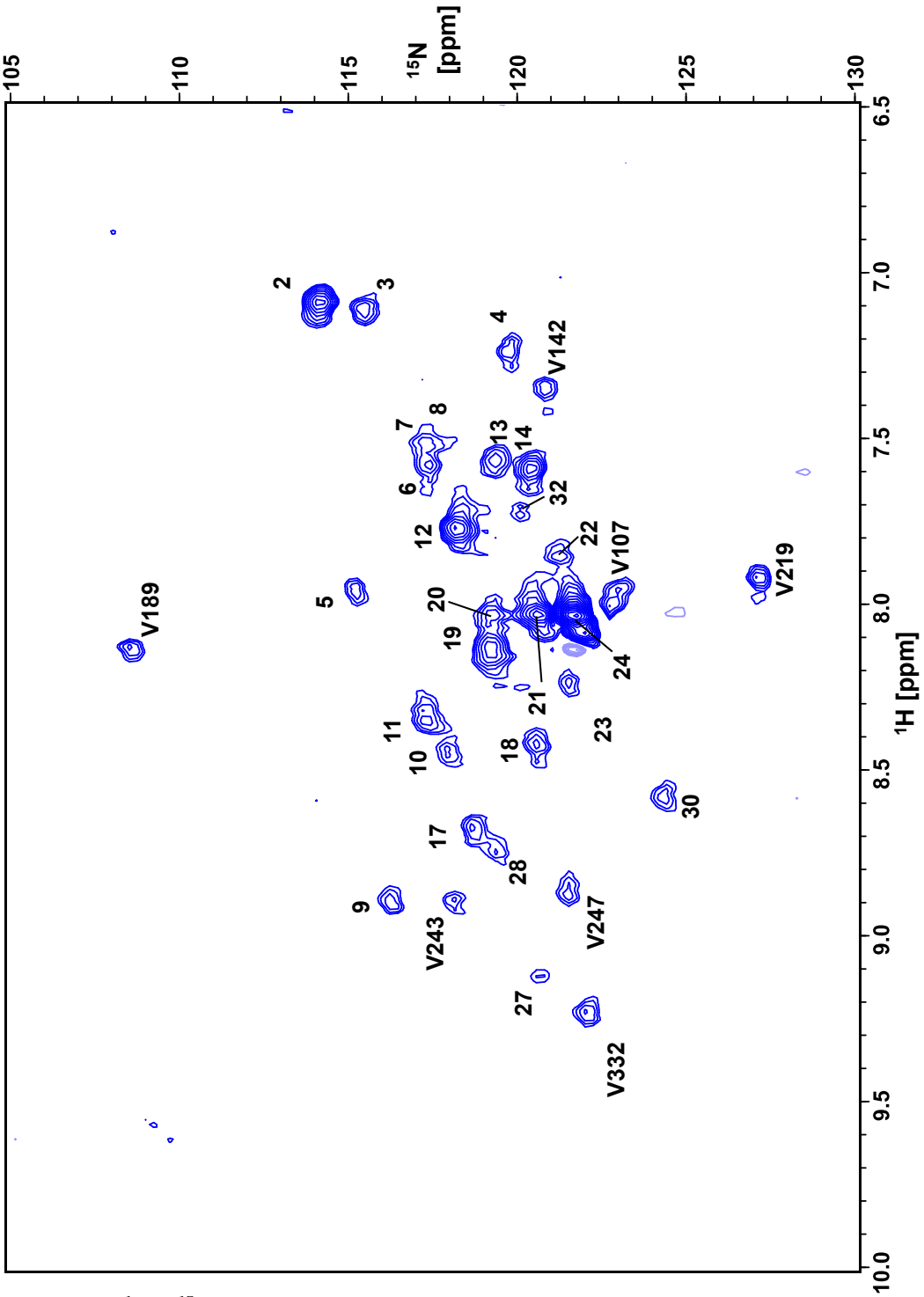


Figure A 6. [^{1}H, ^{15}N]-TROSY spectrum of hADRB1-TS-Twin-Strep1D4 with the antagonist cyanopindolol.

Experimental details described in legend of Figure 30

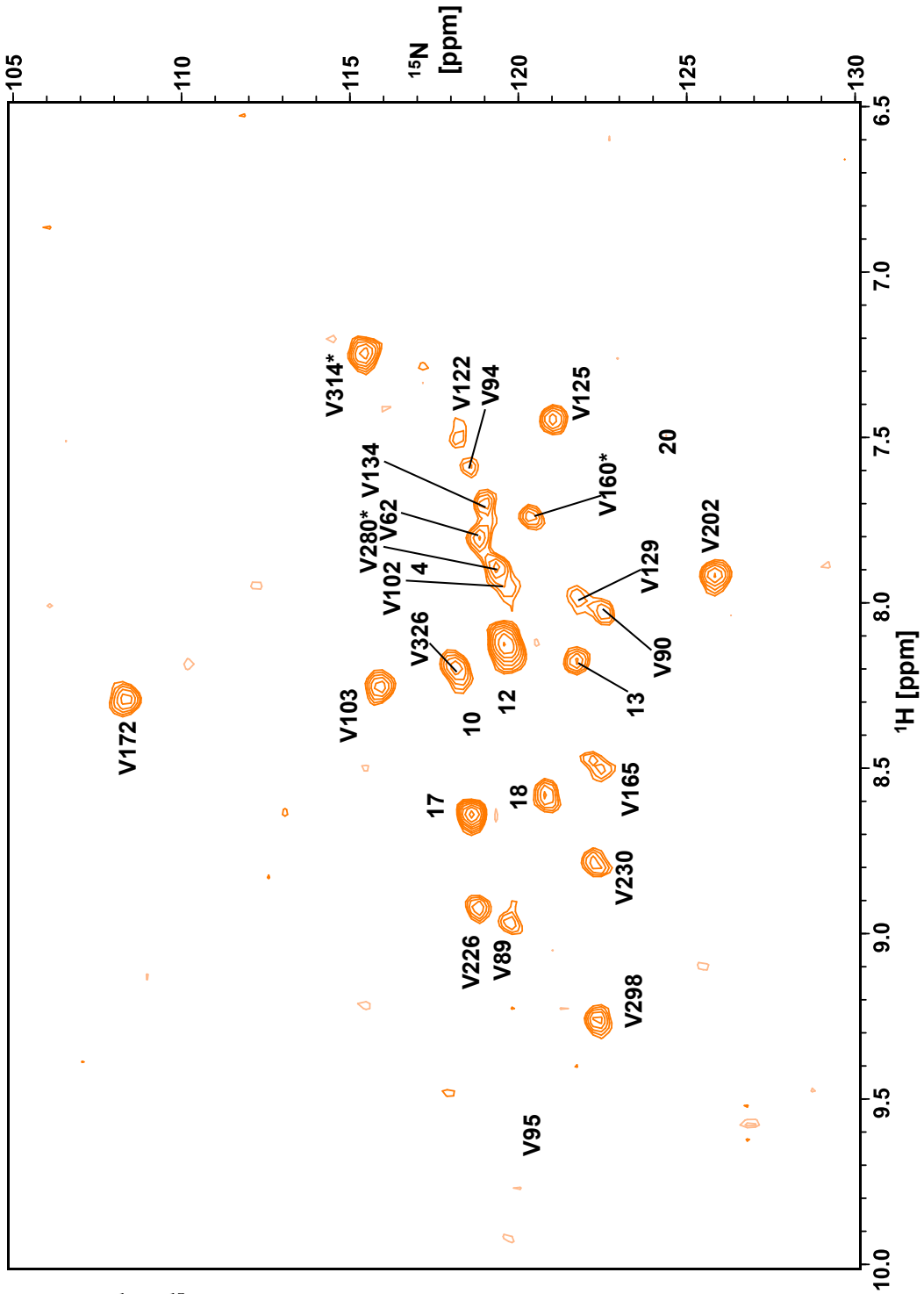


Figure A 7. [^{1}H, ^{15}N]-TROSY spectrum of turkey TS-ADRB1 with the antagonist cyanopindolol.

Experimental details described in and spectrum obtained from (Isogai et al. 2016).

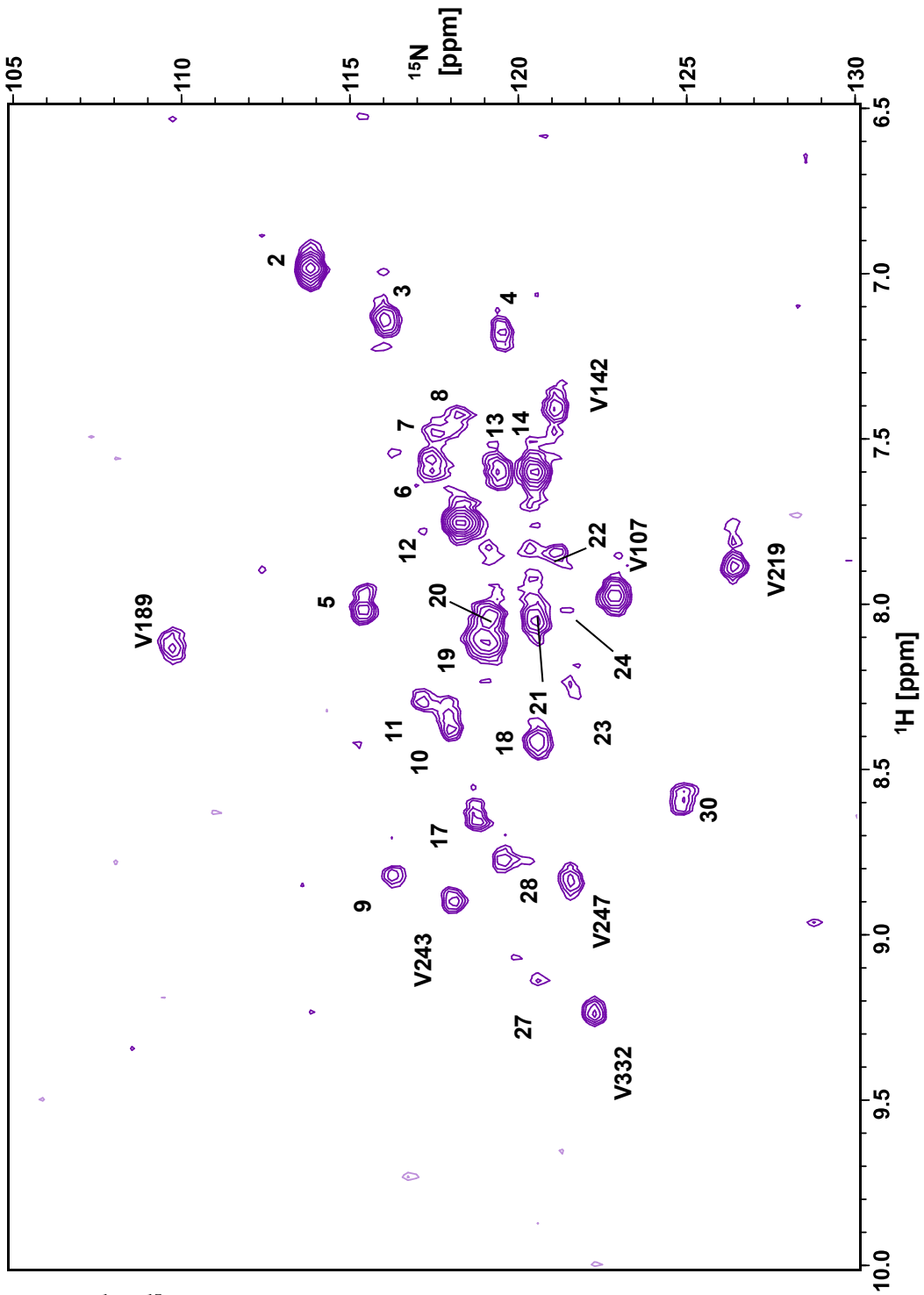


Figure A 8. [^{1}H, ^{15}N]-TROSY spectrum of hADRB1-TS-Twin-Strep with the antagonist alprenolol.

Experimental details described in legend of Figure 34.

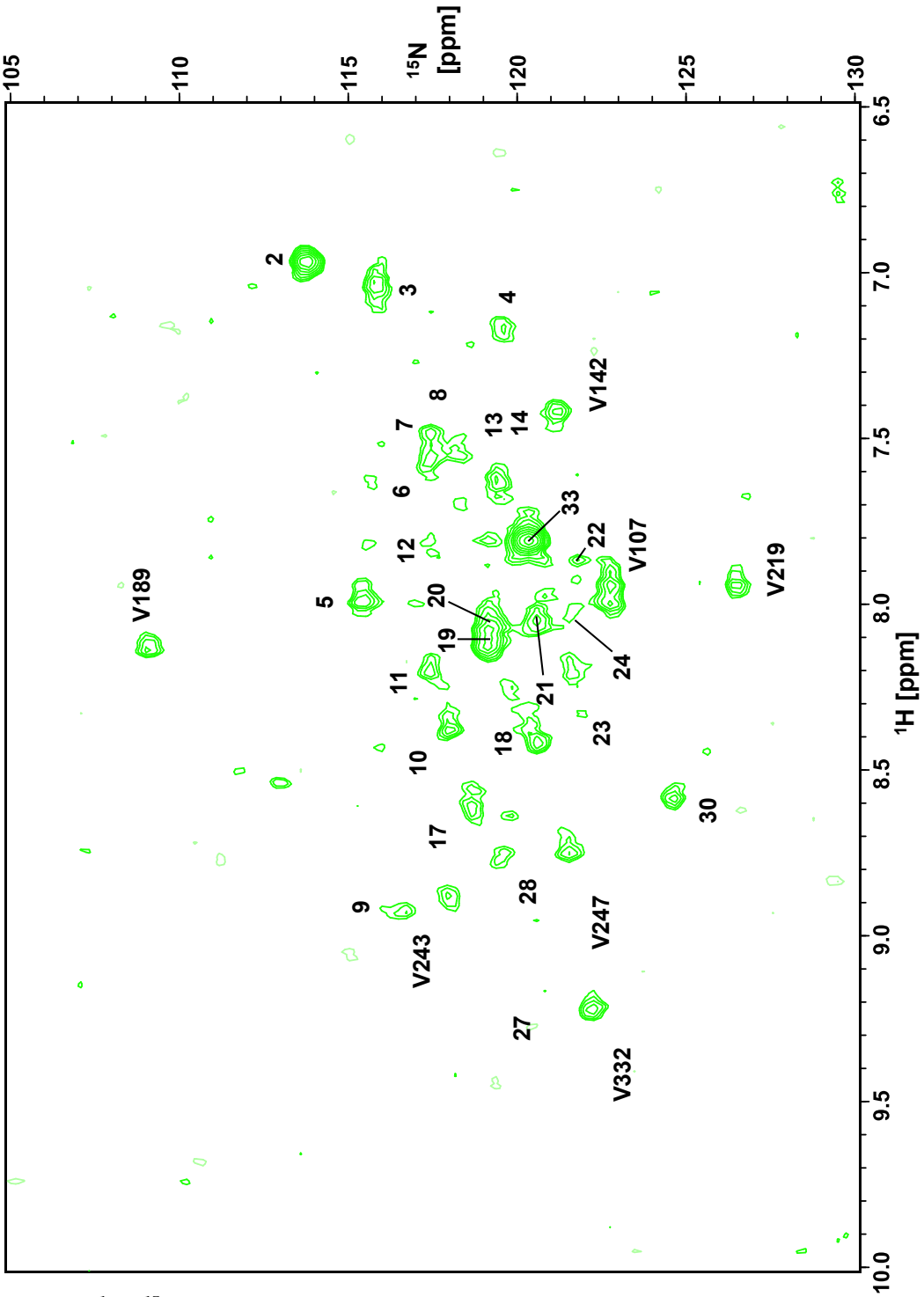


Figure A 9. [^{1}H, ^{15}N]-TROSY spectrum of hADRB1-TS-Twin-Strep with the antagonist S-Propranolol.

Experimental details described in legend of Figure 34 and Figure 35.

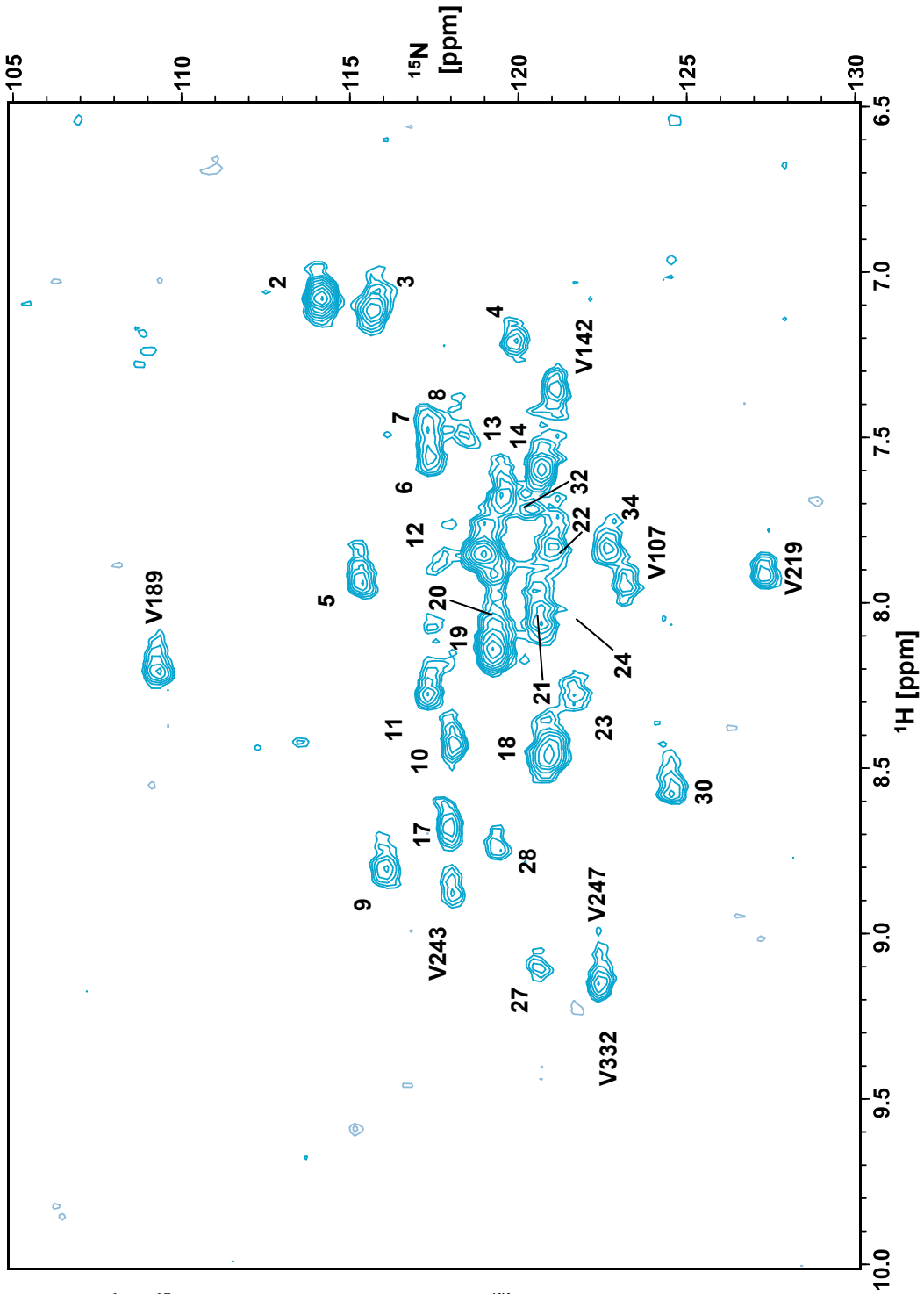


Figure A 10. [^{1}H, ^{15}N]-TROSY spectrum of gp67^{SS}hADRB1-TS-YY-Twin-Strep with the antagonist cyanopindolol.

Experimental details described in legend of Figure 36.

A6. Manuscripts in preparation

Parts of this thesis are being prepared for submission to journals for peer review in modified form. That includes the following manuscripts with the working titles:

- "Tuning the secretory pathway for GPCR over-expression in the Baculovirus/insect cell expression system."

- "Transferring the crystallization construct architecture from the turkey ADRB1 to the human ADRB1."

A7. Curriculum Vitae

CURRICULUM VITAE

JAKOB AASTRUP JØRGENSEN

Cand. Scient (M. Sc.), Biochemistry