

**Predicted signal peptides, and the role  
of the N-terminal tail, at the monoamine  
G-protein coupled receptors  
5-HT<sub>2C</sub> and  $\alpha_{2C}$**

**Jan Anker Jahnsen**



Dissertation for the degree philosophiae doctor (PhD)  
at the University of Bergen

2015

Dissertation date: 6th March 2015

© Copyright Jan Anker Jahnsen

The material in this publication is protected by copyright law.

Year: 2014

Title: Predicted signal peptides, and the role of the N-terminal tail, at the monoamine G-protein coupled receptors 5-HT<sub>2C</sub> and  $\alpha_{2C}$ .

Author: Jan Anker Jahnsen

Print: AIT OSLO AS / University of Bergen

## **Scientific environment**

The work presented here was carried out during 2006-2014 at the Section of Pharmacology at the Department of Clinical Science, University of Bergen, under the supervision of Professor Staffan Uhlén. Both Professor Uhlén and I have been members of the Centre of Pharmacy group, which consists of researchers and PhD-students from Faculty of Mathematics and Natural Sciences and Faculty of Medicine and Dentistry responsible for the education of Masters in the Pharmaceutical Sciences. The funding of the PhD fellowship was provided by The Faculty of Medicine and Dentistry at University of Bergen.

## Acknowledgements

First and foremost I wish to thank my supervisor Professor Staffan Uhlén, who with his enthusiasm for research has been an inspiration to continue working when a seemingly endless amount of laboratory work still needed to be done. Your diligence and effort with regard to teaching the pharmacy students pharmacology has been awe-inspiring. As our research group has only consisted of the two of us, we have worked closely together throughout this period which now is coming to an end. In addition to all I have learnt about research and pharmacology from you, I will miss the late night conversations in the lab about anything from the latest science breakthroughs to the travails of the local underachieving football club. When I started as your PhD student you had just recently moved to Bergen, and it has been both exciting and hard work to build our laboratory to the point where it is now. Now that you are moving back to your family in Uppsala, Sweden, I want to wish you all the best for the future, and hope you will keep in touch with the Centre of Pharmacy in the future.

Next, I wish to thank Dr. Med. Jan Schjøtt leader of RELIS Vest and Adjunct Professor of Section of Pharmacology at the Department of Clinical Science. When my PhD fellowship stipend ended after four years with two papers in the pipeline but none published, you offering me a position as pharmacist at RELIS Vest were vital to the completion of the PhD project. It provided me with the opportunity to stay in close proximity to my supervisor at the Section of Pharmacology, as well as a very interesting and rewarding day job. Your advice and encouragement throughout these years have been very important.

I would like to thank both former and present staff, researchers and PhD students at the Section of Pharmacology and the Bevital foundation. From earlier days in Armauer Hansen Hus to the present in Laboratoriebygget you have all been inclusive, encouraging and always helpful to me.

I would also like to thank my current and former colleagues at RELIS Vest and Section of Clinical Pharmacology at Haukeland University Hospital. You are wonderful people that make going to work exciting and enjoyable, and your encouragement the last few years has been greatly appreciated. This great work environment has much to do with our leaders Dr. Med. Jan Schjøtt and Dr. Med. Bettina Riedel.

I wish to thank current and former PhD students and staff associated with Centre of Pharmacy for contributing to a good environment for research and teaching. The connection with PhD students having the same background and facing the same challenges has been important and helpful.

Last, but not least, I wish to thank my close family: Mom, Dad, Bjarte, Nina, Stine and Andreas as well as all my friends for unconditional support and enduring faith.

Bergen, October 2014.

Jan Anker Jahnsen

*Some people think that science is just all this technology around, but NO it's something much deeper than that. Science, scientific thinking, scientific method is for me the only philosophical construct that the human race has developed to determine what is reliably true.*

*- Sir Harry Kroto, Nobel Laureate in Chemistry, Ask a Nobel Laureate, YouTube, 23/9-2010.*

---

# Contents

<b>SCIENTIFIC ENVIRONMENT .....</b>	<b>3</b>
<b>ACKNOWLEDGEMENTS .....</b>	<b>4</b>
<b>CONTENTS .....</b>	<b>6</b>
<b>ABBREVIATIONS .....</b>	<b>9</b>
<b>ABSTRACT .....</b>	<b>11</b>
<b>LIST OF PUBLICATIONS .....</b>	<b>13</b>
<b>1. INTRODUCTION.....</b>	<b>14</b>
1.1 G-PROTEIN COUPLED RECEPTORS .....	14
1.1.1 <i>G-proteins</i> .....	16
1.1.2 <i>The Rhodopsin family of GPCRs (Class A)</i> .....	18
1.1.3 <i>The serotonin 5-HT<sub>2C</sub> receptor</i> .....	20
1.1.4 <i>The <math>\alpha_{2C}</math>-adrenoceptor</i> .....	23
1.2 MEMBRANE LOCALIZATION OF NEWLY SYNTHESIZED GPCRS: FROM GENETIC CODE TO MATURE PROTEIN .....	27
1.2.1 <i>The protein secretory pathway</i> .....	27
1.2.2 <i>Membrane localization directed by signal peptide or reverse signal anchor</i> ..	28
1.2.3 <i>The basic features of signal peptides</i> .....	31
1.2.4 <i>Protein prediction software: Signal P and TMHMM</i> .....	32
<b>2. AIMS OF THE PRESENT STUDY .....</b>	<b>38</b>
<b>3. SUMMARY OF RESULTS (PAPERS I, II AND III) .....</b>	<b>40</b>
3.1 PAPER I.....	40
3.2 PAPER II.....	42
3.3 PAPER III.....	44
<b>4. GENERAL DISCUSSION .....</b>	<b>47</b>

---

4.1	METHODOLOGICAL CONSIDERATIONS .....	47
4.1.1	<i>Investigating N-termini for signal peptide function with the FLAG epitope</i> .....	48
4.1.2	<i>The N-terminus relevance for expression of cell surface receptors</i> .....	50
4.1.3	<i>The 5-HT<sub>2C</sub> receptor N-terminus' influence on expression</i> .....	51
4.1.4	<i>The <math>\alpha_{2C}</math>-adrenoceptor N-terminus' influence on expression</i> .....	51
4.1.5	<i>Which region of the <math>\alpha_{2C}</math>-adrenoceptor determines selectivity for "bulky" drugs</i> .....	52
4.2	DISCUSSION OF 5-HT <sub>2C</sub> RESULTS .....	53
4.2.1	<i>The 5-HT<sub>2C</sub> receptor has a cleavable signal peptide</i> .....	53
4.2.2	<i>Incorporating epitopes in the N-terminus influence receptor expression</i> .....	54
4.3	DISCUSSION OF ALPHA <sub>2C</sub> -ADRENOCEPTOR RESULTS .....	55
4.3.1	<i>The <math>\alpha_{2C}</math> N-terminus does not contain a 22 amino acid long cleavable signal peptide</i> .....	55
4.3.2	<i>The N-terminus is necessary for efficient expression of mature <math>\alpha_{2C}</math>-adrenoceptors</i> .....	56
4.3.3	<i>The C-terminal half of the <math>\alpha_{2C}</math>-adrenoceptor determines drug selectivity</i> .....	59
5.	<b>CONCLUDING REMARKS</b> .....	60
6.	<b>FUTURE PERSPECTIVES</b> .....	62
6.1	THE 5-HT <sub>2C</sub> N-TERMINUS .....	62
6.1.1	<i>Characterization of the binding epitope for the monoclonal antibody sc-15081 at the 5-HT<sub>2c</sub> receptor</i> .....	62
6.1.2	<i>Investigating the 5-HT<sub>2C</sub> N-terminus for alternative signal peptide cleavage site</i> .....	62
6.2	THE ALPHA <sub>2C</sub> -ADRENOCEPTOR N-TERMINUS .....	63

---

6.2.1	<i>Investigation of alternative signal peptide cleavage site predicted for the</i>	
	<i><math>\alpha_{2C}</math>-adrenoceptor</i> .....	63
7.	<b>REFERENCES</b> .....	<b>65</b>
8.	<b>APPENDIX</b> .....	<b>70</b>



---

## Abbreviations

A complete list of amino acid abbreviations can be seen in appendix 1.

5-HT	5-Hydroxytryptamine (serotonin)
7TM	Seven transmembrane (receptor)
A	Adrenaline (British Approved Name), epinephrine (International Nonproprietary Name)
Cys23Ser	A single nucleotide polymorphism converting cysteine in position 23 to serine
cAMP	Cyclic adenosine monophosphate
CBS	Center for Biological Sequence Analysis
CNS	Central nervous system
COS-7	<u>C</u> V-1 <u>o</u> ri <u>g</u> in with <u>S</u> V40, cell line derived from African green monkey kidney
DAG	Diacyl glycerol
ECL	Extracellular loop
EMA	European Medicines Agency
ER	Endoplasmic reticulum
ERGIC	ER-Golgi intermediate compartment
FDA	Food and Drug Administration (USA)
FSH	Follicle-stimulating hormone
GDP	Guanosine diphosphate
GPCR	G-protein coupled receptor
GTP	Guanosine-5'-triphosphate
HAsp	Influenza A hemagglutinin signal peptide
HEK293	Human embryonic kidney 293 (cell line)
HMM	Hidden Markov model (a statistical probability model)
IP <sub>3</sub>	Inositol trisphosphate
IUPHAR	International Union of Basic and Clinical Pharmacology

$K_i$	Equilibrium dissociation constant of a ligand determined in inhibition studies
LH	Luteinizing hormone
NA	Noradrenaline (British Approved Name), norepinephrine (International Nonproprietary Name)
PCR	Polymerase chain reaction
PIP <sub>2</sub>	Phosphatidylinositol 4,5-bisphosphate
PKC	Protein kinase C
$pK_i$	The negative logarithm to base 10 of the equilibrium dissociation constant $K_i$
PLC	Phospholipase C
PNS	Peripheral nervous system
SNP	Single nucleotide polymorphism
SNRI	Serotonin and norepinephrine reuptake inhibitor
SRP	Signal recognition particle
SSRI	Selective serotonin re-uptake inhibitor
TGN	Trans-Golgi network
TM	Transmembrane (helix)
TMHMM	Prediction of transmembrane helices in proteins (software)
TSH	Thyroid-stimulating hormone

## Abstract

**Background:** G-protein coupled receptors (GPCRs) have seven transmembrane helices and are situated in the cell membrane, where they transduce signals from specific ligands to the interior of the cell. The first step in the path toward a functional GPCR is the synthesis and incorporation of the evolving receptor into the endoplasmic reticulum (ER) membrane. This process is named cotranslational translocation and is directed by a hydrophobic signal sequence located either in the N-terminus or in the first transmembrane segment (TM1). When the signal sequence is located in the N-terminus, it is cleaved off after translocation and is called a signal peptide (SP). When the signal sequence is part of the TM1 it is called a signal anchor. Monoamine GPCRs have in general short N-termini and are expected to use their TM1 as a signal anchor. Two monoamine GPCRs are nevertheless predicted by a SP prediction software to have signal peptides: The 5-HT<sub>2C</sub> receptor and the  $\alpha_{2C}$ -adrenoceptor. For the 5-HT<sub>2C</sub> receptor the consequence of having the predicted SP is that a single nucleotide polymorphism (SNP) will not be present in the mature receptor in the cell membrane. This SNP (Cys23Ser) has in several studies been associated with numerous clinical conditions and outcomes of pharmacotherapy. The  $\alpha_{2C}$ -adrenoceptor is poorly expressed at the cell surface and has a large intracellular pool of receptors. It has been shown previously for other receptors that by adding a cleavable signal peptide sequence immediately upstream to the endogenous N-terminus, the expression levels of  $\beta_2$ - and  $\alpha_{1D}$ -adrenoceptor are greatly enhanced. Consequently, it is seemingly odd that the poorly expressed  $\alpha_{2C}$ -adrenoceptor is predicted to contain a SP.

**Objective:** The primary aim was to determine whether the monoamine GPCRs 5-HT<sub>2C</sub> and  $\alpha_{2C}$  have cleavable signal peptides as predicted. A secondary aim was to determine what relevance the N-termini of the 5-HT<sub>2C</sub> and  $\alpha_{2C}$  receptors have for expression levels of the receptors.

**Materials and methods:** Methods included engineering receptor constructs and chimeras by PCR and transiently transfecting COS-7 and HEK293 cells. Receptor constructs containing FLAG epitope were investigated with the primary antibodies

M1 and M2 in epifluorescence and confocal microscopy. Expression levels of wild type and rebuilt receptor constructs were determined by radioligand binding performed on membrane preparations. For  $\alpha_{2C}$ -adrenoceptors radioligand binding was also performed on whole cells, matching the membranes, to exclude binding to an intracellular pool of receptors.

**Results and Conclusions:** The 5-HT<sub>2C</sub> receptor has a 32 amino acid long cleavable signal peptide, as predicted by its amino acid sequence. When the signal peptide is made non-cleavable by changing one amino acid, the expression level of the receptor is reduced by 70%. We therefore conclude that a 32 amino acid long cleavable signal peptide is participating in the integration of the 5-HT<sub>2C</sub> receptor into the ER membrane. Consequently, the mature receptor does not contain the aforementioned Cys23Ser SNP. The  $\alpha_{2C}$ -adrenoceptor does not possess a 22 amino acid long cleavable signal peptide. Among the  $\alpha_{2C}$ -adrenoceptor constructs, expression was highest for the wild type receptor where the endogenous N-terminus was retained. Furthermore, all attempts at increasing the expression level of the  $\alpha_{2C}$ -adrenoceptor by adding a known SP or by truncating the N-tail, failed. We conclude that the N-terminus is not a major contributor to the low expression level of the  $\alpha_{2C}$ -adrenoceptor.

---

## List of publications

I: **Jahnsen JA**, Uhlén S. The N-terminal region of the human 5-HT<sub>2C</sub> receptor has as a cleavable signal peptide. *Eur J Pharmacol.* 2012 Jun 5;684(1-3):44-50. doi: 10.1016/j.ejphar.2012.03.043.

II: **Jahnsen JA**, Uhlén S. The predicted N-terminal signal sequence of the human  $\alpha_{2C}$ -adrenoceptor does not act as a functional cleavable signal peptide. *Eur J Pharmacol.* 2012 Jun 5;684(1-3):51-8. doi: 10.1016/j.ejphar.2012.03.044.

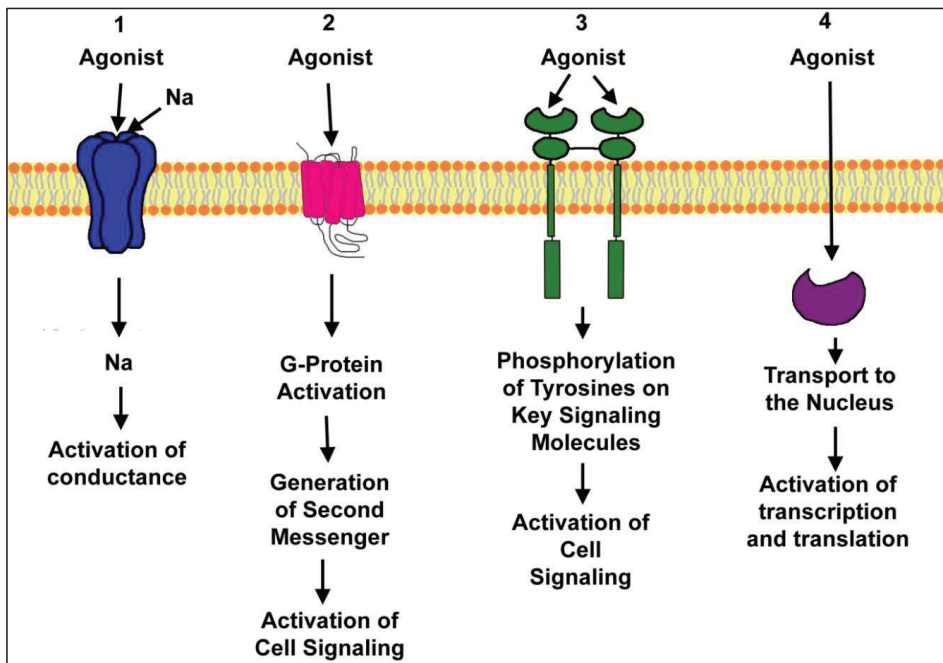
III: **Jahnsen JA**, Uhlén S. The C-terminal half of the  $\alpha_{2C}$ -adrenoceptor determines the receptor's membrane expression level and drug selectivity. *Naunyn Schmiedebergs Arch Pharmacol.* 2013 Dec;386(12):1031-40. doi: 10.1007/s00210-013-0902-z.

*Reprints were made with permission from Elsevier (paper I and II) and Springer (paper III).*

# 1. INTRODUCTION

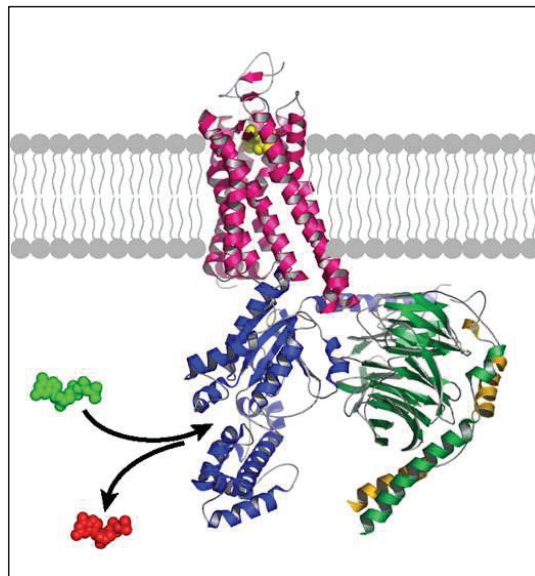
## 1.1 G-PROTEIN COUPLED RECEPTORS

A primary component of cellular life is the ability to communicate. For non-adjacent cell-to-cell communication this always means sending and receiving messages as chemical signals. As the majority of chemical signaling involves hydrophilic ligands unable to penetrate the lipid bilayer of the target cell, there is a need for cell-surface proteins to transduce the signal to the interior of the cell. The proteins that sense the chemical signals are termed receptors and can be divided into the four superfamilies, shown in figure 1: Ligand gated ion channels, G-protein coupled receptors, Kinase linked receptors and Nuclear receptors. All these receptors except the nuclear receptors are localized to the cell surface membrane.



**Figure 1. Overview of the four receptor superfamilies.** 1: Ion channels, 2: G-protein coupled receptors, 3: Kinase linked receptors, 4: Nuclear receptors. The figure illustrates how the receptors are situated in relation to the cell membrane, and how stimulation with an agonist leads to altered cell function. Taken from [1].

The largest family of cell-surface receptors is the G-protein-coupled receptors (GPCRs). The GPCRs play an essential role in modulating various physiological functions, and at least a third of all currently prescribed pharmaceutical drugs are acting on GPCRs [2, 3], with annual revenues in excess of \$40 billion [3], illustrating their importance as drug targets. The common structural features of GPCRs are seven hydrophobic transmembrane helices (TM1-TM7) spanning the cell membrane. The N-terminus is localized extracellular and the C-terminus is localized intracellular. Ligands interact with GPCRs at binding sites between the transmembrane helices and/or at the extracellular surfaces. Intracellular parts of the receptor interact with G-proteins, which is short for heterotrimeric GTP-binding proteins. The general structure of a GPCR interacting with a G-protein can be seen in figure 2.



**Figure 2. A three-dimensional structure of a G-protein coupled receptor embedded in a cell membrane, including the intracellular G-protein subunits.** A GPCR (magenta) in the plasma membrane with a ligand (yellow), and intracellularly the three G-protein subunits (gold, dark green and blue), GTP (bright green) and GDP (red). Taken from [4].

There are 1352 genes coding for physiological receptors in the genome, and of these 901 genes are coding for GPCRs with seven transmembrane helices [5]. The uncertainty whether G-proteins are involved in signal transduction for all GPCRs supports the notion that a more correct term would be 7TM receptors, and the terms GPCRs and 7TM receptors are used interchangeably in the literature. The G-protein-coupled receptors in the human genome form five main families [6]. Since only three of these families/classes are targets for clinically used drugs, the International Union of Basic and Clinical Pharmacology's Committee on Receptor Nomenclature and Drug Classification (NC-IUPHAR), divides the druggable GPCRs into the three classes: A (Rhodopsin), B (Secretin) and C (Glutamate). Excluding the large group of approximately 400 olfactory odorant receptors which are members of the Rhodopsin family, IUPHAR and the British Pharmacological Society (BPS) in its database updated in June 2014 officially recognizes 394 7TM receptors of which 388 are designated GPCRs. 129 of the 388 GPCRs are orphan receptors with no endogenous ligand conclusively determined [7]. A condensed version of the IUPHAR database current to late 2013 is published in the "The concise guide to pharmacology 2013/14: G protein-coupled receptors". As of late 2013 the rhodopsin family (Class A) contained 197 receptors with known ligands and 89 orphans, Class B had 15 receptors with known ligands and no orphans and Class C had 12 receptors with known ligands and 8 orphans [8]. The IUPHAR classification system divides the receptors in families and subtypes based on pharmacology, signal transduction and structure [7].

### **1.1.1 G-proteins**

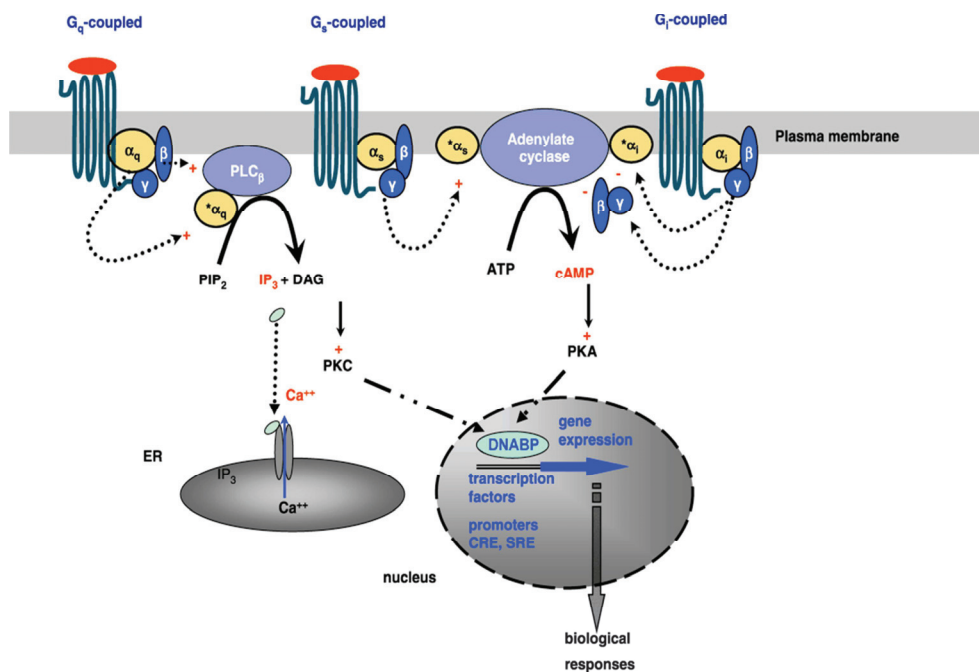
When a ligand binds to a GPCR, the receptor undergoes a conformational change that affects its intracellular coupling to a heterotrimeric GTP-binding protein, i.e. a G-protein. The G-protein is attached to the cell membrane facing the cytosol and consists of the three subunits  $\alpha$ ,  $\beta$  and  $\gamma$ , which make up the two functional entities  $\alpha$  and  $\beta\gamma$ . The conformational change of the receptor induces the  $\alpha$ -subunit to exchange



---

GDP for GTP, with the receptor acting as a direct guanine nucleotide exchange factor. In the classical model this leads to dissociation of the  $\alpha$ -GTP complex and  $\beta\gamma$  complex. The  $\alpha$ -GTP complex will then activate an intracellular second messenger system which will relay the message as intended. When the  $\alpha$ -GTP complex is binding to the intracellular enzymes that produce the second messengers, the GTP is prone for GTPase activity. The hydrolysis of GTP to GDP returns the  $\alpha$ -subunit to the inactive state and reformation of the  $\alpha\beta\gamma$ -complex [9]. The G-proteins are defined by the properties of the  $\alpha$ -subunit, since principal transduction historically is defined by the established  $G\alpha$ -signaling [10].

There are several isoforms of  $\alpha$  and  $\beta\gamma$  subunits, and the 16 mammalian  $\alpha$ -subunits can be divided into four families:  $G\alpha_s$ ,  $G\alpha_{i/0}$ ,  $G\alpha_{q/11}$  and  $G\alpha_{12/13}$ . If the reader wish to see an overview of the different G-proteins, IUPHAR/BPS have made an introduction to GPCRs on the [www.GuideToPharmacology.org](http://www.GuideToPharmacology.org) portal [10]. The two members of the stimulatory  $G\alpha_s$  proteins couples to adenylate cyclase, causing an increase in intracellular cAMP levels. The eight members of the  $G\alpha_{i/0}$  family are often inhibitory in character, for example the three  $G_i$ -proteins harboring  $\alpha_{i-1}$ ,  $\alpha_{i-2}$  and  $\alpha_{i-3}$  inhibit adenylate cyclase and cause a decrease in cAMP levels, thereby kind of having the opposite effect compared to  $G\alpha_s$ -coupling. The four members of the  $G\alpha_{q/11}$  family activates phospholipase  $C_\beta$  ( $PLC_\beta$ ), which in turn results in intramembrane hydrolysis of phosphatidylinositol-4,5-bisphosphate ( $PIP_2$ ) to inositol-1,4,5-triphosphate ( $IP_3$ ) and diacylglycerol (DAG).  $IP_3$  triggers the release of  $Ca^{2+}$  from intracellular stores by binding to the  $IP_3$  receptor, which is a  $Ca^{2+}$  channel, on the endoplasmic reticulum. DAG increases the activity of the enzyme protein kinase C (PKC), which controls the function of other proteins by phosphorylation. The two members of the  $G\alpha_{12/13}$  family regulate the Ras subfamily of Rho proteins. The Rho proteins are small monomeric GTP-binding proteins involved in regulating cell morphology like organelle development, cytoskeletal dynamics and cell movement [9]. The  $G\alpha_s$ ,  $G\alpha_{i/0}$  and  $G\alpha_{q/11}$  families activate the classic GPCR signaling pathways within pharmacology. Some typical second messenger cell responses following agonist binding and receptor activation can be seen in figure 3.



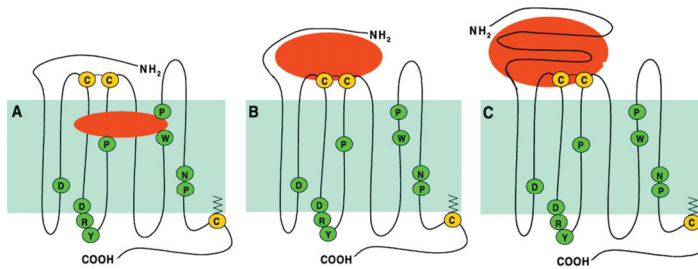
**Figure 3. An overview of classical G-protein signaling and second messenger system for  $G_{\alpha q}$ ,  $G_{\alpha s}$  and  $G_{\alpha i}$ .** Taken from [9].

### 1.1.2 The Rhodopsin family of GPCRs (Class A)

The Rhodopsin family, i.e. Class A, is by far the largest family of GPCRs, comprising more than 90% of the GPCRs, and lends its name from the rhodopsin receptor. Rhodopsin, which is localized to the retina in the eye, is a model system for membrane proteins and has been investigated regarding the molecular mechanism for vision since the early 1980s. In 2000 rhodopsin was the first GPCR to have its crystal structure determined, when Palczewski et al. published the bovine rhodopsin structure [11]. Rhodopsin is activated by photons, which transiently isomerize a molecule of retinal bound within the receptor. Olfactory GPCRs are chemosensory receptors that are activated by exogenous ligands, such as odors. These receptors are very poorly expressed at the cell surface when investigated in cell cultures, which makes characterization of them in traditional pharmacological assays either very difficult or almost impossible [12]. Consequently not much is known about their ligands, and

---

most of the olfactory receptors are therefore orphan receptors with no known ligands. The GPCRs in the body that have endogenous ligands are termed endo-GPCRs [6], and it is among these receptors we find the drug targets so fundamental to human pharmacology and drug therapy [7]. These GPCRs are defined by their endogenous ligands, represented by proteins, peptides, lipids, amino acids, ions, and monoamines derived from amino acids. The Rhodopsin family can be further subdivided into three groups usually termed “1A”, “1B” and “1C” [13]. In group 1A the receptors have small endogenous ligands that bind in a cavity between the transmembrane regions of the receptor. This includes rhodopsin itself as well as the GPCRs activated by monoamines like adrenaline and serotonin, usually designated monoamine receptors. One such monoamine receptor, the  $\beta_2$ -adrenoceptor, is the focus of a sustained effort to elucidate the specific structure, binding and G-protein interaction of druggable GPCRs. In 2007 the crystal structure of the  $\beta_2$ -adrenoceptor in the presence of a partial inverse agonist was published, obtained through two different methodological approaches; the receptor fused to T4 lysozyme [3] or stabilized with the monoclonal antibody Mab5 [14]. Subsequently, in 2013, the adrenaline-activated structure of the  $\beta_2$ -adrenoceptor was published [15], and a general method of using covalent agonists derived from the endogenous monoamine ligands for crystallization of agonist-activated GPCRs was established in 2014 [16]. These recent achievements are likely to accelerate our knowledge of the structure, binding of ligands and function of the individual GPCRs. Although crystal structures of GPCRs are very important for characterizing the binding pocket and conformational changes, the method is not suited to characterize the extracellular N-terminal tail [14]. In group 1B the ligands are peptides and the binding site is extending to the extracellular loops. The 1C group contains GPCRs whose ligands are glycoprotein hormones like thyroid stimulating hormone (TSH), luteinizing hormone (LH) and follicle-stimulating hormone (FSH). These receptors have in common a large extracellular N-terminal tail that is involved in the binding of the hormone. The characteristics of these three groups of Rhodopsin family receptors are shown in figure 4 [9].



**Figure 4. General structure and binding sites for rhodopsin-like family A GPCRs.**

Binding site is indicated in orange. A: Endogenous small-molecule ligands e.g. monoamines. B: Peptide ligands. C: Glycoprotein ligands. Taken from [9].

### 1.1.3 The serotonin 5-HT<sub>2C</sub> receptor

5-hydroxytryptamine (5-HT) is the chemical name for the neurotransmitter and local hormone serotonin. Serotonin is endogenously produced by a two-step conversion of the amino acid tryptophan to serotonin. Its degradation and excretion is mainly through oxidative deamination catalyzed by monoamine oxidase. The presence of the end-product 5-hydroxyindoleacetic acid in urine is an indicator for endogenous 5-HT production. Most of the serotonin is located in the enterochromaffin cells in the wall of the intestine, but it is also present in high concentrations in blood platelets and in the CNS. Serotonin, acting via membrane receptors, is a transmitter substance in the central nervous system (CNS), in the peripheral, including enteric, nervous system (PNS) and in the vasculature [17]. In humans serotonin is identified as the endogenous ligand for at least 12 different GPCR subtypes, as well as for a ligand gated ion channel classified as 5-HT<sub>3</sub>. The human serotonin GPCRs include five 5-HT<sub>1</sub> receptors, three 5-HT<sub>2</sub> receptors and in addition the 5-HT<sub>4,7</sub> receptors [8, 17]

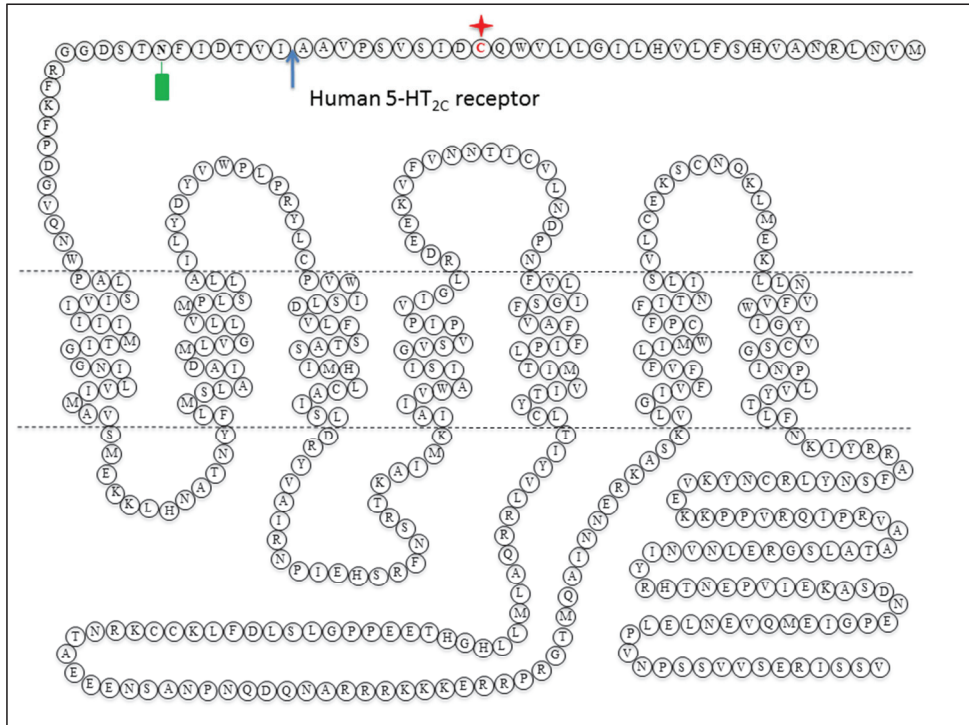
The 5-HT<sub>2</sub> family consists of the 5-HT<sub>2A</sub>, 5-HT<sub>2B</sub> and 5-HT<sub>2C</sub> receptors, which mediate their effects mainly through coupling to the G<sub>q/11</sub> proteins [8]. The three receptors have a close sequence homology and it is a challenge to develop functionally selective agonists [18]. This is of great importance as activation of 5-

---

HT<sub>2A</sub> is associated with hallucinogenic activity and 5-HT<sub>2B</sub> with cardiac valvulopathy [19]. 5-HT<sub>2C</sub> was originally known as 5-HT<sub>1C</sub>, but was later renamed 5-HT<sub>2C</sub>. The receptor was initially designated the “choroid plexus” receptor, but after more sensitive methods were developed, like in situ hybridization, it became clear that the distribution of the 5-HT<sub>2C</sub> receptor was wide-spread in the brain [17]. Because the 5-HT<sub>2C</sub> receptor undergoes extensive mRNA-editing in addition to being spliced, it is possible that 32 different mRNAs and 24 different proteins could exist [17]. When the unedited receptor was compared to two extensively edited forms in a stable cell line, it appeared that the unedited form coupled to both G<sub>q/11</sub>- and G<sub>13</sub>-proteins, while the edited forms only coupled to G<sub>q/11</sub>-proteins [20]. mRNA editing breaks with the theory that “one gene encodes one protein”, and opens up the possibility for major differences between in vivo and in vitro pharmacology for important drug targets [21].

*The Cys23Ser single nucleotide polymorphism in 5-HT<sub>2C</sub>.* There is a wide-spread single nucleotide polymorphism (SNP) termed Cys23Ser (rs6318) for the 5-HT<sub>2C</sub> receptor, with 13% of the Caucasian population having a serine instead of a cysteine at position 23. This polymorphism is situated in the about 57 amino acids long extracellular N-terminus, and has in association studies been predicted to be linked to numerous psychiatric disorders and to varying outcomes from various drug treatments [22, 23]. The N-terminus also contains a glycosylation site at asparagine N39. Also the rat 5-HT<sub>2C</sub> receptor has the N39 glycosylation site, and it has been shown that this receptor is glycosylated [24]. The pharmacological properties of the SNP were extensively investigated by Fentress and colleagues in 2005 [25]. They failed to find any pharmacological consequences for this mutation, and concluded therefore that positive associations between the polymorphism and disease states may be due to some other factor. In addition, we have shown [26] that the SNP probably is not part of the mature receptor since it is cleaved off as part of the signal peptide of the receptor (see figure 5). Nevertheless genetic association studies continue, still in 2014, to link this SNP with diverse conditions as premature ejaculation [27] and

cocaine cue reactivity [28]. The amino acid sequence and topology of 5-HT<sub>2C</sub> is depicted in figure 5, with the cysteine at the SNP site denoted with a red letter and star above it, and the cleavage site of the signal peptide with an arrow.



**Figure 5. The amino acid sequence and topology of the human 5-HT<sub>2C</sub> receptor.** The alignments of transmembrane segments are as predicted by the CBS TMHMM Server 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>), when analyzed without the predicted signal peptide. The predicted cleavage site of the signal peptide is noted with blue arrow. The position of cysteine (C) at the single nucleotide polymorphism site Cys23Ser (rs6318) is noted with a red letter and a star above it. The N-linked glycosylation site N39 in the N-terminus is noted with a green rectangle underneath.

**5-HT<sub>2C</sub> as a drug target.** Lorcaserin is a 5-HT<sub>2C</sub> selective agonist, with the brand name Belviq, approved by the Food and Drug Administration (FDA) in USA as an anti-obesity drug to be used in tandem with diet and exercise. The pharmaceutical company withdrew its application to the European Medicines Agency in 2013

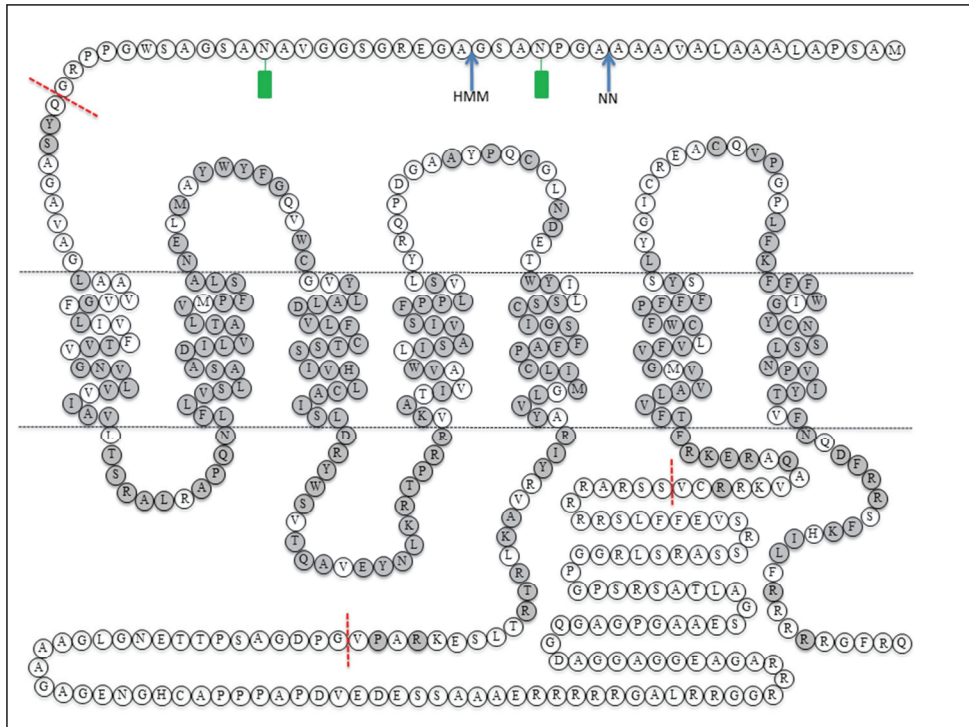
---

because of the agency's concerns over lack of effect and possible adverse effects. The adverse effects of concern were development of tumors in laboratory tests, the risk of valvulopathy and psychiatric disorders like depression [29]. From animal studies it is shown that mice lacking the 5-HT<sub>2C</sub> gene is severely obese and food intake regulation defective [30]. The receptor is attracting interest as a target for an agonist to prevent weight gain induced by antipsychotics [17], especially since antipsychotic drugs that provoke weight gain seem to be inverse agonists at the 5-HT<sub>2C</sub> receptor [31]. 5-HT<sub>2C</sub> agonists are in addition attracting interest as potential treatment for substance abuse [19].

There is also an interest in antagonists selective for 5-HT<sub>2C</sub>. Agomelatine is an anti-depressive drug with a unique mechanism of action, acting as agonist on melatonergic receptors MT1 and MT2 and as an antagonist on 5-HT<sub>2C</sub> [32]. Its efficacy and tolerability when compared to established antidepressants within the SSRI and SNRI groups, has been investigated in a Cochrane Review where the authors concluded the current data was too poor to draw a conclusion [33]. However, the slow pace in the development of new drugs acting at the 5-HT<sub>2C</sub> receptor highlights the need for further studies on the molecular function and physiological importance of this receptor.

#### 1.1.4 The $\alpha_{2C}$ -adrenoceptor

The catecholamines noradrenaline (NA) and adrenaline (A) are the endogenous ligands for the adrenoceptors, which are divided into the three groups  $\alpha_1$ ,  $\alpha_2$  and  $\beta$ , including in total nine subtype receptors:  $\alpha_{1A}$ ,  $\alpha_{1B}$ ,  $\alpha_{1D}$ ,  $\alpha_{2A}$ ,  $\alpha_{2B}$ ,  $\alpha_{2C}$ ,  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  [8]. The human  $\alpha_2$ -adrenoceptor subtypes  $\alpha_{2A}$ ,  $\alpha_{2B}$  and  $\alpha_{2C}$  show highly conserved amino acid sequences, especially in the seven transmembrane helices. The amino acid sequence and topology of the  $\alpha_{2C}$ -adrenoceptor are shown in figure 6, with those amino acids that are identical in the  $\alpha_{2A}$ -adrenoceptor shaded for comparison. The  $\alpha_{2C}$  N-tail sequence has not been aligned with that of  $\alpha_{2A}$  since they are too different.



**Figure 6. The  $\alpha_{2C}$ -adrenoceptor, in comparison with the  $\alpha_{2A}$ -adrenoceptor.** Those amino acids that are identical in the  $\alpha_{2A}$ -adrenoceptor are shaded. No alignments were performed before the dashed (red) line in the N-tail and between the dashed lines (red) in the third intracellular loop, due to the large differences in those regions. The predicted cleavage sites when the N-terminus is analyzed in SignalP 3.0 are noted with (blue) arrows and which prediction method used to obtain the result (HMM = Hidden Markov model, NN = neural networks). The N-linked glycosylation sites N19 and N33 in the N-terminus are noted with a green rectangle underneath. Modified from figure 1 in paper III [34].

Intriguingly the  $\alpha_{2C}$  N-tail is not only unusual compared to the  $\alpha_{2A}$  N-tail, but also when compared to other monoamine GPCRs. The  $\alpha_{2C}$  N-terminus is uncharacteristically long with about 49 amino acids, including a hydrophobic stretch of 11 amino acids. There are two N-linked glycosylation sites at asparagines N19 and N33 in the N-terminus, that both have been verified experimentally [35]. It has also been shown that the receptor expressed at the cell surface is glycosylated at both these sites [36, 37]. In the present study we have verified by our methods that the N19 glycosylation site is not cleaved as part of a signal peptide. Unlike  $\alpha_{2A}$  and  $\alpha_{2B}$ , the



---

$\alpha_{2C}$ -adrenoceptor is poorly expressed at the cell surface and has a prominent intracellular pool of receptors [38]. An exception to the poor expression of the receptor at the cell membrane is in neuroendocrine cell lines where the receptor is efficiently expressed, although still with a significant intracellular pool, suggesting cell line specific properties concerning trafficking from intracellular compartments to the plasma membrane [39]. There is some controversy as to the role of the N-terminus concerning the poor expression of the receptor [38]. One opinion on the matter is that the 11 amino acid hydrophobic stretch in the N-terminus function as an endoplasmic reticulum (ER) retention signal involved in inhibiting intracellular trafficking of the receptor, and thereby inhibiting the export of the  $\alpha_{2C}$ -adrenoceptor to the plasma membrane [35]. Another opinion advocates that there is a retention signal in the C-terminal tail of the  $\alpha_{2C}$ -adrenoceptor, and that this C-terminal retention signal may explain much of the low expression of the receptor at the cell surface [40]. All the  $\alpha_2$ -adrenoceptors couple to G-proteins of the  $G\alpha_i$ -family, and thus have similar signaling properties [7, 8]. Originally there was the perception that  $\alpha_2$ -adrenoceptors were only expressed at presynaptic sites, being responsible for inhibiting NA-release from adrenergic nerves. However, in addition to such presynaptic activity of  $\alpha_2$ -adrenoceptors [41], the current knowledge implies that the majority of  $\alpha_2$ -adrenoceptors are in fact localized post-synaptically [42]. The influence on physiological functions that the  $\alpha_2$ -adrenoceptors are most relevant for, involve  $\alpha_2$ -mediated effects in the central nervous system and the cardiovascular system [43]. In addition, the  $\alpha_2$ -agonist brimonidine is lowering intraocular pressure, and is in clinical use in glaucoma.

The  $\alpha_{2C}$ -adrenoceptor has primarily been detected in the central nervous system [44]. Genetically modified mice strains with varying degrees of receptor deletion, have revealed the role of the  $\alpha_{2C}$ -adrenoceptor as a presynaptic receptor in regulating noradrenaline release [45]. However, it is the  $\alpha_{2A}$ -subtype that is predominant in controlling the release of noradrenaline from sympathetic nerves, while the more evasive  $\alpha_{2C}$  controls release of adrenaline from the adrenal gland [41]. Knock-out mice lacking  $\alpha_{2C}$  receptors exhibit more than two-fold augmented circulating and

urine levels of adrenaline [45]. By analyzing the mice phenotypes it becomes clear that most of the recognized  $\alpha_2$ -mediated pharmacological effects are due to activation of the  $\alpha_{2A}$ -subtype. This includes the antinociceptive, sedative, hypotensive, hypothermic and behavioral effects of  $\alpha_2$ -agonists, as well as a possible positive effect on cognitive functions. The  $\alpha_{2C}$ -subtype has been shown to be involved in modulation of dopaminergic neurotransmission and different behavioral responses, as well as in induction of hypothermia [45].

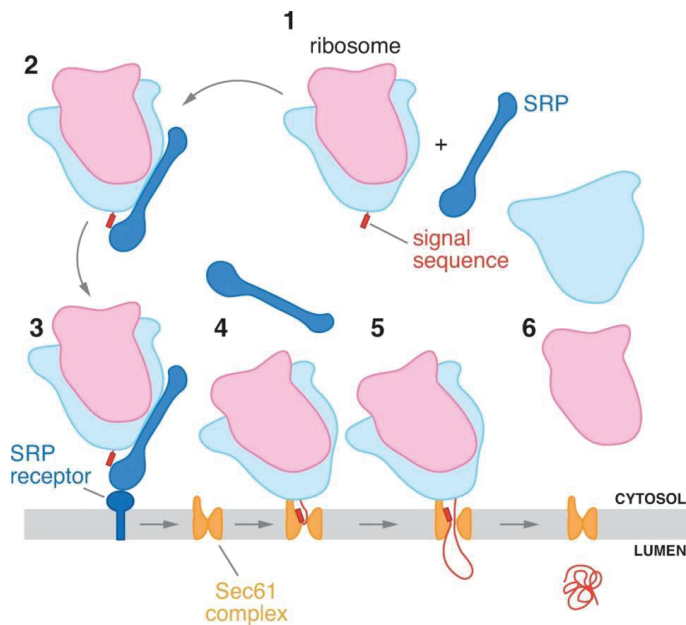
**The  $\alpha_{2C}$ -adrenoceptor as a drug target.** Several physiological effects mediated by  $\alpha_2$ -adrenoceptors seem to involve more than one subtype, with more than one receptor having complementary effects like  $\alpha_{2A}$  and  $\alpha_{2C}$  in controlling chatecolamine release presynaptically [46]. This opens the possibility that an  $\alpha_{2C}$ -selective agonist might be useful, resulting in antinociceptive effect without  $\alpha_{2A}$ -mediated hypotensive and sedative effects. Another strategy is combining an  $\alpha_{2C}$ -antagonist, thereby blocking the  $\alpha_{2C}$ -mediated presynaptic inhibition of noradrenaline and dopamine release, with a drug having monoamine reuptake blocking activity. This could speculatively be useful in treating depression, anxiety and Parkinson's disease [47]. Recently, an  $\alpha_{2C}$ -agonist has been investigated as nasal decongestant in animal models with the aim of developing a decongestant without cardiovascular actions [48].

There is a possibility that the  $\alpha_{2C}$ -adrenoceptor is involved in the pathology of Raynaud Phenomenon, and therefore could be a drug target for treatment of this affliction with an  $\alpha_{2C}$ -antagonist [49]. This disease is recognized by enhanced vasoconstriction induced by cold, emotional stress or exposure to vibrations. It seems like  $\alpha_{2C}$  specifically modulates vascular tone at lower temperature, while being inactive at 37°C. This could be explained by an increased expression or effectiveness of  $\alpha_{2C}$ -adrenoceptors at lower temperatures [40, 49].

## 1.2 MEMBRANE LOCALIZATION OF NEWLY SYNTHESIZED GPCRS: FROM GENETIC CODE TO MATURE PROTEIN

### 1.2.1 The protein secretory pathway

The journey for a GPCR from genetic code to a functional receptor in the cell membrane follows the secretory pathway for proteins. To gain access to the secretory pathway the protein has to have a specific targeting signal that directs the translationally evolving protein to the endoplasmic reticulum (ER) membrane. The cotranslational translocation of a secretory protein is depicted in figure 7.

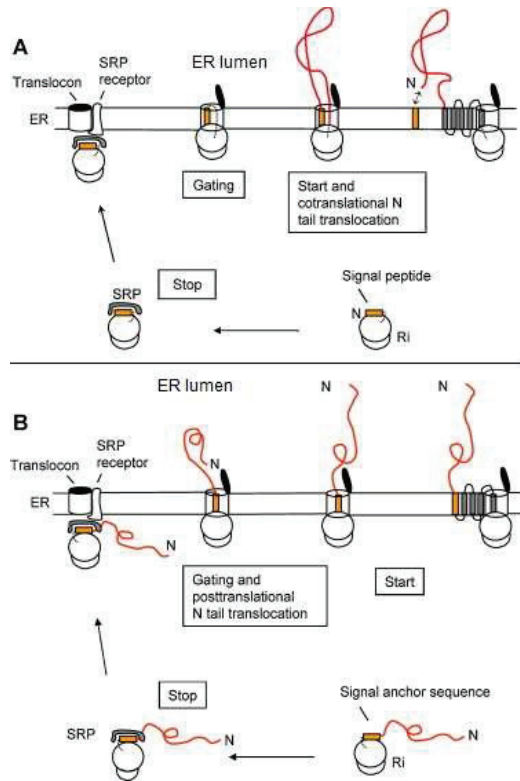


**Fig 7. An overview of how the newly synthesized signal sequence of a secretory protein (red color) interacts with the signal recognition particle (blue) and the Sec61 complex (orange) at the endoplasmic reticulum membrane.** While secretory proteins are completely translocated into the lumen of the ER, membrane-embedded proteins like GPCRs are incorporated in to the ER membrane. Figure taken from [50].

Initially the signal sequence is recognized by the signal recognition particle (SRP), and the complex of ribosome-nascent chain-SRP interacts with the SRP receptor situated at the ER membrane. The SRP dissociates, and a simultaneous protein translation and translocation across the ER membrane ensues, controlled by the Sec61 complex [50]. The Sec61 complex is the core unit of the translocation channel [51], which often is termed the translocon. If the protein is a membrane-bound receptor, it stays embedded in the plasma membrane by means of one or several hydrophobic transmembrane helices. If the protein is secreted, it passes into the ER lumen as depicted in figure 7 and is released extracellularly.

### **1.2.2 Membrane localization directed by signal peptide or reverse signal anchor**

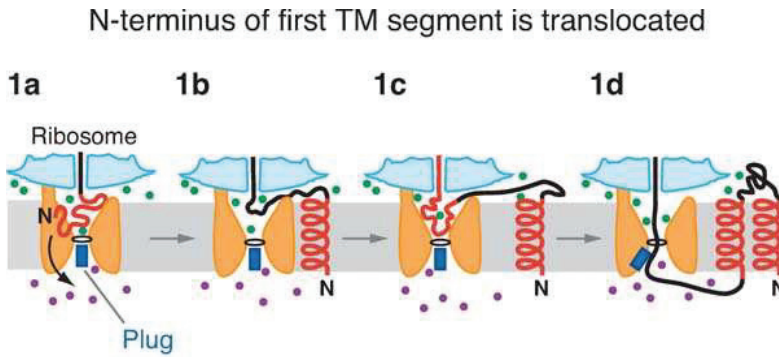
Secretory proteins all have signal peptides and are completely translocated into the lumen of the ER. On the other hand, GPCRs are integrated into the ER membrane with the appearance of a familiar seven transmembrane topology. The signal sequence within the GPCR family can either be a cleavable signal peptide situated in the N-terminus or a reverse signal anchor consisting of the first transmembrane helix as depicted in figure 8. When the targeting function of the signal peptide has been carried out, it is cleaved off by a class of enzymes known as signal peptidases [52]. So, the signal peptide is not present in the mature protein. It is estimated that 90-95% of the GPCRs do not have a signal peptide [51], but use their first transmembrane segment as a reverse signal anchor. As a general rule these GPCRs have short N-terminal tails [53]. Since a general feature of the monoamine GPCRs are short N-termini [9], they are expected to use their first transmembrane segment as a signal anchor.



**Figure 8. Signal peptide function (A) and reverse signal anchor function (B) of a GPCR.** The translocon is the channel regulating transport of proteins across and into the membrane, and consists of large protein complexes with Sec61 as the core unit. The red line with the “N” denotes the N-terminus of the receptor. SRP: Signal recognition particle. Ri: Ribosome. ER: Endoplasmic reticulum. Taken from [51].

The translation and translocation process of the GPCRs, termed cotranslational translocation, is essentially the same whether the receptor possesses a signal peptide or use the first transmembrane segment as a reverse signal anchor. However, a signal peptide has its N-terminus embedded in the membrane, like a transmembrane segment, with its C-terminus in the ER lumen (see panel A in figure 8). After enzymatic cleavage of the signal peptide the mature receptor’s N-terminus is in the ER lumen. For a receptor having a first transmembrane signal anchor, the signal anchor is oriented towards the lumen with the N-terminus in the ER lumen. Thus, the main difference being a posttranslational translocation of the already synthesized N-

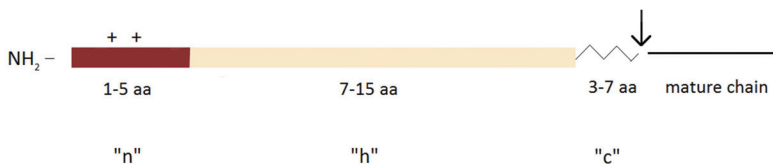
terminus in the case when the first transmembrane segment functions as a reverse signal anchor, as depicted in panel B in figure 8. In both cases the GPCR ultimately gets located in the plasma membrane with an extracellular N-terminus. The core unit of the translocation channel, called translocon in figure 8, is the transmembrane protein complex Sec61 [51]. The Sec61 complex's largest subunit is the  $\alpha$ -subunit, which has ten transmembrane segments [50]. In mammals the other two subunits are termed  $\beta$  and  $\gamma$ , and both have a single transmembrane segment. Each ribosome will associate with four copies of the Sec61 complex, and both the  $\beta$ -subunit and the ribosome are thought to play a part in that interaction. The  $\alpha$ - and  $\gamma$ -subunits are thought to constitute the core of the channel through the ER membrane, with a helical segment of the  $\alpha$ -subunit functioning as a removable plug on the luminal side of the ER (see fig 9). At the cytosolic side the ribosome-Sec61 complex association likely has a permanent opening for the exit of hydrophilic sections of a transmembrane protein. When a sufficiently long and hydrophobic sequence arrives in the translocation channel, a lateral gate opens and the hydrophobic segment is moved into the lipid phase of the membrane [50]. The fully translated and properly folded protein is then transported from the ER membrane by a pathway including formation of membrane vesicles [54], through intracellular compartments like ER-Golgi intermediate compartment (ERGIC), the cis/medial/trans-Golgi apparatus and the trans-Golgi network (TGN) on its way to the cell surface [55]. During or immediately after translocation, receptors with the consensus sequence NxS/T can have carbohydrates attached to its asparagine (N) in what is called N-linked glycosylation [56].



**Figure 9. Schematic illustration of synthesis and translocation of transmembrane proteins into the ER membrane by a ribosome and the Se61 complex/translocation channel.** The removable plug regulating transport into the lumen is depicted in blue, while lateral transport of transmembrane segments into the lipid phase is not shown. Taken from [50].

### 1.2.3 The basic features of signal peptides

The cleavable signal sequence, i.e. the signal peptide, can be divided into three distinct regions based on the properties of the amino acids [57]. The basic design of a signal peptide with the three “n”, “h” and “c” regions can be seen in figure 10.



**Figure 10. The basic design of signal peptides with the three regions and cleavage site indicated.** Modified from the figure presented by Gunnar von Heijne [57].

Starting adjacent to the N-terminal methionine there is a positively charged *n-region*, usually 1-5 amino acids long including positively charged lysine or arginine residues. This region seems to be important for the topological positioning of the signal peptide and thereby ultimately of the N-terminus of the receptor. However, positive charge is not an absolute requirement, as receptor constructs equipped with signal peptides

lacking positive charge except for the N-terminal methionine, or even having a negative charge in this region, may still express mature proteins. Next there is a hydrophobic stretch termed the *h-region*, usually 7-15 amino acids long, that is critical for the signal peptide function. Although mutations leading to charged residues in this area don't necessarily lead to non-expressed proteins, it does severely disrupt the expression efficiency for the protein. Conversely, export-defective proteins with a disrupted h-region have been shown experimentally to obtain an increase in trafficking to the plasma membrane after mutations have increased the hydrophobicity or extended this region [57]. The third region is characterized by a more polar carboxy-terminal domain consisting of 3-7 amino acids, and is termed the *c-region*. This region specifies where, and if, the signal peptide is cleavable, as it is here the signal peptidase enzyme will act. A general rule is that the positions -1 and -3 are particularly important for determining the cleavage site, the so-called “(-3,-1)-rule” [57]. The amino acid in position -1 must be small, essentially alanine, serine, glycine, cysteine, threonine or glutamine. In the -3 position the amino acid must not be aromatic, charged or polar. In addition to these restraints on positions -1 and -3, proline must be absent from position -3 through +1[58]. By systematically exchanging amino acids in this region, studies have shown that it is possible to induce new cleavage sites, and in fact produce proteins with two cleavage sites where cleavage occurs to varying extent at both sites [57].

#### **1.2.4 Protein prediction software: Signal P and TMHMM**

In order to optimize experimental studies as well as avoiding unnecessary laboratory work, the development of advanced computer software predicting protein localization has emerged. Utilizing the known features of signal peptides as described above, it is possible to analyze N-terminal sequences with this kind of software [59]. One such tool is the Signal P software, which is publicly available via the Center for Biological Sequences Analysis (CBS) at the Technical University of Denmark. The software is still evolving from the first version published in 1997 [60], through its current (as of 2014) version Signal P 4.1. Earlier versions are still available online for use. The

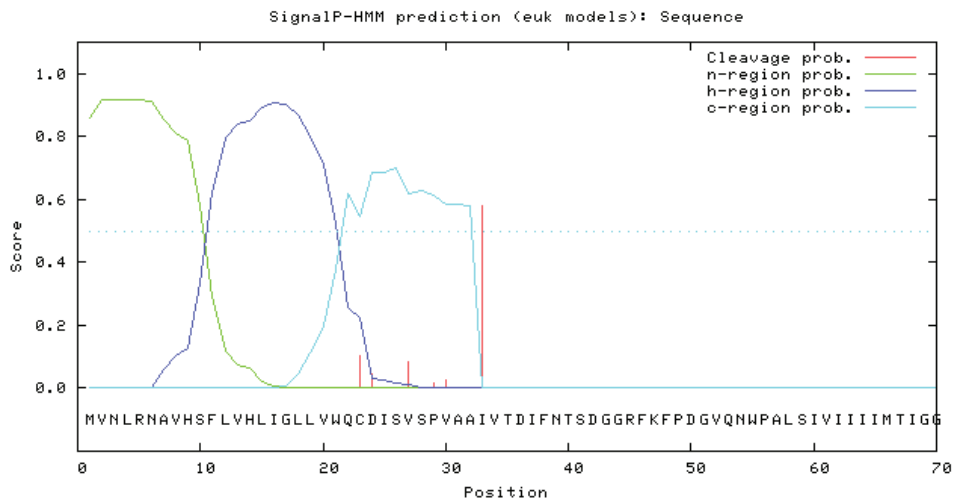


---

third installment of the software, SignalP 3.0, improved the prediction of signal peptides [61] by further developing a method termed the hidden Markov model [62]. In SignalP 3.0 the software can perform analyzes using both the hidden Markov model and the neural networks model, and presents both results. Later versions of SignalP do not use the hidden Markov model, and only present the results using a neural networks method. The result of the prediction analysis will indicate whether the protein has a signal anchor, a signal peptide, or if it is a non-secretory protein. It displays the result as probability of either signal anchor or signal peptide, as well as propose cleavage sites for the signal peptide. There is the possibility of false negatives as well as positives, and especially relevant in the context of analyzing the N-terminal of GPCRs is the possibility that the software confuse the first transmembrane segment for a signal peptide. To avoid this mistake the proteins transmembrane helices can be predicted by using the TMHMM Server also available at the CBS webpage [59]. Examples of such false positives are muscarinic acetylcholine receptor M1 and dopamine receptor D1, where an uncritical analysis of the first 70 amino acids from the N-termini yields predicted signal peptides with cleavage sites proposed at the C-terminal end of the first transmembrane segment (cases marked with \* in table 1). This highlights the importance of taking into consideration the known features of the actual analyzed protein, such as the 7TM structure for GPCRs. If the protein that is investigated is more obscure with unknown features, the risk of misinterpretations makes it even more important to use the predictive theoretical information in tandem with experimental data from the laboratory.

Most GPCRs utilize their first transmembrane domain as a signal anchor, and only 5-10% of GPCRs are estimated to contain cleavable signal peptides [53]. It is estimated that all receptors in the Secretin (class B) [63] and the vast majority in the Glutamate (class C) [51] contain signal peptides. Within the Rhodopsin family (class A) the vast majority of the glycoprotein hormone receptors (subgroup 1C) contain signal peptides. For the luteinizing hormone (LH) receptor it has been shown that inherited mutations in the signal peptide sequence can result in non-cleavage [64]. This will

lead to a reduced expression of functional receptors, and if both parents have LH-inactivating mutations male children will have female phenotype and male pseudohermaphroditism. For the other two Rhodopsin family subgroups, the receptors with small-molecule (1A) and peptide (1B) ligands, signal peptides are thought to be a rare occurrence [51]. A common feature of the signal peptide containing GPCRs is the size of the N-terminus ectodomain [53, 63]. The monoamine GPCRs have in general short N-termini, which in this context should rule out the presence of signal peptides, and to our knowledge no monoamine GPCR have previously been demonstrated to possess a signal peptide. These receptors will rather use the first transmembrane segment as a signal anchor in the translocation of the receptor across the ER membrane. Despite this general rule, the monoamine GPCRs  $\alpha_{2C}$  and 5-HT<sub>2C</sub> are predicted to possess signal peptides when analyzed in Signal P 3.0, as can be seen in table 1. If verified with experimental data, this has implications for the structure of these drug targets at the cell surface. For the 5-HT<sub>2C</sub> receptor the presence of a functional 32 amino acid long signal peptide, would mean the mature receptor expressed at the cell surface lacks the region where the SNP Cys23Ser is situated. In addition, the presence of a signal peptide will influence the expression level of the receptor at the plasma membrane. In the future, perhaps ligands that bind to the cleavage site of the signal peptide may inhibit its cleavage, and thereby decrease the expression of functional receptors at the plasma membrane. The  $\alpha_{2C}$ - as well as the  $\alpha_{1D}$ -adrenoceptor are poorly expressed at the plasma membrane. It's been previously shown that the expression levels of the  $\alpha_{1D}$ -adrenoceptor can be elevated 10-fold by adding a known signal peptide to its N-terminus [65]. This indicates that the unusually long, 95 amino acids, N-terminal tail of the wild type  $\alpha_{1D}$ -adrenoceptor is an obstacle in efficient translocation, since the long N-tail hinders the first transmembrane helix from acting efficiently as a signal anchor. Considering the results from the signal peptide prediction and the potential of the  $\alpha_{2C}$  and 5-HT<sub>2C</sub> receptors as drug targets, the  $\alpha_{2C}$ -adrenoceptor and the 5-HT<sub>2C</sub>-receptor were prime candidates for our present investigations regarding the functional relevance of the N-termini of these two receptors.



**Figure 11.** When analyzing the 5-HT<sub>2C</sub> receptor N-terminus in SignalP 3.0 using the hidden Markov model the result is also presented in graphic form like this plot. The probability score for each of the three signal peptide regions (green, blue and magenta) as well as probability of cleavage at suspected sites (red) can be seen. In this instance the sequence of the first 70 amino acids starting from the N-terminus has been analyzed. The result predicts cleavage at position 33, i.e. the preceding 32 amino acid sequence is predicted to be absent from the mature receptor.

*Signal peptide analysis of selected GPCRs.* A step-by-step procedure for investigating selected GPCRs for presence of signal peptides can be as follows: a) Find the receptor in question at Missouri S&T cDNA Resource Center's website [www.cdna.org](http://www.cdna.org). b) Copy the Genbank accession number (ACC#). c) Enter accession number in the search box at the ncbi databank at [www.ncbi.nlm.nih.gov/nuccore/](http://www.ncbi.nlm.nih.gov/nuccore/). (Step a-c can be substituted by finding the receptor at IUPHARs website [www.guidetopharmacology.org](http://www.guidetopharmacology.org) and clicking the accession number hyperlink) d) Copy the amino acid translation of the gene. e) Test an appropriate length of the translation beginning from the N-terminus in the Signal P 3.0 prediction software at [www.cbs.dtu.dk/services/SignalP-3.0/](http://www.cbs.dtu.dk/services/SignalP-3.0/).

GPCR	SignalP 3.0-HMM result		Accession number	Predicted cleavage site	GPCR Class and ligand
	Signal peptide PRBLTY	Signal anchor PRBLTY			
$\alpha_{2A}$	0.022	0.974	NM_000681.3	Anchor	A/monoamine
$\alpha_{2C}$	0.995	0.004	AY455666	22/23	A/monoamine
5-HT <sub>2A</sub>	0.000	0.116	NM_000621	Non-secretory	A/monoamine
5-HT <sub>2C</sub> -Cys23	0.858	0.057	NM_000868	32/33	A/monoamine
5-HT <sub>2C</sub> -Ser23	0.831	0.060	NM_000868 but Cys > Ser	32/33	A/monoamine
5-HT <sub>2C</sub> -3x HA-tagged	0.000	0.000	28 a.a. insert after start-Met	Non-secretory	A/monoamine
$\alpha_{1D}$	0.002	0.005	NM_000678	Non-secretory	A/monoamine
$\alpha_{1D}$ -SP	0.962	0.015	16 a.a. SP	16/17	A/monoamine
M1*	(0.378)	(0.331)	AF498915	42/43 (false)	A/monoamine
M2	0.007	0.874	AF498916	Anchor	A/monoamine
D1*	(0.530)	(0.442)	NM_000794	46/47 (false)	A/monoamine
D2	0.000	0.943	NM_000795	Anchor	A/monoamine
Corticotropin-releasing factor (CRF) receptor 1	0.986	0.000	AY457172	23/24	A/peptide
CRF receptor 2	0.852	0.000	AY449734	19/20	A/peptide
Endothelin receptor type B	0.999	0.001	AY275463	26/27	A/peptide
$\mu$ opioid receptor	0.001	0.538	AY521028	Anchor	A/peptide
Cannabinoid 1 (CB1) receptor	0.000	0.000	NM_016083	Non-secretory	A/lipid

GPCR	SignalP 3.0-HMM result		Accession number	Predicted cleavage site	GPCR Class and ligand
	Signal peptide PRBLTY	Signal anchor PRBLTY			
CB2 receptor	0.000	0.274	AY242132	Non-secretory	A/lipid
FSH receptor	1.000	0.000	AY429104	17/18	A/glycoprotein
Calcitonin	0.998	0.000	AY430048	24/25	B/peptide
Secretin	1.000	0.000	AY462218	27/28	B/peptide
GABA B2	1.000	0.000	NM_005458	41/42	C/amino acid
Metabotropic glutamate 2	1.000	0.000	NM_000839	18/19	C/amino acid
BA11	1.000	0.000	NM_001702	30/31	Adhesion/ Phosphatidyl-serine
WNT receptor frizzled-4	1.000	0.000	AY462097	36/37	Frizzled/protein

**Table 1. A selection of G-protein coupled receptors analyzed with the signal peptide prediction software SignalP 3.0 using the hidden Markov model.** The first 70 amino acids, counting from the N-terminus, were analyzed for all receptors. Presented is the calculated probability of the receptor having either a signal peptide or signal anchor, and for the predicted signal peptides the most likely cleavage point. A flaw in the software is highlighted by some well characterized GPCRs that are predicted to be non-secretory, like the cannabinoid receptors CB1 and CB2. Also, when analyzing muscarinic acetylcholine receptor M1 and dopamine receptor D1 the software misinterprets the first transmembrane segment for a signal peptide, and proposes a cleavage site within the first transmembrane segment. To denote this mishap M1 and D1 have an asterisk by them and the results are in brackets. Class A = Rhodopsin family, Class B = Secretin family, Class C = Glutamate family, PRBLTY = probability, HMM = hidden Markov model, SP = signal peptide, FSH = follicle-stimulating hormone.

## 2. AIMS OF THE PRESENT STUDY

G-protein coupled receptors that are activated by the monoamine type of endogenous ligands generally have short N-termini, and use their first transmembrane segment as a reverse signal anchor during cotranslational translocation of the receptors into the endoplasmic reticulum membrane. When analyzing the N-termini of the 5-HT<sub>2C</sub> receptor and  $\alpha_{2C}$ -adrenoceptor the protein prediction software Signal P 3.0 concludes that they likely have signal peptides. This is surprising, since these two receptors are the only monoamine GPCRs predicted to have signal peptides. The presence of such signal peptides would have implications for the structure of the mature receptors and for the evaluation of the role of the N-tails at these receptors, which both are considered promising future drug targets. The 5-HT<sub>2C</sub> receptor has a single nucleotide polymorphism within the putative signal peptide sequence. This polymorphism has been proposed to be associated with diverse clinical consequences. The  $\alpha_{2C}$ -adrenoceptor is poorly expressed at the cell surface, while displaying a considerable intracellular pool. This is in contrast to the  $\alpha_{2A}$ -adrenoceptor which is well expressed at the cell surface, and which share a highly conserved amino acid sequence with  $\alpha_{2C}$ , with the exception of the N-tail and the third intracellular loop.

The principal aims of the present project were to:

- i) Determine if the 5-HT<sub>2C</sub> receptor has an N-terminal signal peptide with cleavage site after 32 amino acids.
- ii) Determine if the  $\alpha_{2C}$ -adrenoceptor has an N-terminal signal peptide with cleavage site after 22 amino acids.
- iii) Rebuilding the N-terminal tail of the 5-HT<sub>2C</sub> receptor, as well as that of the  $\alpha_{2C}$ -adrenoceptor, for investigating the influence of the N-terminus on the expression level of these two receptors at the plasma membrane.

---

iii) Create chimeras of the  $\alpha_{2A}$ - and  $\alpha_{2C}$ -adrenoceptors for investigating what parts of the  $\alpha_{2C}$ -adrenoceptor are responsible for: a) its poor expression level at the plasma membrane, and b) for the  $\alpha_{2C}$ -binding-selectivity of ligands.

### 3. SUMMARY OF RESULTS (PAPERS I, II AND III)

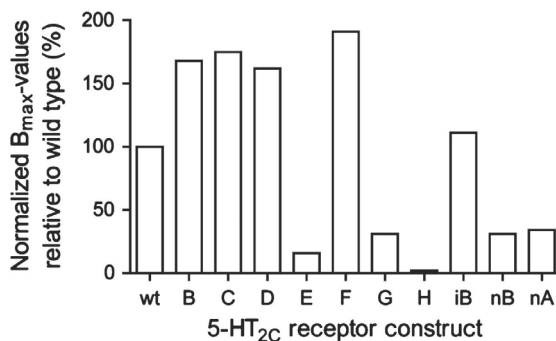
#### 3.1 PAPER I

##### **The N-terminal region of the human 5-HT<sub>2C</sub> receptor has a cleavable signal peptide.**

The principal aim of this paper was to determine if the human 5-HT<sub>2C</sub> receptor has a cleavable signal peptide in its N-terminus. This task was performed by engineering receptor constructs containing an eight amino acid long epitope (FLAG), transiently transfecting COS-7 cells and immunostaining the expressed receptor constructs with the selective antibody M1. The M1 antibody binds to the FLAG epitope only when the epitope is the outmost N-terminal part of the protein. Thereby, inserting the FLAG epitope after the predicted cleavage site makes cleaving detectable by the M1 antibody. The M2 antibody used as control binds to the FLAG epitope regardless of its position in the N-terminus. The 5-HT<sub>2C</sub> receptor has been shown to have a prevalent single nucleotide mutation in its N-terminus where 13% of the Caucasian population has a serine instead of cysteine (Cys23Ser). When analyzing the N-terminus sequence in the protein prediction database SignalP, the results show that the outermost 32 amino acids in its N-terminus are predicted to be cleaved off. This made us interested in validating the prediction results experimentally. At first this was done by checking if the M1 and M2 antibodies were binding to our constructs by qualitative investigation with epifluorescence microscopy. The receptor construct modelling for the wild type receptor (5-HT<sub>2C</sub>-B), i.e. where the only modification was incorporating the epitope immediately downstream of the predicted cleavage site, resulted in positive staining with the selective antibody M1 (5-HT<sub>2C</sub>-B in table 2, paper I). Receptor densities in COS-7 membranes determined by radioligand binding with [<sup>3</sup>H]-mesulergine, were also quantified for all ten engineered constructs. Every binding experiment also contained the wild type receptor with its receptor density set to 100%, and the simultaneously determined densities for the constructs were normalized to the wild type value. The relative expression levels of binding competent receptor constructs supported the qualitative result obtained with



epifluorescence, which indicate that the receptor has a functional signal peptide, as hindering cleavage of the putative signal peptide resulted in a relative expression level of only 30% (construct nA in table 1 and figure 3, both in paper I).



**Figure 12. Relative expression levels of wild type 5-HT<sub>2C</sub> receptor and engineered receptor constructs.** Bar chart is showing binding competent 5-HT<sub>2C</sub> receptors in cell membranes after transient transfection of COS-7 cells, with values normalized to that of the wild type receptor. In paper I, figure 1, an overview of the receptor constructs are shown. This figure is taken from paper I.

To verify the results from epifluorescence and radioligand binding experiments, the immunostaining protocol was optimized for confocal imaging. In addition to validating the previous results (panel a in figure 1, paper I), the confocal imaging experiments also quantified the fluorescence intensity across the cell (panel e and f in figure 1, paper I), thus confirming that we were visualizing receptors expressed at the cell surface. The main conclusion of these experiments is that the human 5-HT<sub>2C</sub> receptor N-terminus has a 32 amino acid long cleavable signal peptide. Consequently the mature receptor does not contain the aforementioned Cys23Ser mutation.

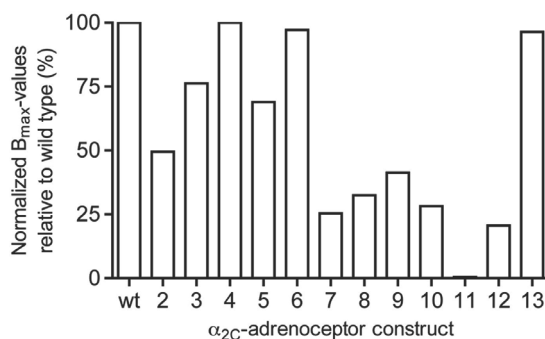
## 3.2 PAPER II

### **The predicted N-terminal sequence of the human $\alpha_{2C}$ -adrenoceptor does not act as a functional cleavable signal peptide.**

The principal aims of this paper was to determine if the human  $\alpha_{2C}$ -adrenoceptor has a cleavable signal peptide (SP) in its N-terminus, and to what extent the N-terminus is contributing to the low expression levels of the receptor. As in paper I, investigating the presence of the predicted cleavable signal peptide was achieved by engineering among others a receptor construct containing the FLAG epitope immediately after the predicted cleavage site in the N-terminus, transiently transfecting COS-7 cells and immunostaining the expressed receptor constructs with the selective antibody M1. It became evident in the preliminary qualitative investigation with epifluorescence microscopy that the receptor construct equipped with FLAG epitope after the predicted signal peptide sequence of the wild type ( $\alpha_{2C}$ -3 in table 2 in paper II) exhibited no staining, i.e. suggesting the predicted signal peptide was not cleaved off. This result was confirmed when the experimental protocol was optimized for investigations with confocal microscopy. As can be seen in figure 4 in paper II, the staining was negative when  $\alpha_{2C}$ -3 was stained with the selective antibody M1 (panel A) while positive when staining with the M2 antibody control (panel B). For the cleavage control construct,  $\alpha_{2C}$ -5, which has the established influenza hemagglutinin signal peptide (HAsp), both M1 and M2 showed positive staining (panel C and D). In panel E and F in figure 4, the presence of the receptors at the cell surface was confirmed by measuring the fluorescence intensity across a cell expressing the control construct  $\alpha_{2C}$ -5. Consequently, the results show that the 22 amino acid long putative endogenous signal peptide is indeed not cleavable, in contrast to the prediction.

To investigate to what extent the N-terminus is contributing to the low expression levels of the receptor; eight additional receptor constructs without FLAG epitope were engineered. These constructs are presented schematically in figure 1 in paper II,

and are termed: 2, 4 and 8-13. Construct 2 contains a known cleavable signal peptide instead of the putative endogenous signal peptide, while in construct 4 the known signal peptide is inserted before the full length wild type receptor. This known signal peptide is the above mentioned influenza hemagglutinin signal sequence (HAsp), which has previously been shown to enhance the expression of  $\alpha_{1D}$ - and  $\beta_2$ -adrenoceptors [65, 66]. Construct 13 has a lysine inserted at amino acid position 3, in an attempt to increase the positive charge of the n-region of the putative signal peptide. A more positively charged n-region could increase the efficiency of translocation [57] and thereby increase the expression level. The remaining constructs (8-12) are to varying degree truncated at the N-terminus, which is a maneuver previously shown to enhance the expression of  $\alpha_{1D}$ -adrenoceptors [65]. Among the receptor constructs without interfering FLAG epitope, constructs 4 (wild type with added HAsp) and 13 (improved endogenous SP) have the same expression levels as the wild type, while the remaining constructs (2 and 8-12) have 50% or less relative expression of binding competent receptors when normalized to the wild type. Figure 3 from paper II with the relative expression levels of the receptor constructs normalized to the wild type is shown below.



**Figure 13. Expression levels of wild type  $\alpha_{2C}$ -adrenoceptor and 12 constructs with modified N-terminus.** Bar chart displaying the  $B_{max}$ -values normalized relative to the wild type  $\alpha_{2C}$ -adrenoceptor. Receptor constructs 3, 5, 6 and 7 have FLAG epitope. Taken from paper II [34].

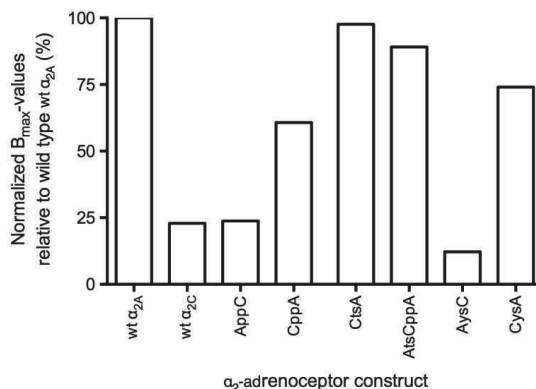
Consequently, the results show that the endogenous N-terminus of the  $\alpha_{2C}$ -adrenoceptor is not fully responsible for the low expression levels of the receptor, but rather is a necessary component for the receptor to be well expressed. This suggests the truncated receptors are not capable to compensate for their deleted N-tail stretches by efficiently utilizing the first transmembrane segment as a signal anchor. See for example the poorly expressed construct 12, where the majority of the N-terminus is truncated. It is especially interesting that none of the two constructs containing the influenza hemagglutinin signal sequence (2 and 4) had increased expression. This means that among our tested constructs the endogenous N-terminus is optimal for expression of the  $\alpha_{2C}$ -adrenoceptor.

### 3.3 PAPER III

The aims in this paper were two-fold: a) Continuing the investigations from paper II into which part of the  $\alpha_{2C}$ -adrenoceptor is responsible for the low expression levels of the receptor, and b) Investigate what parts of the  $\alpha_{2A}$ - and  $\alpha_{2C}$ -adrenoceptor subtypes are most important for drug selectivity between them. In particular investigating whether transmembrane region 1 (TM1) and extracellular loop 2 (ECL2) have a general effect on the affinity for “bulky” drugs at the  $\alpha_{2A}$ -adrenoceptor. To achieve these aims six  $\alpha_{2A}/\alpha_{2C}$ -adrenoceptor chimeras were constructed, the chimeras and the wild type receptors were transiently transfected to HEK293 cells, and the receptor densities in the resulting membrane preparations were determined by radioligand binding. A schematic representation of the chimeric  $\alpha_{2A}/\alpha_{2C}$ -adrenoceptors can be seen in figure 2 in paper III, while the specific conserved amino acids where the chimeric regions are spliced together are shaded in figure 3 in the paper (i.e. “YS”, “TS” and “PP”). These amino acids are also part of the names given to the receptor constructs.

To compare the N-termini's influence on the expression levels, two of the six chimeras had the wild type N-termini preceding the conserved YS-motif exchanged, creating the CysA and AysC chimeras. The expression levels were normalized

relative to the wild type  $\alpha_{2A}$ -adrenoceptor, and are presented in table 1 in paper III [67]. The data from table 1 in paper III is presented below in a bar chart (figure 14).



**Figure 14. Expression levels of wild type  $\alpha_{2A}$ - and  $\alpha_{2C}$ -adrenoceptors and six chimeric constructs.** Relative  $B_{max}$  expression levels of binding competent wild type  $\alpha_{2A}$ - and  $\alpha_{2C}$ -adrenoceptors and six  $\alpha_{2A}/\alpha_{2C}$ -adrenoceptor chimeras, normalized to the wild type  $\alpha_{2A}$ -adrenoceptor. The figure is depicting the data published in table 1 in paper III [67].

While the wild type  $\alpha_{2C}$ -adrenoceptor had a relative expression level of about 23% compared to  $\alpha_{2A}$ , the chimeric receptor containing the N-terminus from  $\alpha_{2A}$ -adrenoceptor (AysC) only had 12%. This indicates that the  $\alpha_{2A}$ -adrenoceptor N-terminus could not overcome the low expression levels of the wild type  $\alpha_{2C}$ -adrenoceptor. The other chimeric receptor constructs focus on the influence of the  $\alpha_{2C}$ -adrenoceptor N-terminus and TM1 region (CtsA), TM2-TM4 region (AtsCppA) and the C-terminal half of the receptor (CppA and AppC). As can be seen in figure 14, the CtsA and AtsCppA chimeras' expression levels are comparable to the wild type  $\alpha_{2A}$ -adrenoceptor, at about 97% and 89% respectively. These two constructs have their C-terminal half taken from  $\alpha_{2A}$ . The lack of expression-modifying effect seems to rule out that the N-terminus and TM1 or TM2-TM4 regions, which in these chimeras are derived from  $\alpha_{2C}$ , have significant impact on the expression level of the

---

wild type  $\alpha_{2C}$ -adrenoceptor. Supporting these observations that the N-terminal half of the  $\alpha_{2C}$ -adrenoceptor does not have a negative effect on the expression of the chimeric receptors, the AppC construct, having the C-terminal half taken from  $\alpha_{2C}$ , is as poorly expressed as the wild type  $\alpha_{2C}$ -adrenoceptor. Consequently, the C-terminal half of the receptor determines the  $\alpha_{2C}$ -adrenoceptor low expression levels at the cell surface.

In addition to the wild type  $\alpha_{2A}$ - and  $\alpha_{2C}$ -adrenoceptors, the chimeras AppC, CppA, CtsA and AtsCppA were used to investigate if the TM1 and ECL2 have a general effect on the affinity for “bulky” drugs at the  $\alpha_{2A}$ -adrenoceptor as proposed by Laurila et al [68]. These chimeras are the same ones as described above (figure 2 and 3 in paper III), and represent the three regions: i) N-terminus-TM1, ii) TM2-TM4 and iii) the C-terminal half from ECL2 onwards. The affinities of seven different ligands were determined for the wild type receptors and for the chimeras. Spiroxatrine, clozapine, spiperone, MK912 and chlorpromazine are all fairly “bulky”  $\alpha_{2C}$ -selective ligands, while atipemazole is rather non-selective and BRL44408 is  $\alpha_{2A}$ -selective. In table 2 the resulting drug affinities are presented with  $pK_i$  values and the drug’s  $K_i$  ratios at  $\alpha_{2A}$  versus the other receptor constructs. Analyzing the results in the table three patterns emerge: 1. Spiroxatrine and MK912 showed high affinity for  $\alpha_{2C}$  and AppC, i.e. the C-terminal half of the receptor is the region determining the drug’s affinities. 2. Spiperone and clozapine showed the highest affinity for receptors containing the C-terminal half of  $\alpha_{2C}$ , but also intermediate affinities for constructs where the TM2-TM4 region was derived from  $\alpha_{2C}$ . 3. Of the five  $\alpha_{2C}$ -selective ligands tested, only chlorpromazine showed higher affinity for CtsA than for wild type  $\alpha_{2A}$ , i.e. the TM1 region of  $\alpha_{2A}$  contribute to the low affinity of chlorpromazine, but not spiroxatrine, clozapine, spiperone or MK912 at  $\alpha_{2A}$  as compared to CtsA. The consequence of these results is that the TM1 and ECL2 regions of the  $\alpha_{2A}$ -adrenoceptor do not have a general effect on the affinity for “bulky” drugs at the  $\alpha_{2A}$ -adrenoceptor.

---

## 4. GENERAL DISCUSSION

### 4.1 METHODOLOGICAL CONSIDERATIONS

The results obtained in paper I, II and III are all based on the same methodological concepts: Rebuilding the wild type receptors by PCR and transiently expressing the mutated receptors in mammalian cell lines COS-7 (papers I and II) and HEK293 (paper III). The receptor constructs expressed at the cell surface are then investigated by immunofluorescence and/or radioligand binding. The radioligand binding experiments are used as a control, in order to determine the expression levels of the receptor constructs that were used in the immunostaining experiments, which, in turn, were performed for investigating the presence of signal peptides. In addition to this the radioligand binding experiments are an independent tool with regard to investigating different receptor construct's efficiency at being expressed at the plasma membrane, by measuring receptor densities in membrane preparations and whole cells. The presence of ligand binding receptors expressed after transfection are chosen as the experimental end-point measurement, presumably reflecting the efficiency of the integration of the receptor into the ER membrane, as this is the critical point when investigating N-terminal function of these receptors. All the radioligand binding experiments include the wild type receptors in addition to the engineered constructs.

In paper I, investigating the 5-HT<sub>2C</sub>-receptor, 10 new constructs were engineered to investigate for the presence and cleavability of the putative signal peptide and its relevance for expression. Seven of the constructs contain an epitope for immunostaining, while three constructs are purely for investigation with radioligand binding.

In paper II, investigating the  $\alpha_{2C}$ -adrenoceptor for a cleavable signal peptide and the N-terminus' effect on expression levels, 12 new constructs were engineered. Four of the constructs contained an epitope for immunostaining, while the remaining eight

constructs are purely for investigating receptor densities at the cell surface with radioligand binding.

In paper III, six chimeric  $\alpha_{2A}$ -/ $\alpha_{2C}$ -adrenoceptor constructs were engineered to investigate expression levels as well as drug subtype selectivity. Consequently none of these receptors are encoded with the epitope, and they are exclusively investigated with radioligand binding.

As the choice of expression vector and cell line can have a significant impact on protein production [69] all 5-HT<sub>2C</sub> constructs are expressed with pcDNA3.1+ in COS-7 cells, while the  $\alpha_{2C}$ -adrenoceptor constructs in paper II are expressed with pCi-neo in COS-7 cells and all constructs in paper III with pcDNA3.1+ in HEK293 cells. By doing this we ensured that the results could be normalized to the wild type receptor for each set of experiments.

#### **4.1.1 Investigating N-termini for signal peptide function with the FLAG epitope**

To investigate whether the putative cleavable signal peptides of the 5-HT<sub>2C</sub> and  $\alpha_{2C}$  receptors in fact are cleaved, we decided to use the well-established technique of building receptor constructs with the encoded FLAG epitope inserted in the N-terminal [70]. The FLAG epitope is an eight amino acid long sequence: DYKDDDDA [66]. Transiently transfecting mammalian cell lines like COS-7 and HEK293 works as a model for how the wild type receptor might work in vivo [71]. The FLAG epitope is encoded in the N-termini at the position where the putative cleavage site is proposed, thereby being the first sequence of the mature receptor if the putative signal peptide is cleaved. The epitope is recognized by a selective primary antibody (M1) that will only bind to the epitope if it is the absolute N-terminus of the receptor. Even the presence of a single methionine encoded by the start codon in front of the FLAG epitope is enough to prevent binding of the M1 antibody [70]. By labeling the selective antibody with a fluorescent secondary



---

antibody, it is possible to visualize the cleaving of the putative signal peptide with fluorescent microscopy. A non-selective primary antibody (M2) that recognizes the epitope regardless of its position in the N-terminus, acts as necessary control to ensure that the epitope-tagged receptor is expressed at the cell surface and can be visualized, independent of whether or not the FLAG epitope has become outmost N-terminal by cleavage of the signal peptide. An alternative strategy could have been to label the primary antibodies M1 and M2 with fluorescent probes, making the use of labeled secondary antibodies unnecessary. However, secondary antibodies offer increased sensitivity through the signal amplification that occurs as multiple secondary antibodies bind to a single primary antibody. In addition, exactly the same amount of the secondary antibody can be added to M1 and M2 treated samples, making background fluorescence identical for the two samples.

Although the technique is well-established, it is necessary to ensure that this model indeed approximates the *in vivo* function of the receptor. This necessitates the engineering of receptor constructs other than the wild type receptor with inserted FLAG epitope. This includes both negative and positive controls for the primary antibodies, as well as controls to ensure the epitope itself does not have an impact on the N-terminal function and subsequently expression of the receptors. The complete lists of 5-HT<sub>2C</sub> receptor and  $\alpha_{2C}$ -adrenoceptor constructs investigated can be seen in figures 1, in paper I and II respectively. Positive control constructs, are represented by constructs having a known cleavable signal peptide instead of the putative endogenous signal peptide, with the cleavage site immediately prior to the epitope. This method has previously been described by Kobilka and co-workers for studies on the  $\beta_2$ -adrenoceptor [66]. In the studies presented in papers I and II, we use both the same signal peptide sequence from influenza hemagglutinin [72] as well as the same epitope as Kobilka's group.

The epitope-tagged receptor constructs were investigated by epifluorescence microscope, which gave us a qualitative result, i.e. positive staining or not. The experiments were performed on cells without permeabilization, as it is the mature

receptors expressed at the cell surface that is of interest to us. Then we wanted to verify that the fluorescence we could see in the epifluorescence microscope indeed was at the cell surface. This was performed by optimizing the techniques used for the epifluorescence microscope, and visualizing the results in a confocal fluorescence microscope. By using a confocal microscope it is possible to visualize the 3D structure of the cell, optically slice-by-slice, from the outermost point of the cell surface through the interior to the cell surface on the opposite side of the cell. By measuring the fluorescent intensity across one such section of a cell it is possible to quantify the difference in fluorescence intensity from the cell membrane to the interior of the cell. This is not meant as a method to exclude the presence of intracellular receptors, as the receptors were not permeabilized, but to verify that we are observing receptors localized to the cell surface.

#### **4.1.2 The N-terminus relevance for expression of cell surface receptors**

Whether the 5-HT<sub>2C</sub> receptor and  $\alpha_{2C}$ -adrenoceptor have cleavable signal peptides or not, their N-termini's relevance for expression levels of the receptors are worth investigating. This is because the presence of a hydrophobic stretch, which is present within the N-tail of these two receptors, clearly is an unusual feature within the monoamine class of GPCRs. By engineering constructs with truncated N-terminal tail [65, 73], or with the addition of the aforementioned signal peptide sequence from influenza hemagglutinin at the N-terminus [65, 66, 73], receptor densities at the plasma membrane have been greatly enhanced for the  $\alpha_{1D}$ - and  $\beta_2$ -adrenoceptors and the cannabinoid receptor CB1. A different cleavable signal peptide named the LRRC32 signal peptide, or Lucy tag, has recently been shown to enhance the cell surface expression of olfactory GPCRs in HEK293T cells when positioned at the N-terminus [12]. Research on these receptors have been hindered by their low or non-existent cell surface expression in cell lines, so this method of increasing expression without changing the structure of the mature receptor could accelerate our knowledge of this very large group of GPCRs. The constructs used in the here presented studies,

---

to reveal the role of the N-terminal tails of the 5-HT<sub>2C</sub>- and  $\alpha_{2C}$ -receptors, are described in the following sections.

#### **4.1.3 The 5-HT<sub>2C</sub> receptor N-terminus' influence on expression**

The major aim of investigating the 5-HT<sub>2C</sub> receptor was determining whether the predicted signal peptide is present and thereby the presence of the single nucleotide polymorphism Cys23Ser in the mature receptor. Consequently the constructs designed for this receptor focuses on changes related to the putative signal peptide. Constructs include clones with the endogenous signal peptide exchanged for the influenza hemagglutinin signal peptide (5-HT<sub>2C</sub>-E in figure 1 in paper I), truncated at the proposed cleavage site (5-HT<sub>2C</sub>-D in figure 1 in paper I), and with single mutation to induce non-cleavage (5-HT<sub>2C</sub>-nA and 5-HT<sub>2C</sub>-nB in figure 1 in paper I). Thereby, we could detect whether the predicted signal peptide could be cleaved, detect the expression level of the receptor when the putative signal peptide was deleted and the expression level of the receptor when the putative signal peptide could no longer be cleaved off.

#### **4.1.4 The $\alpha_{2C}$ -adrenoceptor N-terminus' influence on expression**

Considering the low expression level of the  $\alpha_{2C}$ -adrenoceptor and significant intracellular location, several constructs were made in an attempt to increase the expression of mature ligand binding receptors at the cell surface. The engineered constructs are analogous to earlier successful expression-enhancing changes made on the  $\alpha_{1D}$ -adrenoceptor,  $\beta_2$ -adrenoceptor and CB1 cannabinoid receptor [65, 66, 73]. This includes various truncations of the N-terminus, in addition to adding the influenza hemagglutinin signal peptide to both the wild type receptor and instead of the putative endogenous signal peptide. Due to the large intracellular pool of  $\alpha_{2C}$ -adrenoceptors, we performed radioligand binding studies not only on membrane preparations, but also on whole cells and matching membrane preparations (paper II).

By doing the experiments with whole cells and matching membrane preparations, we could determine if intracellular  $\alpha_{2C}$ -adrenoceptors were binding-competent and to what extent they might influence the saturation experiments of membrane preparations. The  $\alpha_{2A}$ -adrenoceptor is efficiently expressed at the cell surface and has negligible intracellular location, and was used as a model to compare binding to cell surface receptors versus total membrane receptors (including intracellular receptors). In paper III, the N-terminus as well as the first transmembrane helix (TM1), the second to fourth transmembrane helix (TM2-TM4) and the C-terminal half of the  $\alpha_{2C}$ -adrenoceptor, were further investigated with regard to relevance for expression levels. These investigations used the  $\alpha_{2A}$ -adrenoceptor as a model, and the receptor constructs are chimeras of the  $\alpha_{2C}$ - and  $\alpha_{2A}$ -receptors. Thereby we could detect what parts of the  $\alpha_{2C}$ -adrenoceptor had the largest impact for the lower expression level of the  $\alpha_{2C}$ -adrenoceptor compared to the well-expressed  $\alpha_{2A}$ -adrenoceptor.

#### **4.1.5 Which region of the $\alpha_{2C}$ -adrenoceptor determines selectivity for “bulky” drugs**

Scheinin and co-workers investigated the possible role of the N-terminus and first transmembrane domains of  $\alpha_2$ -adrenoceptors on drug subtype selectivity of “bulky” drugs like spiroxatrine and chlorpromazine, that are  $\alpha_{2C}/\alpha_{2B}$ -selective over  $\alpha_{2A}$ -adrenoceptors. This was done by combining chimeric receptor constructs and radioligand binding with receptor modelling [68]. Their conclusions included that the first transmembrane domain of the  $\alpha_{2A}$ -adrenoceptor has an indirect effect on receptor conformation, and thereby decreasing the binding affinity of “bulky” drugs at the wild type  $\alpha_{2A}$ -adrenoceptor. To investigate whether these results could be reproduced and generalized to other “bulky” drugs, the wild type  $\alpha_{2C}$ - and  $\alpha_{2A}$ -adrenoceptors and chimeras exchanging the TM1, TM2-TM4 and the C-terminal half regions of the receptors were investigated for subtype selectivity of seven drugs. Five drugs are fairly bulky and  $\alpha_{2C}$ -selective: Spiroxatrine, clozapine, spiperone, MK912 and chlorpromazine. Atipemazole was chosen because it is rather non-selective and BRL44408 because it is  $\alpha_{2A}$ -selective. The experiments were performed by transient

---

expression of the receptors in HEK293 cells, followed by [<sup>3</sup>H]-RX821002 radioligand binding competition experiments of the tested drugs. The resulting pK<sub>i</sub>-values and the calculated K<sub>i</sub>-values are used to determine which region(s) of the α<sub>2C</sub>-receptor determines the affinity for the tested drugs. The K<sub>i</sub>-values for the individual drugs were compared to their K<sub>i</sub>-value at the α<sub>2A</sub>-receptor, i.e. α<sub>2C</sub>-selectivity is represented by a ratio of above 1 and α<sub>2A</sub>-selectivity by a ratio of below 1. Thereby, we could detect how the three investigated regions of the α<sub>2A</sub>- and α<sub>2C</sub>-adrenoceptors influences drug selectivity for the tested ligands.

## 4.2 DISCUSSION OF 5-HT<sub>2C</sub> RESULTS

### 4.2.1 The 5-HT<sub>2C</sub> receptor has a cleavable signal peptide

The main conclusions from our study of the human 5-HT<sub>2C</sub> receptor is that it has a N-terminal 32 amino acid long cleavable signal peptide, which must be cleaved off for efficient translocation of the receptor to the membrane. To the best of our knowledge this is the first monoamine GPCR identified with a cleavable signal peptide in its N-terminus. As a consequence the mature receptor expressed at the cell surface does not contain the single nucleotide polymorphism Cys23Ser. This mutation occurs among 13% of the Caucasian population, and is being investigated for relevance regarding personal susceptibility to psychiatric disorders and outcome of pharmacological treatments.

As the cleavage of the proposed signal peptide sequence was identified with specific immunolabeling, the investigated construct was the wild type receptor modified with the inclusion of an epitope. This means of course that it strictly speaking is the engineered receptor construct with the FLAG epitope that has been shown to have a cleavable signal peptide. There are however supporting evidence for our conclusion. Thus, in our investigation of receptor densities in membranes with radioligand binding, we show that a very small change, at a very critical site for the functionality of the signal peptide, had a dramatic effect on the expression level of the receptor.

This was done by changing one amino acid at the end of the putative signal peptide in position -1 relative to the cleavage site from alanine to asparagine, and thus cleavage at the proposed site should be hindered as per the (-1, -3)-rule [57]. When compared to the wild type receptor this single amino acid mutation resulted in a 70% reduction of the expression of the receptor. This also implies that the cleavage of the signal peptide sequence is essential for efficient translocation of the wild type receptor. The signal peptide is however not necessary for expression of the receptor, as the construct lacking the 32 amino acids indeed had an increased expression compared to the wild type receptor. This suggests that the receptor can in the absence of a signal peptide in the N-terminus efficiently utilize the first transmembrane segment as a signal anchor. Accordingly, the translocation into the ER of a truncated 5-HT<sub>2C</sub> receptor seems to be transformed from signal peptide-mediated in the wild type receptor to signal anchor-mediated when truncated. In support of our conclusion that 5-HT<sub>2C</sub> has a cleavable signal peptide is a study into 5-HT<sub>2C</sub> receptor dimerization [74], where amino acid sequencing of the N-terminus showed it started with amino acid 33. Also, earlier work with in vitro mRNA transcription and translation of the mouse 5-HT<sub>2C</sub> receptor followed by gel electrophoresis indicated a cleavable signal peptide [75].

#### **4.2.2 Incorporating epitopes in the N-terminus influence receptor expression**

Even though the FLAG epitope is only eight amino acids long and is incorporated in the N-terminus, and engineering constructs with epitopes for immunolabeling is widely used in experiments, our results show that such changes can have substantial impact on receptor expression. By comparing receptor densities in membranes expressing constructs that are differing only by the inclusion of the epitope, it becomes obvious that the epitope itself can influence the expression level of the receptor. With expression levels normalized relative to the wild type receptor, the FLAG epitope situated immediately after the putative signal peptide increased expression levels to about 160% (5-HT<sub>2C</sub>-B in table 1 and figure 1 in paper I). In

---

contrast there was a decrease in expression levels to about 30% when the epitope is situated immediately before the putative signal peptide (5-HT<sub>2C</sub>-G in table 1 and figure 1 in paper I). It is obvious that inserting an epitope before a signal peptide will have a large negative impact since the signal peptide is dependent on its consensus n-h-c structure. For receptors using the first transmembrane segment as a signal anchor the negative impact on expression should be less pronounced. This highlights the importance of corroborating experimental data gathered by immunolabeling epitope-incorporated constructs, with data gained by an alternative method such as radioligand binding studies to assure that the constructs are well expressed and functional.

## 4.3 DISCUSSION OF ALPHA<sub>2C</sub>-ADRENOCEPTOR RESULTS

### 4.3.1 The $\alpha_{2C}$ N-terminus does not contain a 22 amino acids long cleavable signal peptide

As we try to elucidate what distinct functional characteristics the  $\alpha_{2C}$ -adrenoceptor N-terminus might have, we can with confidence say it is unlikely to have a cleavable signal peptide with a length of 22 amino acids. However, since the investigated receptor construct ( $\alpha_{2C}$ -3 in paper II) does contain the FLAG-binding epitope inserted after the putative signal peptide, the downstream environment of the putative signal peptide is altered compared to the wild type receptor. On the other hand, the positive cleavage control construct,  $\alpha_{2C}$ -5 in paper II, had its hemagglutinin signal peptide positioned at the same place as the endogenous putative signal peptide in  $\alpha_{2C}$ -3, and it was properly cleaved off. There is also a possibility that the  $\alpha_{2C}$ -adrenoceptor does have a cleavable signal peptide in its N-terminus, but with a different cleavage site. When analyzing the N-terminus in SignalP 4.1, the latest version of the signal peptide prediction software, the N-terminus is predicted to have two possible cleavage sites, after amino acid number 15 or 22, with a higher predictive value for position 15/16 [76]. In addition, the presence of an N-linked glycosylation site at position 19 of the  $\alpha_{2C}$ -adrenoceptor makes cleaving after 15 amino acids much more probable than

cleaving after 22 amino acids. Further studies outlined in the future perspective chapter would be required to confidently conclude whether the  $\alpha_{2C}$ -adrenoceptor after all has a cleavable signal sequence, in that case with a different cleavage site, presumably after amino acid 15 in its N-terminus.

#### **4.3.2 The N-terminus is necessary for efficient expression of mature $\alpha_{2C}$ -adrenoceptors**

With a total of 13  $\alpha_{2C}$ -adrenoceptor constructs investigated in paper II and an additional 6  $\alpha_{2C}$ - and  $\alpha_{2A}$ -receptor chimeras investigated in paper III, the  $\alpha_{2C}$ -adrenoceptor's N-terminus is thoroughly investigated with regard to influence on expression levels of binding competent receptors. The main conclusion of these investigations is that the  $\alpha_{2C}$ -adrenoceptor N-terminus is necessary for efficient expression of binding competent receptors. So even though the wild type receptor is poorly expressed at the cell surface and the N-terminus is unusually long, all our attempts at increasing its expression level by modifying the N-terminus failed. The common feature of the receptor constructs that have expression levels comparable to the wild type receptor, is that they have an intact endogenous N-terminus.

In contrast to our conclusions regarding the global effect of the N-terminus on expression of binding competent  $\alpha_{2C}$ -adrenoceptors, Angelotti and co-workers have proposed that the hydrophobic stretch of the N-terminus is a retention signal regulating trafficking of the receptor from the endoplasmic reticulum to the cell surface [35]. The differing conclusions regarding the effect of the N-terminus on expression of binding competent receptors as determined by radioligand binding are striking. A detailed discussion of the different interpretation of our results can be read in paper III. While we don't dispute Angelotti and co-workers results from studying the trafficking of epitope-containing receptor constructs, their radioligand binding results (table 3 in [35]) are at odds with ours. They seem to have an increase in the expression of binding competent  $\alpha_{2C}$ -receptor constructs both by truncating the N-terminus and by exchanging the parts of the  $\alpha_{2C}$  N-terminus for the  $\alpha_{2A}$ . Whether



---

there are methodological differences is difficult to interpret from the published data, especially since their table of radioligand binding data for truncated receptors lack the wild type control. Since we are always performing matching transfections and radioligand binding experiments and subsequently normalizing the data to that of the wild type receptor, the expression levels of the different receptor constructs is easily comparable in our data. Another possible extraneous influence on the expression of binding competent receptors is the presence of epitopes in the N-terminus of the receptor. In their trafficking experiments Angelotti and co-workers use receptor constructs containing an epitope in the N-terminus. As we have shown in our study of the 5-HT<sub>2C</sub> receptor the eight amino acid long FLAG epitope incorporated into the N-terminus can have a substantial impact on the expression levels of the receptor.

The publication of paper III prompted Angelotti and Hurt to comment on our results in a letter to editor [77]. We answered this comment from Angelotti and Hurt in a rebuttal letter to the editor [78]. The crux of the disagreement seems to stem from the different data on expression of binding competent  $\alpha_{2C}$ -adrenoceptors as determined by radioligand binding, since we have not studied trafficking of receptors or intracellular location. In support for our radioligand binding data are the work of Laurila and co-workers [68], where a chimeric  $\alpha_{2C}$ -adrenoceptor containing the N-terminus and TM1 from  $\alpha_{2A}$  showed lower expression level than the wild type  $\alpha_{2C}$ -adrenoceptor. However, as they are using stable transfection of CHO cells, the expression levels can be arbitrary depending on the stable transfection process.

The consensus regarding translocation and regulation of trafficking of GPCRs is that the N-terminus and TM1 are important for the transfer of the receptor from the ribosome to the ER [79], while intracellular loops and the C-terminal tail are important with regard to regulation of trafficking of receptors from the ER to the plasma membrane [80]. In support of this general consensus, Motawea and co-workers recently showed that trafficking of  $\alpha_{2C}$ -adrenoceptors from the Golgi compartment to the cell surface was mediated through an arginine-rich (R) region in the C-terminus [40]. The arginine-rich region is not conserved among the  $\alpha_2$ -

adrenoceptor subtypes, and is only present in the  $\alpha_{2C}$ -adrenoceptor where it includes amino acids 454-458, RRRRR, close to TM7. This is in close proximity to the two conserved regions among the  $\alpha_2$ -adrenoceptor subtypes; NPXXYXXF and FXXFXXXF in the end of TM7 and the very beginning of the C-terminus [40]. The NPXXYXXF region is important for the receptor structure and perhaps for Golgi network targeting, while the FXXFXXXF region encodes an endoplasmic reticulum export signal that regulates cell surface expression [81]. Motawea and co-workers showed that the RRRRR motif in the  $\alpha_{2C}$ -adrenoceptor has a dominant-negative effect resulting in an intracellular pool of functional receptors [40]. By mutating all five arginines this ER-Golgi retention sequence lost its effect and resulted in increased expression at the cell surface. The mechanism is suggested to be that filamin-2, an actin-binding protein, binds to the arginine-rich region mentioned and blocks its negative influence, resulting in translocation of the receptor to the cell surface. When filamin is not bound to the receptor it appears that the RRRRR motif is hindering receptor trafficking by negatively dominating over the aforementioned conserved export signal. The same research group published in 2000 a paper connecting the  $\alpha_{2C}$ -adrenoceptor to the Raynaud's Phenomenon [49], and with the identification of the C-terminal retention sequence there is a possibility in the future of targeting this cascade for drug development against this disease. Recently the same group performed in silico modeling to further investigate the interaction between filamin-2 and the RRRRR motif in the  $\alpha_{2C}$ -adrenoceptor and phylogenetic analysis of the motif [82]. The results indicate an interaction between filamin-2 and in particular arginines 454 and 456, and that these interactions have evolved in warm-blooded animals. With regard to an intracellular pool of  $\alpha_{2C}$ -adrenoceptor it remains to be seen whether this C-terminal ER-Golgi retention sequence [40] is in addition to, or excludes, the N-terminal ER retention signal proposed by Angelotti and co-workers [35].

---

### 4.3.3 The C-terminal half of the $\alpha_{2C}$ -adrenoceptor determines drug selectivity

Laurila et al. [68] proposed an interaction between TM1 and ECL2 at the  $\alpha_{2A}$ -adrenoceptor, which would induce a general affinity-lowering effect for bulky drugs. Taking advantage of our chimeric  $\alpha_{2A}/\alpha_{2C}$ -constructs we attempted to verify or disprove this hypothesis. The  $K_i$  values of seven drugs competing for [ $^3$ H]-RX821002 binding were determined at the wild type  $\alpha_{2A}$ - and  $\alpha_{2C}$ -adrenoceptors and at four of the chimeric  $\alpha_{2A}/\alpha_{2C}$ -adrenoceptors. The chimeras were CtsA, CppA, AppC, and AtsCppA (see figure 2 in paper III). The drugs spiroxatrine, clozapine, and chlorpromazine were chosen because these molecules are bulky and  $\alpha_{2C}/\alpha_{2B}$ - over  $\alpha_{2A}$ -selective [68], MK912 because it is bulky and  $\alpha_{2C}$ - but not  $\alpha_{2B}$ - over  $\alpha_{2A}$ -selective [83], BRL44408 because it is  $\alpha_{2A}$ -selective, and atipamezole and RX820102 because they are small and non-selective molecules [67]. Our results show that the  $\alpha_{2C}$ - over  $\alpha_{2A}$ -selectivity of spiroxatrine, spiperone, clozapine, MK912 and chlorpromazine, as well as the  $\alpha_{2A}$ -over  $\alpha_{2C}$ -selectivity of BRL44408, reside mainly in the C-terminal half of the receptors. This half, which is taken from  $\alpha_{2C}$  in AppC, and from  $\alpha_{2A}$  in CppA, includes TM5-7, but not most of TMs 1-4. The full binding pocket of noradrenaline is known to have direct contact sites for noradrenaline at TM3 and TM5 [16]. Thus, all the  $\alpha_{2C}$ -selective ligands, i.e. spiroxatrine, spiperone, clozapine, MK912 and chlorpromazine showed selectivity for the  $\alpha_{2C}$  and AppC receptors over the  $\alpha_{2A}$ , CtsA, CppA, and AtsCppA receptors. Laurila et al. [68] reported somewhat higher affinities for bulky  $\alpha_{2C}$ -selective drugs at CCA (a construct identical to our CtsA-construct) as compared to at  $\alpha_{2A}$ . However, in our hands only chlorpromazine, among the tested  $\alpha_{2C}$ -selective drugs, showed higher affinity for CtsA than for wild type  $\alpha_{2A}$  (2-fold). Our results cast doubt on the hypothesis that there is an  $\alpha_{2A}$ -specific conformational interaction between TM1 and ECL2 in the  $\alpha_{2A}$ -adrenoceptor that would induce a general affinity-lowering effect for all bulky substances. Instead, our results indicate that the C-terminal half of the  $\alpha_{2A}$ - respective  $\alpha_{2C}$ -adrenoceptors, independently of other regions, is the main determinant of drug selectivity between the  $\alpha_{2A}$ - and  $\alpha_{2C}$ -adrenoceptors.

## 5. CONCLUDING REMARKS

We have shown in the present project that:

i) The 5-HT<sub>2C</sub> receptor has a 32 amino acid long cleavable signal peptide. Cleavage of this signal peptide results in a mature receptor that is lacking the amino acid affected by the frequent single nucleotide polymorphism denoted Cys23Ser.

ii) The 5-HT<sub>2C</sub> receptor's expression level is decreased when cleavage of the signal peptide is obstructed experimentally by mutating an amino acid critical for the cleavability.

iii) If the signal peptide is truncated, the 5-HT<sub>2C</sub> receptor's first transmembrane segment functions efficiently as a signal anchor.

iiii) The  $\alpha_{2C}$ -adrenoceptor does not have a 22 amino acid long cleavable signal peptide, but it is unresolved whether there is an alternative cleavage site that, if functional, would generate a 15 amino acid long signal peptide.

iii) The  $\alpha_{2C}$ -adrenoceptor's N-terminal tail does not contribute to the low expression level of this receptor, and if truncated the receptor seems unable of efficiently using its first transmembrane segment as a signal anchor. The  $\alpha_{2C}$ -adrenoceptor's N-terminal is a necessity for efficient expression of binding competent receptors at the cell surface.

iiii) The C-terminal half of the  $\alpha_{2C}$ -adrenoceptor, defined as the region from the end of the fourth transmembrane helix to the receptor's ultimate C-terminus, is responsible for both: a) the low expression level of the receptor at the plasma membrane, and b) the  $\alpha_{2C}$ - over  $\alpha_{2A}$ -selectivity of "bulky" ligands.

As far as we can tell, paper I in this project is the first publication highlighting the presence of an N-terminal cleavable signal peptide in a monoamine G-protein coupled receptor. It is evident from the data presented in this project that the N-

terminus of monoamine GPCRs has a complex importance for trafficking of the receptor from the ribosome to the plasma membrane. This complexity needs to be taken into account when planning and interpreting investigations into these receptors.

## 6. FUTURE PERSPECTIVES

### 6.1 THE 5-HT<sub>2C</sub> N-TERMINUS

#### 6.1.1 Characterization of the binding epitope for the monoclonal antibody sc-15081 at the 5-HT<sub>2c</sub> receptor

In 2011, the binding epitope for a widely used "N-terminally directed" antibody recognizing the 5-HT<sub>2C</sub> receptor was disclosed, the epitope being amino acids 27-45 in the N-terminal tail [84]. We wish to prove or disprove whether this antibody, SR-2C Antibody (N-19): sc-15081 from Santa Cruz Biotechnology, recognizes a truncated 5-HT<sub>2C</sub> receptor, i.e. the receptor where the 32 amino acids long signal peptide has been cleaved off. The immunofluorescence study will be performed essentially the same way as in our previous study presented in paper I [26], examining the receptor constructs A, D, and E from that study. The hypothesis would be that the sc-15081 antibody recognizes the truncated receptor. Otherwise, 5-HT<sub>2C</sub> receptors identified in publications with this antibody must be uncleaved receptors, meaning there might exist two populations of 5-HT<sub>2C</sub> receptors in the body, namely the fraction with cleaved, and the fraction with uncleaved signal peptide.

#### 6.1.2 Investigating the 5-HT<sub>2C</sub> N-terminus for alternative signal peptide cleavage site

If the antibody described in the section above does not detect the 32/33-cleaved 5-HT<sub>2C</sub> receptor, which lacks the amino acids 27-32 present in the antigenic epitope, then additional experiments are warranted to explain why this antibody in fact has been shown to label 5-HT<sub>2C</sub> receptors in several immune-histological studies [84]. The latest version of the Signal P prediction software is version 4.1. When analyzing the 5-HT<sub>2C</sub> N-terminus in Signal P 4.1 the software predicts there to be a signal peptide with cleavage site after amino acid 22, as opposed to after amino acid 32 with version 3.0. This opens up the possibility of multiple cleavage sites, which would be interesting to investigate further for this receptor that already is known for

---

undergoing extensive RNA editing. In addition to this, cleavage after 22 amino acids would produce a mature protein with the aforementioned SNP Cys23Ser as the first amino acid. Although functional studies of the SNP concluded with it having no pharmacological relevance for the receptor [25], a mutation in the +1 position of a signal peptide could be interesting to investigate. Exchanging cysteine for serine at position 23 does not affect the prediction results in either of the Signal P versions noteworthy. The hypothesis would be that the 5-HT<sub>2C</sub> receptor is functionally expressed with different structures, meaning that the N-terminus has more than one cleavage site.

## 6.2 THE ALPHA<sub>2C</sub>-ADRENOCEPTOR N-TERMINUS

### 6.2.1 Investigation of alternative signal peptide cleavage site predicted for the $\alpha_{2C}$ -adrenoceptor

The study is motivated by the results obtained in paper II and developments in the signal peptide prediction software SignalP. In paper II we show that the predicted 22 amino acid long signal peptide is not a functional signal peptide, and that the endogenous N-terminus tail is essential for expression of binding competent receptors. When analyzing the  $\alpha_{2C}$  N-terminus in SignalP 4.1 with the setting optimized for sensitivity, the software predicts there to be a signal peptide with cleavage site after the 15th amino acid. Only one new construct has to be made and tested, in order to verify or disprove an alternative signal peptide cleavage site for the  $\alpha_{2C}$ -adrenoceptor. We will produce a new  $\alpha_{2C}$ -adrenoceptor construct, with the FLAG-epitope inserted after 15 amino acids in the N-terminal of the receptor. The construct will be tested the same way as in our previous study in paper II [34]. In addition, we will make a non-cleavable  $\alpha_{2C}$ -constructs, where alanine (A) at position 15 in the N-terminal tail “MASPALAAALAVAAAAAGPNASG-AGER...” will be exchanged for asparagine (N). The expression levels for the constructs will be tested by radioligand binding. The hypothesis would be that the  $\alpha_{2C}$ -adrenoceptor N-

---

terminus indeed has a cleavable signal peptide, although with a different cleavage site than we investigated in paper II.



---

## 7. REFERENCES

1. Pharmacology Corner. *Types of drug receptors (figure)*. <http://media.pharmacologycorner.com/wp-content/uploads/2009/01/types-of-drug-receptors.jpg> (accessed July 2014).
2. Overington, J.P., B. Al-Lazikani, and A.L. Hopkins. *How many drug targets are there?* Nat Rev Drug Discov, 2006. **5**(12): p. 993-6.
3. Cherezov, V., et al. *High-Resolution Crystal Structure of an Engineered Human  $\beta_2$ -Adrenergic G Protein-Coupled Receptor*. Science, 2007. **318**(5854): p. 1258-65.
4. Oldham, W. *Nobel in Chemistry reveals VU ties that bind: receptor-G-protein (figure)*. <http://news.vanderbilt.edu/2012/10/nobel-in-chemistry-research-reveals-vu-ties-that-bind/> (accessed July 2014).
5. Almen, M.S., et al. *Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin*. BMC Biol, 2009. **7**: 50.
6. Fredriksson, R., et al. *The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints*. Mol Pharmacol, 2003. **63**(6): p. 1256-72.
7. The IUPHAR/BPS Guide to PHARMACOLOGY. *Database content*. <http://www.guidetopharmacology.org/targets.jsp> (updated 20<sup>th</sup> June 2014).
8. Alexander, S.P., et al. *The Concise Guide to PHARMACOLOGY 2013/14: G protein-coupled receptors*. Br J Pharmacol, 2013. **170**(8): p. 1459-581.
9. Jacoby, E., et al. *The 7 TM G-protein-coupled receptor target family*. ChemMedChem, 2006. **1**(8): p. 761-82.
10. The IUPHAR/BPS Guide to PHARMACOLOGY. *Introduction to G Protein-Coupled Receptors*. <http://guidetopharmacology.org/pdfs/GPCRs.pdf> (accessed July 2014).
11. Palczewski, K., et al. *Crystal structure of rhodopsin: A G protein-coupled receptor*. Science, 2000. **289**(5480): p. 739-45.
12. Shepard, B.D., et al. *A cleavable N-terminal signal peptide promotes widespread olfactory receptor surface expression in HEK293T cells*. PLoS One, 2013. **8**(7): e68758.
13. Bockaert, J. and J.P. Pin. *Molecular tinkering of G protein-coupled receptors: an evolutionary success*. EMBO J, 1999. **18**(7): p. 1723-9.
14. Rasmussen, S.G.F., et al. *Crystal structure of the human  $\beta_2$  adrenergic G-protein-coupled receptor*. Nature, 2007. **450**(7168): p. 383-7.
15. Ring, A.M., et al. *Adrenaline-activated structure of  $\beta_2$ -adrenoceptor stabilized by an engineered nanobody*. Nature, 2013. **502**(7472): p. 575-9.
16. Weichert, D., et al. *Covalent agonists for studying G protein-coupled receptor activation*. Proc Natl Acad Sci U S A, 2014. **111**(29): p. 10744-8.
17. Hannon, J. and D. Hoyer. *Molecular biology of 5-HT receptors*. Behav Brain Res, 2008. **195**(1): p. 198-213.
18. Barnes NM, R. Andrade, J. Bockaert, et al. *5-Hydroxytryptamine receptors: Introduction*. IUPHAR/BPS Guide to PHARMACOLOGY. <http://guidetopharmacology.org/GRAC/FamilyIntroductionForward?familyId=1> (accessed August 2014).

19. Higgins, G.A., E.M. Sellers, and P.J. Fletcher. *From obesity to substance abuse: therapeutic opportunities for 5-HT<sub>2C</sub> receptor agonists*. Trends Pharmacol Sci, 2013. **34**(10): p. 560-70.
20. Price, R.D., et al. *RNA editing of the human serotonin 5-HT<sub>2C</sub> receptor alters receptor-mediated activation of G<sub>13</sub> protein*. J Biol Chem, 2001. **276**(48): p. 44663-8.
21. Decher, N., M.F. Netter, and A.K. Streit. *Putative impact of RNA editing on drug discovery*. Chem Biol Drug Des, 2013. **81**(1): p. 13-21.
22. Drago, A. and A. Serretti. *Focus on HTR2C: A Possible Suggestion for Genetic Studies of Complex Disorders*. Am J Med Genet B Neuropsychiatr Genet, 2009. **150b**(5): p. 601-37.
23. Gunes, A., et al. *Further evidence for the association between 5-HT<sub>2C</sub> receptor gene polymorphisms and extrapyramidal side effects in male schizophrenic patients*. Eur J Clin Pharmacol, 2008. **64**(5): p. 477-82.
24. Backstrom, J.R., et al. *Identification of rat serotonin 5-HT<sub>2C</sub> receptors as glycoproteins containing N-linked oligosaccharides*. Brain Res Mol Brain Res, 1995. **33**(2): p. 311-8.
25. Fentress, H.M., et al. *Pharmacological properties of the Cys23Ser single nucleotide polymorphism in human 5-HT<sub>2C</sub> receptor isoforms*. Pharmacogenomics J, 2005. **5**(4): p. 244-54.
26. Jahnsen, J.A. and S. Uhlen. *The N-terminal region of the human 5-HT<sub>2C</sub> receptor has as a cleavable signal peptide*. Eur J Pharmacol, 2012. **684**(1-3): p. 44-50.
27. Janssen, P.K., et al. *The 5-HT<sub>2C</sub> receptor gene Cys23Ser polymorphism influences the intravaginal ejaculation latency time in Dutch Caucasian men with lifelong premature ejaculation*. Asian J Androl, 2014. **16**(4): p. 607-10.
28. Anastasio, N.C., et al. *Variation within the serotonin (5-HT) 5-HT<sub>2C</sub> receptor system aligns with vulnerability to cocaine cue reactivity*. Transl Psychiatry, 2014. **4**: e369.
29. European Medicines Agency. *Withdrawal of the marketing authorisation application for Belviq (lorcaserin)*. [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Medicine\\_QA/2013/05/WC500143811.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Medicine_QA/2013/05/WC500143811.pdf) (accessed August 2014).
30. Tecott, L.H., et al. *Eating disorder and epilepsy in mice lacking 5-HT<sub>2C</sub> serotonin receptors*. Nature, 1995. **374**(6522): p. 542-6.
31. Rauser, L., et al. *Inverse Agonist Actions of Typical and Atypical Antipsychotic Drugs at the Human 5-Hydroxytryptamine<sub>2c</sub> Receptor*. J Pharmacol Exp Ther, 2001. **299**(1): p. 83-9.
32. European Medicines Agency. *Valdoxan : EPAR - Product Information* (last updated June 2014). <http://www.ema.europa.eu>.
33. Guaiana, G., et al. *Agomelatine versus other antidepressive agents for major depression*. Cochrane Database Syst Rev, 2013. **12**: Cd008851.
34. Jahnsen, J.A. and S. Uhlén, *The predicted N-terminal signal sequence of the human  $\alpha_{2C}$ -adrenoceptor does not act as a functional cleavable signal peptide*. Eur J Pharmacol, 2012. **684**(1-3): p. 51-8.
35. Angelotti, T., et al. *Regulation of G-protein coupled receptor traffic by an evolutionary conserved hydrophobic signal*. Traffic, 2010. **11**(4): p. 560-78.
36. Wozniak, M. and L.E. Limbird. *The Three  $\alpha_2$ -Adrenergic Receptor Subtypes Achieve Basolateral Localization in Madin-Darby Canine Kidney II Cells via Different Targeting Mechanisms*. J Biol Chem, 1996. **271**(9): p. 5017-24.
37. Jeyaraj, S.C., et al. *Cooling Evokes Redistribution of  $\alpha_{2C}$ -Adrenoceptors from Golgi to Plasma Membrane in Transfected Human Embryonic Kidney 293 Cells*. Mol Pharmacol, 2001. **60**(6): p. 1195-200.

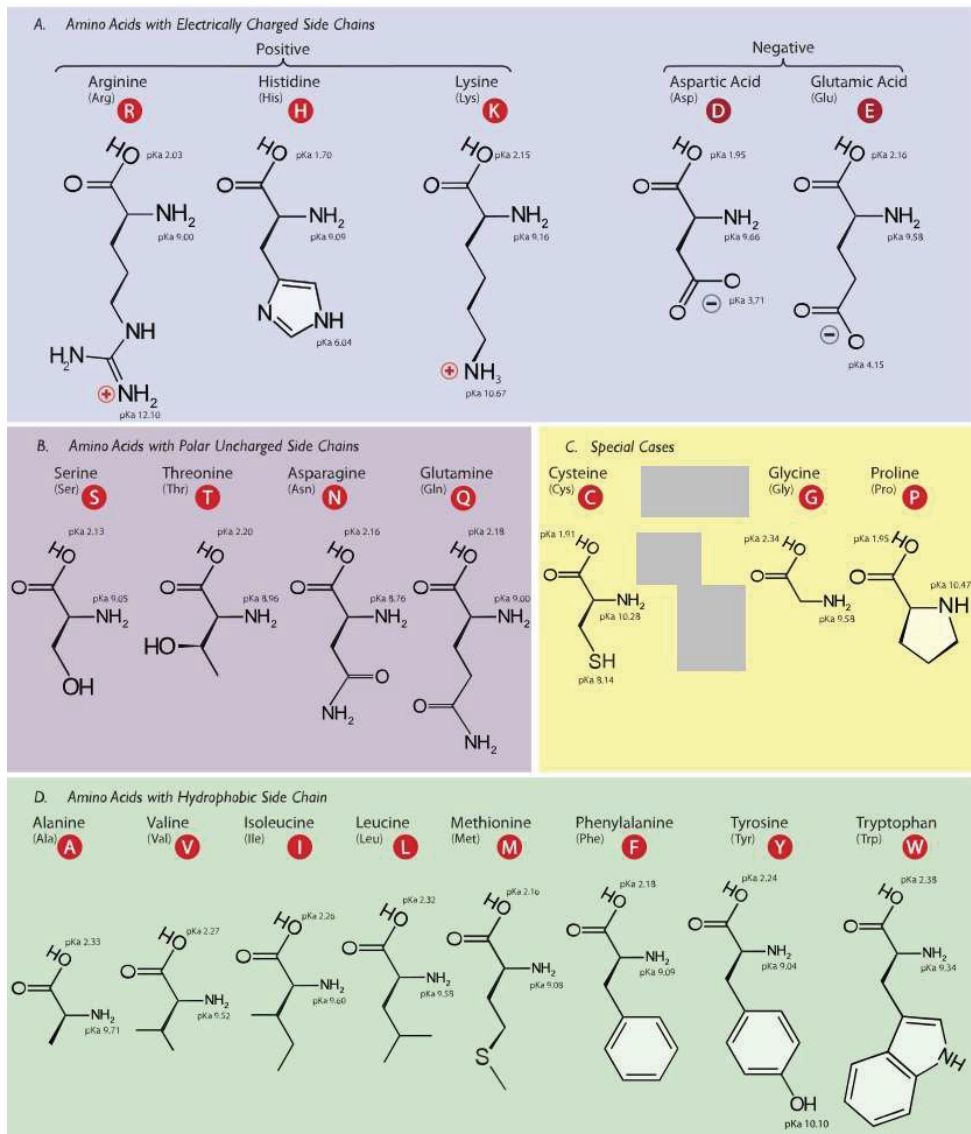
38. Chotani, M.A. and N.A. Flavahan. *Intracellular  $\alpha_{2C}$ -adrenoceptors: storage depot, stunted development or signaling domain?* Biochim Biophys Acta, 2011. **1813**(8): p. 1495-503.
39. Hurt, C.M., F.Y. Feng, and B. Kobilka. *Cell-type Specific Targeting of the  $\alpha_{2C}$ -Adrenoceptor. Evidence for the organization of receptor microdomains during neuronal differentiation of PC12 cells.* J Biol Chem, 2000. **275**(45): p. 35424-31.
40. Motaweia, H.K., et al. *Cyclic AMP-Rap1A signaling mediates cell surface translocation of microvascular smooth muscle  $\alpha_{2C}$ -adrenoceptors through the actin-binding protein filamin-2.* Am J Physiol Cell Physiol, 2013. **305**(8): C829-45.
41. Trendelenburg, A.U., et al. *All three  $\alpha_2$ -adrenoceptor types serve as autoreceptors in postganglionic sympathetic neurons.* Naunyn Schmiedeberg's Arch Pharmacol, 2003. **368**(6): p. 504-12.
42. Docherty, J.R. *Subtypes of functional  $\alpha_1$ - and  $\alpha_2$ -adrenoceptors.* Eur J Pharmacol, 1998. **361**(1): p. 1-15.
43. Gyires, K., et al.  *$\alpha_2$ -Adrenoceptor subtypes-mediated physiological, pharmacological actions.* Neurochem Int, 2009. **55**(7): p. 447-53.
44. Uhlen, S., et al. *Autoradiographic studies of central  $\alpha_{2A}$ - and  $\alpha_{2C}$ -adrenoceptors in the rat using [ $^3$ H]MK912 and subtype-selective drugs.* Brain Res, 1997. **770**(1-2): p. 261-6.
45. Philipp, M. and L. Hein. *Adrenergic receptor knockout mice: distinct functions of 9 receptor subtypes.* Pharmacol Ther, 2004. **101**(1): p. 65-74.
46. Gilsbach, R. and L. Hein. *Are the pharmacology and physiology of  $\alpha_2$  adrenoceptors determined by  $\alpha_2$ -heteroreceptors and autoreceptors respectively?* Br J Pharmacol, 2012. **165**(1): p. 90-102.
47. Quaglia, W., et al.  *$\alpha_{2C}$ -adrenoceptor modulators: a patent review.* Expert Opin Ther Pat, 2011. **21**(4): p. 455-81.
48. Jia, Y., et al. *Pharmacological evaluation of selective  $\alpha_{2C}$ -adrenergic agonists in experimental animal models of nasal congestion.* J Pharmacol Exp Ther, 2014. **349**(1): p. 75-84.
49. Chotani, M.A., et al. *Silent  $\alpha_{2C}$ -adrenergic receptors enable cold-induced vasoconstriction in cutaneous arteries.* Am J Physiol Heart Circ Physiol, 2000. **278**(4): H1075-83.
50. Osborne, A.R., T.A. Rapoport, and B. van den Berg. *Protein translocation by the Sec61/SecY channel.* Annu Rev Cell Dev Biol, 2005. **21**: p. 529-50.
51. Schulein, R., et al. *Functional significance of cleavable signal peptides of G protein-coupled receptors.* Eur J Cell Biol, 2012. **91**(4): p. 294-9.
52. Tuteja, R. *Type I signal peptidase: an overview.* Arch Biochem Biophys, 2005. **441**(2): p. 107-11.
53. Wallin, E. and G. von Heijne. *Properties of N-terminal tails in G-protein coupled receptors: a statistical study.* Protein Eng, 1995. **8**(7): p. 693-8.
54. Jensen, D. and R. Schekman. *COPII-mediated vesicle formation at a glance.* J Cell Sci, 2011. **124**(Pt 1): p. 1-4.
55. Wu, G. *Regulation of post-Golgi traffic of G protein-coupled receptors.* Subcell Biochem, 2012. **63**: p. 83-95.
56. Gavel, Y. and G. von Heijne. *Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering.* Protein Eng, 1990. **3**(5): p. 433-42.
57. von Heijne, G. *The signal peptide.* J Membr Biol, 1990. **115**(3): p. 195-201.
58. von Heijne, G. *A new method for predicting signal sequence cleavage sites.* Nucleic Acids Res, 1986. **14**(11): p. 4683-90.

59. Emanuelsson, O., et al. *Locating proteins in the cell using TargetP, SignalP and related tools*. Nat Protoc, 2007. **2**(4): p. 953-71.
60. Nielsen, H., et al. *Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites*. Protein Eng, 1997. **10**(1): p. 1-6.
61. Bendtsen, J.D., et al. *Improved prediction of signal peptides: SignalP 3.0*. J Mol Biol, 2004. **340**(4): p. 783-95.
62. Nielsen, H. and A. Krogh. *Prediction of signal peptides and signal anchors by a hidden Markov model*. Proc Int Conf Intell Syst Mol Biol, 1998. **6**: p. 122-30.
63. Couvineau, A., C. Rouyer-Fessard, and M. Laburthe. *Presence of a N-terminal signal peptide in class II G protein-coupled receptors: crucial role for expression of the human VPAC1 receptor*. Regul Pept, 2004. **123**(1-3): p. 181-5.
64. Mitri, F., et al. *A novel compound heterozygous mutation of the luteinizing hormone receptor -implications for fertility*. J Assist Reprod Genet, 2014. **31**(7): p. 787-94.
65. Petrovska, R., et al. *Addition of a signal peptide sequence to the  $\alpha_{1D}$ -adrenoceptor gene increases the density of receptors, as determined by [ $^3$ H]-prazosin binding in the membranes*. Br J Pharmacol, 2005. **144**(5): p. 651-9.
66. Guan, X.M., T.S. Kobilka, and B.K. Kobilka. *Enhancement of membrane insertion and function in a type IIIb membrane protein following introduction of a cleavable signal peptide*. J Biol Chem, 1992. **267**(31): p. 21995-8.
67. Jahnsen, J.A. and S. Uhlen. *The C-terminal half of the  $\alpha_{2C}$ -adrenoceptor determines the receptor's membrane expression level and drug selectivity*. Naunyn Schmiedebergs Arch Pharmacol, 2013. **386**(12): p. 1031-40.
68. Laurila, J.M., et al. *Involvement of the first transmembrane segment of human  $\alpha_2$ -adrenoceptors in the subtype-selective binding of chlorpromazine, spiperone and spiroxatrine*. Br J Pharmacol, 2011. **164**(5): p. 1558-72.
69. Parham, J.H., T. Kost, and J.T. Hutchins. *Effects of pCIneo and pCEP4 expression vectors on transient and stable protein production in human and simian cell lines*. Cytotechnology, 2001. **35**(3): p. 181-7.
70. Einhauer, A. and A. Jungbauer. *The FLAG peptide, a versatile fusion tag for the purification of recombinant proteins*. J Biochem Biophys Methods, 2001. **49**(1-3): p. 455-65.
71. Kallal, L. and J.L. Benovic. *Fluorescence microscopy techniques for the study of G protein-coupled receptor trafficking*. Methods Enzymol, 2002. **343**: p. 492-506.
72. Jou, W.M., et al. *Complete structure of the hemagglutinin gene from the human influenza A/Victoria/3/75 (H3N2) strain as determined from cloned DNA*. Cell, 1980. **19**(3): p. 683-96.
73. Andersson, H., et al. *Membrane assembly of the cannabinoid receptor 1: impact of a long N-terminal tail*. Mol Pharmacol, 2003. **64**(3): p. 570-7.
74. Mancia, F., et al. *Ligand sensitivity in dimeric associations of the serotonin 5HT<sub>2C</sub> receptor*. EMBO Rep, 2008. **9**(4): p. 363-9.
75. Abramowski, D. and M. Staufenbiel. *Identification of the 5-hydroxytryptamine<sub>2C</sub> receptor as a 60-kDa N-glycosylated protein in choroid plexus and hippocampus*. J Neurochem, 1995. **65**(2): p. 782-90.
76. Petersen, T.N., et al. *SignalP 4.0: discriminating signal peptides from transmembrane regions*. Nat Methods, 2011. **8**(10): p. 785-6.
77. Hurt, C.M. and T. Angelotti. *Expression and trafficking of functional G protein-coupled receptors are related, yet distinct, concepts*. Naunyn Schmiedebergs Arch Pharmacol, 2014. **387**(10): p. 1009-12.
78. Uhlen, S. and J.A. Jahnsen. *Perspective on the role of the N-terminal tail of the  $\alpha_{2C}$ -adrenoceptor*. Naunyn Schmiedebergs Arch Pharmacol, 2014. **387**(10): p. 1013-4.

- 
79. von Heijne, G. *The membrane protein universe: what's out there and why bother?* J Intern Med, 2007. **261**(6): p. 543-57.
  80. Dong, C., et al. *Regulation of G protein-coupled receptor export trafficking.* Biochim Biophys Acta, 2007. **1768**(4): p. 853-70.
  81. Bermak, J.C., et al. *Regulation of transport of the dopamine D1 receptor by a new membrane-associated ER protein.* Nat Cell Biol, 2001. **3**(5): p. 492-8.
  82. Pawlowski, M., et al. *In silico modeling of human  $\alpha_{2C}$ -adrenoreceptor interaction with filamin-2.* PLoS One, 2014. **9**(8): e103099.
  83. Uhlen, S., et al. *[<sup>3</sup>H]RS79948-197 binding to human, rat, guinea pig and pig  $\alpha_{2A}$ ,  $\alpha_{2B}$ - and  $\alpha_{2C}$ -adrenoceptors. Comparison with MK912, RX821002, rauwolscine and yohimbine.* Eur J Pharmacol, 1998. **343**(1): p. 93-101.
  84. Bubar, M.J., S.J. Stutz, and K.A. Cunningham, *5-HT<sub>2C</sub> receptors localize to dopamine and GABA neurons in the rat mesoaccumbens pathway.* PLoS One, 2011. **6**(6): e20508.

## 8. APPENDIX

An overview of the 20 amino acids with abbreviations and structure.



The figure is a modified version of a figure made by Dan Cojocari (Department of Medical Biophysics, University of Toronto, Canada) published in the Wikimedia Commons database.