CHALMERS
UNIVERSITY OF TECHNOLOGY

# Optimization of HCAR1 and binding partners for cryo-EM studies

Master's thesis in Biotechnology

LUCAS VON BRÖMSEN

DEPARTMENT OF BIOLOGY AND BIOLOGICAL ENGINEERING
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2022
www.chalmers.se

Master's thesis 2022

# Optimization of HCAR1 and binding partners for cryo-EM studies

LUCAS VON BRÖMSEN


CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Biology and Biological Engineering
*Division of Industrial Biotechnology*
Chalmers University of Technology
Gothenburg, Sweden 2022

Optimization of HCAR1 and binding partners for cryo-EM studies
LUCAS VON BRÖMSEN



Supervisor: Linda Johansson, Department of Medicinal Biochemistry and Cell Biology, Gothenburg University
Examiner: Johan Larsbrink, Department of Biology and Biological Engineering, Chalmers University of Technology

Master's Thesis 2022
Department of Biology and Biological Engineering
Division of Industrial Biotechnology
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2022

Optimization of HCAR1 and binding partners for cryo-EM studies
LUCAS VON BRÖMSEN
Department of Biology and Biological Engineering
Chalmers University of Technology

# Abstract

The G protein-coupled receptor HCAR1 has recently been linked to several different types of cancer, and silencing of this gene has led to decreased cancer cell growth and proliferation. In a clinical setting and structure-based drug design, the structure of this protein would be of immense importance for drug development; no experimental structures of HCAR1 have been determined however. This study aimed to investigate a novel method, the Nb6 method, that utilizes a nanobody for obtaining the inactive-state structure of the receptor. In parallel, a conventional experiment for obtaining the active-state structure of the receptor by using a G protein was performed. The results concluded that there was a major difference in expression and stability of HCAR1 depending on what host system was used. Additionally, a stable and pure form of HCAR1 was obtained in large quantities by using the BacMam system. The nanobody was also purified in large quantities, the purity was however of questionable quality. Complex formation of HCAR1 and the nanobody was evaluated but it was determined that optimization of all processes needed further improvement and no definitive conclusion as to if there was an interaction between HCAR1 and the nanobody could be deduced. HCAR1 in combination with G protein for obtaining the active-state structure could not be obtained in a stable form, though this was likely due to the host system this protein was expressed in. Several implementations are suggested as to how to improve purification and stability of the proteins which would result in stable proteins suitable for structural determination with cryo-EM.

Keywords: HCAR1, GPR81, lactate receptor, Nb6, BacMam, GPCR.

# Acknowledgements

Firstly, I would like to thank my supervisor Linda Johansson at the University of Gothenburg for whom without I would never have had the chance to do this interesting project that have taught me a great deal. Not only did I get the possibility to delve deeper into the field of structural biology, a passion of mine, but also learn more about science as a whole and combine youthful curiosity and awe with matured ingenuity and discipline. Thank you for the possibility to do this project in your research group and the lessons I have learnt.

Furthermore, I would like to express my deepest gratitude to the members of the Johansson research group I had the pleasure to work with, Petra Båth, Cecilia Safari and Daniel Hedberg. Many thanks to you for the great support and helping me with many aspects of the project as well as making me feel like home in your group.

Lastly, I would also like to thank my family and friends who have always been there for me and helped me along the journey in any was possible and given me the strength to continue no matter the circumstances.

Lucas von Brömsen, Gothenburg, June 2022

# List of Acronyms

Below is the list of acronyms that have been used throughout this thesis listed in alphabetical order:

| | |
|---|---|
| Cryo-EM | Cryogenic Electron Microscopy |
| GPCR | G Protein-coupled Receptor |
| H2R | Histamine receptor 2 |
| HCAR1 | Hydroxycarboxylic Acid Receptor 1 |
| HPLC | High-performance liquid chromatography |
| ICL3 | Intracellular Loop 3 |
| KOR | $\kappa$-opioid Receptor |
| MOR | $\mu$-opiod Receptor |
| Nb6 | Nanobody 6 |
| NTSR1 | Neurotensin Receptor 1 |
| SDS-PAGE | Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis |
| SSTR2 | Somatostatin Receptor 2 |
| WB | Western blot |

# Contents

# 1
# Introduction

One of the greatest challenges in modern healthcare is cancer. It is the second most common cause of death after cardiovascular diseases, with about one of every six deaths in the world being cause by some type of cancer [1]. During the last decade however, there have been great developments in cancer therapy and reducing the mortality rate of cancer. For example, total cancer death rate went down from 200.8 to 146.2 deaths per 100 000 population in the United States between 1999 to 2019, a 27 % decrease in overall cancer mortality rate [2]. There is however a large discrepancy between different types of cancers with some attributing more to reduce overall cancer mortality rate. Lung cancer and breast cancer, the two most common types of cancers, have had their death rates reduced from 58.9 to 33.4 and 31.6 to 19.4 per 100 000 population respectively in the United States between 1992 to 2019 [3][4]. Other less common cancer types though, such as liver cancer and pancreatic cancer, have not seen similar improvements to death rate with death rates increasing from 3.9 to 6.6 and 10.7 to 11.1 per 100 000 population respectively in the US during the same period [5][6]. The reason for the differences in attempting to improve mortality rate over time may be attributed to several different causes, such as discrepancies in funding of research, screening efforts, and inherent differences in symptoms and severity of particular cancer types. The bottom line however is that cancer still remains a severe issue in society and many additional implementations and innovations are needed to further decelerate cancer death rate for all types of cancers regardless if they have seen improvements to death rate or not.

There are various viable options that can be targeted and improved for reducing cancer mortality rate, such as screening, biomarker testing, surgery, chemotherapy, targeted therapy etc. [7]. In common for many of these alternatives is that they utilize chemicals, drugs or biomolecules for treating or identifying cancer. The target for these molecules are often elements that are unique for cancer cells, e.g. specific genes or proteins involved in cancer growth and proliferation. Hence, when designing these molecules that target cancer cells, structural and functional information regarding both the drug and the target are of utmost importance and can alleviate efforts to quickly and effectively generate potential cancer treatments through a structure-based drug design.

One protein that has recently been observed to have a part in various cancer types is HCAR1. It is a membrane protein that has been found to be upregulated in various cancers including breast, cervical, colon, pancreatic and lung [8]. Recent studies have shown the importance of HCAR1 for cancer growth and proliferation, and si-

lencing of this gene has led to death of these cancer cells, supporting the potential for this protein as a possible therapeutic target for several different types of cancers [9]. The structure for this protein is undiscovered and an experimental structure would immensely aid in a clinical setting and a structure-based drug design.

## 1.1 Aims

The aim of the thesis was to attempt to elucidate the inactive-state structure of the protein HCAR1 using a novel method referenced as the Nb6 method that has been used for obtaining inactive-state structures of G-protein coupled receptors (GPCRs). Additionally, it was also an aim to attempt to elucidate the active-state structure of the protein HCAR1 using a conventional method with a G protein. Lastly, an aim of the project was to optimize the overall purification process for HCAR1 and related constructs to ensure proteins were sufficiently stable and present in large quantities such that they would be suitable for cryogenic electron microscopy (cryo-EM) studies .

# 2
# Theory

## 2.1 G protein-coupled receptors

G protein-coupled receptors, also known as GPCRs, are one of the most important and clinically relevant protein families. This is signified by the fact that they are either involved in or are the target of approximately 34% all FDA approved drugs, while only representing $\sim$ 4% of all human genes [10]. They constitute a majority of the membrane proteins and have several different functions. The most prominent function of GPCRs is to act as a mediator to cellular responses caused by hormones and neurotransmitters. Therefore a wide variety of GPCRs are needed to be able to transmit signals for many different molecules. Despite this, GPCRs generally contain conserved motifs and are very similar to each other from a three dimensional structure point of view [11]. Common for GPCRs is that they all have seven transmembrane regions that are $\alpha$-helical and between these segments are alternating intracellular or extracellular loops. Because of this, they are sometimes referred to as 7TM receptors, and the different segments can be communicated as numerical regions that are shared between all GPCRs, e.g. TM5 or ICL3 to describe the fifth transmembrane region or the third intracellular loop respectively [12]. GPCRs are also often classified depending on their structural features and protein sequence, and can belong to family A, family B, family C, adhesion or Frizzled/Taste 2. Proteins from the same family are very structurally similar but also functionally akin and are often activated by related pathways or mechanisms. The greatest differences between GPCRs are in the loop regions and the ligand-binding site, while transmembrane regions are often similar and share conserved features, especially when belonging to the same subfamily [12].

GPCRs are often thought of as bimodal switches, but can in reality exist in several different conformations and have complex regulatory pathways. The primary state of a GPCR is the inactive state. This is the conformation of the GPCR when it is not bound to any agonist and thus is in a position in which it cannot perform its intended function to transmit signals through coupling with G proteins. There does however exist a basal agonist-independent activity where activation of the GPCR and its G protein can occur without agonist binding [13]. Activation of the GPCR and its corresponding G protein is though most often obtained by an agonist binding to the GPCR which changes the conformation of the GPCR, making it able to couple with a G protein and resulting in a downstream signaling cascade. This conformation where the GPCR actuates a signal transduction pathway is the active

state of the GPCR [14]. Similarly to the inactive state which could be activated without agonist binding, the active state can also be attained through G-protein-independent pathways in which the G-protein is not necessary for activation [15]. Additionally, it has also been observed that GPCRs can form intermediate state conformations between the process of inactive state to active state, which can be stable. Evidence suggests that some GPCRs may have many different conformational intermediates that can be functionally active, and hence there may be many different native conformations that can be responsible for a functionally active state of the GPCR [14][15]. The inactive and active state are thus regulated in various ways that are still poorly understood with the potential for many functionally active native conformations as well.

## 2.2 HCAR1

Hydroxycarboxylic Acid Receptor 1, also known as GPR81 or HCAR1, is a G-protein coupled receptor belonging to family A of GPCRs [16]. It is the receptor for lactate in humans and has several different functions. It is upregulated in adipocytes, where its activation induces the inhibition of lipolysis via G protein dependent activation [17]. Additionally, recent studies have indicated that HCAR1 is expressed in neurons where it presynaptically downmodulates neuronal activity through activation of $G_i\alpha$-protein-dependent pathways [18][19]. A very prominent role described for HCAR1, however, is in various cancers including colon, breast, lung, pancreatic and cervical, where it is found to be upregulated [8]. Several studies have confirmed that silencing or down-regulation of HCAR1 leads to decreased tumor cell proliferation, thus supporting a role of HCAR1 in tumor microenvironment metabolism [9][20].

The endogenous ligand l-lactate is an agonist [21], thus displaying G protein dependent activation within the cells and very few small molecule ligands have been synthesized for this receptor. To the author's knowledge, there are currently no inactivating – and thus downregulating – ligands of HCAR1 that have been reported in publications. These inactivating ligands would be of major advantage in a clinical setting, since inactivation of HCAR1 in tumor cells may aid in decreasing tumor cell proliferation. For efficient development of inactivating ligands, the inactive structure of HCAR1 would be an essential tool in this endeavour, though the active state structure is also a major target. There is currently however no experimental structure of either the active or inactive structure of HCAR1. Only a predicted structure of HCAR1 has been constructed using AlphaFold version 2.0, which can be observed in Fig 2.1.

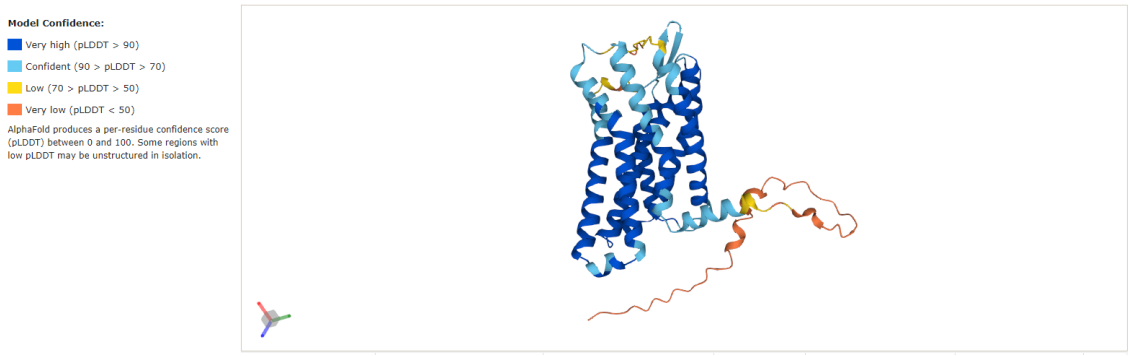



**Figure 2.1:** *Predicted structure of HCAR1 constructed from AlphaFold version 2.0 [22].*

## 2.3 Structural determination of GPCRs

For structural determination of GPCRs, X-ray crystallography has generally been the preferred method of choice. The first GPCR was structurally determined in 2000, and up until 2017 all 188 structures of GPCRs that had been deposited in the Protein Data Bank (PDB) had been obtained using X-ray crystallography [23]. But in 2017 the first structure of a GPCR using cryogenic electron microscopy (cryo-EM) for structural determination was attained [24]. This heralded a new era of structural determination of GPCRs, and rapid advancements in cryo-EM led to a shift with cryo-EM being the dominant technique. Between Jan-Jul 2021, 99 GPCR structures were deposited in the PDB and 78% of these were resolved with cryo-EM [26].

The reason that X-ray crystallography was the conventional method up until 2017 was that cryo-EM was simply too unrefined of a technique and did not result in sufficiently high resolutions for the structures. Technical limitations was the impediment of technique, but advancements in mainly single-particle cryo-EM transformed the method to a tractable choice for GPCRs [27][28]. Today cryo-EM is more favored than X-ray crystallography for structural determination of GPCRs and the trend is only increasing. There are several reasons to why cryo-EM is preferred in this field. Firstly, sample preparation is much easier for cryo-EM than X-ray crystallography. After purification of the protein, X-ray crystallography requires the protein to be crystallized, a process that takes several days or weeks not accounting for optimizing crystallization conditions and iteratively improving the construct [25]. In contrast, cryo-EM sample preparation is performed in minutes and additionally does not require as stable of a construct as in X-ray crystallography [26]. Furthermore, in X-ray crystallography, well-diffracting crystals are required for structural determination. This often results in heavy engineering efforts including removing large regions, doing mutations and removing post-translational modifications, all in an attempt to optimize the crystallization process [29]. This process can take several years. Cryo-EM in comparison does not have this necessity and can be conducted on proteins with unstructured regions and large post-translational modifications, all while requiring less sample as well as being of lower quality and a less homogeneous sample [26]. Lastly, a significant advantage of cryo-EM over X-ray crystallography

for structural determination of GPCRs is the ability of cryo-EM to capture conformational dynamics. When vitrifying purified proteins on grids in cryo-EM, the protein retains its conformational landscape and it is possible to obtain important structural information and dynamics of the receptor from different discrete conformations in the sample originating from the same data set [26]. This is especially crucial for GPCRs which are characterized by their transient states and large dynamic flexibility [30]. It is also possible to analyse or predict continuous movements and motions of receptors in cryo-EM data sets using novel computational methods [31]. X-ray crystallography lacks many of these attributes due to the necessity of crystal formation and consequently lattice contacts potentially imposing structural or dynamical constraints on the receptor. This results in difficulties studying flexibility and conformational dynamics by X-ray crystallography, even if it is possible to some extent [32].

There are however some distinct advantages using X-ray crystallography in structure determination of GPCRs. The greatest advantage is the possibility to obtain several different structures of the same receptor quickly, which can be valuable in structure-based drug design [33]. When you have obtained a crystal with the receptor bound to a ligand and derived a structure from this, it is possible to grow crystals with the protein bound to different ligands in similar conditions with a high probability of success. Using molecular replacement, it is feasible to determine multiple high-resolution structures very quickly [34]. This process may require hours for X-ray crystallography, while in cryo-EM every structure has to be independently determined which may take several days or weeks [26]. Another advantage is that while cryo-EM is preferred for active state structures where the receptor is bound to G-protein or arrestins, X-ray crystallography currently has a slight advantage for inactive state structures of GPCRs. One big reason for this is that inactive state structures are usually too small for cryo-EM and X-ray crystallography may resolve them better [35]. That may however soon change with the development of a new technique to determine the structure of inactive state GPCRs with cryo-EM, the Nb6 method.

## 2.4 The Nb6 method

Recently, a novel method has been developed for structural determination of GPCRs in the inactive state by cryo-EM, which will be referred to as the Nb6 method. The Nb6 method utilizes a camelid single-chain antibody, also known as nanobody [36], for stabilization of the complex. The nanobody used in this method is known as Nanobody 6, or Nb6. The purpose of the nanobody is to stabilize the complex but also to act as a fiducial marker for projection alignment when collecting data [37]. When constructing electron densities map in cryo-EM, many projections from different angles are required to generate a 3D structure [38]. Large masses or electron densities assist in this aim by having recognizable features. This is also a reason why cryo-EM is limited by the size of the protein and have difficulties in initial alignment at low resolutions for low mass proteins [40]. Generally, the solution to this problem is adding binding partners such as G protein or $\beta$-arrestin to obtain a higher mass

and overcome this issue, but this often results in the active state structure which may not always be desirable [26]. However, some nanobodies have been observed to bind rigidly to specific GPCRs and lock the protein in a certain conformation [41]. Despite the small size of nanobodies (∼12 to 15 kDa), some of them have features that strongly stabilizes the complex and sufficiently recognizable features to make them suitable fiducial markers [42].

One nanobody that has been observed to have a strong interaction with a specific GPCR is Nb6. It binds to the $\kappa$-opioid receptor (KOR) in the third intracellular loop, ICL3, of the receptor [42]. A rigid, high-affinity bond is formed that locks the receptor in an inactive state by creating several hydrogen bonds and potentiating a TM5-TM6 interaction that stabilizes the inactive state of the receptor. The way in which Nb6 binds to ICL3 also results in a certain conformation where ICL3 is locked and allosterically blocking G protein-coupling, ensuring that the inactive state is maintained [42]. What is remarkable though about this interaction is that it has been observed that even after transferring the Nb6-binding sequence of KOR to another GPCR, the high-affinity interaction is retained [37]. Since Nb6 is a suitable fiducial marker that can be used for cryo-EM, this permits structural determination of a multitude of different GPCRs and can hugely facilitate structural determination of inactive state GPCRs with this technique. In fact, the same researchers that discovered the interaction, have successfully determined the inactive state structure of four GPCRs from family A: NTSR1, MOR, H2R and SSTR2 [37]. SSTR2 was additionally in contrast to the others previously uncharacterized and in the unliganded (apo) state, and all structures were resolved at a high resolution. For MOR, two point mutations sufficed to enable Nb6 binding due to a close homology with KOR, while SSTR2 required changing 15 residues, NTSR1 switching 31 residues and H2R switching 42 residues with that of the KOR ICL3 region and TM5/TM6 regions [37]. Optimally, as few residues as possible are changed to be certain that the original structure and function from the receptor you are studying are retained. Depending on if there is little homology of your receptor with regions of ICL3, TM5 and TM6 of KOR, it may be necessary to change many residues for a stronger interaction to Nb6. The residues that were changed by the researchers for NTSR1, MOR, H2R and SSTR2 compared to HCAR1 in this study can be observed in Fig 2.2.

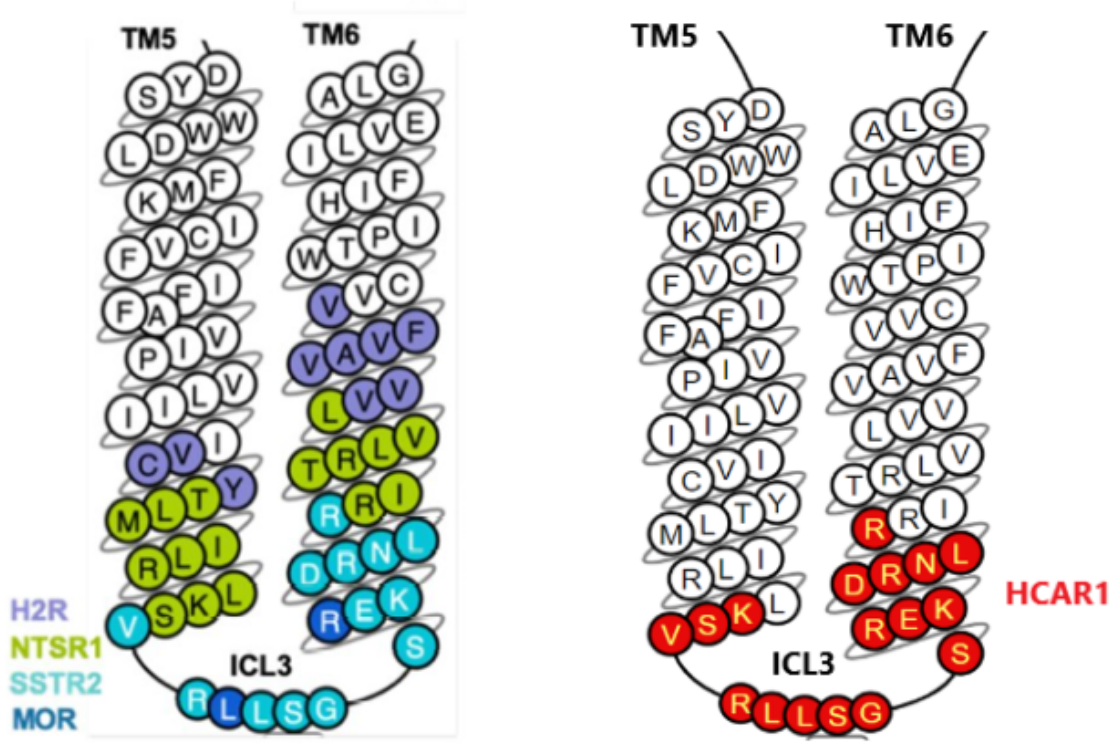



**Figure 2.2:** *Residues from KOR transferred to the respective protein from NTSR1, MOR, H2R and SSTR2 (left) [39] compared to HCAR1 (right).*

## 2.5 Baculovirus expression system

For study and structure determination of proteins, large quantities are desirable. Preferably, production of mammalian proteins, which are often the target in academia and industry, would be produced naturally in mammalian cells for protein production since this ensures proper folding and post-translational modifications. For many different proteins, mammalian cells are therefore the apparent choice [43]. For GPCR expression however, mammalian cells have been observed to produce relatively small quantities [44]. Although there have been some improvements in increasing production [45], it is generally much more expensive and time-consuming compared to some other host systems [44]. Instead, one host system that has been observed to be especially good at expressing GPCRs while also being able to do most post-translational modifications required for mammalian proteins are insect cells. These have been found to be able to produce up to 25-600 times more receptors per cell compared to naturally producing mammalian cells when using the baculovirus expression system, a system designed to increase recombinant protein expression [46]. The baculovirus expression system can also be applied to mammalian cells, called BacMam. However, this method using mammalian cells is not as established as for insect cells and is also not as attractive since mammalian cells are more expensive to culture and can generally not grow in as high densities as insect cells, leading to lower protein production [47]. Thus, the baculovirus expression system using insect cells is generally preferred globally for heterologous protein production of GPCRs,

although it is hugely dependent on the specific protein being investigated and there does not exist a single, general method that is optimal for all proteins.

For expression of GPCRs in insect cells, the baculovirus expression system can be utilized. This method is dependent on a cloned baculoviral DNA, also known as bacmid, for proper expression. The principle for this method has been explained previously [48]. Briefly, a gene of interest for recombinant expression is cloned into a donor plasmid. The plasmid can then recombine with the parent bacmid to generate an expression bacmid that contains the gene of interest. This method also operates by a site-specific transposition of the gene, i.e. a specific sequence of the donor plasmid that includes your gene will transpose with a region of the parent bacmid such that it is controlled by a different promoter. This is achieved by a helper plasmid which codes for the necessary transposases. This will result in that *Escherichia coli* clones that contain the parent bacmid, will gain an antiobiotic marker from the donor plasmid as well as lose its inherent *lacZ* marker. Clones containing the desired expression bacmid can hence easily be selected. The viral DNA can be isolated, and the bacmid DNA is transfected to insect cells resulting in recombinant virus production. The baculovirus can then infect other uninfected insect cells which will lead to a high frequency of infection and thus a large expression of the introduced gene will be attained. A flow chart over the whole process is shown in Fig 2.3. The BacMam system is practically identical, with the exception that the infection step is performed on mammalian cells instead of insect cells and the bacmid contains a different promoter for expression [47].

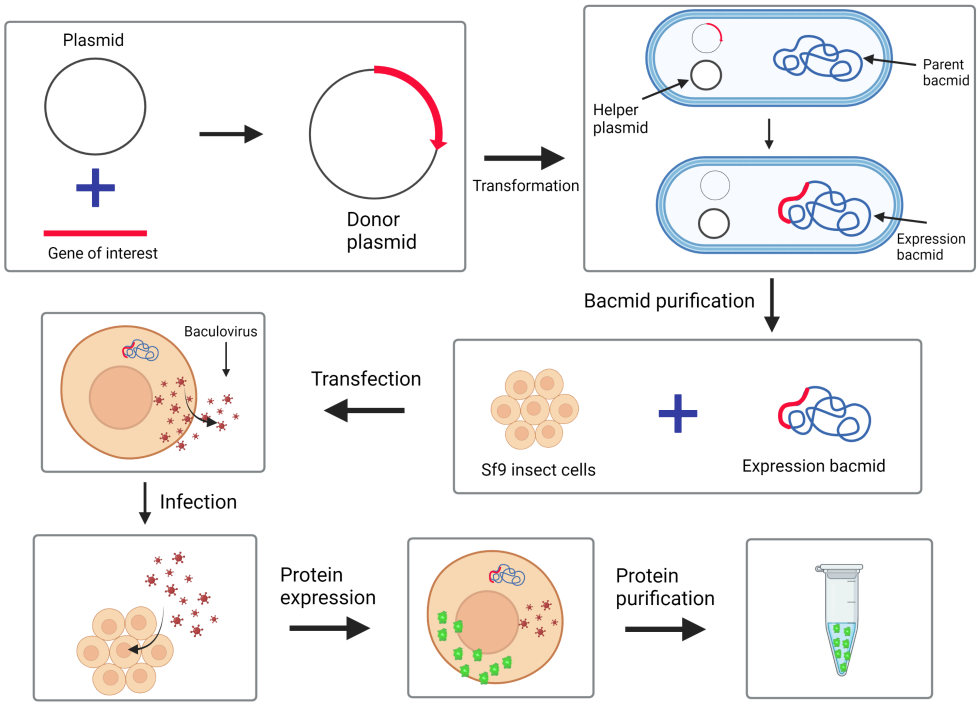



**Figure 2.3:** *Flow chart over the baculovirus expression system. Created with BioRender.com*

# 3
# Methods

## 3.1 Cloning

The human HCAR1 gene and the Nb6 gene were ordered from Eurofins and had been codon optimized for expression in Sf9 insect cells. The HCAR1 gene contained a 10x His-tag and a FLAG epitope (DYDKKK) located at the N-terminus, followed by a HRV 3C protease cleavage site. The C-terminus was truncated and fused to green fluorescent protein (GFP). The Nb6 binding region of KOR had also been transferred to the HCAR1 gene with residues R204$^{5.66}$ to K216$^{6.31}$ of HCAR1 being exchanged for K254$^{5.66}$ to R270$^{6.31}$ from KOR. The Nb6 gene contained a 6x His-tag at the C-terminus followed by GFP. Amplification of the genes were performed by a PCR reaction with Q5 polymerase and the program and primers used are listed in Appendix A.1 and A.2 for HCAR1 and Nb6 respectively. The pFastBac1 vector was used for cloning of both genes.

Similarly to the HCAR1 construct described above that was used for the baculovirus expression system in insect cells, a second HCAR1 construct was generated that was codon optimized for mammalian cells, specifically HEK293F cells. This construct would be used in the BacMam system instead of the baculovirus expression system for insects and will hereafter be denoted as HCAR1-BacMam while the other construct will be referred to as simply HCAR1. The other differences for HCAR1-BacMam compared to HCAR1 apart from codon optimization was that HCAR1-BacMam had yellow fluorescent protein (YFP) instead of GFP, as well as the mutation D7.49N that had been found to be stabilizing for the inactive state conformation based on previous results from a colleague (unpublished results). Primers and PCR program used for this construct are listed in Appendix A.3.

An HCAR1/G-protein construct was also designed for obtaining the active state structure. The Nb6 binding region of the HCAR1 construct was exchanged for the native sequence of HCAR1 (optimized for insect cells) using KLD site-directed mutagenesis for insertion of the original sequence. For the G-protein, the subunits $G_\alpha$(i), $G_\beta$ and $G_\gamma$ were expressed to form a complete G-protein. These subunits contained no tags or fusion partners except $G_\beta$ which contained a 6x His-tag at the N-terminus. $G_\alpha$ was expressed using the pFastBac1 vector while $G_\beta$ and $G_\gamma$ were co-expressed in the pFastBacDual vector. Primers and PCR program used for this construct are listed in Appendix A.4.

## 3.2 Expression and purification of HCAR1 and HCAR1-BacMam

*E. coli* Top10 cells were transformed by heatshock with the pFastBac1 vector containing the HCAR1 gene and incubated overnight at 37°$C$, 190 rpm in 5 ml LB-media containing 100 µg/ml ampicillin. The plasmid was purified (QIAprep Spin Miniprep Kit) and then transformed into *E. coli* DH10Bac cells to generate a bacmid. The bacmid was purified using an optimized protocol, outlined in Appendix B.1. Purified bacmid was diluted to 1000 ng/µl and used for transfection of insect cells.

For HCAR1, Sf9 insect cells were transfected at a density of 2.5 millions cells/ml with 5 µl of the bacmid. Sf-900™ II SFM (ThermoFisher) was used as media and Expifectamine (ThermoFisher) used as transfection reagent. Transfected cells were grown for 96 hours (300 rpm, 37°$C$), and the media containing the virus was then used to infect insect cells for a higher expression compared to solely transfection. For infection, 40 ml of cells were grown, infected at a density of 2 million cells/ml with a virus amount of 50 µl/ml. 48 hours after infection, cells were centrifuged and pellets were frozen at -80°$C$.

HCAR1-BacMam was transfected with the same procedure described above to generate virus. This virus was then used to infect HEK293F cells cultivated in FreeStyle™ 293 Expression Medium (ThermoFisher). Cells were infected at a density of 2 million cells/ml with a virus titer of 80 µl/ml for a total volume of cells of 100 ml. After 24 hours, sodium butyrate was added to a concentration of 10 mM. 72 hours after infection, cells were centrifuged and pellets were frozen at -80°$C$.

For all purifications described hereafter, all following steps and centrifugations were performed at 4°$C$ if not specified otherwise. HCAR1 and HCAR1-BacMam were both purified using the same protocol described here.

25 ml of low salt buffer (10 mM HEPES pH 7.5, 10 mM $MgCl_2$, 20 mM KCl) was added to the cell pellet along with protease inhibitor. The mixture was homogenized by douncing 25 times and was then centrifuged at 50 000 x g for 25 min. The supernatant was discarded, and 25 ml high salt buffer (10 mM HEPES pH 7.5, 10 mM $MgCl_2$, 20 mM KCl, 1 M NaCl) was added to the pellet along with protease inhibitor and dounced 25 times. The solution was centrifuged at 50 000 x g for 25 min. The supernatant was discarded, and 10 ml of low salt buffer with protease inhibitor was added to the pellet and dounced 25 times. 10 ml of solubilization buffer (50 mM HEPES pH 7.5, 1600 mM NaCl, 10 % glycerol (v/v), 40 µM lactate, 1%/0.2% DDM/CHS) was added to the solution along with protease inhibitor, 1 mg/ml iodoacetamide and 40 µM lactate. The solution was incubated and stirred for 2.5 hours and was then centrifuged at 50 000 x g for 25 min. The supernatant was transferred to a Falcon tube and 300 µl slurry of TALON beads were added, along with 5 mM imidazole. The mixture was incubated overnight and centrifuged at 700 x g for 5 min. The supernatant was removed, 1 ml of wash 1 buffer (50 mM HEPES pH 7.5, 800 mM NaCl, 10 mM $MgCl_2$, 10 % glycerol (v/v), 10 mM

imidazole, 8 mM ATP, 40 µM lactate, 1%/0.2% DDM/CHS) was added, incubated for 20 min and then centrifuged at 700 x g for 3 min. This was then repeated once. The beads were then washed with 0.5 ml of wash 2 buffer (50 mM HEPES pH 7.5, 800 mM NaCl, 10 % glycerol (v/v), 20 mM imidazole, 40 µM lactate, 0.05%/0.01% DDM/CHS) and incubated for 20 min then centrifuged at 700 x g for 3 min. This was repeated once. The suspension was then added to columns and eluted using gravity flow elution in Micro Bio-Spin™ columns (Bio-Rad). Elution buffer (25 mM HEPES pH 7.5, 800 mM NaCl, 10 % glycerol (v/v), 220 mM imidazole, 40 µM lactate, 0.01%/0.002% DDM/CHS)) was added to the beads and the sample was eluted in 100 µl fractions except the first fraction which was 50 µl. The fractions containing purified receptor was pooled and used for further analysis or complex formation, in this study up to four fractions were collected since practically all of the sample had been eluted by this point. The first fraction were omitted in presenting of the results since essentially no protein would have had the time to be eluted in this fraction.

## 3.3 Expression of Nb6

Nb6 was transformed using the same method as HCAR1 described above and the bacmid was isolated using the same protocol. Transfection and infection of the Sf9 cells were performed identically as for HCAR1 with the exception that a virus titer of 40 µl/ml was used for a volume of cells of 40 ml for infection.

25 ml of low salt buffer was added to the pellet along with protease inhibitor. The mixture was homogenized by douncing 25 times and was then centrifuged at 4730 rpm for 1 h. A 700 µl slurry of TALON beads was added to the supernatant, along with 10 mM imidazole. The mixture was incubated overnight and then centrifuged at 700 x g for 5 min. The supernatant was removed, 1 ml of wash 1 buffer without detergent, ATP or lactate was added, incubated for 20 min and then centrifuged at 700 x g for 3 min. This was then repeated once. The beads were then washed with 0.5 ml of wash 2 buffer without detergent or lactate and incubated for 20 min then centrifuged at 700 x g for 3 min. This was repeated once. The suspension was then added to columns and eluted using gravity flow elution. Elution buffer without detergent or lactate was added to the beads with the same fractionation scheme eplained before. Fractions containing protein could be used for further analysis or complex formation, however only one fraction of Nb6 was used in the complex formation due to a too high fluorescence in another fraction to be measured accurately.

## 3.4 Expression and purification of HCAR1/G protein complex

HCAR1 that had the Nb6 binding region substituted for the original sequence was transformed and the bacmid isolated as previously described. This was also the case for $G_\alpha$, $G_\beta$ and $G_\gamma$. The only difference was that $G_\beta$ and $G_\gamma$ were co-expressed in the pFastBacDual vector, though this had no impact on the method used. Ad-

ditionally, a bacmid containing the antibody ScFv16 which is a GPCR/G-protein complex-stabilizing antibody was generated. Transfection for all of these constructs was performed identically as above and they were transfected separately. For infection, the different constructs were co-infected at titers of 50 µl/ml for all viruses except the $G_\alpha$ virus which was infected with a titer of 100 µl/ml. A total volume of cells of 60 ml was used for infection.

Purification of HCAR1/G-protein complex was performed as following. 20 ml of low salt buffer was added to the pellet along with protease inhibitor. The mixture was homogenized by douncing 25 times and after this 50 µM lactate and 25 mU/ml apyrase was added to the mixture. The mixture was incubated at room temperature for 1.5 hours with gentle stirring. 20 ml solubilization buffer was added to the mixture and was then incubated for 2.5 hours at $4°C$. All purification steps after this was performed identically as described for HCAR1 and HCAR1-BacMam.

## 3.5 Complex formation of HCAR1-BacMam and Nb6

The second fraction of purified Nb6 was added to the pooled fractions of HCAR1-BacMam at a molar ratio of 2:1. The mixture was incubated overnight (∼16 h) at $4°C$. To evaluate if there had been a complex formation, a high-performance liquid chromatography (HPLC) analysis was performed. However, it was not certain that even if an interaction was formed it would be distinguishable on a chromatogram alone due to the small size of Nb6. Therefore an additional purification step was performed on the complex and was then evaluated on a Western blot (WB) to ascertain if there was an interaction. HCAR1-BacMam contains a FLAG-tag while Nb6 does not, hence if a FLAG purification is performed on the complex and both HCAR1-BacMam and Nb6 is thereafter observed on a Western blot, it indicates that there is an interaction between the two proteins. Theoretically, only a His-WB needs to be performed since both HCAR1-BacMam and Nb6 contain His-tags, but HCAR1-BacMam has in our experiments been found to give very weak signals on His-WB and therefore an anti-FLAG-WB was also performed.

All steps for this FLAG purification process were performed at $4°C$. The complex was added to 100 µl of FLAG resin supplemented with 10 mM $CaCl_2$. The mixture was incubated for 1 hour and then centrifuged at 700 x g for 3 min. The supernatant was removed, 1 ml of wash 1 buffer without ATP or imidazole and instead 10 mM $CaCl_2$ was added, incubated for 20 min and then centrifuged at 700 x g for 3 min. This was then repeated twice. The suspension was then added to columns and eluted using gravity flow elution. Elution buffer with imidazole swapped for FLAG-peptide at a concentration of 100 µg/ml was added to the resin and eluted with the same fractionation scheme explained before.

## 3.6 HPLC, SDS-PAGE and Western blot analysis

All HPLC analyses in this study were of the type size-exclusion chromatography (SEC). The column used in HPLC was an AdvanceBio SEC 300A column (Agilent) and the buffer used was SEC buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 0.05%/0.01% DDM/CHS).

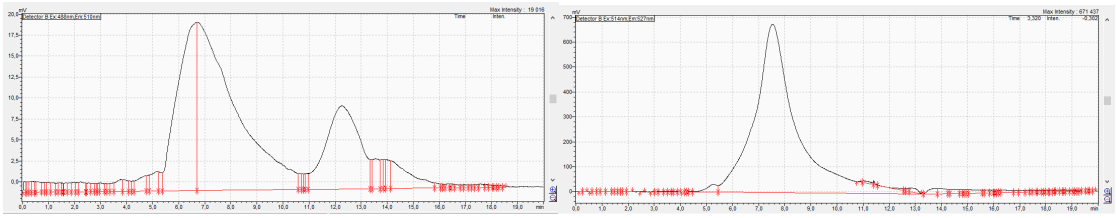
For the sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot protocols, see Appendix B.2 and B.3.

# 4

# Results

## 4.1 HCAR1 compared to HCAR1-BacMam

Both HCAR1 in insect cells and HCAR1 in mammalian cells were tested to evaluate if one host system was better for expressing HCAR1 and a better choice to obtain large quantities. In Fig 4.1, chromatograms for both systems are compared. HCAR1-BacMam used a 2.5x larger volume of cells but they are otherwise directly comparable. It could be observed that HCAR1-BacMam worked significantly better in regard to intensity, peak shape (symmetrical and monodsiperse peak, indicative of a homogenous protein) and expected retention time. Furthermore, a small peak is observed in Fig 4.1 (a) which is likely a product from proteolytic cleavage of the HCAR1 construct containing GFP, and hence an impurity. HCAR1-BacMam was therefore used for complex formation. Highest intensity fractions are shown in Fig 4.1, all fractions are shown in Appendix C.1.



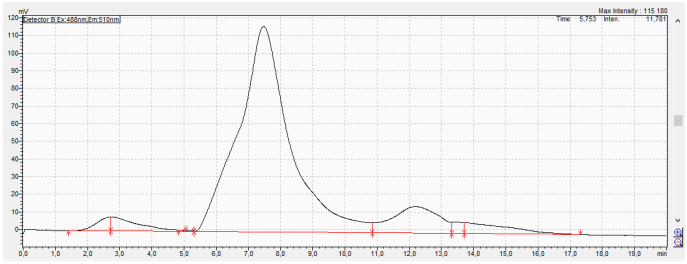
((a)) ((b))

**Figure 4.1:** *Comparison of highest intensity fraction for* ***(a)*** *HCAR1 (Max. intensity: 19 mV)* ***(b)*** *HCAR1-BacMam (Max. intensity: 671 mV).*

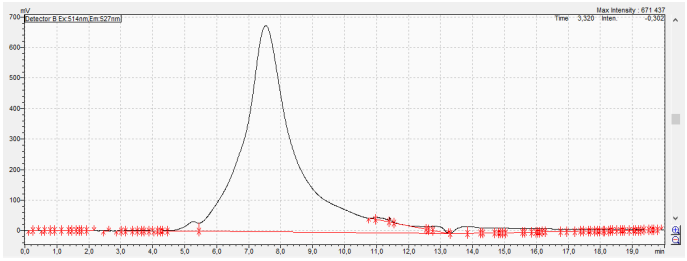
## 4.2 HCAR1-BacMam, Nb6 and complex HPLC data

The chromatograms for HCAR1-BacMam with the control can be observed in Fig 4.2. The control has an expected peak shape and retetion time (expected retention time of ∼7.5 min) indicating the purification protocol was successful seen from Fig 4.2 (a). For all samples purified using this protocol, the first fraction is excluded since the fraction volume was smaller than the column volume and hence practically no protein will be able to be eluted in this fraction. This was confirmed by the chromatograms of the first fraction as well (unpublished data). For the second to

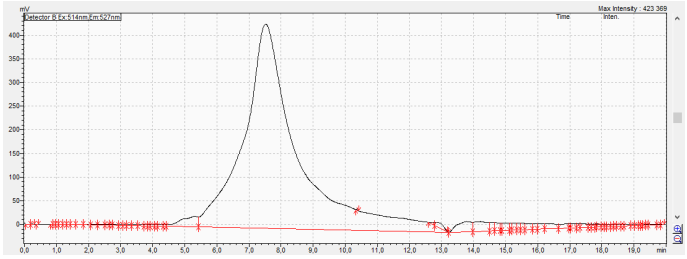
fourth fraction of HCAR1-BacMam, a sharp peak was observed at 7.5 min (Fig 4.2 (b)-(d)). Usually only the few fractions after the first contain protein and then a sharp drop in intensity is observed due to all protein being eluted. This behaviour was also observed here with fraction 2 containing much protein, fraction 3 some protein and fraction 4 basically none. Hence only fractions containing high amounts of protein, in this case fraction 2 & 3, were pooled and either used for Western blot analysis or complex formation.



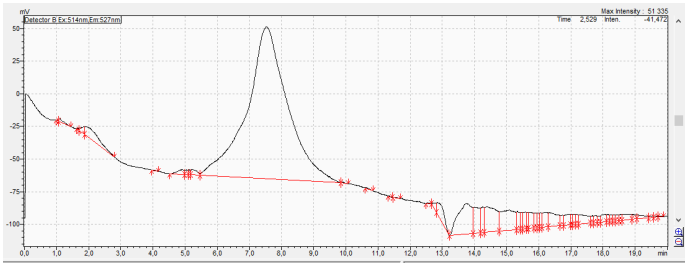
(a)



(b)



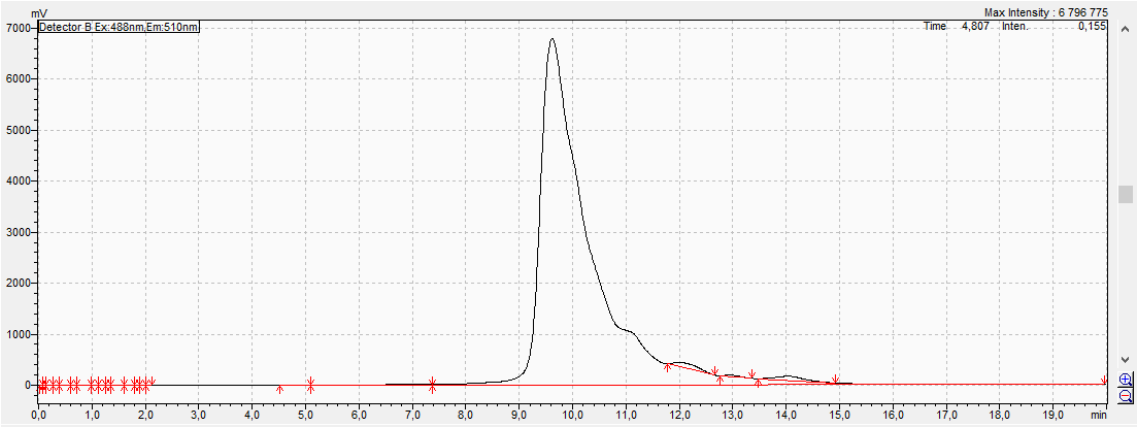
(c)



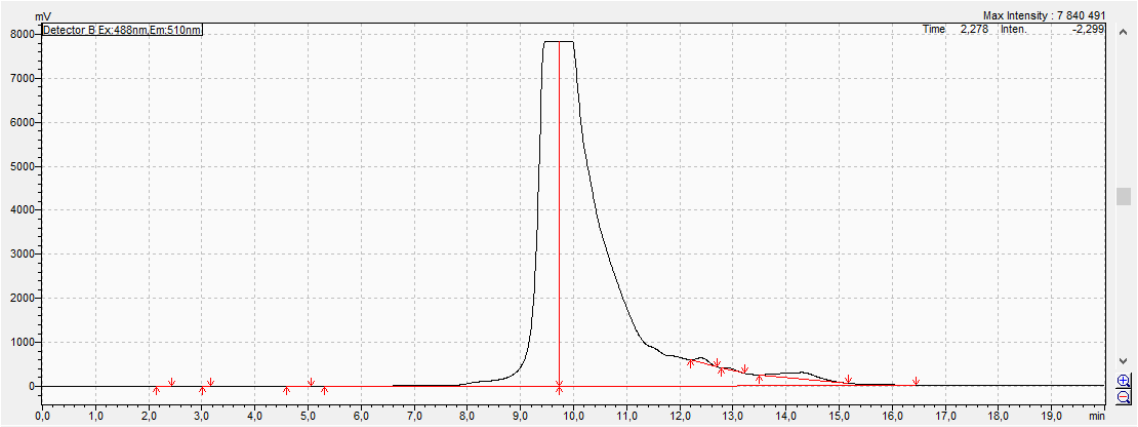
(d)

**Figure 4.2:** *Chromatograms of the control and HCAR1-BacMam fractions.* ***(a)*** *Control (Max. intensity: 115 mV).* ***(b)*** *Fraction 2 of HCAR1-BacMam (Max. intensity: 671 mV).* ***(c)*** *Fraction 3 of HCAR1-BacMam (Max. intensity: 423 mV).* ***(d)*** *Fraction 4 of HCAR1-BacMam (Max. intensity: 51 mV).*

The chromatograms for Nb6 are shown in Fig 4.3. Because Nb6 is a soluble protein and not a GPCR and this was the first time a nanobody construct was expressed in our lab, a control was not prepared alongside the Nb6 sample. The protocol was however modified to be optimized for soluble proteins and the results indicate that the purification was successful. Observing Fig 4.3, the same behaviour as for HCAR1-BacMam can be noted with the first two fractions containing much protein and a sharp drop in intensity seen from the fourth fraction and onward. The peak for the Nb6 sample was located at ∼9.5 min. This is approximately where we expect the nanobody to elute due to its small size. Additionally, it can be noted that the intensity was extremely high for the nanobody as seen in Fig 4.3 (a)-(b). In the third fraction the intensity was too high to be measured accurately by the fluorescence detector. Therefore Nb6 was not pooled and only the second fraction was used for analysis and complex formation.

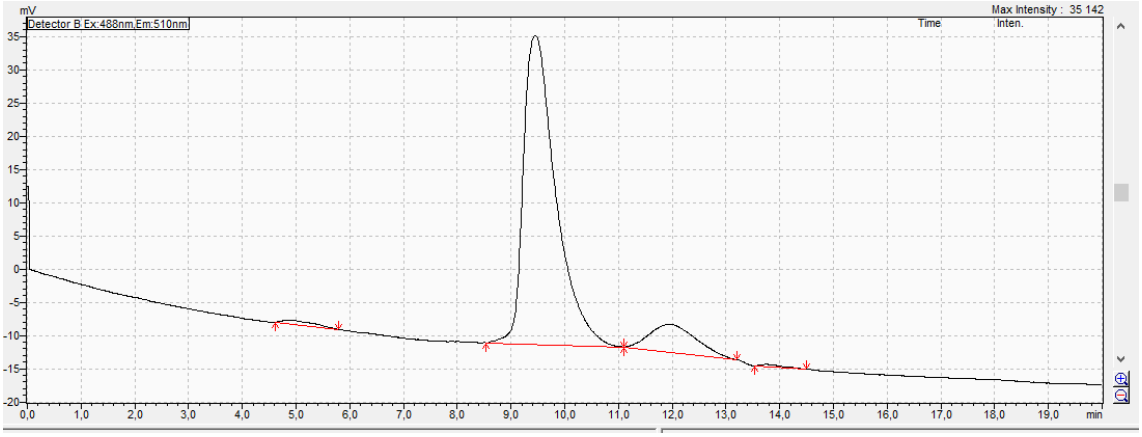



(a)




(b)

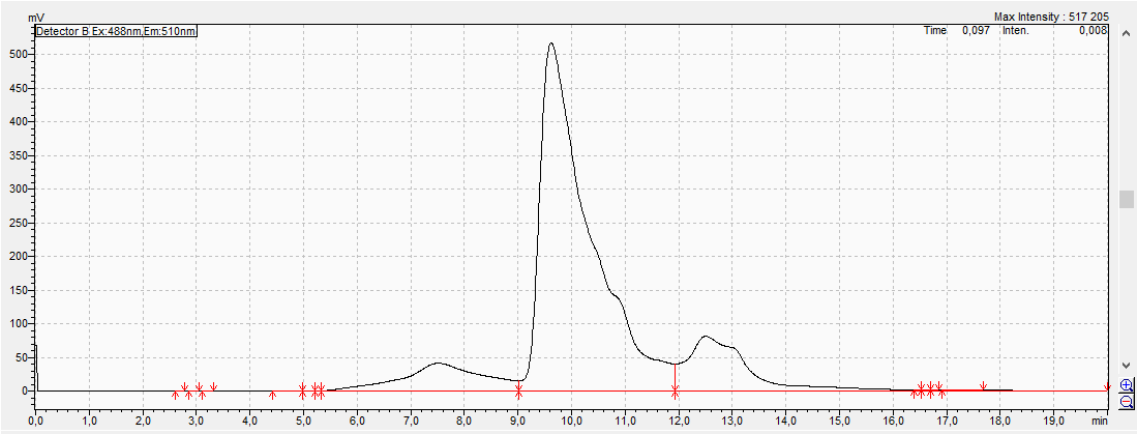
**Figure 4.3:** *Chromatograms of the Nb6 fractions.* ***(a)*** *Fraction 2 of Nb6 (Max. intensity: 6800 mV).* ***(b)*** *Fraction 3 of Nb6 (Max. intensity: 7840 mV).*



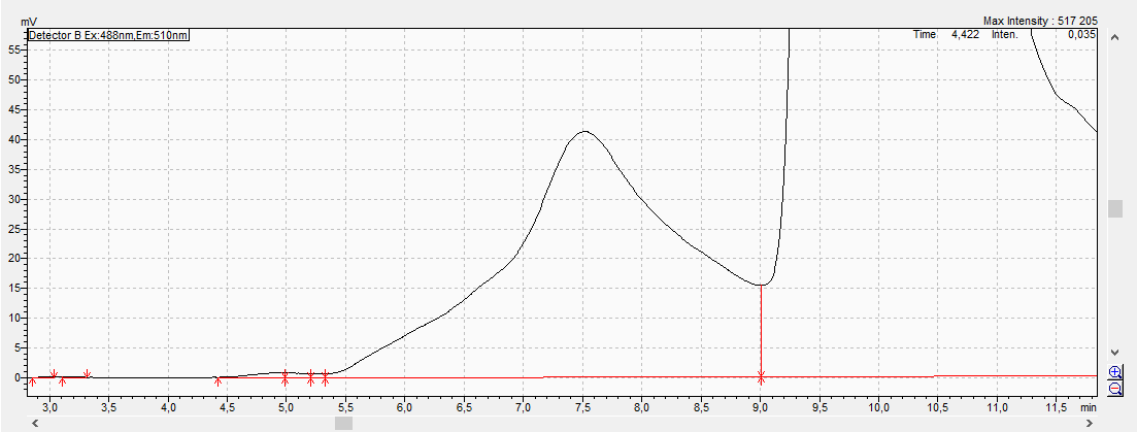

(c)

**Figure 4.3:** *Chromatograms of the Nb6 fractions.* ***(c)*** *Fraction 4 of Nb6 (Max. intensity: 35 mV).*

HCAR1-BacMam and Nb6 were incubated together and an HPLC analysis was performed on the potential complex. The chromatogram data for this is shown in Fig 4.4. Nb6 was in abundance as seen by Fig 4.4 (a). A zoomed in picture of the peak where we expect the complex to be eluted near is shown in Fig 4.4 (b). From this, it can be observed that the peak does not appear to be shifted and has a retention time of ∼7.5 min, the same as HCAR1-BacMam. This indicates an interaction between HCAR1-BacMam and Nb6 has not occurred. However, it is possible an interaction between the two proteins has developed and it is not distinguishable on a chromatogram alone. Additionally, the shape of the peak is slightly different from the peak for HCAR1-BacMam with a not as sharp peak and a blunt beginning of the peak, suggesting the presence of a larger species than HCAR1-BacMam alone (e.g. complex formation or aggregate of HCAR1-BacMam). To ascertain if an interaction had formed, a series of gels and Western blots were progressively done to confirm whether an interaction was attained and also to verify that HCAR1-BacMam and Nb6 were expressed correctly.




(a)




(b)

**Figure 4.4:** *Chromatograms of the potential complex.* ***(a)*** *Complex.* ***(b)*** *Zoomed in image of expected complex peak (Max. intensity: 42 mV).*

An HPLC analysis was also performed after the FLAG purification since this step was essential to ensure if there was an interaction. The chromatograms showed no clear result though with no appreciable quantities of HCAR1-BacMam or the presumed complex and no clear conclusion could be drawn, see Appendix C.3 for chromatograms.

## 4.3 Western blots or SDS-PAGE of HCAR1-BacMam, Nb6 and complex

An anti-FLAG Western blot for HCAR1-BacMam can be observed in Fig 4.5. HCAR1-BacMam is observed to have a clear band at the same size as the control protein A2A, a previously purified protein approximately the same size of HCAR1. This indicates that HCAR1-BacMam was successfully purified. The presumed complex is also shown in the image as a control and comparison to HCAR1-BacMam alone.

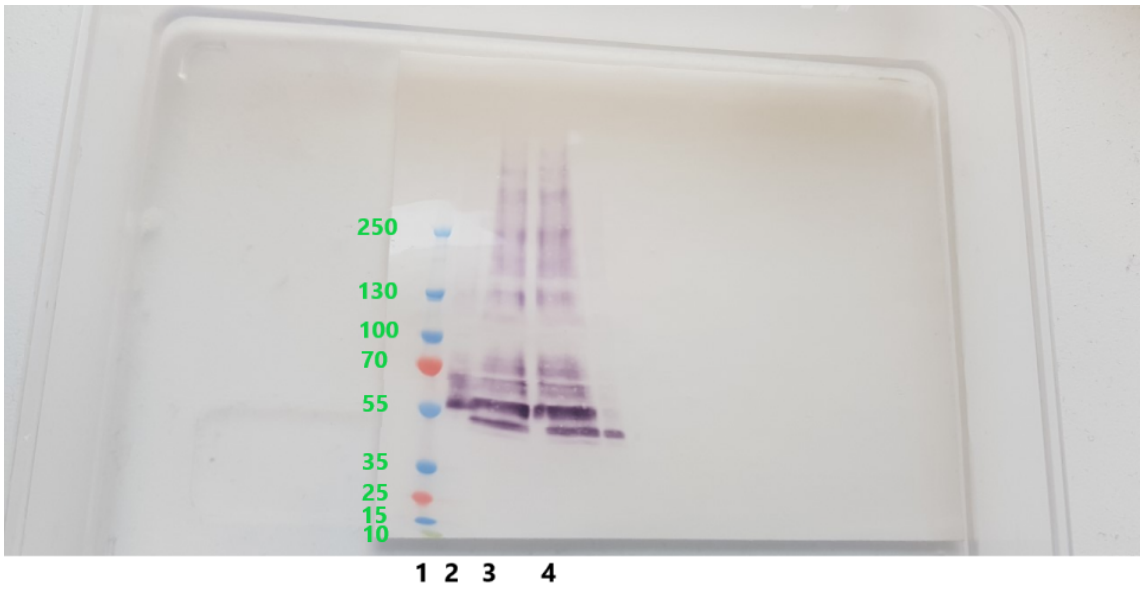
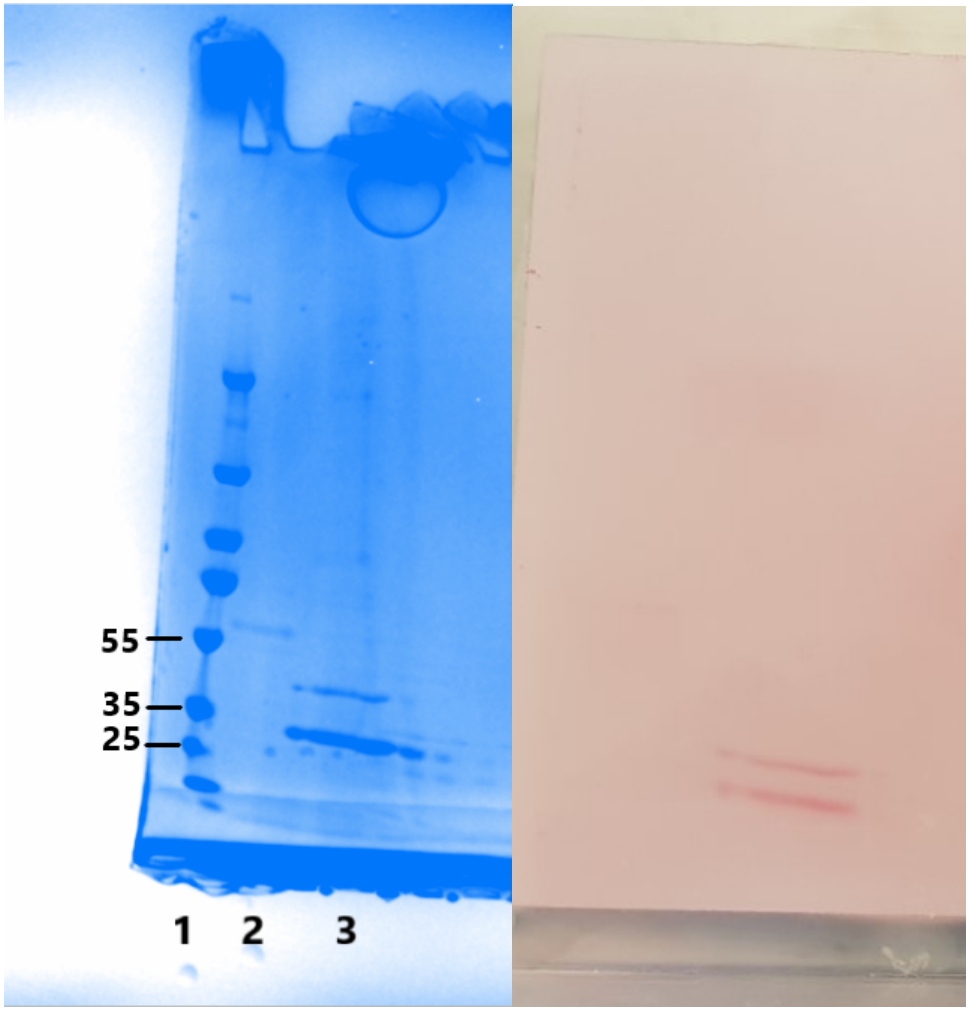



**Figure 4.5:** *Anti-FLAG Western blot. 1 is a reference protein ladder, 2 is A2A (a control protein approximately the same size of HCAR1), 3 is HCAR1-BacMam and 4 is the presumed complex.*

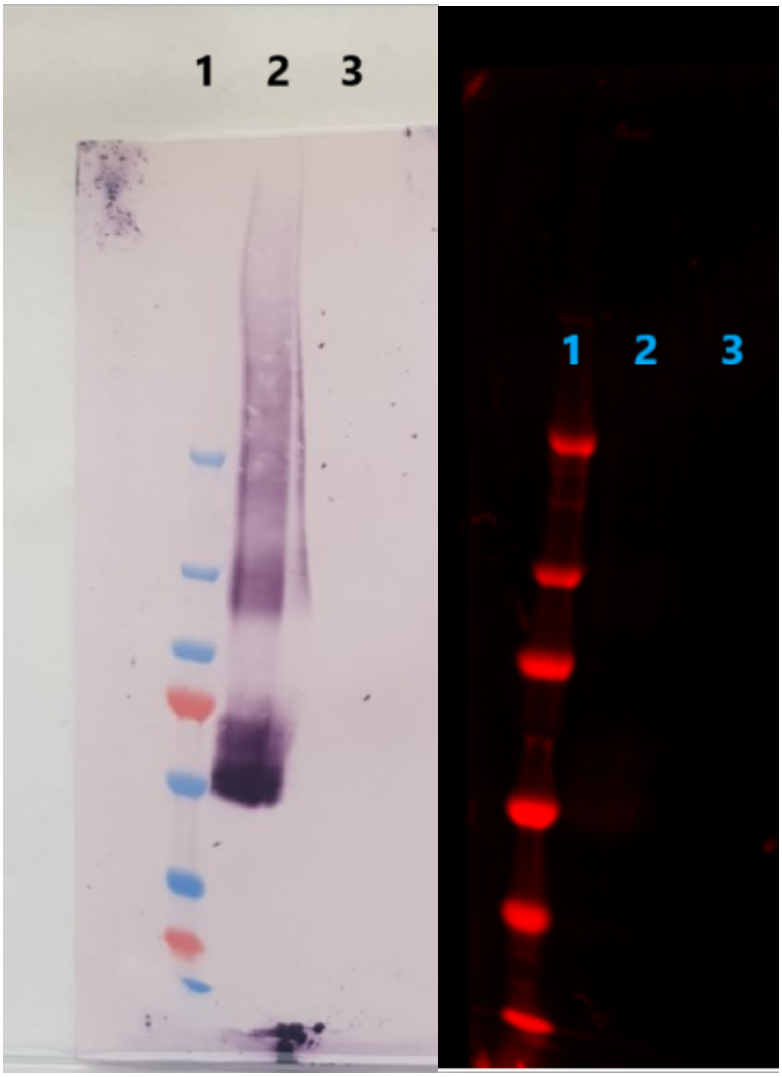
An SDS-PAGE was performed on Nb6, shown in Fig 4.6. A2A is faintly seen in lane 2, while two clear bands are detected in lane 3; one at ∼30 kDA and one at ∼40 kDa. The band at 40 kDa is likely Nb6 followed by GFP (Nb6 - 13 kDa, GFP - 27 kDa). For the other band however it could either be pure GFP that has been cleaved or a dimer of Nb6. An additional step was performed in which a His-WB was stained to visually detect if both bands contained a His-tag. From Fig 4.6 (b), it is observed that both bands contain a his-tag. However, since the His-tag is located at the C-terminus for Nb6, it is possible that the band is either a dimer of Nb6 or GFP with a His-tag as a result of cellular proteolysis leaving GFP with the His-tag. No clear conclusion could be made from these results.




((a)) ((b))

**Figure 4.6:** ***(a)*** *SDS-PAGE. 1 is a reference protein ladder, 2 is A2A and 3 is Nb6.* ***(b)*** *Stained His-WB of Nb6.*

Two Western blots were also performed on the presumed complex after the FLAG purification since this was essential to verify if there was an interaction between HCAR1-BacMam and Nb6. No clear result was obtained though due to low presence of the proteins which could not be detected on either of the Western blots, as seen in Fig 4.7. Unfortunately the control for the His-WB was poor and had degraded, however the WB was done correctly and this is evident by that this analysis was done on the same gel as the HCAR1/G protein sample, see Appendix C.4 for full image.




((a)) ((b))

**Figure 4.7:** ***(a)*** *Anti-FLAG Western blot. 1 is a reference protein ladder, 2 is a control and 3 is the presumed complex after the FLAG purification.* ***(b)*** *His Western blot. 1 is a reference protein ladder, 2 is a control and 3 is the presumed complex after the FLAG purification.*

## 4.4 HCAR1/G protein data of HPLC and Western blot

A chromatogram of the second fraction of HCAR1/G protein is shown in Fig 4.8. Fraction 3 and 4 are shown in Appendix C.2.

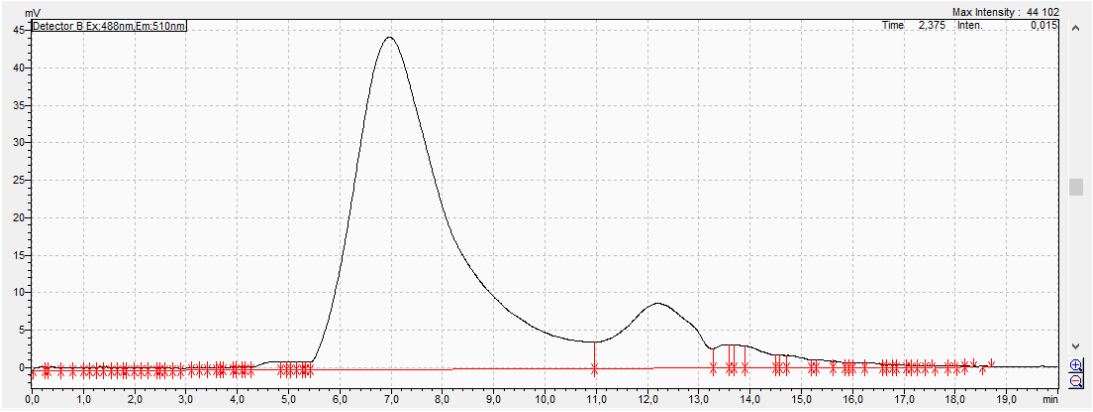
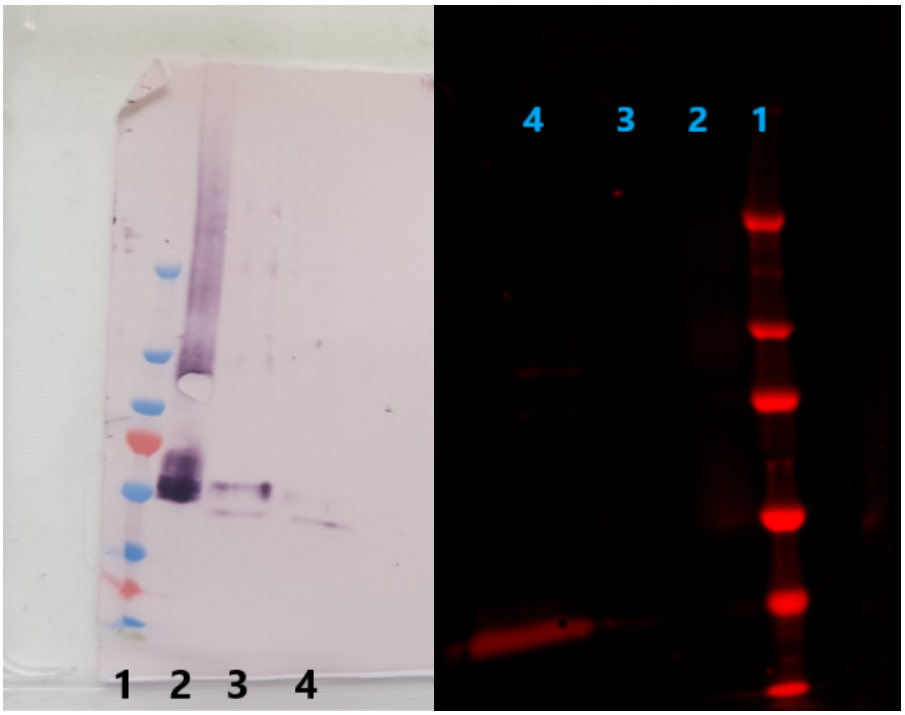



**Figure 4.8:** *Chromatogram of fraction 2 of HCAR1/G protein.*

It can be noted that the intensity of the peak is very low. The retention time is however shifted toward the left, indicating the presence of a larger complex. This could be the HCAR1/G protein complex, though it is more likely it is aggregate or an incorrectly expressed form of HCAR1 due to the large similarity with HCAR1 as previously shown in Fig 4.1 (a). To establish the origin of the observed shift in retention time, two Western blots were performed on the potential HCAR1/G protein complex. An anti-FLAG-WB to detect the presence of HCAR1 and a His-WB to assess if $G_\beta$ is expressed in large quantities. This is shown in Fig 4.9.




((a)) ((b))

**Figure 4.9:** ***(a)*** *Anti-FLAG-WB. 1 is a reference protein ladder, 2 is a control, 3 is HCAR1 and 4 is HCAR1/G protein.* ***(b)*** *His-WB. 1 is a reference protein ladder, 2 is a control, 3 is HCAR1 and 4 is HCAR1/G protein.*

In the anti-FLAG-WB, the HCAR1 protein in the potential HCAR1/G protein complex could be observed, however the intensity was very faint, consistent with the low intensity from the chromatogram. The subunit $G_\beta$ could however be observed in the His-WB (Fig 4.9 (b)), indicating the presence of this subunit and a potential G protein. The control could not be observed in the His-WB, though this is likely due to a poor control that was used. To ascertain that the other subunits were also correctly expressed, an SDS-PAGE was performed to evaluate if they were expressed.

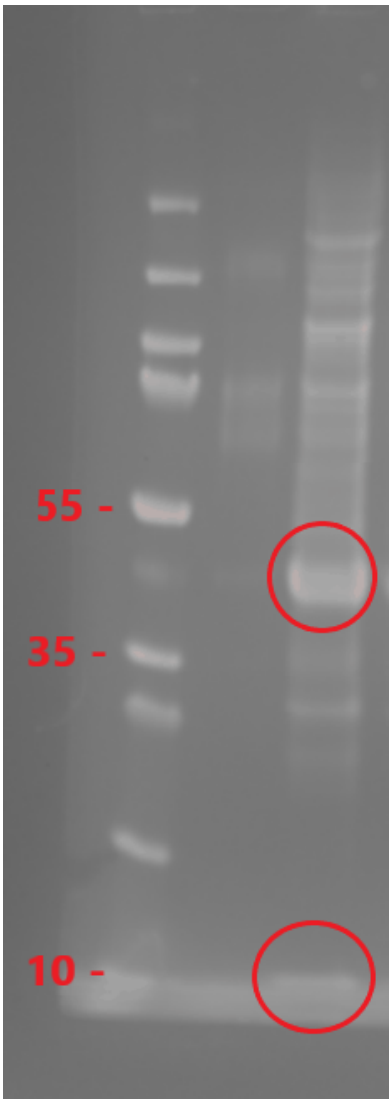



**Figure 4.10:** *SDS-PAGE of HCAR1/G protein.*

Observing Fig 4.10, a strong band was observed at ∼40 kDa, probably corresponding to $G_\alpha$ (40 kDa) and $G_\beta$ (38.5 kDa). A band could also be observed at ∼10 kDa, where we expect $G_\gamma$ (8 kDa) to be, even if it is difficult to distinguish. This suggests the subunits of the G protein were completely expressed and hence the G protein was also likely expressed in a fully functional form.

# 5
# Discussion

HCAR1 could be expressed in both insect cells and mammalian cells. It could however be noted that HCAR1 expressed in insect cells were significantly inferior compared to expression in a human, mammalian cell line in this study in terms of quantity and expected retention time, corresponding to stability. The reason for this is still not known. It may be due to innate differences between insect cells and mammalian cells, such as the membrane composition or presence of other molecules that may interact with HCAR1 and affect its stability and expression, or due to some post-translational modification unavailable for insect cells. Because GPCRs generally are highly expressed in insect cells as well, this indicates that the issue is HCAR1-specific, and more research is required to understand the underlying cause. Additional studies are also required to validate that the BacMam system is more suited for HCAR1 expression and may be a more tractable choice for high-protein expression and functional and structural studies of this protein.

HCAR1-BacMam was highly expressed and had a retention time equal to the expected, indicating HCAR1-BacMam was in a native and stable form in large quantities. The anti-FLAG-WB in Fig 4.5 also confirmed that this was the case. Expression of HCAR1-BacMam suitable for complex formation and structural determination was hence obtained.

Nb6 was observed to have an extremely high fluorescence intensity seen in Fig 4.3. The retention time was expected given the small size of Nb6, however the beginning of the peak was very sharp and vertical, and small peaks were present after the initial one. This could indicate that there may be a different species of Nb6, an impurity or something similar. An SDS-PAGE was done to see if Nb6 was pure. A band of ∼40 kDa was seen, very likely Nb6 followed by GFP which was the desired construct. A smaller band of ∼30 kDa was however also obtained and is likely the reason behind the sharp incline in the chromatogram. A His-WB that was stained for visual detection was performed to evaluate the origin of this band. It was believed to either be free GFP or a dimer of Nb6. The band was observed to still remain on the His-WB. It could therefore be a dimer of Nb6 that contains the His-tag. However, since the His-tag was located on the C-terminus of the Nb6 gene, it is also possible that this is GFP containing the His-tag which would lead to a similar molecular weight of ∼30 kDa. The existence of this band is likely due to a site near the C-terminus of Nb6 that is susceptible to proteolytic cleavage, but based from the analyses made no conclusion of what this is could be made and a different analysis such as mass spectrometry would be required to determine this.

Pure HCAR1-BacMam and Nb6 were combined in an attempt to form a complex that could be structurally determined. Both HCAR1-BacMam and Nb6 are observed in Fig 4.4, however no discernible shift in retention time for HCAR1-BacMam was observed that was expected if an interaction had formed. Furthermore, the intensity of the peak near 7.5 min was surprisingly low based on the high intensity of HCAR1-BacMam alone. It is possible that some of HCAR1-BacMam could have degraded or the stability of the protein could have been affected. The shape of the peak was relatively blunt as well, and some of the protein may have formed aggregate or possibly a complex with Nb6 since it is possible that this interaction would not be distinguishable on the chromatogram. Thus, a FLAG purifiation had to be performed to remove excess Nb6, and if Nb6 was still observed on a His-WB after purification, it would indicate that there was an interaction. However, neither HCAR1-BacMam or Nb6 was seen on Western blots (Fig 4.7). Although this is likely due to the fact that the total protein quantity was too low to be detected. Practically no HCAR1-BacMam was present after the FLAG purification (Appendix C.3), and hence even if there was an interaction between HCAR1-BacMam and Nb6 it would not be seen. Therefore, these steps have to be optimized to guarantee that there is sufficient material for after complex formation and FLAG purification to be able to definitively conclude if there is an interaction between the two proteins.

Processes requiring additional optimization to definitively reach a conclusion on a potential interaction is: Nb6 purification, complex formation and FLAG purification of the complex. Regarding Nb6 purification, it is desirable to eliminate the existence of the second band present on the His-WB. Potential solutions for this could be increased concentration of protease inhibitor or change the construct design with a His-tag at the N-terminus. Nb6 was also in this study produced in Sf9 insect cells while it was in previous studies expressed in bacteria, potentially affecting purity or stability [37]. Concerning complex formation, it was surprising that the intensity of the peak near 7.5 min in Fig 4.4 had dropped notably. The reason as to why is still unclear, though optimization of this could include a different incubation time, incubating with a more purified form of Nb6 as well as concentrating HCAR1-BacMam before complex formation to attain more material. Lastly, the FLAG purification for HCAR1-BacMam may need to be further optimized since most optimization of purification processes in this project concerned the initial His purification for HCAR1-BacMam and Nb6. Possible modifications could be to use more resin or to optimize the addition of $CaCl_2$ in the different steps, though it is probable that if the other suggestions previously described are implemented it would positively affect the FLAG purification of the complex and lead to higher quantities as a natural consequence.

The HCAR1/G protein complex method in this study is in contrast to the Nb6 method a conventional method that has been ubiquitously applied to a multitude of different GPCRs. It was therefore more likely that this approach would lead to a higher chance of determining the structure of HCAR1, in this case the active-state structure. This aim was though not achieved and still requires substantial opti-

mization and development in this study to reach. From Fig 4.8 and Fig 4.9, it was observed that HCAR1/G protein had a low intensity in the chromatogram and was barely seen on the WB. Notably though, HCAR1/G protein had a very similar shape and retention time as HCAR1 that was expressed in insect cells. Since HCAR1/G protein was also expressed in insect cells, it is therefore highly plausible that the low intensity is due to the fact this was expressed in insect cells that were previously concluded to not be optimal for expression of HCAR1 constructs. There is however a shift in retention time to an earlier timepoint in the chromatogram which is expected if there would be an interaction with the G protein, though it is likely due to an unstable form of HCAR1 or aggregate. The subunits of the G protein were however correctly expressed most likely. $G_\beta$ is observed in Fig 4.9 and the other subunits can be seen on the SDS-PAGE in Fig 4.10 and have thus likely formed a complete G protein. If this construct was designed and optimized for expression in mammalian cells using the BacMam system, it is likely it could be structurally determined after some optimization and lead to the first resolved active-state structure of HCAR1.

Lastly, an important point is also regarding the screening process for the Nb6 method. The fundamental principle of the method is based upon the interaction between Nb6 and the transferred region from KOR that recognizes the nanobody. It is impossible to know beforehand if there will be an interaction of Nb6 with your engineered construct. Therefore, the same researchers that developed this method simultaneously developed a bioluminescence resonance energy transfer (BRET)-based sensor that can screen if there is an interaction with your modified GPCR and Nb6 [37]. This allows for an efficient and rapid screening where many different constructs can be screened simultaneously to evaluate if there is a strong interaction. Conversely, in this study, only one construct was evaluated that was derived from homology based modeling and on the predicted structure of HCAR1, which may not even be accurate. A BRET-based sensor or some other technique that can measure binding affinity (e.g. Monolith) is ideally used for efficient screening of constructs. Though this requires a stable and pure form of the nanobody and receptor, of which optimized protocols adapted to our lab and equipment have not been constructed. Therefore focus was on optimizing these constructs to obtain stable forms in large quantities before optimizing the screening procedure which would be a large and time-consuming project in itself.

Ultimately, the procedures were not sufficiently optimized to generate a pure and stable complex that could be structurally determined. Further research and effort in this study are needed to ensure sufficiently stable and pure proteins, however, some suggested solutions are outlined in this report. Should this be implemented, there is a high possibility that a first resolved structure of HCAR1 could be obtained, most likely an active-state structure of HCAR1. The Nb6 method has been demonstrated to be a quick and effective method to obtain inactive-state structure of GPCRs, though it is still a new method and needs to be further investigated and developed for a wide-spread implementation in research. No conclusion on the applicability of the Nb6 method for HCAR1 could be made in this study due to issues in the purification process.

# 6
# Conclusion

The aim of the project was to elucidate either the active-state structure or inactive-state structure of HCAR1, as well as optimize the purification process for HCAR1 and related constructs. The first aim regarding resolving a structure of HCAR1 could not be obtained in this project, however, clear suggestions on implementations how to reach this have been presented in the discussion. The aim regarding optimization of the purification process was partly successful where for example a clear result on optimal host system production as well as a favorable purification protocol for HCAR1 is demonstrated. Some constructs such as Nb6 though requires additional optimization before it would be suitable for complex formation of HCAR1 and cryo-EM studies. Altogether, there is still considerable work of optimization left before a resolved structure is feasible, it is however within reach and clear propositions on how to attain this have been presented.

# References


[1] WHO. Cancer. 2022. Retrieved 20 April 2022, Updated 3 February 2022. Available from: https://www.who.int/news-room/fact-sheets/detail/cancer

[2] CDC. An Update on Cancer deaths. 2022. Retrieved 20 April 2022, Updated 28 February 2022. Available from: https://www.cdc.gov/cancer/dcpc/research/update-on-cancer-deaths/index.htm

[3] National Cancer Institute. Cancer Stat Facts: Lung and Bronchus Cancer. Bethesda: National Institute of Health. Retrieved 23 April 2022. Available from: https://seer.cancer.gov/statfacts/html/lungb.html

[4] National Cancer Institute. Cancer Stat Facts: Female Breast Cancer. Bethesda: National Institute of Health. Retrieved 23 April 2022. Available from: https://seer.cancer.gov/statfacts/html/breast.html

[5] National Cancer Institute. Cancer Stat Facts: Liver and Intrahepatic Bile Duct Cancer. Bethesda: National Institute of Health. Retrieved 23 April 2022. Available from: https://seer.cancer.gov/statfacts/html/livibd.html

[6] National Cancer Institute. Cancer Stat Facts: Pancreatic Cancer. Bethesda: National Institute of Health. Retrieved 23 April 2022. Available from: https://seer.cancer.gov/statfacts/html/pancreas.html

[7] National Cancer Institute. Types of Cancer Treatment. Bethesda: National Institute of Health. Retrieved 23 April 2022. Available from: https://www.cancer.gov/about-cancer/treatment/types

[8] Roland CL, Arumugam T, Deng D, Liu SH, Philip B, Gomez S, Burns WR, Ramachandran V, Wang H, Cruz-Monserrate Z, Logsdon CD. Cell surface lactate receptor GPR81 is crucial for cancer cell survival. Cancer Res. 2014 Sep 15;74(18):5301-10. doi: 10.1158/0008-5472.CAN-14-0319

[9] Lee YJ, Shin KJ, Park SA, Park KS, Park S, Heo K, Seo YK, Noh DY, Ryu SH, Suh PG. G-protein-coupled receptor 81 promotes a malignant phenotype in breast cancer through angiogenic factor secretion. Oncotarget. 2016 Oct 25;7(43):70898-70911. doi: 10.18632/oncotarget.12286

[10] Kooistra AJ, Mordalski S, Pándy-Szekeres G, Esguerra M, Mamyrbekov A, Munk C, Keserű GM, Gloriam DE. GPCRdb in 2021: integrating GPCR sequence, structure and function. Nucleic Acids Res. 2021 Jan 8;49(D1):D335-D343. doi: 10.1093/nar/gkaa1080

[11] Nagarathnam B, Kannan S, Dharnidharka V, Balakrishnan V, Archunan G, Sowdhamini R. Insights from the analysis of conserved motifs and permitted amino acid exchanges in the human, the fly and the worm GPCR clusters. Bioinformation. 2011;7(1):15-20. doi: 10.6026/97320630007015

[12] Rosenbaum DM, Rasmussen SG, Kobilka BK. The structure and function of G-protein-coupled receptors. Nature. 2009 May 21;459(7245):356-63. doi: 10.1038/nature08144

[13] Weis WI, Kobilka BK. The Molecular Basis of G Protein-Coupled Receptor Activation. Annu Rev Biochem. 2018 Jun 20;87:897-919. doi: 10.1146/annurev-biochem-060614-033910

[14] Deupi X, Kobilka B. Activation of G protein-coupled receptors. Adv Protein Chem. 2007;74:137-66. doi: 10.1016/S0065-3233(07)74004-4

[15] Kobilka BK, Deupi X. Conformational complexity of G-protein-coupled receptors. Trends Pharmacol Sci. 2007 Aug;28(8):397-406. doi: 10.1016/j.tips.2007.06.003

[16] Ge H, Weiszmann J, Reagan JD, Gupte J, Baribault H, Gyuris T, Chen JL, Tian H, Li Y. Elucidation of signaling and functional activities of an orphan GPCR, GPR81. J Lipid Res. 2008 Apr;49(4):797-803. doi: 10.1194/jlr.M700513-JLR200

[17] Liu C, Wu J, Zhu J, Kuei C, Yu J, Shelton J, Sutton SW, Li X, Yun SJ, Mirzadegan T, Mazur C, Kamme F, Lovenberg TW. Lactate inhibits lipolysis in fat cells through activation of an orphan G-protein-coupled receptor, GPR81. J Biol Chem. 2009 Jan 30;284(5):2811-2822. doi: 10.1074/jbc.M806409200

[18] de Castro Abrantes H, Briquet M, Schmuziger C, Restivo L, Puyal J, Rosenberg N, Rocher AB, Offermanns S, Chatton JY. The Lactate Receptor HCAR1 Modulates Neuronal Network Activity through the Activation of G and G Subunits. J Neurosci. 2019 Jun 5;39(23):4422-4433. doi: 10.1523/JNEUROSCI.2092-18.2019

[19] Briquet M, Rocher AB, Alessandri M, Rosenberg N, de Castro Abrantes H, Wellbourne-Wood J, Schmuziger C, Ginet V, Puyal J, Pralong E, Daniel RT, Offermanns S, Chatton JY. Activation of lactate receptor HCAR1 down-modulates neuronal activity in rodent and human brain tissue. J Cereb Blood Flow Metab. 2022 Mar 3:271678X221080324. doi: 10.1177/0271678X221080324

[20] Wagner W, Ciszewski WM, Kania KD. L- and D-lactate enhance DNA repair and modulate the resistance of cervical carcinoma cells to anticancer drugs via histone deacetylase inhibition and hydroxycarboxylic acid receptor 1 activation. Cell Commun Signal. 2015 Jul 25;13:36. doi: 10.1186/s12964-015-0114-x

[21] Liu C, Wu J, Zhu J, Kuei C, Yu J, Shelton J, Sutton SW, Li X, Yun SJ, Mirzadegan T, Mazur C, Kamme F, Lovenberg TW. Lactate inhibits lipolysis in fat cells through activation of an orphan G-protein-coupled receptor, GPR81. J Biol Chem. 2009 Jan 30;284(5):2811-2822. doi: 10.1074/jbc.M806409200

[22] Jumper J et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021 Aug;596(7873):583-589. doi: 10.1038/s41586-021-03819-2

[23] The G protein database, GproteinDb. Pandy-Szekeres G, Esguerra M, Hauser AS, Caroli J, Munk C, Pilger S, Keseru GM, Kooistra AJ, Gloriam DE.

Nucleic acids research, 2022, 50:D518-D525. Available from: https://gpcrdb.org/structure/

[24] Liang YL, Khoshouei M, Radjainia M, Zhang Y, Glukhova A, Tarrasch J, Thal DM, Furness SGB, Christopoulos G, Coudrat T, Danev R, Baumeister W, Miller LJ, Christopoulos A, Kobilka BK, Wootten D, Skiniotis G, Sexton PM. Phase-plate cryo-EM structure of a class B GPCR-G-protein complex. Nature. 2017 Jun 1;546(7656):118-123. doi: 10.1038/nature22327

[25] Manjasetty BA, Büssow K, Panjikar S, Turnbull AP. Current methods in structural proteomics and its applications in biological sciences. 3 Biotech. 2012;2(2):89-113. doi:10.1007/s13205-011-0037-1

[26] García-Nafría J, Tate CG. Structure determination of GPCRs: cryo-EM compared with X-ray crystallography. Biochem Soc Trans. 2021 Nov 1;49(5):2345-2355. doi: 10.1042/BST20210431

[27] Bai XC, McMullan G, Scheres SH. How cryo-EM is revolutionizing structural biology. Trends Biochem Sci. 2015 Jan;40(1):49-57. doi: 10.1016/j.tibs.2014.10.005

[28] Cheng Y. Single-particle cryo-EM-How did it get here and where will it go. Science. 2018 Aug 31;361(6405):876-880. doi: 10.1126/science.aat4346

[29] Warne T, Serrano-Vega MJ, Tate CG, Schertler GF. Development and crystallization of a minimal thermostabilised G protein-coupled receptor. Protein Expr Purif. 2009 Jun;65(2):204-13. doi: 10.1016/j.pep.2009.01.014

[30] Latorraca NR, Venkatakrishnan AJ, Dror RO. GPCR Dynamics: Structures in Motion. Chem Rev. 2017 Jan 11;117(1):139-155. doi: 10.1021/acs.chemrev.6b00177

[31] Punjani A, Fleet DJ. 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. J Struct Biol. 2021 Jun;213(2):107702. doi: 10.1016/j.jsb.2021.107702

[32] Nass Kovacs G, Colletier JP, Grünbein ML, Yang Y, Stensitzki T, Batyuk A, Carbajo S, Doak RB, Ehrenberg D, Foucar L, Gasper R, Gorel A, Hilpert M, Kloos M, Koglin JE, Reinstein J, Roome CM, Schlesinger R, Seaberg M, Shoeman RL, Stricker M, Boutet S, Haacke S, Heberle J, Heyne K, Domratcheva T, Barends TRM, Schlichting I. Three-dimensional view of ultrafast dynamics in photoexcited bacteriorhodopsin. Nat Commun. 2019 Jul 18;10(1):3177. doi: 10.1038/s41467-019-10758-0

[33] Congreve M, de Graaf C, Swain NA, Tate CG. Impact of GPCR Structures on Drug Discovery. Cell. 2020 Apr 2;181(1):81-91. doi: 10.1016/j.cell.2020.03.003

[34] Evans P, McCoy A. An introduction to molecular replacement. Acta Crystallogr D Biol Crystallogr. 2008 Jan;64(Pt 1):1-10. doi: 10.1107/S0907444907051554

[35] Congreve M, de Graaf C, Swain NA, Tate CG. Impact of GPCR Structures on Drug Discovery. Cell. 2020 Apr 2;181(1):81-91. doi: 10.1016/j.cell.2020.03.003

[36] Bao G, Tang M, Zhao J, Zhu X. Nanobody: a promising toolkit for molecular imaging and disease therapy. EJNMMI Res. 2021 Jan 19;11(1):6. doi: 10.1186/s13550-021-00750-5

[37] Michael J. Robertson, Makaía Papasergi-Scott, Feng He, Alpay B. Seven, Justin G. Meyerowitz, Ouliana Panova, Maria Claudia Peroto, Tao Che,

Georgios Skiniotis. Structure Determination of Inactive-State GPCRs with a Universal Nanobody. bioRxiv [Preprint]. Cited 5 June 2022. 2021.11.02.466983; doi: https://doi.org/10.1101/2021.11.02.466983

[38] Sigworth FJ. Principles of cryo-EM single-particle image processing. Microscopy (Oxf). 2016 Feb;65(1):57-67. doi: 10.1093/jmicro/dfv370

[39] Michael J. Robertson, Makaía Papasergi-Scott, Feng He, Alpay B. Seven, Justin G. Meyerowitz, Ouliana Panova, Maria Claudia Peroto, Tao Che, Georgios Skiniotis. Structure Determination of Inactive-State GPCRs with a Universal Nanobody. bioRxiv 2021.11.02.466983. Fig 1.c, Construct design & evaluation for inactive state GPCR structure determination.

[40] Henderson, R. (1995). The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. Quarterly Reviews of Biophysics, 28(2), 171-193. doi:10.1017/S003358350000305X

[41] Cheloha RW, Harmand TJ, Wijne C, Schwartz TU, Ploegh HL. Exploring cellular biochemistry with nanobodies. J Biol Chem. 2020 Nov 6;295(45):15307-15327. doi: 10.1074/jbc.REV120.012960

[42] Che T, English J, Krumm BE, Kim K, Pardon E, Olsen RHJ, Wang S, Zhang S, Diberto JF, Sciaky N, Carroll FI, Steyaert J, Wacker D, Roth BL. Nanobody-enabled monitoring of kappa opioid receptor states. Nat Commun. 2020 Mar 2;11(1):1145. doi: 10.1038/s41467-020-14889-7

[43] Khan KH. Gene expression in Mammalian cells and its applications. Adv Pharm Bull. 2013;3(2):257-63. doi: 10.5681/apb.2013.042

[44] Milić D, Veprintsev DB. Large-scale production and protein engineering of G protein-coupled receptors for structural studies. Front Pharmacol. 2015 Mar 31;6:66. doi: 10.3389/fphar.2015.00066

[45] Almo SC, Love JD. Better and faster: improvements and optimization for mammalian recombinant protein production. Curr Opin Struct Biol. 2014 Jun;26:39-43. doi: 10.1016/j.sbi.2014.03.006

[46] Massotte D. G protein-coupled receptor overexpression with the baculovirus-insect cell system: a tool for structural and functional studies. Biochim Biophys Acta. 2003 Feb 17;1610(1):77-89. doi: 10.1016/s0005-2736(02)00720-4

[47] Dukkipati A, Park HH, Waghray D, Fischer S, Garcia KC. BacMam system for high-level expression of recombinant soluble and membrane glycoproteins for structural studies. Protein Expr Purif. 2008 Dec;62(2):160-70. doi: 10.1016/j.pep.2008.08.004

[48] Kost TA, Condreay JP, Jarvis DL. Baculovirus as versatile vectors for protein expression in insect and mammalian cells. Nat Biotechnol. 2005 May;23(5):567-75. doi: 10.1038/nbt1095

# A

# Appendix - primers

In this appendix, primers and PCR programs for the different constructs are listed.

## A.1 Primers and PCR program for HCAR1

The primers used was for amplifying the HCAR1 gene.

- Forward primer: 5'-TACCGCATGCTATAACGGTAGCTGCTG-3'
- Reverse primer: 5'-TCGTGTCGACCACCGGTCGTTGG-3'

PCR program:

1. Heat lid to $110°C$
2. $98°C$ for 30 sec.
3. 35x cycle
   $98°C$ for 10 sec.
   $60 - 72°C$ gradient for 30 sec.
   $72°C$ for 3 min 30 sec.
4. $72°C$ for 5 min.
5. Store at $8°C$.

## A.2 Primers and PCR program for Nb6

The primers used was for amplifying the Nb6 gene.

- Forward primer: 5'-GATCGGATCCATGGCCCAGGTGCAA-3'
- Reverse primer: 5'-ACCTGTCGACCGCTTCCGGCTC-3'

PCR program:

1. Heat lid to $110°C$
2. $98°C$ for 30 sec.
3. 35x cycle
   $98°C$ for 10 sec.
   $55 - 65°C$ gradient for 30 sec.
   $72°C$ for 20 sec.
4. $72°C$ for 5 min.

5. Store at $8°C$.

## A.3 Primers and PCR program for HCAR1-BacMam

The primers used was for site-directed mutagenesis of a modified HCAR1 template using a KLD reaction.

- Forward primer: 5'-CCGGGAGAAGGACAGAAACCTG-AGAAAAGCCACACGCTTCATC-3'
- Reverse primer: 5'-CTGCCGCTCAGCAGTCTCACGCTC-TTCAGTGACCACACTATCTTAAAG-3'

PCR program:
1. Heat lid to $110°C$
2. $98°C$ for 30 sec.
3. 35x cycle
   $98°C$ for 10 sec.
   $60 - 72°C$ gradient for 30 sec.
   $72°C$ for 3 min 30 sec.
4. $72°C$ for 5 min.
5. Store at $8°C$.

## A.4 Primers and PCR program for HCAR1/G-protein

The primers used was for site-directed mutagenesis of a modified HCAR1 template using a KLD reaction.

- Forward primer: 5'-TCGTCAGGCTCGTATGAAGAAGGCAACCAGATTTATTATG-3'
- Reverse primer: 5'-GCCAGCTGCTGACGACGACGCAAACTCCAAACTATCTTAAAAC-3'

PCR program:
1. Heat lid to $110°C$
2. $98°C$ for 30 sec.
3. 35x cycle
   $98°C$ for 10 sec.
   $60 - 72°C$ gradient for 30 sec.
   $72°C$ for 3 min 30 sec.
4. $72°C$ for 5 min.
5. Store at $8°C$.

# B

# Appendix - protocol

## B.1 Bacmid purification protocol

**Day 1 - Bacmid generation**

1. Thaw DH10BAC cells on ice and transfer to 14 ml round-bottomed tubes.
2. Add 200 ng of plasmid to each tube.
3. Incubate on ice for 20 min.
4. Heatshock at 42°$C$ for 45 sec.
5. Place back on ice for 2 min.
6. Add 150 µl of SOC medium to each tube.
7. Shake for 4 hrs at 37°$C$.
8. Plate 10 µl on triple antibiotic plates (with IPTG and Bluo-Gal).
9. Incubate at 37°$C$ for 48 hours.

**Day 3 - Inoculation from plates**

1. From a single colony confirmed to have a white phenotype on plates, inoculate a liquid culture (LB medium) containing 50 µg/ml kanamycin, 7 µg/ml gentamicin and 10 µg/ml tetracycline. Grow at 37°$C$ overnight.

**Day 4 - Bacmid purification**

1. Spin cells at 3700 rpm, 10 min, 4°$C$. Discard supernatant.
2. Add 300 µl of P1 buffer to each tube, vortex and move to Eppendorf tubes.
3. Add 300 µl of P2 buffer to each tube, invert and incubate at room temperature.
4. Add 300 µl of P3 buffer and invert until clear. Incubate on ice for 5-7 min.
5. Spin at max speed, room temperature, 10 min, leading to a big white pellet.
6. Transfer 800 µl supernatant to an Eppendorf tube.
7. Add 700 µl of 100% isopropanol.
8. Invert and incubate on ice for 5 min.
9. Spin for 15 min at max speed and room temperature. Discard supernatant.
10. Add 500 µl of 70% EtOH and invert until pellet is loose.
11. Spin 5 min at max speed and room temperature.
12. Aspirate off EtOH with vacuum suction.

13. Leave tubes open to dry for 10-15 min.
14. Add 60 µl MQ water to tubes carefully. Store at $4°C$.

## B.2 Western blot protocol

Western blot transfer was done according to the manufacturer's instruction (Bio-Rad).

ANTI-FLAG-WB

1. Transfer protein to the PVDF membrane.
2. Incubate the membrane in blocking buffer for 1 hr.
3. Wash the membrane 3x times with PBS/TWEEN (5 minutes each time).
4. Incubate the membrane with primary antibody - Monoclonal ANTI-FLAG® M2-Alkaline Phosphatase (Sigma-Aldrich) - 1:1000 in blocking buffer for 1 hr.
5. After incubating the membrane in primary antibody, wash 3x times with PBS/TWEEN (5 minutes each time).
6. Treat the membrane with Substrate, SIGMAFAST BCIPÓ/NBT tablet. Add 10 ml of water and vortex until dissolved.
7. Incubate the membrane in substrate mixture for 10-30 minutes until color development.

His-WB

1. Transfer protein to a LF PVDF membrane.
2. Incubate the membrane in blocking buffer for 1 hr.
3. Wash the membrane 3x times with PBS/TWEEN (5 minutes each time).
4. Incubate the membrane with primary antibody - Penta-His Tag Monoclonal Antibody (ThermoFisher) - 1:1000 in blocking buffer for 1 hr.
5. After incubating the membrane in primary antibody, wash 3x times with PBS/TWEEN (5 minutes each time).
6. Incubate the membrane with secondary antibody - Goat anti-Mouse IgG Antibody, Alexa Fluor™ 700 (ThermoFisher) - 1:1000 in blocking buffer for 1 hr.
7. Wash the membrane 3x times with PBS/TWEEN (5 minutes each time).
8. Image the membrane with a fluorescence imaging system.
9. (Optional) Add 0.1% w/v Ponceau S in 5% v/v acetic acid for visual detection of proteins.

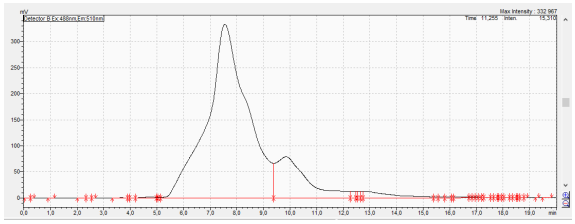
## B.3 SDS-PAGE protocol

NuPAGE™ 4 to 12%, Bis-Tris (ThermoFisher) gels were prepared according to the manufacturer's instruction. The gel was run at 225 V for 35 min. The gel was washed in SimplyBlue™ SafeStain (ThermoFisher) for 20 min, then washed twice for 20 min each time with MQ water.

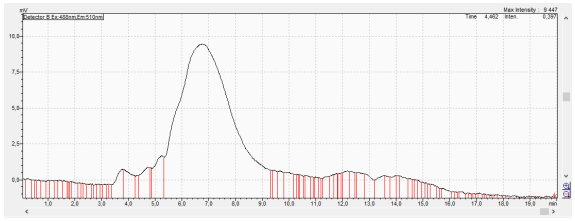
# C

# Appendix - results

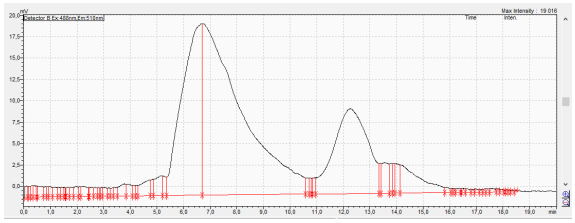
## C.1 HCAR1 chromatograms



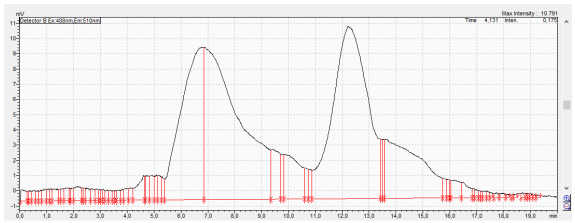
(a)



(b)



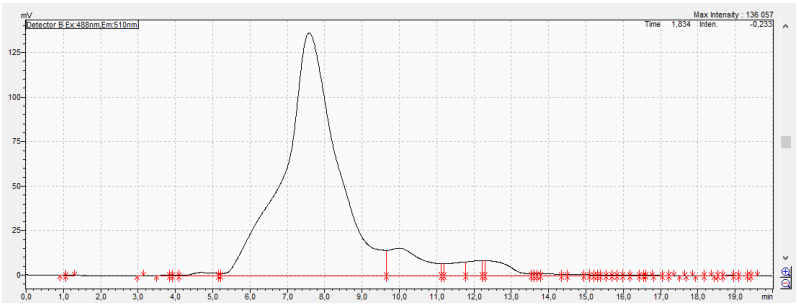
(c)



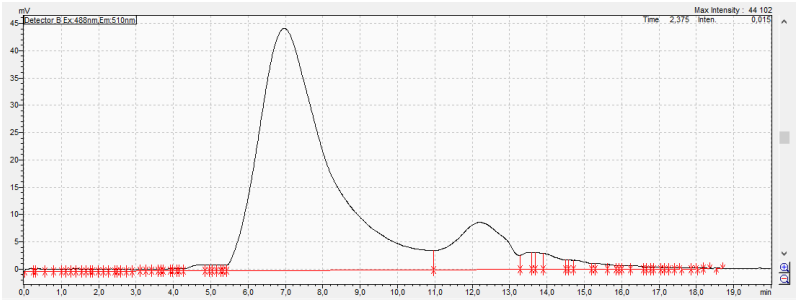
(d)

**Figure C.1:** *Chromatograms of the control and HCAR1 fractions.* ***(a)*** *Control.* ***(b)*** *Fraction 2 of HCAR1.* ***(c)*** *Fraction 3 HCAR1.* ***(d)*** *Fraction 4 of HCAR1.*

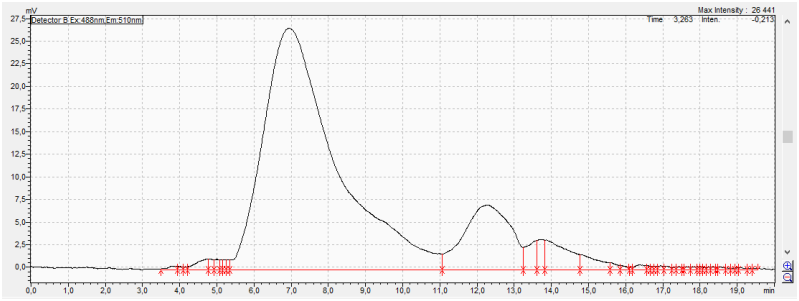
## C.2 HCAR1/G protein chromatograms



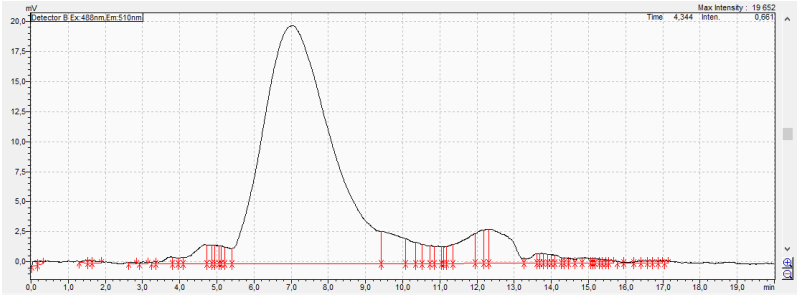
(a)



(b)



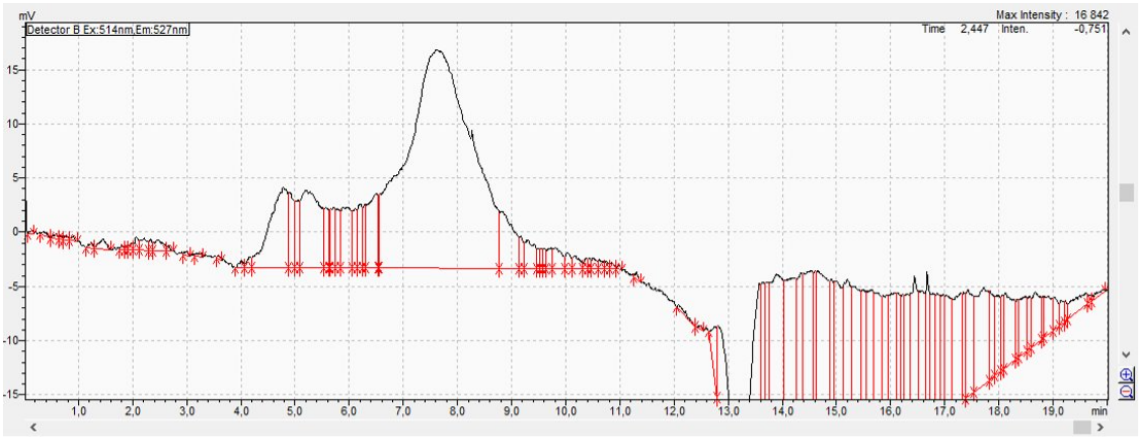
(c)



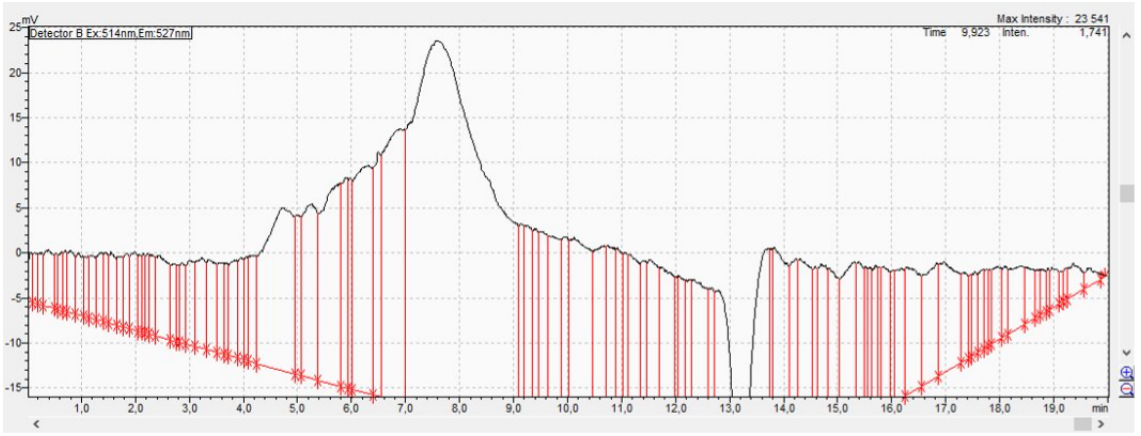
(d)

**Figure C.2:** *Chromatograms of the control and HCAR1/G protein fractions.* ***(a)*** *Control.* ***(b)*** *Fraction 2 of HCAR1/G protein.* ***(c)*** *Fraction 3 HCAR1/G protein.* ***(d)*** *Fraction 4 of HCAR1/G protein.*

## C.3 Chromatograms of FLAG purified complex




(a)




(b)

**Figure C.3:** *Chromatograms of FLAG purified complex.* ***(a)*** *Fraction 2 of FLAG purified complex.* ***(b)*** *Fraction 3 of FLAG purified complex.*

## C.4 His-WB of HCAR/G protein and complex

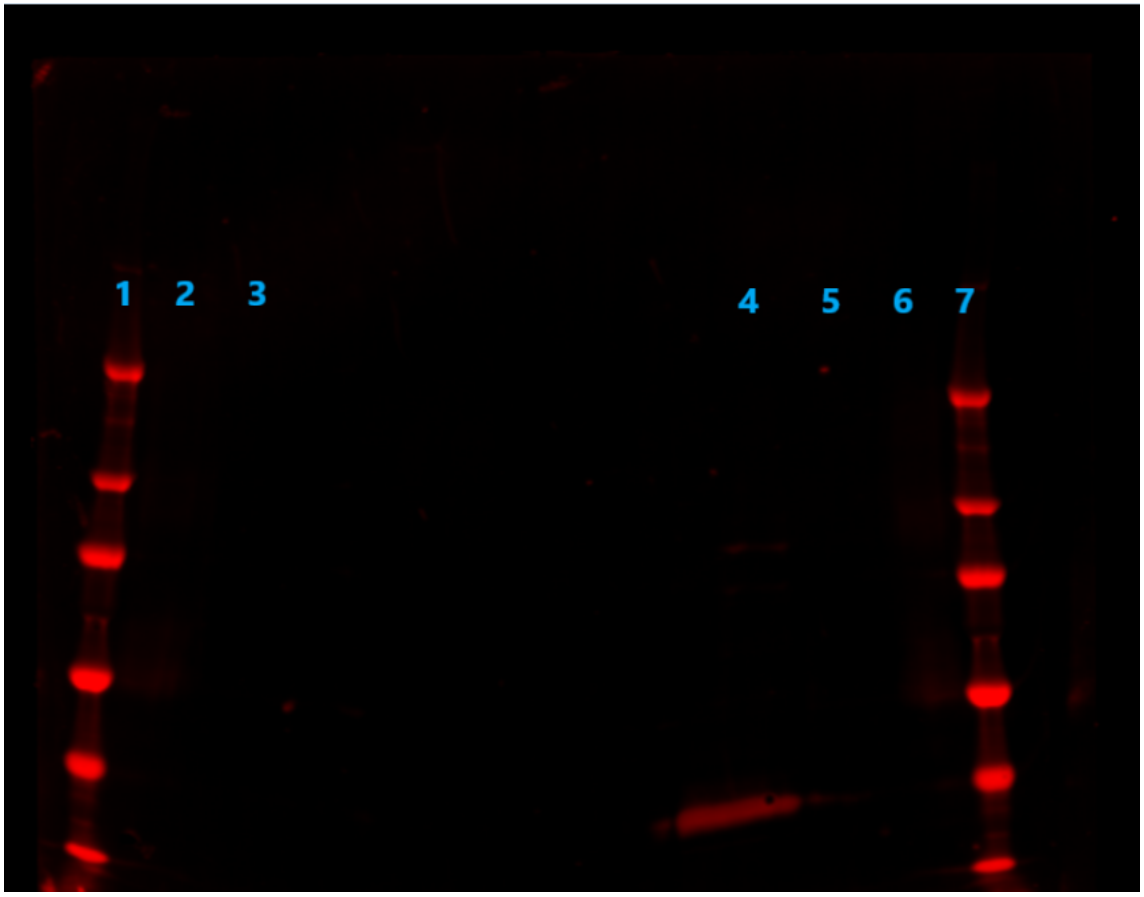



**Figure C.4:** *Full image of His-WB. 1 is a reference protein ladder, 2 is a control, 3 is the complex after the FLAG purification, 4 is HCAR1/G protein, 5 is HCAR1, 6 is a control and 7 is a reference protein ladder.*

**DEPARTMENT OF BIOLOGY AND BIOLOGICAL ENGINEERING**
**CHALMERS UNIVERSITY OF TECHNOLOGY**
Gothenburg, Sweden
www.chalmers.se


CHALMERS
UNIVERSITY OF TECHNOLOGY