

UNIVERSITY OF TURIN

UNIVERSITÀ DEGLI STUDI DI TORINO



**IN-SILICO SCREENING and EXPERIMENTAL
VALIDATION of NEW ALLOSTERIC MODULATORS
for the
CALCIUM-SENSING RECEPTOR**

submitted by

Wolfgang Schlattl

A Thesis in the Field of Bioinformatics

for the degree of

Doctor of Philosophy in Complex Systems for Life Sciences
(PhD)

Supervisor: Prof. Macro Botta

Co-supervisor: Luisa Pugliese, PhD

PhD. Cycle XXXII

Matr. 839432

2020

This thesis is dedicated to my family.



Acknowledgements

I want to thank my supervisor, Prof. Marco Botta from the University of Torino, for his support during my studies. Special thanks go to my supervisor Luisa Pugliese, PhD who allowed me to work for her at S.A.F.AN.-BIOINFORMATICS. She offered me the PhD position as Early Stage researcher in the Horizon 2020 Marie Curie project “CaSR Biomedicine – Calcium-sensing Receptor” and had a significant impact on mentoring and discussing every point of my work. The use of her in-silico profiling tool SAFAN-ISPSM contributed to the successful completion of this thesis. My gratitude also goes to Prof. Enikő Kally, who initiated the “CaSR Biomedicine” program. She introduced my fellow ESR colleagues and me to the scientific network of the calcium-sensing receptor. She also helped me in executing my experiments at the Department of Pathophysiology and Allergy Research in Vienna. Taha Elajnaf, the ESR fellow at the Medical University of Vienna, helped me master my cell culture experiments and introduced me to GraphPad prism, data analysis, and data representation. Iris Mos, the ESR fellow at the University of Copenhagen, trained and introduced me to the IP-1 assay. She provided me with the CaSR transfected HEK293 cells to perform the experimental validation. I want to thank Prof. Hans Bräuner-Osborne from the University of Copenhagen for the opportunity to carry out my academic secondment at his laboratory and the funding of the experiments. My gratitude also goes to Univ. Prof. Dr. Ulla Derhaschnig from the Department of Clinical Pharmacology at the Medical University of Vienna. She gave me the opportunity for a clinical secondment and provided me with deep insight into clinical trials. Sergio Rabellino helped me execute my calculations at the Occam C3S cluster, and I am thankful for all his suggestions. I want to thank the consortium of the CaSR biomedicine project for their support and the organisation and contribution to the ESR schools, feedback from the annual meetings and the journal club, especially Dr. Jutta Pauschenwein. I am also grateful for the lessons of Andres Hollow-Tas in Management Systems. Special thanks go to all the members of the CaSR Biomedicine consortium and all ESR fellows. Finally, I thank my wife Verena for her support throughout the years, for her love and the gift we share by our beloved daughter Valentina.

Abstract

The extracellular calcium-sensing receptor (CaSR) is a member of the class C G-protein coupled receptors. This membrane protein is expressed throughout the body but foremost in the parathyroid gland. In the kidney, its primary function is to maintain the extracellular calcium level in the normal range. The project aim is to discover new allosteric modulators for CaSR. A new compound could become the next potential therapeutic agents for the treatment of CaSR related calciotropic (familial hypocalciuric hypercalcemia, autosomal dominant hypocalcaemia, neonatal severe primary hyperparathyroidism) and non-calciotropic diseases (diabetes, Alzheimer's disease, cardiovascular disease or cancer). The discovery process included the in-silico screening of small molecule libraries with the ligand-based method SAFAN-ISPSM. The used compound library combines the database entries of DrugBank, FooDB, HerDing and TCM database@Taiwan. A substructure search was performed on the ZINC15 database with a set of CaSR related fragments obtained from SAFAN-ISPSM. A structure-based pharmacophore model was created to identify potential compounds on ZINCPharmer. The combined library of over 250000 entries was screened with SAFAN-ISPSM and resulted in 3925 potential compounds with a significant predicted binding affinity constant (pK). The screening results were sorted and filtered according to the evaluation parameters from SAFAN-ISPSM (compound-target ranking, pK, SAFAN score, the similarity of compounds and similarity of fragments). Homology models were created in the active and inactive state for molecular docking experiments with AutoDock Vina. The results helped to identify compounds able to enter the binding pocket in the transmembrane domain. In the end, a selection of 8 purchasable compounds was tested experimentally with an IP-1 Gq assay. The experimental validation revealed the discovery of the new potent positive allosteric modulator 3-([1-(naphthalen-1-yl)ethyl]amino)methyl-phenol.

Table of content

Acknowledgements	4
Abstract	5
Table of content	6
1 Introduction and background:	9
1.1 Aim of the thesis.....	9
1.2 The human extracellular calcium-sensing receptor (CaSR)	9
1.3 The function of CaSR in calcium homeostasis:	10
1.4 Ligands of CaSR	11
1.5 Structure of the extracellular calcium-sensing receptor (CaSR).....	12
1.6 G-protein coupled receptor (GPCR).....	14
1.7 CaSR signalling.....	15
1.8 CaSR related pathophysiology	17
1.8.1 Diseases related to gain-of-function mutations.....	17
1.8.2 Diseases related to loss-of-function mutations	17
1.9 Drug discovery and development of calcilytics and calcimimetics overview 18	
1.10 Which therapeutics are on the market acting on CaSR	19
2 Results:.....	21
Chapter I: Virtual screening for allosteric modulators of the extracellular calcium- sensing receptor (CaSR)	21
2.1 Structure prediction with I-TASSER	27
2.2 Structure prediction with GPCR-I-TASSER.....	34
2.3 Homology modelling of CaSR with YASARA	38
2.4 Homology modelling with YASARA for virtual screening.....	46
Chapter II: Discovery of novel allosteric modulator for the extracellular calcium- sensing receptor by in-silico profiling with SAFAN-ISPSM.....	63
2.5 Rationale:.....	63

2.6	Virtual profiling with SAFAN-ISPSM	63
2.7	Second virtual profiling approach with SAFAN-ISPSM	73
2.7.1	Experimental setup for the IP-1 Gq assay	108
2.7.2	Results	109
2.7.3	Conclusion	110
3	Discussion:	112
4	Materials and methods	114
4.1	Computer-aided drug discovery methods (CADD).....	114
4.1.1	Homology modelling.....	114
4.1.2	In-silico profiling with SAFAN-ISPSM.....	119
4.1.3	File formats	121
4.2	The validation set for virtual screening:	124
4.2.1	Negative allosteric modulators for the human extracellular calcium-sensing receptor.....	124
4.2.2	Positive allosteric modulators for the human extracellular calcium-sensing receptor.....	147
4.3	Compounds for experimental validation.....	162
4.4	Kits.....	162
4.5	Materials for the IP-1 Gq assay	163
4.6	IP-1 Gq assay	164
5	Bibliography:.....	166
6	Declaration	180
7	Abbreviation	180
8	Table of Figures	181
9	Appendix:	192
10	Funding.....	195
11	Publication	195
12	Curriculum Vitae.....	197

1 Introduction and background:

1.1 Aim of the thesis

This PhD is part of the Marie-Skłodowska Curie European Training Network, CaSR Biomedicine, funded by the European Community. The common goal of this programme is to increase the knowledge about the extracellular calcium-sensing receptor.

This PhD's objective is to identify ligands of the human extracellular calcium-sensing receptor (CaSR) for the treatment of related diseases. In this thesis, the combination of different computational methods helps select relevant compounds for the experimental validation. The resulting active compounds represent new lead structures for developing new therapeutics or chemical agents for future studies of CaSR.

1.2 The human extracellular calcium-sensing receptor (CaSR)

The extracellular calcium (Ca^{2+}_e) homeostasis in the body is regulated and maintained by the extracellular calcium-sensing receptor (CaSR). This protein inhibits the secretion of the calcium releasing parathyroid hormone (PTH) and urinary calcium reabsorption. The bovine parathyroid form of this class C G-protein coupled receptor (GPCR) was first isolated from *Xenopus laevis* oocytes and described in 1993 (1).

This receptor is expressed throughout the body. In the skin, it promotes keratinocyte differentiation. CaSR modulates acid secretion in the stomach. The receptor influences the differentiation and recruitment of osteoblasts and osteoclasts in bones. In the colon, it controls toxin-mediated fluid secretion, but it is most expressed in parathyroid glands and kidneys (Figure 1.1) (2).

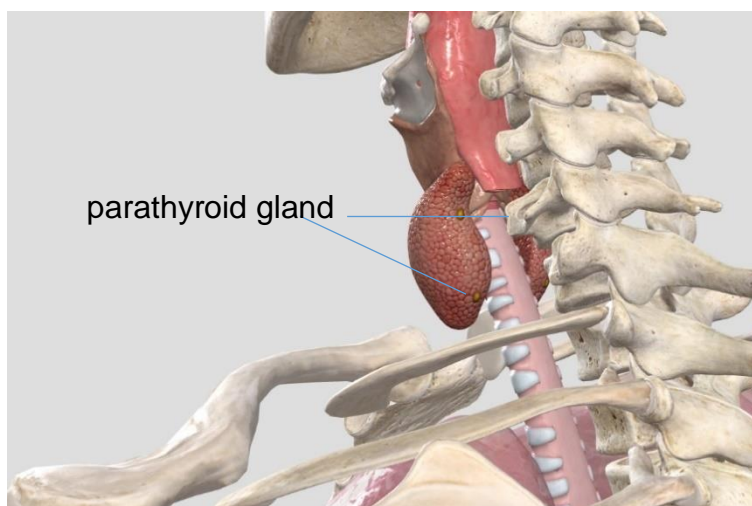


Figure 1.1: The extracellular calcium-sensing receptor is mainly expressed in the parathyroid gland and kidney

1.3 The function of CaSR in calcium homeostasis:

The dominating role of CaSR is reacting to changes in free extracellular calcium levels and keep it in a narrow range of 1.1 to 1.4 mM. It is vital for a variety of biological functions. At low blood calcium levels, CaSR is inactive, PTH is secreted from thyroid C cells and stimulates calcium reabsorption in the kidney, calcium release in the bones and 1,25 - vitamin D₃ synthesis, which follows calcium uptake in the intestines to elevate the systemic calcium levels. But if the calcium concentration is too high, CaSR gets activated and inhibits PTH secretion until it drops again. Its fundamental role is to protect the body from hypercalcaemia (3).

Parathyroid hormone is a potent bone anabolic agent. But only when administered by daily injections of short durations. A prerequisite for an anabolic response. In theory, calcilytics could be used for the treatment of osteoporosis patients. Orally administered calcilytics as microemulsion induce higher levels of PTH and fall back to baseline within minutes. Osteoporosis is characterised by bone mass loss, higher risk of bone fracture and bone fragility. The equilibrium of bone formation by osteoblasts and bone deterioration by osteoclasts is out of balance, leading to osteoporosis. Currently, treatment is possibilities with medicines like bisphosphonate to stop further bone mass loss. However, a lot of patients are diagnosed when the disease is already in an advanced stage. There would be a need for bone anabolic, and calcilytics appear to be one. Full-length PTH or N-terminal fragment (teriparatide) is approved to stimulate bone reconstruction. The

administration of these peptides is required daily. Stimulation of endogenous PTH secretion can be achieved by negative modulation of the cell surface receptor CaSR. Calcilytics mimic the state of hypocalcaemia. It is documented that elevated levels of PTH only results in higher bone mass if the doses are administered transiently.

The effect does not persist longer than 2-4 hours. Permanent elevated PTH levels stimulate osteoblasts and osteoclasts and lead to a higher bone turnover and not to increased bone mass. The negative allosteric modulator NPS-2143 is a calcilytic with a too long half-life by stimulating PTH secretion. Not useable for the treatment of osteoporosis. Ronacaleret, a derivate of NPS-2143, shows better kinetics. A proof of concept study showed that calcilytics stimulated bone formation markers in postmenopausal women but did not affect the resorption markers (4). New calcilytics with a different pharmacological profile could become potential drugs for treating osteoporosis and other bone- and mineral-related disorders such as hyperparathyroidism.

1.4 Ligands of CaSR

CaSR signalling is affected by a variety of ligands. The endogenous ligands of CaSR are extracellular calcium, magnesium, L-tryptophan and spermine. In addition, this receptor interacts with a series of di- and trivalent cations (potency rank $Gd^{3+} > Ca^{2+} > Mg^{2+}$) as well as polyamines (spermine > spermidine > putrescine) and drugs like etelcalcedite, and evocalcet (5).

Compounds that interact with CaSR are categorised as Type I or Type II ligands. Type I ligands bind to the orthosteric binding sites in the extracellular domain, and Type II ligands interact allosterically with CaSR in the transmembrane domain.

Depending on the modulators' activity, the type II ligands are positive (PAM) or negative allosteric modulators (NAM). PAMs decrease the setpoint of CaSR and receptor activation initiates at lower extracellular calcium concentrations. NAMs make receptor activation more difficult. As a result, a normal response requires higher extracellular calcium concentrations. PAMs and NAMs do not activate the receptor themselves, but they shift the calcium-response curve's potency either to

the left or the right (Figure 1.2) (2). Known allosteric modulators are listed on page 124.

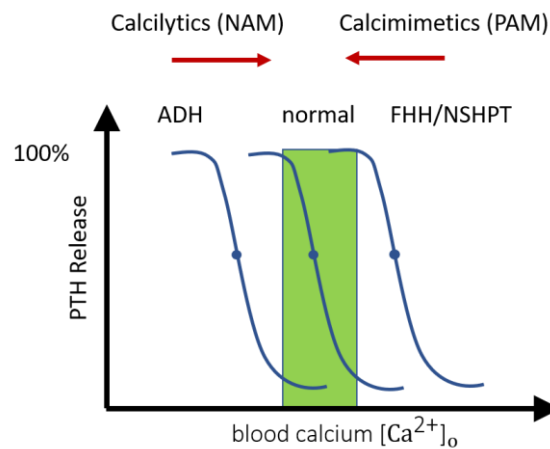


Figure 1.2: Calcium-response curve of CaSR; positive allosteric modulators (PAMs) cause a potency shift to lower calcium levels and negative allosteric modulators a shift to higher levels inducing a change of CaSR's setpoint

1.5 Structure of the extracellular calcium-sensing receptor (CaSR)

The structure of proteins is defined at four different levels. The primary structure is represented by a linear sequence of up to twenty different L-amino acids linked by peptide bonds (Figure 1.3).

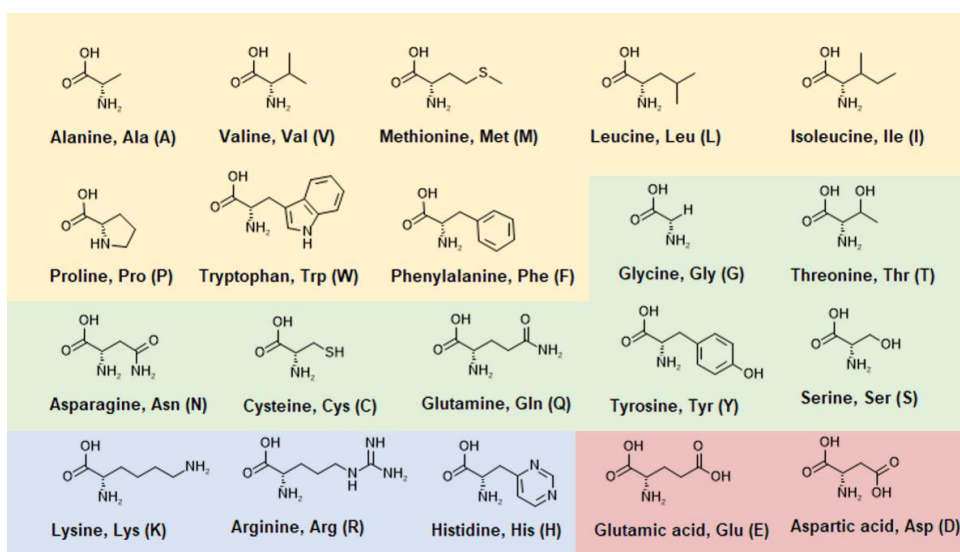


Figure 1.3: essential amino acids; yellow: hydrophobic; green: neutral; blue: alkaline; red: acidic

Local interactions of the backbone form secondary structure elements like α -helices, beta β -sheets or random coil. The tertiary structure is the functional form of a monomeric protein, including long-range interactions from the amino acid side chains. If more than one protein is required to achieve the functional form, the assembly is organised in a quaternary structure (6).

The primary sequence of the human extracellular calcium-sensing receptor (CaSR) is encoded in gene 3q13.3-21 and consists of 8 exons. Exon 2 to 7 express as 1078 amino acids long protein (7) (Figure 1.4). The functional form of CaSR is expressed as a homodimer and does not contain the 20 amino acid signal peptide at the N-terminus.

```
>sp|P41180|CASR_HUMAN
MAFYSCCWVLLALTWHTSA YGPDQRAQKKGDI ILGGLFPIHFGVAAKDQDLKSRPESVE C
IRYNFRGFRWLQAMIFAIEEINSSPALLPNLTLGYRIFDT CNTVSKALEATLSFVAQNKI
DSLNLDEF C N C SEHIPSTIAVVGATGSGVSTAVANLLGLFYIPQVSYASSSRLLSNKNQF
KSFLRTIPNDEHQATAMADIEYFRWNWVGTTAADDYGRPGIEKFREEAEERDI C IDFS
ELISQYSDEEEIQHVVEVIQNSTAKVIVVFSSGPDLEPLIKEIVRRNITGKIWLASEAWA
SSSLIAMPQYFHVVGGTIGFALKAGQIPGFREFLKKVHPRKSVHNGFAKEFWEETFN C HL
QEGAKGPLPVDTFRLRGHEESGDRFSNSSTAFRPL C TGDENISSVETPYIDYTHLRISYNV
YLAVYSIAHALQDIYT C LPGRGLFTNGS C ADIKKVEAWQVLKHLRHLNFTNNMGEQVTFD
ECGDLVGNYSIINWHLSPEDGSIVFKEVGYYNVYAKKGERLFINEEKILWSGFSREV PFS
N C SRD C LAGTRKGIIEGEPT C C F E C V E C PDGEYSDET DASA C N K C PDDFWSNENHTS C IA
KEIEFLSWTEPF GIALTLFAVLGIFLTAFLVGVFI KFRNTPIVKATNRE LSYLLLFSLLC
CFSSSLFFIGEPQDWT C RLRQ PAFGISFVLCISCILVKTNRVLLVFEAKIPTSFHRKWWG
LNLQ FLLVFLCTFMQIVICVIWLYTAPPSSYRNQELEDEIIFIT C HEGS LMALGFLIGYT
CLLAAICFFFAFKSRKLPENFNEAKFITFSMLIFFIVWISFIPAYAST YGKFVSAVEVIA
ILAASFGLLACIFFNKIYIILFKPSRNTIEEVR CSTAAHAFKVAARATLRRSNVSRKRSS
SLGGSTGSTPSSSISSKSNS EDPFPQPERQKQQQPLALTQQEQQQQPLTLPQQQRSQQQP
RCKQKVI FGS GTVTFSLSFDEPQKNAMAHNRNTHQNSLEAQKSSDTLTRHQPLLPQCGE
TDLDLTVQETGLQGPVGGDQRPEVEDPEELSPALVVSSSQSFVISGGGSTVTENVVNS
```

Figure 1.4: Primary sequence of the human extracellular calcium-sensing receptor (hCaSR) in FASTA format retrieved from UniProt (8) (<http://www.uniprot.org/uniprot/P41180.fasta>), cyan: signal peptide, yellow: cysteine part of a cysteine-bridge, green: transmembrane helices

The receptor consists of four domains. Outside the cell is the bilobed extracellular domain (ECD) also called 'Venus flytrap' because of its shape. The extracellular domain is attached to a heptahelical transmembrane domain (7TM), representing all members of the GPCR superfamily. It connects to the intracellular domain (ICD) interacting predominantly with the G-proteins G_i and G_q (Figure 1.5). The ECD is 612 amino acids long and includes the orthosteric binding sites for calcium and the

cysteine-rich domain (CRD). The two protomers of CaSR covalently connect on top of the homodimer via disulphide bridges at the cysteines 129 and 131. The ECD has 11 potential N-glycosylation sites, and for normal receptor expression, the receptor gets partially glycosylated in the Golgi apparatus. Orthosteric ligands interact with the ECD and the allosteric with the 7TM. The ECD's crystal structure revealed four calcium-binding sites and binding sites for L-tryptophan and phosphate (9).

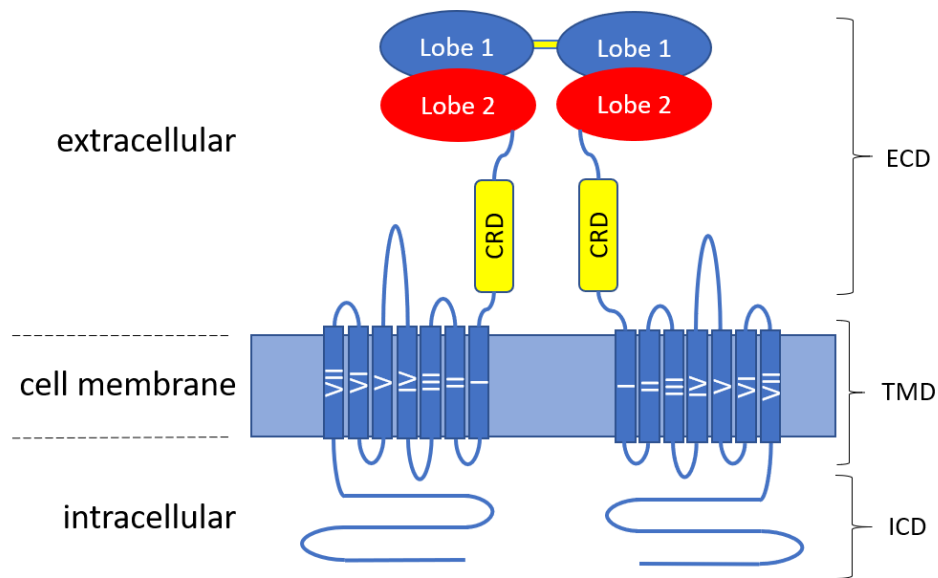


Figure 1.5: Schematic representation of CaSR's structure in the inactive state (homology model created with YASARA); ECD: extracellular domain, CRD: cysteine-rich domain, TMD: transmembrane domain, ICD: intracellular domain

1.6 G-protein coupled receptor (GPCR)

CaSR's seven-transmembrane domain makes it a member of the G-protein coupled receptor (GPCR) superfamily. The GRAFS classification scheme divides this superfamily into five families. The most prominent family is the rhodopsin (R) or class A family. Besides, there is the adhesion (A)/secretin (S) or class B family, glutamate or class C family and the frizzled/smoothed family/taste type 2 family. CaSR is a Class C GPCR members besides eight metabotropic glutamate receptors, three taste type-1 receptors, two GABA-B receptors and eight orphan receptors with unknown endogenous ligand (Figure 1.6) (10).

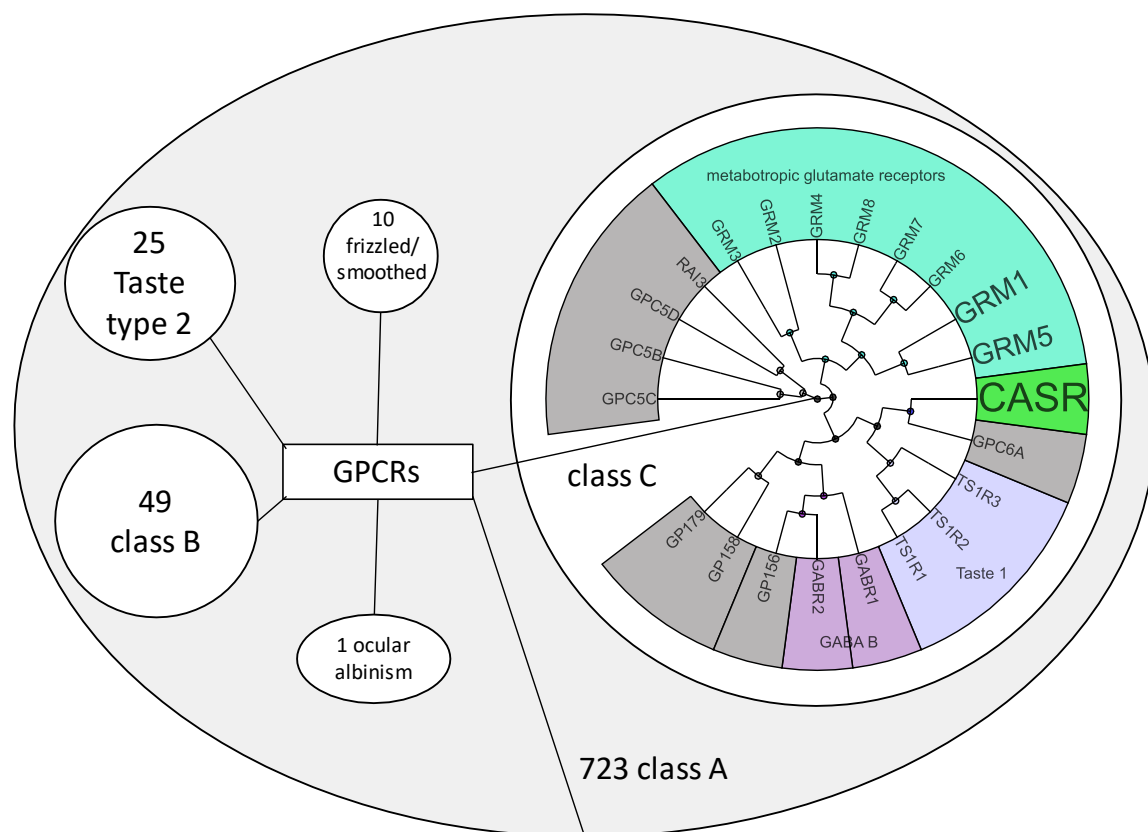


Figure 1.6: GPCRs overview: The human GPCRs are clustered into different protein families; class A, B, C, frizzled/smoothed, taste type 2 and ocular albinism; The GPCR class C and the calcium-sensing receptor (green); (phylogenetic tree created with 831 GPCR entries from UniProt (8) at GPCRdb with the UPGMA distance calculation method (11))

GPCRs are integrated into the cell membrane and transmit extracellular ligand stimulated signals into a variety of cells. Many diseases are related to GPCR malfunction, and that is why over 40% of all FDA approved drugs on the market are targeting this receptor family (12).

1.7 CaSR signalling

CaSR can stimulate an extensive and complex array of signalling pathways upon activation. Its primary signalling pathways are inhibiting the adenylate cyclase via Gi/Go proteins and the stimulation of the phospholipase C (PLC) and phospholipase D (PLD) pathways through G_q/G₁₁ and G_{12/13}, respectively. Besides, CaSR also transduces through MAPK pathways like ERK1/2, p38 MAPK, JNK, PI-3K, PI-4K, Rho and the EGF receptor (5).

In general, the G-proteins are heterotrimeric proteins which consist of a G_{α} subunit and a $G_{\beta\gamma}$ subunit. As depicted in Figure 1.7, the G_{α} -subunit binds the guanidine nucleotide GPD in the resting state, and the heterotrimeric complex is bound to a GTP exchange factor (GEF) like CaSR (1). Upon activation of GEF by an agonist, a conformational change is induced, and the GDP in the G_{α} -subunit is exchanged by GPT (2). The G_{α} subunit and $G_{\beta\gamma}$ subunit become activated and stimulate diverse signalling pathways (3). After the receptor deactivation, the GTP is exchanged again by GDP (4) and the complex forms again the resting state (5) (13).

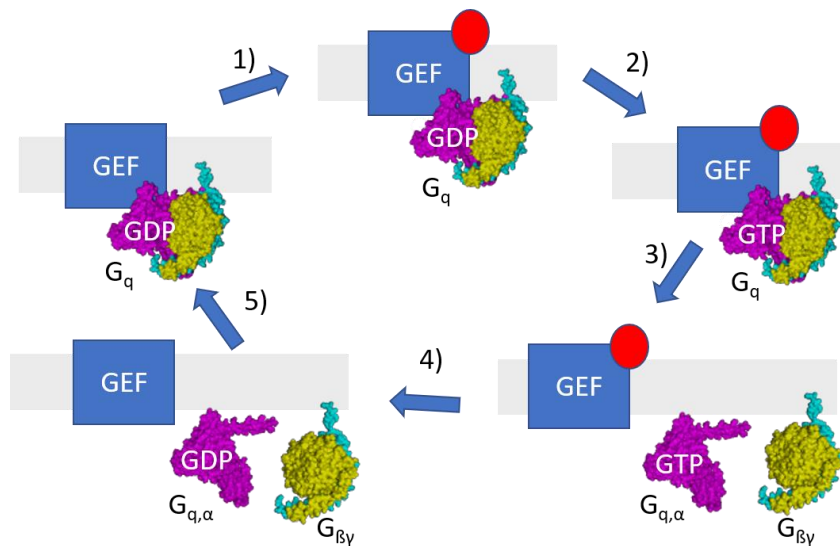


Figure 1.7: G-protein activation and deactivation cycle

1.8 CaSR related pathophysiology

The pathophysiology of CaSR is mainly characterised by gain-of-function or loss-of-function mutations.

1.8.1 Diseases related to gain-of-function mutations

Autosomal dominant hypocalcaemia (ADH type 1) is based on gain-of-function mutation within CaSR. In the case of ADH type 1, biochemical tests show low calcium and PTH levels, hypercalciuria with nephrolithiasis and nephrocalcinosis, intracranial calcifications and seizures. The pathophysiological cause is a lowered set point for the calcium level. Most activating mutations are identified in the ECD of CaSR, which cause an increased affinity to calcium. Most patients are asymptomatic (14).

Sporadic idiopathic hypoparathyroidism is suggested with activating mutation and ranges from mild to severe cases (2).

Barter syndrome type V is characterised through insufficient reabsorption of sodium and chloride in the kidney. This disease causes hypocalcaemic metabolic alkalosis, high levels of renin and aldosterone (14).

Besides the activating mutations of CaSR, there are also activating autoantibodies causing sporadic hypoparathyroidism in patients with autoimmune polyendocrine syndrome type 1 (15).

1.8.2 Diseases related to loss-of-function mutations

Familial hypocalciuric hypercalcaemia (FHH1) is an autosomal dominant disease caused by heterozygous loss-of-function mutation in CaSR. Most of the patients are asymptomatic. They have hypercalcaemia but a normal urinary calcium excretion (2).

Neonatal severe primary hyperparathyroidism (NSHPT) is a rare disease characterised by extreme hypercalcaemia, severe hyperparathyroidism related to parathyroid hyperplasia, and skeletal demineralisation, respiratory distress and hypotonia. NSHPT is the homozygous form of FHH (2).

1.9 Drug discovery and development of calcilytics and calcimimetics overview

The drug discovery process for the extracellular calcium-sensing receptor had a bumpy start because the receptor's very existence was controversial before it was successfully cloned in 1993s (1). Subsequent, it was possible to design assays for high throughput screening. One of the first assays, which is still in use, is based on the measurement of Gq/PLC/IP₃ induced intracellular calcium mobilisation in CaSR transfected HEK-293 cells or CHO cells. Besides, a more selective [³H]IP radio- or IP1 fluorescence immunoassay is used to detect CaSR induced IP1 accumulation (16,17). In the beginning, the screening was low-throughput, but it changed after the implementation of the Fluorescent Imaging Plate Reader (FLIPR™) technology. In the 1980s, the starting point to design a drug was very difficult because the natural ligands of CaSR are calcium ions. The question was asked what acts like calcium but is not calcium and the answer was spermidine (EC₅₀ = 150µM) which showed activity apart from ions. Other polyamines were identified to act on CaSR like putrescine and hexacyclen (EC₅₀ = 20µM). Primarily orthosteric acting ligands were categorised as Type I calcimimetic or calcilytics, and compounds that bind to the allosteric binding site are part of category Type II. The discovery of allosteric modulators for CaSR began with the screening of approved drugs and resulted in identifying the non-selective calcium channel blocker fendiline. Nowadays, a variety of different compound classes have been discovered. On the side of calcimimetics, there are phenyl- and naphthylamines (*Table 4.4, Table 4.5, Table 4.6, Table 4.9, Table 4.14, Table 4.22, Table 4.31, Table 4.32* starting on page 126) including tecalcet, cinacalcet and evocalcet; arylalkylamines (*Table 4.34, Table 4.35, Table 4.36*); urea-based compounds (*Table 4.28, Table 4.29, Table 4.30*); and 1-(benzothiazole-2-yl)phenylethanols like

AC-265347 (Table 4.27). The main indication is for secondary hyperparathyroidism (SHPT) (18).

On the side of the calcilytics, the compounds are amino alcohols (Table 4.10, Table 4.11; Table 4.19) like ronacaleret and encaleret; quinazolinones (Table 4.2, Table 4.12); pyrimidiones (Table 4.3; Table 4.16; Table 4.17); pyrimidines/pyridines (Table 4.8); benzimidazoels (Table 4.15); and diaminocyclohexanes (Table 4.7) including calhex (18).

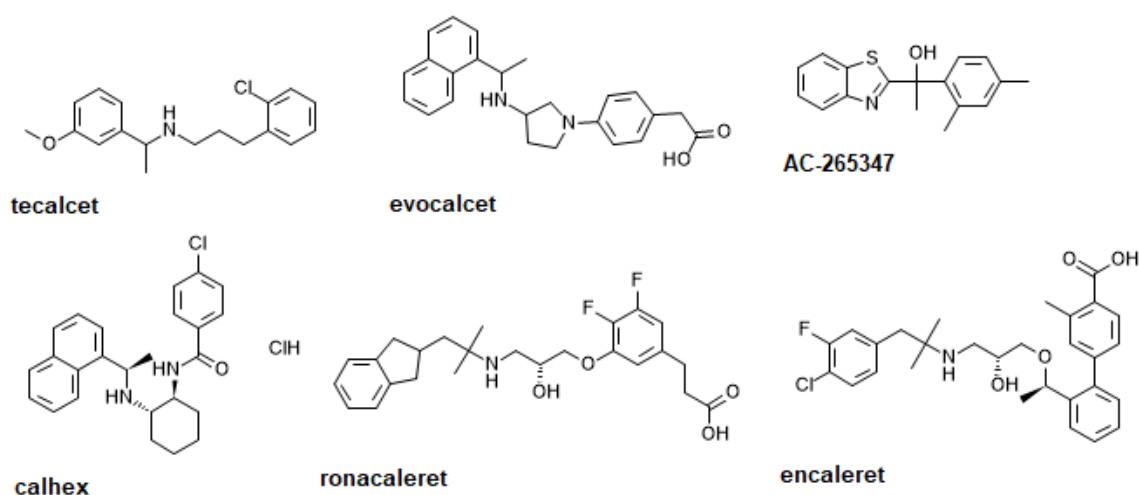


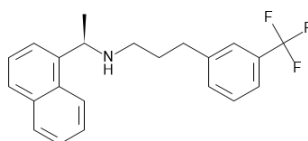
Figure 1.8 allosteric modulators of CaSR; first row: calcimimetics type II; second row: calcilytics

An overview of the different scaffold is presented in the section of the validation set on page 124.

1.10 Which therapeutics are on the market acting on CaSR

At the time, there are three drugs approved for the market: Cinacalcet, Etecalcetide and Evocalcet.

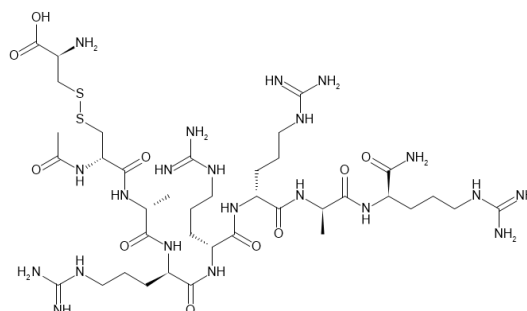
Cinacalcet (Mimpara®, Sensipar®):



[(1R)-1-(naphthalen-1-yl)ethyl]({3-[3-(trifluoromethyl)phenyl]propyl})amine

Cinacalcet is an allosteric modulator approved for secondary hyperparathyroidism (SHPT) in patients with chronic kidney disease (CKD) on dialysis, for hypercalcemia because of parathyroid carcinoma and for patients who are unable to undergo parathyroidectomy suffering from severe hypercalcaemia including primary hyperparathyroidism (19). Adverse drug effects for cinacalcet include nausea, vomiting, hypocalcaemia and hypoparathyroidism (20,21).

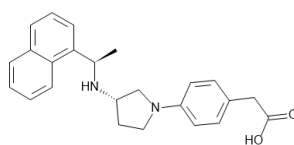
Etelcalcetide (Parsabiv®)



N-acetyl-D-cysteinyl-D-alanyl-D-arginyl-D-arginyl-D-arginyl-D-alanyl-D-argininamide (1->1')-disulfide compound with L-cysteine

Etelcalcetide acts as direct CaSR agonist. It is an octapeptide containing 7 D-amino acids, has a higher potency than Cinacalcet and is hypothesised to have less adverse effects. However, studies showed an increase in frequency and degree of QT interval prolongation (20).

Evocalcet



2-{4-[(3S)-3-[(1R)-1-(naphthalen-1-yl)ethyl]amino]pyrrolidin-1-yl]phenyl}acetic acid

Evocalcet is also a calcimimetic used for secondary hyperparathyroidism. It causes fewer gastrointestinal symptoms as cinacalcet or etelcalcetide.

The biggest issue with the currently approved drugs is that they cause gastrointestinal intolerance and reduce the patient's adherence to using the drug (22).

2 Results:

This project aims to discover novel allosteric modulators for the extracellular calcium-sensing receptor (CaSR). The hypothesis is that in-silico screening of prepared compound libraries will lead to potential candidates. The top-scoring compounds are evaluated by in-vitro experiments determining allosteric activity.

Chapter I: Virtual screening for allosteric modulators of the extracellular calcium-sensing receptor (CaSR)

In this thesis, the drug discovery process of new allosteric modulators started with the virtual screening of approved drugs. In the best-case scenario, potential compounds could be used for drug repurposing and the treatment of CaSR related disease.

The virtual screening is performed with molecular docking experiments using AutoDock VINA (23) within the YASARA software package (24). Molecular docking is a structure-based method that requires a three-dimensional model of CaSR and a compound library for screening. Docking experiments provide predictions about the binding affinities between the receptor and the screening compounds. The predicted binding affinities get ranked for selecting potential candidates, and the obtained complexes are analysed for CaSR-ligand interactions. The top-scoring compounds are subject to experimental evaluation in an IP-1 Gq assay.

In the first attempt, the free accessible online database DrugBank (version 5.0.10) was selected as a compound library for the docking experiments. This database lists molecular information about drugs, their chemical structure, pharmacology, interactions, products on the market, chemical identifiers, references, clinical trials and properties (25). The database contains 8486 small molecules and 1782 biologicals, but only the small molecules are used for screening (Figure 2.1).

DRUGBANK - SMALL MOLECULES

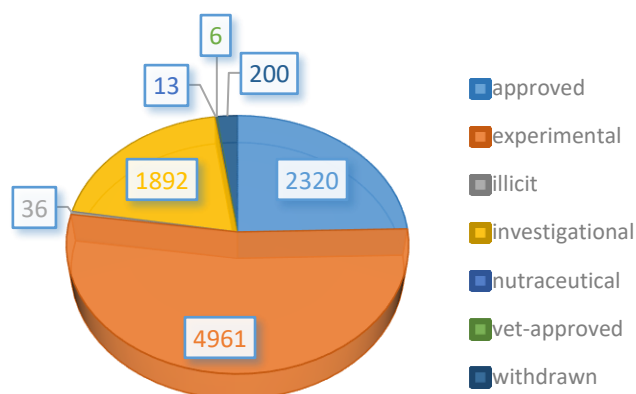


Figure 2.1 DrugBank database version 5.0.10 with 8486 small molecule entries

Besides, a model of CaSR is required for the docking experiments, and there are several methods to obtain it. Usually, a protein structure is determined either experimentally or by protein structure predictions using computational methods. Experimental methods include X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy and cryo-electron microscopy (6).

X-ray crystallography is a method to determine a protein's structure by applying X-rays to a protein crystal. It means the protein should be able to form a crystal, but there is not a universal crystallisation technique for all the proteins. It depends on many parameters like temperature, pH, vibrations, protein and precipitant concentration, flexibility, and protein purity to find the perfect condition. Supersaturation is the key to form the protein crystals either of the precipitant or the protein. The required X-rays are created by accelerated electrons (synchrotron) that collide with an anode metal. The protein rotates in a goniometer machine and is bombarded with X-rays that scatter at the electrons of the protein. A diffraction pattern is a result containing structural information. Atomic resolution is possible because the protein is present in the crystal in repeated symmetric units, amplifying the signal and making it detectable. The diffraction pattern from X-ray crystallography follows Bragg's law of constructive coherent interference (*Equation 1*) (6).

$$n\lambda = 2d \sin \theta$$

Equation 1

The X-ray wavelength is described by λ , the interplane distance is d , and the angle between the incident ray and the planes is θ . Charged coupled devices (CCD) are used to detect the X-rays, and diffraction intensity data is recorded. The required phase information is obtained by methods like molecular replacement or isomorphous replacement. Modern software is used to process the amplitude and phase information with a fast Fourier transform method. The outcome is an electron density map showing the 3D contour of the protein. A quality measure of the model is defined by the R-value, which represents higher quality if it is closer to zero. This method gives the highest atomic resolution for proteins, but it also inheres several limitations. This method's prerequisite is a soluble protein, crystallisable and can be obtained in large quantity. It means that proteins larger than 250 kDa are usually not accessible with this method (6).

Nuclear magnetic resonance spectroscopy (NMR) measures the absorption of electromagnetic radiation of H^1 , C^{13} or N^{15} nuclei in the presents of a strong magnetic field. Each of these isotopes has a detectable resonance radio frequency that shifts the nuclei from the ground to the excited spinning state. A resonance spectrum can be recorded by changing the electromagnetic frequency or the applied magnetic field. Depending on the conditions, the spectra give either structural information of near neighbouring atoms interconnected by covalent bonds or the proximity of sequential distant residues via the Nuclear Overhauser effect, if the nuclei are not further apart than 5 Angstrom. These techniques are used to reconstruct the 3D structure of the protein. This method requires soluble proteins not larger than 2-40 kDa, which reflects its limitation. The resolution is not so high compared to X-ray crystallography, but the dynamic behaviour can be investigated (6).

Cryo-electron microscopy utilises electron beams with a wavelength of 0.02 Angstrom to determine protein structure. This technique allows the visualisation of cell architecture and complex structures up to several mega Daltons. Recent advances in this field made it possible to solve AAV2 capsid (3,9 MDa) structure at a resolution of 1.8 Angstrom. The sample in solution is spread evenly over a carbon-based film and frozen in liquid nitrogen. The frozen sample is transferred to cryo-transfer holders, and several thousand images are taken with low electron

doses. The detected cryo-EM micrographs are computationally reconstructed to obtain a 3D structure (6).

Protein structures which have not been solved experimentally yet can be predicted by different computational methods including ab initio methods, homology modelling or threading (6).

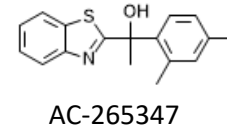
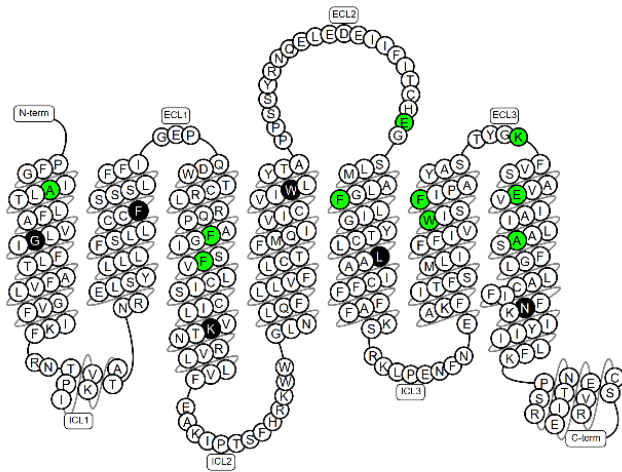
Ab initio methods use quantum mechanics to predict the protein structure, starting from the primary sequence without any reference template. Unfortunately, current equations in quantum mechanics cannot be solved for complex structures. Therefore, several assumptions are required for the prediction of proteins with not more than 150 residues. This method is used if no data are available about similar structures. But the accuracy is the bottleneck of these methods (6).

Homology modelling is based on the evolutionary principle that protein structures may be conserved which share high sequences similarity. The target protein sequence can be used to search the Protein Data Bank for homologous protein structures. The best match is identified by sequences alignment should be above 40% sequence identity. An initial model is built with insertion, deletions and residue substitutions from the template and optimised to the final model. It is a common method for protein structure prediction and requires a homologous published template (6).

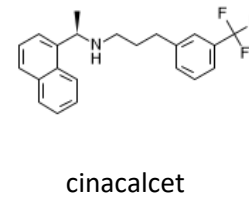
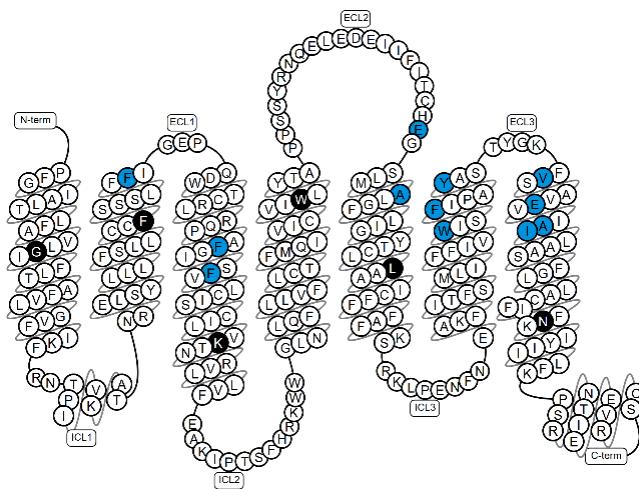
Threading methods are similar to homology modelling but without the need for homologous sequences. An evaluation parameter like a quasi-energy function is used to align the protein sequence to the backbone structure. This method is based on the principle that the number of unique folds is smaller than the number of proteins in nature. Subunits of proteins may fold independently of the protein itself. The prediction considers parameters like the amino acid position, the secondary structure, solvent accessibility and the proximity of sequential distant amino acids (6).

The extracellular calcium-sensing receptor is a transmembrane protein composed of three domains, a big extracellular domain, the heptahelical transmembrane domain and the intracellular domain. A prior mutagenesis study reveals the binding pocket of known allosteric modulators in the upper part of the transmembrane region (Figure 2.2) (26).

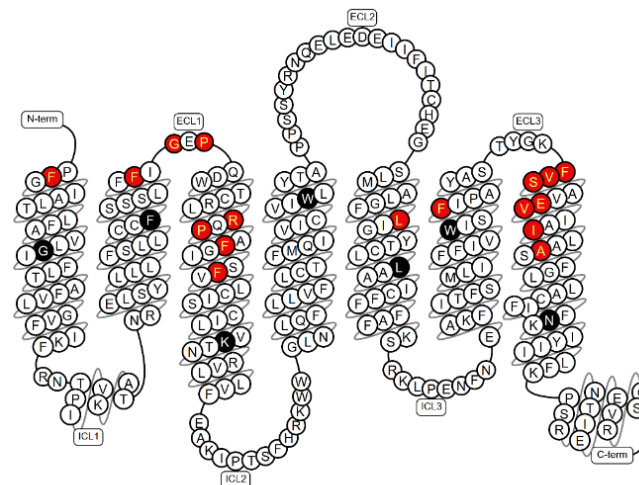
A



B



C



The study of Leach et al. presents the interaction profile of two positive allosteric modulators (AC-265347 and cinacalcet) and one negative allosteric modulator NPS-2143. The alanine scan shows the residues F612, A615, F684, F688, E767, F775, L776, W818, F821, K831, E837, A840 and A844 are relevant for the interaction with AC-265347 (Figure 2.2-A). Cinacalcet interacts with the residues F612, F668, G670, P672, R680, P682, F684, F688, L776, F821, F832, V833, S834, E837, V838, I841 and A844 (Figure 2.2-B). The alanine mutations reveal for NPS-2143 a significant change in binding affinity for the residues F668, F684, F688, E767, A772, W818, F821, Y825, V833, E837, I841 and A840 (Figure 2.2-C). The modulators' differences in binding affinity were determined with Fluo-4 in an intracellular calcium mobilisation assay using full-length c myc CaSR expressed in FlpIn HEK293 TRex cells. The conclusion of this study is that the binding profiles of different allosteric chemotypes are overlapping, but they are not identical (26).

Based on this study's results, it was evident that a model of CaSR requires at least the transmembrane to perform virtual screening. An experimentally determined structure was searched at the Protein Data Bank. However, the hit list contained only X-ray structures of CaSR's extracellular domain in the active (PDB: 5K5S) and inactive state (PDB: 5FBK, 5FBH, 5K5T) (9,28). Under these circumstances, the decision was made to use protein structure prediction methods to obtain a model of CaSR.

2.1 Structure prediction with I-TASSER

The state-of-the-art online server I-TASSER was used for protein structure prediction of CaSR. It is the top-scoring online protein structure prediction server according to the CASP experiments VII to XIII (29–35). The Critical Assessment of protein Structure Prediction, CASP, is held every two years since 1994. The goal is to determine the state-of-the-art protein structure prediction tools. Researchers worldwide participate in the competition and are provided with the primary sequence of unpublished protein structures. After the competition is over, the predictions are compared to the experimentally determined structures. The results are ranked and published in the categories contacts in protein structures, template-free modelling, template-based modelling, structure refinement, accuracy estimation, protein assemblies and data assisted modelling (35).

I-TASSER is an Iterative Threading ASSEmbly Refinement method consisting of four main steps (36). In the first step, structural templates are identified by the meta-threading server LOMETS through a nonredundant structure library. It generates three-dimensional models based on the input sequence and combines the information of 11 threading programs (CEthreader, FFAS3D, HHpred, HHsearch, MUSTER, Neff-Muster, PPAS, PRC, PROSPECT2, SP3 and SparksX) (37). The theory behind these methods is a profile-to-profile alignment, and each program focuses on a different structural feature. These methods are combined with increasing the detection of templates. The regions of the aligned template are clustered into threading-aligned and threading-unaligned. In the second step, the aligned fragments are extracted for the full-length model and reassembled from the templates. The unaligned parts are calculated by *ab initio* folding. Replica-exchange Monte Carlo simulations are performed for the reassembly and folding step and supported by an optimised knowledge-based force field. The force field contains information about generic statistical potentials, hydrogen-bonding networks and threading-based restraints from LOMETS. In the third step, the structures are clustered, and the conformation with the lowest energy becomes the starting point of a new assembly simulation. The final models are generated after the structures undergo a two-step energy minimisation approach. In the end, the

quality of the models is described by a confidence score (C - score) ranging from -5 to 2 (Equation 2).

$$C - score = \ln \left(\frac{M/M_{tot}}{\langle RMSD \rangle} * \frac{1}{9} \sum_{i=1}^9 \frac{Z_i}{Z_i^{cut}} \right) \quad \text{Equation 2}$$

In the C-score calculation, M represents the multiplicity of structures in the SPICKER cluster. The SPICKER program clusters different lowest free-energy conformations that result from Monte Carlo simulations. M_{tot} is the total number of structures used during the clustering. $\langle RMSD \rangle$ is the average deviation of the used decoys relative to the centroids of the cluster. The Z values correspond to the highest Z-score of the threading program (i) in LOMETS and the defined cutoff Z^{cut} of the same LOMETS program (12,36). Higher values express better models. The score's foundation is the significance of threading alignments and the density of structure clustering (36). The program ResQ (38) determines the quality per residue and the B factor of the target. The fourth step includes the functional annotation of the best models. Functional templates are identified from the BioLip database, and the information is extracted about ligand-binding sites, enzyme commission and gene ontology (36,39).

The full-length extracellular calcium-sensing receptor's protein structure prediction was executed with the FASTA sequence from UniProt (Figure 1.4) as input at the I-TASSER website (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>). The calculation was executed with default settings, finished after 40 hours and generated five models. In the process, I-TASSER identified the extracellular region of the metabotropic glutamate receptor 3 (PDB: 2E4X, 2E4U, 2E4V), the transmembrane domain of the glutamate receptor 1 (PDB: 4OR2), the heterotetrameric NMDA receptor ion channel (PDB: 4PE5), the extracellular domain of CaSR in the active state (PDB: 5K5S), the exodeoxyribonuclease V from Escherichia coli (PDB: 1W36-B) and the extracellular domain of CaSR in the inactive state (PDB: 5K5T) as top-ten ranking threading templates (Table 2.1).

Table 2.1 Threading templates identified by LOMETS from the PDB library and used for structure prediction by I-TASSER; ident. 1: sequence identity as the ratio of the aligned regions between the threading templates to the query sequence ranging from 0 to 1; ident. 2: sequence identity as the ratio of the whole template chain to the query sequence ranging from 0 to 1; Cov: coverage representing the number of aligned residues compared to the whole query sequence; Norm. Z-score: normalised Z-score of the threading alignments, good alignments correspond to a Z-score >1, thread. pro.: threading program of the meta-server LOMETS giving the result

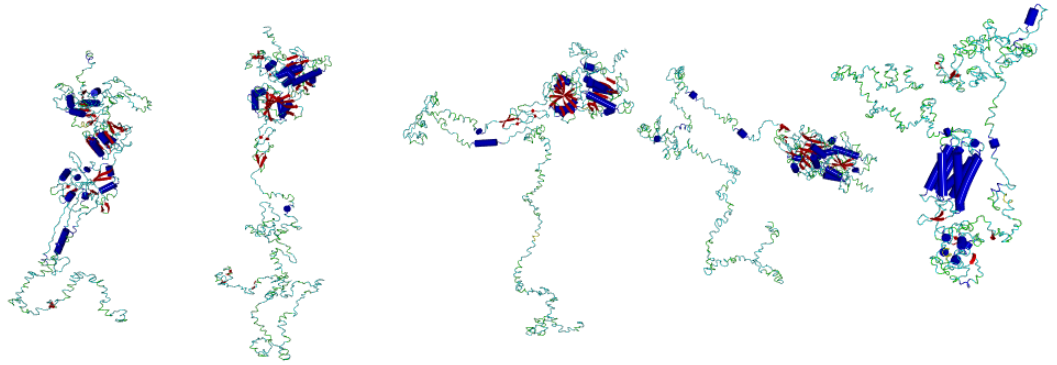
Rank	PDB Hit	Ident. 1	Ident. 2	Cov	Norm. Z-score	thread. pro.
1	2e4xB	0.35	0.18	0.47	2.25	MUSTER
2	4or2A	0.25	0.09	0.33	4.43	FFAS03
3	2e4uA	0.33	0.18	0.47	4.66	FFAS-3D
4	4pe5A	0.13	0.17	0.63	2.37	Neff-PPAS
5	2e4vB	0.34	0.18	0.47	2.73	SPARXS-X
6	4or2A	0.30	0.07	0.24	4.28	pGenTHREADER
7	5k5sA	1.00	0.49	0.49	3.65	HHSEARCH2
8	1w36B	0.08	0.18	0.77	2.22	PROSPECT2
9	5k5tA	1.00	0.53	0.54	5.72	
10	5k5sA	1.00	0.49	0.49	3.05	HHSEARCH I

The resulting sequence alignment in Figure 2.3 shows multiple alignments for template 4OR2 and 5K5S. Templates 1, 3, 5, 7, 9 and 10 cover the first 600 amino acids representing the extracellular domain. The structure of the metabotropic glutamate receptor 1 (PDB: 4OR2-A) was used in two different alignments (2 and 6). It matches the residues in the region 500 to 860 and covers the heptahelical domain of CaSR at a low sequence identity of 30%. Template number 4 is aligned with the extracellular and transmembrane domain of CaSR, and template 8 covers the whole sequence but has the lowest local alignment of all (Ident. 1: 0,08).



Figure 2.3 *HHHHH alignment results of I-TASSER including the top-scoring threading templates with the primary sequence of CaSR; sed. str.: secondary structure prediction of CaSR sequence with regions denoted as random coil (C), helix (H) or beta-sheet (S); colouring scheme: ClustalX (40)*

A visual inspection of the full-length models reveals disordered structures and the quality of the models is quite bad with c-scores below -2.88. The c-score correspond to the estimated quality of the resulting structure. It ranges from -5 to 2, and higher values represent better models.



model 1	model 2	model 3	model 4	model 5
c-score: -2,89	-2,88	-2,94	-2,98	-3,06

Figure 2.4 I-TASSER results of predicted full-length CaSR models, c-score: confidence score representing the model quality in the range between -5 to 2 where better models have higher values

Model 5 shows a heptahelical domain but has the worst c-score of all and seems not reliable. The extracellular domain of CaSR is solved as an X-ray crystal structure and is part of the threading templates, but the structure is only partly present in model one to four.

The estimated local accuracy provides another evaluation parameter. I-TASSER estimates the local accuracy by support vector regression, and it depends on the coverage of the threading alignment, the divergence to the I-TASSER simulation decoys, the secondary structure prediction and the solvent accessibility. A benchmark test of the local accuracy shows an average error of 2.21 Angstrom and gives a Pearson correlation coefficient of 0.7. The local accuracy increases by higher threading alignment coverage, in the regions of helices and sheets, and for buried residues at a threshold of 25% (38).

The models' residue-specific quality is estimated based on the root mean square deviation (RMSD) in Angstrom (Equation 3).

$$RMSD = \frac{\sum_{n=1}^n R_i * R_i}{n} \quad \text{Equation 3}$$

R_i represents the connecting vector between the atom pairs n in space (24). The models' estimated local accuracy diverges at best by 4 Angstrom from the native structure (Figure 2.5). On average model 1 differ by 7.2 Angstrom, model 2 by 10.4 Angstrom, model 3 by 10.3 Angstrom, model 4 by 10.9 Angstrom and model 5 by

11.2 Angstrom. The loops have an RMSD of 20 Angstrom and more around residue 127 and 380 in the extracellular domain, although there are atomic coordinates in the crystal structure records (PDB: 5K5S, 5K5T). The best-estimated quality has model 1, but it does not show a seven-transmembrane domain. For all the models the lowest deviation has the extracellular domain to residue 600, excluding some areas for the loops. The deviation increases at the transmembrane domain region between residue 600 and 860 to an average of 10 Angstrom. Model 5 shows a heptahelical domain, but the local accuracy is about 6 Angstrom. The worst local accuracy has the intracellular domain ranging from residue 860 to 1078 and deviates on average by 20 Angstrom. This circumstance results from the alignment with only one template (PDB: 1w36-B) to this region that shares 18% sequence identity with the whole protein.

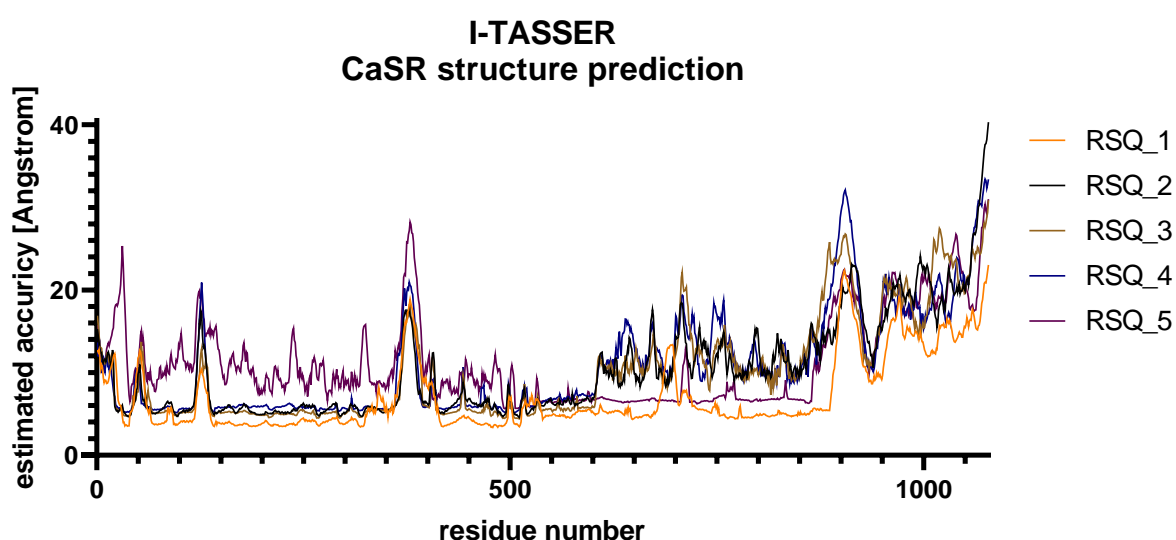


Figure 2.5 Residue-specific quality (RSQ) estimating the deviation between the predicted model and the native structure resulting from I-TASSER

The conclusion is that the created CaSR models do not have the correct fold because of the big deviation of accuracy to an estimated native structure. The models from I-TASSER are not reliable and cannot be used for docking experiments. As a consequence, a different method is required to predict the structure of CaSR.

2.2 Structure prediction with GPCR-I-TASSER

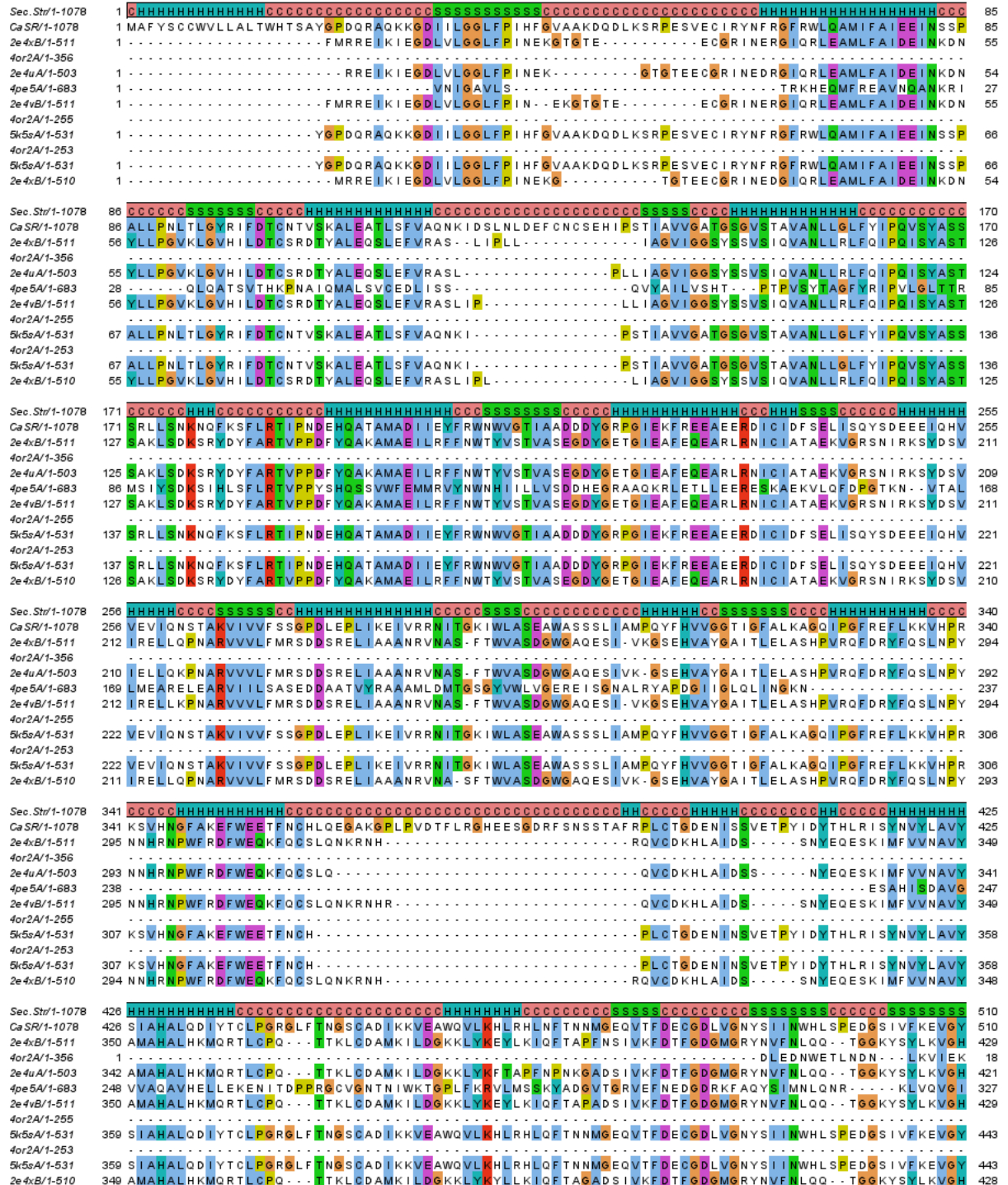
The next structure prediction attempt involved the use of the hybrid structure assembly algorithm GPCR-I-TASSER. This online server is the extended version of I-TASSER and appears to be more appropriate to generate a model of the class C GPCR CaSR. The improvement does include the use of a GPCR-specific database (GPCR-RD). This database contains data about experimental contacts and helix orientation gathered by literature and data mining to increase the structural assembly's accuracy. Also, in the absence of a homology template, the structure is predicted via ab initio folding method from scratch. The prediction of a new model is supported by new GPCR and transmembrane specific energy functions as part of the I-TASSER force field. It should improve the structural assembly and the refinement of GPCR models generated by threading or ab initio predictions. The benchmark test of GPCR-I-TASSER showed that 20 out of 24 transmembrane helix assemblies were predicted with the correct fold (12).

The structure prediction procedure started with the FASTA sequence of CaSR (Figure 1.4) as input at the GPCR-I-TASSER online server website (<https://zhanglab.ccmb.med.umich.edu/GPCR-I-TASSER/>). The default settings were used for the prediction. After 40 hours, the calculation finished and generated five models. GPCR-I-TASSER identified the same threading templates as I-TASSER but without the exodeoxyribonuclease V from Escherichia coli (PDB: 1w36-B) and the extracellular domain of CaSR in the inactive state (PDB: 5K5T) (Table 2.2).

Table 2.2 Threading templates used for structure prediction by GPCR-I-TASSER; ident. 1: sequence identity as the ratio of the aligned regions between the threading templates to the query sequence ranging from 0 to 1; ident. 2: sequence identity as the ratio of the whole template chain to the query sequence ranging from 0 to 1; Cov: coverage representing the number of aligned residues compared to the whole query sequence; Norm. Z-score: normalised Z-score of the threading alignments, good alignments correspond to a Z-score >1, thread. pro.: threading program of the meta-server LOMETS giving the result

Rank	PDB Hit	Ident. 1	Ident. 2	Cov	Norm. Z-score	thread. pro.
1	2e4xB	0.35	0.18	0.47	2.25	MUSTER
2	4or2A	0.24	0.09	0.33	4.91	FFAS03
3	2e4uA	0.33	0.18	0.47	4.66	FFAS-3D
4	4pe5A	0.13	0.17	0.63	2.37	Neff-PPAS
5	2e4vB	0.34	0.18	0.47	2.73	SPARKS-X
6	4or2A	0.30	0.07	0.24	4.28	pGenTHREADER
7	5k5sA	1.00	0.49	0.49	3.65	HHSEARCH2
8	4or2A	0.29	0.07	0.23	4.49	FFAS03
9	5k5sA	1.00	0.49	0.49	3.05	HHSEARCH I
10	2e4xB	0.35	0.18	0.47	3.72	NEFF-PPAS

The threading templates are aligned multiple times to the primary structure of CaSR (Figure 2.6). However, the intracellular domain is not covered without the template 1w26-B and is predicted by ab initio calculations.



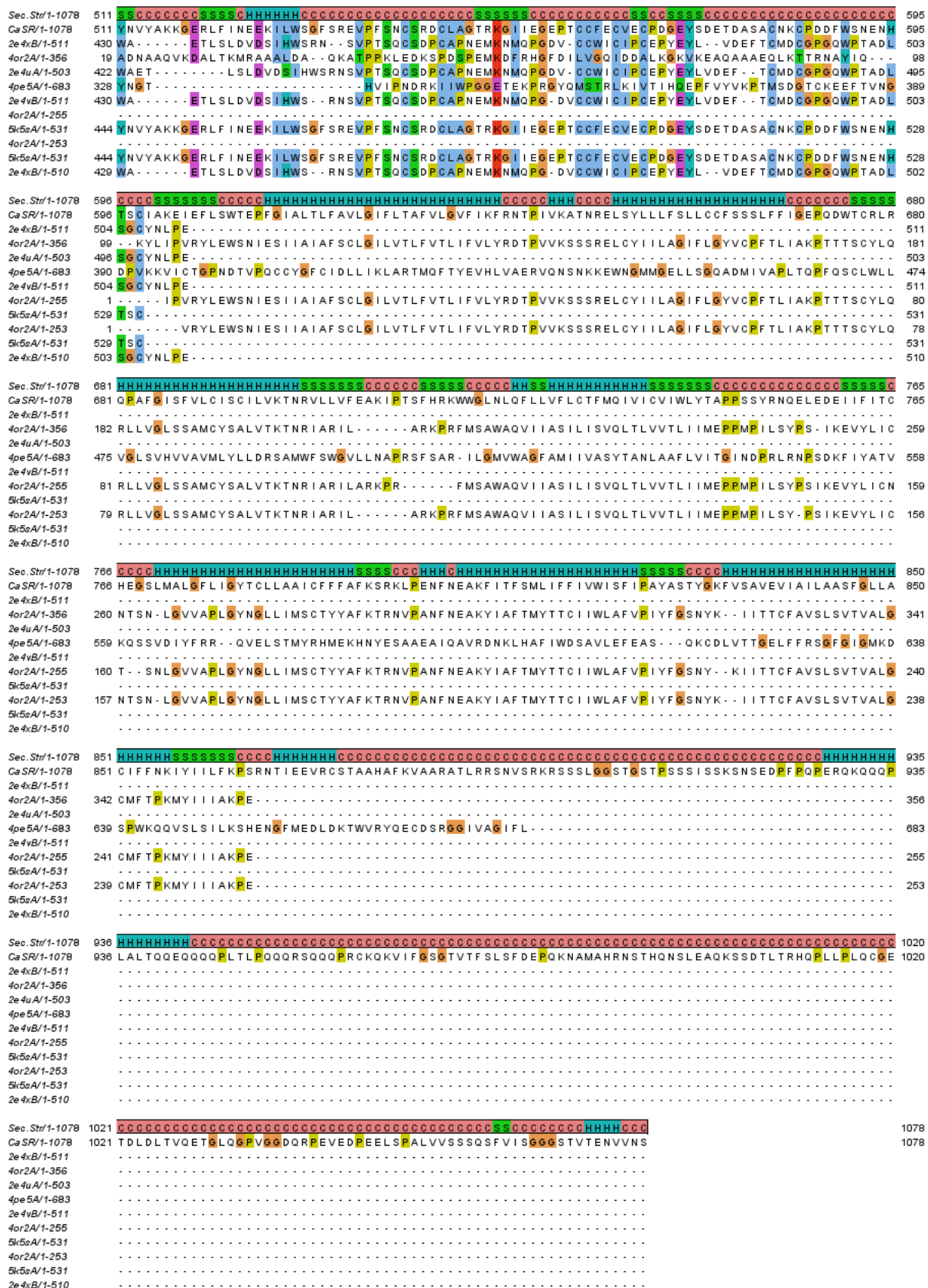


Figure 2.6 sequence alignment results of GPCR-I-TASSER including the top-scoring threading templates with the primary sequence of CaSR; sed. str.: secondary structure prediction of CaSR sequence with regions denoted as random coil (C), helix (H) or beta-sheet (S); colouring scheme: ClustalX (40)

The resulting models from GPCR-I-TASSER are disordered and have low c-scores between -3,73 and -2,72 (Figure 2.7). Model 1 to 4 contain the extracellular domain of CaSR from the 5K5S template but do not show a transmembrane domain. In the last model, the transmembrane appears similar to the I-TASSER result, but the model also has the worst confidence score.

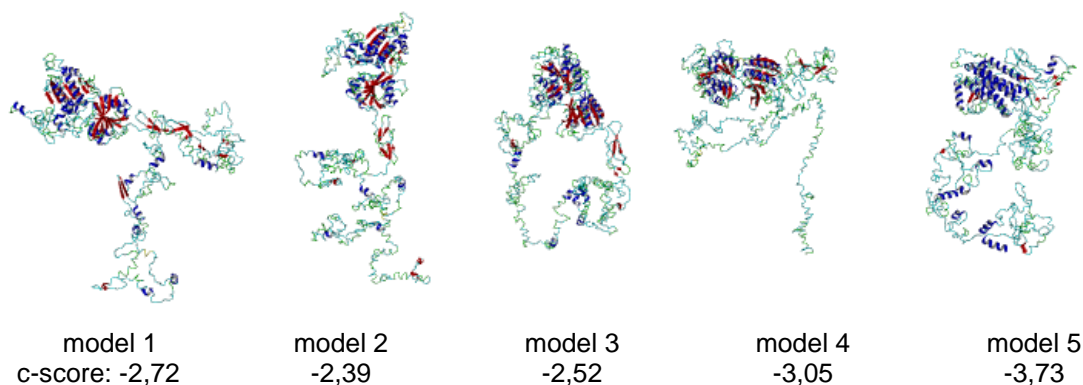


Figure 2.7 GPCR-I-TASSER results of predicted full-length CaSR models, c-score: confidence score representing the model quality in the range between -5 to 2 where better models have higher values

The models' estimated structural accuracy deviates at least 3,5 Angstrom from the native structure (Figure 2.8). Although model 1 has the highest estimated accuracy around 7 Angstrom on average, it seems not applicable for docking experiments without a seven-transmembrane. The same situation applies to model 2, 3 and 4. Model 5 does have a heptahelical domain, but with an average deviation of 11.9 Angstrom from the estimated native structure, it is not reliable.

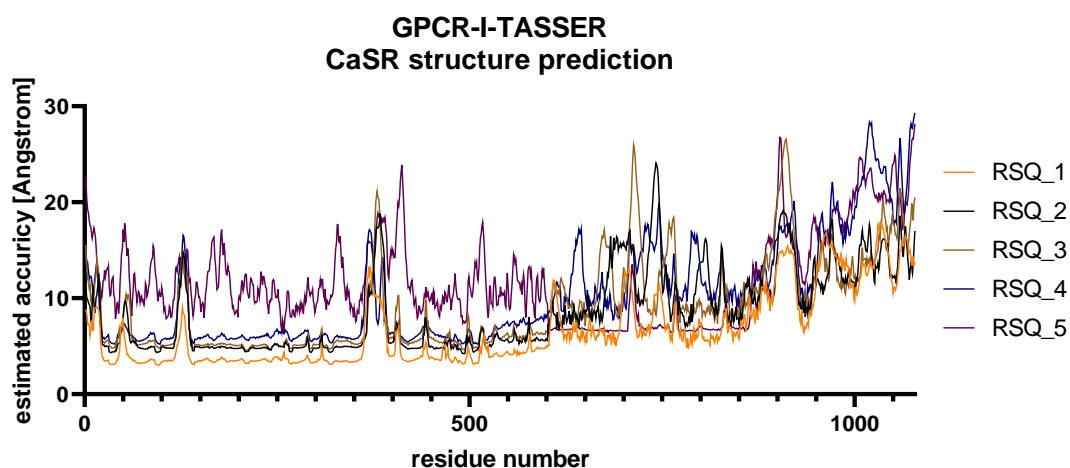


Figure 2.8 Residue-specific quality (RSQ) estimating the deviation between the predicted model and the native structure resulting from GPCR-I-TASSER.

Consequently, the predicted models from I-TASSER and GPCR-I-TASSER are not applicable, and the modelling process requires a different method.

2.3 Homology modelling of CaSR with YASARA

The virtual screening of DrugBank requires a structural model of CaSR to perform molecular docking experiments. The online threading servers I-TASSER and GPCR-I-TASSER could not generate a reliable model with the correct fold.

YASARA Structure was used to create the model of CaSR. YASARA's homology modelling protocol (www.yasara.org/hm_build.mcr) allows many manual interventions to guide the modelling process (page 114) (41). In the first attempt, the full sequence of CaSR was selected in FASTA format (Figure 1.4) to predict the three-dimensional structure with YASARA's homology modelling module (version 16.9.23.L.64). The experiment was executed with slow modelling speed to achieve the best prediction. This setting considers many templates and alignments, loops and side chains are thoroughly sampled; the modelling process runs simulated annealing minimisations and builds a hybrid model (41). Homology modelling was started without manually provided templates. Three PSI-BLAST iterations were performed with the input sequence to extract a position-specific scoring matrix (PSSM) from UniRef90 (42) and to search the Protein Data Bank (43) for hits. The resulting list contains only structures with a maximum E-value of 0.5. The five top-scoring templates were selected to start the modelling process. The oligomerisation parameter was left at the default setting of four which generated at most a tetrameric model. Each template was aligned five times to generate alignment variations. The loops were sampled fifty times to identify the best conformation and the termini were extended by up to ten missing residues.

The search for homology templates resulted in the identification of 58 hits (Table 2.3). YASARA selected four templates for homology modelling based on the parameter 'total score'. The templates include chain A of the human extracellular calcium-sensing receptor's extracellular domain in active form (PDB: 5K5S-A; released 2016-08-03), chain A of the human extracellular calcium-sensing receptor's extracellular domain in the inactive form (PDB: 5K5T-A; released 2016-08-03), chain A of the human calcium-sensing receptor's extracellular domain (PDB: 5FBK-A; released 2016-06-22) and chain A of the human class C G protein-coupled metabotropic glutamate receptor 1 (PDB: 4OR2-A; released 2014-03-19).

Table 2.3 YASARA's template hit list from homology modelling with CaSR's sequence; temp: template; Cov: target coverage; Res: resolution in Angstrom; total score = (BLAST alignment score) x (WHAT_CHECK quality score in the PDBFinder2 database) x (target coverage), the quality score range is between terrible (0.0) and perfect (1.0) (41,44,45); (*): template structure was discarded because the total score is less than 30% of the maximum and does not provide any additional residues to the alignment

Temp	Total score	BLAST E-value	Cov	PDB-ID	Res. [Å]	Temp	Total score	BLAST E-value	Cov	PDB-ID	Res. [Å]
1	600.14	0	49%	5K5S-A	2.60	*	2.18	3e-05	29%	5EWJ-B	-
2	591.52	0	54%	5K5T-A	3.10	*	1.96	0.006	32%	5KC9-A	-
3	536.13	0	45%	5FBK-A	2.10	*	1.95	2e-06	35%	1JDP-A	-
*	116.31	1e-155	46%	2E4U-A	-	*	1.82	1e-05	34%	3SAJ-A	-
*	115.10	1e-124	39%	5C5C-A	-	*	0.86	1e-06	9%	4TLM-D	-
*	106.02	5e-128	40%	4XAR-A	-	*	0.80	7e-07	7%	4Q6B-A	-
*	104.88	3e-127	40%	3SM9-A	-	*	0.77	0.019	6%	3IPC-A	-
*	104.76	1e-132	41%	1EWK-B	-	*	0.77	0.0002	7%	3TJE-F	-
*	103.14	6e-128	39%	3KS9-A	-	*	0.66	0.02	6%	3IP5-A	-
*	98.24	2e-120	39%	4XAQ-B	-	*	0.64	0.042	7%	1USG-A	-
*	90.74	2e-126	40%	3LMK-B	-	*	0.63	0.02	6%	3IP7-A	-
*	86.05	2e-120	40%	5CNJ-A	-	*	0.60	2e-07	10%	5IOU-B	-
*	70.18	9e-128	40%	2E4Z-A	-	*	0.59	0.5	9%	4C9E-H	-
4	59.32	4e-95	33%	4OR2-A	2.80	*	0.55	0.042	7%	1USK-C	-
*	31.17	7e-48	24%	4OO9-A	-	*	0.53	5e-16	9%	4UQQ-B	-
*	30.47	1e-47	24%	5CGD-A	-	*	0.51	6e-05	10%	4PE5-B	-
*	18.17	2e-19	35%	4MS3-B	-	*	0.51	0.019	6%	1QO0-A	-
*	17.77	2e-20	37%	4MR8-A	-	*	0.46	0.5	7%	4RSU-L	-
*	15.35	3e-10	32%	3QEK-B	-	*	0.38	6e-05	10%	5FXH-D	-
*	12.59	6e-12	31%	5EWM-A	-	*	0.36	0.018	6%	1QNL-A	-
*	10.60	5e-16	31%	3QLU-C	-	*	0.34	6e-05	10%	5FXK-B	-
*	9.69	2e-10	30%	3Q41-A	-	*	0.34	0.04	7%	2LBP-A	-
*	7.42	2e-11	55%	5IPR-C=D	-	*	0.32	8e-05	10%	5FXI-D	-
*	5.92	3e-11	31%	3OLZ-B	-	*	0.28	1e-15	3%	5CMM-A	-
*	5.73	0.001	30%	3HUT-A	-	*	0.27	0.37	7%	4C9R-D	-
*	3.09	0.13	36%	1DP4-A	-	*	0.15	1e-09	2%	4TLM-A	-
*	2.76	2e-05	31%	4GPA-A	-	*	0.02	0.0002	1%	5KCA-A	-
*	2.71	0.009	31%	4N0Q-B	-	*	0.02	1e-09	0%	4PE5-C	-
*	2.66	0.02	27%	3IP6-A	-	*	0.01	6e-09	0%	5FXI-A	-

The other structures were discarded because the total score was not higher than 30% of the maximum or the templates did not provide any additional residues to the alignment. The secondary structure was predicted with PSI-Pred (46) using the PSSM from PSI-BLAST as input. The resulting secondary structure prediction assists YASARA during alignment corrections and loop modelling (41).

secondary structure prediction with PSI-Pred

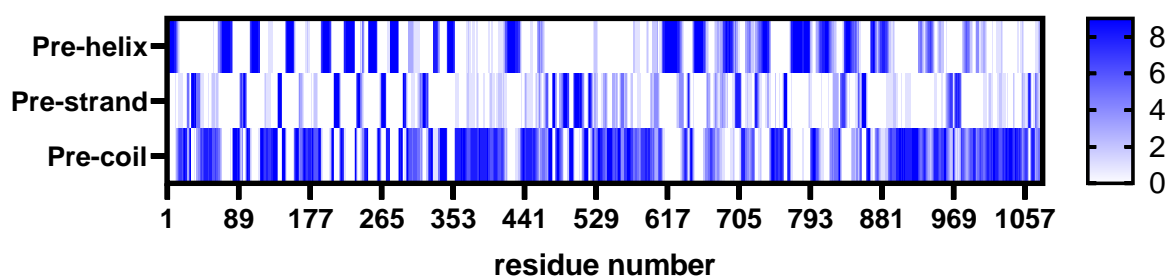


Figure 2.9 Heatmap of the secondary structure prediction of CaSR's sequence with PSI-Pred (47); the colour intensity reflects the estimated probability of each secondary structure element graded between 0 (minimum) and 9 (maximum)

Seq Str/1-1116
Target/1-1078
Temp. 1/1-532
Temp. 2/1-528
Temp. 3/1-573
Temp. 4/1-571
Temp. 5/1-568
Temp. 6/1-567
Temp. 7/1-567
Temp. 8/1-488
Temp. 9/1-484
Temp. 10/1-481
Temp. 11/1-485
Temp. 12/1-485
Temp. 13/1-322
Temp. 14/1-327
Temp. 15/1-324
Temp. 16/1-319
Temp. 17/1-331

Seq Str/1-1116
Target/1-1078
Temp. 1/1-532
Temp. 2/1-528
Temp. 3/1-573
Temp. 4/1-571
Temp. 5/1-568
Temp. 6/1-567
Temp. 7/1-567
Temp. 8/1-488
Temp. 9/1-484
Temp. 10/1-481
Temp. 11/1-485
Temp. 12/1-485
Temp. 13/1-322
Temp. 14/1-327
Temp. 15/1-324
Temp. 16/1-319
Temp. 17/1-331

Seq Str/1-1116
Target/1-1078
Temp. 1/1-532
Temp. 2/1-528
Temp. 3/1-573
Temp. 4/1-571
Temp. 5/1-568
Temp. 6/1-567
Temp. 7/1-567
Temp. 8/1-488
Temp. 9/1-484
Temp. 10/1-481
Temp. 11/1-485
Temp. 12/1-485
Temp. 13/1-322
Temp. 14/1-327
Temp. 15/1-324
Temp. 16/1-319
Temp. 17/1-331

Seq Str/1-1116
Target/1-1078
Temp. 1/1-532
Temp. 2/1-528
Temp. 3/1-573
Temp. 4/1-571
Temp. 5/1-568
Temp. 6/1-567
Temp. 7/1-567
Temp. 8/1-488
Temp. 9/1-484
Temp. 10/1-481
Temp. 11/1-485
Temp. 12/1-485
Temp. 13/1-322
Temp. 14/1-327
Temp. 15/1-324
Temp. 16/1-319
Temp. 17/1-331

Seq Str/1-1116
Target/1-1078
Temp. 1/1-532
Temp. 2/1-528
Temp. 3/1-573
Temp. 4/1-571
Temp. 5/1-568
Temp. 6/1-567
Temp. 7/1-567
Temp. 8/1-488
Temp. 9/1-484
Temp. 10/1-481
Temp. 11/1-485
Temp. 12/1-485
Temp. 13/1-322
Temp. 14/1-327
Temp. 15/1-324
Temp. 16/1-319
Temp. 17/1-331

Seq Str/1-1116
Target/1-1078
Temp. 1/1-532
Temp. 2/1-528
Temp. 3/1-573
Temp. 4/1-571
Temp. 5/1-568
Temp. 6/1-567
Temp. 7/1-567
Temp. 8/1-488
Temp. 9/1-484
Temp. 10/1-481
Temp. 11/1-485
Temp. 12/1-485
Temp. 13/1-322
Temp. 14/1-327
Temp. 15/1-324
Temp. 16/1-319
Temp. 17/1-331

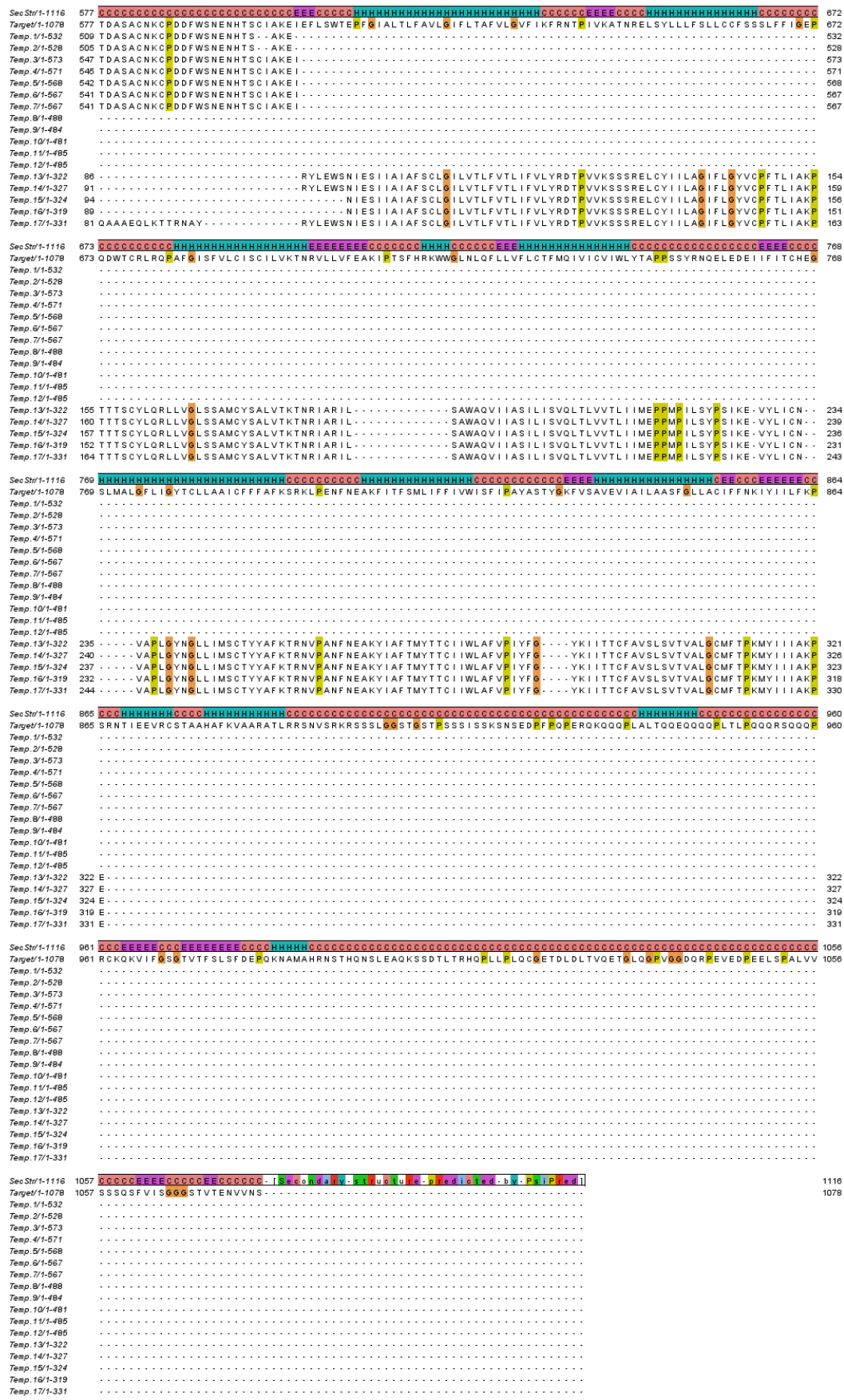


Figure 2.10 Sequence alignment between the 17 generated homology models and the primary sequence of the full-length CaSR; Sec. str.: secondary structure predicted with PSI-Pred (46)

The prediction in Figure 2.9 shows seven alpha-helices between residue 600 and 865 that correspond to the transmembrane region. It also shows that the extracellular domain contains numerous helices and beta-strands, consistent with the crystal structures (PDB: 5K5S; 5K5T). The intracellular domain is predicted to be mainly random coil. It would explain why the structure of this domain has not been determined yet.

The next step of the homology modelling process includes the creation of the target sequence profile. It is the result of a multiple sequence alignment that is based on 588 identified UniRef90 sequences. The results show that seventeen models are created out of four template structures with different alignments (Figure 2.10). The sequence alignment shows that the signal peptide was skipped including the first twenty amino acids. It also reveals that model one to twelve cover the target sequence up to position 600. It corresponds to the extracellular domain. Model thirteen to seventeen have an alignment covering part of the extracellular domain and mostly the transmembrane domain between residue 600 and 865. However, none of the templates shows a sequence alignment above residue 865. The consequence is that the intracellular domain is not present in any model. The number of aligned residues is listed in [Table 2.4](#).

Table 2.4 homology modelling results of 17 created models; seq. id.: sequence identity; seq. sim.: sequence similarity; Residues: the position of the modelled residues

model	oligomeric state	aligned residues	seq. id. [%]	seq. sim. [%]	loops added	Z-score	model ID	residues	comment
1	homodimer	532	99,60	99,80	6	-1,155	5K5S--01	20-602	satisfactory
2	homodimer	528	99,20	99,40	8	-1,237	5K5S--02	20-602	satisfactory
3	homodimer	573	99,70	99,80	6	-1,150	5K5T--01	21-603	satisfactory
4	homodimer	571	98,80	98,90	8	-1,154	5K5T--02	21-603	satisfactory
5	homodimer	568	98,80	98,90	7	-1,140	5K5T--03	21-603	satisfactory
6	homodimer	567	99,10	99,30	7	-1,149	5K5T--04	21-603	satisfactory
7	homodimer	567	98,90	99,50	7	-1,247	5K5T--05	21-603	satisfactory
8	homodimer	488	99,60	99,60	4	-0,675	5FBK--01	21-539	good
9	homodimer	484	99,20	99,20	6	-0,779	5FBK--02	21-539	good
10	homodimer	481	99,00	99,00	6	-0,722	5FBK--03	21-539	good
11	homodimer	485	98,40	98,60	6	-0,076	5FBK--04	21-539	good
12	homodimer	485	99,20	99,40	6	-0,849	5FBK--05	21-539	good
13	homodimer	322	27,00	48,10	10	-2,199	4OR2--01	347-865	poor
14	homodimer	327	26,30	47,70	10	-1,845	4OR2--02	347-865	satisfactory
15	homodimer	324	26,20	47,20	11	-1,844	4OR2--03	347-865	satisfactory
16	homodimer	319	27,00	47,60	11	-2,358	4OR2--04	347-865	poor
17	homodimer	331	27,80	49,80	12	-2,221	4OR2--05	347-865	poor

After the alignment, missing loops are added, and all models are created as homodimers. The side-chains are introduced, optimised and fine-tuned. An energy minimisation step followed including a combination of steepest descent and simulated annealing with restrained backbone atoms. The resulting half refined models were saved, and the quality of the model was determined by calculating the Z-scores (Equation 4). The Z-score is the weighted sum of dihedral, packing 1D, packing 3D and overall structure parameter and is calculated by applying the knowledge-based force field YASARA 2 (41).

$$\text{Z-score} = 0.145 \times \text{Dihedrals} + 0.390 \times \text{Packing1D} + 0.465 \times \text{Packing 3D} \quad \text{Equation 4}$$

The dihedrals variable describes the normality of all dihedral angles according to the force field. The variable packing1D expresses the normality of one-dimensional distance-dependent packing interactions of all molecules. Packing3D represents the normality of 3D direction-dependent packing interactions for common atom types found in proteins. All parameters depend on the YASARA 2 force field (24).

In the next step, simulated annealing minimisations are performed with complete unrestrained atoms to generate full refined models. The Z-scores are calculated again as before and compared to the half-refined models. The higher value of the overall Z-score corresponds to the better and selected model ([Table 2.4](#)). Terminal introduced residues are not considered for the Z-score calculation. Residues outside the maximum added termini length are skipped because ab initio predictions are not part of YASARA's homology modelling procedure. In the end, the quality Z-score of each model is determined per residue to allow a better interpretation of the results ([Figure 2.12](#)) (41). Residues with a Z-score below -2 have poor quality and values below -4 are considered bad.

The visualisation confirms that model one to twelve contains only the extracellular domain and model thirteen to seventeen, mainly the transmembrane domain ([Figure 2.12](#)). The evaluation of the models shows that the Z-score fluctuates between -2 and 2. The C-terminal area is modelled bad in model one and two and shows a Z-score below -4. Similar bad regions are in model thirteen around residue 488 and model fourteen around residue 468 and 758. The graphic representation

also reveals that the models thirteen to seventeen have an extracellular domain which is very different from the models one to twelve. These structures appear because YASARA did not identify and discard the T-lysozyme attached to the N-terminus of the transmembrane of the metabotropic glutamate receptor 1 (PDB: 4OR2). The conclusion is that the models thirteen to seventeen are not reliable up to residue 600. The orientation of the protomers is odd in model three to seven. The models show that the C-termini are not pointing in the same direction as expected. It means the quaternary structure is not correct of the homodimers.

In addition to the 17 models, YASARA tried to generate a hybrid model with a better Z-score. The creation process starts with an initial model that is considered as most suitable. Bad regions are replaced by equivalent residues from models with higher quality. Finally, 34 different hybrid models were created, and the best has an overall quality Z-score of -0,950 with 610 residues as a homodimer. Compared to the seventeen models, the hybrid model does not have the best Z-score ([Table 2.4](#)). But it has the best combination of Z-score and residue length and was selected by YASARA as the final model.

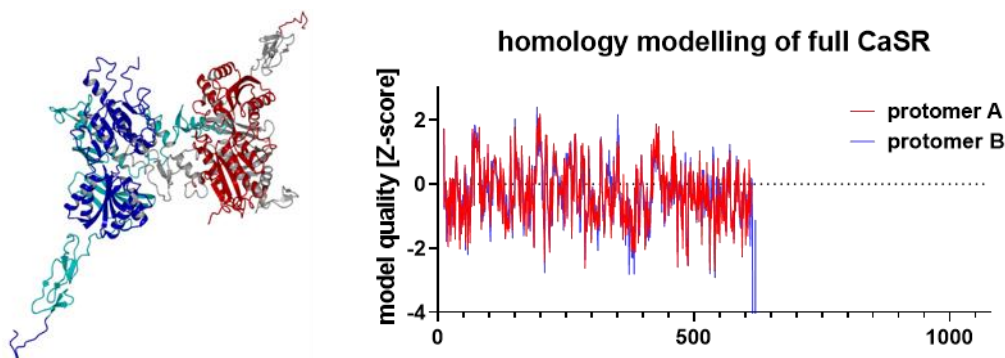


Figure 2.11 Final model of the homology modelling process based on the full-length CaSR sequence and created with YASARA Structure (41)

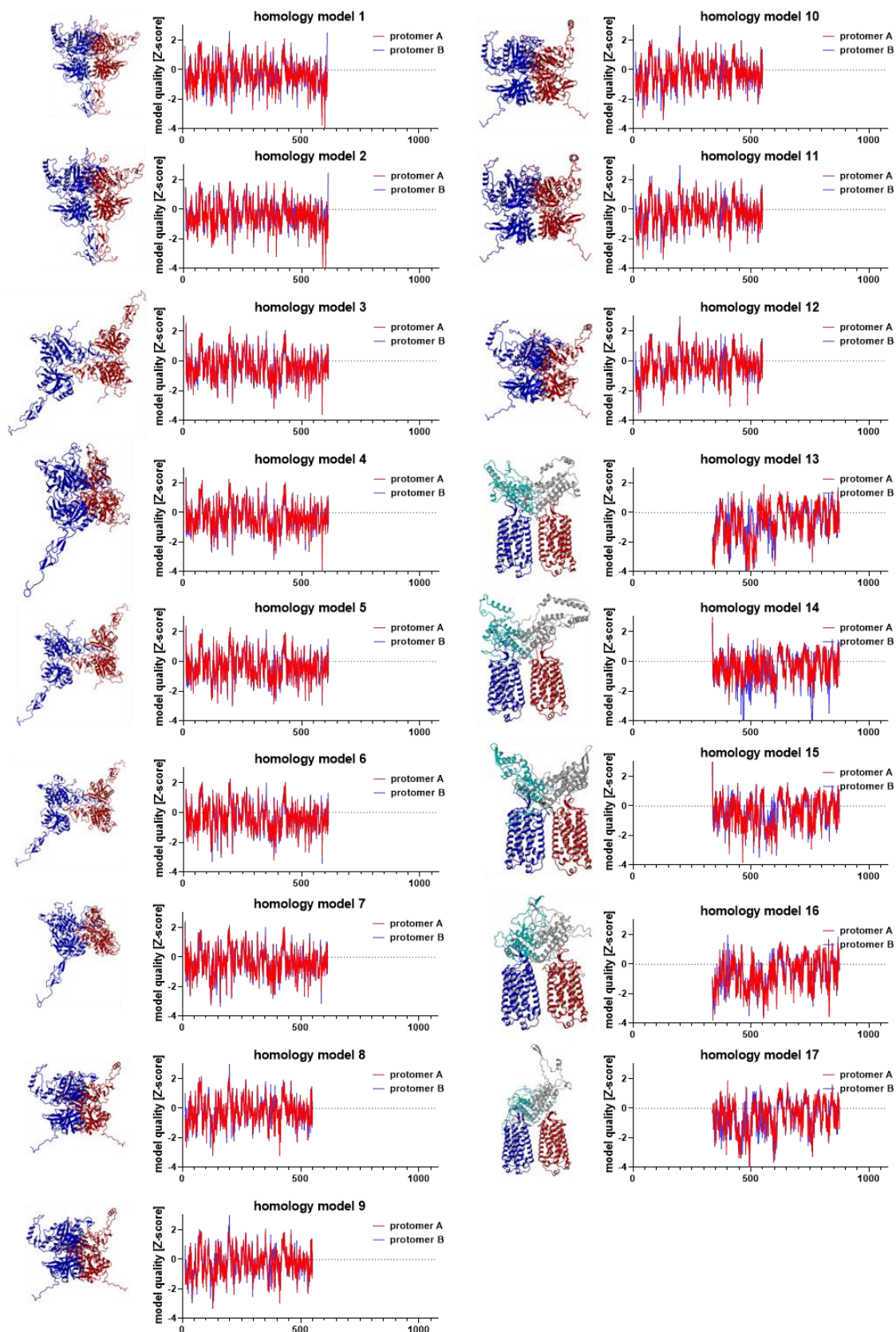


Figure 2.12: Intermediate homology models based on CaSR's primary sequence and generate with YASARA structure, results are shown in two columns, on the left is a graphic of the model and on the right the calculated model quality per residue (Z-score); Z-score < -2: poor; Zscore < -4: bad; blue: CaSR protomer A; red: CaSR protomer B, cyan and grey: incorrect homology model part based on the T-lysozyme

2.4 Homology modelling with YASARA for virtual screening

At this point, the previous automated structure prediction methods did not provide a reliable model of CaSR for virtual screening. The idea was to reconstruct the model of CaSR's transmembrane published by Leach et al. (26) but without the helix alignment gaps as mentioned in the supplemental material. After several setbacks, the modelling procedure for customised GPCRs was used according to Constanzi et al. (10). The described procedure focuses on homology modelling of class A GPCRs. Since bovine rhodopsin (48) (PDB: 1F88) was published as first GPCR in 2000, it was the only template to generate a homology model. But in the last two decades, several other GPCR crystal structures were solved and published. The workflow was adapted to create a homology model of CaSRs transmembrane.

The customised modelling of CaSR was initiated with the search for templates on the Protein Data Bank (43). In 2017 the search was performed with the keywords 'class C GPCR transmembrane' and gave the results listed in [Table 2.5](#).

Table 2.5 Class C GPCR templates of the transmembrane region; GRM1: metabotropic glutamate receptor 1; GRM5: metabotropic glutamate receptor 5; oligo.state: oligomerisation state; ref.: reference

GPCR	PDB ID	oligo. state	resolution	released	ref.
GRM1	4OR2	homodimer	2.80 Å	2014-03-19	(49)
GRM5	4O09	monomer	2.60 Å	2014-07-02	(50)
GRM5	5CGC	monomer	3.10 Å	2015-08-12	(51)
GRM5	5CGD	monomer	2.60 Å	2015-08-12	(51)

The crystal structures were downloaded in PDB format (page 122) and prepared to create a homology model of the seven-transmembrane domain containing only the alpha-helices. It is noteworthy that the template structures are all crystallised with negative allosteric modulators, and the probability is very high that they represent an inactive conformation. At the given time, no PDB entry could be found of a class C GPCR in the active form. Active class A GPCR structures are available but were not considered because of average sequence identity of 10.9% and an RMSD of more than 2.0 Angstrom (PDB: 3QAK, 3SN6, 3PQR) (52). The preparation of the templates involved the removal of non-protein entries, the fused T-lysozyme and the loops. The resulting templates contained all the residues of the seven-transmembrane alpha-helices. A structure-based sequence alignment

was downloaded from GPCRdb (11) between CaSR and the structures of the metabotropic glutamate receptor subtype 1 and 5. The final alignment is depicted in Figure 2.13, and it was used for the homology modelling experiment (page 114).

```

CaSR_7TM      611 PFGIALTLFAVLGIFLTAFLVGVFIKFNRELSYLLLFSLCCFSSSLFFIDWTCRLRQPAFGIS 687
T001_4O09A   578 PAPIAAVVFACLGLLATLFTVTVFIIYSRELCYIILAGICLGYLCTFCLIQIYCYLQRIGIGLS 654
T002_4OR2A   591 IESIIAIAFSCGLILVTLFVTLIFVLYSRELCYIILAGIFLGYVCPFTLITTS CYLQRLLVGLS 687
T003_4OR2B   591 IESIIAIAFSCGLILVTLFVTLIFVLYSRELCYIILAGIFLGYVCPFTLITTS CYLQRLLVGLS 687
T004_5CGCA   578 PAPIAAVVFACLGLLATLFTVTVFIIYSRELCYIILAGICLGYLCTFCLIQIYCYLQRIGIGLS 654
T005_5CGDA   578 PAPIAAVVFACLGLLATLFTVTVFIIYSRELCYIILAGICLGYLCTFCLIQIYCYLQRIGIGLS 654

CaSR_7TM      688 FVLCISCVLTKNRLVLLVFGLNLFQFLVFLCTFMQIVICVIWLYTALMALGFLIGYTCLLAAIC 787
T001_4O09A   655 PAMSYSALVTKTYRAARILSACAQLVIAFILICIQLGIIIVALFIMENLGVVAPLGYNGLLILAC 754
T002_4OR2A   668 SAMCYSALVTKTNRIARILSAWAQVVIASILISVQLTLVVTLIIMENLGVVAPLGYNGLLIMSC 787
T003_4OR2B   668 SAMCYSALVTKTNRIARILSAWAQVVIASILISVQLTLVVTLIIMENLGVVAPLGYNGLLIMSC 787
T004_5CGCA   655 PAMSYSALVTKTYRAARILSACAQLVIAFILICIQLGIIIVALFIMENLGVVAPLGYNGLLILAC 754
T005_5CGDA   655 PAMSYSALVTKTYRAARILSACAQLVIAFILICIQLGIIIVALFIMENLGVVAPLGYNGLLILAC 754

CaSR_7TM      788 FFFAFKSEAKFITFSMLIFFVWISFIPAYASFVSAVEVIAILAASFGLLACIFFNKIYIILF 862
T001_4O09A   755 TFYAFKTEAKYIAFTMYTTCIIWLA芙蓉PIYFGYKIIITMCFSVLSLTSATVALGCMFVPKVYIILA 827
T002_4OR2A   768 TYYAFKTEAKYIAFTMYTTCIIWLA芙蓉PIYFGYKIIITTCFVSLSVTVALGCMFVTPKMYIIIA 840
T003_4OR2B   768 TYYAFKTEAKYIAFTMYTTCIIWLA芙蓉PIYFGYKIIITTCFVSLSVTVALGCMFVTPKMYIIIA 840
T004_5CGCA   755 TFYAFKTEAKYIAFTMYTTCIIWLA芙蓉PIYFGYKIIITMCFSVLSLTSATVALGCMFVPKVYIILA 827
T005_5CGDA   755 TFYAFKTEAKYIAFTMYTTCIIWLA芙蓉PIYFGYKIIITMCFSVLSLTSATVALGCMFVPKVYIILA 827

```

Figure 2.13 Sequence alignment between the seven individual helices of CaSR and the templates 4O09-A, 4OR2-A, 4OR2-B, 5CGC-A and 5CGD-A (created with Jalview 2.11.1.3) (40)

Homology modelling was performed with YASARA structure (version 17.1.17.L.64). The modelling speed was set to 'slow' to create a model with the highest quality score. The (PSI)-BLAST and E-value parameter were redundant because the sequence alignment was provided manually. The homodimer (PDB: 4OR2) was split into two molecules giving five heptameric template structures. Five alignment variations were created per sequence. The termini extension was omitted, and the sampling of the loops was not required to create the seven alpha-helical model.

YASARA automatically predicted the secondary structure with PSI-PRED (46) based on the target sequence (Figure 2.14, first line). The prediction deviates slightly from the pure alpha-helical template structures. The modelling program characterised the alignment to be ambiguous and created multiple alignment variations, as depicted in Figure 2.14. The alignment shows that YASARA introduced unnecessary gaps in the sequences of the created models. The alignment was unchanged for template 4OR2A variant 2, 4OR2B variant 2 and 5CGCA variant 3. During the homology modelling procedure, a half-refined and full-refined model was created for each alignment.

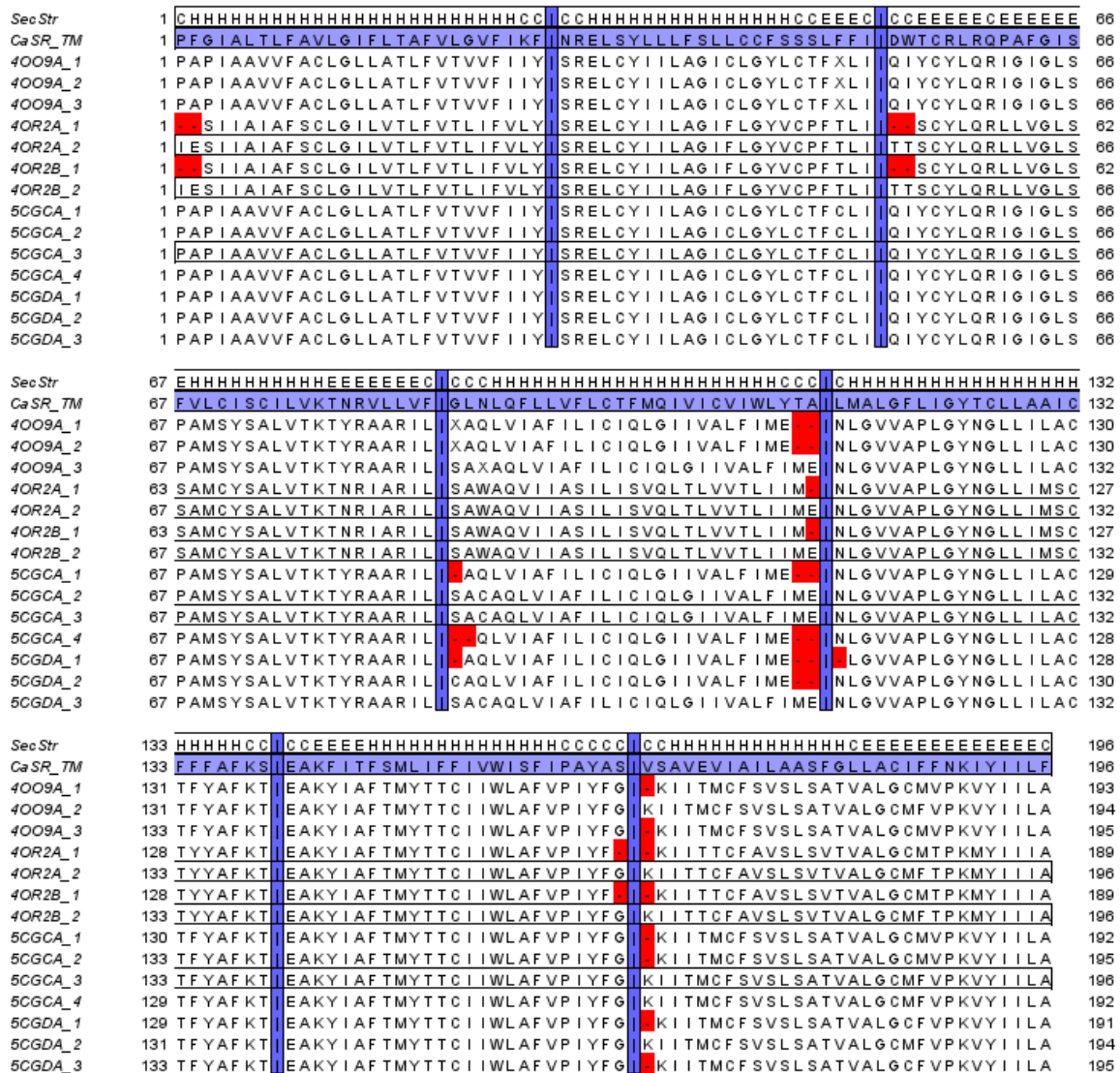


Figure 2.14 Sequence alignment of the resulting homology models with the sequence of CaSR's seven transmembrane alpha-helices; sec. str.: secondary structure prediction with PSI-PRED (46); CaSR_TM: target sequence of the calcium-sensing receptor seven alpha-helical structures of the transmembrane; blue underlined vertical lines (|): delimiters between the seven helical sequences; red: incorrectly introduced gaps

Table 2.6 homology modelling results of 14 created models; seq. id.: sequence identity; seq. sim.: sequence similarity based on the BLOSUM62 score (Figure 4.3); residues: the position of the modelled residues

model	oligomeric state	aligned residues	seq. id. [%]	seq. sim. [%]	loops added	Z-score	model ID	residues	comment
1	heteroheptamer	187	29.9	51.3	3	0.142	4009A_1	1-196	optimal
2	heteroheptamer	188	30.9	52.1	1	0.607	4009A_1	1-196	optimal
3	heteroheptamer	189	30.7	50.8	2	0.448	4009A_1	1-196	optimal
4	heteroheptamer	183	28.4	52.5	6	0.474	4OR2A_1	3-196	optimal
5	heteroheptamer	190	28.9	51.6	0	0.583	4OR2A_2	1-196	optimal
6	heteroheptamer	183	28.4	52.5	6	0.637	4OR2B_1	3-196	optimal
7	heteroheptamer	190	28.9	51.6	0	0.637	4OR2B_2	1-196	optimal
8	heteroheptamer	186	30.1	51.6	4	0.637	5CGCA_1	1-196	optimal
9	heteroheptamer	189	30.7	50.8	2	0.372	5CGCA_2	1-196	optimal

10	heteroheptamer	190	31.6	51.6	0	0.893	5CGCA_3	1-196	optimal
11	heteroheptamer	186	31.2	52.7	2	0.694	5CGCA_4	1-196	optimal
12	heteroheptamer	185	30.8	52.4	5	0.599	5CGDA_1	1-196	optimal
13	heteroheptamer	188	30.9	52.1	1	1.018	5CGDA_2	1-196	optimal
14	heteroheptamer	189	31.2	51.3	2	0.543	5CGDA_3	1-196	optimal

The fourteen models' results reveal a quite low sequence identity of 28.4 to 31.6% between the helices of CaSR and the template structures (Table 2.6). But the sequence similarities are acceptable in the range of 50.8 to 52.7%. The homology model's quality also depends on the sequence identity and the sequence length between the template and the target structure. The sequence identity between CaSR and metabotropic glutamate receptor 1 or 5 is close to the twilight zone (Figure 2.15), making it difficult to create a reliable homology model.

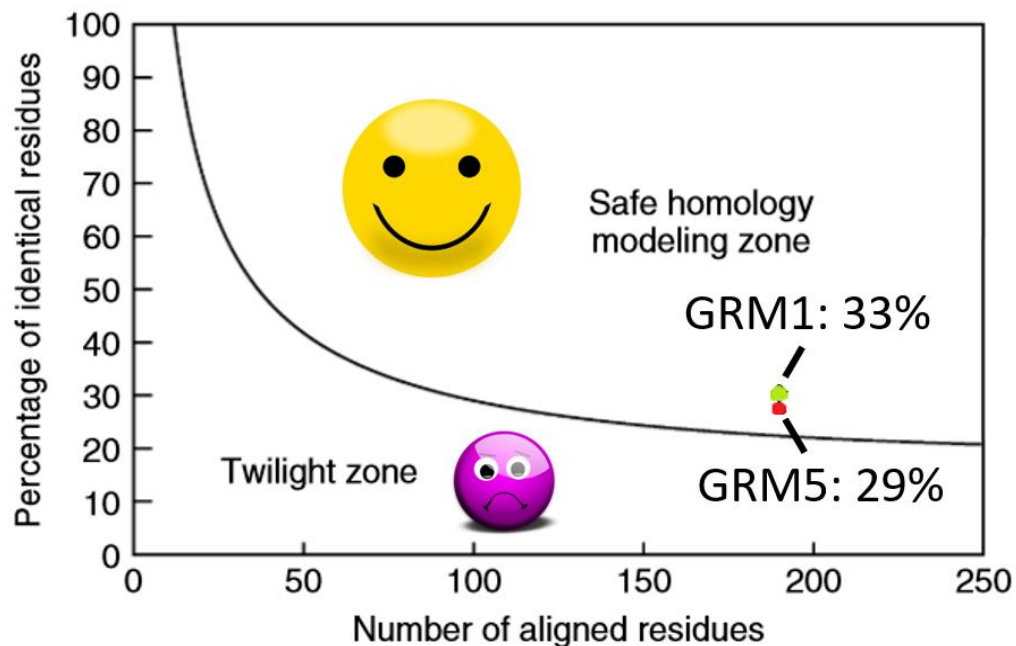


Figure 2.15 Relationship between sequence identity and sequence length to create a reliable homology model (modified image from (53)); dots: sequence identity between metabotropic glutamate receptor 1 (GRM1) or metabotropic glutamate receptor 5 (GRM5) to CaSR

In the end, YASARA created a hybrid model out of the fourteen with a quality Z-score of 1.068. It was rejected because of the mismatching alignment in most of the cases.

Model 5, 7 and 10 listed in Table 2.6 were used as templates in a subsequent homology modelling run to create a hybrid model. The setting was the same as before but without the parameter for multiple alignment variants. The new run did

not generate a hybrid model but noted that it was impossible to gain a model with further improvements. Looking at the quality Z-score results hardly reveals any difference between the models from the first run (Figure 2.16). Finally, model 7 was selected based on metabotropic glutamate receptor 1 because it had the most improvement in the second run to a Z-score of 0.902.

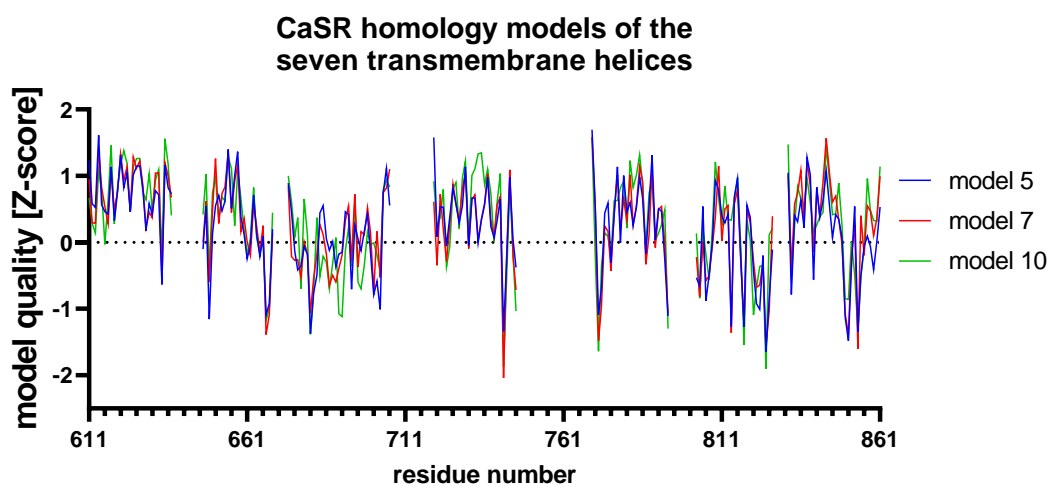


Figure 2.16 estimated model quality per residue of the homology models created for the CaSR transmembrane helices

The next step in the homology modelling workflow includes loop modelling. The loops were modelled one after the other onto the seven alpha-helices. The N-terminal extracellular residues 601-611 were created based on template 4OR2-A (metabotropic glutamate receptor 1). Although it is linked to a T-lysozyme, this template was selected, and the impact is unknown on the native structure. It became the best option because the N-terminal residues in the GRM5 templates form a less probable alpha-helix. The intracellular loops 1 and 2 and the extracellular loops 1, 2 and 3 are aligned with the template 4OR2-A. Only the intracellular loop 3 was based on the template structure 4OO9. The missing loops were created with the 'BuildLoop' comment in YASARA (24). This command scans a set of non-redundant PDB entries with a minimum sequence identity of 90%. Top scoring sequences are selected based on overlapping anchor atoms with the lowest RMSD. The loops are covalently integrated at the N-terminal or C-terminal end or fused with the protein's backbone (54). The alignment for the loops is shown in Figure 2.17.

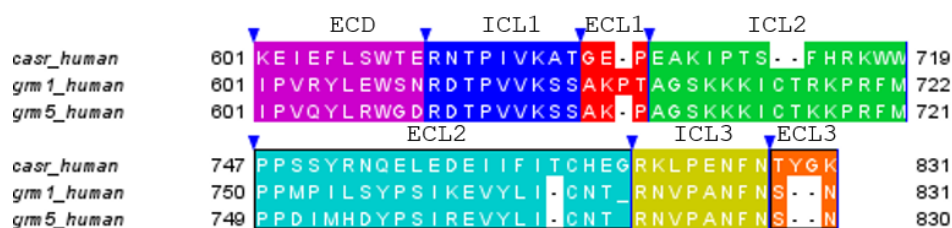


Figure 2.17 sequence alignment between the loops of CaSR, metabotropic glutamate receptor 1 (GRM1) and 5 (GRM5); ECD: extracellular domain residues; ICL: intracellular loop; ECL: extracellular loop

After the loop modelling, a conserved disulfide bridge was placed between residue Cys677 from helix III and Cys765 from extracellular loop 2. The model was energy minimised using steepest decent minimisation with the NOVA force field in vacuo (proteiN Optimisation in VAcuo) (55). This force field uses a force cutoff at 10.5 Angstrom, and it omits solvent molecules. Other force fields overestimate the resulting electrostatic interaction like lysine and glutamine without solvent and try to compensate this circumstance by setting the net charge to 0. In the NOVA force field, the net charge was optimal at 0.205 after optimisation (24).

After the minimisation step, the residues F668, R680, L776, F821 and V838 around the binding pocket were finetuned, according to Leach et al. (26). But the pocket volume had not the right size to accommodate the negative allosteric modulator NPS-2143. This issue was faced by sampling the conformation of the extracellular loop 2 twelve times, and a structure was selected that had an open binding pocket (Figure 2.18). Docking experiments were executed to recreate the published CaSR – NPS-2143 complex (26).

Docking is performed with a customised version of AutoDock VINA (23,24). This program makes it possible to dock with flexible parts of the protein and uses receptor ensembles to explore the receptor conformational space with different networks of sidechain rotamers. The rotameric states are explored with the SCWALL method (Side-Chain conformations With All available methods) (41). It initiates the optimisation process with the SCWRL method (Side-Chains optimisation With a Rotamer Library). The starting conformations are created with a simple repulsion energy function that optimises Van der Waals forces like hydrophobic core residues. Electrostatic surface interactions are improved with the steepest descent minimisation using the YASARA2 force field. YASARA requires at least a receptor target, one ligand and a simulation box surrounding residues of

interest to predict the desired receptor-ligand interaction. The docking poses are clustered into different groups if the RMSD is bigger than the default cutoff of 5 Angstrom. The interactions get scored according to the binding energy in kcal/mol; however, AutoDock VINA may create identical poses. The interaction can be judged by the ligand efficiency parameter. It represents the binding energy per heavy atom, and if the value is ≥ 0.35 , the ligand is considered a good binder (24). The results also contain the dissociation constant in pM and list the contacting receptor residues. The docking experiments were performed with a modified version of macros provided by YASARA (24).

At first, the docking experiments showed that the negative allosteric modulator was predicted to interact only with the receptor surface. Also, the seven helices without the loops showed the same result. After the loop was sampled for an open conformation, the ligand got into the binding pocket but did not interact with the predicted residues at the bottom of it. Different settings were explored to reconstruct the published model involving different combinations of energy minimisations and the docking cell's adaption. However, none of the created receptor ensembles showed an interaction as expected.

The binding pocket was inspected again, and the residues compared to the initial template of the metabotropic glutamate receptor 1 (PDB: 4OR2-B). The sequence alignment showed that only two out of seventeen residues are identical in the binding pocket. But the group of Leach et al. compared the binding pockets between the different class C GPCRs and concluded that they are in a similar area within the seven transmembranes (26). Based on this information, the strategy changed to reconstruct the CaSR – NPS2143 receptor complex. The published complex of GMR1 (PDB: 4OR2) was used as a template for structural alignment. In the PDB 4OR2, the metabotropic glutamate receptor 1 is in complex with the negative allosteric modulator FITM (4-fluoro-N-methyl-N-{4-[6-(propan-2-ylamino)pyrimidin-4-yl]-1,3-thiazol-2-yl}benzamide). The alignment method MUSTANG (56) was used for the structural aligned of the CaSR model to the template complex. It is one of YASARAs alignment methods, and it tries to superimpose as many equivalent atoms as possible at a minimum RMSD. The structure of NPS-2143 was reconstructed in YASARA based on the SMILE (page 123) obtained from PubChem (PubChem CID: 6918446) (57). NPS-2143 was

structurally aligned to FITM of the GRM1 complex. The template complex was removed in the next step and resulted in an aligned complex between CaSR and NPS-2143. The backbone atoms of CaSR were fixed, and the remaining structure was energetically minimized. The resulting receptor complex of CaSR and NPS-2143 shows that the allosteric modulator reaches into the middle of the transmembrane and is not stuck on top of it (Figure 2.18-A). The interaction profile of NPS-2143 is in accordance with Leach et al. (26). It depicts the allosteric modulator's interaction to residue F668, F684, F688, F821 and V833 (Figure 2.18-A) which are important for the binding affinity. It also reveals that the residues F612, E670, P672, R680, L682, L776, F832, S834, V838 and A844 are also relevant for the binding affinity but do not interact with the ligand directly.

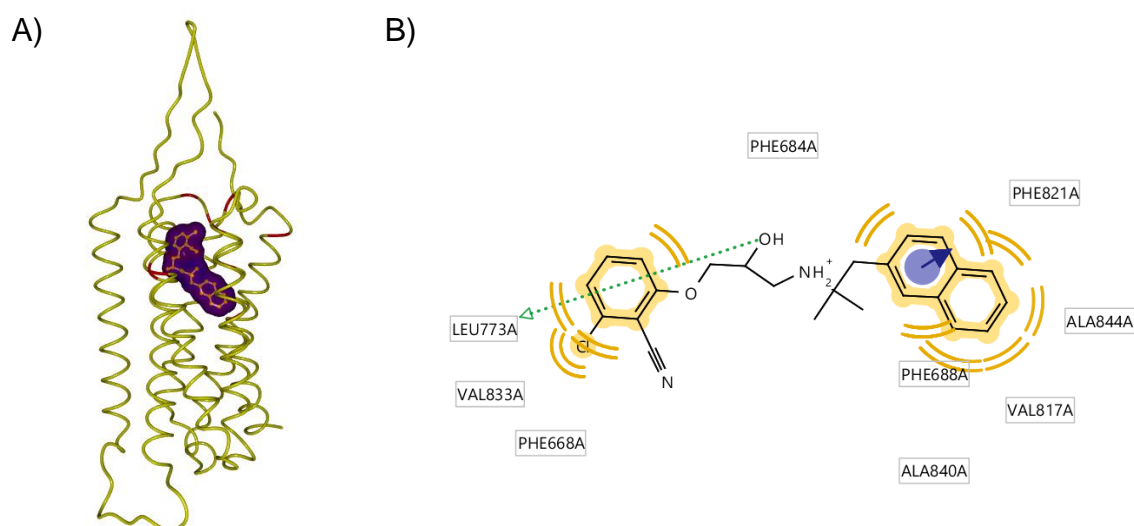


Figure 2.18 CaSR - NPS-2143 receptor complex with an open conformation for extracellular loop 2, A) NPS-2143 transmembrane complex after structural alignment of CaSR with GRM1, purple: binding pocket; B) interaction profile of NPS-2143 with CaSR, yellow: hydrophobic interactions, green: hydrogen donor interaction; blue: aromatic $\pi - \pi$ interaction

The quality of the final model was evaluated again by calculating the estimated quality Z-score per residue. The comparison to the initial model 7 shows that the loops between the alpha-helices have a poor quality (Z-score < -2), but it should not influence the docking experiments (Figure 2.19-A). Another quality check includes a 2D plot of the backbone angles phi and psi. The torsion angle phi is between the C-C α atoms and psi is between the backbone's C α -N atoms. The resulting Ramachandran plot is characterized by areas that correspond to secondary structure elements. In the region around -150° to -45° phi and 90° to 180° psi, it is most likely that the residue is to part of a beta-sheet. Alpha-helices

tend to appear in the area of -110° to -30° phi and -60° to 30° psi. A smaller area around 50° to 70° phi and 20° to 50° psi is most likely occupied by residues of a left-handed helix. The depiction of the phi and psi angles is an easy and quick way to evaluate a model's quality. The results of for Ramachandran plot for the homology model shows that 89,2% of the residues have a psi and phi angle in the allowed areas (Figure 2.19-B; red), 10% are in the allowed regions (Figure 2.19-B, yellow) and two residues are in the disallowed regions. Residue R638 is part of the intracellular loop 1 and residue L770 is the anchor residue for extracellular loop 2. Both residues are not part of the binding pocket and cannot interfere with docked ligands. Energy minimization of these and adjacent residues moved the angles even further into the disallowed regions.

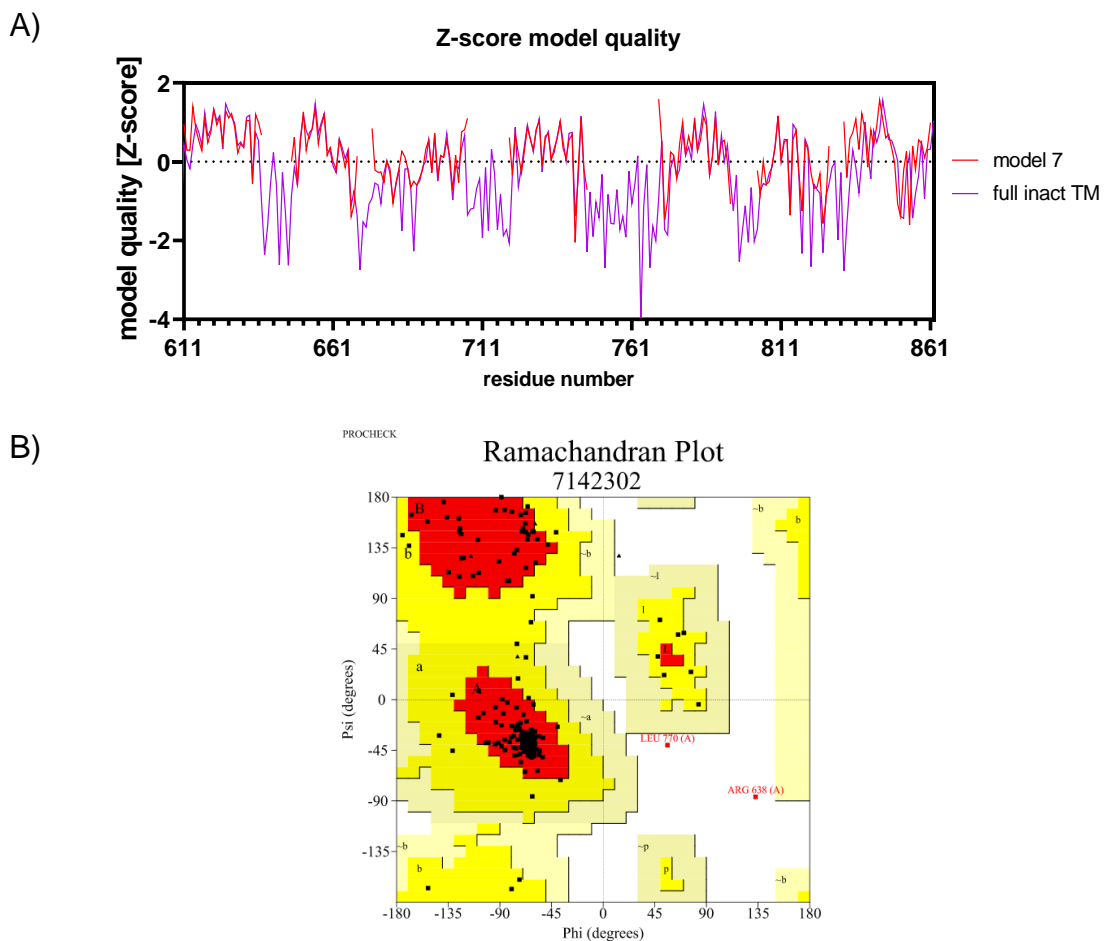


Figure 2.19 Quality check of CaSR transmembrane homology model; A) Z-score quality check per residue; B) Ramachandran Plot with PROCHECK (58)

The documentation states, if the homology model was created with YASARA, further refinement steps are not required because it is most likely that the quality of the model gets worse. The refinement macro of YASARA was applied to the

model to determine if any part of the model could be improved (24). The procedure involves a 500 ps simulation of the model using the knowledge-based YASARA 2 force field at 298 K, pH = 7.4 and a density of 0.997. Twenty snapshots were saved, and several parameters were determined to judge the quality of the models. The model of snapshot 18 returned the lowest energy and the highest quality. But the evaluation of the models per residue shows that the quality hardly changes compared to the starting structure (Figure 2.20). As a result, the refined models were rejected.

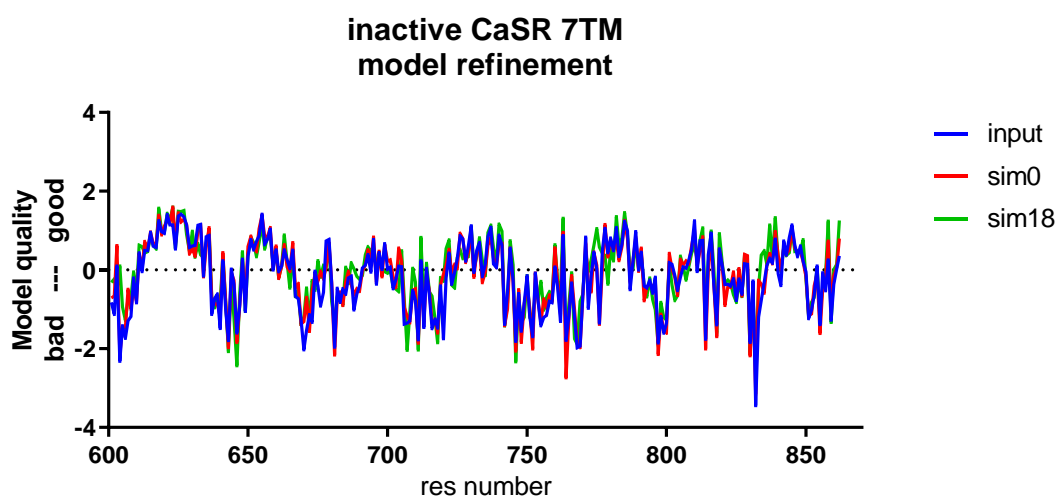


Figure 2.20 model refinement of CaSR's homology model in the inactive state with YASARA's refinement macro (24)

The created model represents the inactive form of CaSR's transmembrane because it is based on template PDB: 4OR2. Before the homology model could be used for virtual screening, it was validated by docking experiments with 363 known negative allosteric modulators (NAM) (page 124). The validation set was extracted from the ChEMBL database (59). The ligands were docked eight times with a minimum RMSD of 5 Angstrom against the rigid receptor with AutoDock VINA. The analysis was performed by evaluating the linear regression with the Pearson correlation r (60,61) (Equation 5).

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{[\sum_{i=1}^n (x_i - \bar{x})^2][\sum_{i=1}^n (y_i - \bar{y})^2]}}$$

Equation 5

The x and y values represent the predicted and experimental binding affinity entries. The Pearson correlation can have values between -1 and 1, corresponding to absolute anticorrelation or correlation between the variables, respectively. If the results are close to zero, it means that the two variables have no linear relationship at all.

The docking experiment's first result looked very promising with a Pearson correlation of 0,92 for predicting the experimental value within a 95% confidence interval. Unfortunately, this result could not be repeated. Instead, multiple reruns confirmed that the prediction has a Pearson correlation of 0.16. It means the predictions based on this model have no correlation with the true values and are random. The conclusion is that the model cannot be used for virtual screening because the scoring function and the predicted dissociation constant, are not reliable.

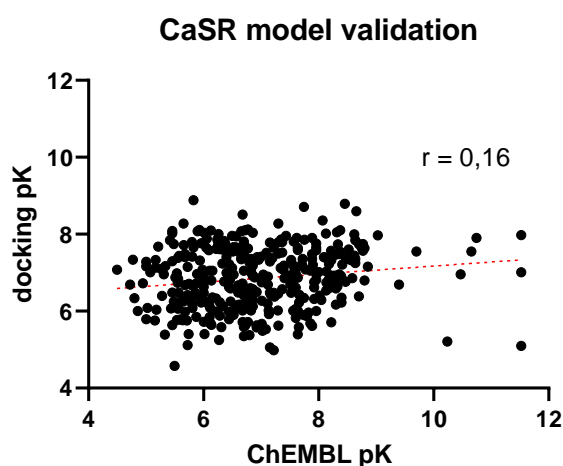


Figure 2.21 CaSR model validation with 363 negative allosteric modulators. Pearson correlation coeff.: 0.16;

The interaction pattern shows that all NAMs interact with at least 2 of the 12 binding-affinity relevant residues (668, 684, 688, 767, 772, 818, 821, 825, 833, 837, 840 and 841). But only 46,8% (170/363) NAMs reach the bottom of the binding pocket. One option to increase the correlation between the predictions and the true values was to create multiple receptor conformations with molecular dynamics simulations (MD). Three 5 nanosecond MD simulations were executed with a cell membrane and run with the YASARA macro (www.yasara.org/md_runmembrane.mcr) (24) using different numbers of random seed. The macro

automatically scans for exposed transmembrane helices. It identifies helices longer than 16 amino acids including more than seven hydrophobic ones with at least three having an accessible surface area > 30% relative to the maximum. The Y-axis of the helices is determined by summing up the least-square lines through the C α atoms, and the membrane is oriented perpendicular to it. The membrane was placed to cover the largest number of exposed hydrophobic residues and has a width of 28 Angstrom. The membrane composition is 100% phosphatidylethanolamine, and it was enclosed by a simulation cell of size 75.11 x 109.12 x 75.11 Angstrom. The next step, the protein gets scaled to 0.9 along the XZ-axes and lipid were deleted that bumped into the protein at a distance closer than 0.75 Angstrom. The scaling was slowly reversed with a short simulation at 298 K in vacuo. The fixed protein was scaled every 200 femtoseconds by 1.02 while the membrane could move with restrains to an ideal geometry. It was achieved by pulling lipid residues into the membrane with atoms beyond 21.5 Angstrom from the membrane centre and pushing phosphorus atoms away from the membrane centre if they are closer than 14 Angstrom. The AMBER14 force field was used with Lipid14/GAFF/AM1BCC parameters for non-standard residues (62–65). After the protein reached its original size again, the pK a s were predicted of the protein side chains (66). The protonation states get assigned for a pH of 7.4. The simulation cell is filled with TIP3P water molecules, 0.9% NaCl and counter ions during the main simulation (67). Long-range interactions are calculated with the Particle mesh Ewald (PME) method with an 8 Angstrom force cutoff for non-bonded forces (68). The time steps were set to 4 femtoseconds, and the simulations were performed with constrained hydrogens at constant pressure and temperature (NPT ensemble) according to the setting of (69). The membrane was restrained during the first 250 picoseconds in the pressure equilibration phase before the main simulation run until 5 nanoseconds.

The results show that the system is equilibrated after the first 100 picoseconds (Figure 2.22) according to the RMSD of all the protein atoms (Equation 3).

Molecular dynamics simulations with CaSR's TM in the inactive state

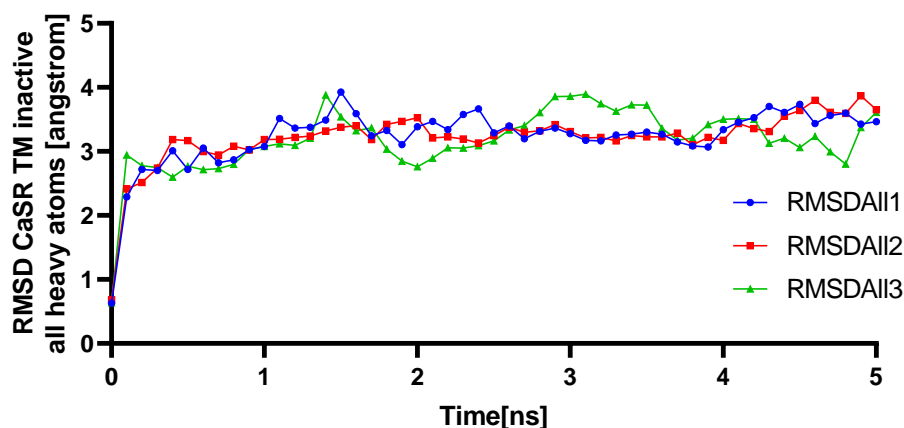
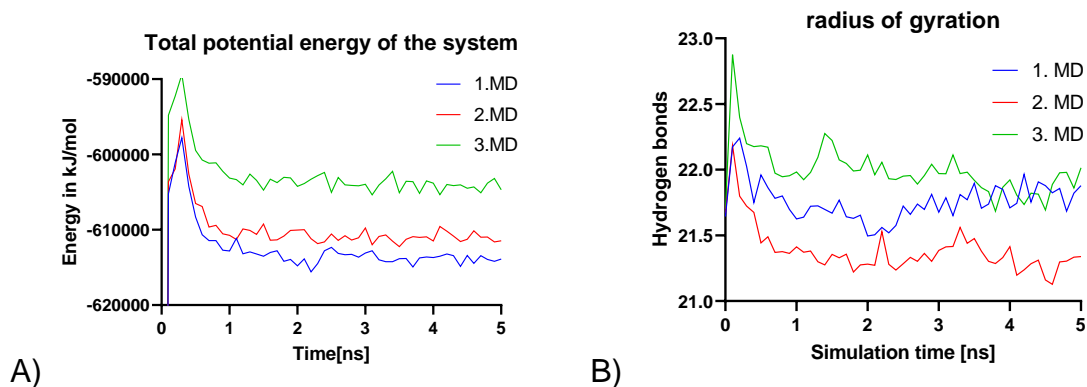


Figure 2.22 RMSD of three 5 ns long molecular dynamics simulations with the transmembrane model

The simulations analysis indicates that the whole system reaches the equilibrium of potential energy after 1.0 ns (Figure 2.23-A). The radius of gyration shows that the average distance of the model atoms does not change significantly (Figure 2.23-B). This result reflects the compactness of the seven-transmembrane domain, and bigger fluctuations are not expected (Equation 6).

$$R_{gyr,Mass} = \sqrt{\frac{\sum_{i=1}^N Mass_i (\vec{R}_i - \vec{C})^2}{\sum_{i=1}^N Mass_i}} \quad \text{Equation 6}$$

The radius of gyration (R_{gyr}) gives the average distance of the atom positions (R_i) relative to the centre of mass (C) covering N atoms (24). The analysis also shows no dramatic changes in the number of hydrogens within the model and the solvent (Figure 2.23-C&D).



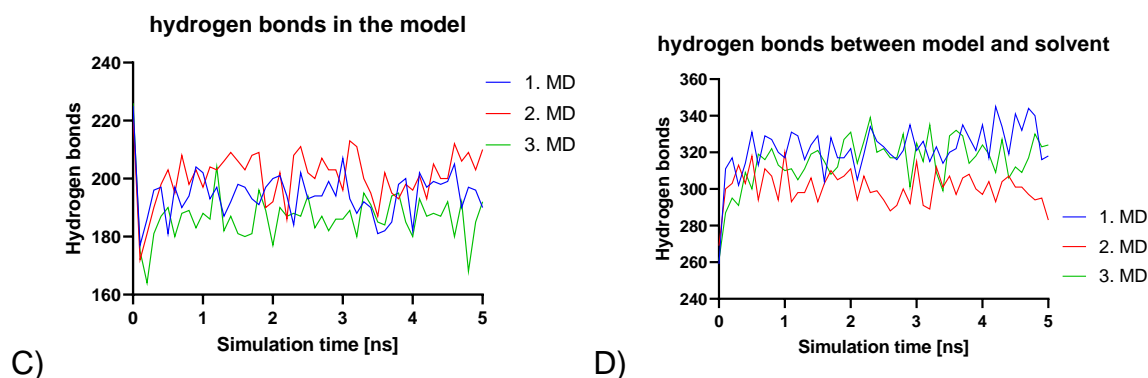


Figure 2.23 Results of three 5ns molecular dynamics simulations; A) total potential energy of the system as a function of simulation time; B) radius of gyration representing the fluctuations of all model atoms around the centre of mass; C) number of hydrogen bonds within the model

The three molecular dynamics simulation results were utilized to select several snapshots of the model at local minima and maxima of RMSD (Figure 2.22). These generated conformers were used for prospective validation studies performing docking experiments with the same set of 363 NAMs, and the results are presented in Figure 2.24.

The docking experiments' outcome shows that the predicted binding affinities by AutoDock VINA hardly correlate with the experimental data. The initial model has a Pearson correlation of 0.16, and the predictions do not follow a linear correlation but are randomly distributed around it. The conformers perform better, but the best model (Figure 2.24; 3.MD, snapshot 1.4 ns) correlates only $r = 0.33$ between the predictions and the measured results.

In the end, all the generated models of the extracellular calcium-sensing receptor could not be used for reliable virtual screening of DrugBank and the discovery of a new potential drug candidate.

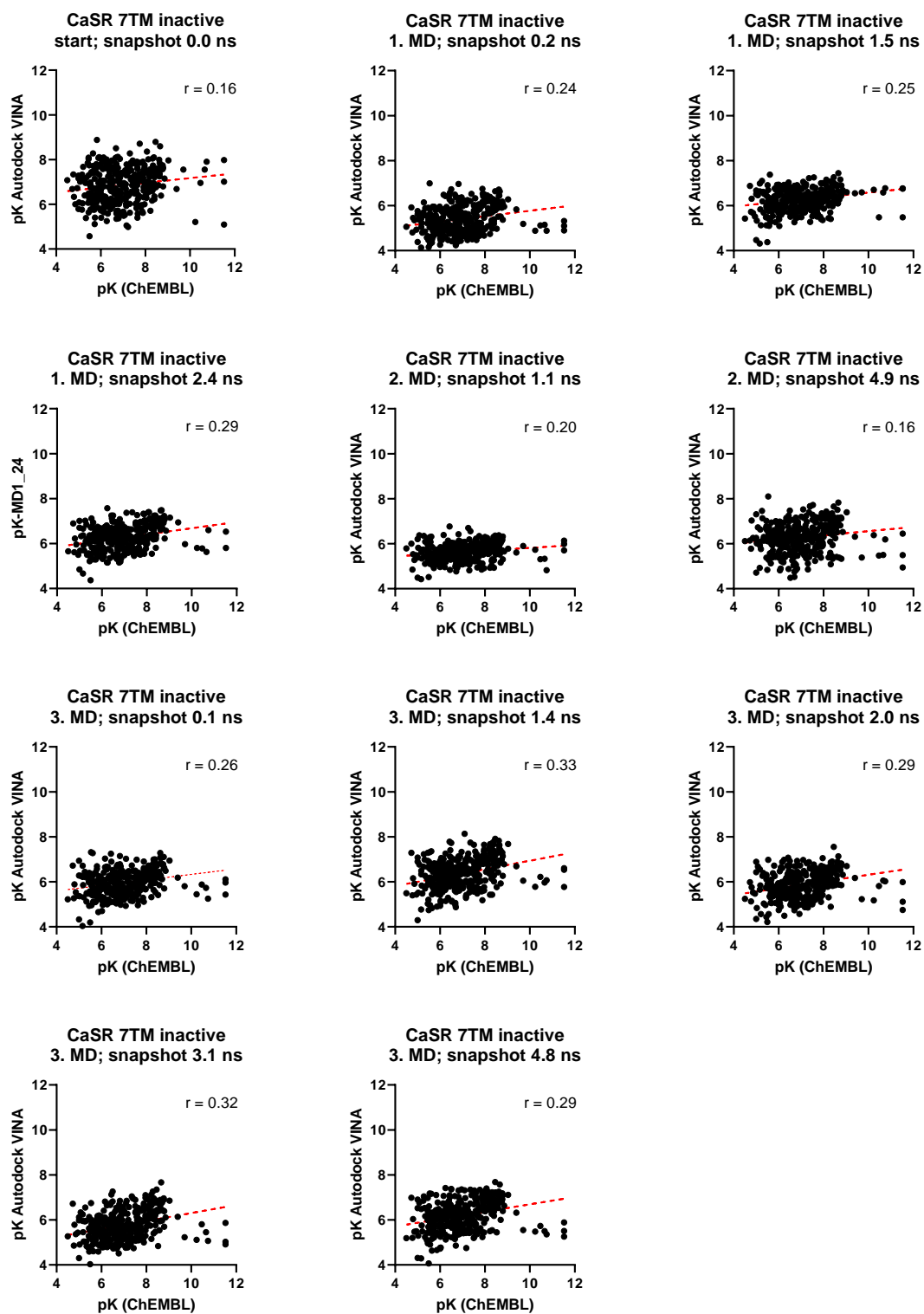


Figure 2.24 Prospective study with eleven conformers of CaSR's transmembrane homology model by applying docking experiments with the validation set of 363 negative allosteric modulators; r : Pearson correlation (Equation 5)

Discussion:

The discovery of a novel allosteric modulator for the extracellular calcium-sensing receptor (CaSR) involved the virtual screening of the compound library DrugBank. Molecular docking with AutoDock VINA was the method of choice. It requires a model of CaSR and studies of Leach et al. (26) pointed out that the binding pocket for allosteric modulators of CaSR is in the upper part of the transmembrane domain. Until now, there is no experimentally determined structure available of CaSR's transmembrane. Structure prediction methods like I-TASSER and GPCR-I-TASSER were tested to create a model of CaSR but failed. Homology modelling with YASARA was performed based on the FASTA sequence of CaSR. It did not generate a useful model for virtual screening.

The big challenge for automated methods is the low sequence identity of 30% between CaSR and other class C GPCR members. The PDB structures of the glutamate receptor 1 and 5 require manual preparation for homology modelling. The adequate multisequence alignment was downloaded from GPCRdb. During the modelling process, many parameters were tested to gain a model with the highest quality. The alpha-helices were modelled first showing optimal modelling results. The model's structural alignment with the template shows that only 2 out of 17 residues are conserved in the binding pocket. The correct loop conformations' modelling was quite difficult because the templates are fused with a T-lysozyme either at the intracellular or extracellular side. The impact on the loop conformations is unknown and could represent artefacts. Finally, the best model of CaSR was validated with 363 known negative allosteric modulators from ChEMBL by molecular docking experiments. Unfortunately, the results showed nearly no correlation ($r = 0.16$) between the predicted and experimental binding affinities. Besides, refinement methods could hardly improve the quality of the model, and the best one showed a Pearson correlation of 0.33. Initial screening runs of small test sets indicated that about 2200 compounds could be docked within one day. The discovery of potential hits would not be feasible by screening large compound libraries of more than 50000 compounds. At this point, the virtual screening approach was rejected, and a different method was required.

Chapter II: Discovery of novel allosteric modulator for the extracellular calcium-sensing receptor by in-silico profiling with SAFAN-ISPSM

2.5 Rationale:

The aim of this project is the discovery of novel allosteric modulators for the human extracellular calcium-sensing receptor CaSR. The hypothesis is that such modulators could become potential agents for treating CaSR-related diseases (page 17). Previous work did not result in reliable screening hits using structure-based docking experiments. The new strategy involved the screening of small molecules libraries with the ligand-based in-silico profiling tool SAFAN-ISPSM (Structural.And.Functional.Analysis & In-Silico Profiling of Small Molecules) (70) (page 119).

2.6 Virtual profiling with SAFAN-ISPSM

Classical virtual screening approaches usually predict binding affinities of one target against multiple compounds. However, SAFAN-ISPSM allows the quantitative prediction of binding affinities (pK) between a compound library and 4500 targets from fifteen protein classes using a fragment-weight assignment algorithm. The predictions are based on re-evaluated ChEMBL23 binding data (59) by comparing the created compound- and fragment-fingerprints with the Tanimoto similarity. More details about the method are explained on page 119. The profiling process is initiated with compounds in SMILE format (page 123).

The first profiling attempt included a compound library with 9097 small molecule entries of DrugBank Version 5.1.0 (25) and 26490 entries of FooDB Version 1.0 (71). Peptides were not part of the profiling. FooDB is a database containing information about the chemical and biological properties of phytochemicals. It was chosen to identify potential nutritional compounds that act on CaSR. Before the compounds were run with SAFAN-ISPSM, over 13000 most diverse compounds were extracted from the ChEMBL23 database to evaluate the ligand-based method (Figure 2.25). The evaluation was executed with the leave-one-out-cross-validation

method and should hint on how well the predictions correlate with the true experimental values. One difficulty was that some interactions were listed multiple times by different experimental setups. The measured binding affinity (pK) often deviated by more than one order of magnitude.

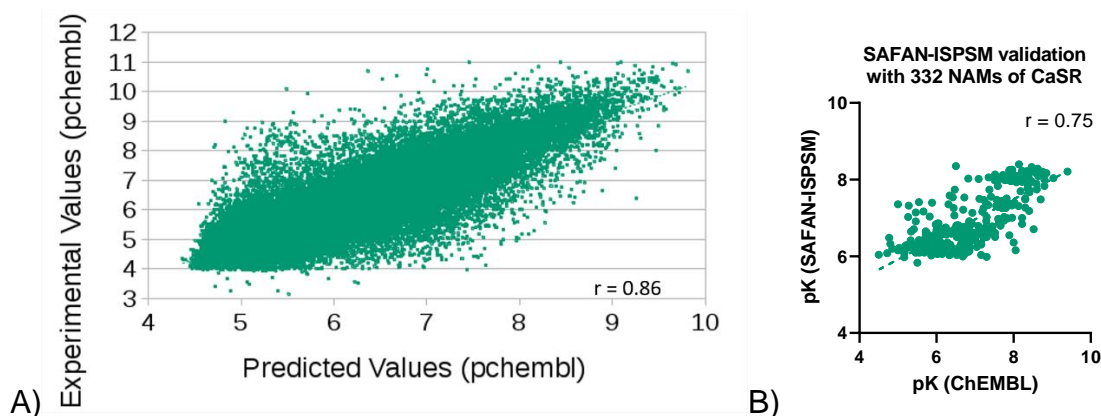


Figure 2.25 In-silico profiling with SAFAN-ISPSM; A) leave-one-out-cross-validation of the method with over 13 000 maximum diverse compounds randomly selected from the ChEMBL23 database (DOI: 10.6019/CHEMBL.database.23); [this image was provided by S.A.F.AN.-BIOINFORMATICS (70)]; B) leave-one-out-cross-validation with 332 known negative allosteric modulators (NAM) for CaSR

The validation (Figure 2.25-A) shows that the predicted binding affinities have a Pearson correlation of $r = 0.86$ with the data listed in ChEMBL23. Most interactions are within an error of one order of magnitude. The leave-one-out-cross validation with 332 known negative allosteric modulators gave a Pearson correlation of $r = 0.75$. This result led to the conclusion that identifying a novel allosteric modulator is far more probable than with molecular docking. However, ligand-based methods are limited by the Tanimoto similarity approach because only compounds are ranked at the top that shares significant similarity with known calcimimetics and calcilytics. It is unrealistic to identify a completely novel scaffold with this method.

Nevertheless, the compound library was profiled with SAFAN-ISPSM Version 2018/06-b1 on the C3S Occam supercomputer (72) using a node with 24 CPUs. 9097 small molecule entries of DrugBank were profiled in six hours and 26490 entries from FooDB in ten hours. SAFAN-ISPSM created a profiling list for every compound. An example is shown in Table 2.7. Interactions are only kept if the predicted pK is greater than 4. The results are ranked by predicted pK values starting with the highest.

Table 2.7 SAFAN-ISPSM version 2018/06 profiling results of DrugBank entry DB08656; position: target ranking position in the profiling list; target: predicted target that interacts with the input compound; pK: predicted binding affinity; alltan: Tanimoto similarity between the input compound and the most similar ChEMBL compound; fragtan: Tanimoto similarity between the fragment fingerprints and the ChEMBL fragment fingerprints; round: compounds without (0) or with (1) a binding affinity relevant fragment (BARF)

position	target	pK	alltan	fragtan	round
1	CAH12	7.756	0.519	0.651	0
2	CAH2	7.506	0.522	0.651	0
3	CAH9	7.506	0.515	0.651	0
4	ADRB2	7.476	0.559	0.666	0
5	DRD3	7.143	0.525	0.696	0
6	TRPV1	6.956	0.623	0.664	0
7	UROK	6.858	0.589	0.604	0
8	CAH1	6.757	0.504	0.651	0
9	ADRB1	6.745	0.544	0.707	0
10	CASR	6.717	0.514	1.000	0
...
60	1CYSP_nomam	4.67	0.641	0.739	0

The identification of potential allosteric modulators started by searching the profiling lists for interactions with CaSR. In the end, 186 small molecules of DrugBank and 210 compounds from FooDB were predicted to bind CaSR significantly.

The next step included the experimental evaluation of the most promising compounds. They were tested with a G_q-IP1 assay for allosteric activity. However, from the 396 identified small molecules, only one had CaSR listed as the primary target. The idea was to test more than one compound, but the budget did not allow to evaluate all of them. Another requirement was to select compounds that could be bought from stock and did not have to be synthesised. In the first selection process (Figure 2.26), compounds were chosen with a pK higher than 6.0 and a CaSR target position below 22. Also, entries were preferred with a round equal to 1. This parameter indicates that the compound contains a binding affinity relevant fragment (BARF). In the end, 52 compounds fulfilled these criteria (Figure 2.26).

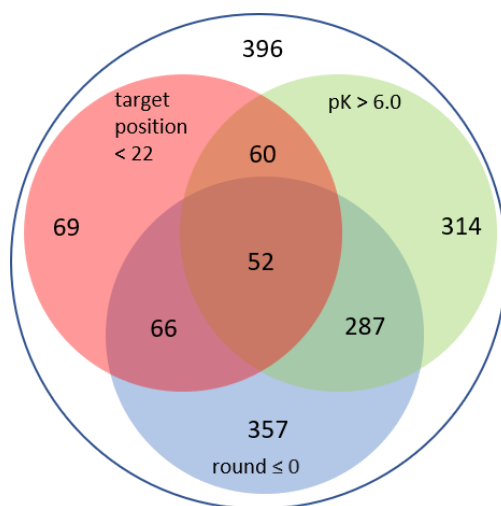
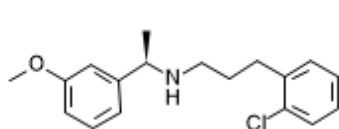
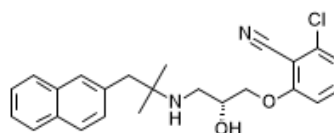


Figure 2.26 Venn diagram of 396 compounds interacting with CaSR and the pre-selection process including the parameters target position, pK and round

The resulting compounds were selected for docking experiments with the created model of CaSR in the inactive state (Figure 2.18). Docking was performed to clarify if the compounds fit into the binding pocket and interact with relevant residues (Figure 2.2). The final selection considered only compounds ready for purchase; they were not too expensive and had a short shipping time (Table 2.8). The final selection included the nonsteroidal selective estrogen receptor modulator toremifene citrate, the cholinesterase inhibitor rivastigmine tartrate, the food and flavour ingredient 1-phenyl-1-propanol and the beta-adrenergic receptor antagonists bupranolol, bucindolol, alprenolol and oxprenolol. The beta-blockers share a common scaffold with the aryloxy propanolamines (Table 4.4). The experiments were performed with the known positive allosteric modulator NPS-R568 and the negative allosteric modulator NPS-2143 as a reference (Figure 2.27).



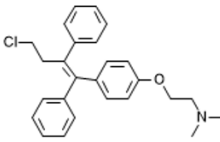
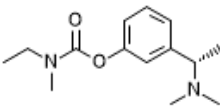
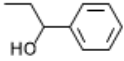
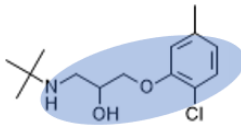
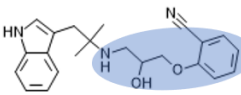
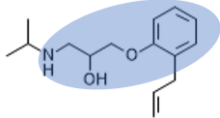
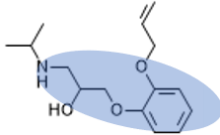
Tecalcet (NPS-R-568); pK: 7.10 (80 nM)



NPS-2143; pK: 7.37 (43 nM)

Figure 2.27 Reference compound tecalcet and NPS-2143 for the experimental validation with an IP1-Gq assay

Table 2.8 first screening subset for experimental validation; blue: aryloxy propanolamine scaffold

<p>1. Toremifene citrate</p> <p><u>IUPAC:</u> (2-{4-[(1Z)-4-chloro-1,2-diphenylbut-1-en-1-yl]phenoxy}ethyl) dimethylamine</p> <p><u>Smile:</u> CN(C)CCOC1=CC=C(C=C1)C(=C(\CCCl)C1=CC=CC=C1)\C1=CC=CC=C1</p> <p><u>predicted pK:</u> 7.51 (31 nM)</p>	
<p>2. Rivastigmine</p> <p><u>IUPAC:</u> 2,3-dihydroxybutanedioic acid; 3-[1-(dimethylamino)ethyl]phenyl N-ethyl-N-methylcarbamate</p> <p><u>Smile:</u> CCN(C)C(=O)OC1=CC=CC(=C1)[C@H](C)N(C)C</p> <p><u>predicted pK:</u> 6.31 (490 nM)</p>	
<p>3. 1-Phenyl-1-propanol</p> <p><u>IUPAC:</u> 1-phenylpropan-1-ol</p> <p><u>Smile:</u> CCC(O)C1=CC=CC=C1</p> <p><u>predicted pK:</u> 6.62 (240 nM)</p>	
<p>4. Bupranolol</p> <p><u>IUPAC:</u> 1-(tert-butylamino)-3-(2-chloro-5-methylphenoxy)propan-2-ol</p> <p><u>Smile:</u> CC1=CC(OCC(O)CNC(C)(C)C)=C(Cl)C=C1</p> <p><u>predicted pK:</u> 6.34 (457 nM)</p>	
<p>5. Bucindolol</p> <p><u>IUPAC:</u> 2-(2-hydroxy-3-[[1-(1H-indol-3-yl)-2-methylpropan-2-yl]amino]propoxy)benzonitrile</p> <p><u>Smile:</u> CC(C)(CC1=CNC2=CC=CC=C12)NCC(O)COC1=CC=CC=C1C#N</p> <p><u>predicted pK:</u> 6.71 (195 nM)</p>	
<p>6. Alprenolol</p> <p><u>IUPAC:</u> 1-[2-(prop-2-en-1-yl)phenoxy]-3-[(propan-2-yl)amino]propan-2-ol</p> <p><u>Smile:</u> CC(C)NCC(O)COC1=CC=CC=C1CC=C</p> <p><u>predicted pK:</u> 6.39 (407 nM)</p>	
<p>7. Oxprenolol</p> <p><u>IUPAC:</u> 1-[2-(prop-2-en-1-yloxy)phenoxy]-3-[(propan-2-yl)amino]propan-2-ol</p> <p><u>Smile:</u> CC(C)NCC(O)COC1=CC=CC=C1OCC=C</p> <p><u>predicted pK:</u> 6.11 (776 nM)</p>	

Experimental validation with the Cisbio IP-1 Gq assay

The experimental evaluation of the in-silico profiling results was executed with the cisbio IP-1 Gq assay. This assay was chosen because at higher calcium concentrations ($EC_{50} = 4,75 \text{ mM Ca}^{2+}$) CaSR is activated and mainly transduces through G_i and G_q proteins. The heterotrimeric G_q -protein dissociates and acts on phospholipase C β (PLC- β) which converts phosphatidylinositol-4,5-bisphosphate (PIP_2) into diacylglycerol and D-myo-inositol 1,4,5-triphosphate (IP_3) (**Fehler! Verweisquelle konnte nicht gefunden werden.**). In this assay, LiCl inhibits the degradation of IP_3 to D-myo-inositol and leads to IP_1 accumulation.

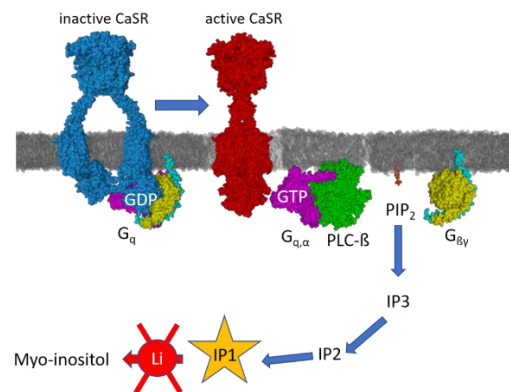


Figure 2.28 CaSR Gq activation and IP1 accumulation, GDP: guanosine diphosphate; GTP; guanosine triphosphate; PLC β : phospholipase C β ; IP₃: D-myo-inositol 1,4,5-triphosphate; IP₂: inositol-4-5-phosphate; IP₁: inositolphosphate

The increasing IP_1 concentration is detected with the homologous time-resolved fluorescence (HTRF) technology. HTRF combines time-resolved measurements (TR) with Förster resonance energy transfer (FRET) (73). FRET is based on an energy transfer between a donor and acceptor dye if they are close. The result is the emission of a specific concentration- and distance-dependent fluorescence signal. In the assay, cells generate native IP_1 through G_q signalling. The native unlabelled IP_1 is extracted and mixed with acceptors (d2-labelled IP_1) and donor molecules (anti- IP_1 -Tb-cryptate). The native IP_1 competitively binds to the donor and reduces the acceptor-donor FRET signal inversely proportional. The fluorescence signal is measured at 620 and 665 nm using an HTRF compatible reader (74). The ratio of the two signals is calculated according to Equation 7, to eliminate background noise.

$$Ratio = \frac{Signal\ 665\ nm}{Signal\ 620\ nm} \times 10^4$$

Equation 7

The Gq-IP1 measurements were executed according to the protocol on page 164. The predicted pK-SAFAN helped to choose the first ligand concentrations for the experiment. Still, it did not provide any information if the ligands would act as agonists/antagonists or as positive/negative allosteric modulators. The first experimental setup included the measurement at 0.01-, 1- and 100-times ligand concentration without calcium ($[Ca^{2+}] = 0\ mM$) (Figure 2.29). The different concentrations were chosen because it was not clear if and how much the predicted pK deviates from the true value.

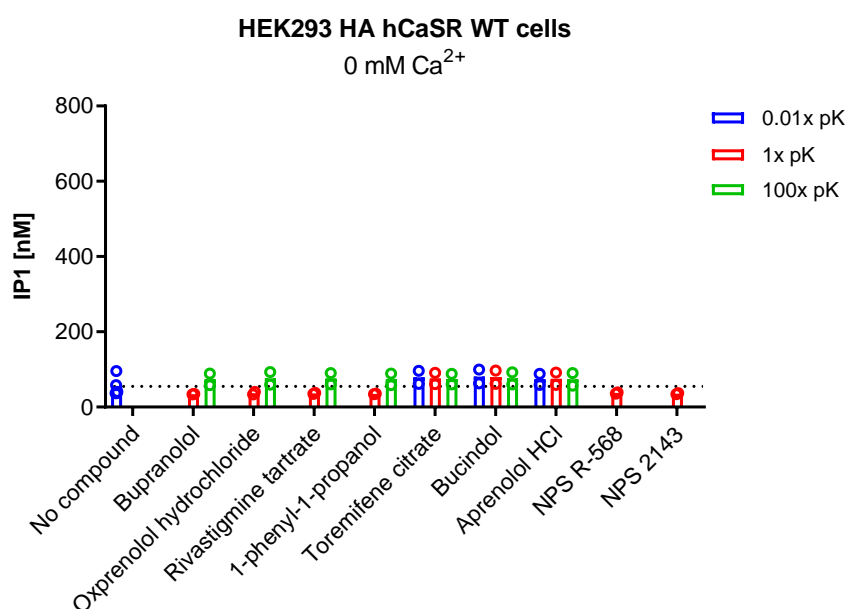


Figure 2.29 IP-1 accumulation assay to determine the agonistic activity of the ligands

The results show that no compound act as an agonist at CaSR up to 100 x pK concentration. It also demonstrates that the allosteric modulators NPS-R-568 and NPS-2143 do not act as an agonist but shift the potency of orthosteric ligands. The next setup involved the measurement at different calcium concentrations. The setpoint for the human wildtype CaSR was identified at 4.5 mM extracellular calcium. Therefore, measurements were performed at 2.5 mM calcium to identify positive allosteric modulators.

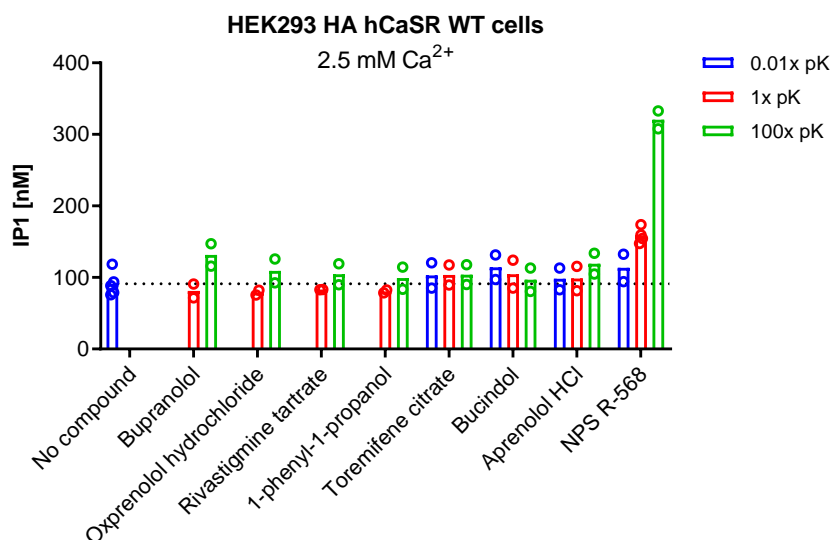


Figure 2.30 IP1 accumulation assay with the first compound library at 2.5 mM calcium to identify positive allosteric modulators of CaSR

The HTRF readout shows that the known positive allosteric modulator NPS-R-568 give a significant signal at 1- and 100 times pK concentration. It also shows that the tested compounds do not have positive allosteric modulator activity. The last setup was chosen with 7.5 mM calcium to identify negative allosteric modulators (Figure 2.31).

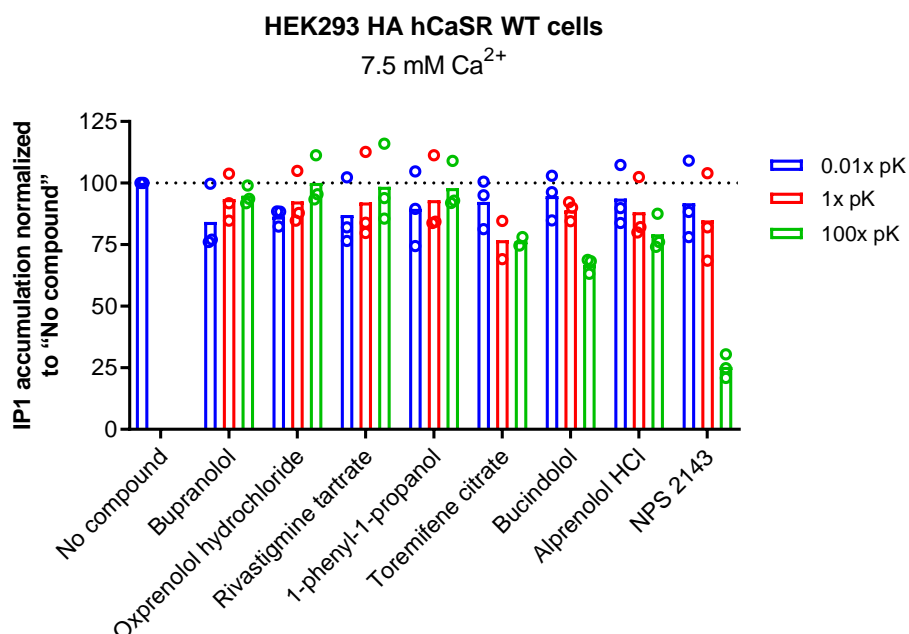


Figure 2.31 IP1 accumulation assay with the first compound library at 7.5 mM calcium to identify negative allosteric modulators of CaSR (normalized to no compound)

The normalized HTRF results indicate that none of the compounds has a similar potency as the reference NPS-2143. Toremifene citrate and bucindolol show insignificant negative allosteric activity at high concentrations. CaSR gets modulated from both compounds at 100 times pK concentration which corresponds to 3100 nM for toremifene and 19500 nM for bucindolol. At such high concentrations, interactions with other targets are very probable and could interfere with the results. A brief structure-activity relationship analysis reveals that toremifene hardly shows any similarity to known calcilytics of the ChEMBL database. Still, bucindolol shares a bunitrolol scaffold with NPS-2143 and ChEMBL371936 (Figure 2.32, blue) (Table 4.4). The calcilytic NPS-2143 is the most potent one ($IC_{50} = 43$ nM) (75). It has a linear shape containing a 2-naphthyl at the tert-butylamino end and a 2-chloro group at the benzonitrile. Previous studies showed that the R-enantiomers of the aminoalcohols are more potent and selective for CaSR.

In contrast, the S-enantiomers are more potent at adrenergic receptors of the class A GPCRs. Any changes to the geminal dimethyls lead to loss of potency (76). The compound CHEMBL371936 is still very potent with an IC_{50} of 100nM. The differences to NPS-2143 are the missing chlorine, the S-chirality in the aminoalcohol linker and the (1H)-indol-2-yl structure instead of the naphthalen-2-yl. The non-selective adrenoceptor blocker bucindolol is nearly identical to CHEMBL371936. But the (1H)-indol-3-yl group deviates from the linear shape and results in a tremendous loss of potency. The influence of chirality could not be tested for bucindolol because only the racemic mixture was available for the experiments.

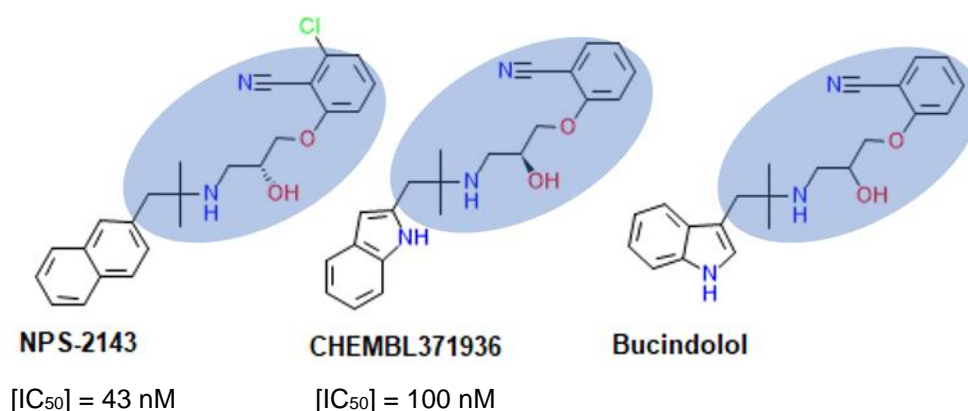


Figure 2.32 Comparison of bucindolol with known allosteric modulators reveals a common butrinolol scaffold [2-(3-(tert-Butylamino)-2-hydroxypropoxy)benzonitrile] (blue)

Conclusion of the first compound profiling with SAFAN-ISPSM

SAFAN-ISPSM allows fast profiling of huge compound databases compared to molecular docking experiments. Also, the predicted binding affinities have a higher correlation to the experimental values ($r = 0.75$) and additional parameters of SAFAN-ISPSM help select top-ranking compounds. Unfortunately, only seven out of 396 predicted compounds could be tested with a G_q -IP1 assay. The experiments showed that none of the compounds has a comparable potency as the known calcilytic NPS-2143. Minimal modulator activity showed the non-steroidal selective estrogen receptor modulator toremifene and the non-selective adrenoreceptor blocker bucindolol. The scaffold of toremifene is not similar to any known allosteric modulator of CaSR but bucindolol is nearly identical to the known entry CHEMBL371936. The conclusion is that the predicted binding affinities deviate at least two orders of magnitude from the true values. This difference could be the comparison to experimental binding affinities that were determined in various setups (cell lines, methods, CaSR variants). Besides, chirality should play a more important role to predict the binding affinity. A closer look at the Tanimoto similarity revealed a slight overrepresentation of the compound atoms. Consequently, the SAFAN-ISPSM algorithm was adapted for the next profiling attempt. The high profiling speed leads to the conclusion to use a larger compound library that increases the chance of identifying a novel allosteric modulator.

2.7 Second virtual profiling approach with SAFAN-ISPSM

The results of the first in-silico profiling and the experimental validation lead to the conclusion that the Tanimoto similarity calculation of SAFAN-ISPSM needed to be optimized. One obstacle was the overrepresentation of compound atoms resulting in deviating Tanimoto similarities. Also, the chirality was taken into account with a higher priority for the predicted binding affinity. The difficulty was also the correct representation of the compound structure with the Tanimoto fingerprint. The best fingerprint generation is a delicate operation that balances between a too general and a too specific match. The algorithm should also identify identical compounds that are written as a different generic SMILE (page 123). This point is very important in case of tautomers (Figure 2.33).

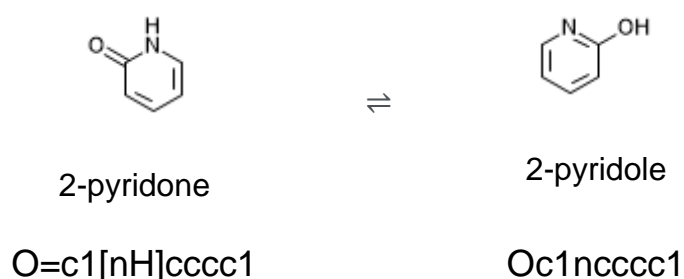


Figure 2.33 keto-enol tautomerism of 2-pyridone

A limitation of many similarity search programs is the unprecise match between two identical chemical compounds. 2-pyridone and 2-pyridole have different smiles, but it represents the same compound from a chemical point of view. Openbabel is a chemical toolbox is widely used in bioinformatics. It is open-source software and allows working with different chemical file formats (77). But it has its limitations. The Tanimoto similarity calculated with Openbabel version 2.3.2 of the example above results in a Tanimoto coefficient of 0.42. This mismatch would lead to wrong predictions and missed potential compounds. SAFAN-ISPSM was adapted to identify tautomers as different representations of the same compound. The method was evaluated with over 21000 maximum diverse compounds involving more than 65000 interactions. The leave-one-out-cross-validation results are a Pearson Correlation of 0.91 for SAFAN-ISPSM version 2019/06-b2 (Figure 2.34).

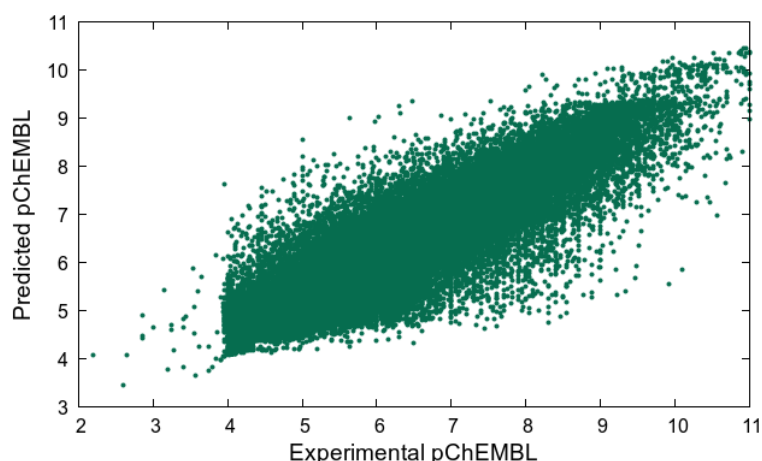


Figure 2.34 SAFAN-ISPSM leave-one-out cross-validation of 21090 diverse compounds involving 65936 interactions (Pearson Correlation $r = 0.91$) (70)

Leave-one-out cross-validation of SAFAN-ISPSM with CaSR ligands

The performance of SAFAN-ISPSM was tested before the virtual profiling of any new compound library. The binding affinities and SMILES of 551 calcimimetics and calcilytics were extracted from the ChEMBL database. The evaluation was executed with leave-one-out cross-validation to identify the correlation between the experimental determined pChEMBL values (negative logarithm of IC₅₀, XC₅₀, AC₅₀, Ki, EC₅₀, K_d or potency) (59) and the predicted binding affinities of SAFAN-ISPSM. The Pearson correlation is 0,81 between the ChEMBL entries and the prediction from SAFAN-ISPSM (Figure 2.35). The evaluation also resulted in 41 false-negative predictions because the calculated binding affinities were below the cut-off ($pK \geq 4,0$) (70).

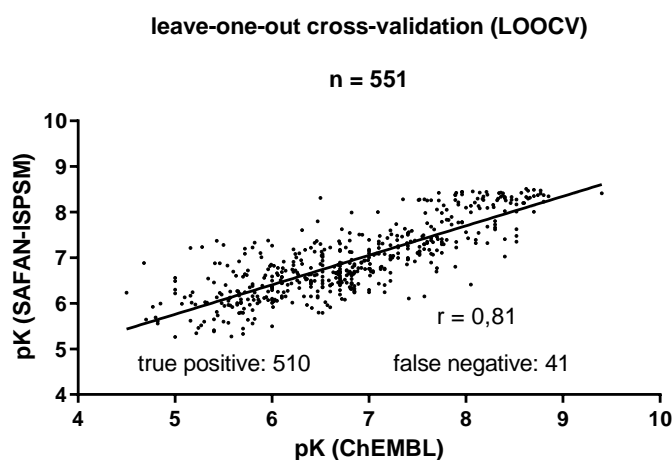


Figure 2.35 Leave-one-out cross-validation of the ChEMBL dataset for CaSR-ligand interactions

Proof-of-concept validation of SAFAN-ISPSM

The proof-of-concept validation was performed as an internal control of SAFAN-ISPSM and to identify the true positives. The idea was to test the method if it could correctly detect known compounds that are usually excluded from the dataset during the leave-one-out-cross validation. Therefore, all smiles were kept during the proof-of-concept validation, but only the IDs were changed of the known CaSR ligands for the pK prediction. The predicted pK values had a Pearson correlation of 0,945 to the ChEMBL validation set. The deviation results from multiple pK values for the same interaction and a slight mismatch between the structures in SAFAN-ISP fingerprints. Consequently, 55 out of 1000 compound binding affinities would not be identified correctly. The predictions show that those interactions deviate up to two orders of magnitudes from the true value.

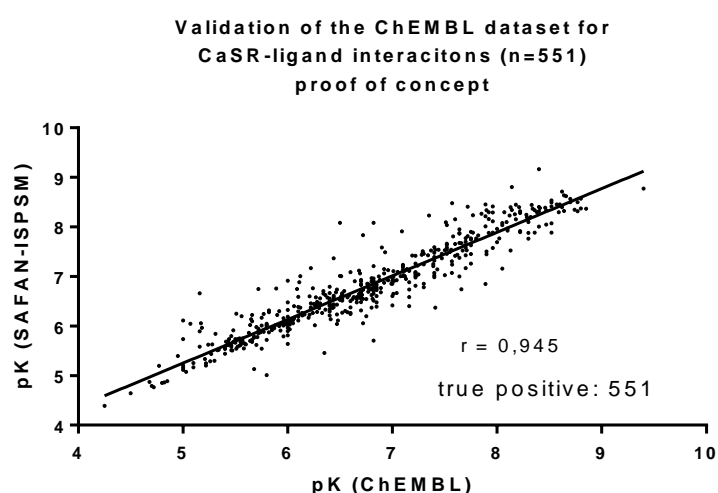


Figure 2.36 Proof of concept validation with the ChEMBL data set for CaSR-ligand interactions

In general, the limitation of virtual screening or profiling methods is the used compound library's size. Also, compound similarity calculations make it impossible to predict a new compound with a binding affinity beyond the listed entries of the dataset (78).

New compound library for virtual profiling

The idea for the next profiling attempt was to use an updated version of SAFAN-ISPSM and a larger compound library to increase the probability of identifying a novel allosteric modulator. The new compound library included 9480 entries of DrugBank Version 5.1.2 (25) and the 26490 entries of FooDB Version 1.0 (79). Other databases were searched, and 4245 entries of the HerDing database (80) also became part of the compound library. This database lists 4880575 herb-chemical-gene relationships, 88217 herb-chemical relationships, 228904 chemical-gene relationships, 19476 herb names, 6655 chemical names and 16762 gene names (80). The database contains 4222 PubChem IDs, which could be used to extract the corresponding SMILES from the PubChem database (57). The entries with the PubChem CID 54742560 and PubChem CID 54742561 were split into individual compounds, and aluminium oxide (PubChem CID 9989226) was excluded. During another search for additional small molecules, 36115 of 57423 entries were available for profiling from the collection of traditional Chinese medicines (TCM) listed at TCM database@Taiwan (81). The SMILES of the TCM database@Taiwan version 2014-01-31 were downloaded from the ZINCdatabase (82). At the time, the ZINC12 database was the latest version which contains more than 35 million purchasable compounds (83). This compound library represented a huge opportunity to identify potential ligands but was too big to profile all entries.

One option at the database website allows the search for compounds with common substructures. The fragment-weight-assignment algorithm version A2 from SAFAN-ISPSM identified 19 binding affinity relevant fragments (BARF) for CaSR (70). The BARFs were used for a substructure search at the ZINC12 database and resulted in 63400 hits representing 57832 unique compounds. Shortly after, the ZINC15 database went online and was searched with 33 BARFs resulting from the fragment weight assignment algorithm version B2 (70). The new ZINC15 database contains more than 750 million purchasable compounds ready for virtual screening (84). The substructure search resulted in 247619 entries in total, which corresponds to 173645 unique entries. Another option was found to identify potential compounds at the ZINC database with the structure-based search engine ZINCPharmer.

ZINCPharmer searches more than 20 million compounds at the ZINC database, with a total of more than 200 million conformations using a pharmacophore model (83). This online interface supports pharmacophore models generated with LigandScout, MOE, Pharmer, LigBuilder, Pharmagist and a lot of other structure programs and their file formats (78,85–88). From the list of programs, the free available structure-based software LigBuilder 2.0 was used to create the pharmacophore models (78). LigBuilder is a de-novo design program and requires the target structure of the macro molecule. It uses the program Cavity 1.0 to identify binding pockets and creates pharmacophore models. In the next step different strategies like the growing mode, linking mode or exploring mode can be selected to generate new compounds. The program was tested but only resulted in linear aliphatic compounds. These structures did not seem to be reasonable to act as allosteric modulators at CaSR. Therefore, LigBuilder was only used to create the pharmacophore models and to search ZINCPharmer for potential ligands. Luckily, new template structures of the metabotropic glutamate receptor 5 were published of the intracellular domain truncated receptors in the active and inactive state (PDB: 6N51; 6N52). The new structures made it possible for the first time to model the seven-transmembrane domain of CaSR in the active state. The hypothesis was that a combination of all published template structures would result in improved homology models in the active and inactive state. The new models could help to identify potential active compounds with ZINCPharmer and could be used during the final selection process.

Homology modelling of CaSR in the inactive state

At the time, there has not been a crystal structure of the whole receptor and homology modelling was the method of choice to create a three-dimensional model. As described on page 12, the extracellular calcium-sensing receptor is composed of three major domains. The N-terminus consists of the extracellular domain build up by two lobes and a cysteine-rich domain forming a Venus-flytrap motif (ECD, residue 20-612). The middle part of the protein includes seven-transmembrane alpha-helices (7TM) as well as three extracellular loops (EL1 – 3) and three intracellular loops (ICL1 – 3) (residues 613-862). The C-terminal part of

the protein involves the intracellular domain (ICD, residues 863-1078). The binding pocket for CaSR's allosteric modulators was identified by mutagenesis studies and homology modelling in the upper part of the seven-transmembrane domain between helix III, V, VI and VII (26). The relevant residues for the negative allosteric modulation activity are depicted in Figure 2.2. The inspection of former models' binding pocket (Figure 2.18) revealed that the extracellular loops interact with allosteric modulators. However, the extracellular loops were not consistent in the crystal structures of metabotropic glutamate receptor 1 and 5. It was not clear if they form an open or closed conformation. The right loop conformer could act as a filter for allosteric modulators and improve the selection of compounds that can reach into the binding pocket. The result of a literature search revealed that GPCR loop modelling is impaired by imprecise homology modelling of the seven-transmembrane helical cores. The precision improves by constraints like neighbouring loops, resulting in a decrease of conformational space (89). This concept became very important after the publication of the latest crystal structures in 2019. The structures represent the extracellular and seven-transmembrane domain of the metabotropic glutamate receptor 5 in the active (PDB: 6N51) and inactive state (PDB: 6N52). The new crystal structures made it possible to identify the conformation of ECL 2 the inactive as well as the active state. Further constraints for the intracellular loops of the inactive CaSR were established by creating a full model in complex with the human Gq-protein. The complex was based on known GPCR G-protein complexes utilising structural alignment. Protein structure prediction was used for the intracellular domain. The modelling procedure for CaSR was performed according to Constanzi's guidelines (10) and extended for a full receptor Gq-protein complex in the inactive state (Figure 2.37).

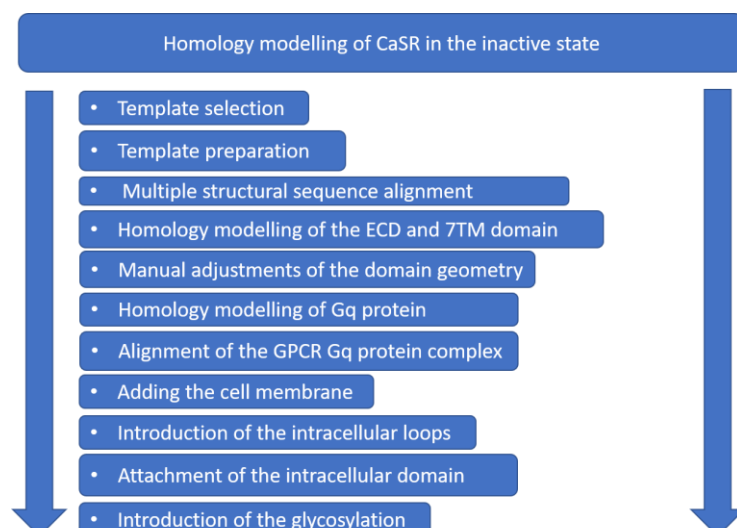


Figure 2.37 Homology modelling process of the full-length CaSR – Gq protein complex

The homology modelling of the inactive CaSR conformer was performed with YASARA Structure Version 19.1.27.W.64 (24) (page 114). The procedure started with the template structure search and identification of relevant entries at Protein Data Bank (43) and Protein Data Bank in Europe (90). Eight crystal structures Table 2.9 were selected as templates for homology modelling of CaSR in the inactive conformation.

Table 2.9 Class C GPCR templates; CASR: extracellular calcium sensing receptor; GRM1: metabotropic glutamate receptor 1; GRM5: metabotropic glutamate receptor 5; oligo.state: oligomerisation state; aa.: number of amino acids; ref.: reference

GPCR	PDB ID	oligo. state	resolution	aa.	released	ref.
CASR	5K5T	homodimer	3.10 Å	1156	2016-05-23	(9)
GRM1	4OR2-A	protomer	2.80 Å	239	2014-03-19	(49)
GRM1	4OR2-B	Protomer	2.80 Å	239	2014-03-19	(49)
GRM5	6FFI	monomer	2.20 Å	210	2018-03-07	(50)
GRM5	6FFH	monomer	2.65 Å	210	2018-03-07	(50)
GRM5	5CGD	monomer	2.60 Å	210	2015-08-12	(51)
GRM5	5CGC	monomer	3.10 Å	210	2015-08-12	(51)
GRM5	6N52	homodimer	4.00 Å	1530	2019-01-23	(91)

The alignment of CaSRs sequence (Figure 1.4) and the templates was created at the GPCR database (11) between CaSR, mGlu1 and mGlu5. The structural sequence alignment was used for homology modelling and made a secondary structure prediction obsolete (Figure 2.38). The template structures required preparation before they could be used for modelling experiments. The templates

were downloaded in PDB format except for PDB 5K5T. CaSR's extracellular domain was downloaded as a homodimer in the biological assembly 2 format.

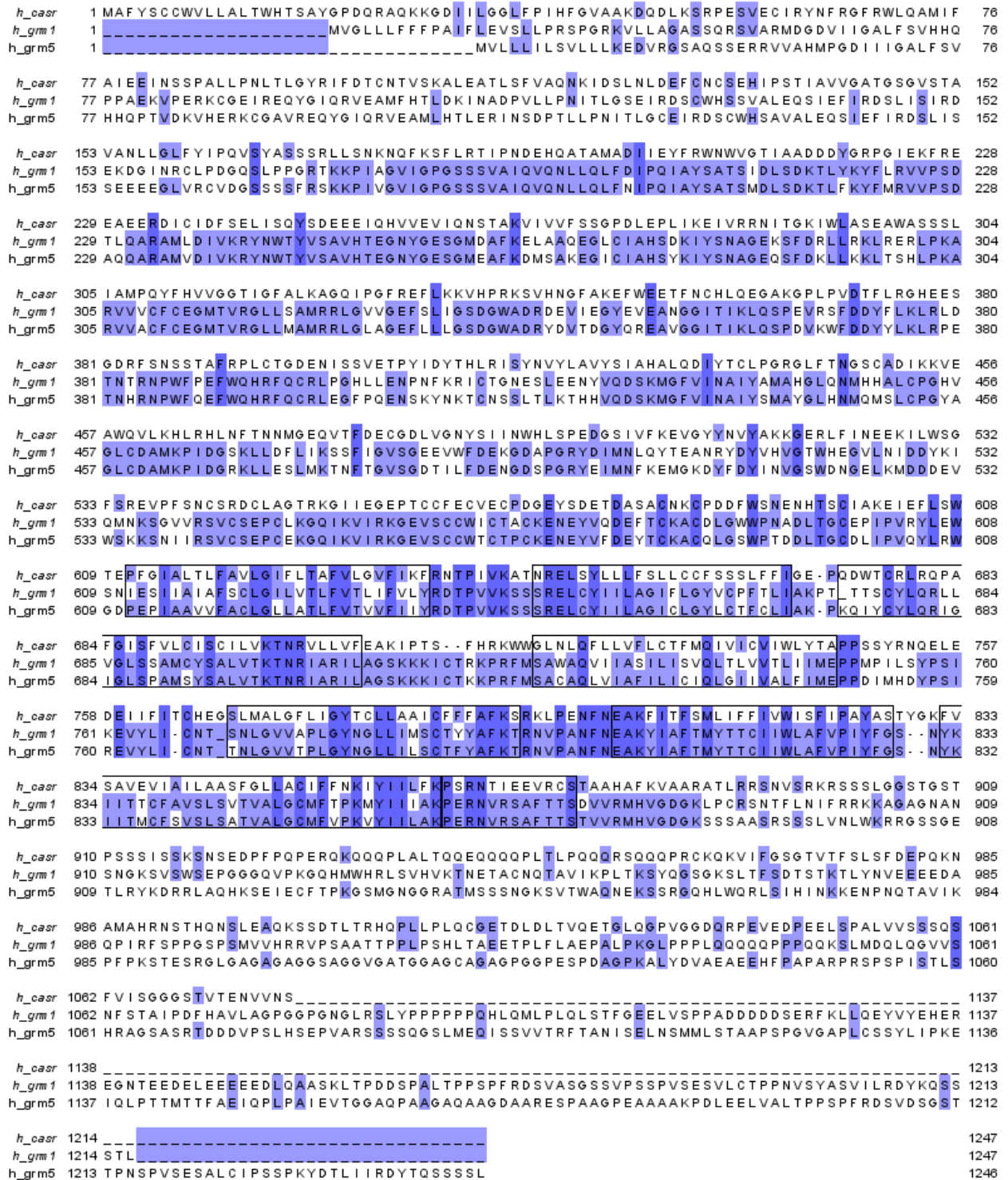


Figure 2.38 Structural alignment of CaSR, mGlu1 and mGlu5 sequence acquired from GPCRdb (11), blue: sequence identity at the aligned position, boxes represent the transmembrane helices

All structures were cleaned from non-protein entries; missing hydrogens were added and saved as template structures. The homodimer of mGlu1 (PDB: 4OR2) was split up into protomers. The extracellular loops ECL1 (residue 650-653), ICL2 (residue 686-687) and ECL3 (residue 696-701) were removed as well as the remaining intracellular domain (residue 808-809) and the thermostabilising fusion protein BRIL (92). The structure of mGlu₅ (PDB 6FFI) was modified by removing residues 568-577, 637-639, 673-840, 876-895, 955-956 and 989-992. The remaining structure included the seven-transmembrane alpha-helices, ICL1 and ICL3 (93). The same structural elements were left after removing residues 568-577, 637-639, 673-678, 1002-1680, 1716-1724, 1728-1735, 1795-1796 and 1829-1832 from PDB 6FFH (93). The structures PDB 5CGD and 5CGD were prepared by deleting residues 1002-1161, 568-577, 637-639, 673-680, 716-735, 795-796 and 829-831 (51). PDB 6N52 (mGlu₅) is the latest published crystal structure in the inactive conformation. It consists of an ECD-7TM homodimer. The template preparation included removing ECL1, ECL3 and ICD (residues 637-639 795-796 829-832). Finally, all template structures ([Table 2.9](#)) were prepared for homology modelling.

Homology modelling of the CaSR's ECD-7TM domain in the inactive state

The modelling process started with the eight prepared structures and the structural alignment from GPCRdb between CaSR, mGlu₁ and mGlu₅ (Figure 2.38). The alignment was used to create the alignment file between the template sequences in Jalview (40) (Figure 2.39). The automated homology modelling procedure was set to 'slow' to achieve the best model quality. PSI-BLAST and E-value were not required because the templates were provided manually. One model was built for every template structures with the predefined alignment. Alignment variations were prohibited and the sampling of different loop conformations and the extension of missing residues at the termini.

```

      10      20      30      40      50      60      70      80
Ca SR/1-823 YGF DQRAQKKGD I I L G G L F P I H F G V A A K D Q D L K S R P E S V E C I R Y N F R G F R W L Q A M I F A I E E I N S S P A L L P N L T L G Y R I F D T C N
5K57/1-575 - G P D Q R A Q K K G D I I L G G L F P I H F G V A A K D Q D L K S R P E S V E C I R Y N F R G F R W L Q A M I F A I E E I N S S P A L L P N L T L G Y R I F D T C N
4OR2A/1-233 .....
4OR2B/1-233 .....
6FFV1-203 .....
6FFH1-203 .....
5CDG/1-203 .....
5CGC/1-203 .....
6N52/1-753 ..... R V V A H M P G D I I I G A L F S V H H G P T V D K V H E R K K G A V R E Q Y G I Q R V

      90      100      110      120      130      140      150      160
Ca SR/1-823 T V S K A L E A T L S F V A Q N K I D S L N L D E F C N C S E H I P S T I A V V G A T G S G V S T A V A N L L G L F Y I P Q V S Y A S S R L L S N K N Q F K S F L R
5K57/1-575 T V S K A L E A T L S F V A Q N K I D S L N L D E F C N - - - - P S T I A V V G - T G S G V S T A V A N L L G L F Y I P Q V S Y A S S R L L S N K N Q F K S F L R
4OR2A/1-233 .....
4OR2B/1-233 .....
6FFV1-203 .....
6FFH1-203 .....
5CDG/1-203 .....
5CGC/1-203 .....
6N52/1-753 E A M L H T L E R I N S D P T L L P N I T L G C E I R D S C W H S A V A L E Q S I E F I R D S L I S S E E - - - - - S F R S K K P I V G V I - - G S S

      170      180      190      200      210      220      230      240
Ca SR/1-823 T I P N D E H Q A T A M A D I I E Y F R W N W V G T I A A D D D Y G R P G I E K F R E E A E E R D I C I D F S E L I S Q Y S D E E E I Q H V V E V I Q N S T A K V I V
5K57/1-575 T I P N D E H Q A T A M A D I I E Y F R W N W V G T I A A D D D Y G R P G I E K F R E E A E E R D I C I D F S E L I S Q Y S D E E E I Q H V V E V I Q N S T A K V I V
4OR2A/1-233 .....
4OR2B/1-233 .....
6FFV1-203 .....
6FFH1-203 .....
5CDG/1-203 .....
5CGC/1-203 .....
6N52/1-753 S V A I Q V Q N L L Q L F N I P Q I A Y S A T S M D L S D K T L F K Y F M R V V P S D A Q Q A R A M V D I V K R Y N W T Y V S A V H T E S N Y G E S G M E A F K D M S

      250      260      270      280      290      300      310      320      330
Ca SR/1-823 V F S S G P D L E P L I K E I V R R N I T G K I W L A S E A W A S S S L I A M P Q Y F H V V G G T I G F A L K A G Q I P G F R E F L K K V H P R K S V H N G F A K E F
5K57/1-575 V F S S G P D L E P L I K E I V R R N I T G K I W L A S E A W A S S S L I A M P Q Y F H V V G G T I G F A L K A G Q I P G F R E F L K K V H P R K S V H N G F A K E F
4OR2A/1-233 .....
4OR2B/1-233 .....
6FFV1-203 .....
6FFH1-203 .....
5CDG/1-203 .....
5CGC/1-203 .....
6N52/1-753 A K E G I C I A H S Y K I Y S N A G E Q S F D K L L K K L T S H - - K A R V V A C F C E G M T V R G L L M A M R R L G L A G E F L L L G S D G W A D R Y D V T D G Y Q

      340      350      360      370      380      390      400      410
Ca SR/1-823 W E E T F N C H L Q E G A K G P L P V D T F L R G H E E S G D R F S N S S T A F R P L C T G D E N I S S V E T P Y I D Y T H L R I S Y N V Y L A V Y S I A H A L Q D I
5K57/1-575 W E E T F N C H L Q E G A K - P L P V D T F L R G H E E S G D R F S Q S S T A F R P L C T G D E N I S S V E T P Y I D Y T H L R I S Y N V Y L A V Y S I A H A L Q D I
4OR2A/1-233 .....
4OR2B/1-233 .....
6FFV1-203 .....
6FFH1-203 .....
5CDG/1-203 .....
5CGC/1-203 .....
6N52/1-753 R E A V G G I T I K L Q S P D V K W F D D Y Y L K L R P E T N H R N P W F Q E F W Q H R F Q C R L E G F P Q E N S K Y N K T C N S S L T L K T H H V Q D S K M G F V I

      420      430      440      450      460      470      480      490
Ca SR/1-823 Y T C L P G R G L F T N G S C A D I K K V E A W Q V L K H L R H L N F T N N M G E Q V T F D E C G D L V G N Y S I I N W H L S P E D G S I V F K E V G Y Y N V Y A K K
5K57/1-575 Y T C L P G R G L F T - G S C A D I K K V E A W Q V L K H L R H L N F T N N M G E Q V T F D E C G D L V G N Y S I I N W H L S P E D G S I V F K E V G Y Y N V Y A K K
4OR2A/1-233 .....
4OR2B/1-233 .....
6FFV1-203 .....
6FFH1-203 .....
5CDG/1-203 .....
5CGC/1-203 .....
6N52/1-753 N A I Y S M A Y G L H N M Q M S L C P G Y A G L C D A M - - I D G R K L L E S L M K T N F T G V S G D T I L F D E N G D S P G R Y E I M N F K E M G K D Y F D Y I N V

      500      510      520      530      540      550      560      570      580
Ca SR/1-823 G E R L F I N E E K I L W S G F S R E V P F S N C S R D C L A G T R K G I I E G E P T C C F E C V E C P D G E Y S D E T D A S A C N K C P D D F W S N E N H T S C I A
5K57/1-575 G E R L F I N E E K I L W S G F S R E V P F S N C S R D C L A G T R K G I I E G E P T C C F E C V E C P D G E Y S D E T D A S A C N K C P D D F W S N E N H T S C I A
4OR2A/1-233 .....
4OR2B/1-233 .....
6FFV1-203 .....
6FFH1-203 .....
5CDG/1-203 .....
5CGC/1-203 .....
6N52/1-753 G S W D N G E L K M D D E V W S K S N I I R S V C S E P C E K G Q I K V I R K G E V S C W C T C T P C K E N E Y V F D E Y T C K A C Q L G S W P T D D L T G C D L

```

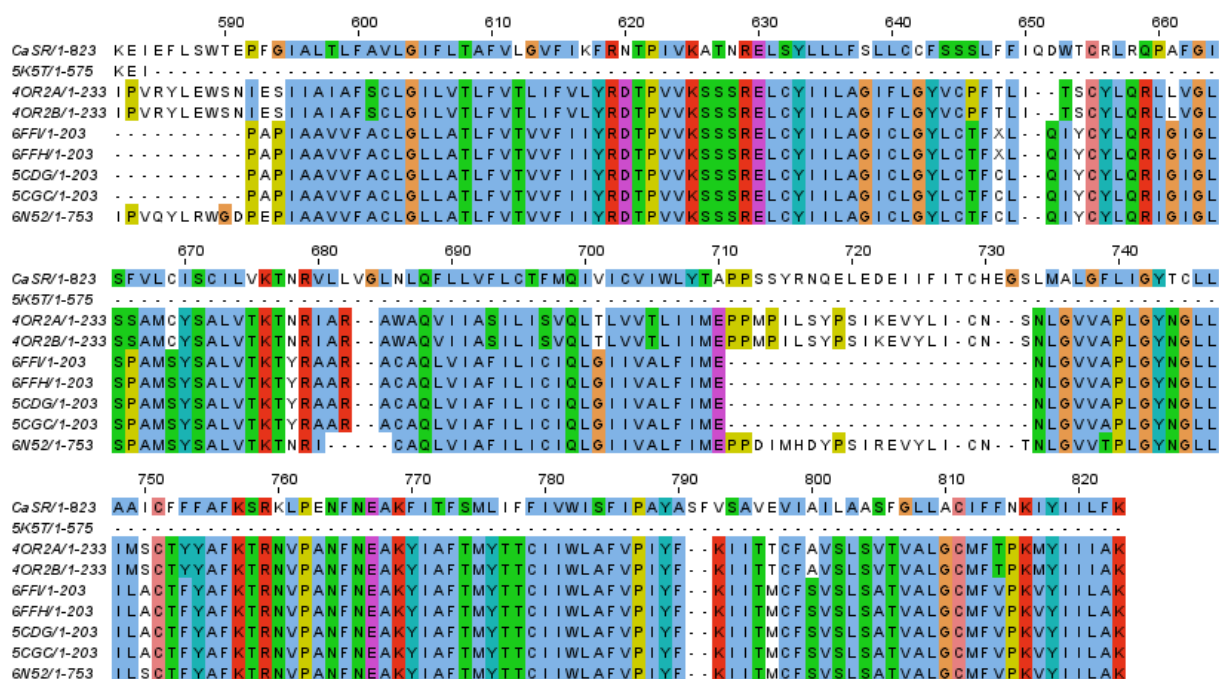


Figure 2.39 Multiple sequence alignment for homology modelling of the ICD truncated CaSR in the inactive conformation; template sequences of CaSR-ECD, mGlu1 and mGlu5 transmembrane were aligned according to the multiple structure alignment available of the GPCRdb website (Figure 2.38). ECL1, ICL2 and ECL3 were omitted in the template structures.

The outcome was eight homology models based on the template structures and one hybrid model. The eight homology models' results are listed in [Table 2.10](#), and the model quality is scored according to the knowledge-based Z-score (Figure 2.40). The Z-Score is defined as the magnitude of standard deviations on average between the model and a high-resolution X-ray structure. High scoring models are more reliable by equal sequence length. The formula of YASARA's Z-score is given in equation 3. The results show that model number 5 has the highest Z-Score of all the transmembrane templates and is based on PDB 6FFH ($Z = -0.275$). Only the first model covers the residues of the extracellular domain, which is based on the crystal structure of CaSR (5K5T). But the results also indicate that model number 8 covers most residues between the target sequence and the templates with the worst Z-score of all the models ($Z = -2.666$). YASARA created a hybrid model out of the eight. It represents the final model with a Z-score of -0,735 and covers 822 residues of the ICD truncated CaSR. The model's major setback was the unrealistic geometry between the extracellular domain and the seven-transmembrane domain (Figure 2.41 A). Besides, the missing TYR20 was added with the 'BuildLoop' command in YASARA (54).

Table 2.10 homology modelling results of 8 models; seq. id.: sequence identity; seq. sim.: sequence similarity based on the BLOSUM62 score (Figure 4.3); residues: the position of the modelled residues

model	oligomeric state	aligned residues	seq. id. [%]	seq. sim. [%]	loops added	Z-score	model ID	residues	comment
1	homodimer	575	99.7	99.8	6	-0.831	5K5T	2-584	good
2	monomer	233	31.8	54.5	6	-0.587	4OR2_A	582-823	good
3	monomer	233	31.8	54.5	6	-0.610	4OR2_B	582-823	good
4	monomer	203	34.5	55.2	5	-0.306	6FFI	592-823	good
5	monomer	203	34.5	55.2	5	-0.275	6FFH	592-823	good
6	monomer	203	34.5	55.2	5	-0.421	5CGD	592-823	good
7	monomer	203	34.5	55.2	5	-0.474	5CGC	592-823	good
8	homodimer	753	14.1	27.9	10	-2.666	6N52	40-823	poor

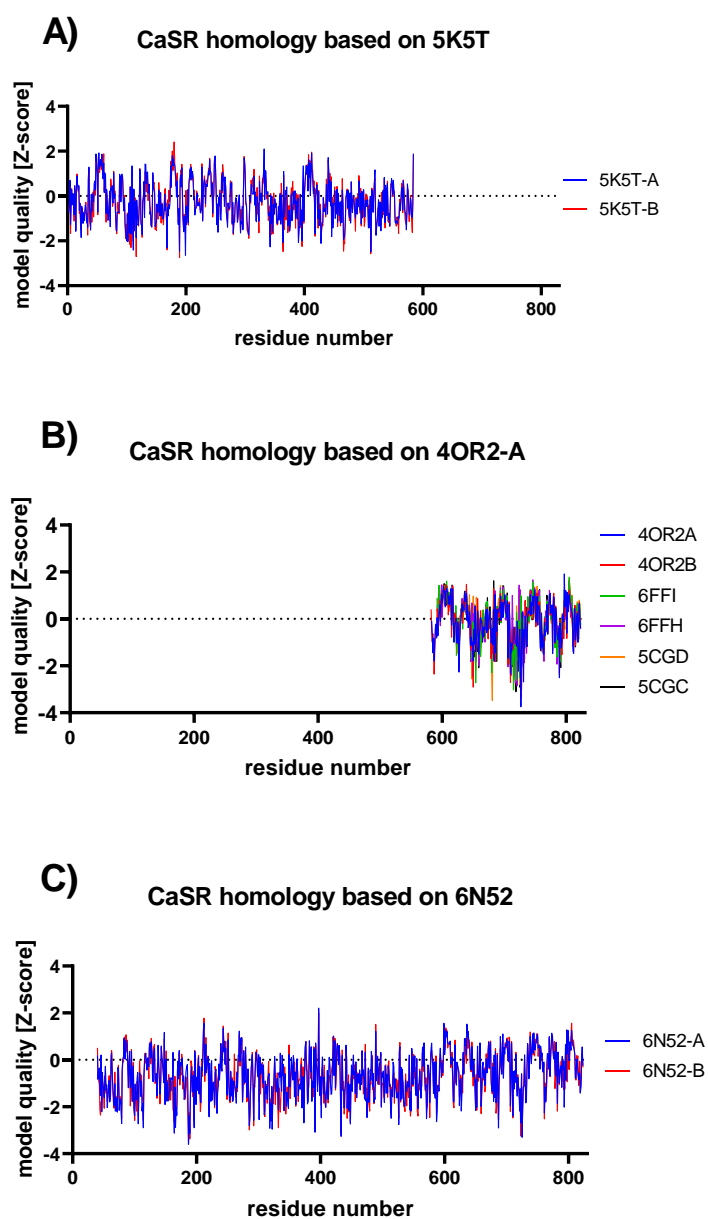


Figure 2.40 Model quality of the created homology models; A) Z-score quality of the extracellular domain; B) Z-score quality of the seven-transmembrane domain; C) Z-score quality of the ICD truncated CaSR

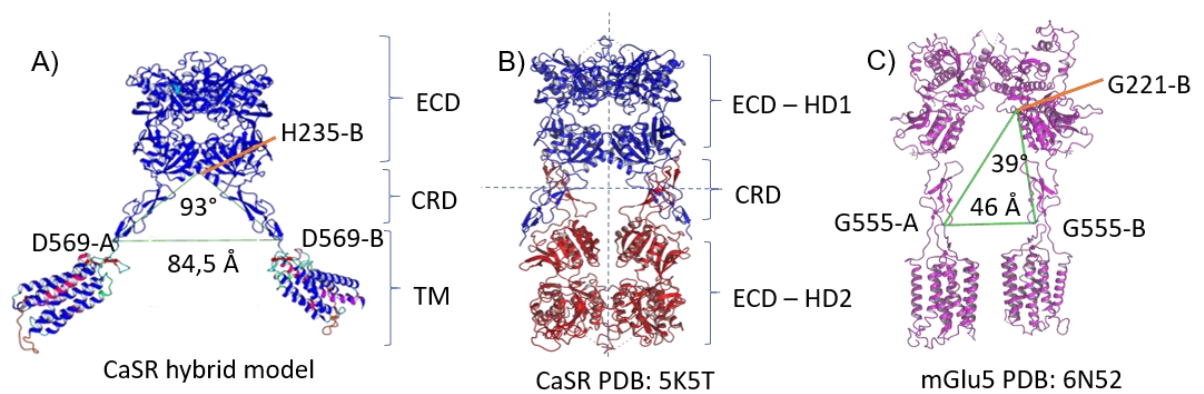


Figure 2.41 Geometric comparison of the final model to crystal structures A) Final homology model generated with YASARA based on the predefined multiple sequence alignment from GPCRdb (11), ECD: extracellular domain, CRD: cysteine-rich domain, TM: transmembrane domain, the CRDs of the homodimer opens up by 93° using D588 from protomer A and H254 and D588 from protomer B as reference points; blue: initial model, different colours: hybrid parts; B) the ECD of CaSR (blue) is staged in the crystal structure PDB: 5K5T, the CRD interacts with the ECD of a second homodimer (red) and with its CRD, forming an unnatural angle between the CRDs; C) crystal structure PDB: 6N52 representing mGlu5 in the inactive state, this structure contains the ECD, CRD and the TM, the geometry of this receptor suggests that the protomers are arranged parallel and that the TMs are not in direct contact, but the parallel orientation seems to be the consequence of the membrane

Figure 2.41 A shows the CaSR hybrid model with the blue template and the hybridisation parts in an alternative colour. In this model, the tertiary structure of the domains is identical to the template structures. The ECD geometry and the CDRs of both protomers result from the overrated distance between the CRDs of the initial CaSR-ECD structure (PDB: 5K5T). In the initial ECD structure, shown in Figure 2.41-B, the 7TM domains are missing, and the distance of 84,5 Angström seems to be the result of the crystallisation process. The asymmetric unit contains two staged CaSR homodimers of the ECD in a compact form. In this formation, the CRD interact artificially with the ECD of the geminal homodimer. The result of this staging is the angle of 93° between the CRDs and a distance of 84,5 Å between the protomers D569.

In comparison to the crystal structure of mGlu5 in the inactive form (PDB: 6N52), the CRDs and TMs appear to arrange in a parallel form because of the presents of the cellular transmembrane. The model quality check indicates that the hybrid model has the best quality score, including the most residues (Figure 2.42-A). The Ramachandran plot shows that the model has some residues in the disallowed region, including ASN83, SER366 and ARG396. Those residues are all positioned in the extracellular domain, and they are not relevant for the docking or pharmacophore modelling.

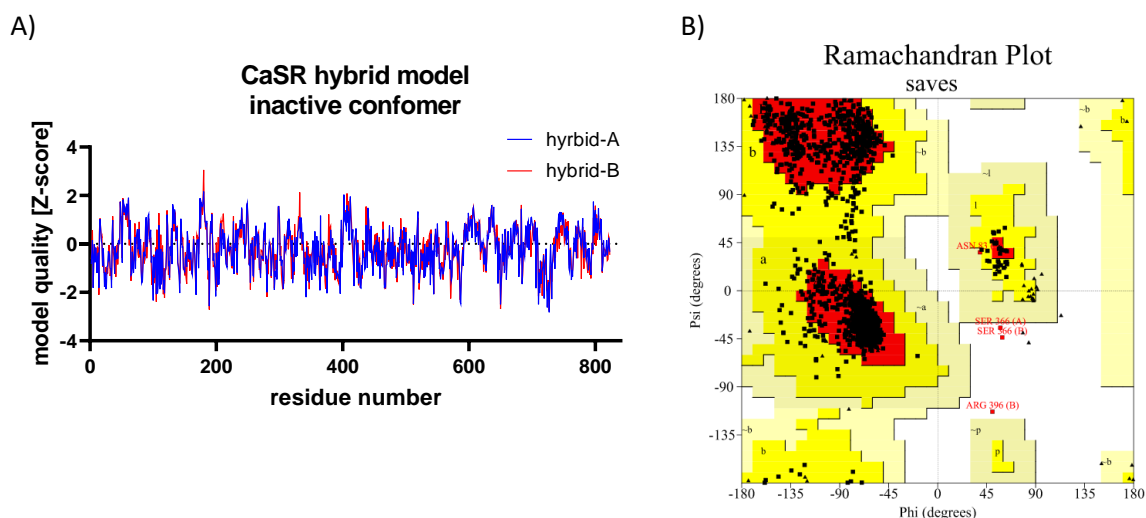


Figure 2.42 Hybrid homology model of the ICD truncated CaSR in the inactive conformation; A) Z-score model quality of the hybrid model protomers per residue; B) Ramachandran Plot of the hybrid model; residues in the disallowed region: ASN83, SER366; ARG396

In general, the final model showed a faulty geometry of the individual domains, which is not reasonable. The angles of the random coil linking the ECD and the 7TM (residues 599-605) were manually altered to achieve a similar geometry to the crystal structure of mGlu₅ (PDB: 6N52) (Figure 2.41-C). As a result of the new parallel 7TM domains, it was necessary to optimise the conformation of extracellular loops 1 and 3 and intracellular loop 2. The residues in those loops were energetically minimised with YASARA's 'OptimizeLoop' command (54). The model quality check of the adapted hybrid model revealed that the manual alterations did not impair the model. The Ramachandran plot's data also confirms that fewer residues are in the disallowed region Figure 2.43.

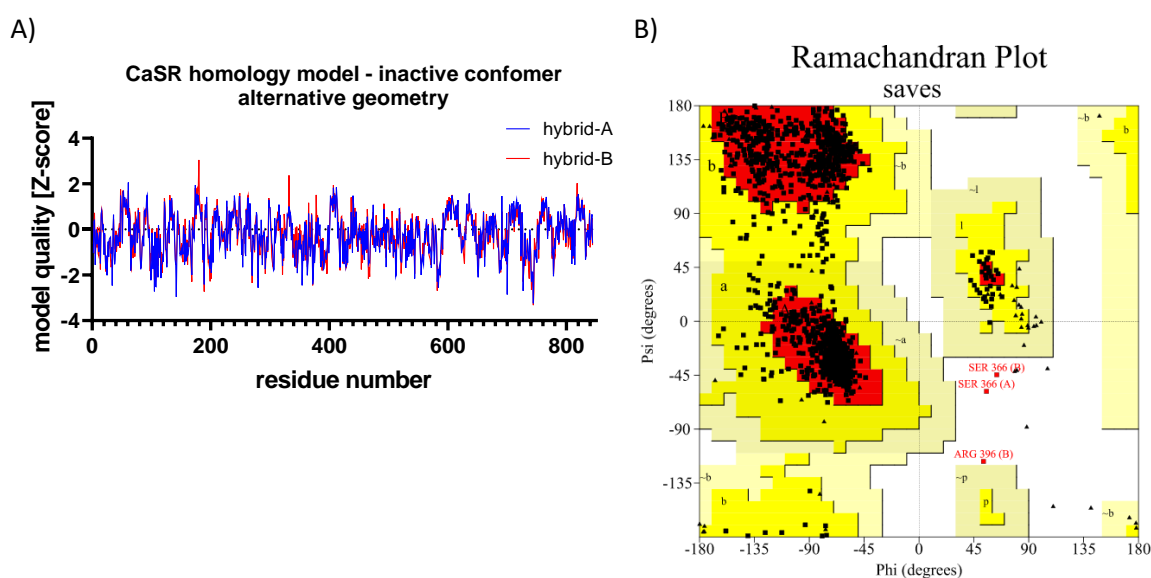


Figure 2.43 Hybrid homology model of the ICD truncated CaSR with an alternative geometry; A) Z-score model quality of the adapted hybrid model protomers per residue; B) Ramachandran Plot of the hybrid model; residues in disallowed regions: SER366 and ARG396

Homology modelling of the heterotrimeric Gq protein

The model of the heterotrimeric Gq-protein is based on the crystal structure PDB 3AH8. The homology modelling procedure was chosen to gain full length proteins. The structure was cleaned from all non-protein entries. The alpha, beta, and gamma subunit sequence were aligned with the protein sequence from UniProt (GNAQ_human: P50148; GBB1_human: P62873; GBG2_human: P59768) (Figure 2.44). This alignment was used to create the model with the cleaned crystal structure. Homology modelling was performed with slow speed. The provided alignment file made the PSI-BLAST iterations and E-value threshold irrelevant. The crystal structure was used as the only template in a heterotrimeric oligomerisation state. Alternative alignments were avoided as well as a sampling of loop conformations. The termini were extended by up to ten missing residues.

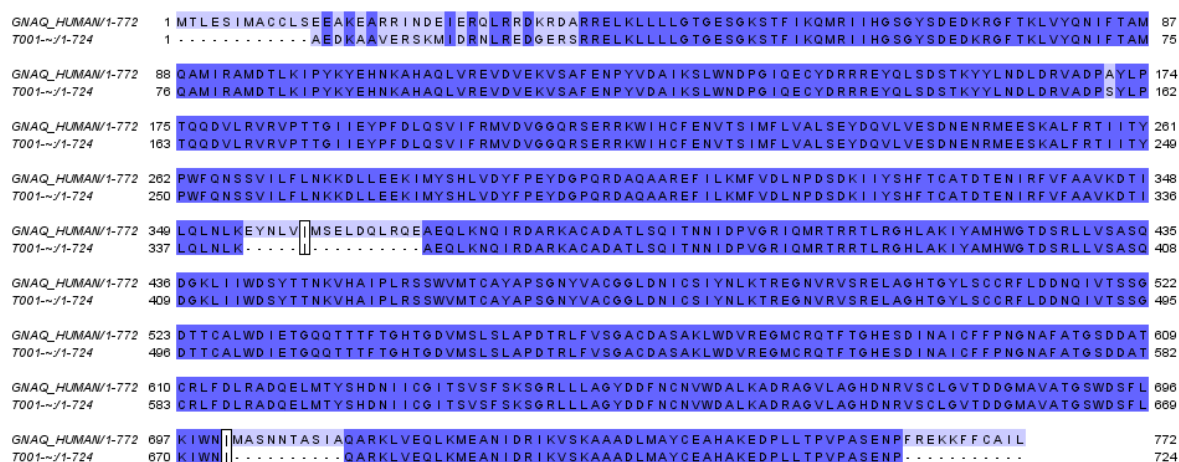


Figure 2.44 Sequence alignment of the heterotrimeric Gq-protein between the target sequence and the template (PDB: 3AH8)

The sequence of the final model was checked again for missing residues. The N-terminal residues of G alpha q (residue 1-3) and the C-terminal residues of the gamma subunit (residue 70-71) were supplemented with the 'BuildLoop' command of YASARA (54). The C-termini of the G alpha q subunit was optimised and resulted in a helical structure. It is the only model and has a quality score of -0.776, and the corresponding Z-score per residue is depicted in Figure 2.45. The final model was used to create a CaSR Gq protein complex.

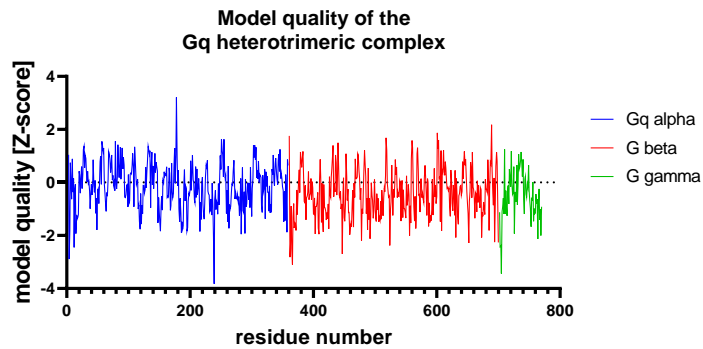


Figure 2.45 Model quality of the Gq heterotrimeric complex with an overall Z-score of -0.776

Alignment of the CaSR Gq-protein complex

The CaSR Gq-protein complex's modelling is based on the assumption that different G-proteins have a similar orientation at the GPCR G-protein interface (94). The structure of a rhodopsin Gi-protein complex (PDB: 6CMO) was used to create the CaSR Gq-protein complex by superimposition with the method SHEBA (95). The CaSR ECD-7TM model was prepared for the complex by deleting the intracellular loop 2 first because the loop would have bumped into the Gq-protein. The modified CaSR ECD-7TM model was structurally aligned with rhodopsin to superimpose the seven transmembrane domains, and the Gq-protein model was aligned with the Gi-protein in the complex. Finally, the rhodopsin Gi-protein complex was removed, and the CaSR Gq-protein complex remained (Figure 2.46).

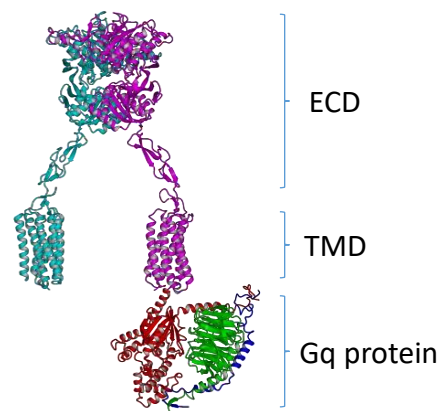


Figure 2.46 CaSR Gq-protein complex after structural alignment to a rhodopsin Gi complex; CaSR homodimer (cyan, magenta); G alpha q: red; G beta: green; G gamma: blue; ECD: extracellular domain; TMD: seven-transmembrane domain

The CaSR Gq-protein complex was inserted into a cell membrane to limit the conformational space of the intracellular and extracellular loops of the transmembrane domain. The introduction of the membrane was executed with the 'md_runmembrane' macro from YASARA (96). The macro was stopped after the membrane was introduced and before an energy minimisation was started. The next step included the removal of intracellular loop 2 to construct the CaSR G-protein complex. It was reconstructed with the 'BuildLoop' command. In the loop modelling process, 16 different conformations were generated for residue 704 to 720. This loop's secondary structure was predefined to form an alpha-helix for residue 704 – 708, and the remaining residues formed random coil. Different loop conformations were inspected, and one with the least impact on the starting structure was finally selected. At this point, the only part missing was the intracellular domain to gain the full-length model. Homology modelling of this domain was not possible because the most homologous structures had less than 18% sequence similarity. The conclusion was to use the online server I-TASSER for protein structure prediction (page 27). ECD-truncated CaSR protein sequence was used from residue 608-1078 covering the transmembrane and the intracellular domain (Figure 2.47).

```
>ECD-truncated CaSR
WTEPFGIALTLFAVLGIFLTAFLVGLVFIKFRNTPIVKATNRELSYLLLFSLCCFSSSLF
FIGEPQDWTCLRQPAFGISFVLCISCIILVKTNRVLLVFEAKIPTSFHRKWWGLNLQFLL
VFLCTFMQIVICVIWLYTAPPSSYRNQELEDEIIFITCHEGSLMALGFLIGYTCLLAAIC
FFFAFKSRKLPENFNEAKFITFSMLIFFIVWISFIPAYASTYGKFVSAVEVIAILAASFG
LLACIFFNKIYIILFKPSRNTIEEVRCSTAAHAFKVAARATLRRSNVSRKRSSSLGGSTG
STPSSSISKSNSEDPFPQPERQKQQQPLALTQQEQQQQPLTLPQQQRSQQQPRCKQKVI
FGSGTVTFSLSFDEPQKNAMAHNRNTHQNSLEAQKSSDTLTRHQPLLPLQCGETDLDLTV
QETGLQGPVGGDQRPEVEDPEELSPALVVSSSQSFVISGGGSTVTENVVNS
```

Figure 2.47 Primary sequence of ECD truncated CaSR

This part of CaSR's sequence has been chosen to test the seven-transmembrane's correct prediction as a reference. The sequence has been aligned to six different protein structures (Table 2.11). The global sequence identity is quite low and ranges from 0.15 to 0.18, making it rather difficult to predict the native structure (Figure 2.47; residue 863-1078).

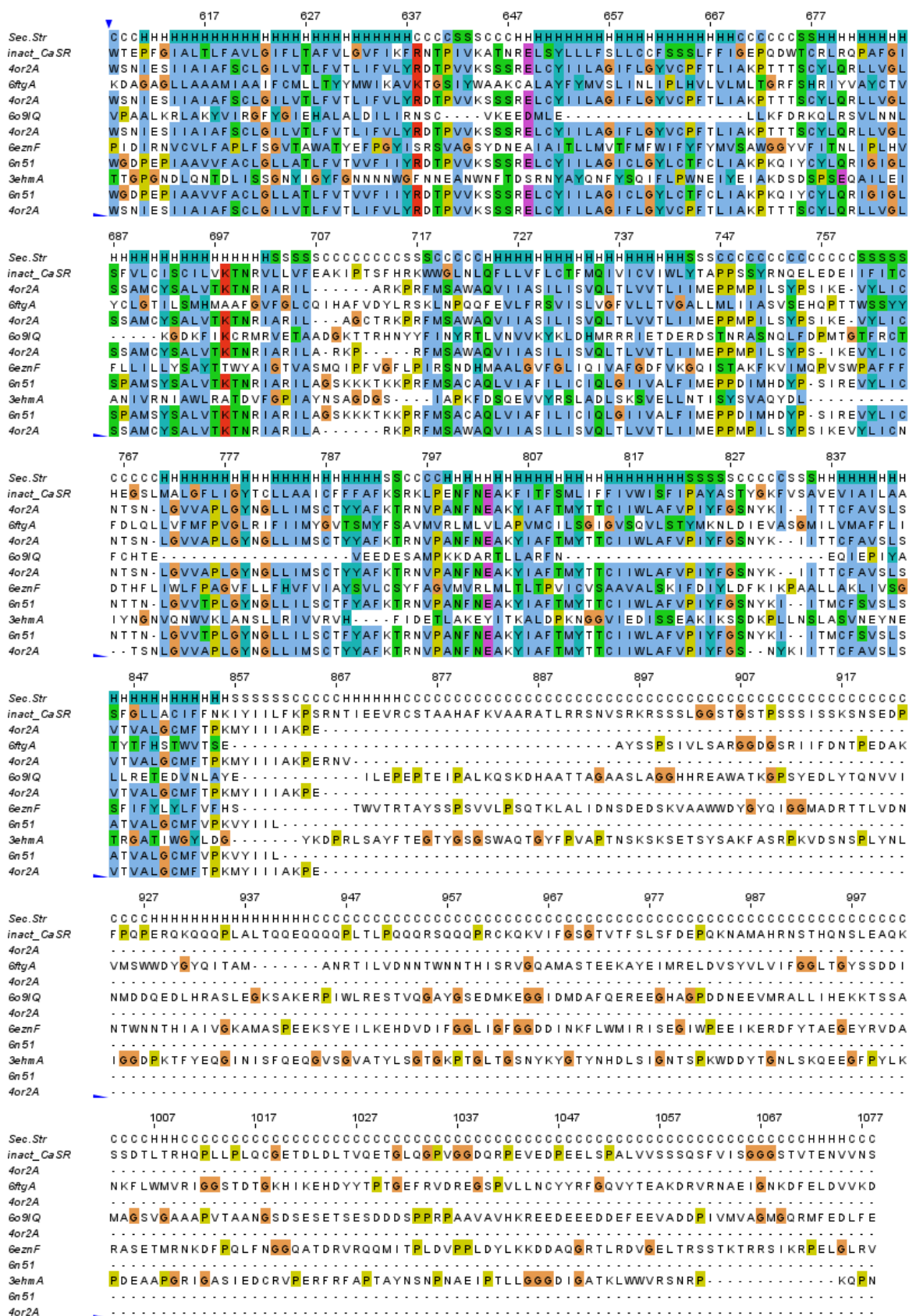


Figure 2.48 sequence alignment results of I-TASSER including the top-scoring threading templates with the primary sequence of ECD-truncated CaSR; sed. str.: secondary structure prediction of CaSR sequence with regions denoted as random coil (C), helix (H) or beta-sheet (S); colouring scheme: ClustalX (40)

Table 2.11 Threading templates identified by LOMETS from the PDB library and used for structure prediction by I-TASSER; ident. 1: sequence identity as the ratio of the aligned regions between the threading templates to the query sequence ranging from 0 to 1; ident. 2: sequence identity as the ratio of the whole template chain to the query sequence ranging from 0 to 1; Cov: coverage representing the number of aligned residues compared to the whole query sequence; Norm. Z-score: normalized Z-score of the threading alignments, good alignments correspond to a Z-score >1, thread. pro.: threading program of the meta-server LOMETS giving the result

Rank	PDB Hit	Ident. 1	Ident. 2	Cov	Norm. Z-score	thread. pro.
1	4OR2A	0.30	0.17	0.53	1.71	MUSTER
2	6FTGA	0.08	0.17	0.90	1.86	PROSPECT2
3	4OR2A	0.30	0.17	0.54	3.42	FFAS-3D
4	6O91Q	0.07	0.15	0.82	1.77	PROSPECT2
5	4OR2A	0.30	0.17	0.53	4.57	SPARXS-X
6	6EZNF	0.08	0.17	0.98	1.72	PROSPECT2
7	6N51	0.31	0.18	0.53	5.01	HHSEARCH2
8	3EHMA	0.09	0.16	0.92	1.70	PROSPECT2
9	6N51	0.31	0.18	0.53	3.80	HHSEARCH I
10	4OR2A	0.30	0.17	0.53	2.64	Neff-PPAS

I-TASSER predicted five structures of the ECD-truncated CaSR and four had a seven-alpha helical domain. The structures' quality check shows that all the models are estimated to be at least five Angstrom apart from the native structure. The intracellular domain quality is even worse because of the low sequence similarity with the template structures (Figure 2.49).

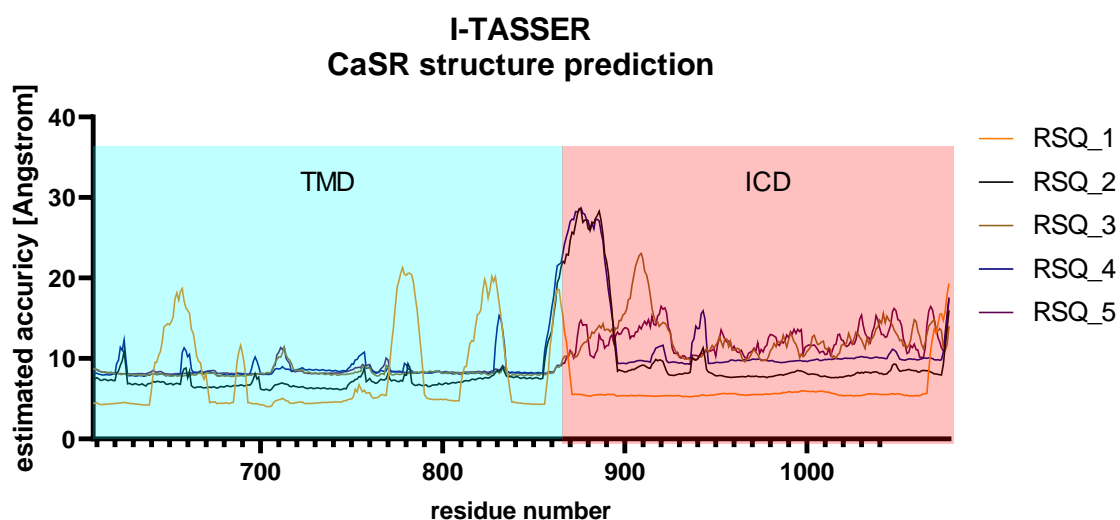


Figure 2.49 Residue-specific quality (RSQ) of the ECD-truncated CaSR estimating the deviation between the predicted models and the native structure resulting from I-TASSER; TMD: transmembrane domain; ICD: intracellular domain

Model 2 was chosen for the full CaSR model because of the globular form, the highest c-score of -1,57 and the accessible Thr888 phosphorylation site (97). Model 2 from I-TASSER was aligned to the seven-transmembrane of the ICD-truncated CaSR-Gq protein complex. The residues 608-862 were removed from Model 2, and the remaining structure was linked to the CaSR-complex. The linking residues between the seven-transmembrane and the intracellular domain were energy minimized to remove bumps.

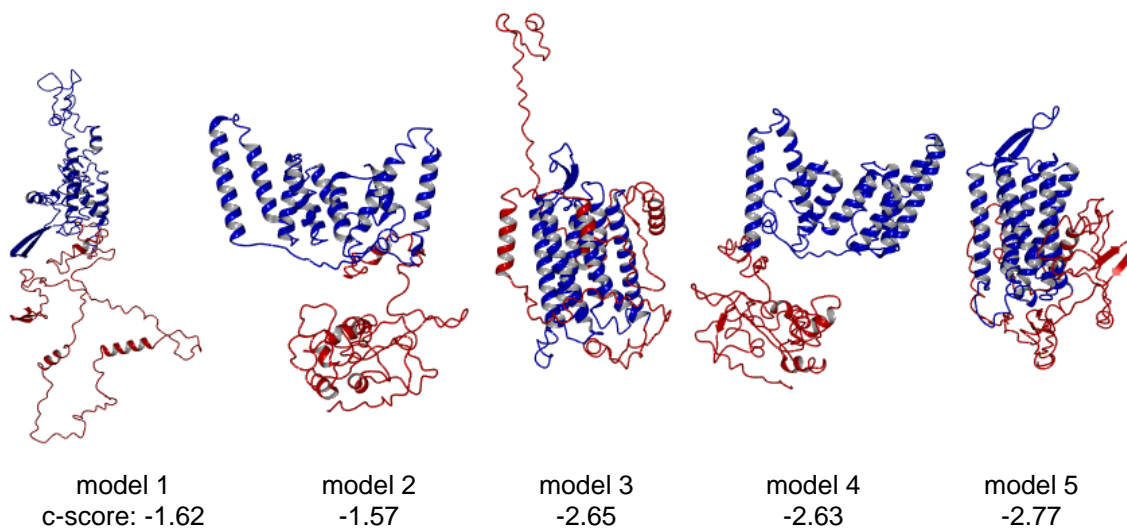


Figure 2.50 2.51 I-TASSER results of predicted ECD-truncated CaSR models, c-score: confidence score representing the model quality in the range between -5 to 2 where better models have higher values; blue: transmembrane domain; red: intracellular domain

The last step of the homology modelling process involved the introduction of N-acetylglucosamine molecules at the glycosylation sites. The molecules were extracted from CaSR's crystal structure (PDB: 5K5T) and linked to the corresponding asparagine sidechain at position 261, 287, 446, 468, 488, 547 and 594. The final model is depicted in Figure 2.52.

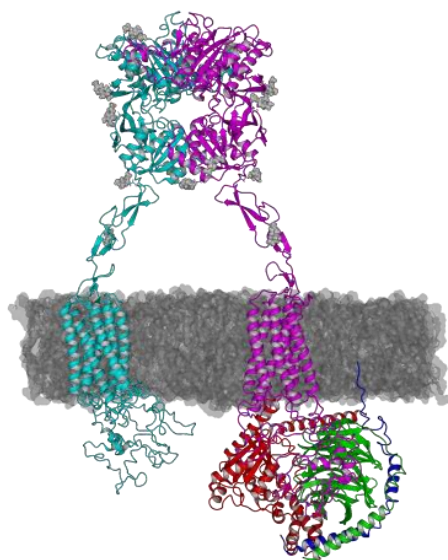


Figure 2.52 The final model of the full-length extracellular calcium-sensing receptor in complex with the Gq-protein; cyan and magenta: CaSR homodimer, red: Gq alpha subunit, green: G beta subunit; blue: G gamma subunit

The model was used for prospective validation studies with the latest ChEMBL entries of negative allosteric modulators for CaSR (n = 368). The results should show how well the new models predict the binding affinities with Autodock VINA. Before the docking experiments, three different molecular dynamics simulations were performed with the CaSR-Gq complex to create multiple receptor conformations (Figure 2.53). The used parameters were the same as for the first molecular dynamics simulations with the homology model of the transmembrane domain (page 58). RMSD fluctuations of the transmembrane domain show that the equilibrium is reached within five nanoseconds.

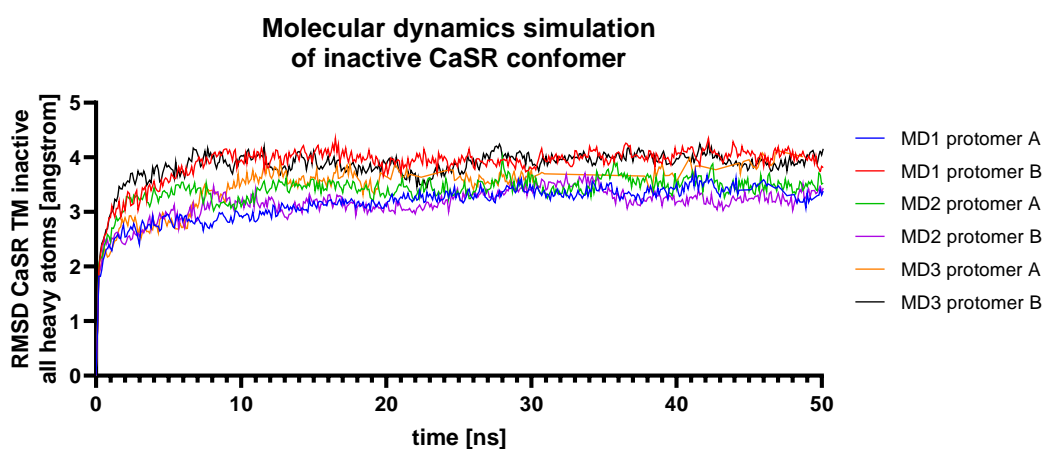


Figure 2.53 RMSD fluctuations of three 50ns long molecular dynamics simulations of the CaSR-Gq protein complex transmembrane domain

The docking experiments with 368 negative allosteric modulators show that the new model does not perform better than the old one. The predicted binding affinities hardly correlate with the ChEMBL values. The highest Pearson correlation of 0.31 was found with the conformer at the starting point of the molecular dynamics simulation (0.0 ns) (Figure 2.54). Consequently, the model was not used to predict the binding affinity between a ligand and the transmembrane of CaSR with AutoDock VINA. The model was used for structure-based pharmacophore modelling with LigBuilder 2.0 and to select potential candidates by evaluating the interaction profile.

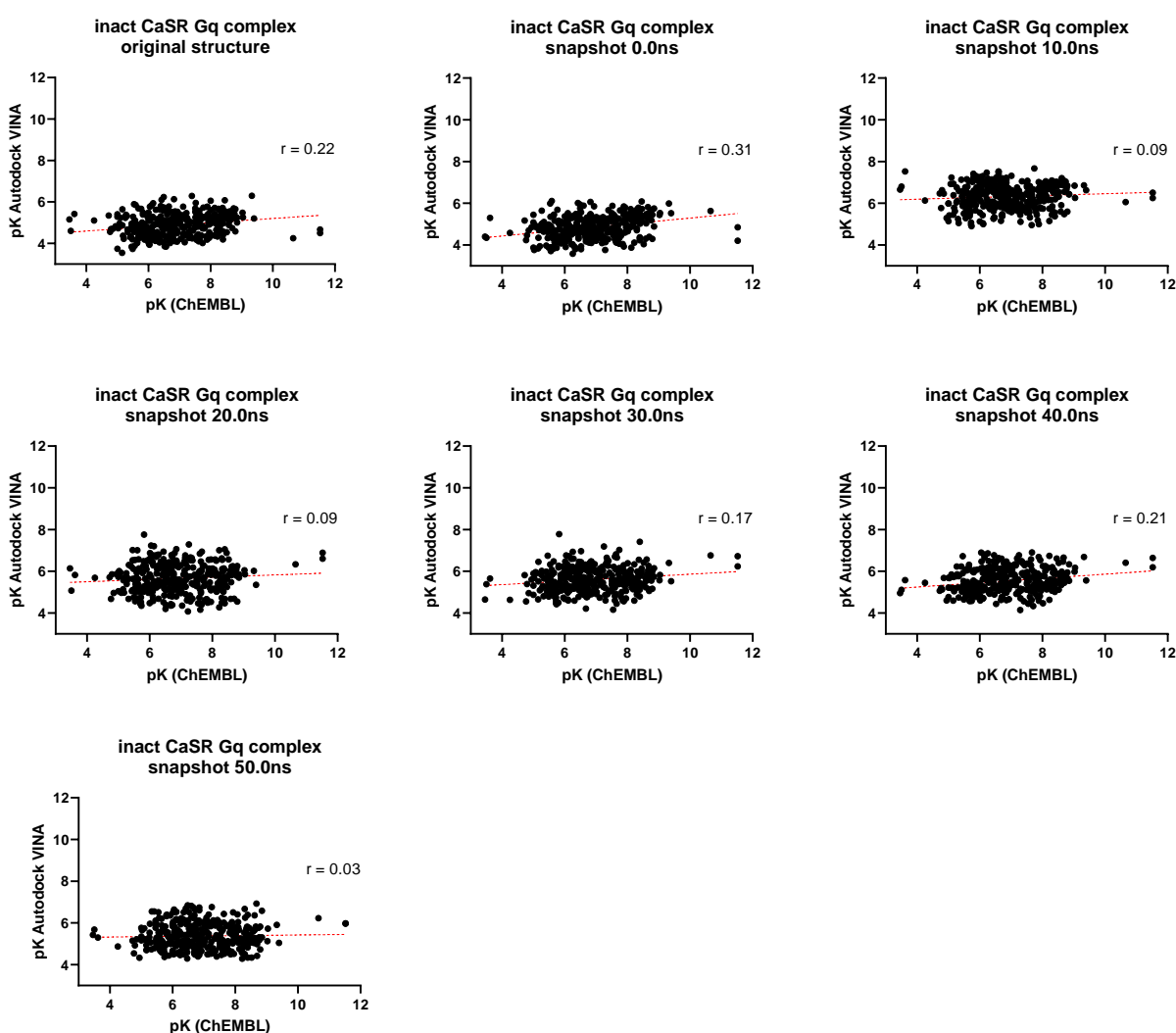


Figure 2.54 CaSR model complex validation with 368 negative allosteric modulators of different; r : Pearson coefficient (Equation 5)

The docking experiments

shows that the correlation between the predicted pK values from AutoDock VINA and the ChEMBL entries perform at the same level ($r = 0.31$) or worse than the first model. However, the extracellular loops structure is based on more information because of the new metabotropic glutamate receptor 5 crystal structure (PDB: 6N52). The improved structure around the binding pocket should increase the probability to select a potential active allosteric modulator.

Homology modelling of CaSR in the active conformation

The homology modelling process of CaSR in the active state started by searching crystal structures at the Protein Data Bank (43). The result listed three crystal structures of CaSR's extracellular domain (PDB: 5FBH, 5FBK, 5K5S) and one crystal structure of metabotropic glutamate receptor 5 (PDB: 6N51) including the extracellular domain, cysteine-rich domain and transmembrane domain as a homodimer (9,28,91). The template structures were downloaded from the Protein Data Bank (43). All non-protein entries were removed as well as the nanobodies Nb43 in PDB 6N51. Homodimers were split into the individual protomers ([Table 2.12](#)).

Table 2.12 Class C GPCR templates; CASR: extracellular calcium sensing receptor; GRM5: metabotropic glutamate receptor 5; oligo.state: oligomerisation state; aa.: number of amino acids; ref.: reference

GPCR	PDB ID	oligo. state	resolution	aa.	released	ref.
CASR	5FBH-B	monomer	2.70 A	470	2016-06-22	(28)
CASR	5FBH-A	monomer	2.70 A	472	2016-06-22	(28)
CASR	5FBK-A	monomer	2.10 A	487	2016-06-22	(28)
CASR	5FBK-B	monomer	2.10 A	467	2016-06-22	(28)
CASR	5K5S-A	monomer	2.60 A	531	2016-08-03	(9)
CASR	5K5S-B	monomer	2.60 A	537	2016-08-03	(9)
GRM5	6N51-B	monomer	4.00 A	793	2015-08-12	(91)
GRM5	6N51-A	monomer	4.00 A	791	2019-01-23	(91)

The template sequences were saved to create the alignment file (Figure 2.55). The homology model was created with the prepared template structures and the alignment file utilizing the automated modelling procedure of YASARA (24). The modelling process was performed at 'slow' speed to create models with the highest Z-score. The PSI-BLAST and E-value parameter were irrelevant because seven template structures were provided manually. Each template was used only once to create a monomeric model. Loop sampling was not part of the modelling process, and also the extension of short termini was avoided.

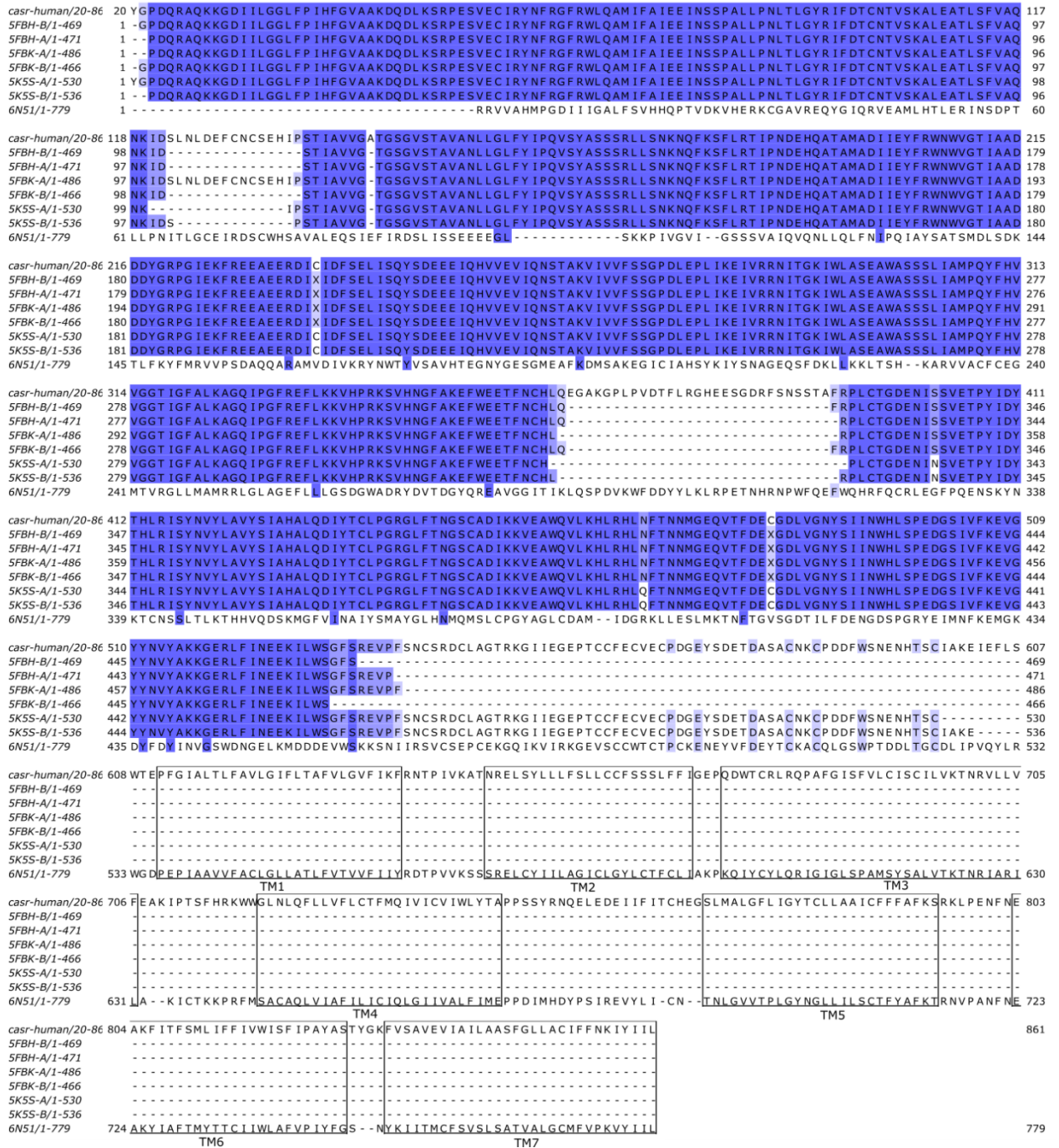


Figure 2.55 Structural alignment for homology modelling of CaSR in the active conformation, TM: transmembrane alpha-helix, blue: sequence identity

The result of the modelling process was a single protomer of ICD-truncated CaSR in the active state with a Z-score of -0.968. The model contains the extracellular domain and the transmembrane domain in a straight arrangement (Figure 2.56-A). The model's quality was estimated to be good, and the corresponding Z-score

mainly ranges between -2 and 2 (Figure 2.56-B). The structural alignment of two CaSR protomers to any of the template structures resulted in a faulty homodimer model.

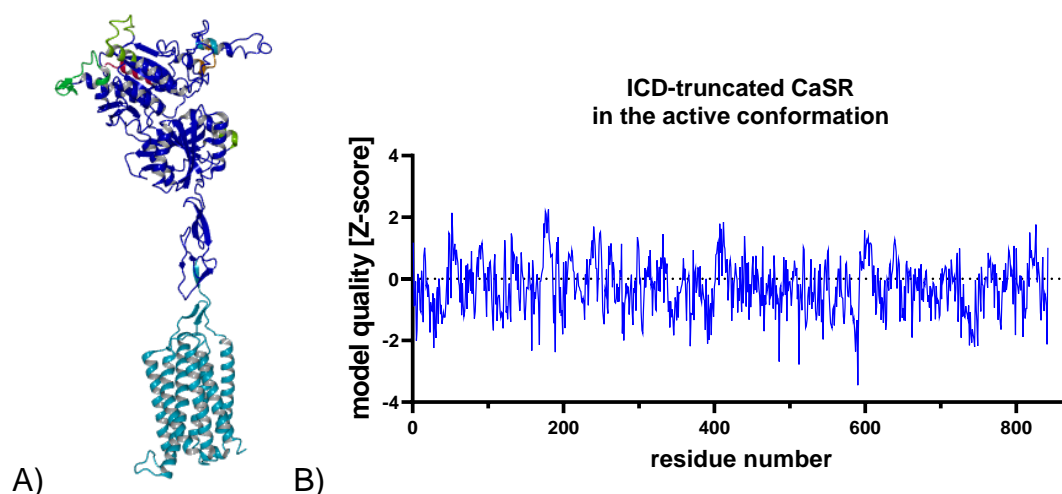


Figure 2.56 Homology model in the active conformation A) final ICD-truncated CaSR model; A) B) quality Z-score per residue of the ICD-truncated CaSR model

The final strategy involved a new CaSR homology model of a transmembrane homodimer. The created ICD-truncated CaSR protomers should be aligned to the transmembrane dimer to create the homodimer. The starting point of the new homology model was the prepared template of GMR5. The extracellular domain (residue 1 to 568) was removed, and the sequence was used to create the alignment file (Figure 2.57).

<i>casr-human/20-86</i>	20	E I E F L S W T E P F G I A L T L F A V L G I F L T A F V L G V F I K F R N T P I V K A T N R E L S Y L L L F S L L C C F S S S L F F I G E P Q D W	93
<i>6N51/1-253</i>	1	P V Q Y L R W G D P E P I A A V V F A C L G L L A T L F V T V V F I I Y R D T P V V K S S R E L C Y I I L A G I C L G Y L C T F C L I A K P K O I	74
		TM1	TM2
<i>casr-human/20-86</i>	94	T C R L R Q P A F G I S F V L C I S C I L V K T N R V L L V F E A K I P T S F H R K W W G L N L Q F L L V F L C T F M Q I V I C V I W L Y T A P P S	167
<i>6N51/1-253</i>	75	Y C Y L Q R I G I G L S P A M S Y S A L V T K T N R I A R I L A - - K I C T K K P R F M S A C A Q L V I A F I L I C I O L G I I V A L F I M E P P D	146
		TM3	TM4
<i>casr-human/20-86</i>	168	S Y R N Q E L E D E I I F I T C H E G S L M A L G F L I G Y T C L L A A I C F F A F K S R K L P E N F N E A K F I T F S M L I F F I V W I S F I P	241
<i>6N51/1-253</i>	147	I M H D Y P S I R E V Y L I - C N - - T N L G V V T P L G Y N G L L I L S C T F Y A F K T I R N V P A N F N E A K Y I A F T M Y T T C I I W L A F V P	217
		TM5	TM6
<i>casr-human/20-86</i>	242	A Y A S T Y G K F V S A V E V I A I L A A S F G L L A C I F F N K I Y I I L	279
<i>6N51/1-253</i>	218	I Y F G S - - N Y K I I T M C F S V S L S A T V A L G C M F V P K V Y I I L	253
		TM7	

Figure 2.57 Structural alignment of CaSR transmembrane domain to mGlu₅

The resulting homology model of the transmembrane dimer was used to align the CaSR protomers to a homodimer. The disulphide bridges between the protomers at cysteine 129 and 131 were introduced again, and the corresponding loops were minimized (residue 121-135). The created homodimer of CaSR's extracellular

domain and the transmembrane domain was aligned with the full model of CaSR in the inactive state. The intracellular domain was transferred to the model in the active state, and the missing intracellular loop 2 was introduced. Finally, the N-acetylglucosamine molecules from the crystal structures were linked to create the final homology model of the full-length CaSR in the active state. The overall model quality is -2.646 of CaSR's homodimer in the active state.

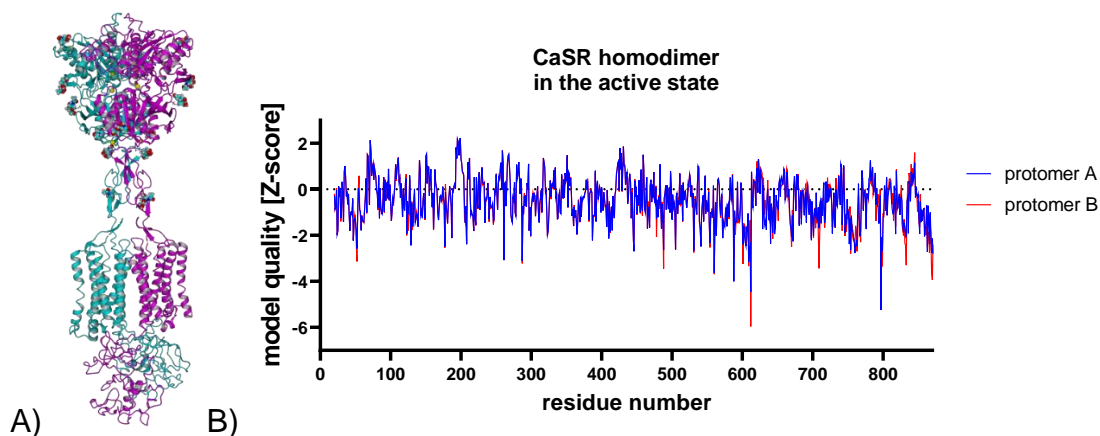


Figure 2.58 Homology model of the full-length CaSR in the active state; A) depiction of the homodimer model; B) quality Z-score per residue of CaSR in the active state

The model quality per residue revealed some major deviations to the estimated native structure. The more significant uncertainties were located in the loops of the transmembrane. The model was used in three 50 ns long molecular dynamics simulations to remove potential bumps close to the binding pocket and create different conformers every 10 ns. The equilibrium is reached after 5 ns (Figure 2.59).

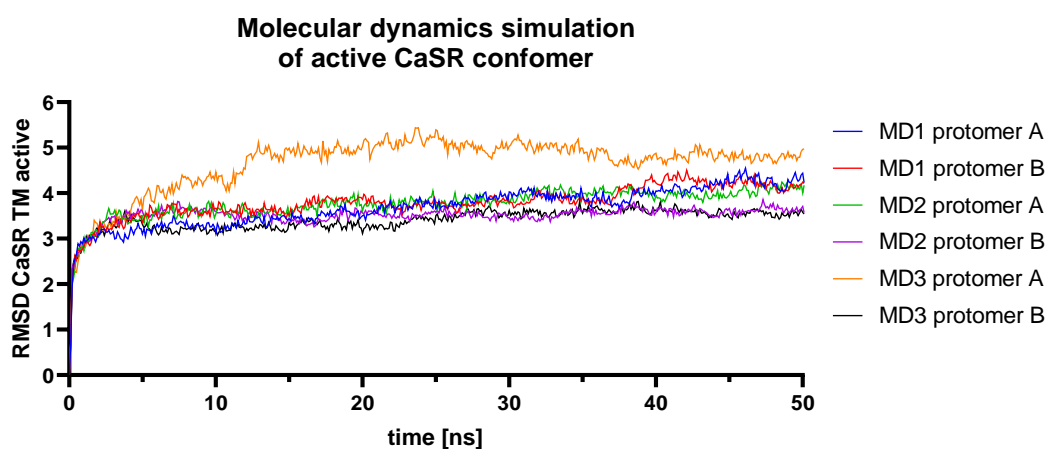


Figure 2.59 RMSD fluctuations of three 50ns long molecular dynamics simulations of the CaSR transmembrane (residue 601-863) in the active conformation

Seven conformers of CaSR in the active state were validated with 236 positive allosteric modulators, and the outcome is depicted in Figure 2.60 utilizing docking experiments with AutoDock VINA.

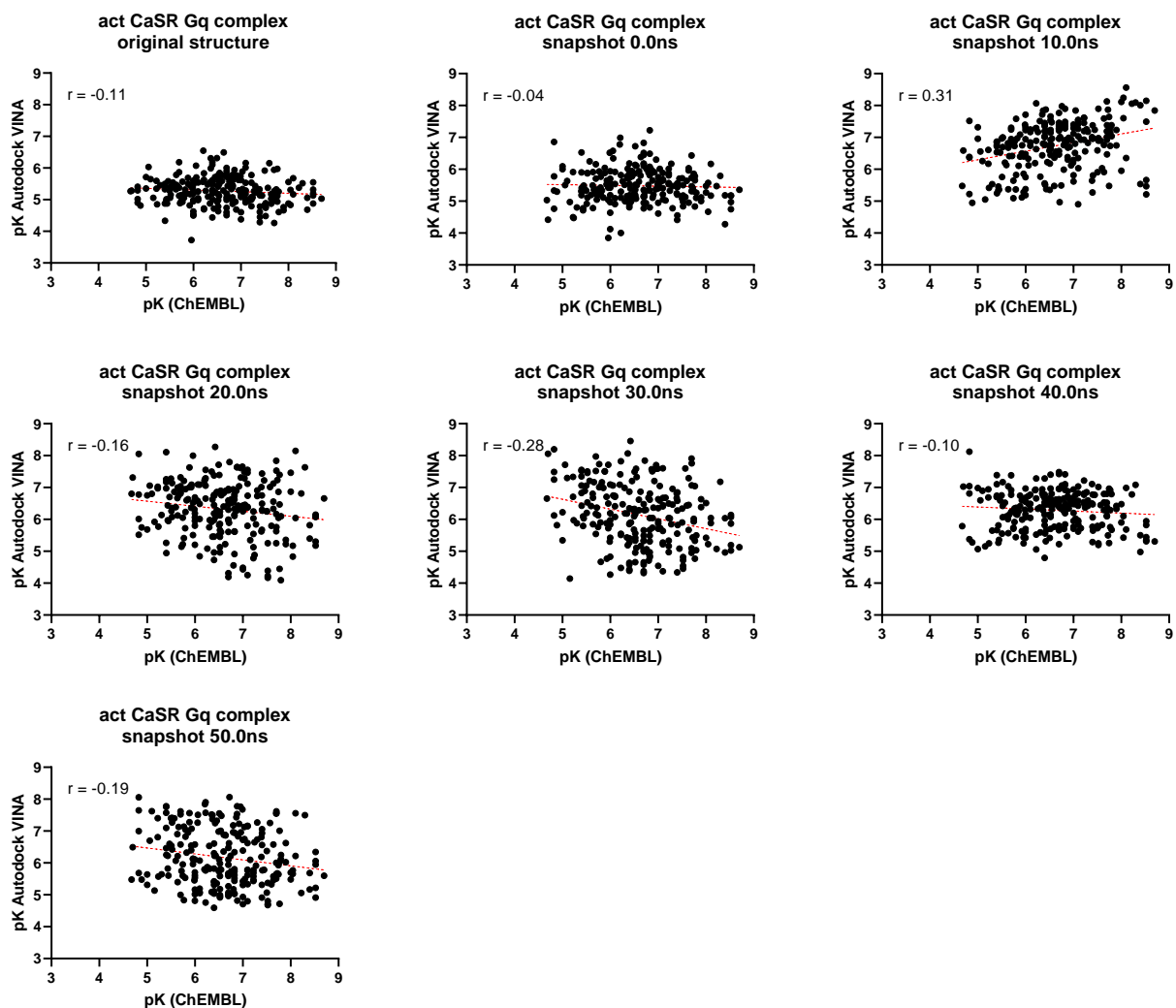


Figure 2.60 Model validation of seven CaSR conformers in the active state with the ChEMBL validation set of 236 positive allosteric modulators

The validation results show that the correlation for the predicted binding affinities ($r = 0.31$) is not better than the correlation for the first model ($r = 0.33$). It is also odd that most of the conformers had a negative correlation. Consequently, the generated models can not be used for docking experiments to perform virtual screening. However, the best model (Figure 2.60; 10ns) was created with a new crystal structure and has been used for structure-based pharmacophore modelling and final compound selection by analysing the interaction profile.

Pharmacophore modelling

The compound library's last dataset was created by searching the ZINC database with the online interface ZINCPharmer. The webserver requires pharmacophore models as input and supports file formats from LigandScout (.pml), MOE (.ph4), Pharmer (.json), LigBuilder (.txt) and PharmaGist (.mol2) (78,83,85–88). Pharmacophore models are the three-dimensional arrangement of functional features representing non-covalent interactions between a macromolecule and a ligand. The method is complementary to molecular docking and allows the identification of new scaffolds. The chemical features include hydrophobic, aromatic, hydrogen donor, hydrogen acceptor, positive ion and negative ion interactions. Besides, exclusion spheres are used to limit the conformational space of ligands and filter compounds that would cause atomic clashes (85,98,99). In the case of CaSR, the created homology models without a ligand were used to extract a structure-based pharmacophore model and search the ZINC database for potential compounds. The pharmacophore models had to be created of the binding pocket in the upper part of the seven-transmembranes. The method of choice was the free tool LigBuilder 2.0. The de-novo design program LigBuilder 2.0 identifies the binding pocket of a protein and creates a pharmacophore model with the CAVITY 1.0 algorithm (78). The homology models of CaSR's transmembrane domain were used as input for LigBuilder 2.0. The homology models were used in four different states to create four different pharmacophore models: active undocked, active pre-docked, inactive undocked and inactive pre-docked. The standard input file of LigBuilder had to be modified, and the location of the homology models was defined to create the pharmacophore models. The detection mode was set to 0 to search the whole seven-transmembrane domain for the cavity. Each time the CAVITY 1.0 algorithm was executed, 15 different pharmacophore models were created. The corresponding proteins were visualized in YASARA, and the models were selected displaying the cavity in the upper part

of the seven-transmembrane between helix III, V, VI and VII (

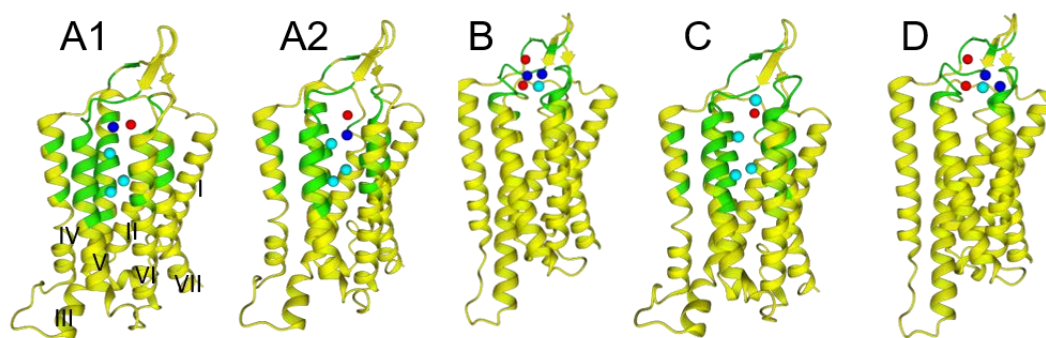


Figure 2.61). The pharmacophore models were stored as TXT files, and the content of all four is listed in appx. 9.1.

ZINCPharmer search

The pharmacophores from LigBuilder 2.0 composed of seven to eleven features. LigBuilder does not allow the direct validation of the pharmacophore model. Different combinations of features were selected and used to search for potential compounds at ZINCPharmer. The minimum requirement for the ZINCPharmer search are three features, but this would lead to millions of hits. Different combinations showed that six features and more give no results at all. The search for every pharmacophore model was performed with five features that gave the largest number of hits and contained the largest number of known allosteric modulators for CaSR (Figure 2.61).

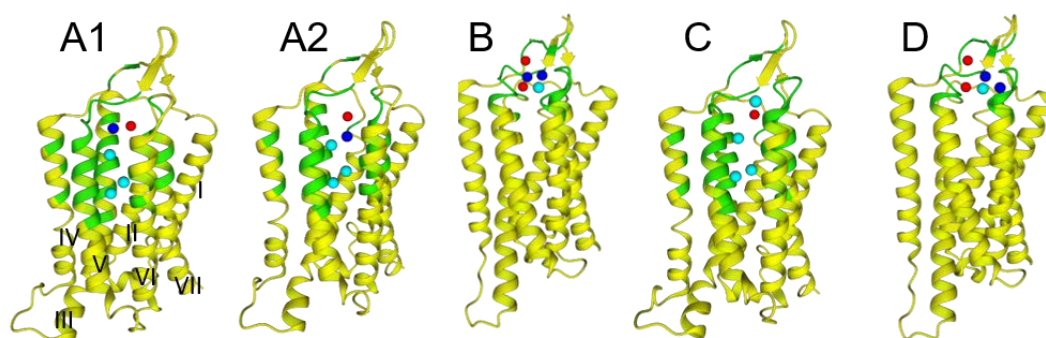


Figure 2.61 structure-based pharmacophore models of CaSR's allosteric binding site; green: residues identified by Cavity 1.0 to participate in the binding pocket. red: hydrogen acceptor, blue: hydrogen donor, cyan: hydrophobic (representation done with YASARA); A1 and A2 are the pharmacophores of the active

model, with different feature selection, B inactive model, C active pre-docked model, D inactive pre-docked model

The search results for every pharmacophore model is summarized in Table 2.13.

Table 2.13 ZINCPharmer search results

pharmacophore model	screening hits
A1	776
A2	1897
B	195
C	211
D	263

The combined compound library has 3342 entries from the ZINCPharmer search, and 1657 entries were ready for sale.

Combined compound library for virtual profiling:

The first virtual profiling attempt had limited success in identifying active allosteric modulators with a significant binding affinity for CaSR. The new strategy involved an updated version of SAFAN-ISPSM, emphasising chirality and reducing atomic overrepresentation in the Tanimoto calculation. Part of the new screening attempt was to increase the number of profiled compounds to increase the probability of discovering a potent allosteric modulator. The new compound library contains the entries of DrugBank (version 5.1.2), FooDB (version 1.0), HerDing database and TCM Database@Taiwan. Additional compounds were identified from the search of ZINC12 database with BARFs from SAFAN-ISPSM's fragment-weight-assignment algorithm version A2 and ZINC15 database from the search with BARFs from SAFAN-ISPSM's fragment-weight-assignment algorithm version B2. The last compound list was obtained by searching ZINC15 with structure-based pharmacophore models via ZINCPharmer. The new compound library contains 251632 small molecules, and an overview is shown in Figure 2.62.

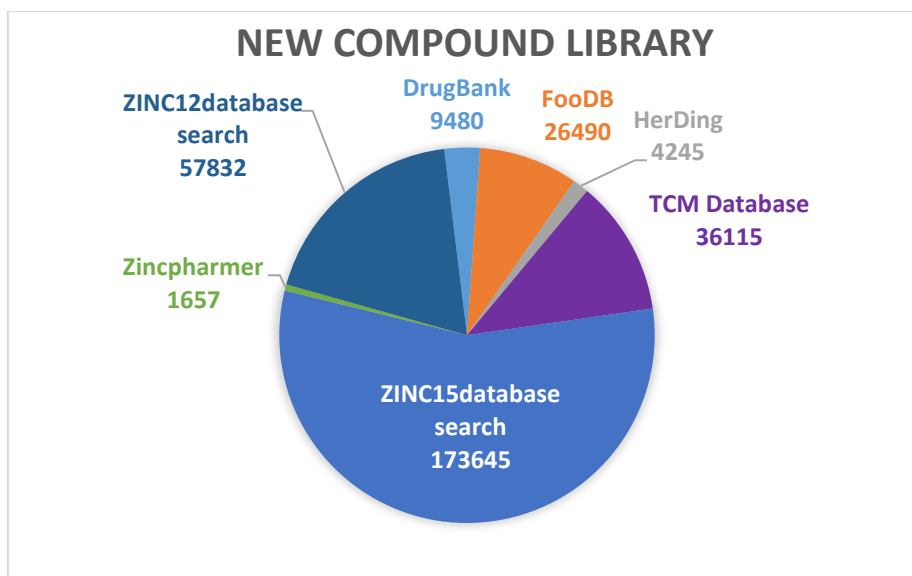


Figure 2.62 new compound library for the second profiling attempt with SAFAN-ISPSM 2019/06-b2

Virtual profiling of the new compound library with SAFAN-ISPSM

After the new compound library was available, the 251632 small molecules were profiled with SAFAN-ISPSM 2019/06-b2. The virtual-profiling was executed at the C3S supercomputer OCCAM (72) and took six days using 156 CPUs. The screening of 251632 small molecules resulted in 3924 predicted CaSR-ligand interactions with a $pK \geq 4.0$.

The experimental validation of the screening results required the selection of a practical subset of top-ranking ligands. The subset is characterised by different selection criteria and filters (Figure 2.63).

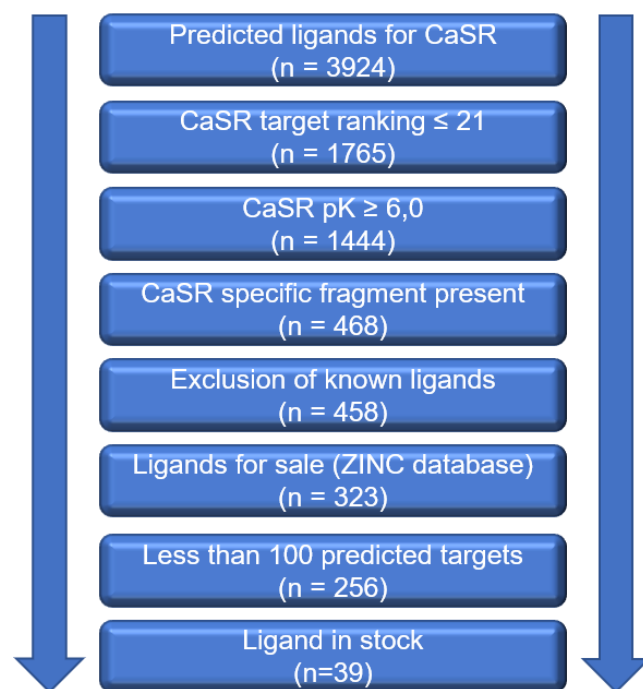


Figure 2.63 selection process of the screening results for experimental validation

The ligand-profiling results of SAFAN-ISPSM were considered for the selection which includes data about predicted CaSR targets ranking, binding affinities (pK), fragment-similarity to known compounds (fragtan), compound similarity to known compounds (alltan) and the presents of BARFs. The confidence of the prediction is expressed by the SAFAN-score ranging from 0 to 1 (Table 2.14).

Table 2.14 SAFAN-ISPSM version 2019/06-B2 profiling results of ZINC entry ZINC000000036062; position: target ranking position in the profiling list; target: predicted target that interacts with the input compound; pK: predicted binding affinity; alltan: Tanimoto similarity between the input compound and the most similar ChEMBL compound; fragtan: Tanimoto similarity between the fragment fingerprints and the ChEMBL fragment fingerprints; round: compounds without (0) or with (1) a binding affinity relevant fragment (BARF); SAFAN-score: confidence score of the prediction depending on alltan and fragtan

position	target	pK	alltan	fragtan	round	SAFAN-score
1	CASR	6.705	0.508	0.833	0	0.831
2	AK1C3	6.353	0.496	0.750	0	0.816
3	NOS1	6.091	0.482	0.639	0	0.841
4	MK01	5.951	0.502	0.756	0	0.827
5	PPIA	5.821	0.526	0.708	0	0.829
...
21	DPOLB	4.501	0.570	0.754	0	0.865

The profiling results are ranked by predicted binding affinities between the ligand and the predicted targets. If the binding affinity with CaSR was not listed within the top 21, the compound was not considered which left 1765 small molecules out of

3924. The CaSR-ligand interaction should have a binding affinity of at least 6.0, valid for 1444 ligands (100). In the next selection step, only compounds were considered that contained a CaSR specific BARF which is indicated by a round of 1. Ten already known CaSR ligands were excluded from the 468 remaining compounds. The compound synthesis was not considered for experimental validation. The 458 compounds were checked at the ZINC database, and 323 were for sale. All the compounds were excluded, which were predicted to interact with more than 100 targets ($pK \geq 4,0$; $n = 256$) and did not enter the binding pocket of CaSR in the active or inactive state after docking experiments. The compound should also be listed at the ZINC database as “in stock” for immediate purchase which left 39 potential candidates. The idea was to avoid compound series for the experimental validation. Therefore, the similarity of the remaining compounds was analysed to already known CaSR ligands. The ChEMBL database was searched for all CaSR ligands across all species, and a validation set of 908 compounds was created. The validation set and the remaining ligands were clustered with multidimensional scaling and a similarity cut-off equal to 0.9 at the ChemMine Tools website (101).

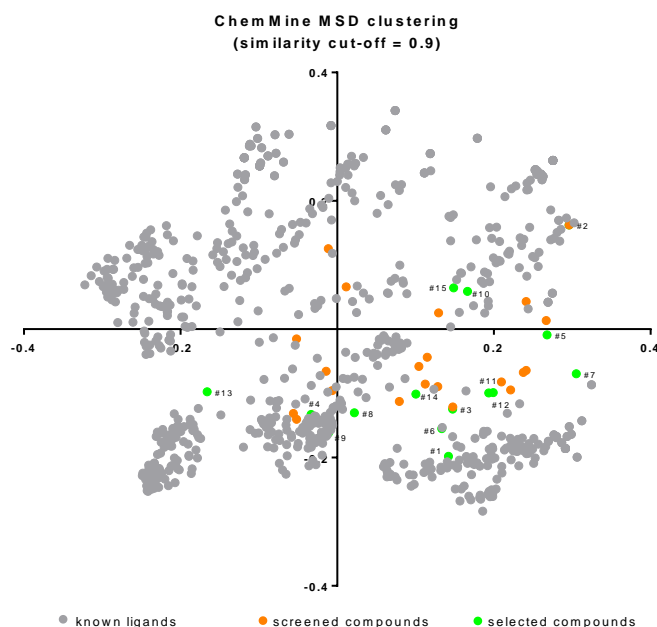


Figure 2.64 Multidimensional scaling of the remaining screening compounds ($n = 39$) with known CaSR ligands ($n = 908$) to identify dissimilar compounds

The clustering helped in the selection of 15 dissimilar compounds. In the last selection step, the 15 divers screening compounds were evaluated by their SAFAN fragment similarity and their SAFAN compound similarity to known CaSR ligands (Figure 2.65).

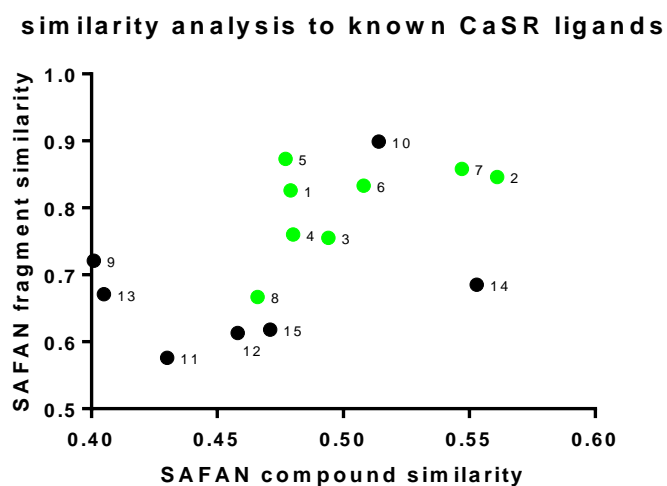
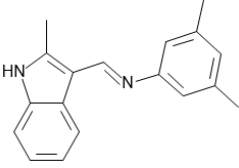
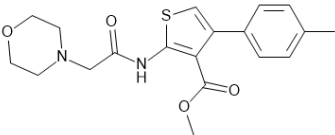
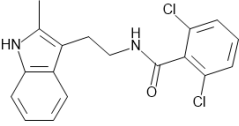
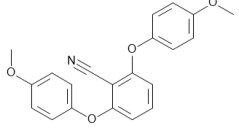
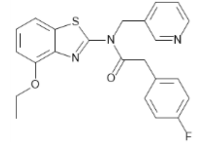
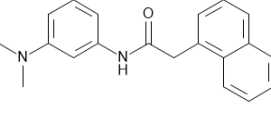
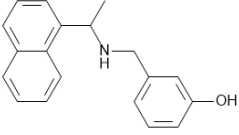
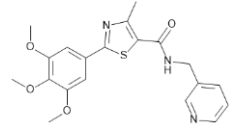


Figure 2.65 SAFAN similarity analysis of divers screening compounds to known CaSR ligands; green: selected screening compounds for experimental validation with an IP-1 Gq HTRF assay

The compounds with the highest fragment similarity (fragtan) and compound similarity (alltan) were selected for the experimental validation. However, compound 10 and 14 were excluded because they were not available for testing at a reasonable time. Finally, the selection process resulted in 8 potential allosteric modulators for CaSR, evaluated with an IP-one Gq assay (Table 2.15).

Table 2.15 second screening subset for experimental validation, 8 out of 3924 predicted ligands

<p>1. ZINC00000065809</p> <p><u>IUPAC:</u> (E)-N-(3,5-dimethylphenyl)-1-(2-methyl-1H-indol-3-yl)methanimine</p> <p><u>Smile:</u> <chem>Cc1[nH]c2ccccc2c1\C=N\c1cc(C)cc(C)c1</chem></p> <p><u>predicted pK:</u> 6.54</p>	
<p>2. ZINC000020394024</p> <p><u>IUPAC:</u> methyl 4-(4-methylphenyl)-2-[2-(morpholin-4-yl)acetamido] thiophene-3-carboxylate</p> <p><u>Smile:</u> <chem>COC(=O)c1c(NC(=O)CN2CCOCC2)sc1-c1ccc(C)cc1</chem></p> <p><u>predicted pK:</u> 6.48</p>	
<p>3. ZINC00000091620</p> <p><u>IUPAC:</u> 2,6-dichloro-N-[2-(2-methyl-1H-indol-3-yl)ethyl]benzamide</p> <p><u>Smile:</u> <chem>Cc1[nH]c2ccccc2c1CCNC(=O)c1c(Cl)cccc1Cl</chem></p> <p><u>predicted pK:</u> 6.37</p>	
<p>4. ZINC000000139716</p> <p><u>IUPAC:</u> 2,6-bis(4-methoxyphenoxy)benzonitrile</p> <p><u>Smile:</u> <chem>COc1ccc(Oc2cccc(Oc3ccc(OC)cc3)c2C#N)cc1</chem></p> <p><u>predicted pK:</u> 6.25</p>	
<p>5. ZINC000010020294</p> <p><u>IUPAC:</u> N-(4-ethoxy-1,3-benzothiazol-2-yl)-2-(4-fluorophenyl)-N-[(pyridin-3-yl)methyl]acetamide</p> <p><u>Smile:</u> <chem>CCOc1cccc2sc(nc12)N(Cc1cccnc1)C(=O)Cc1ccc(F)cc1</chem></p> <p><u>predicted pK:</u> 7.07</p>	
<p>6. ZINC000000036062</p> <p><u>IUPAC:</u> N-[3-(dimethylamino)phenyl]-2-(naphthalen-1-yl)acetamide</p> <p><u>Smile:</u> <chem>CN(C)c1cccc(NC(=O)Cc2cccc3ccccc23)c1</chem></p> <p><u>predicted pK:</u> 6.71</p>	
<p>7. ZINC000020136919</p> <p><u>IUPAC:</u> 3-({[1-(naphthalen-1-yl)ethyl]amino}methyl)phenol</p> <p><u>Smile:</u> <chem>CC(NC1CCCC(O)C1)C1CCCC2CCCC12</chem></p> <p><u>predicted pK:</u> 6.94</p>	
<p>8. ZINC000012324922</p> <p><u>IUPAC:</u> 4-methyl-N-[(pyridin-3-yl)methyl]-2-(3,4,5-trimethoxyphenyl)-1,3-thiazole-5-carboxamide</p> <p><u>Smile:</u> <chem>COC1CC(OC)C(OC)C1C(=O)NCc1cccnc1</chem></p> <p><u>predicted pK:</u> 6.59</p>	

2.7.1 Experimental setup for the IP-1 Gq assay

The IP-1 accumulation was determined for every compound at 0.2 mM, 2.5 mM and 7.5 mM calcium concentration compared to the positive allosteric modulator NPS-R568 and the negative allosteric modulator NPS-2143.

The experiments for IP1 accumulation measurements were performed with Flp-In HEK293 HA hCaSR WT cells and empty HEK293 cells as a negative control. They were provided by Iris Mos from the University of Copenhagen. The cells were seeded in HTRF 96-low volume plates (50000/well) and grown for 18 hours at 37°C and 5% CO₂ in DMEM [-Ca²⁺, -Mg²⁺]. The compounds were prepared at the same day of the measurement with 100-fold predicted pK concentration in ligand buffer (Ca²⁺ + Mg²⁺ free HBSS, 20 mM HEPES, 20 mM LiCl, pH = 7.4) at 0,28 (V/V) % DMSO. The cells were washed at the day of the experiment with assay buffer (Ca²⁺ + Mg²⁺ free HBSS, 20 mM HEPES, pH = 7.4). 50µl ligand solution was used to preincubate the cells in triplicates for 30 min at 37°C. The cells were washed and stimulated at 0.2 mM, 2.5 mM or 7.5 mM calcium concentration for 30 minutes at 37°C. The cells were rewashed with assay buffer and lysed with 30 µl lysis and detection buffer (Cisbio) for 30 minutes at room temperature. In the end, 30 µl assay buffer was added, and 10 µl of lysed cells was transferred to a 96-well low volume reader plate (CisBio) and mixed with 10µl anti-IP1 cryptate TB conjugate and IP1-d2 conjugate in assay buffer (1:1:38) fluorophore solution. After sealing the plates were incubated at room temperature for one hour in the dark. Finally, the plates were excited at 340 nm and read with a homologous time-resolved fluorescence (HTRF) reader (infinite M1000pro) at 665 nm and 620 nm. The HTRF ratio was calculated according to Equation 7, and with IP1 standards (Cisbio), a standard curve was recorded to determine the measured IP1 concentration.

The experiments for the determination of compound curves were also executed according to the protocol of accumulation measurements. But instead the calcium concentration was fixed at EC50 = 4,75mM and the compound concentrations were measured at 30 µM, 100 x pK concentration, 3µM, 1µM, 0.3µM, 0.1µM, 0.03µM and 0.01µM in triplicates.

The IP1 accumulation assay to determine the compounds' calcium curve was executed according to the protocol of the accumulation assay. However, the compound concentration was fixed at 1 μM , and the calcium concentration was measured at 10 mM, 6 mM, 5 mM, 4.5 mM, 3.5 mM, 2.5 mM, 1 mM and 0.5mM calcium.

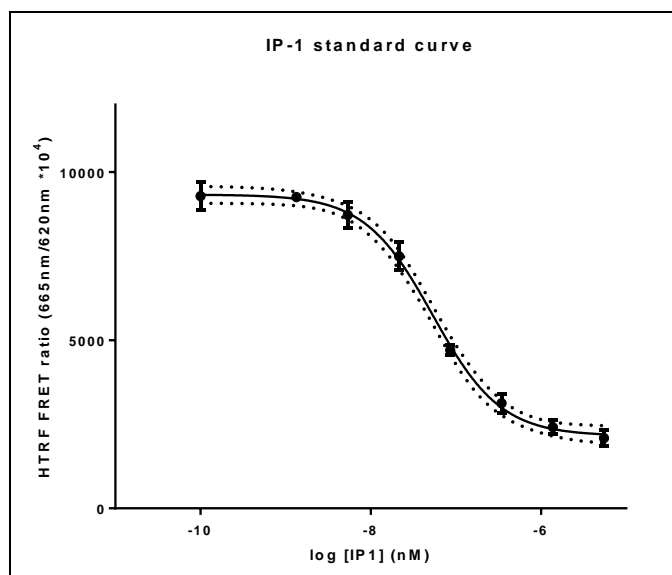


Figure 2.66 *IP1 standard curve with IP1 standards (CisBio)*

2.7.2 Results

The most significant results were obtained at 2,5 mM calcium (Figure 2.67-A). The IP1 accumulation assays revealed that compound 7, 3-({[1-(naphthalen-1-yl)ethyl]amino}methyl)phenol, shows a similar positive allosteric activity at the HEK293 HA hCaSR WT cells as the positive allosteric modulator NPS-R568. The other compounds showed no significant difference to the vehicle. The allosteric effects were also proven to be CaSR specific because no significant response was detected in the HEK293 cells (Figure 2.67-B). The positive allosteric effect of compound 7 is also demonstrated by the potency shift of the calcium curve to the right with compound 7 compared to the calcium curve only (Figure 2.67-C). The results of the IP1 accumulation assay at different compound concentrations show that compound 7 is less potent to shift the calcium curve to the right as the standard NPS-R568 (Figure 2.67 D). The binding affinity of compound 7 was determined at $pEC_{50} = 5.5$ and is 1.4 order of magnitude off from the predicted value ($pK = 6,935$). The difference in the order of magnitude for compound 5 between is 1,9 between

the predicted binding affinity ($pK = 7,0$) and the experimentally measured one ($pEC_{50} = 5,1$).

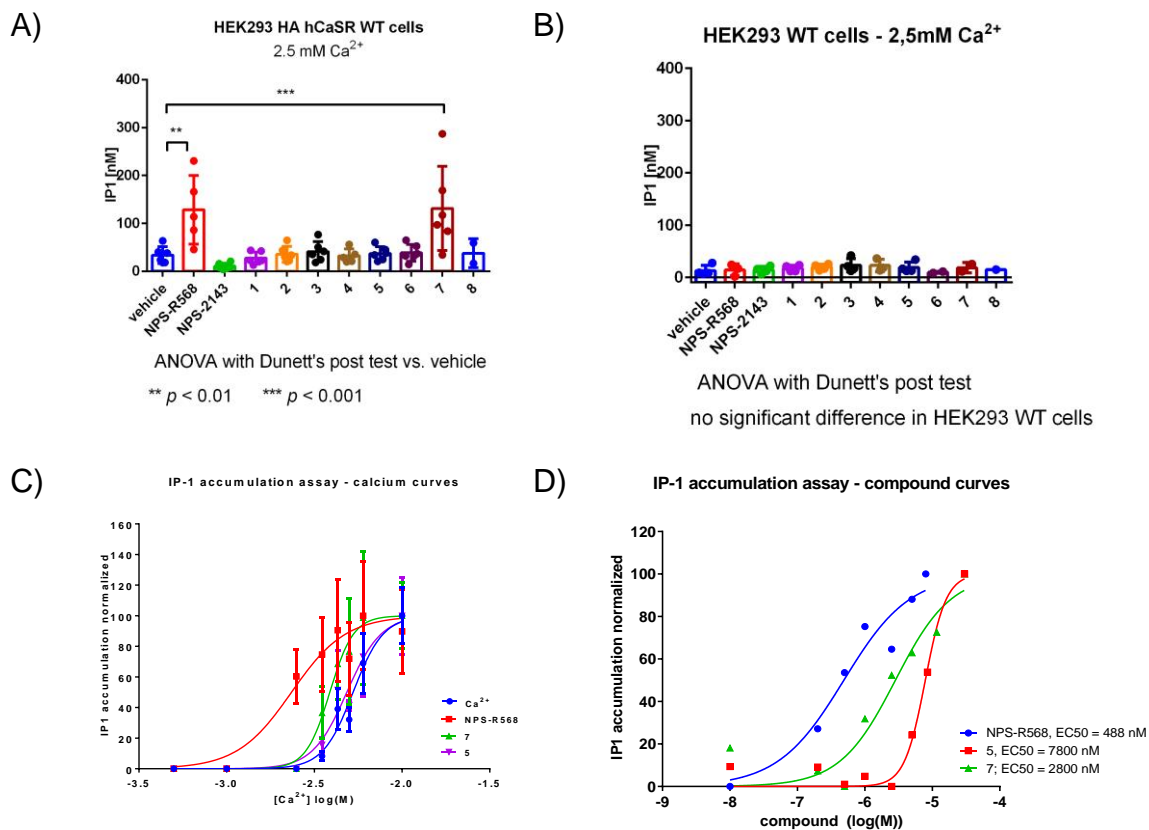


Figure 2.67 IP-1 accumulation results of 8 screening compounds against the positive allosteric modulator NPS-R568 and negative allosteric modulator NPS-2143, A) measurement in HEK293 HA hCaSR WT cells, B) experiments in HEK293 WT cells; data represents mean \pm standard deviation of 6 independent experiments in triplicates, The statistical analysis was performed with one-way ANOVA with Dunnett's post-test vs vehicle. C) IP-1 accumulation assay at different calcium concentrations shows a potency shift of the positive allosteric modulator NPS-R568 to the right as well as for compound 7, indicating positive allosteric activity at CaSR; D) IP-1 accumulation assay at different compound concentrations indicate a weaker allosteric modulator activity of compound 7 in comparison to the standard NPS-568

2.7.3 Conclusion

The search for a novel allosteric modulator with the ligand-based in-silico profiling tool SAFAN-ISPSM version 2019/06-B2 was conducted with the new compound library consisting of over 250000 small molecules. The screening results revealed about 4000 potential ligands for the human extracellular calcium-sensing receptor. Due to practical limitations, a subset of 8 compounds was chosen for experimental validation with an IP1 Gq assay. This subset resulted from a selection and filtering

process with parameters provided by SAFAN-ISPSM (compound-target ranking, predicted binding affinity, presents of an interaction specific fragment, fragment similarity and compound similarity). Finally, the IP1 Gq assay revealed that 3-([1-(naphthalen-1-yl)ethyl]amino)methylphenol ($EC_{50} = 5.5$) has positive allosteric activity but with less potency at CaSR in comparison to the positive allosteric modulator NPS-R568 ($EC_{50} = 6.3$). The calculation of the compound similarity between the discovered compound and approved calcimimetics shows the highest similarity to Evocalcet (Tanimoto = 0,37; Figure 2.68). The similarity is quite low to known CaSR ligands, but SAFAN-ISPSM managed to identify it as one out of eight compounds.

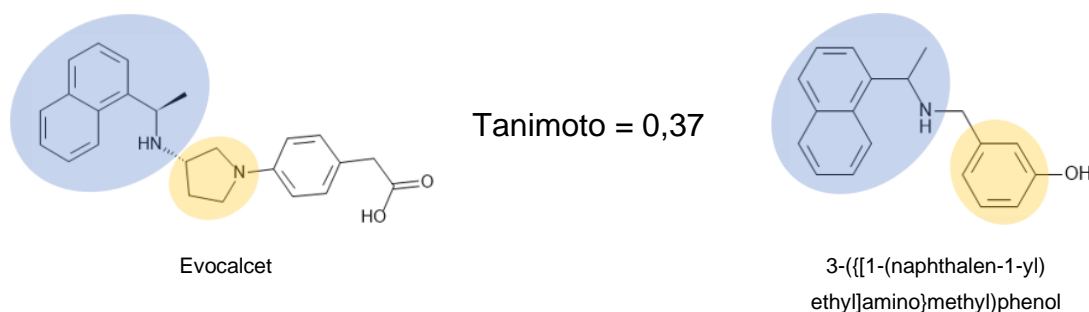


Figure 2.68 Compound similarity between the approved Evocalcet and the discovered positive allosteric modulator

3-([1-(naphthalen-1-yl)ethyl]amino)methylphenol shares the 1-(naphthalen-1-yl)ethylamine substructure with Evocalcet and Cinacalcet (Figure 1.8) which is relevant for the positive allosteric activity (Figure 2.68, blue). Prior studies identified this moiety to have the greatest impact on the quantitative structure-activity relationship and that the potency rank follows 1-naphthyl \gg meta-methoxyphenyl $>$ phenyl. Changes to the secondary amine results in a loss of potency, and depending on the assay, the R-enantiomer is far more potent than the S-enantiomer (18). The rest of the molecule is very tolerant to modification. Another similarity to Evocalcet is the ring substructure attached to the secondary amine, which increases the molecule's rigidity. The IP-1 assay was performed with the racemic form of 3-([1-(naphthalen-1-yl)ethyl]amino)methylphenol and the next step should be the determination of the EC_{50} value of the individual enantiomers. Another point that was not addressed in this study is the determination of the compound's potential biased activity. CaSR interacts with different signalling partners, and in theory, it could have a bigger effect on one of them in comparison.

The discovery of 3-([1-(naphthalen-1-yl)ethyl]amino)methylphenol is the starting point to determine whether or not this compound could become the next therapeutic agents for the treatment of CaSR related diseases. Also, the hitlist of SAFAN-ISPSM could contain other potential CaSR ligands which should be evaluated in future experiments.

3 Discussion:

This thesis's main goal was the discovery of a novel allosteric modulator as a potential agent for the treatment of extracellular calcium-sensing receptor (CaSR) related diseases. The idea was to use computer-aided drug discovery methods for virtual screening of big compound libraries to gain a list of potential allosteric ligands. The potency of these compounds was evaluated with the selective IP-1 Gq assay. Prior studies identified the binding site for allosteric modulators in the upper part of the transmembrane of this G-protein coupled receptor (26). The first attempt for virtual screening involved docking experiments with a model of CaSRs transmembrane domain. The creation of a model with the automated web-based threading server I-TASSER did not result in reliable structures. Instead, homology models in the inactive state were manually created based on the crystal structures of metabotropic glutamate receptor 1 and 5 using YASARA (24). The validation of the models with known allosteric modulators resulted at best in a Pearson correlation of $r = 0.33$. This result leads to strategy change and the use of the ligand-based profiling tool SAFAN-ISPSM. This program is designed for ligand-profiling and predicts the binding affinities against multiple targets. The validation with 13000 maximum divers compounds showed that the Pearson correlation of SAFAN-ISPSM is $r = 0.86$. The program was also evaluated with 332 calcilytics for CaSR and had a Pearson correlation of $r = 0.75$. In the first attempt, the entries of DrugBank and FooDB were profiled with SAFAN-ISPSM 2018/06-B1 and 396 compounds were identified to have a significant binding affinity for CaSR. A subset of seven compounds was selected for experimental validation with an IP-1 Gq assay. The outcome was that toremifene and bucindolol showed negative allosteric modulation activity at 100 fold predicted concentrations but had not a comparable potency as the reference NPS-2143. In a second attempt, the algorithms of SAFAN-ISPSM were optimized to be more sensitive for chirality and less impaired

by atom overrepresentation in the fingerprints. Besides, a new compound library was generated containing more than 250000 small molecules. The library is a combination of different databases including DrugBank, FooDB, HerDing and TCM database. In addition, the binding affinity relevant fragments for CaSR were used from SAFAN-ISPSM to search the ZINC database for potential candidates. Structure-based pharmacophore models were extracted from homology models in the active and inactive conformation. They were used to search ZINC15 database via the ZINCPharmer interface. The new version of SAFAN-ISPSM 2019/06-B2 was validated with about 21000 maximum diverse small molecules and showed a Pearson correlation of 0.91 between the predicted and experimental binding affinities. The correlation for 551 allosteric modulators of CaSR was 0.81 executed with the leave-one-out cross-validation. The profiling of the 250000 small molecules was performed within six days at the OCCAM C3S supercomputer (72) and resulted in 3924 potential allosteric modulators. After a rigorous selection process, a subset of eight compounds was experimentally evaluated with an IP1-Gq assay. The results showed that compound 3-({[1-(naphthalen-1-yl)ethyl]amino}methyl)phenol has a significant positive allosteric modulator activity in comparison with the reference NPS-R568. Structural analysis shows that the found compound is a derivative of Evocalcet and shares the 1-(naphthalen-1-yl)ethylamine substructure, relevant to the activity. The predicted binding affinity for 3-({[1-(naphthalen-1-yl)ethyl]amino}methyl)phenol is at least more than one magnitude of order apart from the experimental tested one. The conclusion is that one out of eight compounds was identified to have a significant allosteric modulator activity for CaSR. Further studies could identify even more active compounds from the 4000 predicted small molecules and could be the starting point of the drug development process.

4 Materials and methods

4.1 Computer-aided drug discovery methods (CADD)

4.1.1 Homology modelling

Homology modelling, also known as comparative modelling, is a structure-based method in bioinformatics. It allows the modelling of experimentally unsolved protein structures by comparing the target's primary sequence to homologous protein entries in the Protein Data Bank (template).

Homology modelling with YASARA

The minimum requirement for YASARA to build a homology model is the target sequence in the FASTA format (Figure 4.1). Besides, it is also possible to provide template structures and alignments manually.

Homology modelling with YASARA starts with a defined number of PSI-BLAST iterations of the FASTA sequence against the UniRef90 database. The result is a position-specific scoring matrix (PSSM), and many parameters influence it. The selected E-value sets the threshold for considering a template sequence. Lower E-values correspond to more significant sequence alignments and reduce the number of templates. Higher numbers of PSI-BLAST iterations include more distantly related sequences. The Protein Data Bank (PDB) (43) is screened with the generated PSSM, excluding protein purification tags to avoid false positives. Suitable templates are identified and sorted by the alignment score, structural quality checks and manual intervention. The selection process is programmed to prefer high-resolution structures over higher sequence identity. The final homology modelling step depends on the entered maximum number of relevant template structures, alignments per template, alignment variations, the oligomerization state, loop conformations and the number of residues for terminal loop extensions. YASARA provides the option to use its PSSP database (Profiles from sequence and structurally related proteins) to create a structural alignment (24). With this option, it applies the SSALN substitution matrix (102) to identify structurally related

proteins. The secondary structure of the target sequence gets predicted with the online server PSIPRED (24,47). The selection process keeps only top-scoring templates of high quality. Alternative alignments result from a stochastic pairwise approach (103). Oligomeric states are also considered for the final models unless it is defined otherwise. YASARA enables the extension of missing loops and termini by up to 10 residues. Predictions are not reliable for larger gaps.

The subsequent model refinement uses a simple repulsive energy function, electrostatic and knowledge-based packing interactions to optimize the side chains. It also involves the sampling of various loop conformations, and the best is selected. The hydrogen-bonding network is optimized, considering pH and ligands. In the next step, the model is energy minimized with an explicit solvent shell using the latest version of YASARA's knowledge-based force field YASARA 2 (41). It involves a steep descent minimization using a 10,5 Angstrom force cut-off to remove any bumps and uses the Particle Mesh Ewald algorithm (104) to treat long-range electrostatic interactions. Besides, simulated annealing finishes after the minimization reaches convergence, and the energy improves less than 0,05 kJ/mol (timestep 2 fs, atom velocities scaled down by 0,9 every 10th step) (24). The system moves to the next local minimum, which does not have to be the global one.

The homology modelling process uses only one energy minimization step. It includes automatic parametrization of ligands with AutoSMILES. The solvation shell is primarily for the screening of electrostatic interactions. The energy minimization takes place in two parts. In part one, a half-refined model results after minimizing the model with fixed backbone atoms. In the second part, the whole system gets minimized.

The final selection depends on the highest overall quality Z-score of all generated models. YASARA also builds a hybrid model composed of the best-scored models. Low-quality regions get replaced by higher-scoring segments from other models. The final model with the highest Z-score can have a quality score ranging from optimal (>0) to terrible (< -5).

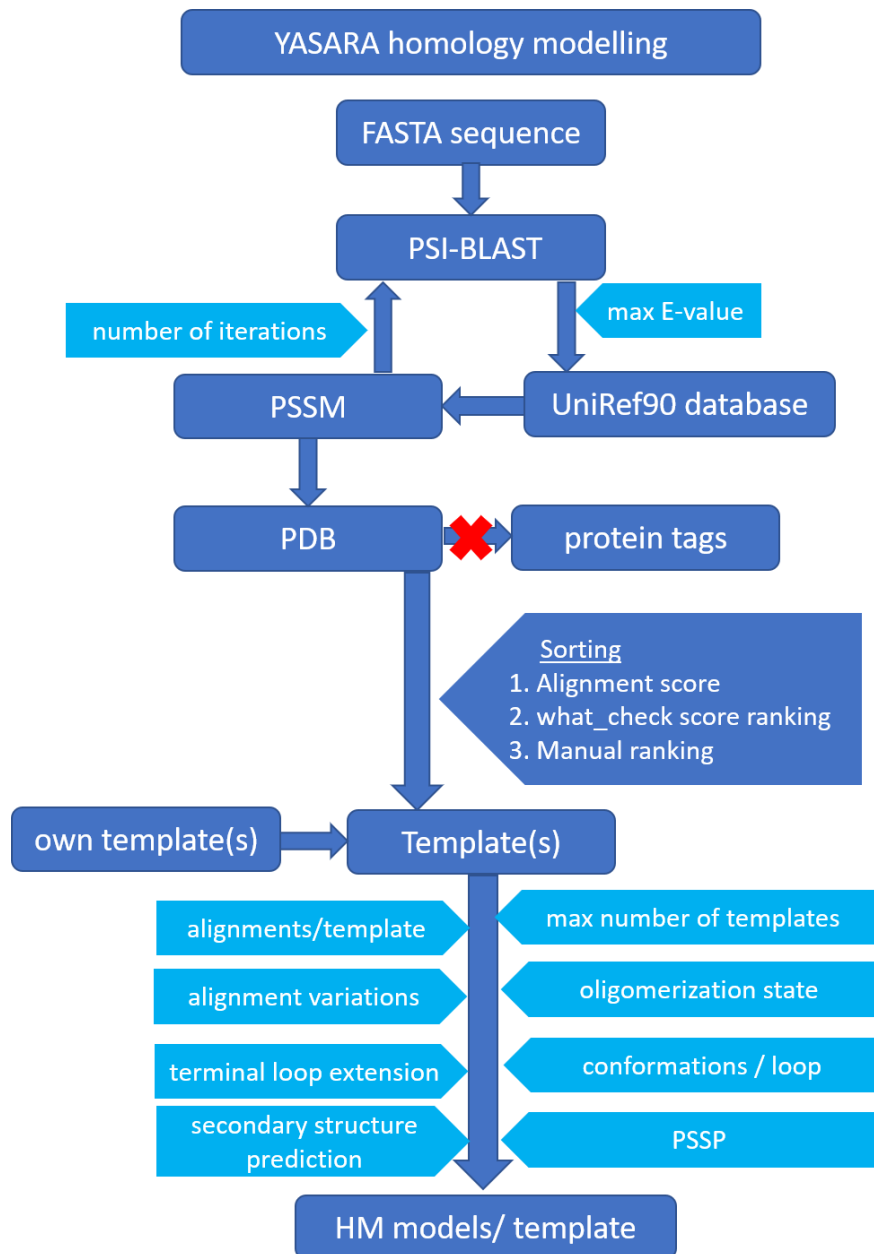
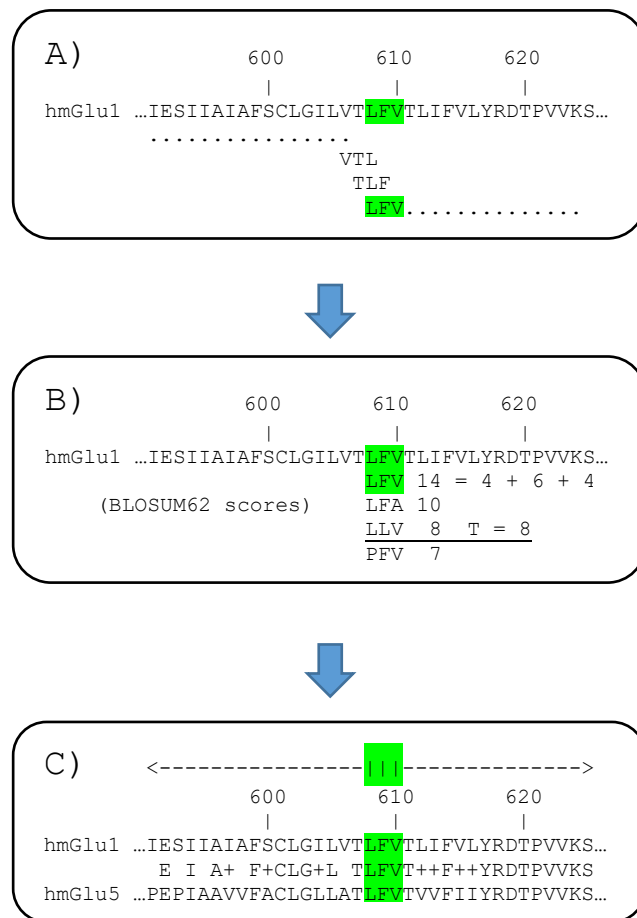


Figure 4.1 YASARA homology modelling workflow, the description is in the text, PSI-BLAST: position-specific iterative basic local alignment search tool, PSSM: position-specific scoring matrix, PDB: Protein Data Bank, HM: homology modelling; PSSP: profiles from sequence and structurally related proteins

BLASTp: Basic Local Alignment Search Tool for proteins

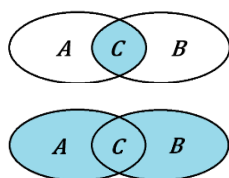
BLASTp is a fast heuristic tool that approximates the alignment between sequences and identifies homologous proteins. BLAST operates in three major steps (*Figure 4.2*). First, it generates high scoring words with a fixed length (W) derived from the target sequence. Second, the tool uses the high scoring words to scan a template sequence or database for hits. The hits get scored according to the BLOSUM62 substitution matrix (*Figure 4.3*). Only results remain with a score higher than a threshold T . In the third step, residues are added or removed from words until the resulting segment score does not increase anymore. If the segment score is higher than the cut-off S , the match gets into the result list. Finally, the aligned sequences get sorted from small to big expect-values (E values) that correspond to the chance of random matches. Lower E values indicate a higher significance for the match (105).



4.1.2 In-silico profiling with SAFAN-ISPSM

In the field of computer-aided drug discovery (CADD), the methods are divided into two main areas, structure-based drug discovery (SBDD) and ligand-based drug discovery (LBDD). As the name suggests, the requirement in SBDD is the information about the three-dimensional structure of the target. The most prominent methods in this field are homology modelling, molecular docking experiments and molecular dynamics simulations. However, if the structure of the target is not known and cannot be modelled, methods of LBDD are used to discover new lead structures (111). SAFAN-ISPSM is a ligand-based drug discovery method, and its main application is in-silico profiling of small molecules. Usually, classical virtual screening methods predict the binding affinities (pK) of numerous compounds against one target, the results get ranked, and the top-scoring compounds are selected for experimental evaluation. But in the case of in-silico profiling, one compound can be screened against multiple targets. In this case, the most potential target is predicted for every compound, and the results make it possible to estimate the side effects and toxicity. The value of pK describes the strength of the interaction and is defined as the negative decadic logarithm of the dissociation constant K_D . It represents a ligand concentration at which half of the ligands are bound to the target in the complex. The lower the dissociation constant K_D , the higher is the corresponding binding affinity pK (e.g. $K_D = 1 \text{ nM} \triangleq \text{pK} = 9$). The pK describes a compound's nature, if it is a strong or weak binder but does not give any information about the activity. Therefore, other indicators are used to describe the concentration of half-maximum inhibition (IC_{50}) or the concentration of half-maximum stimulation (EC_{50}). The prediction of SAFAN-ISPSM is based on re-evaluated binding data from the ChEMBL database (59). The pChEMBL value defines the interactions listed on ChEMBL. It is a mixture of metrics representing the concentration of different half-maximum activities (IC_{50} , EC_{50} , AC_{50} , XC_{50} , K_i , K_d or Potency) although not all the values are directly comparable (59). SAFAN-ISPSM requires the input structure of the compound library in SMILE format (page 123). A fragmentation process is initiated and breaks down the one-dimensional representation of the structure into smaller substructures. SAFAN-ISPSM describes them with a unique fingerprint based on a combination of the smallest unit called 'string'. The strings are divided into four categories: aliphatic, aromatic, ramification and chiral. The fragments of the unknown compound get compared to

a list of known structures. The identification of binding affinity relevant fragments (BARF) helps determine the corresponding target, and the binding affinity is determined with a fragment weight assignment algorithm. The predictions of LBDD methods are based on the similarity measurements between known and unknown structures. The similarity in SAFAN-ISPSM is calculated with the Tanimoto coefficient (S) (Equation 8). The fingerprint of the target structure is compared to the entries of the knowledge-based database. The Tanimoto coefficient (S) is defined by the fraction of the common part (c) between two structures and all the strings in the fingerprints without counting the overlapping common part twice (A + B - c).



$$S = \frac{c}{A + B - c} \quad \text{Equation 8}$$

The Tanimoto coefficient can have a value between 0 and 1. However, a Tanimoto of 1.0 does not automatically mean that two compared structures are identical. It means that both compounds are represented by the same substructures, which could be arranged differently. This situation applies to tautomeric forms of the same compound. While open-source programs like Open Babel (77) struggle to identify two tautomers to be the same, SAFAN-ISPSM avoids this mismatch even if the compound is divided into fragments. Another difficulty represents chiral atoms. The challenge starts with the SMILE format because chiral atoms are not defined absolutely (R, S) but as anticlockwise (@) and clockwise (@@). The reconstruction of the absolute configuration becomes a challenge if not all the relevant atoms are present in the fragment. This problem is addressed in a later version of SAFAN-ISPSM by calculating several Tanimoto coefficients considering different aspects of the compared structures. In the end, the most probable binding affinity is calculated from the different results by a support vector machine algorithm (SVM) utilizing WEKA (112).

The results of SAFAN-ISPSM is a target hit list for the profiled compound including ranked interaction position, name of the target, predicted binding affinity, the similarity of compounds (alltan) and similarity of fragments (fragtan). The indicator round defines if a BARF was present in the compound (round = 1) or not (round =

0). The SAFAN-score represents the confidence in the calculated pK based on alltan and fragtan.

Table 4.1: SAFAN-ISPSM representative results version 06/2019; first column: target gene name; alltan: Tanimoto similarity between compounds; fragtan: Tanimoto similarity between fragments; round: presents of a binding affinity relevant fragment (BARF); SAFAN-score: confidence score of the results based on alltan and fragtan

position	target	pK	alltan	fragtan	round	safan-score
...
125	HCD2	5.371	0.612	0.871	0	0.830
126	ICAM1	5.368	0.564	1.000	0	0.892
127	ITAL	5.368	0.564	1.000	0	0.832
128	ITB2	5.368	0.564	1.000	0	0.820
129	CASR	5.368	0.470	1.000	0	0.879
130	CP1A2	5.366	0.687	0.882	0	0.899
131	ADA1B	5.366	0.569	0.945	0	0.895
132	NOS2	5.348	0.623	0.875	0	0.805
133	1CYSP_nomam	5.347	0.613	0.873	0	0.879
134	BRD4	5.342	0.473	0.877	0	0.888
135	TPO	5.337	0.728	?	0	0.821
...

The profiling hit list starts with the highest pK prediction. Ideally, the similarity to other compounds (alltan) should be low to find a new scaffold. The similarity of fragments should be high (fragtan) to identify active compounds. If the index round is 1, the compound contains a binding-affinity relevant fragment, and the probability is higher to discover an active compound. The SAFAN-score can have values between 0 and 1, and it represents the probability that the prediction is close to the true value.

4.1.3 File formats

FASTA file format

The FASTA format is widely used to represent biological sequences, and it was first presented in 1985 by William Pearson and David Lipman (113). The format contains the sequence of proteins or nucleic acids stored in plain text. It starts with a smaller than sign (>) and a description of the protein or nucleic acid. A (|) symbol separates additional meta-data. The second line contains the protein sequence in one-letter code. It is possible to store multiple sequences in one file. The FASTA

format is primarily used for sequence comparisons in bioinformatics. Figure 1.4 shows the FASTA sequence of CaSR.

PDB file format

The PDB file format contains a text-based 3D structure description of experimentally determined macromolecules first published on the Protein Data Bank database (43). Each line of this file starts with a record describing the following information. In the first line is the HEADER with the PDB-ID consisting of 4 alphanumeric characters and additional information. The following lines contain metadata like the TITLE and information about the compounds (COMPND); the SOURCE, keywords (KEYWDS), the method for the structure determination (EXPDTA), the AUTHOR, literature citation (JRNL) and others. Usually, there is a section of general remarks (REMARK) listing information about the experimental setting. The most important lines contain the coordinates of the individual protein atoms (ATOM) or atoms of other molecules (HETATM). These lines have the information about the atom number, atom type, amino acid, protein chain letter, residue number in the protein, XYZ coordinates, occupancy, temperature factor (B-factor) and the element. At the bottom, each PDB file closes with an END. Several other record types appear in PDB files, and the whole description can be found at the www.wwpdb.org website (108). An excerpt of the PDB file 5K5T.pdb is shown below.

```
HEADER      SIGNALING PROTEIN                               23-MAY-16   5K5T
TITLE       CRYSTAL STRUCTURE OF THE INACTIVE FORM OF HUMAN CALCIUM-SENSING
TITLE       2 RECEPTOR EXTRACELLULAR DOMAIN
COMPND      MOL_ID: 1;
... [lines skipped]
SOURCE      MOL_ID: 1;
SOURCE      2 ORGANISM_SCIENTIFIC: HOMO SAPIENS;
... [lines skipped]
KEYWDS      VENUS FLYTRAP MODULE, CYSTEINE RICH DOMAIN, HOMODIMER, SIGNALING
KEYWDS      2 PROTEIN
EXPDTA      X-RAY DIFFRACTION
AUTHOR      Y.GENG, L.MOSYAK, I.KURINOV, H.ZUO, E.STURCHLER, T.C.CHENG, P.SUBRAMANYAM,
... [lines skipped]
JRNL        AUTH   Y.GENG, L.MOSYAK, I.KURINOV, H.ZUO, E.STURCHLER, T.C.CHENG,
... [lines skipped]
REMARK      2 RESOLUTION.      3.10 ANGSTROMS.
... [lines skipped]
ATOM        1  N   GLY  A  21           8.462  -6.889  -71.589  1.00112.57      N
ANISOU      1  N   GLY  A  21          14382  14826  13563   -202    281    256      N
ATOM        2  CA  GLY  A  21           8.791  -5.962  -70.513  1.00112.05      C
ANISOU      2  CA  GLY  A  21          14307  14677  13588   -152    312    272      C
ATOM        3  C   GLY  A  21          10.273  -5.927  -70.188  1.00115.62      C
... [lines skipped]
HETATM     4568  O1  SO4  A  702          11.102  -23.635  -50.011  1.00  81.34      O
HETATM     4569  O2  SO4  A  702          10.438  -23.019  -47.785  1.00  84.58      O
... [lines skipped]
END
```

SMILES

SMILES stand for Simplified Molecular Input Line Entry System. It is a compact language for the one-dimensional description of a chemical structure or reaction. It is based on a simple vocabulary and involves only a few grammar rules. This representation is very useful in data processing of big chemical libraries because SMILES can be treated as words. Each structure can be represented by a unique SMILE, also known as 'canonical SMILE', although the rules allow multiple descriptions of the same structure. SMILES that contain chiral information are called 'isomeric SMILES' (114). The isomeric SMILE of the negative allosteric modulator is, for example:

```
CC(C)(CC1=CC2=CC=CC=C2C=C1)NC[C@H](COC3=C(C(=CC=C3)Cl)C#N)O
```

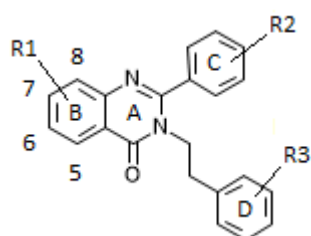
In this description, the structure is represented with no spaces and includes implicit hydrogens. SMILES are created based on five general rules defining atoms, bonds, connecting atoms in a ring, and disconnections. Atoms are described by their element symbols. Elements themselves are written by the symbol in square brackets. The second letter of elements is written in lower case. Aromatic atoms like carbons, sulfur, oxygen and nitrogen are written in lower case and aliphatic atoms in upper case. Atoms with abnormal valences are denoted in brackets. Hydrogens and charges need to be mentioned in brackets as H, + and -. Bonds are described by the symbols – (single); = (double) or # (triple). Single and aromatic bonds are all omitted for adjacent atoms. Branches are denoted by enclosing the concerning atoms in parentheses. Cyclic structures are described by breaking one bond of the ring and label the created terminal atoms with the same number. Disconnected structures have a period symbol "." between their compound SMILES. Additional symbols can be used for further specifications of the structure. The characters / and \ are used to distinguish E- and Z- configurations around a double bond. @ and @@ characters are written before a chiral atom and are together enclosed in brackets. @ and @@ describe the anticlockwise and clockwise character of the attached atoms but do not define the absolute chirality. A detailed description of all the rules is listed at <https://daylight.com/dayhtml/doc/theory/theory.smiles.html> (115).

4.2 The validation set for virtual screening:

The validation set was downloaded from the ChEMBL database (59). It contains a list of compounds showing activity at CaSR in human, mice, rats and cows. Entries with a clearly defined binding affinity were considered for the validation set. The compounds displayed here are divided into negative and positive allosteric modulators and were used to validate homology models and the predictions of SAFAN-ISPSM (70). At the end of every table is the interaction profile of the most potent derivative created by molecular docking with the model on page 53. The interactions are colour coded and represent hydrophobic interactions (yellow), hydrogen donor bonds (green), hydrogen acceptor bond (red), positive ionizable areas (blue star), negative ionizable areas (red star), and aromatic π - π interactions (blue ring) (85).

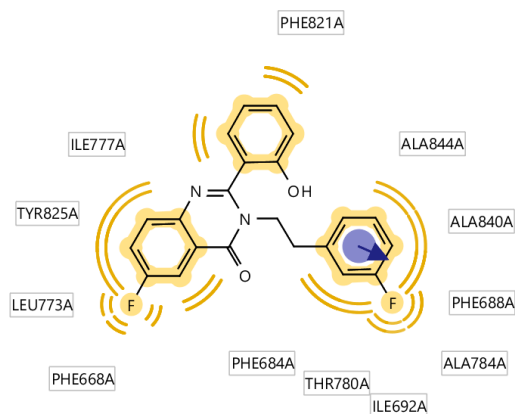
4.2.1 Negative allosteric modulators for the human extracellular calcium-sensing receptor

Table 4.2: 4(3H)-quinazolinones; negative allosteric modulators tested with fluo-3 in vitro with HEK 293 4.0-7 cells in a calcium concentration-dependent manner, compounds were tested with the fluorescence indicator Biotium by increasing extracellular calcium from 1.0 to 1.3 mM, the signal was normalized to no compound, all compounds were tested as duplicates at eight different concentrations with 30 μ M as the highest (59,116)



2- phenyl- 3- phenethyl- 4(3H)- quinazolinone

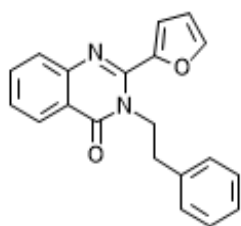
CHEMBL-ID	R1	R2	R3	IC50 [nM]
CHEMBL180391	H	H	H	14000
CHEMBL179333	H	3-OH	H	2800
CHEMBL182107	H	2,5-di-OH	H	1600
CHEMBL178442	7-Cl	2-OH	3-F	1600
CHEMBL180250	H	2-OH	3-Cl	800
CHEMBL179907	8-methyl	2-OH	H	700
CHEMBL361316	7-F	2-OH	H	600
CHEMBL180057	6-methyl	2-OH	3-F	520
CHEMBL361333	6-Cl	2-OH	H	300
CHEMBL360226	5-methyl	2-OH	3-F	250
CHEMBL361623	5-methyl	2-OH	H	250
CHEMBL183287	6-F	2-OH	H	210
CHEMBL180463	6-F	2-OH	3-F	190



CHEMBL180463

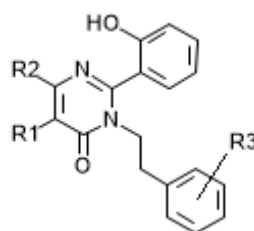
Figure 4.4 CHEMBL 180463, interaction map of the most potent 4(3H)-quinazolinones ($IC_{50} = 190$ nM) docked to CaSRs homology model on page 53

Table 4.3: 4(3H)-pyrimidinone; negative allosteric modulators tested with fluo-3 in vitro with HEK 293 4.0-7 cells in a calcium concentration-dependent manner (117)



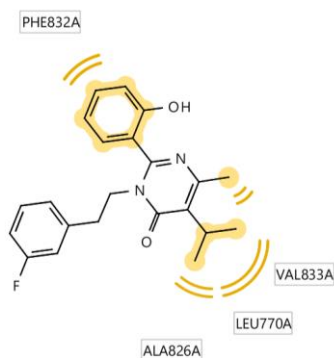
Chembl180336 (NPS 53574)

$IC_{50} = 3500$ nM



2-(2-hydroxyphenyl)- 3-phenethyl- 4(3H)-pyrimidinone

pyrimidinones	R1	R2	R3	IC_{50} [nM]
CHEMBL195176	H	methyl	H	1900
CHEMBL373306	H	methyl	2-F	1600
CHEMBL194063	H	methyl	3-F	1100
CHEMBL195045	methyl	CF ₃	H	800
CHEMBL371828	methyl	methyl	H	200
CHEMBL370503		-(CH) ₃ -	H	200
CHEMBL195375		-(CH) ₃ -	3-F	200
CHEMBL193498	methyl	methyl	3-F	170
CHEMBL194490		-(CH) ₄ -	3-F	160
CHEMBL369998		-(CH) ₄ -	H	140
CHEMBL195566	propyl	methyl	3-F	140
CHEMBL195305	cyclopropyl	methyl	3-F	120
CHEMBL195859	ethyl	methyl	H	100
CHEMBL194764	ethyl	methyl	2-F	97
CHEMBL195711	isopropyl	methyl	3-F	95

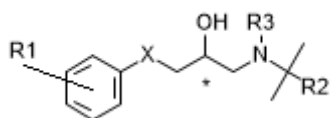


CHEMBL195711

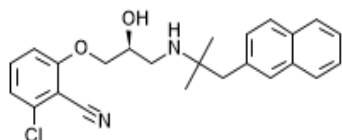
Figure 4.5 CHEMBL195711, interaction map of the most potent 4(3H)-pyrimidinone ($IC_{50} = 95 \text{ nM}$) docked to CaSRs homology model on page 53

A hydrogen bond between 2(2-hydroxyphenyl) and N1 could influence the activity.

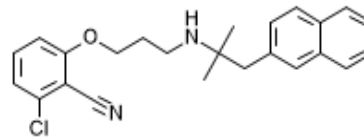
Table 4.4 Aryloxy propanolamines; IC_{50} determined in human TT cells by Fluorescent Imaging Plate Reader (FLIPR™) measuring the inhibition of intracellular calcium (118)



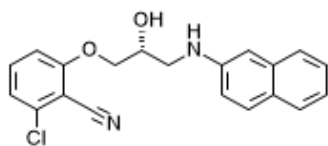
CHEMBL-ID	X	R1	R2	R3	*	IC_{50} [nM]
CHEMBL434937	O	3-CH ₃ ; 4-SO ₂ CH ₃	4-methoxyphenyl	H	R,S	11360
CHEMBL200248	O	2-SO ₂ CH ₃	phenyl	H	S	9430
CHEMBL199473	S	2,3-di-Cl	4-F-phenyl	H	R	6180
CHEMBL201065	O	4-OCF ₃	4-F-phenyl	H	R	3300
CHEMBL201381	O	3-methoxyphenyl	4-methoxyphenyl	H	S	1390
CHEMBL199363	O	2-CN; 3-Cl	4-F-phenyl	H	R	600
CHEMBL370811	O	2-CN ; 3-Cl	naphth-2-yl	CH ₃	R	270
CHEMBL371936	O	2-CN	1H-indol-2-yl	H	S	100
CHEMBL199894	O	2-CN	1-benzothiophen-2yl	H	S	60
CHEMBL200311	O	2-CN; 3-Cl	naphth-2-yl	H	S	50



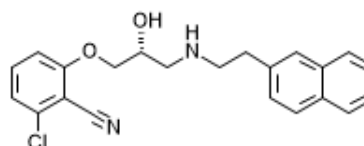
CHEMBL200311 (NPS-2143) ; $IC_{50} = 50 \text{ nM}$



CHEMBL199785; $IC_{50} = 880 \text{ nM}$



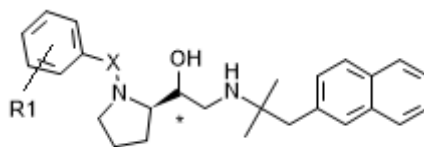
CHEMBL200423; $IC_{50} = 4600 \text{ nM}$



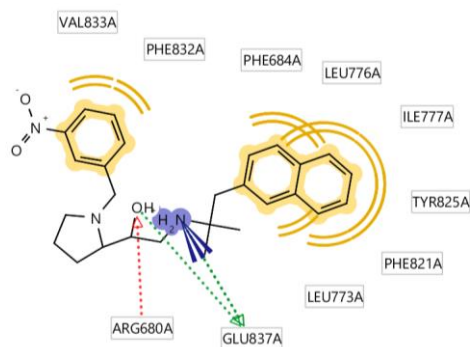
CHEMBL372821; $IC_{50} = 3860 \text{ nM}$

Figure 4.6: impact of structural modifications on the structure-activity-relationship of NPS-2143 (118,119)

Table 4.5: (R)-1-arylmethylpyrrolidin-2-yl ethanolamines in human TT cells measuring the inhibition of intracellular calcium release by FLIPR™ (118)



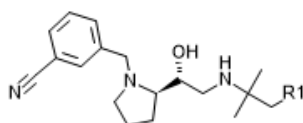
CHEMBL-ID	X	R1	*	IC50 [nM]
CHEMBL383842	CO	3-CN	R	8430
CHEMBL200878	CH ₂	3-COOH	R	6840
CHEMBL425394	CH ₂	3-phenyl	R	2470
CHEMBL263828	CH ₂	3-Br	R	1570
CHEMBL200422	CH ₂	2-OH, 3-methoxy	R	1560
CHEMBL199964	CH ₂	2-CN	S	1270
CHEMBL265362	CH ₂	3-methoxy	R	1190
CHEMBL200297	CH ₂	3-OH	R	1160
CHEMBL201085	CH ₂	2-OH	R	970
CHEMBL199529	CH ₂	4-NO ₂	R	790
CHEMBL201137	CH ₂	3-F	R	750
CHEMBL200690	CH ₂	2,3-di-F	R	360
CHEMBL200361	CH ₂	2-CN	R	340
CHEMBL199548	CH ₂	4-CN	R	250
CHEMBL203479	CH ₂	2-CN, 3-Cl	R	240
CHEMBL199716	CH ₂	3-CN	R	190
CHEMBL197863	CH ₂	2-NO ₂	R	160
CHEMBL199495	CH ₂	2-OH, 3-F	R	140
CHEMBL372321	CH ₂	2-OH, 3- NO ₂	R	140
CHEMBL372820	CH ₂	3-NO ₂	R	140



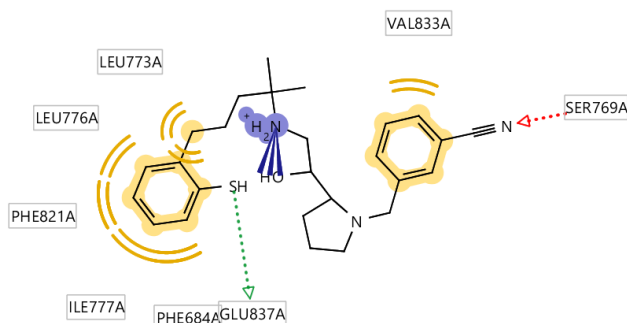
CHEMBL372820

Figure 4.7 CHEMBL372820, interaction map of the most potent (R)-1-arylmethylpyrrolidin-2-yl ethanolamines (IC₅₀ = 140 nM) docked to CaSRs homology model on page 53

Table 4.6: 3-[[[(2R)-2-[(1R)-2-[[1-aryl-2-methylpropan-2-yl]amino]-1-hydroxyethyl]-pyrrolidin-1-yl]methyl]benzo-nitriles in human TT cells measured by FLIRP (118)



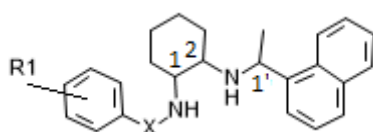
CHEMBL-ID	R1	IC50 [nM]
CHEMBL383609	(1H)-benzimidazol-2-yl	2840
CHEMBL370671	4-methoxyphenyl	2670
CHEMBL381685	1,3-benzothiazol-2-yl	1350
CHEMBL200041	1-benzothiophen-3-yl	1000
CHEMBL538155	1-indol-2-yl	440
CHEMBL200312	1-benzothiophen-2-yl	270



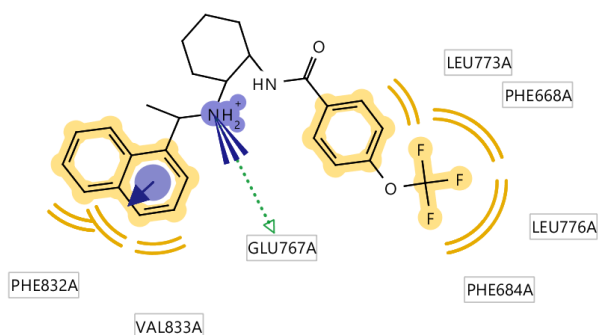
CHEMBL200312

Figure 4.8 CHEMBL200312, interaction map of the most potent pyrrolidin-1-yl]methyl]benzo-nitriles (IC50 = 270 nM) docked to CaSRs homology model on page 53

Table 4.7: N1-Benzoyl-N2-[1-(1-naphthyl)ethyl]-trans-1,2-diaminocyclohexanes; inhibition of [³H]IP accumulation produced by Ca²⁺ (9mM) in CHO cells expressing rat cloned CaSR by the test compounds (120)



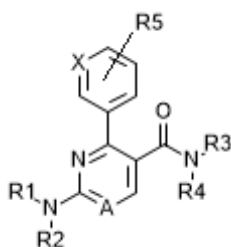
CHEMBL-ID	R1	X	1	2	1'	IC50 [nM]
CHEMBL559185	4-methoxy	SO ₂	R	S	R,S	8000
CHEMBL537426	4-OCF ₃	SO ₂	R	S	R,S	5400
CHEMBL537427	4-OCF ₃	SO ₂	S	R	R	2600
CHEMBL556781	4-OCF ₃	CO	R	S	S	1700
CHEMBL557567	4-OCF ₃	CO	R	S	R,S	900
CHEMBL537870	3,4-di-Cl	CO	S	R	R	330
CHEMBL536975	4-OCF ₃	CO	S	R	R	250



CHEMBL536975

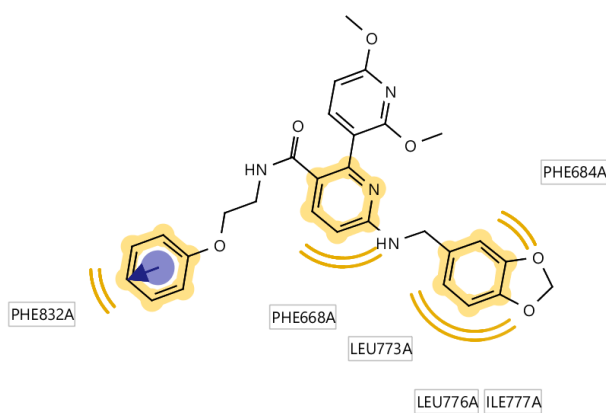
Figure 4.9 CHEMBL536975, interaction map of the most potent N1-Benzoyl-N2-[1-(1-naphthyl)ethyl]-trans-1,2-diaminocyclohexanes (IC₅₀ = 250 nM) docked to CaSRs homology model on page 53

Table 4.8: Trisubstituted pyrimidines/pyridines; IC₅₀ determined in HEK-293 cells using FLIRP assays to detect the inhibition of intracellular calcium release (121)



CHEBML-ID	A	X	NR1R2	NR3R4	R5	IC ₅₀ [nM]
CHEMBL472731	N	C	N-benzyl, N-methyl	N-(2-phenoxyethyl)	4-CF ₃	18900
CHEMBL453360	N	C	N-(2-methylbenzyl)	N-(2-phenoxyethyl)	3,4,5-trimethoxy	10000
CHEMBL473345	N	C	N-benzyl, N-methyl	N-(2-phenoxyethyl)	3,4-dimethoxy	6900
CHEMBL454158	N	C	N-(2-Cl-benzyl)	N-(2-phenoxyethyl)	3,4,5-trimethoxy	4900
CHEMBL523682	N	C	N-(4-methoxybenzyl)	N-(2-phenoxyethyl)	3,4,5-trimethoxy	3100
CHEMBL523701	N	C	N-(4-methoxyphenethyl)	N-(2-phenoxyethyl)	3,4,5-trimethoxy	2800
CHEMBL454375	N	C	N-(3-methylbutyl)	N-(2-phenoxyethyl)	3,4,5-trimethoxy	2200
CHEMBL472406	N	C	N-benzyl, N-ethyl	N-3-phenylpropyl	3,4,5-trimethoxy	2100
CHEMBL511485	N	C	N-benzyl, N-ethyl	N-4-phenylbutyl	3,4,5-trimethoxy	1800
CHEMBL494646	N	C	N-methyl, N-phenethyl	N-(2-phenoxyethyl)	3,4,5-trimethoxy	1500
CHEMBL455406	C	C	N-benzyl, N-ethyl	N-(2-phenoxyethyl)	2,4-dimethoxy	1050
CHEMBL494441	N	C	N-(3-methoxyphenethyl)	N-(2-phenoxyethyl)	3,4,5-trimethoxy	1000
CHEMBL454374	N	C	N-benzyl	N-(2-phenoxyethyl)	3,4,5-trimethoxy	730
CHEMBL510598	N	C	N-benzyl, N-methyl	N-(2-phenoxyethyl)	3,4,5-trimethoxy	470

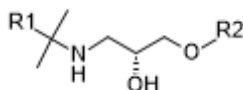
CHEMBL494647	N	C	N-phenethyl	N-(2-phenoxyethyl)	3,4,5-trimethoxy	400
CHEMBL462661	N	N	N-benzyl, N-ethyl	N-(2-phenoxyethyl)	2,4-dimethoxy	360
CHEMBL458515	N	N	N-(1,3-benzodioxol-5-yl)-methyl	1,2,3,4-tetrahydroisoquinolin-2-yl	2,4-dimethoxy	360
CHEMBL493228	N	C	N-(3-methoxybenzyl)	N-(2-phenoxyethyl)	3,4,5-trimethoxy	340
CHEMBL494442	N	C	N-(2-methoxyphenethyl)	N-(2-phenoxyethyl)	3,4,5-trimethoxy	340
CHEMBL458516	N	N	N-(1,3-benzodioxol-5-yl)-methyl	N-(2-phenoxyethyl)	2,4-dimethoxy	220
CHEMBL508063	N	C	N-benzyl, N-ethyl	N-(2-phenoxyethyl)	3,4,5-trimethoxy	200
CHEMBL460726	C	N	N-(2-Cl-benzyl)	N-(2-phenoxyethyl)	2,4-dimethoxy	197
CHEMBL514434	C	C	N-(2-Cl-benzyl)	N-(2-phenoxyethyl)	3,4,5-trimethoxy	162
CHEMBL454386	C	C	N-benzyl, N-ethyl	N-(2-phenoxyethyl)	3,4,5-trimethoxy	154
CHEMBL495263	N	C	1(S)-N-(1-benzyl-2-hydroxyethyl)	N-(2-phenoxyethyl)	3,4,5-trimethoxy	150
CHEMBL509410	N	C	N-(1,3-benzodioxol-5-yl)-methyl	N-(2-phenoxyethyl)	3,4,5-trimethoxy	150
CHEMBL454385	N	N	N-(1,3-benzodioxol-5-yl)-methyl	N-(4-Cl-phenylprop-2-yl)	2,4-dimethoxy	150
CHEMBL515828	C	N	1(S)-N-(1-benzyl-2-hydroxyethyl)	N-(2-phenoxyethyl)	2,4-dimethoxy	140
CHEMBL451383	C	C	N-(1,3-benzodioxol-5-yl)-methyl	N-(2-phenoxyethyl)	3,4,5-trimethoxy	130
CHEMBL457071	C	C	1(S)-N-(1-benzyl-2-hydroxyethyl)	N-(2-phenoxyethyl)	3,4,5-trimethoxy	76
CHEMBL493638	N	C	N-(2-Cl-benzyl)	N-(2-phenoxyethyl)	3,4,5-trimethoxy	70
CHEMBL457249	C	N	N-(1,3-benzodioxol-5-yl)-methyl	N-(2-phenoxyethyl)	2,4-dimethoxy	60



CHEMBL457249

Figure 4.10 CHEMBL457249, interaction map of the most potent trisubstituted pyrimidines/pyridines; IC₅₀ = 60 nM) docked to CaSRs homology model on page 53

Table 4.9: phenoxy analogues of NPS-2143 and ronacalcet, IC50 determined in HEK cells, human liver microsome and rat liver (16)



CHEMBL-ID	R1	R2	IC50 [nM]
CHEMBL1783765	3-F-4-methylbenzyl	3-(3-carboxyl-tricyclo-[3.2.1.02,4]octan-6yl)phenyl	3500
CHEMBL1783936	Indan-2-ylmethyl	(1R)-1-[2-(3-carboxyl-tricyclo-[3.2.1.02,4]octan-6yl)-phenyl]ethyl	3500
CHEMBL1783763	Indan-2-ylmethyl	3-(3-carboxyl-tricyclo-[3.2.1.02,4]octan-6yl)phenyl	900
CHEMBL1783932	1-oxoisindolin-2-yl	2-cyano-5-(3-carboxyl-tricyclo-[3.2.1.02,4]octan-6yl)phenyl	350
CHEMBL1783766	3-F-4-methylbenzyl	2-fluoro-5-(3-carboxyl-tricyclo-[3.2.1.02,4]octan-6yl)phenyl	300
CHEMBL1198855	Indan-2-ylmethyl	2,3-difluoro-5-(3-carboxypropyl)-phenyl	146
CHEMBL1783762	Indan-2-ylmethyl	2,3-difluoro-5-(3-carboxyl-tricyclo-[3.2.1.02,4]octan-6yl)phenyl	56
CHEMBL1783934	4-Cl-2-F-benzyl	(1R)-1-[2-(3-carboxyl-tricyclo-[3.2.1.02,4]octan-6yl)-phenyl]ethyl	50
CHEMBL1783933	3-F-4-methylbenzyl	(1R)-1-[2-(3-carboxyl-tricyclo-[3.2.1.02,4]octan-6yl)-phenyl]ethyl	25
CHEMBL1783931	3-F-4-methylbenzyl	2-cyano-5-(3-carboxyl-tricyclo-[3.2.1.02,4]octan-6yl)phenyl	23
CHEMBL1783935	3-Cl-4-methylbenzyl	(1R)-1-[2-(3-carboxyl-tricyclo-[3.2.1.02,4]octan-6yl)-phenyl]ethyl	18
CHEMBL1783764	Indan-2-ylmethyl	2-cyano-5-(3-carboxyl-tricyclo-[3.2.1.02,4]octan-6yl)phenyl	6

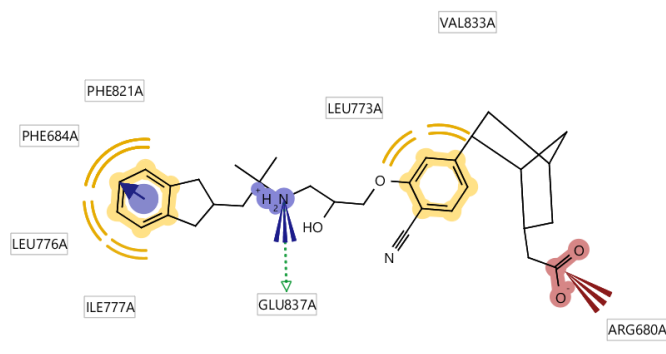
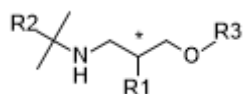
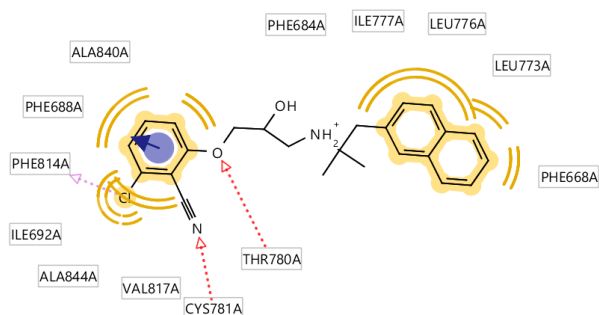


Figure 4.11 CHEMBL1783764, interaction map of the most potent phenoxy analogues of NPS-2143 and ronacalcet (IC50 = 6 nM) docked to CaSRs homology model on page 53

Table 4.10: amino alcohol-based parathyroid hormone secretagogues; IC50 determined in HEK-293 cells using FLIRP assays to detect the inhibition of intracellular calcium release and ³H radioimmunoassay for binding studies (17)



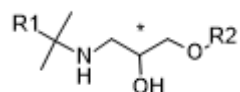
CHEMBL-ID	R1	R2	R3	(*)	IC50 [nM]
CHEMBL482873	OH	4-methoxybenzyl	2-propyl	R,S	10000
CHEMBL482462	OH	4-methoxybenzyl	propyl	R,S	7000
CHEMBL4106263	OH	4-methoxybenzyl	2-methoxyphenyl	R,S	3500
CHEMBL482874	OH	4-methoxybenzyl	butyl	R,S	3200
CHEMBL483080	OH	4-methoxyphenethyl	phenyl	R,S	2780
CHEMBL488337	OH	4-methoxybenzyl	phenyl	R,S	2300
CHEMBL488737	OH	4-methoxybenzyl	2,3-dichlorophenyl	R,S	2300
CHEMBL487511	OH	4-methoxybenzyl	2-naphthyl	R,S	1300
CHEMBL488997	OH	4-methoxybenzyl	2-chlorophenyl	R,S	1000
CHEMBL521907	OH	4-methoxybenzyl	3-chlorophenyl	R,S	1000
CHEMBL488738	OCH ₃	4-methoxybenzyl	2,3-dichlorophenyl	R,S	950
CHEMBL520396	H	4-methoxybenzyl	2,3-dichlorophenyl	R,S	930
CHEMBL487508	OH	4-methoxybenzyl; N-methyl	2,3-dichlorophenyl	S	300
CHEMBL488736	OH	1-naphthylmethyl	phenyl	R,S	210
CHEMBL491251	OH	4-methoxybenzyl	2,3-dichlorophenyl	R,S	200
CHEMBL504479	OH	4-methoxybenzyl	2-cyanophenyl	R,S	140
CHEMBL523475	OH	4-methoxybenzyl	3-chloro-2-cyanophenyl	R,S	70
CHEMBL48750	OH	4-methoxybenzyl	2,3-dichlorophenyl	R	68
CHEMBL2112075	OH	2-naphthylmethyl	3-Cl-2-cyanophenyl	R	3



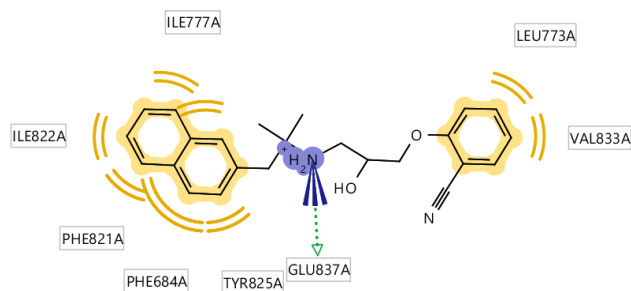
CHEMBL2112075

Figure 4.12 CHEMBL2112075, interaction map of the most potent amino alcohol-based parathyroid hormone secretagogues (IC50 = 3 nM) docked to CaSRs homology model on page 53

Table 4.11: Amino alcohol-based parathyroid hormone secretagogues; IC50 determined in HEK-293 cells using FLIRP assays to detect the inhibition of intracellular calcium release (122)



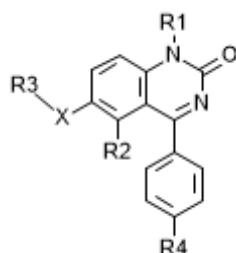
CHEMBL-ID	R1	R2	(*)	IC50 [nM]
CHEMBL486278	4-methoxybenzyl	2-(3H)-oxobenzo[D]imidazol-4-yl	S	44074
CHEMBL572163	2-naphthylmethyl	4-(3-carboxypropanyl)phenyl	R	44047
CHEMBL570593	2-naphthylmethyl	2-cyano-4-(3-carboxypropanyl)phenyl	R	0.2
CHEMBL180672	2-naphthylmethyl	3-chloro-2-cyanophenyl	R	0.058
CHEMBL1204009	2-naphthylmethyl	2-cyano-4-[3-(ethylcarboxypropyl)]phenyl	R	0.034
CHEMBL382741	2-naphthylmethyl	2-cyanophenyl	R	0.003



ChEMBL382741

Figure 4.13 CHEMBL382741, interaction map of the most potent amino alcohol-based parathyroid hormone secretagogues (IC50 = 0.003 nM) docked to CaSRs homology model on page 53

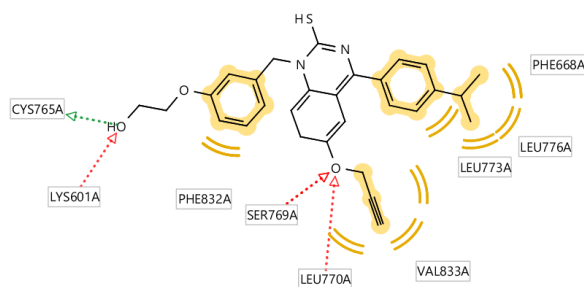
Table 4.12: 1-Alkyl-4-phenyl-6-alkoxyquinazolin-2(1H)-one; IC50 determined in CCL39 fibroblasts using FLIRP assays to detect the inhibition of intracellular calcium release (4)



CHEMBL-ID	R1	R2	R3	R4	X	IC50 [nM]
CHEMBL598638	Benzyl	allyl	H	2-propyl	O	44077
CHEMBL598436	3-methylbenzyl	H	2-propynyl	2-propyl	O	44044
CHEMBL604506	3-methoxybenzyl	H	2-propynyl	2-propyl	O	44014

CHEMBL603248	4-carboxy-methoxypiperbenzyl	H	2-propynyl	2-propyl	O	43994
CHEMBL597800	3-ethoxy-4-methoxybenzyl	H	2-propynyl	2-propyl	O	43922
CHEMBL598016	benzyl	propyl	H	2-propyl	O	43988
CHEMBL591240	4-methylbenzyl	H	2-propynyl	2-propyl	O	43983
CHEMBL597208	benzyl	cyclopropyl-methyl	H	2-propyl	O	43953
CHEMBL591475	Benzyl	H	2-propynyl	2-propyl	O	43925
CHEMBL597194	3-carboxy-methoxybenzyl	H	2-propynyl	2-propyl	O	43924
CHEMBL590539	2-(4-methylpiperazin-1-yl)-N-3-(methylene-phenyl)acetamide	H	2-propynyl	2-propyl	O	43923
CHEMBL597396	3-(2-hydroxy-ethoxy)benzyl	H	2-propynyl	2-propyl	O	43897
CHEMBL597400	4-ethylbenzyl	H	2-propynyl	2-propyl	O	43865
CHEMBL597401	4-methoxybenzyl	H	2-propynyl	2-propyl	O	43865
CHEMBL590560	2-[4-(3-dimethylaminopropyl)-piperazin-1-yl]-N-3-(methylene-phenyl)-acetamide	H	2-propynyl	2-propyl	O	43838
CHEMBL597011	2-propyl	H	H	2-propyl	O	15900
CHEMBL604302	2-propyl	H	benzyl	2-propyl	N	10000
CHEMBL599229	2-propyl	H	3-butenyl	2-propyl	O	6600
CHEMBL599869	2-propyl	H	methyl	phenyl	O	5500
CHEMBL599227	2-propyl	H	butyl	2-propyl	O	3700
CHEMBL599665	2-propyl	H	methyl	cyclopropyl	O	3300
CHEMBL599039	2-propyl	H	N-ethyl;N-ethyl	2-propyl	N	2900
CHEMBL599025	2-propyl	H	2-propyl	2-propyl	O	2700
CHEMBL608721	2-propyl	H	Ethyl	2-propyl	O	2200
CHEMBL599664	2-propyl	H	methyl	Ethyl	O	2100
CHEMBL599868	2-propyl	H	methyl	t-butyl	O	2100
CHEMBL597206	2-propyl	H	methyl	2-propyl	O	1400
CHEMBL599467	2-propyl	H	Propyl	2-propyl	O	1100
CHEMBL599858	2-propyl	H	Propyl	2-propyl	N	1100
CHEMBL599857	2-propyl	H	N-ethyl;N-ethyl	2-propyl	N	520
CHEMBL589977	2-propyl	H	Ethyl	2-propyl	N	460
CHEMBL610511	2-propyl	H	Allyl	2-propyl	N	230
CHEMBL599228	2-propyl	H	Allyl	2-propyl	O	180

CHEMBL591241	2-propyl	H	2-butynyl	2-propyl	O	170
CHEMBL597812	2-propyl	H	2-propynyl; 2-propynyl	2-propyl	N	120
CHEMBL599040	2-propyl	H	Allyl	2-propyl	N	70
CHEMBL597811	2-propyl	H	2-propynyl	2-propyl	N	30
CHEMBL604713	4-(2-propyl)benzyl	H	2-propynyl	2-propyl	O	27
CHEMBL596789	4-ethoxybenzyl	H	2-propynyl	2-propyl	O	19
CHEMBL590552	2-[4-(2-methoxyethyl)-piperazin-1-yl]-N-3-(methylene-phenyl)-acetamide	H	2-propynyl	2-propyl	O	13
CHEMBL598637	3-ethylbenzyl	H	2-propynyl	2-propyl	O	13
CHEMBL592176	2-propyl	H	2-propynyl	2-propyl	O	9
CHEMBL597397 [1H-quinazolin-2-thione]	2-hydroxyethoxy	H	2-propynyl	2-propyl	O	0.4



CHEMBL597397

Figure 4.14 CHEMBL597397, interaction map of the most potent 1-Alkyl-4-phenyl-6-alkoxyquinazolin-2(1H)-one ($IC_{50} = 0.4$ nM) docked to CaSRs homology model on page 53

Table 4.13: Parathyroid hormone mimetics; IC50 determined in HEK-293 cells using FLIRP assays to detect the inhibition of intracellular calcium release (119)

CHEMBL-ID	IC50 [nM]
CHEMBL1084514	520
CHEMBL180030	300
CHEMBL1083694	97
CHEMBL1084513	73
CHEMBL2368751	25

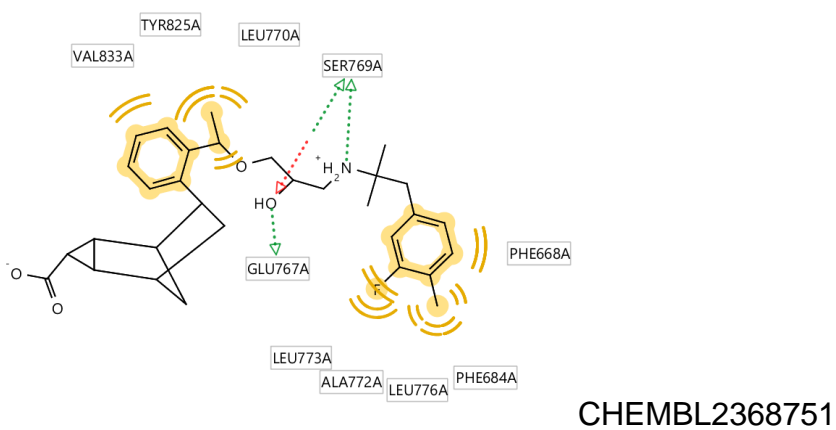
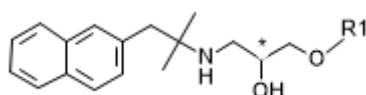
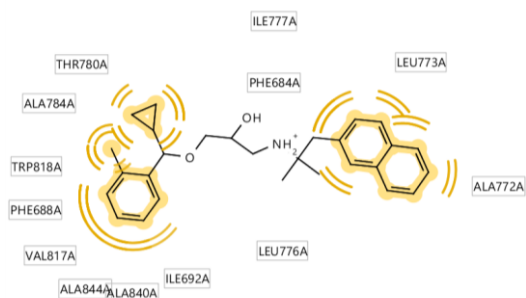


Figure 4.15 CHEMBL2368751, interaction map of the most potent parathyroid hormone mimetics (IC50 = 25 nM) docked to CaSRs homology model on page 53

Table 4.14: Aminopropanediol derivatives; IC50 values determines in PC12h cells (123)



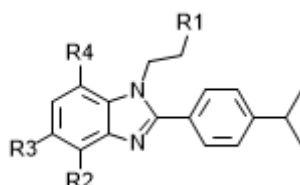
CHEMBL-ID	R1	(*)	IC50 [nM]
CHEMBL1083949	2-phenyl-2-propyl	R	2240
CHEMBL1085987	2-phenyl-1-propyl	R	810
CHEMBL1085335	1-(S)-cyclopropyl-(2-methylphenyl)methyl	S	760
CHEMBL1083306	Benzhydryl	R	730
CHEMBL1084201	1-cyclohexyl(phenyl)methyl	R	610
CHEMBL1082719	2,2-dimethyl-1-phenylpropyl	R	570
CHEMBL1084756	Benzyl	R	560
CHEMBL1082379	1-naphthylethyl	R	550
CHEMBL1083314	1-(2-nitrophenyl)ethyl	R	540
CHEMBL1083950	1-phenylprop-2-yl	R	440
CHEMBL1086472	Pyridin-2-yl-ethyl	R	420
CHEMBL1082710	1-naphthylethyl	R	360
CHEMBL1085334	1-(R)-cyclopropyl-(2-methylphenyl)methyl	S	310
CHEMBL1083951	3-phenylbutyl	R	300
CHEMBL1083308	1-(3-chlorophenyl)ethyl	R	220
CHEMBL1084757	1-phenylethyl	R	220
CHEMBL1082712	1-phenyl-2-propenyl	R	210
CHEMBL1083610	1-(2-methylphenyl)ethyl	R	200
CHEMBL1082717	1-phenyl-2-methylpropyl	R	200
CHEMBL1084200	1-cyclopentyl-1-phenylmethyl	R	200
CHEMBL1082998	1-(3-methylphenyl)ethyl	R	170
CHEMBL1082718	1-phenylpentyl	R	170
CHEMBL1082787	1-(2-cyanophenyl)ethyl	R	160
CHEMBL1082306	1-phenylbutyl	R	150
CHEMBL1082999	1-(4-methylphenyl)ethyl	R	140
CHEMBL1083307	1-(2-chlorophenyl)ethyl	R	140
CHEMBL1083313	1-(2-methoxyphenyl)ethyl	R	120
CHEMBL1082711	1-phenylpropyl	R	92
CHEMBL1085807	(S)-cyclopropyl-(2-methylphenyl)methyl	R	83
CHEMBL1083289	cyclobutylphenylmethyl	R	74
CHEMBL1083288	Cyclopropylphenylmethyl	R	70
CHEMBL1086281	Cyclopropyl-(2-methoxyphenyl)methyl	R	41
CHEMBL1082788	Cyclopropyl-(2-methylphenyl)methyl	R	37



CHEMBL1082788

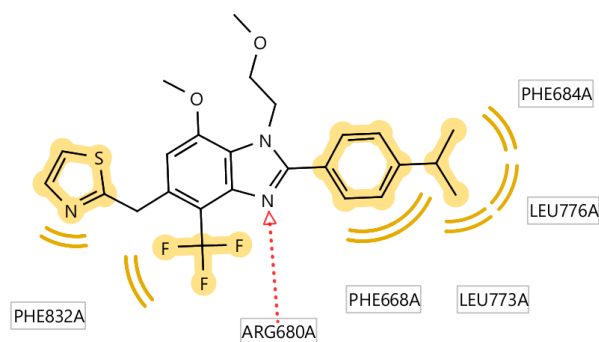
Figure 4.16 CHEMBL1082788, interaction map of the most potent aminopropanediol derivatives; ($IC_{50} = 37 \text{ nM}$) docked to CaSRs homology model on page 53

Table 4.15: penta-substituted benzimidazoles; IC_{50} determined in hamster fibroblasts using FLIRP assays to detect the inhibition of intracellular calcium release (124)



CHEMBL-ID	R1	R2	R3	R4	IC_{50} [nM]
CHEMBL1223771	Methoxy	CF_3	2-methyl-sulfinylpyridin-3-ylmethyl	Methoxy	43984
CHEMBL1223714	Methoxy	CF_3	2-methyl-mercaptopyridin-3-ylmethyl	Methoxy	43983
CHEMBL1223299	Methoxy	H	H	H	17000
CHEMBL1223298	Ethyl	Methoxy	H	H	4690
CHEMBL1223221	Ethyl	2-propyl	H	H	2550
CHEMBL1223301	Methoxy	H	H	Methyl	2100
CHEMBL1223222	Methoxy	2-propyl	H	H	1950
CHEMBL1223302	Methoxy	Methoxy	H	Methyl	1900
CHEMBL1223300	Methoxy	H	H	Methoxy	1900
CHEMBL1223370	Methoxy	Methoxy	H	Methoxy	1820
CHEMBL1223437	Methoxy	Br	H	Br	1000
CHEMBL1223440	Methyl	Methyl	H	Methoxy	790
CHEMBL1223371	Methoxy	Methoxy	H	Cl	770
CHEMBL1222425	Methoxy	H	CF_3	Methoxy	360
CHEMBL1222491	Methoxy	Br	3,4-dimethyl-phenylmethyl	Methoxy	340
CHEMBL1223439	Methoxy	Methyl	H	Methoxy	210
CHEMBL1222424	Methoxy	Br	Ethyl	Methoxy	190
CHEMBL1223438	Dimethylamino	BR	H	Methoxy	190
CHEMBL1222426	Methoxy	Br	Phenyl	Methoxy	60
CHEMBL1223369	Methoxy	CF_3	H	Methoxy	51
CHEMBL1223303	Methoxy	Cl	H	Methoxy	50

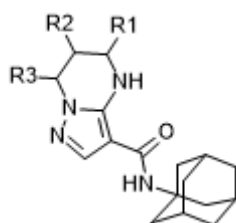
CHEMBL1222490	Methoxy	Br	3,4-dimethoxy-benzyl	Methoxy	43
CHEMBL1222489	Methoxy	CF ₃	2-methoxy-benzyl	Methoxy	26
CHEMBL1223509	Methoxy	Br	Cyano	Methoxy	25
CHEMBL1223367	Methoxy	Br	H	Methoxy	22
CHEMBL1223508	Methoxy	CF ₃	BR	Methoxy	16
CHEMBL1223711	Methoxy	Br	Pyridin-3-ylmethyl	Methoxy	15
CHEMBL1223505	Methoxy	Br	Br	Methoxy	14
CHEMBL1223368	Methoxy	I	H	Methoxy	13
CHEMBL1223507	Methoxy	Br	CF ₃	Methoxy	11
CHEMBL1223506	Methoxy	I	Br	Methoxy	6
CHEMBL1222492	Methoxy	Br	2-methylsulfinyl-benzyl	Methoxy	5
CHEMBL1222427	Methoxy	Br	Benzyl	Methoxy	5
CHEMBL1223712	Methoxy	CF ₃	2-methoxy-pyridin-3yl-methyl	Methoxy	4
CHEMBL1223713	Methoxy	Br	2-methyl-mercaptopyridin-3-ylmethyl	Methoxy	4
CHEMBL1223772	Methoxy	CF ₃	1,3-thiazol-2-ylmethyl	Methoxy	2



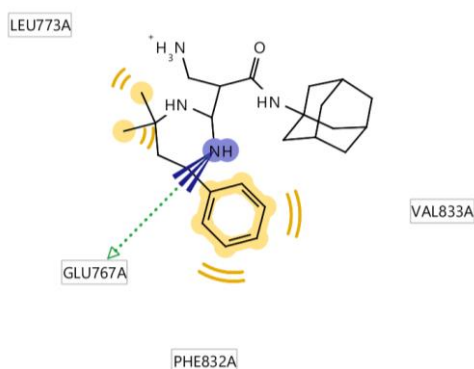
CHEMBL1223772

Figure 4.17 CHEMBL1223772, interaction map of the most potent penta-substituted benzimidazoles ($IC_{50} = 2$ nM) docked to CaSRs homology model on page 53

Table 4.16: tetrahydropyrazolopyrimidine derivatives; IC50 measured with a GTP-binding assay (125)



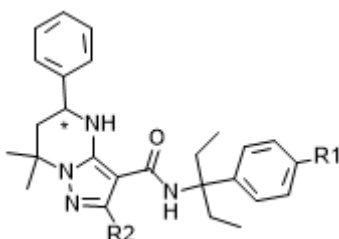
CHEMBL-ID	R1	R2	R3	IC50 [nM]
CHEMBL1290216	(S)-phenyl;H	H;H	H; H	53000
CHEMBL1290007	(S)-phenyl;H	H;H	(R)-phenyl	700
CHEMBL1290329	Phenyl	H	Dimethyl	580
CHEMBL1289892	2-furanyl;H	H;H	(R)-CF ₃	400
CHEMBL1289679	(S)-3-Cl-phenyl;H	H;H	(R)-CF ₃	320
CHEMBL1289573	(S)-2-Cl-phenyl;H	H;H	(R)-CF ₃	110
CHEMBL1289346	(S)-phenyl;H	H;H	(R)-CF ₃	97
CHEMBL1290106	(S)-phenyl;H	H;H	(S)-methyl	76
CHEMBL1290328	(S)-phenyl;H	H; H	dimethyl	10



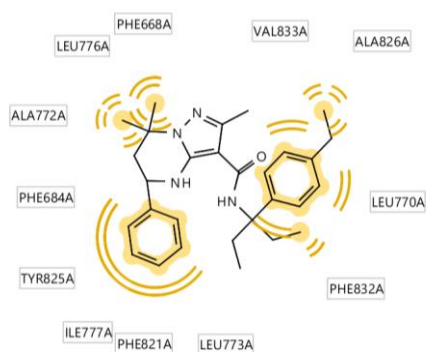
CHEMBL1290328

Figure 4.18 CHEMBL1290328, interaction map of the most potent tetrahydropyrazolopyrimidine derivatives (IC₅₀ = 10 nM) docked to CaSRs homology model on page 53

Table 4.17: tetrahydropyrazolopyrimidine; IC50 measured with a GTP-binding assay (126)



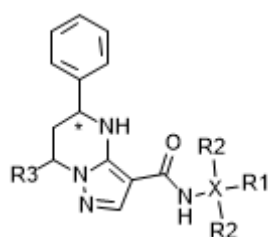
CHEMBL-ID	R1	R2	(*)	IC50 [nM]
CHEMBL1689062	Acetyl	methyl	S	44047
CHEMBL1689041	Methyl	Ethyl	R; S	44018
CHEMBL1689046	Methyl	methylmercapto	R;S	44016
CHEMBL1689057	Methoxy	Methyl	S	44013
CHEMBL1689051	2-propyl	Methyl	S	44013
CHEMBL1689058	Methoxy	Methyl	S	43986
CHEMBL1689047	Methyl	Methyl	S	43923
CHEMBL1689049	H	Methyl	S	43892
CHEMBL1688099	Methyl	H	S	43892
CHEMBL1689055	H	Methyl	S	43863
CHEMBL1689050	Ethyl	Methyl	S	43865
CHEMBL1687956	CL	Methyl	S	43836
CHEMBL1689040	Methyl	Methyl	R;S	43834
CHEMBL1689061	Carboxy	Methyl	S	320
CHEMBL1689044	Methyl	Phenyl	R;S	81
CHEMBL1689045	Methyl	Benzyl	R;S	45
CHEMBL1689048	Methyl	Methyl	R	31
CHEMBL1689107	1-hydroxyethyl	Methyl	S	13
CHEMBL1689043	Methyl	Butyl	R;S	11
CHEMBL1689056	F	Methyl	S	11
CHEMBL1689054	Methoxy	Methyl	S	11
CHEMBL1689042	Methyl	propyl	R;S	6
CHEMBL1689053	Ethoxy	Methyl	S	3
CHEMBL1689052	Methoxy	Methyl	S	2
CHEMBL1689060	Ethyl	methyl	S	0.94



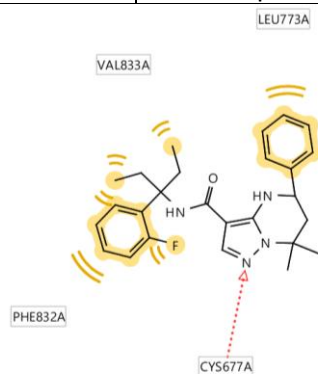
CHEMBL1689060

Figure 4.19 CHEMBL 1689060, interaction map of the most potent tetrahydropyrazolopyrimidine (IC50 = 0.94 nM) docked to CaSRs homology model on page 53

Table 4.18: Short-acting oral calcilytics; IC50 determined in PC12h cells (127)



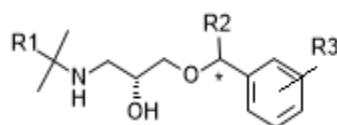
CHEMBL-ID	R1	R2	R3	X	(*)	IC50 [nM]
CHEMBL1688097	4-methylphenyl	ethyl	dimethyl	C	R,S	44078
CHEMBL1689810	4-CF ₃ -phenyl	ethyl	dimethyl	C	R,S	73989
CHEMBL1688087	4-methoxyphenyl	ethyl	dimethyl	C	R,S	43988
CHEMBL1688088	4-benzyloxyphenyl	ethyl	dimethyl	C	R,S	43959
CHEMBL1687946	4-methylphenyl	ethyl	dimethyl	C	R,S	43956
CHEMBL1689809	4-CF ₃ -phenyl	methyl	dimethyl	C	R,S	43954
CHEMBL1689816	4-Cl-phenyl	ethyl	dimethyl	C	R,S	43928
CHEMBL1689811	4-CF ₃ -phenyl	propyl	dimethyl	C	R,S	43925
CHEMBL1689815	Phenyl	ethyl	dimethyl	C	R,S	43924
CHEMBL1689818	4-F-phenyl	ethyl	dimethyl	C	R,S	43897
CHEMBL1688103	Ethyl-4-benzoate	ethyl	dimethyl	C	R,S	43866
CHEMBL1688101	4-CF ₃ -phenyl	H	(R)- CF ₃	C	S	2900
CHEMBL1688102	H	phenyl	(R)- CF ₃	C	S	380
CHEMBL1689814	4-CF ₃ -phenyl	cyclohexyl	dimethyl	C	R,S	150
CHEMBL1289345	-	-	(R)- CF ₃	adamantyl	S	40
CHEMBL1688091	4-(phenylacetate)	ethyl	dimethyl	C	R,S	31
CHEMBL1688100	4-methylphenyl	ethyl	dimethyl	C	S	31
CHEMBL1689813	4-CF ₃ -phenyl	2-propyl	dimethyl	C	R,S	30
CHEMBL1688085	3-F-phenyl	ethyl	dimethyl	C	R,S	25
CHEMBL1689812	4-CF ₃ -phenyl	butyl	dimethyl	C	R,S	23
CHEMBL1689817	3-Cl-phenyl	ethyl	dimethyl	C	R,S	17
CHEMBL1688090	4-(benzylacetate)	ethyl	dimethyl	C	R,S	14
CHEMBL1688092	4-(phenylpropionate)	ethyl	dimethyl	C	R,S	12
CHEMBL1688086	2-F-phenyl	ethyl	dimethyl	C	R,S	12



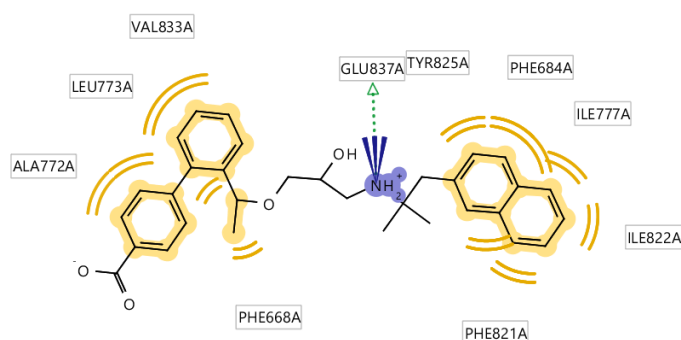
CHEMBL1688086

Figure 4.20 CHEMBL1688086, interaction map of the most potent short-acting oral calcilytics (IC50 = 12 nM) docked to CaSRs homology model on page 53

Table 4.19: Aminopropanediol derivatives; IC50 determined in PC12h cells (127)



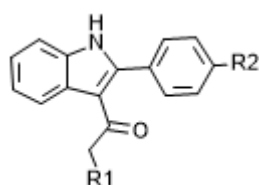
CHEMBL-ID	R1	R2	R3	(*)	IC50 [nM]
CHEMBL1672965	2-naphthylmethyl	methyl	2-carboxymethyl	R	2200
CHEMBL1672967	2-naphthylmethyl	methyl	4-(4-benzoic acid)		1500
CHEMBL1672962	2-naphthylmethyl	cyclopropyl	3-carboxy		960
CHEMBL1672966	2-naphthylmethyl	methyl	2-(2-carboxy-ethyl)	R	24
CHEMBL1086282	2-naphthylmethyl	cyclopropyl	2-methyl	R	23
CHEMBL1672971	3-F-4-Cl-phenyl	methyl	2-(4-benzoic acid)	R	23
CHEMBL1672960	2-naphthylmethyl	cyclopropyl	Hydroxymethyl	R	18
CHEMBL1672968	2-naphthylmethyl	methyl	3-(4-benzoic acid)		18
CHEMBL1672973	3-F-4-Cl-phenyl	methyl	2-(2-methyl-4-benzoic acid)	R	12
CHEMBL1672969	2-naphthylmethyl	methyl	2-(4-benzoic acid)		11
CHEMBL1672972	2-naphthylmethyl	methyl	2-(2-methyl-4-benzoic acid)	R	5
CHEMBL1672970	2-naphthylmethyl	methyl	2-(4-benzoic acid)	R	4



CHEMBL1672970

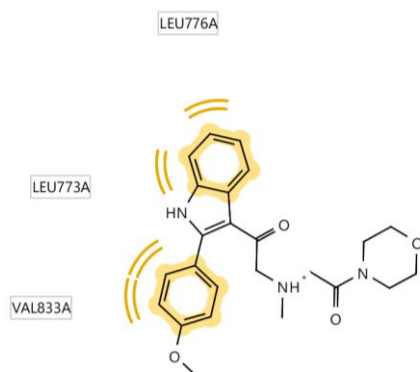
Figure 4.21 CHEMBL1672970, interaction map of the most potent aminopropanediol derivatives (IC50 = 4 nM) docked to CaSRs homology model on page 53

Table 4.20: 2-phenylindole; IC50 determined with a FRET-based inositol monophosphate assay (128)



CHEMBL-ID	R1	R2	IC50 [nM]
CHEMBL3633665	N-methyl,N-2-morpholin-4-ylethyl	H	347000

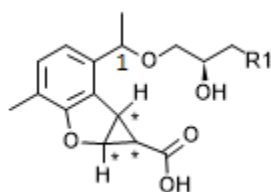
CHEMBL3633650	N-methyl,N-2-(morpholin-4yl)-2-oxoethyl	H	316000
CHEMBL3633663	3-(morpholin-4-yl-carbonyl)piperidin-1-yl	H	240000
CHEMBL3633656	N-methyl,N-2-(morpholin-4yl)-2-oxoethyl	methoxy	56200



CHEMBL3633656

Figure 4.22 CHEMBL3633656, interaction map of the most potent 2-phenylindole ($IC_{50} = 56200$ nM) docked to CaSRs homology model on page 53

Table 4.21: dihydro benzofuran cyclopropane derivatives; IC_{50} measured in cell-based FLIPR assays (129)



CHEBML-ID	R1	(*)	1	IC_{50} [nM]
CHEMBL3827475	[1-(2,3-dihydro-1,4-benzodioxin-6yl)-2-methylpropan-2-yl]amino	1S,1As,6bR	R	44064
CHEMBL3827570	[1-(3,4-dimethylphenyl)-2-methylpropan-2-yl]amino	1S,1As,6bR	R	44029
CHEMBL3827041	[1-(3-F-4-methylmercaptophenyl)-2-methylpropan-2-yl]amino	1S,1As,6bR	R	44009
CHEMBL3828483	[1-(2-naphthyl)-2-methylpropan-2-yl]amino	1R,1Ar,6bS	R	44003
CHEMBL3828187	[1-(2-Cl-4-methoxyphenyl)-2-methylpropan-2-yl]amino	1S,1As,6bR	R	44001
CHEMBL3827830	(2-methyl-5-phenylpentan-2-yl)amino	1S,1As,6bR	R	43949
CHEMBL3827304	[1-(1-benzothiophen-5-yl)-2-methylpropan-2-yl]amino	1S,1As,6bR	R	22890
CHEMBL3827259	[1-(2-naphthyl)-2-methylpropan-2-yl]amino	1S,1As,6bR	R	16619
CHEMBL3828329	2-benzyl-2-methylpyrrolidin-1-yl	1S,1As,6bR	R	1332

CHEMBL3827473	2-benzylpyrrolidin-1-yl	1S,1As,6bR	R	489.1
CHEMBL3827410	[1-(quinoline-6-yl)-2-methylpropan-2-yl]amino	1S,1As,6bR	R	457.1
CHEMBL3828182	[1-(4-Cl-3-F-phenyl)-2-methylpropan-2-yl]amino			202.7
CHEMBL3827912	[1-(2,3-dihydro-1H-inden-2-yl)-2-methylpropan-2-yl]amino			117.3
CHEMBL3827356	[1-(4-Cl-3-F-phenyl)-2-methylpropan-2-yl]amino	1R,1Ar,6bS	R	77.2
CHEMBL3827723	[1-(2-naphthyl)-2-methylpropan-2-yl]amino			64.3
CHEMBL3827154	[1-(4-Cl-3-F-phenyl)-2-methylpropan-2-yl]amino	1S,1As,6bR	R	58.7
CHEMBL3827577	(2-methyl-4-phenylsulfanylbutan-2-yl)amino	1S,1As,6bR	R	50.8
CHEMBL3828154	[1-(2,3-dihydro-1H-inden-2-yl)-2-methylpropan-2-yl]amino	1R,1Ar,6bS	R	45.7
CHEMBL3827209	[1-(2-methyl-(4-methylsulfanylphenyl)propan-2-yl]amino	1S,1As,6bR	R	39.3
CHEMBL3827736	[1-(2,3-dihydro-1H-inden-2-yl)-2-methylpropan-2-yl]amino	1S,1As,6bR	R	20

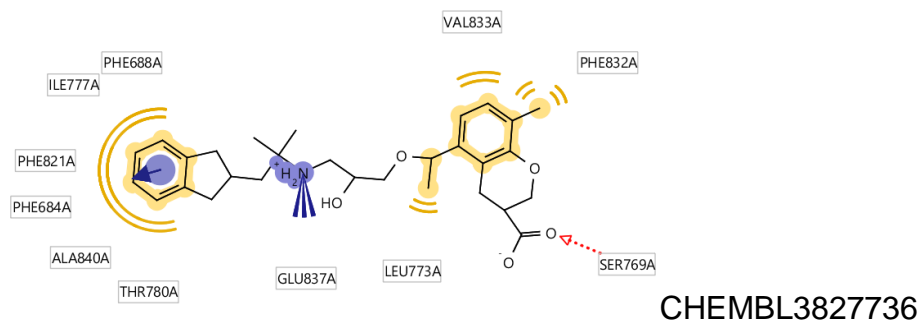
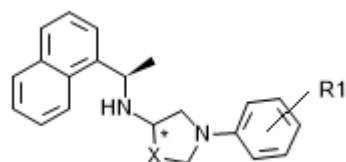


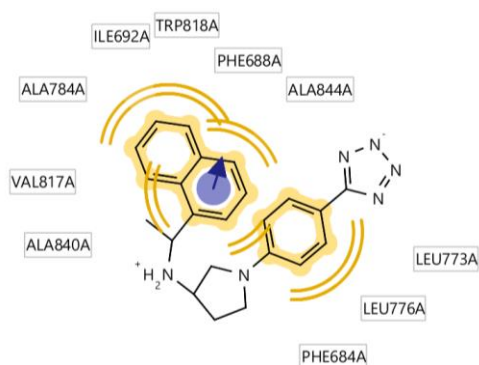
Figure 4.23 CHEMBL3827736, interaction map of the most potent dihydro benzofuran cyclopropane derivatives ($IC_{50} = 20$ nM) docked to CaSRs homology model on page 53

Table 4.22: evocalcet derivatives; IC_{50} measured in human liver microsomes (130)



CHEBML-ID	R1	X	(*)	IC_{50} [nM]
CHEMBL4130093	3-(trifluoromethoxy)	C_2H_4	S	44074
CHEMBL4128835	4-carboxymethyl	CH_2	S	44073
CHEMBL4129011	4-(trifluoromethoxy)	CH_2	S	44050
CHEMBL4125724	3-(trifluoromethoxy)	CH_2	S	44023

CHEMBL4125688	3-CF ₃	CH ₂	R	44021
CHEMBL4128542	3-(trifluoromethoxy)	C ₂ H ₄	R	43922
CHEMBL4126057	3-(trifluoromethoxy)	CH ₂	R	43894
CHEMBL4127153	3-CF ₃	CH ₂	S	43843
CHEMBL4125917	3-CF ₃	C ₂ H ₄	R	43832
CHEMBL4129371	4-(trifluoromethoxy)	CH ₂	R	116
CHEMBL4126482	4-carboxy	CH ₂	S	8
CHEMBL4126877	4-(N-2-hydroxyethylamide)	CH ₂	S	0.91
CHEMBL4126450	4-(2H-tetrazol-5-yl)	CH ₂	S	0.47



CHEMBL4126450

Figure 4.24 CHEMBL4126450, interaction map of the most potent evocalcet derivatives ($IC_{50} = 0,47$ nM) docked to CaSRs homology model on page 53

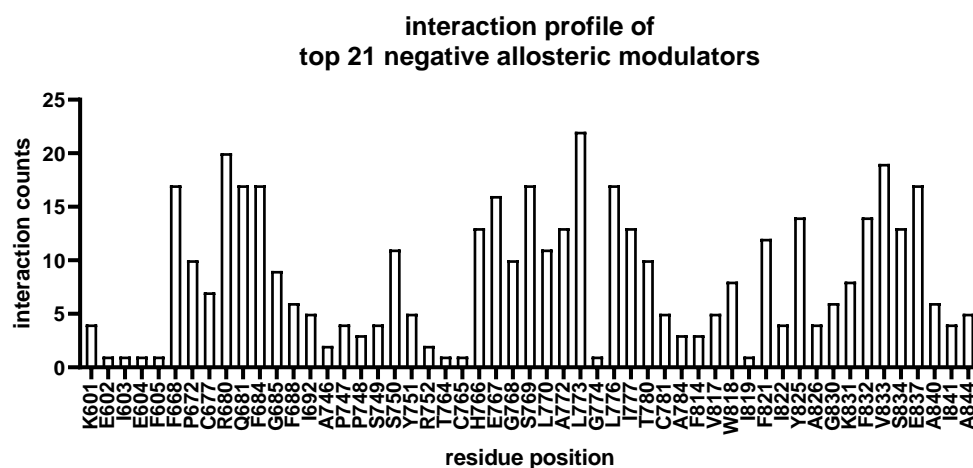
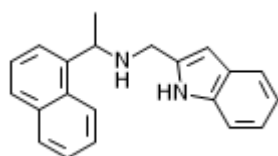


Figure 4.25 histogram of residue interactions of CaSR transmembrane with the top-scoring negative allosteric modulators from 21 different scaffolds, docking done with AutoDock VINA, model was used from page 46, interaction profile (CHEMBL180463; CHEMBL195711; CHEMBL195711; CHEMBL195711; CHEMBL200312; CHEMBL536975; CHEMBL457249; CHEMBL1783764; CHEMBL2112075; CHEMBL382741; CHEMBL597397; CHEMBL2368751; CHEMBL1082788; CHEMBL1223772M; CHEMBL1290328; CHEMBL1689060; CHEMBL1688086; CHEMBL1672970; CHEMBL3633656; CHEMBL3827736; CHEMBL4126450)

4.2.2 Positive allosteric modulators for the human extracellular calcium-sensing receptor

Table 4.23: calindol; IC₅₀ measured in CHO cells with myo-[³H]inositol (131)

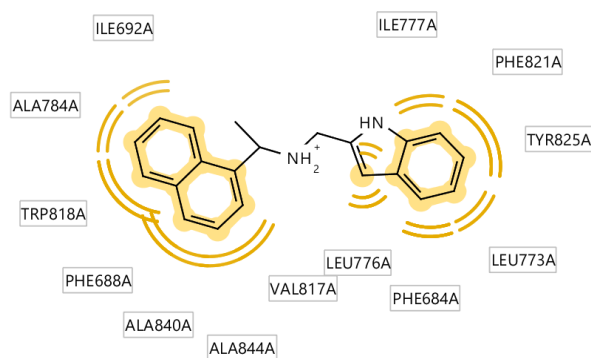
(R)- calindol



CHEMBL2092942, EC₅₀= 310 nM, HEK cells, human, (R)

EC₅₀= 1000 nM, CHO, rat

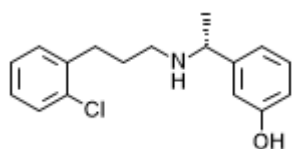
(R,S) calindol CHEMBL543875, EC₅₀= 1300 nM, HEK cells, human, (R,S)



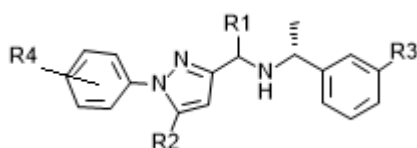
CHEMBL2092942

Figure 4.26 CHEMBL2092942, interaction map of (R) calindol, IC₅₀ = 310 nM) docked to CaSRs homology model on page 53

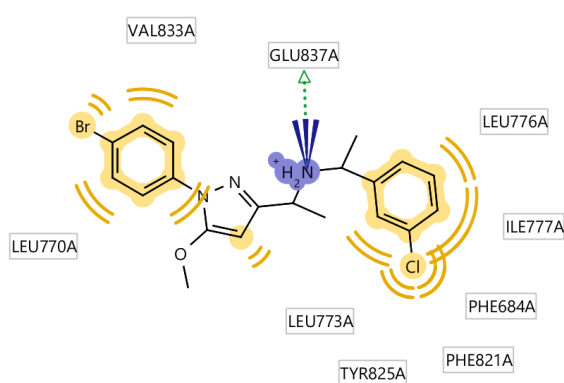
Table 4.24: 1-(1-Phenyl-1H-pyrazol-3-yl)methanamines; EC₅₀ measured in HEK293 cells (132)



CHEMBL292376; EC₅₀ = 194 nM



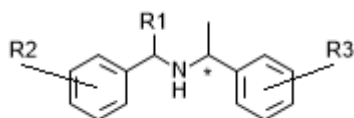
CHEMBL-ID	R1	R2	R3	R4	EC50 [nM]
CHEMBL572362	(R)-propan-2-yl	H	H	3-CF ₃	2340
CHEMBL583469	(R)-methyl	methoxy	Cl	4-carboxyl	1920
CHEMBL571048	H	H	H	H	1580
CHEMBL577306	(R)-ethyl	H	H	3-CF ₃	1180
CHEMBL583522	(R)-methyl	H	H	3-CF ₃	233
CHEMBL572363	(R)-methyl	H	F	3-CF ₃	148
CHEMBL568485	(R)-methyl	H	Cl	3-CF ₃	92
CHEMBL569863	(R)-methyl	methoxy	H	3-CF ₃	70
CHEMBL577333	(R)-methyl	methoxy	Cl	4-Cl	57
CHEMBL571476	(R)-methyl	methoxy	Cl	3-CF ₃	41
CHEMBL569182	(R)-methyl	methoxy	Cl	Br	23



ChEMBL569182

Figure 4.27 CHEMBL569182, interaction map of the most potent 1-(1-Phenyl-1H-pyrazol-3-yl)methanamines (IC₅₀ = 23 nM) docked to CaSRs homology model on page 53

Table 4.25: Positive allosteric modulators with a dibenzylamine core; EC₅₀ determined in a FLIPR assay (133)



CHEMBL-ID	R1	R2	R3	(*)	EC50 [nM]
CHEMBL1224191	H	2-Cl	3-methoxy	R	3700
CHEMBL1224425	(S)-methyl	3-(4-CF ₃ -phenyl)-4-methoxy	H	R	3000
CHEMBL1224260	H	3-(2-F-phenyl)	3-methoxy	R	2700
CHEMBL1224261	H	3-(2-methylphenyl)	3-methoxy	R	2200
CHEMBL1224192	H	3-(2-Cl-phenyl)	3-methoxy	R	1700
CHEMBL1224423	H	3-(4-CF ₃ -phenyl)-4-methoxy	4-F	R	1700
CHEMBL1224262	H	2-methoxy	3-methoxy	R	1400
CHEMBL1224342	H	3-(4-CF ₃ -phenyl)-4-methoxy	3-F	R	1300

CHEMBL4224190	H	H	3-methoxy	R	1200
CHEMBL1224263	H	3-(4-methoxyphenyl)	3-methoxy	R	870
CHEMBL1224264	H	3-(4-CF ₃ -phenyl)	3-methoxy	R	620
CHEMBL1224193	H	4-methoxy	3-methoxy	R	370
CHEMBL1224344	H	3-(4-CF ₃ -phenyl)-4-Cl	3-methoxy	R	280
CHEMBL1224343	H	3-(4-CF ₃ -phenyl)-4-methoxy	3-methyl	R,S	160
CHEMBL1224426	H	3-(4-CF ₃ -phenyl)-4-(CF ₃ -oxy)	H	R	110
CHEMBL1224341	H	3-(4-CF ₃ -phenyl)-4-methoxy	H	R	27
CHEMBL1224340	H	3-(4-CF ₃ -phenyl)-4-methoxy	3-methoxy	R	23
CHEMBL1224424	(R)-methyl	3-(4-CF ₃ -phenyl)-4-methoxy	H	R	17

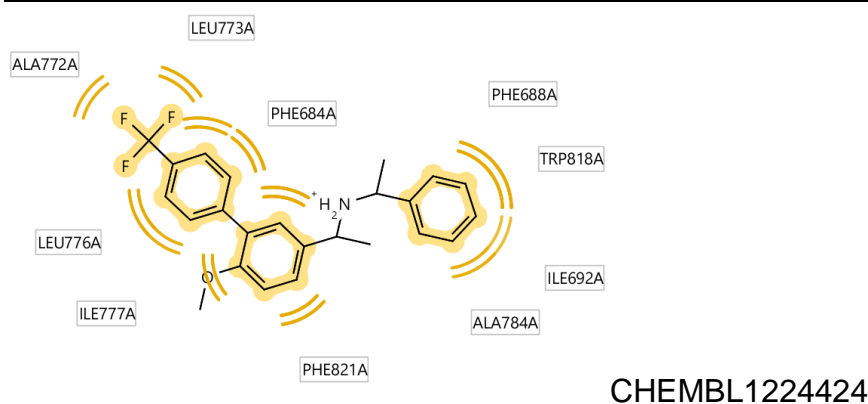
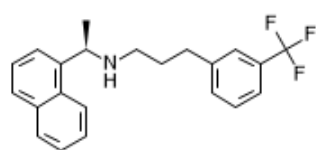


Figure 4.28 CHEMBL1224424, interaction map of the most potent dibenzylamine ($IC_{50} = 17$ nM) docked to CaSRs homology model on page 53

Table 4.26: cinacalcet; EC_{50} measured in a functional cell-based assay (134)



CHEMBL1200776; $EC_{50} = 19.95$ nM; cinacalcet

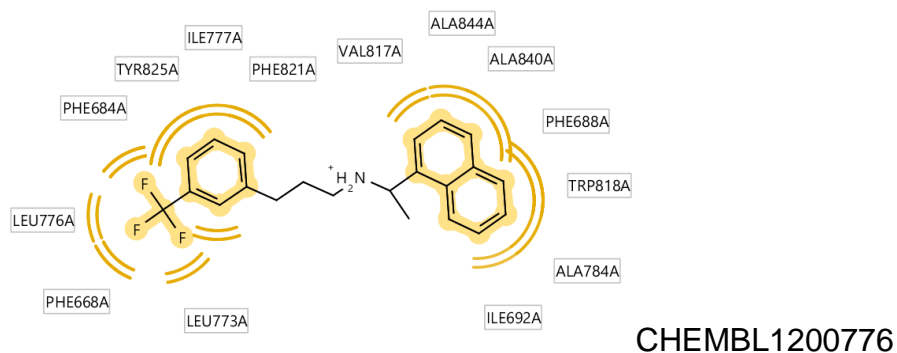
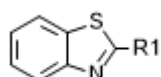


Figure 4.29 CHEMBL1200776, interaction map of cinacalcet ($IC_{50} = 19.95$ nM) docked to CaSRs homology model on page 53

Table 4.27: 1-(benzothiazole-2-yl)1-phenylethanol; EC50 measured in a functional cell-based assay (134)



CHEMBL-ID	R1	EC50 [nM]
CHEMBL1256367	1-hydroxy-1-(2,4-dimethylphenyl)ethyl	44190
CHEMBL1256330	1-hydroxy-1-(2-methoxyphenyl)ethyl	3981.07
CHEMBL1256486	1-(2,4-dimethylphenyl)ethyl	2511.89
CHEMBL1258390	1-hydroxy-(2-naphthyl)ethyl	630.96
CHEMBL1256301	1-hydroxy-1-(4-Cl-phenyl)	398.11
CHEMBL1255602	1-(1-hydroxy-5,7-dimethyl-1,2,3,4-tetrahydronaphthyl)	316.23
CHEMBL1256331	1-hydroxy-1-(2-methylphenyl)ethyl	251.19
CHEMBL1258389	1-hydroxy-1-(2,4-dimethylphenyl)-2-cyclopropylethyl	199.53
CHEMBL1256366	1-hydroxy-1-(4-F-3-methylphenyl)ethyl	158.49
CHEMBL1256303	1-hydroxy-1-(3,4-di-Cl-phenyl)ethyl	100
CHEMBL1256404	1-hydroxy-1-(2,4-di-Cl-phenyl)ethyl	100
CHEMBL1256402	(S)-1-hydroxy-1-(2,4-dimethylphenyl)ethyl	79.43
CHEMBL1256403	1-hydroxy-1-(2-methyl-4-Cl-phenyl)ethyl	31.62
CHEMBL1256368	(R)-1-hydroxy-1-(2,4-dimethylphenyl)ethyl	10

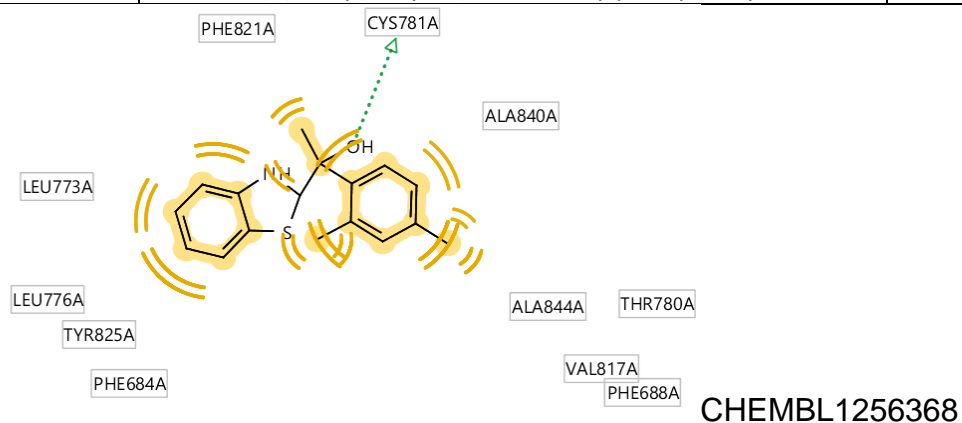
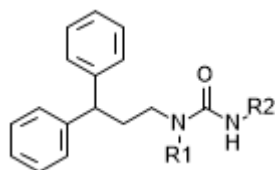


Figure 4.30 CHEMBL1256368, interaction map of the most potent 1-(benzothiazole-2-yl)1-phenylethanol (IC50 = 10 nM) docked to CaSRs homology model on page 53

Table 4.28: Trisubstituted ureas; EC50 determined in CHO cells with a Luciferase assay (135)



CHEMBL-ID	R1	R2	EC50 [nM]
CHEMBL2346778	H	H	20000

CHEMBL2346767	H	H	15000
CHEMBL2346775	H	H	15000
CHEMBL2346779	H	4-methoxybenzyl	15000
CHEMBL2346781	H	2(morpholin-4-yl)ethyl	10000
CHEMBL2346785	2-cylcohexylethyl	3-(methoxycarbonyl)phenyl	10000
CHEMBL2346766	H	3-methoxyphenyl	8000
CHEMBL2346768	H	2-Cl-phenyl	6000
CHEMBL2346666	H	(1R)-1-phenylethyl	6000
CHEMBL2346769	H	4-Cl-phenyl	4000
CHEMBL2346770	H	3-Cl-phenyl	4000
CHEMBL2346773	H	4-methylphenyl	4000
CHEMBL2346774	H	4-phenylphenyl	4000
CHEMBL2346780	H	2-Cl-benzyl	4000
CHEMBL2346772	H	4-(ethoxycarbonyl)phenyl	3000
CHEMBL2346777	H	3-CF ₃ -4-Cl-phenyl	3000
CHEMBL2346771	H	3-(methoxycarbonyl)phenyl	2000
CHEMBL2346776	H	4-Cl-3-(methoxy-carbonyl)phenyl	2000
CHEMBL2346783	benzyl	3-(methoxycarbonyl)phenyl	2000
CHEMBL2349579	2-morpholin-4-ylethyl	3-carboxyphenyl	2000
CHEMBL2346782	2-methylpropyl	3-(methoxycarbonyl)phenyl	1500
CHEMBL2349593	2-morpholin-4-ylethyl	3-methylsulfonylphenyl	1000
CHEMBL2349574	2-morpholin-4-ylethyl	3-(2H-tetrazol-5-ylamino)phenyl	1000
CHEMBL2349582	2-morpholin-4-ylethyl	3-(3,3-dimethylbut-2-yloxy-carbonyl)phenyl	1000
CHEMBL2346784	2-[(2-methylpropan-2yl)oxy]-2-oxoethyl	3-(methoxycarbonyl)phenyl	800
CHEMBL2349588	2-morpholin-4-ylethyl	3-(hydroxymethyl)phenyl	800
CHEMBL2349592	2-morpholin-4-ylethyl	3-methylsulfinylphenyl	800
CHEMBL2349596	2-morpholin-4-ylethyl	3-cyanophenyl	800
CHEMBL2346765	2-morpholin-4-ylethyl	3-(1H-benzimidazol-2-ylamino)phenyl	600
CHEMBL2349587	2-morpholin-4-ylethyl	3-acetylphenyl	550
CHEMBL2349578	3-(4-methoxycarbonylpiperazin-1yl)propan-1-yl	3-(methoxycarbonyl)phenyl	400
CHEMBL2349594	2-morpholin-4-ylethyl	3-(CF ₃ -sulfonyl)phenyl	400
CHEMBL2349597	2-morpholin-4-ylethyl	3-CF ₃ -phenyl	400
CHEMBL2349589	2-morpholin-4-ylethyl	3-(methoxymethyl)phenyl	400
CHEMBL2349585	2-morpholin-4-ylethyl	3-(4,4,5,5,5-pentafluoropent-1-oxycarbonyl)phenyl	400
CHEMBL2349591	2-morpholin-4-ylethyl	3-(CF ₃ -sulfonyl)phenyl	350
CHEMBL2349605	2-morpholin-4-ylethyl	3-(1,3-thiazol-2yl)phenyl	300
CHEMBL2349598	2-morpholin-4-ylethyl	3-methoxyphenyl	300
CHEMBL2349599	2-morpholin-4-ylethyl	3-(dimethylamino)phenyl	300
CHEMBL2349583	2-morpholin-4-ylethyl	3-(2-methylpropan-2yloxy-caronyl)phenyl	300

CHEMBL2349581	2-morpholin-4-ylethyl	3-(phenylmethoxy-carbonyl)phenyl	200
CHEMBL2349576	2-morpholin-4-ylethyl	3-(1-methyltetrazol-5-yl)phenyl	200
CHEMBL2349600	2-morpholin-4-ylethyl	4-phenylphenyl	200
CHEMBL2349601	2-morpholin-4-ylethyl	3-(3-pyridin-3-ylphenyl)	200
CHEMBL2349595	2-morpholin-4-ylethyl	3-(trifluoromethoxy)phenyl	200
CHEMBL2349584	2-morpholin-4-ylethyl	3-(2-hydroxyethoxy-carbonyl)phenyl	150
CHEMBL2349580	2-morpholin-4-ylethyl	3-(propan-2-yloxy-carbonyl)phenyl	150
CHEMBL2349586	2-morpholin-4-ylethyl	3-(3-propanoylphenyl)	150
CHEMBL2346787	2-morpholin-4-ylethyl	3-(methoxycarbonyl)phenyl	150
CHEMBL2349573	2-morpholin-4-ylethyl	3-(1,3-benzothiazol-2-yl)phenyl	150
CHEMBL2349602	2-morpholin-4-ylethyl	3-(furan-2-yl)phenyl	150
CHEMBL2349604	2-morpholin-4-ylethyl	3-(3-methyl-1,2-oxazol-5yl)phenyl	150
CHEMBL2349590	2-morpholin-4-ylethyl	3-methylsulfanylphenyl	130
CHEMBL2349577	3-(4-methylpiperazin-1-yl)propan-1-yl	3-(methoxycarbonyl)phenyl	100
CHEMBL2346786	2-morpholin-4-ylethyl	3-(methoxycarbonyl)phenyl	100
CHEMBL2349575	2-morpholin-4-ylethyl	3-(2-methyltetrazol-5-yl)phenyl	60
CHEMBL2349603	2-morpholin-4-ylethyl	3-(1,3-oxazol-5-yl)phenyl	60
CHEMBL2346788	3-(piperazin-1-yl)propan-1-yl	3-(methoxycarbonyl)phenyl	60

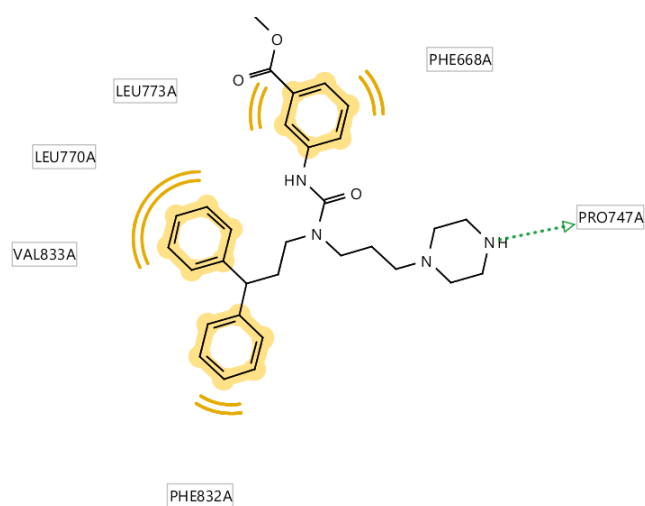
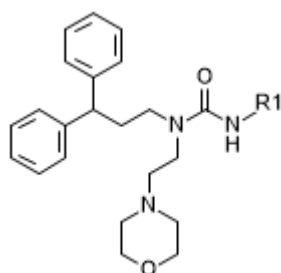
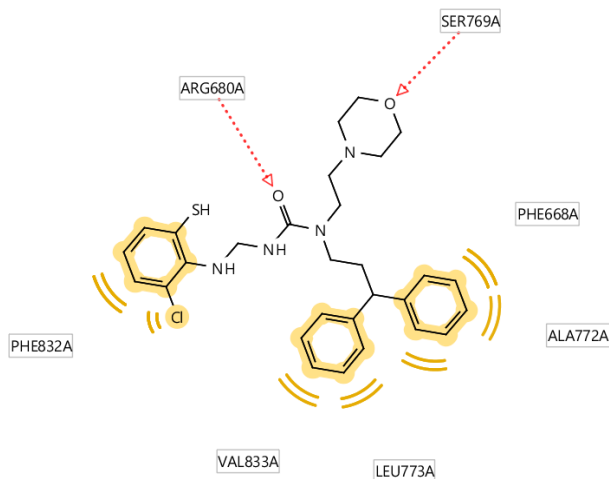


Figure 4.31 CHEMBL2346788, interaction map of the most potent trisubstituted ureas ($IC_{50} = 60$ nM) docked to CaSRs homology model on page 53

Table 4.29: Trisubstituted ureas: EC50 determined in CHO cells with a Luciferase assay (135)



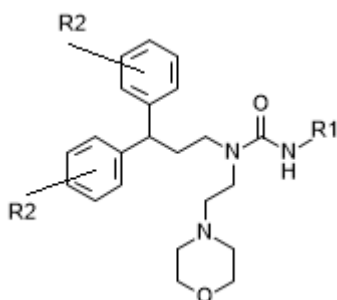
CHEMBL-ID	R1	EC50 [nM]
CHEMBL2377733	3,4-dimethyl-1,2-oxazol-5-yl	1700
CHEMBL2375385	6-ethoxy-1,3-benzothiazol-2-yl	1000
CHEMBL2377726	pyridin-3-yl	1000
CHEMBL2377742	4-methyl-1,3-benzothiazol-2-yl	600
CHEMBL2377728	Pyrazin-2-yl	450
CHEMBL2377729	6-methoxy-3-methylpyrimidin-2-yl	400
CHEMBL2377734	Methoxycarbon-2-ylthiophen-3-yl	400
CHEMBL2377738	1-benzylbenzimidazol-2-yl	300
CHEMBL2377731	1,3-thiazol-2-yl	300
CHEMBL2377730	4-(1,3-thiazol-2-yl)pyrimidin-2yl	200
CHEMBL2377732	5-methyl-1,2-oxazol-3-yl	200
CHEMBL2375386	6-fluoro-1,3-benzothiazol-2-yl	200
CHEMBL2377750	6-ethoxycarbonyl-1,3-benzothiazol-2-yl	200
CHEMBL2377751	6-(dimethylaminoethylaminocarbonyl)-1,3-benzothiazol-2-yl	200
CHEMBL2377737	1-methylbenzimidazol-2-yl	150
CHEMBL2375383	6-carboxy-1,3-benzothiazol-2-yl	100
CHEMBL2377745	6-trifluoromethoxy-1,3-benzothiazol-2-yl	90
CHEMBL2377727	4,6-dimethylpyridin-2-yl	80
CHEMBL2377747	6-hydroxymethyl-1,3-benzothiazol-2-yl	80
CHEMBL2377752	6-methylsulfonyl-1,3-benzothiazol-2-yl	80
CHEMBL2377739	5-Cl-1,3-benzoxazol-2-yl	80
CHEMBL2377746	6-Cl-1,3-benzothiazol-2-yl	60
CHEMBL2377736	1H-benzimidazol-2-yl	60
CHEMBL2377744	6-methoxy-1,3-benzothiazol-2-yl	40
CHEMBL2377741	4-methoxy-1,3-benzothiazol-2-yl	30
CHEMBL2377743	5-(methanesulfonamido)-1,3-benzothiazol-2-yl	30
CHEMBL2377748	6-(methanesulfonamido)-1,3-benzothiazol-2-yl	30
CHEMBL2377749	6-methyl-1,3-benzothiazol-2-yl	30
CHEMBL2377735	1,3-benzothiazol-2-yl	20
CHEMBL2377740	4-Cl-1,3-benzothiazol-2-yl	20



CHEMBL2377740

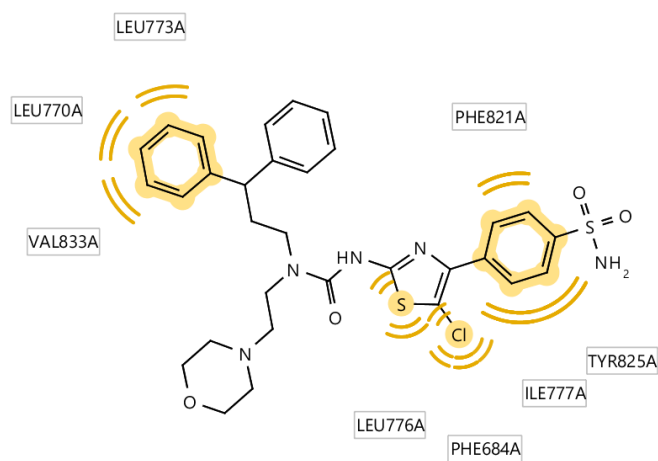
Figure 4.32 CHEMBL2377740, interaction map of the most potent trisubstituted ureas ($IC_{50} = 20$ nM) docked to CaSRs homology model on page 53

Table 4.30: urea-based calcimimetics; EC_{50} determined in rat liver microsomes (136)



CHEBML-ID	R1	R2	EC_{50} [nM]
CHEMBL3093413	5-phenyl-1,3,4-oxdiazol-2-yl	H	1030
CHEMBL3093412	6-phenyl-pyridin-2-yl	H	300
CHEMBL3093414	5-phenyl-1,3,4-thiadiazol-2-yl	H	286
CHEMBL3093415	5-phenyl-1,2,4-thiadiazol-3-yl	H	175
CHEMBL3091483	5-cyano-4-phenyl-1,3-thiazol-2-yl	H	157
CHEMBL3093418	5-fluoro-4-phenyl-1,3-thiazol-2-yl	H	53
CHEMBL3093417	4-phenyl-1,3-thiazol-2-yl	H	36
CHEMBL3093416	3-phenyl-1,2,4-thiadiazol-5-yl	H	23
CHEMBL3093430	3-(4-methanesulfamidophenyl)-1,2,4-thiadiazol-5-yl	4-F	20
CHEMBL3093426	5-chloro-4-(4-sulfamoylphenyl)-1,3-thiazol-2-yl	3-F	19
CHEMBL3093420	5-chloro-4-phenyl-1,3-thiazol-2-yl	H	18
CHEMBL3093419	5-methylsulfonyl-4-phenyl-1,3-thiazol-2-yl	H	17
CHEMBL3093411	4-[4-(methanesulfonamido)phenyl]-1,3-thiazol-2-yl	H	16
CHEMBL3093427	5-chloro-4-(4-sulfamoylphenyl)-1,3-thiazol-2-yl	4-F	10
CHEMBL3093428	5-Cl-4-[4-(methanesulfonamido)phenyl]-1,3-thiazol-2-yl	4-F	9
CHEMBL3093429	3-(4-sulfamoylphenyl)-1,2,4-thiadiazol-5-yl	4-F	9

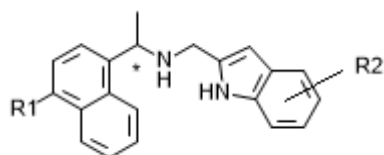
CHEMBL3093423	3-[4-(methanesulfonamido)phenyl]-1,2,4-thiadiazol-5-yl	H	6
CHEMBL3093425	5-chloro-4-(4-sulfamoylphenyl)-1,3-thiazol-2-yl	2-F	4
CHEMBL3093424	3-(4-sulfamoylphenyl)-1,2,4-thiadiazol-5-yl	H	3
CHEMBL3093421	5-Cl-4-[4-(methanesulfonamido)phenyl]-1,3-thiazol-2-yl	H	3
CHEMBL3093422	5-chloro-4-(4-sulfamoylphenyl)-1,3-thiazol-2-yl	H	2



CHEMBL3093422

Figure 4.33 CHEMBL3093422, interaction map of the most potent urea-based calcimimetics ($IC_{50} = 2 \text{ nM}$) docked to CaSRs homology model on page 53

Table 4.31: Calindol derivatives; EC_{50} measured in CHO cells with an IP accumulation assay (137)



CHEBML-ID	R1	R2	(*)	EC_{50} [nM]
CHEMBL3753729	H	4-phenyl	R	136
CHEMBL1801356	H	H	R	132
CHEMBL3753320	F	H	R,S	131
CHEMBL3753595	H	5-hydroxy	R	122.2
CHEMBL3754639	H	4-hydroxy	R	102
CHEMBL3753306	H	7-nitro	R	20

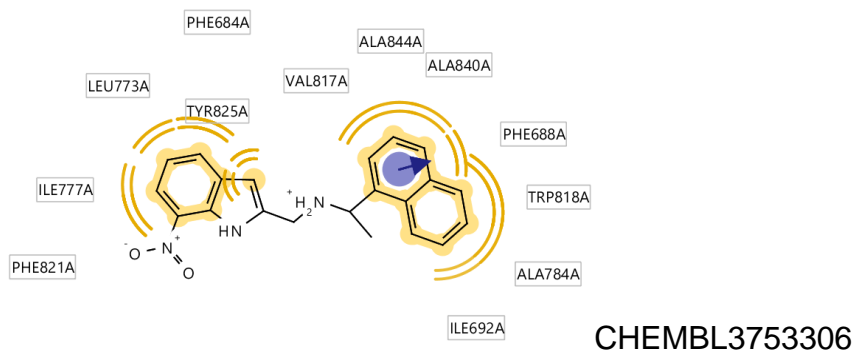
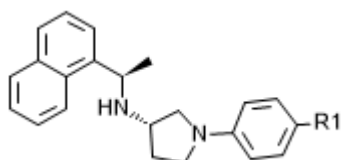
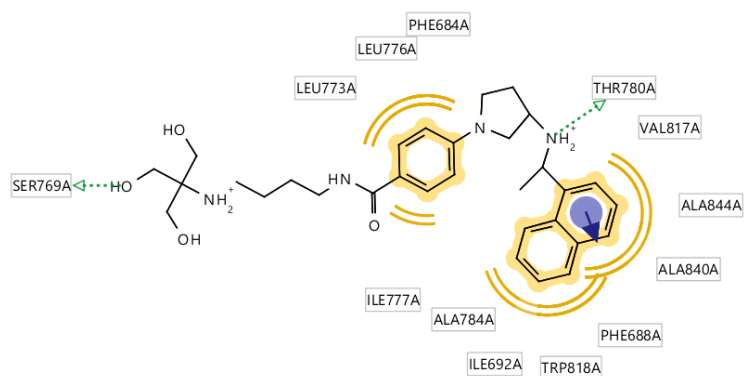


Figure 4.34 CHEMBL3753306, interaction map of the most potent calindol derivatives ($IC_{50} = 20 \text{ nM}$) docked to CaSRs homology model on page 53

Table 4.32: Evocalcet derivatives; EC_{50} measured in human liver microsomes (130)



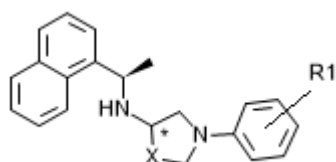
CHEBML-ID	R1	EC_{50} [nM]
CHEMBL4060025	N-(2-aminoethyl)amide	21885
CHEMBL4067601	N,N-[2-(dimethylamino)ethyl]sulfonamide	21885
CHEMBL4101097	N-[2-[[1,3-dihydroxy-2-(hydroxymethyl)propan-2yl]amino]ethyl]amide	7943
CHEMBL4084641	N-[2-(4,5-dicarboxyimidazole-1-yl)ethyl]sulfonamide	7943
CHEMBL4086045	N-[2-(4,5-dicarboxyimidazole-1-yl)ethyl]amide	5012
CHEMBL4076239	phosphonic acid	3162.28
CHEMBL4069946	2-carboxyl-1-oxoethylsulfonamide	1584.89
CHEMBL4091950	3-carboxy-1-oxopropan-1-ylsulfonamide	501.19
CHEMBL4097717	(1R)-1,2-dicarboxyethylamide	398.11
CHEMBL4079960	carboxylic acid	158.49
CHEMBL4062970	N,N,N-[2-(trimethylamine)ethyl]amide	125.89
CHEMBL4070562	1-carboxy-5-aminopentan-1-ylamide	100
CHEMBL4068107	Carboxymethylamide	79.43
CHEMBL4071600	3-(4-carboxypiperidin-1-yl)propan-1-ylamide	50
CHEMBL4074358	N-[1,3-dihydroxy-2-(hydroxymethyl)propan-2-yl]amide	39.81
CHEMBL4092377	2-sulfoethylamide	39.81
CHEMBL4094738	Carboxymethylsulfonamide	39.81
CHEMBL4082125	N-methyl-N-[(1S,2R,3R,4S)-pentahydroxypentyl]amide	31.62
CHEMBL4072950	Sulfonamide	19.95
CHEMBL4100132	N-[4-[[1,3-dihydroxy-2-(hydroxymethyl)propan-2yl]amino]butyl]amide	15.85



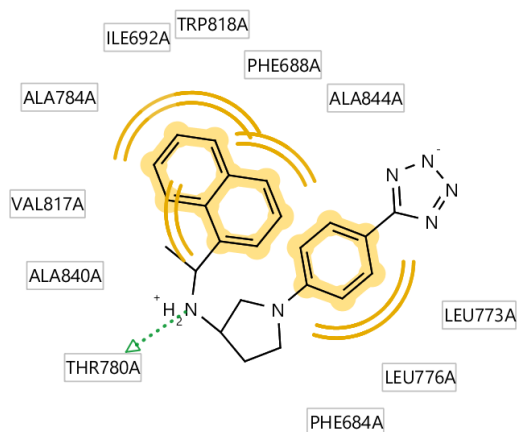
CHEMBL4100132

Figure 4.35 CHEMBL4100132, interaction map of the most potent Evocalcet derivatives ($IC_{50} = 15.85$ nM) docked to CaSRs homology model on page 53

Table 4.33: 3-aminopyrrolidine; EC_{50} measured in Madin-Darby canine kidney cells using a FLIPR assay (138)



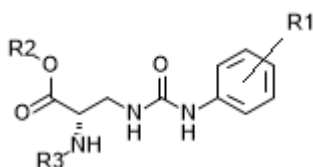
CHEBML-ID	R1	X	(*)	EC_{50} [nM]
CHEMBL4127295	4-carboxy	CH ₂	R	8800
CHEMBL4130141	4-carboxy	C ₂ H ₄	S	1100
CHEMBL4130093	4-trifluoromethoxy	C ₂ H ₄	S	600
CHEMBL4125724	3-trifluoromethoxy	CH ₂	S	600
CHEMBL4127153	3-CF ₃	CH ₂	S	450
CHEMBL4125688	3-CF ₃	CH ₂	R	430
CHEMBL4125917	3-CF ₃	C ₂ H ₄	R	380
CHEMBL4126057	3-trifluoromethoxy	CH ₂	R	320
CHEMBL4126482	4-carboxy	CH ₂	S	290
CHEMBL4129371	4-trifluoromethoxy	CH ₂	R	230
CHEMBL4128542	3-trifluoromethoxy	C ₂ H ₄	R	190
CHEMBL4129011	4-trifluoromethoxy	CH ₂	S	120
CHEMBL4128835	Carboxymethyl	CH ₂	S	93
CHEMBL4126877	4-(2-hydroxyethyl)amide	CH ₂	S	50
CHEMBL4126450	2-H-tetrazol-5-yl	CH ₂	S	17



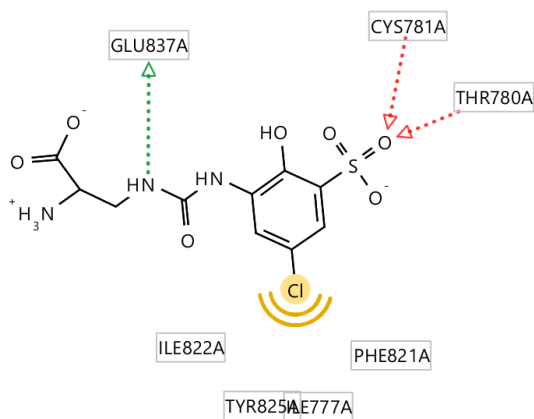
CHEMBL4126450

Figure 4.36 CHEMBL4126450, interaction map of the most potent 3-aminopyrrolidine ($IC_{50} = 17 \text{ nM}$) docked to CaSRs homology model on page 53

Table 4.34: Alkylamine derivatives; (139)



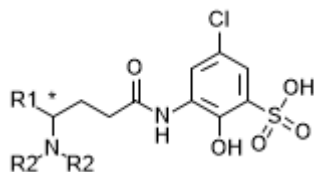
CHEBML-ID	R1	R2	R3	EC50 [nM]
CHEMBL3923504	3-acetamido, 2-OH, 5-sulfo	H	H	21100
CHEMBL3958734	3-COOH, 4-CH ₃	H	H	15000
CHEMBL3931232	2-methoxy, 3-COOH	H	H	5800
CHEMBL3922323	2-methyl, 5-COOH	H	H	5000
CHEMBL3903207	3-sulfo, 4-methoxy	H	H	3700
CHEMBL3919534	3-COOH, 5-methoxy	H	H	3600
CHEMBL3928924	3-Cl, 5-COOH	H	H	2300
CHEMBL3946636	3-nitro, 5-COOH	H	H	1800
CHEMBL3925295	3-COOH	H	H	1600
CHEMBL3967566	2,4-dimethyl, 5-sulfo	H	H	1100
CHEMBL3937957	2-OH, 3-COOH	H	H	1000
CHEMBL3964061	2-methoxy, 5-sulfo	H	H	1000
CHEMBL3950758	2-F, 5-sulfo	H	H	490
CHEMBL3974155	3-Cl, 4-methyl, 5-sulfo	H	methyl	313
CHEMBL3889990	3-sulfo	H	H	18
CHEMBL3972910	2-methyl, 3-Cl, 5-sulfo	H	H	3
CHEMBL3937671	2-OH, 3-sulfo, 5-Cl	methyl	H	3
CHEMBL3980928	2-OH, 3-sulfo, 5-Cl	H	H	3



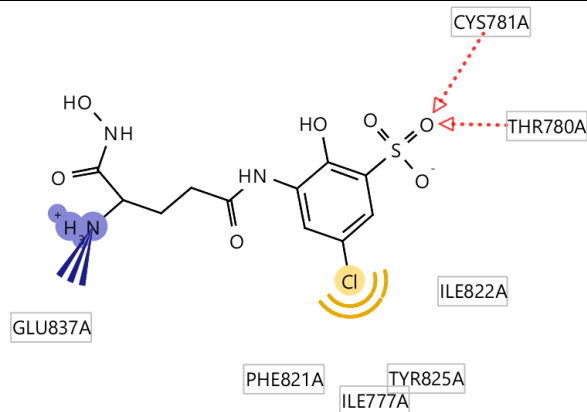
CHEMBL3980928

Figure 4.37 CHEMBL398092, interaction map of the most potent alkylamine derivatives; ($IC_{50} = 3 \text{ nM}$) docked to CaSRs homology model on page 53

Table 4.35: Alkylamine derivatives: (139)



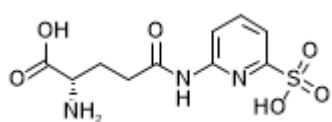
CHEBML-ID	R1	R2	(*)	EC50 [nM]
CHEMBL3975131	cycloheptylaminocarbonyl	H	R,S	3400
CHEMBL3949209	morpholin-4-ylcarbonyl	H	R,S	3100
CHEMBL3962913	benzenesulfonamidocarbonyl	H	S	2300
CHEMBL3955623	carboxy	methyl	S	1800
CHEMBL3986052	2-(3-Cl-phenyl)ethylaminocarbonyl	H	S	1200
CHEMBL3897239	5-benzoyl-1,3,4-oxdiazol-2yl	H	S	880
CHEMBL3957364	methanesulfonamido	H	S	390
CHEMBL3971904	2-(2-phenylacetyl)hydrazinylcarbonyl	H	S	250
CHEMBL3943365	5-methyl-1,3,4-oxadiazol-2yl	H	S	180
CHEMBL3974015	hydroxyaminocarbonyl	H	S	39



CHEMBL3974015

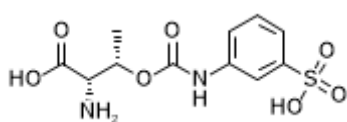
Figure 4.38 CHEMBL3974015, interaction map of the most potent alkylamine derivatives ($IC_{50} = 39 \text{ nM}$) docked to CaSRs homology model on page 53

Table 4.36: Alkylamine derivatives: (139)



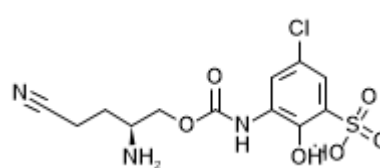
CHEMBL3941122

EC50 = 13000 nM



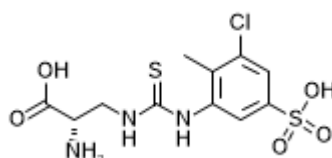
CHEMBL3966753

EC50 = 7000 nM



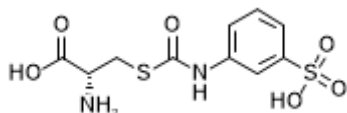
CHEMBL3948703

EC50 = 1100 nM



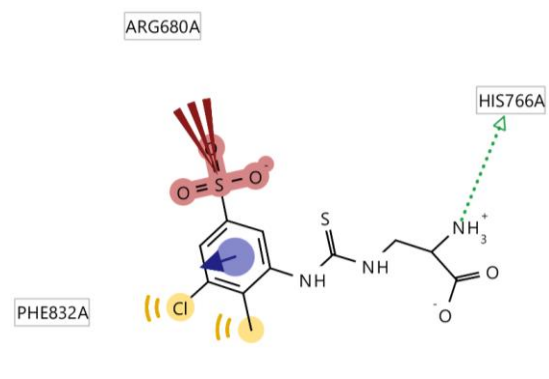
CHEMBL3978873

EC50 = 40 nM



CHEMBL3895461

EC50 = 4 nM



CHEMBL3895461

Figure 4.39 CHEMBL3895461, interaction map of the most potent alkylamine derivatives ($IC_{50} = 4$ nM) docked to CaSRs homology model on page 53

**interaction profile of
top 16 positive allosteric modulators**

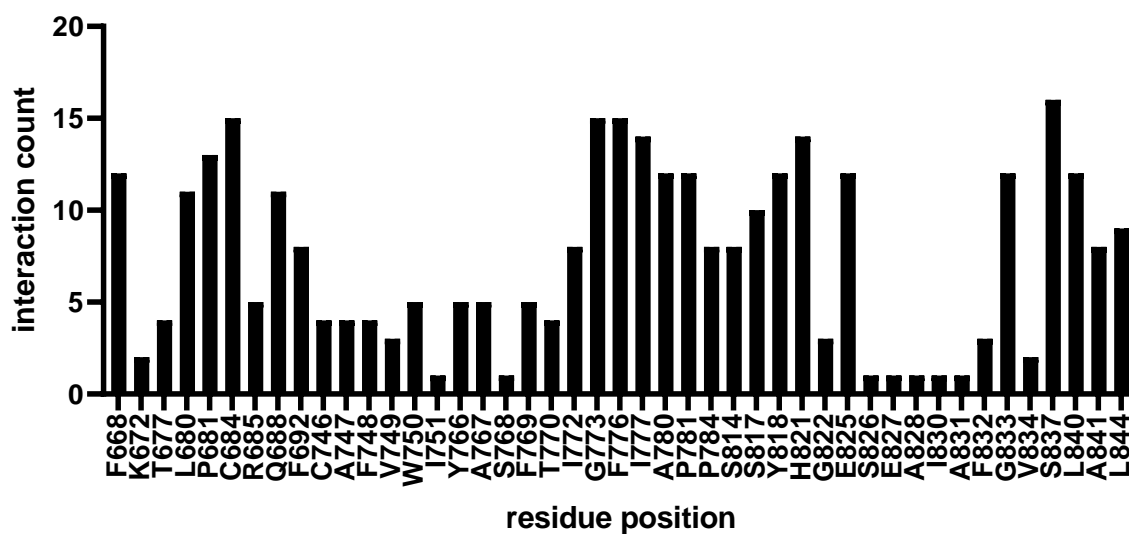


Figure 4.40 histogram of residue interactions of CaSR transmembrane with the top-scoring positive allosteric modulators from 16 different scaffolds, docking done with AutoDock VINA, model was used from page 46, interaction profile; CHEMBL2092942; CHEMBL292376; CHEMBL569182; CHEMBL180672; CHEMBL1224424; CHEMBL1200776; CHEMBL1256368; CHEMBL2346788; CHEMBL2377740; CHEMBL3093422; CHEMBL3753306; CHEMBL4100132; CHEMBL4126450; CHEMBL3980928; CHEMBL3974015; CHEMBL3895461

4.3 Compounds for experimental validation

Table 4.37 purchased compounds for the first experimental validation of the SAFAN-ISPSM screening results for potential allosteric modulators of CaSR

CID:	Name:	screening ID:	vendor:	Catalogue number:
1	Toremifene citrate	DB00539	Key Organics	KS-5242
2	Bucindolol	DB12752	Cayman Chemicals	21070
3	Alprenolol HCl	DB00866	Alsachim	SVI-ALS-18-132
4	Bupranolol	DB08808	Abcam	Ab141132
5	Rivastigmine	DB00989	Key Organics	KS-1188
6	Oxprenolol HCl	DB01580	Key Organics	KS-1152
7	1-phenyl-1-proanol	FDB008267	Key Organics	STR02714

Compounds were dissolved in DMSO to create 10 mM stock solutions and stored at -20°C.

Table 4.38 purchased compounds for the second experimental validation of the SAFAN-ISPSM screening results for potential allosteric modulators of CaSR

CID:	ZINC database ID:	vendor:	Catalogue number:
1	ZINC00000065809	Otava Ltd.	0127440833
2	ZINC000020394024	ChemBridge Corporation	9052361
3	ZINC00000091620	ChemBridge Corporation	8879399
4	ZINC000000139716	Maybridge Ltd. Cat num.	CD06887SC
5	ZINC000010020294	Life Chemicals Inc.	F2894-0192
6	ZINC00000036062	ChemBridge Corporation	6049907
7	ZINC000020136919	ENAMINE Ltd. Cat. num.	Z86194874
8	ZINC000012324922	Eximed	EiM07-30473

Compounds were dissolved in DMSO to create 20 mM stock solutions and stored at -20°C.

4.4 Kits

IP-one Gq HTRF Kit: Cisbio, Cat. No. 62IPAPEB

4.5 Materials for the IP-1 Gq assay

DMSO: Dimethyl sulfoxide D8418-500ml, Sigma

PDL: Poly-D-lysine hydrobromide, Sigma, P7886-50MG

Trypsin: GIBCO Trypsin, 25300-054

Cytoplates-one: CytoOne Multiwell Plate with Lid 96, CC7682-7596

Microplates: 96 well microplates Greiner Bio-One, 655101

Reagents reservoir: Diversified Biotech Reagent Reservoirs

Lysis & detection buffer: Lysis & Detection Buffer 6 CisBio 200ml, 62CL6FDF

low volume plates: HTRF 96 well low volume plate Cisbio, 66PL96005

Count slides: BioRad Counting Slides

Trypan blue: Trypan Blue 0,4% 0,85% NaCl Lonza, 17-942E

DMEM: Gibco DMEM 500ml, 41965-039

Glutamax: GlutaMax Gibco 100ml, 35050-038

Penstrep: GIBCO PENSTREP 100ml, 15140-122

Pyruvate: Sigma Na pyruvate 100ml, 58636-100ml

HBSS: Gibco HBSS (10x) no calcium no magnesium, 14185045

HEPES: GIBCO HEPES (1M), 15630080

ViewSeal: Sigma-Aldrich Greiner ViewSeal for 96plates, transp., Z617571-100EA

HTRF reader: Tecan infinite M1000Pro

Incubator: Heraeus Cytoperm 2

Cell counter: Bio Rad, TC10 Automated Cell Counter

pH meter: 827 pH lab, Metrohm

4.6 IP-1 Gq assay

The IP-1 accumulation was determined for every compound at different calcium concentrations (0.2 – 7.5 mM) in comparison to the reference compounds NPS-R568 (positive allosteric modulator) and NPS-2143 (negative allosteric modulator).

The experiments for IP1 accumulation measurements were performed with Flp-In HEK293 HA hCaSR WT cells and reference cells as a negative control (Flp-In HEK293 myc GABA-A delta cells/ empty HEK293 cells). The cells were provided by Iris Mos from the University of Copenhagen. The assay was executed in two days according to the IP-1 accumulation protocol.

On the first day, HTRF 96-low volume plates were prepared to achieve a count of 50000 cells per well. The preparation started by mixing a PDL stock solution (4 mg/ml) in PBS. Each well was incubated with 50µl PDL stock solution at 37°C for 30 minutes to reduce the cells' detachment. During the incubation time, the cells' confluence was checked with a passage number never higher than 25. The cells are grown in media [DMEM (1x), + glutaMAX (1%), + pyruvate (1%), + dFBS (10%), + pen-strep (1%)]. After the confluence was determined, the media was removed, and the cells were washed with PBS. The PBS was removed and the cells detached after 2ml trypsin was added and an incubation time of one minute at room temperature. The resuspended cell solution was diluted 1:10 in media, and 10µl of the solution was mixed with 10µl trypan blue for cell counting. The cells were spun down and resuspended in media to generate a 5 million cell per millilitre solution. The solution was diluted with media to create at least 10 ml of 500000 cells/ml solution. The PDL incubated plate was washed twice with 100 µl/well PBS and seeded with 100 µl/well cell solution (50000 cells/well). The seeded cells were incubated for 18 hours at 37°C and 5% CO₂.

On the second day, the confluence was checked for the wells. Preincubation solutions were prepared containing the ligand at final concentration in assay buffer (HBSS [- Ca²⁺, - Mg²⁺]; + 20 mM HEPES; pH = 7.4) and 0.2 mM Ca²⁺ to avoid cell death without CaSR stimulation. The DMSO concentrations of all ligand solutions were equalized to avoid bias at a maximum of 0.28 (V/V) %. DMSO concentrations above 1% hamper cell growth. Ligand stimulation solutions were prepared to have the same composition as the preincubation solutions but with

20 mM LiCl and calcium at the desired stimulation concentration. The cells were washed with 100 μ l/well of assay buffer and then treated with 50 μ l preincubation solution at 37°C for 30 minutes. The solution was removed, the cells were washed again with 100 μ l assay buffer, and 50 μ l of the stimulation solutions were added. The plates were incubated at 37°C for 30 minutes. The cells were rewashed with assay buffer and lysed with 30 μ l lysis and detection buffer (Cisbio) for 30 minutes at room temperature. In the end, 30 μ l assay buffer was added, 10 μ l of the solution was transferred to a 96-well low volume reader plate (CisBio) and mixed with 10 μ l anti-IP1 cryptate TB conjugate and IP1-d2 conjugate in assay buffer (1:1:38) fluorophore solution. After sealing, the plates were incubated at room temperature for one hour in the dark. Finally, the plates were excited at 340 nm and read with a homologous time-resolved fluorescence (HTRF) reader (PerkinElmer EnVision or infinite M1000pro) at 665 nm and 620 nm. The HTRF ratio was calculated according to Equation 7. The standard curve was recorded with IP1 standards (Cisbio) to determine the measured IP1 concentration (Figure 2.66).

The compound curves were recorded according to the IP-one accumulation protocol, but the cells were always stimulated a fixed calcium concentration (EC50 = 4,75mM). The IP-1 concentrations were measured at 30 μ M, 100 x pK concentration, 3 μ M, 1 μ M, 0.3 μ M, 0.1 μ M, 0.03 μ M and 0.01 μ M compound concentration in triplicates. The calcium curves of the compounds were determined according to the protocol of the accumulation assay. However, the compound concentrations were fixed at 1 μ M, and the calcium concentration was measured at 10 mM, 6 mM, 5 mM, 4.5 mM, 3.5 mM, 2.5 mM, 1 mM and 0.5mM calcium.

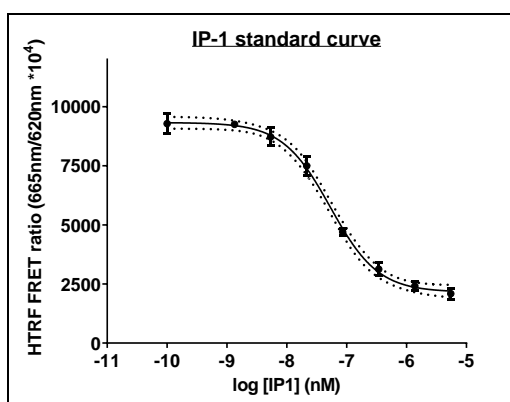


Figure 4.41 IP1 standard curve with IP1 standards (CisBio)

5 **Bibliography:**

1. Brown EM, Gamba G, Riccardi D, Lombardi M, Butters R, Kifor O, et al. Cloning and characterization of an extracellular Ca(2+)-sensing receptor from bovine parathyroid. *Nature* [Internet]. 1993;366(6455):575–80. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8255296>
2. Alfadda TI, Saleh AMA, Houillier P, Geibel JP. Calcium-sensing receptor 20 years later. *Am J Physiol Cell Physiol* [Internet]. 2014;307(3):C221-31. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4121584&tool=pmcentrez&rendertype=abstract>
3. Zhang C, Miller CL, Brown EM, Yang JJ. The calcium sensing receptor: from calcium sensing to signaling. *Sci China Life Sci.* 2015 Jan;58(1):14–27.
4. Widler L, Altmann E, Beerli R, Breitenstein W, Bouhelal R, Buhl T, et al. 1-Alkyl-4-phenyl-6-alkoxy-1H-quinazolin-2-ones: a novel series of potent calcium-sensing receptor antagonists. *J Med Chem.* 2010 Mar;53(5):2250–63.
5. Bikle D, Bräuner-Osborne H, Brown EM, Chang W, Conigrave A, Hannan F, et al. Calcium-sensing receptor (version 2019.4) in the IUPHAR/BPS Guide to Pharmacology Database. *IUPHAR/BPS Guid to Pharmacol CITE* [Internet]. 2019;2019(4). Available from: <https://doi.org/10.2218/gtopdb/F12/2019.4>.
6. *Frontiers in Protein Structure, Function, and Dynamics* - Google Books [Internet]. [cited 2020 Nov 12]. Available from: [https://books.google.it/books?id=UCzvDwAAQBAJ&pg=PP6&dq=protein+structure+determination+methods&hl=de&sa=X&ved=2ahUKEwjtyOT7PzsAhVF_qQKHSszC6YQ6AEwAHoECAIQAg#v=onepage&q=protein structure determination methods&f=false](https://books.google.it/books?id=UCzvDwAAQBAJ&pg=PP6&dq=protein+structure+determination+methods&hl=de&sa=X&ved=2ahUKEwjtyOT7PzsAhVF_qQKHSszC6YQ6AEwAHoECAIQAg#v=onepage&q=protein%20structure%20determination%20methods&f=false)
7. Hendy GN, Canaff L. Calcium-sensing receptor gene: Regulation of expression. *Frontiers in Physiology.* 2016.
8. Bateman A. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids*

- Res. 2019;
9. Geng Y, Mosyak L, Kurinov I, Zuo H, Sturchler E, Cheng TC, et al. Structural mechanism of ligand activation in human calcium-sensing receptor. *Elife*. 2016;5(JULY).
 10. Costanzi S. Homology modeling of class a g protein-coupled receptors. *Methods Mol Biol*. 2012;
 11. Pándy-Szekeres G, Munk C, Tsonkov TM, Mordalski S, Harpsøe K, Hauser AS, et al. GPCRdb in 2018: Adding GPCR structure models and ligands. *Nucleic Acids Res*. 2018;
 12. Zhang J, Yang J, Jang R, Zhang Y. GPCR-I-TASSER: A Hybrid Approach to G Protein-Coupled Receptor Structure Modeling and the Application to the Human Genome. *Structure*. 2015;
 13. Hille B, Dickson E, Kruse M, Falkenburger B. Dynamic Metabolic Control of an Ion Channel. In *Academic Press*; 2014. p. 219–47.
 14. Brown EM, Rosen CJ, Mulder JE. Disorders of the calcium-sensing receptor: Familial hypocalciuric hypercalcemia and autosomal dominant hypocalcemia - UpToDate [Internet]. 2019 [cited 2020 May 29]. Available from: <https://www.uptodate.com/contents/disorders-of-the-calcium-sensing-receptor-familial-hypocalciuric-hypercalcemia-and-autosomal-dominant-hypocalcemia>
 15. Sarin R, Tomar N, Ray D, Gupta N, Sharma YD, Goswami R. Absence of pathogenic calcium sensing receptor mutations in sporadic idiopathic hypoparathyroidism. *Clin Endocrinol (Oxf)*. 2006 Sep;65(3):359–63.
 16. Balan G, Bauman J, Bhattacharya S, Castrodad M, Healy DR, Herr M, et al. The discovery of novel calcium sensing receptor negative allosteric modulators. *Bioorg Med Chem Lett* [Internet]. 2009;19(12):3328–32. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19442519>
 17. Marquis RW, Lago AM, Callahan JF, Trout REL, Gowen M, DeMar EG, et al. Antagonists of the Calcium Receptor I. Amino Alcohol-Based Parathyroid Hormone Secretagogues. *J Med Chem* [Internet]. 2009 Jul 9 [cited 2018 May

10];52(13):3982–93.

Available

from:

<http://www.ncbi.nlm.nih.gov/pubmed/19492813>

18. Nemeth EF, Van Wagenen BC, Balandrin MF. Discovery and Development of Calcimimetic and Calcilytic Compounds. In: Progress in Medicinal Chemistry. 2018.
19. Patel J, Bridgeman MB. Etecalcetide (Parsabiv) for secondary hyperparathyroidism in adults with chronic kidney disease on hemodialysis. P and T. 2018.
20. Eidman KE, Wetmore JB. Treatment of secondary hyperparathyroidism: How do cinacalcet and etealcetide differ? Seminars in Dialysis. 2018.
21. Bover J, Ureña P, Ruiz-García C, da Silva A, Lescano P, del Carpio J, et al. Clinical and practical use of calcimimetics in dialysis patients with secondary hyperparathyroidism. Clin J Am Soc Nephrol. 2016;
22. Fukagawa M, Shimazaki R, Akizawa T. Head-to-head comparison of the new calcimimetic agent evocalcet with cinacalcet in Japanese hemodialysis patients with secondary hyperparathyroidism. Kidney Int. 2018;
23. Oleg T, Arthur J. O. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. J Comput Chem. 2010;
24. Krieger E, Vriend G. YASARA View - molecular graphics for all devices - from smartphones to workstations. Bioinformatics. 2014;30(20):2981–2.
25. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: A major update to the DrugBank database for 2018. Nucleic Acids Res. 2018;
26. Leach K, Gregory KJ, Kufareva I, Khajehali E, Cook AE, Abagyan R, et al. Towards a structural understanding of allosteric drugs at the human calcium-sensing receptor. Cell Res. 2016;26(5):574–92.
27. Ballesteros JA, Weinstein H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. Methods Neurosci. 1995;

28. Zhang C, Zhang T, Zou J, Miller CL, Gorkhali R, Yang JY, et al. Structural basis for regulation of human calcium-sensing receptor by magnesium ions and an unexpected tryptophan derivative co-agonist. *Sci Adv.* 2016;2(5):e1600241.
29. Moult J, Fidelis K, Kryshtafovych A, Rost B, Hubbard T, Tramontano A. Critical assessment of methods of protein structure prediction - Round VII. In: *Proteins: Structure, Function and Genetics.* 2007.
30. Moult J, Fidelis K, Kryshtafovych A, Rost B, Tramontano A. Critical assessment of methods of protein structure prediction-Round VIII. *Proteins Struct Funct Bioinforma.* 2009;
31. Moult J, Fidelis K, Kryshtafovych A, Tramontano A. Critical assessment of methods of protein structure prediction (CASP)-round IX. *Proteins Struct Funct Bioinforma.* 2011;
32. Moult J, Fidelis K, Kryshtafovych A, Schwede T, Tramontano A. Critical assessment of methods of protein structure prediction (CASP) - round x. *Proteins Struct Funct Bioinforma.* 2014;
33. Moult J, Fidelis K, Kryshtafovych A, Schwede T, Tramontano A. Critical assessment of methods of protein structure prediction: Progress and new directions in round XI. *Proteins Struct Funct Bioinforma.* 2016;
34. Moult J, Fidelis K, Kryshtafovych A, Schwede T, Tramontano A. Critical assessment of methods of protein structure prediction (CASP)—Round XII. *Proteins Struct Funct Bioinforma.* 2018;
35. Kryshtafovych A, Schwede T, Topf M, Fidelis K, Moult J. Critical assessment of methods of protein structure prediction (CASP)—Round XIII. *Proteins: Structure, Function and Bioinformatics.* 2019.
36. Yang J, Zhang Y. I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Res.* 2015;
37. Zheng W, Zhang C, Wuyun Q, Pearce R, Li Y, Zhang Y. LOMETS2: improved meta-threading server for fold-recognition and structure-based function annotation for distant-homology proteins. *Nucleic Acids Res.* 2019;

38. Yang J, Wang Y, Zhang Y. ResQ: An Approach to Unified Estimation of B-Factor and Residue-Specific Error in Protein Structure Prediction. *J Mol Biol.* 2016;
39. Yang J, Roy A, Zhang Y. BioLiP: A semi-manually curated database for biologically relevant ligand-protein interactions. *Nucleic Acids Res.* 2013;
40. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics.* 2009;
41. Krieger E, Joo K, Lee J, Lee J, Raman S, Thompson J, et al. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins: Structure, Function and Bioinformatics.* 2009.
42. Suzek BE, Huang H, Mcgarvey P, Wu CH, Bateman A. UniRef: comprehensive and non-redundant UniProt reference clusters. 2007;23(10):1282–8.
43. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank (www.rcsb.org). *Nucleic Acids Res.* 2000;
44. Hooft RWW, Vriend G, Sander C, Abola EE. Errors in protein structures [3]. *Nature.* 1996.
45. Joosten RP, Te Beek TAH, Krieger E, Hekkelman ML, Hooft RWW, Schneider R, et al. A series of PDB related databases for everyday needs. *Nucleic Acids Res.* 2011;
46. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol.* 1999;
47. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. *Bioinformatics.* 2000;
48. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, et al. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science (80-).* 2000;

49. Wu H, Wang C, Gregory KJ, Han GW, Cho HP, Xia Y, et al. Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science* (80-). 2014;
50. Doré AS, Okrasa K, Patel JC, Serrano-Vega M, Bennett K, Cooke RM, et al. Structure of class C GPCR metabotropic glutamate receptor 5 transmembrane domain. *Nature*. 2014;
51. Christopher JA, Aves SJ, Bennett KA, Doré AS, Errey JC, Jazayeri A, et al. Fragment and Structure-Based Drug Discovery for a Class C GPCR: Discovery of the mGlu5 Negative Allosteric Modulator HTL14242 (3-Chloro-5-[6-(5-fluoropyridin-2-yl)pyrimidin-4-yl]benzotrile). *J Med Chem*. 2015;
52. Ye L, Van Eps N, Zimmer M, Ernst OP, Scott Prosser R. Activation of the A_{2A} adenosine G-protein-coupled receptor by conformational selection. *Nature*. 2016;
53. Krieger E, Nabuurs SB, Vriend G. Chapter 25: homology modeling. In: *Structural Bioinformatics*. 2003.
54. Canutescu AA, Dunbrack RL. Cyclic coordinate descent: A robotics algorithm for protein loop closure. *Protein Sci*. 2003;
55. Krieger E, Koraimann G, Vriend G. Increasing the precision of comparative models with YASARA NOVA - A self-parameterizing force field. *Proteins Struct Funct Genet*. 2002;
56. Konagurthu AS, Whisstock JC, Stuckey PJ, Lesk AM. MUSTANG: A multiple structural alignment algorithm. *Proteins Struct Funct Genet*. 2006;
57. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, et al. PubChem 2019 update: Improved access to chemical data. *Nucleic Acids Res*. 2019;
58. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr*. 1993;
59. Gaulton A, Hersey A, Nowotka ML, Patricia Bento A, Chambers J, Mendez D, et al. The ChEMBL database in 2017. *Nucleic Acids Res*. 2017;

60. Mukaka M M. A guide to appropriate use of Correlation coefficient in medical research. *Melawi Med.* 2012.
61. Thafar M, Raies A Bin, Albaradei S, Essack M, Bajic VB. Comparison Study of Computational Prediction Tools for Drug-Target Binding Affinities. *Frontiers in Chemistry.* 2019.
62. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput.* 2015;
63. Dickson CJ, Madej BD, Skjevik ÅA, Betz RM, Teigen K, Gould IR, et al. Lipid14: The amber lipid force field. *J Chem Theory Comput.* 2014;
64. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general Amber force field. *J Comput Chem.* 2004;
65. Jakalian A, Jack DB, Bayly CI. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J Comput Chem.* 2002;
66. Krieger E, Nielsen JE, Spronk CAEM, Vriend G. Fast empirical pKa prediction by Ewald summation. *J Mol Graph Model.* 2006;
67. Krieger E, Darden T, Nabuurs SB, Finkelstein A, Vriend G. Making optimal use of empirical energy functions: Force-field parameterization in crystal space. *Proteins Struct Funct Genet.* 2004;
68. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. *J Chem Phys.* 1995;
69. Krieger E, Vriend G. New ways to boost molecular dynamics simulations. *J Comput Chem.* 2015;
70. Pugliese L. S.A.F.AN. BIOINFORMATICS – Computational results with experimental reliability [Internet]. 2019 [cited 2020 May 24]. Available from: <http://safan-bioinformatics.it/>
71. Scalbert A, Andres-Lacueva C, Arita M, Kroon P, Manach C, Urpi-Sarda M, et al. Databases on food phytochemicals and their health-promoting effects.

- Journal of Agricultural and Food Chemistry. 2011.
72. Aldinucci M, Bagnasco S, Lusso S, Pasteris P, Rabellino S, Vallero S. OCCAM: A flexible, multi-purpose and extendable HPC cluster. In: Journal of Physics: Conference Series. 2017.
 73. Degorce F, Card A, Soh S, Trinquet E, Knapik GP, Xie B. HTRF: A technology tailored for drug discovery - A review of theoretical aspects and recent applications. Current Chemical Genomics. 2009.
 74. Garbison KE, Heinz BA, Lajiness ME. IP-3/IP-1 Assays [Internet]. Eli Lilly & Company and the National Center for Advancing Translational Sciences; 2004. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22553873>
 75. He G, Acuff T, Ryan W, Yang C, Douthit M, Bowman R, et al. Inhibitory effects of calcium antagonists on alpha-adrenoceptor-mediated contraction in the human internal mammary artery. Br J Clin Pharmacol. 1994;
 76. Nemeth EF. The search for calcium receptor antagonists (calcilytics). J Mol Endocrinol. 2002;29(1):15–21.
 77. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: An Open chemical toolbox. J Cheminform. 2011;
 78. Yuan Y, Pei J, Lai L. LigBuilder 2: A practical de novo drug design approach. J Chem Inf Model. 2011;
 79. FooDB [Internet]. [cited 2019 May 31]. Available from: www.foodb.ca
 80. Choi W, Choi CH, Kim YR, Kim SJ, Na CS, Lee H. HerDing: Herb recommendation system to treat diseases using genes and chemicals. Database. 2016;
 81. Chen CYC. TCM Database@Taiwan: The world's largest traditional Chinese medicine database for drug screening In Silico. PLoS One. 2011;
 82. ZINC TCM Database @ Taiwan [Internet]. [cited 2018 Dec 10]. Available from: <https://zinc.docking.org/catalogs/tcmnp/>
 83. Koes DR, Camacho CJ. ZINCPharmer: Pharmacophore search of the ZINC database. Nucleic Acids Res. 2012;

84. Sterling T, Irwin JJ. ZINC 15 - Ligand Discovery for Everyone. *J Chem Inf Model*. 2015;
85. Wolber G, Langer T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J Chem Inf Model*. 2005;
86. Molecular Operating Environment (MOE) 2013.08. Molecular Operating Environment (MOE), 2013.08; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7. *Mol Oper Environ (MOE)*, 201308; Chem Comput Gr Inc, 1010 Sherbooke St West, Suite #910, Montr QC, Canada, H3A 2R7, 2013. 2016;
87. Koes DR, Camacho CJ. Pharmer: Efficient and exact pharmacophore search. *J Chem Inf Model*. 2011;
88. Schneidman-Duhovny D, Dror O, Inbar Y, Nussinov R, Wolfson HJ. PharmaGist: a webserver for ligand-based pharmacophore detection. *Nucleic Acids Res*. 2008;
89. Arora B, Coudrat T, Wootten D, Christopoulos A, Noronha SB, Sexton PM. Prediction of Loops in G Protein-Coupled Receptor Homology Models: Effect of Imprecise Surroundings and Constraints. *J Chem Inf Model*. 2016;
90. Gutmanas A, Alhroub Y, Battle GM, Berrisford JM, Bochet E, Conroy MJ, et al. PDBe: Protein data bank in Europe. *Nucleic Acids Res*. 2014;
91. Koehl A, Hu H, Feng D, Sun B, Zhang Y, Robertson MJ, et al. Structural insights into the activation of metabotropic glutamate receptors. *Nature*. 2019;
92. Wu H, Wang C, Gregory KJ, Han GW, Cho HP, Xia Y, et al. Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science [Internet]*. 2014;344(6179):58–64. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24603153>
93. Christopher JA, Orgován Z, Congreve M, Doré AS, Errey JC, Marshall FH, et al. Structure-Based Optimization Strategies for G Protein-Coupled Receptor (GPCR) Allosteric Modulators: A Case Study from Analyses of New

- Metabotropic Glutamate Receptor 5 (mGlu5) X-ray Structures. *J Med Chem*. 2019;
94. Flock T, Hauser AS, Lund N, Gloriam DE, Balaji S, Babu MM. Selectivity determinants of GPCR-G-protein binding. *Nature*. 2017;
 95. Jung J, Lee B. Protein structure alignment using environmental profiles. *Protein Eng*. 2000;
 96. Krieger E. md_runmembrane.mcr [Internet]. 2019 [cited 2019 Jun 7]. Available from: http://www.yasara.org/md_runmembrane.mcr
 97. McCormick WD, Atkinson-Dell R, Campion KL, Mun HC, Conigrave AD, Ward DT. Increased receptor stimulation elicits differential calcium-sensing receptorT888 dephosphorylation. *J Biol Chem*. 2010;285(19):14170–7.
 98. Roy K, Kar S, Das RN. Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment. *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*. 2015.
 99. Proekt A, Hemmings HC. Mechanisms of drug action. In: *Pharmacology and Physiology for Anesthesia: Foundations and Clinical Application*. 2018.
 100. Ward DT, Riccardi D. New concepts in calcium-sensing receptor pharmacology and signalling. Vol. 165, *British Journal of Pharmacology*. 2012. p. 35–48.
 101. Backman TWH, Cao Y, Girke T. ChemMine tools: An online service for analyzing and clustering small molecules. *Nucleic Acids Res*. 2011;
 102. Qiu J, Elber R. SSALN: An alignment algorithm using structure-dependent substitution matrices and gap penalties learned from structurally aligned protein pairs. *Proteins Struct Funct Genet*. 2006;
 103. Mückstein U, Hofacker IL, Stadler PF. Stochastic pairwise alignments. *Bioinformatics*. 2002;
 104. Darden T, York D, Pedersen L. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J Chem Phys*. 1993;

105. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;
106. CLC Manuals - clcsupport.com [Internet]. [cited 2020 Oct 1]. Available from: http://resources.qiagenbioinformatics.com/manuals/clcmainworkbench/800/index.php?manual=How_does_BLAST_work.html
107. Henikoff S, Henikoff JG. Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci U S A.* 1992;
108. Introduction to Pharmaceutical Bioinformatics - Jarl E. S. Wikberg - Google Books [Internet]. [cited 2020 Oct 1]. Available from: <https://books.google.at/books?id=KbKsoAEACAAJ&dq=introduction+to+pharmaceutical+bioinformatics+2011&hl=de&sa=X&ved=2ahUKEwivoY704JP sAhUymIsKHe0NDagQ6AEwAHoECAAQAQ>
109. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research.* 1997.
110. Bhagwat M, Aravind L. PSI-BLAST tutorial. *Methods Mol Biol* [Internet]. 2007 Jan 11 [cited 2020 Oct 2];395:177–86. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK2590/>
111. Vázquez J, López M, Gibert E, Herrero E, Javier Luque F. Merging ligand-based and structure-based methods in drug discovery: an overview of combined virtual screening approaches. *Molecules.* 2020.
112. Frank E, Hall MA, Witten IH. The WEKA Workbench. Online Appendix for “Data Mining: Practical Machine Learning Tools and Techniques.” In: Morgan Kaufmann. 2016.
113. Lipman DJ, Pearson WR. Rapid and sensitive protein similarity searches. *Science* (80-). 1985;
114. Daylight Cheminformatics [Internet]. [cited 2020 May 23]. Available from: <https://www.daylight.com/smiles/index.html>
115. Weininger D. SMILES, a Chemical Language and Information System: 1: Introduction to Methodology and Encoding Rules. *J Chem Inf Comput Sci.*

- 1988;
116. Shcherbakova I, Balandrin MF, Fox J, Ghatak A, Heaton WL, Conklin RL. 3H-Quinazolin-4-ones as a new calcilytic template for the potential treatment of osteoporosis. *Bioorganic Med Chem Lett.* 2005;15(6):1557–60.
 117. Shcherbakova I, Huang G, Geoffroy OJ, Nair SK, Swierczek K, Balandrin MF, et al. Design, new synthesis, and calcilytic activity of substituted 3H-pyrimidin-4-ones. *Bioorganic Med Chem Lett.* 2005;15(10):2537–40.
 118. Gavai A V., Vaz RJ, Mikkilineni AB, Roberge JY, Liu Y, Lawrence RM, et al. Discovery of novel 1-arylmethyl pyrrolidin-2-yl ethanol amines as calcium-sensing receptor antagonists. *Bioorganic Med Chem Lett.* 2005;
 119. Allen JG, Fotsch C, Babij P. Emerging targets in osteoporosis disease modification. Vol. 53, *Journal of Medicinal Chemistry.* 2010. p. 4332–53.
 120. Kessler A, Faure H, Petrel C, Rognan D, Césarío M, Ruat M, et al. N1-benzoyl-N2-[1-(1-naphthyl)ethyl]-trans-1,2-diaminocyclohexanes: Development of 4-chlorophenylcarboxamide (Calhex 231) as a new calcium sensing receptor ligand demonstrating potent calcilytic activity. *J Med Chem.* 2006;
 121. Yang W, Ruan Z, Wang Y, Van Kirk K, Zhengping M, Arey BJ, et al. Discovery and structure-activity relationships of trisubstituted pyrimidines/pyridines as novel calcium-sensing receptor antagonists. *J Med Chem.* 2009;
 122. Marquis RW, Lago AM, Callahan JF, Rahman A, Dong X, Stroup GB, et al. Antagonists of the calcium receptor. 2. Amino alcohol-based parathyroid hormone secretagogues. *J Med Chem.* 2009;52(21):6599–605.
 123. Shinagawa Y, Inoue T, Hirata K, Katsushima T, Nakagawa T, Matsuo Y, et al. New aminopropandiol derivatives as orally available and short-acting calcium-sensing receptor antagonists. *Bioorganic Med Chem Lett.* 2010;
 124. Gerspacher M, Altmann E, Beerli R, Buhl T, Endres R, Gamse R, et al. Penta-substituted benzimidazoles as potent antagonists of the calcium-sensing receptor (CaSR-antagonists). *Bioorganic Med Chem Lett.* 2010;

125. Yoshida M, Mori A, Inaba A, Oka M, Makino H, Yamaguchi M, et al. Synthesis and structure-activity relationship of tetrahydropyrazolopyrimidine derivatives--a novel structural class of potent calcium-sensing receptor antagonists. *Bioorg Med Chem* [Internet]. 2010 Dec 15 [cited 2018 May 10];18(24):8501–11. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0968089610009582>
126. Yoshida M, Mori A, Morimoto S, Kotani E, Oka M, Notoya K, et al. Novel and potent calcium-sensing receptor antagonists: Discovery of (5R)-N-[1-ethyl-1-(4-ethylphenyl)propyl]-2,7,7-trimethyl-5-phenyl-4,5,6,7-tetrahydropyrazolo[1,5-a]pyrimidine-3-carboxamide monotosylate (TAK-075) as an orally active bone anabolic agent. *Bioorganic Med Chem*. 2011;
127. Shinagawa Y, Inoue T, Katsushima T, Kiguchi T, Ikenogami T, Ogawa N, et al. Discovery of a potent and short-acting oral calcilytic with a pulsatile secretion of parathyroid hormone. *ACS Med Chem Lett*. 2011;
128. Johansson H, Boesgaard MW, Nørskov-Lauritsen L, Larsen I, Kuhne S, Gloriam DE, et al. Selective Allosteric Antagonists for the G Protein-Coupled Receptor GPRC6A Based on the 2-Phenylindole Privileged Structure Scaffold. *J Med Chem*. 2015;
129. Liang GB, Zhou C, Huo X, Wang H, Yang X, Huang S, et al. Discovery of novel dihydrobenzofuran cyclopropane carboxylic acid based calcium sensing receptor antagonists for the treatment of osteoporosis. *Bioorg Med Chem Lett* [Internet]. 2016;26(16):4077–80. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27397499>
130. Miyazaki H, Ikeda Y, Sakurai O, Miyake T, Tsubota R, Okabe J, et al. Discovery of evocalcet, a next-generation calcium-sensing receptor agonist for the treatment of hyperparathyroidism. *Bioorganic Med Chem Lett*. 2018;
131. Kessler A, Faure H, Petrel C, Ruat M, Dauban P, Dodd RH. N2-Benzyl-N1-(1-(1-naphthyl)ethyl)-3-phenylpropane-1, 2-diamines and conformationally restrained indole analogues: Development of calindol as a new calcimimetic acting at the calcium sensing receptor. *Bioorganic Med Chem Lett*. 2004;
132. Poon SF, St Jean DJ, Harrington PE, Henley C, Davis J, Morony S, et al.

- Discovery and optimization of substituted 1-(1-phenyl-1H-pyrazol-3-yl) methanamines as potent and efficacious type II calcimimetics. *J Med Chem.* 2009;
133. Harrington PE, St. Jean DJ, Clarine J, Coulter TS, Croghan M, Davenport A, et al. The discovery of an orally efficacious positive allosteric modulator of the calcium sensing receptor containing a dibenzylamine core. *Bioorganic Med Chem Lett.* 2010;
 134. Gustafsson M, Jensen J, Bertozzi SM, Currier EA, Ma JN, Burstein ES, et al. Discovery of a class of calcium sensing receptor positive allosteric modulators; 1-(benzothiazol-2-yl)-1-phenylethanol. *Bioorganic Med Chem Lett.* 2010;
 135. Temal T, Jary H, Auberval M, Lively S, Guédin D, Vevvert JP, et al. New potent calcimimetics: I. Discovery of a series of novel trisubstituted ureas. *Bioorganic Med Chem Lett.* 2013;
 136. Wehn PM, Harrington PE, Carlson TJ, Davis J, Deprez P, Fotsch CH, et al. Metabolism-guided discovery of a potent and orally bioavailable urea-based calcimimetic for the treatment of secondary hyperparathyroidism. *Bioorganic Med Chem Lett.* 2013;
 137. Kiefer L, Beaumard F, Gorojankina T, Faure H, Ruat M, Dodd RH. Design and synthesis of calindol derivatives as potent and selective calcium sensing receptor agonists. *Bioorganic Med Chem.* 2016;
 138. Sparks SM, Spearing PK, Diaz CJ, Cowan DJ, Jayawickreme C, Chen G, et al. Identification of potent, nonabsorbable agonists of the calcium-sensing receptor for GI-specific administration. *Bioorganic Med Chem Lett.* 2017;27(20):4673–7.
 139. US9253997B2 - Alkylamine derivative - Google Patents [Internet]. [cited 2020 Nov 30]. Available from: <https://patents.google.com/patent/US9253997>

6 Declaration

At this moment, I declare that this thesis “IN-SILICO SCREENING and EXPERIMENTAL VALIDATION of NEW ALLOSTERIC MODULATORS for the CALCIUM-SENSING RECEPTOR” was composed by myself and has not been submitted for any previous degree. The work presented here is my own except where indicated by references.

7 Abbreviation

BARF	binding affinity relevant fragment
CaSR	extracellular calcium-sensing receptor
pK	binding affinity
IP-1	inositol-1-phosphate
GPCR	G-protein coupled receptor
Ca ²⁺ _e	extracellular calcium
PTH	parathyroid hormone
PAMs	positive allosteric modulators
NAMs	negative allosteric modulators
ECD	extracellular domain
7TM	seven alpha-helical transmembrane domain
CRD	cysteine-rich domain,
TMD	transmembrane domain,
ICD	intracellular domain
GABA-B	gamma aminobutyric acid receptor
PLC	phospholipase C
PLD	phospholipase D
GEF	GTP exchange factor
ADH type 1	Autosomal dominant hypocalcaemia
FHH1	Familial hypocalciuric hypercalcemia
NSHPT	Neonatal severe primary hyperparathyroidism
SHPT	secondary hyperparathyroidism
CKD	chronic kidney disease
DMSO	dimethyl sulfoxide
HPT	hyperparathyroidism
VS	virtual screening
TCM	traditional Chinese medicines
EL	extracellular loops
ICL	intracellular loops

mGlu ₁	metabotropic glutamate receptor 1
mGlu ₅	metabotropic glutamate receptor 5
HTRF	homologous time-resolved fluorescence
HEK293	human embryonic kidney cells

8 Table of Figures

Figure 1.1: The extracellular calcium-sensing receptor is mainly expressed in the parathyroid gland and kidney	10
Figure 1.2: Calcium-response curve of CaSR; positive allosteric modulators (PAMs) cause a potency shift to lower calcium levels and negative allosteric modulators a shift to higher levels inducing a change of CaSR's setpoint	12
Figure 1.3: essential amino acids; yellow: hydrophobic; green: neutral; blue: alkaline; red: acidic	12
Figure 1.4: Primary sequence of the human extracellular calcium-sensing receptor (hCaSR) in FASTA format retrieved from UniProt (8) (http://www.uniprot.org/uniprot/P41180.fasta), cyan: signal peptide, yellow: cysteine part of a cysteine-bridge, green: transmembrane helices	13
Figure 1.5: Schematic representation of CaSR's structure in the inactive state (homology model created with YASARA); ECD: extracellular domain, CRD: cysteine-rich domain, TMD: transmembrane domain, ICD: intracellular domain	14
Figure 1.6: GPCRs overview: The human GPCRs are clustered into different protein families; class A, B, C, frizzled/smoothed, taste type 2 and ocular albinism; The GPCR class C and the calcium-sensing receptor (green); (phylogenetic tree created with 831 GPCR entries from UniProt (8) at GPCRdb with the UPGMA distance calculation method (11))	15
Figure 1.7: G-protein activation and deactivation circle	16
Figure 1.8 allosteric modulators of CaSR; first row: calcimimetics type II; second row: calcilytics	19
Figure 2.1 DrugBank database version 5.0.10 with 8486 small molecule entries	22
Figure 2.2 Snake diagrams showing the transmembrane of the extracellular calcium-sensing receptor and the interacting residues with allosteric modulators.	

black: most conserved residue and a reference point in each transmembrane for class C GPCRs defined as position X.50 according to the Ballesteros-Weinstein numbering system (27); A) green: interacting residues with AC-265347, B) blue: interacting residues with cinacalcet, C) red: interacting residues with NPS-2143 according to the study (26).....25

Figure 2.3 sequence alignment results of I-TASSER including the top-scoring threading templates with the primary sequence of CaSR; sed. str.: secondary structure prediction of CaSR sequence with regions denoted as random coil (C), helix (H) or beta-sheet (S); colouring scheme: ClustalX (40) 31

Figure 2.4 I-TASSER results of predicted full-length CaSR models, c-score: confidence score representing the model quality in the range between -5 to 2 where better models have higher values 32

Figure 2.5 Residue-specific quality (RSQ) estimating the deviation between the predicted model and the native structure resulting from I-TASSER 33

Figure 2.6 sequence alignment results of GPCR-I-TASSER including the top-scoring threading templates with the primary sequence of CaSR; sed. str.: secondary structure prediction of CaSR sequence with regions denoted as random coil (C), helix (H) or beta-sheet (S); colouring scheme: ClustalX (40) 36

Figure 2.7 GPCR-I-TASSER results of predicted full-length CaSR models, c-score: confidence score representing the model quality in the range between -5 to 2 where better models have higher values 37

Figure 2.8 Residue-specific quality (RSQ) estimating the deviation between the predicted model and the native structure resulting from GPCR-I-TASSER..... 37

Figure 2.9 Heatmap of the secondary structure prediction of CaSR's sequence with PSI-Pred (47); the colour intensity reflects the estimated probability of each secondary structure element graded between 0 (minimum) and 9 (maximum)).. 39

Figure 2.10 Sequence alignment between the 17 generated homology models and the primary sequence of the full-length CaSR; Sec. str.: secondary structure predicted with PSI-Pred (46) 41

Figure 2.11 Final model of the homology modelling process based on the full-length CaSR sequence and created with YASARA Structure (41)..... 44

Figure 2.12: Intermediate homology models based on CaSR's primary sequence and generate with YASARA structure, results are shown in two columns, on the left is a graphic of the model and on the right the calculated model quality per

residue (Z-score); Z-score < -2: poor; Zscore < -4: bad; blue: CaSR protomer A; red: CaSR protomer B, cyan and grey: incorrect homology model part based on the T-lysozyme.....	45
Figure 2.13 Sequence alignment between the seven individual helices of CaSR and the templates 4OO9-A, 4OR2-A, 4OR2-B, 5CGC-A and 5CGD-A (created with Jalview 2.11.1.3) (40).....	47
Figure 2.14 Sequence alignment of the resulting homology models with the sequence of CaSR's seven transmembrane alpha-helices; sec. str.: secondary structure prediction with PSI-PRED (46); CaSR_TM: target sequence of the calcium-sensing receptor seven alpha-helical structures of the transmembrane; blue underlined vertical lines (): delimitators between the seven helical sequences; red: incorrectly introduced gaps	48
Figure 2.15 Relationship between sequence identity and sequence length to create a reliable homology model (modified image from (53)); dots: sequence identity between metabotropic glutamate receptor 1 (GRM1) or metabotropic glutamate receptor 5 (GRM5) to CaSR	49
Figure 2.16 estimated model quality per residue of the homology models created for the CaSR transmembrane helices	50
Figure 2.17 sequence alignment between the loops of CaSR, metabotropic glutamate receptor 1 (GRM1) and 5 (GRM5); ECD: extracellular domain residues; ICL: intracellular loop; ECL: extracellular loop	51
Figure 2.18 CaSR - NPS-2143 receptor complex with an open conformation for extracellular loop 2, A) NPS-2143 transmembrane complex after structural alignment of CaSR with GRM1, purple: binding pocket; B) interaction profile of NPS-2143 with CaSR, yellow: hydrophobic interactions, green: hydrogen donor interaction; blue: aromatic $\pi - \pi$ interaction	53
Figure 2.19 Quality check of CaSR transmembrane homology model; A) Z-score quality check per residue; B) Ramachandran Plot with PROCHECK (58)	54
Figure 2.20 model refinement of CaSR's homology model in the inactive state with YASARA's refinement macro (24).....	55
Figure 2.21 CaSR model validation with 363 negative allosteric modulators. Pearson correlation coeff.: 0.16;	56
Figure 2.22 RMSD of three 5 ns long molecular dynamics simulations with the transmembrane model	58

Figure 2.23 Results of three 5ns molecular dynamics simulations; A) total potential energy of the system as a function of simulation time; B) radius of gyration representing the fluctuations of all model atoms around the centre of mass; C) number of hydrogen bonds within the model.....	59
Figure 2.24 Prospective study with eleven conformers of CaSR's transmembrane homology model by applying docking experiments with the validation set of 363 negative allosteric modulators; r: Pearson correlation (Equation 5)	60
Figure 2.25 In-silico profiling with SAFAN-ISPSM; A) leave-one-out-cross-validation of the method with over 13 000 maximum diverse compounds randomly selected from the ChEMBL23 database (DOI: 10.6019/CHEMBL.database.23); [this image was provided by S.A.F.AN.-BIOINFORMATICS (70)]; B) leave-one-out-cross-validation with 332 known negative allosteric modulators (NAM) for CaSR	64
Figure 2.26 Venn diagram of 396 compounds interacting with CaSR and the pre-selection process including the parameters target position, pK and round.....	66
Figure 2.27 Reference compound tecalcet and NPS-2143 for the experimental validation with an IP1-Gq assay	66
Figure 2.28 CaSR Gq activation and IP1 accumulation, GDP: guanosine diphosphate; GTP; guanosine triphosphate; PLC β : phospholipase C β ; IP ₃ : D-myoinositol 1,4,5-triphosphate; IP ₂ : inositol-4-5-phosphate; IP ₁ : inositolphosphate	68
Figure 2.29 IP-1 accumulation assay to determine the agonistic activity of the ligands.....	69
Figure 2.30 IP1 accumulation assay with the first compound library at 2.5 mM calcium to identify positive allosteric modulators of CaSR	70
Figure 2.31 IP1 accumulation assay with the first compound library at 7.5 mM calcium to identify negative allosteric modulators of CaSR (normalized to no compound)	70
Figure 2.32 Comparison of bucindolol with known allosteric modulators reveals a common butrinolol scaffold [2-(3-(tert-Butylamino)-2-hydroxypropoxy)benzotrile] (blue).....	72
Figure 2.33 keto-enol tautomerism of 2-pyridone.....	73
Figure 2.34 SAFAN-ISPSM leave-one-out cross-validation of 21090 diverse compounds involving 65936 interactions (Pearson Correlation $r = 0.91$) (70).....	74

<i>Figure 2.35</i>	<i>Leave-one-out cross-validation of the ChEMBL dataset for CaSR-ligand interactions.....</i>	<i>74</i>
<i>Figure 2.36</i>	<i>Proof of concept validation with the ChEMBL data set for CaSR-ligand interactions.....</i>	<i>75</i>
<i>Figure 2.37</i>	<i>Homology modelling process of the full-length CaSR – Gq protein complex</i>	<i>79</i>
<i>Figure 2.38</i>	<i>Structural alignment of CaSR, mGlu1 and mGlu5 sequence acquired from GPCRdb (11), blue: sequence identity at the aligned position, boxes represent the transmembrane helices.....</i>	<i>80</i>
<i>Figure 2.39</i>	<i>Multiple sequence alignment for homology modelling of the ICD truncated CaSR in the inactive conformation; template sequences of CaSR-ECD, mGlu1 and mGlu5 transmembrane were aligned according to the multiple structure alignment available of the GPCRdb website (Figure 2.38). ECL1, ICL2 and ECL3 were omitted in the template structures.</i>	<i>83</i>
<i>Figure 2.40</i>	<i>Model quality of the created homology models; A) Z-score quality of the extracellular domain; B) Z-score quality of the seven-transmembrane domain; C) Z.score quality of the ICD truncated CaSR</i>	<i>84</i>
<i>Figure 2.41</i>	<i>Geometric comparison of the final model to crystal structures A) Final homology model generated with YASARA based on the predefined multiple sequence alignment from GPCRdb (11), ECD: extracellular domain, CRD: cysteine-rich domain, TM: transmembrane domain, the CRDs of the homodimer opens up by 93° using D588 from protomer A and H254 and D588 from protomer B as reference points; blue: initial model, different colours: hybrid parts; B) the ECD of CaSR (blue) is staged in the crystal structure PDB: 5K5T, the CRD interacts with the ECD of a second homodimer (red) and with its CRD, forming an unnatural angle between the CRDs; C) crystal structure PDB: 6N52 representing mGlu5 in the inactive state, this structure contains the ECD, CRD and the TM, the geometry of this receptor suggests that the protomers are arranged parallel and that the TMs are not in direct contact, but the parallel orientation seems to be the consequence of the membrane.....</i>	<i>85</i>
<i>Figure 2.42</i>	<i>Hybrid homology model of the ICD truncated CaSR in the inactive conformation; A) Z-score model quality of the hybrid model protomers per residue; B) Ramachandran Plot of the hybrid model; residues in the disallowed region: ASN83, SER366; ARG396</i>	<i>86</i>

Figure 2.43 Hybrid homology model of the ICD truncated CaSR with an alternative geometry; A) Z-score model quality of the adapted hybrid model protomers per residue; B) Ramachandran Plot of the hybrid model; residues in disallowed regions: SER366 and ARG396	86
Figure 2.44 Sequence alignment of the heterotrimeric Gq-protein between the target sequence and the template (PDB: 3AH8)	87
Figure 2.45 Model quality of the Gq heterotrimeric complex with an overall Z-score of -0.776	88
Figure 2.46 CaSR Gq-protein complex after structural alignment to a rhodopsin Gi complex; CaSR homodimer (cyan, magenta); G alpha q: red; G beta: green; G gamma: blue; ECD: extracellular domain; TMD; seven-transmembrane domain	88
Figure 2.47 Primary sequence of ECD truncated CaSR	89
Figure 2.48 sequence alignment results of I-TASSER including the top-scoring threading templates with the primary sequence of ECD-truncated CaSR; sed. str.: secondary structure prediction of CaSR sequence with regions denoted as random coil (C), helix (H) or beta-sheet (S); colouring scheme: ClustalX (40)	90
Figure 2.49 Residue-specific quality (RSQ) of the ECD-truncated CaSR estimating the deviation between the predicted models and the native structure resulting from I-TASSER; TMD: transmembrane domain; ICD: intracellular domain	91
Figure 2.50 2.51 I-TASSER results of predicted ECD-truncated CaSR models, c-score: confidence score representing the model quality in the range between -5 to 2 where better models have higher values; blue: transmembrane domain; red: intracellular domain	92
Figure 2.52 The final model of the full-length extracellular calcium-sensing receptor in complex with the Gq-protein; cyan and magenta: CaSR homodimer, red: Gq alpha subunit, green: G beta subunit; blue: G gamma subunit.....	93
Figure 2.53 RMSD fluctuations of three 50ns long molecular dynamics simulations of the CaSR-Gq protein complex transmembrane domain.....	93
Figure 2.54 CaSR model complex validation with 368 negative allosteric modulators of different; r: Pearson coefficient (Equation 5).....	94
Figure 2.55 Structural alignment for homology modelling of CaSR in the active conformation, TM: transmembrane alpha-helix, blue: sequence identity.....	96

Figure 2.56 Homology model in the active conformation A) final ICD-truncated CaSR model; A) B) quality Z-score per residue of the ICD-truncated CaSR model	97
Figure 2.57 Structural alignment of CaSR transmembrane domain to mGlu ₅ .	97
Figure 2.58 Homology model of the full-length CaSR in the active state; A) depiction of the homodimer model; B) quality Z-score per residue of CaSR in the active state	98
Figure 2.59 RMSD fluctuations of three 50ns long molecular dynamics simulations of the CaSR transmembrane (residue 601-863) in the active conformation	98
Figure 2.60 Model validation of seven CaSR conformers in the active state with the ChEMBL validation set of 236 positive allosteric modulators	99
<i>Figure 2.61 structure-based pharmacophore models of CaSR's allosteric binding site; green: residues identified by Cavity 1.0 to participate in the binding pocket. red: hydrogen acceptor, blue: hydrogen donor, cyan: hydrophobic (representation done with YASARA); A1 and A2 are the pharmacophores of the active model, with different feature selection, B inactive model, C active pre-docked model, D inactive pre-docked model</i>	<i>101</i>
Figure 2.62 new compound library for the second profiling attempt with SAFAN-ISPSM 2019/06-b2.....	103
Figure 2.63 selection process of the screening results for experimental validation	104
Figure 2.64 Multidimensional scaling of the remaining screening compounds (n = 39) with known CaSR ligands (n = 908) to identify dissimilar compounds	105
Figure 2.65 SAFAN similarity analysis of divers screening compounds to known CaSR ligands; green: selected screening compounds for experimental validation with an IP-1 Gq HTRF assay	106
Figure 2.66 IP1 standard curve with IP1 standards (CisBio).....	109
Figure 2.67 IP-1 accumulation results of 8 screening compounds against the positive allosteric modulator NPS-R568 and negative allosteric modulator NPS-2143, A) measurement in HEK293 HA hCASR WT cells, B) experiments in HEK293 WT cells; data represents mean ± standard deviation of 6 independent experiments in triplicates, The statistical analysis was performed with one-way ANOVA with Dunnett's post-test vs vehicle. C) IP-1 accumulation assay at different calcium concentrations shows a potency shift of the positive allosteric modulator	

NPS-R568 to the right as well as for compound 7, indicating positive allosteric activity at CaSR; D) IP1 accumulation assay at different compound concentrations indicate a weaker allosteric modulator activity of compound 7 in comparison to the standard NPS-568.....	110
Figure 2.68 Compound similarity between the approved Evocalcet and the discovered positive allosteric modulator.....	111
Figure 4.1 YASARA homology modelling workflow, the description is in the text, PSI-BLAST: position-specific iterative basic local alignment search tool, PSSM: position-specific scoring matrix, PDB: Protein Data Bank, HM: homology modelling; PSSP: profiles from sequence and structurally related proteins	116
Figure 4.2 BLASTp workflow; A: creation of high scoring words with fixed length (e.g. W = 3); B: scoring matches according to BLOSUM62 and selection over cut-off T; C) extension of the high scoring word to the final segment score and sorting according to the E-value (105,106)	118
Figure 4.3 BLOSUM62 substitution matrix for proteins; The description of the rate a character changes in the sequence over time. (107,108)	118
Figure 4.4 ChEMBL180463, interaction map of the most potent 4(3H)-quinazolinones (IC ₅₀ = 190 nM) docked to CaSRs homology model on page 53	125
Figure 4.5 ChEMBL195711, interaction map of the most potent 4(3H)-pyrimidinone (IC ₅₀ = 95 nM) docked to CaSRs homology model on page 53..	126
Figure 4.6: impact of structural modifications on the structure-activity-relationship of NPS-2143 (118,119)	126
Figure 4.7 ChEMBL372820, interaction map of the most potent (R)-1-arylmethylpyrrolidin-2-yl ethanolamines (IC ₅₀ = 140 nM) docked to CaSRs homology model on page 53	127
Figure 4.8 ChEMBL200312, interaction map of the most potent pyrrolidin-1-yl[methyl]benzo-nitriles (IC ₅₀ = 270 nM) docked to CaSRs homology model on page 53	128
Figure 4.9 ChEMBL536975, interaction map of the most potent N1-Benzoyl-N2-[1-(1-naphthyl)ethyl]-trans-1,2-diaminocyclohexanes (IC ₅₀ = 250 nM) docked to CaSRs homology model on page 53.....	129

Figure 4.10 CHEMBL457249, interaction map of the most potent trisubstituted pyrimidines/pyridines; (IC ₅₀ = 60 nM) docked to CaSRs homology model on page 53.....	130
Figure 4.11 CHEMBL1783764, interaction map of the most potent phenoxy analogues of NPS-2143 and ronacalceret (IC ₅₀ = 6 nM) docked to CaSRs homology model on page 53.....	131
Figure 4.12 CHEMBL2112075, interaction map of the most potent amino alcohol-based parathyroid hormone secretagogues (IC ₅₀ = 3 nM) docked to CaSRs homology model on page 53.....	132
Figure 4.13 CHEMBL382741, interaction map of the most potent amino alcohol-based parathyroid hormone secretagogues (IC ₅₀ = 0.003 nM) docked to CaSRs homology model on page 53.....	133
Figure 4.14 CHEMBL597397, interaction map of the most potent 1-Alkyl-4-phenyl-6-alkoxyquinazolin-2(1H)-one (IC ₅₀ = 0.4 nM) docked to CaSRs homology model on page 53.....	135
Figure 4.15 CHEMBL2368751, interaction map of the most potent parathyroid hormone mimetics (IC ₅₀ = 25 nM) docked to CaSRs homology model on page 53	136
Figure 4.16 CHEMBL1082788, interaction map of the most potent aminopropanediol derivatives; (IC ₅₀ = 37 nM) docked to CaSRs homology model on page 53.....	138
Figure 4.17 CHEMBL1223772, interaction map of the most potent penta-substituted benzimidazoles (IC ₅₀ = 2 nM) docked to CaSRs homology model on page 53.....	139
Figure 4.18 CHEMBL1290328, interaction map of the most potent tetrahydropyrazolopyrimidine derivatives (IC ₅₀ = 10 nM) docked to CaSRs homology model on page 53.....	140
Figure 4.19 CHEMBL1689060, interaction map of the most potent tetrahydropyrazolopyrimidine (IC ₅₀ = 0.94 nM) docked to CaSRs homology model on page 53.....	141
Figure 4.20 CHEMBL1688086, interaction map of the most potent short-acting oral calcilytics (IC ₅₀ = 12 nM) docked to CaSRs homology model on page 53.....	142

Figure 4.21 CHEMBL1672970, interaction map of the most potent aminopropandiol derivates (IC ₅₀ = 4 nM) docked to CaSRs homology model on page 53	143
Figure 4.22 CHEMBL3633656, interaction map of the most potent 2-phenylindole (IC ₅₀ = 56200 nM) docked to CaSRs homology model on page 53	144
Figure 4.23 CHEMBL3827736, interaction map of the most potent dihydro benzofuran cyclopropane derivates IC ₅₀ = 20 nM) docked to CaSRs homology model on page 53	145
Figure 4.24 CHEMBL4126450, interaction map of the most potent evocalcet derivates (IC ₅₀ = 0,47 nM) docked to CaSRs homology model on page 53	146
Figure 4.25 histogram of residue interactions of CaSR transmembrane with the top-scoring negative allosteric modulators from 21 different scaffolds, docking done with AutoDock VINA, model was used from page 46, interaction profile (CHEMBL180463; CHEMBL195711; CHEMBL195711; CHEMBL195711; CHEMBL200312; CHEMBL536975; CHEMBL457249; CHEMBL1783764; CHEMBL2112075; CHEMBL382741; CHEMBL597397; CHEMBL2368751; CHEMBL1082788; CHEMBL1223772M; CHEMBL1290328; CHEMBL1689060; CHEMBL1688086; CHEMBL1672970; CHEMBL3633656; CHEMBL3827736; CHEMBL4126450)	146
Figure 4.26 CHEMBL2092942, interaction map of (R) calindol, IC ₅₀ = 310 nM) docked to CaSRs homology model on page 53	147
Figure 4.27 CHEMBL569182, interaction map of the most potent 1-(1-Phenyl-1H-pyrazol-3-yl)methanamines (IC ₅₀ = 23 nM) docked to CaSRs homology model on page 53	148
Figure 4.28 CHEMBL1224424, interaction map of the most potent dibenzylamine (IC ₅₀ = 17 nM) docked to CaSRs homology model on page 53	149
Figure 4.29 CHEMBL1200776, interaction map of cinacalcet (IC ₅₀ = 19.95 nM) docked to CaSRs homology model on page 53	149
Figure 4.30 CHEMBL1256368, interaction map of the most potent 1-(benzothiazole-2-yl)1-phenylethanol (IC ₅₀ = 10 nM) docked to CaSRs homology model on page 53	150
Figure 4.31 CHEMBL2346788, interaction map of the most potent trisubstituted ureas (IC ₅₀ = 60 nM) docked to CaSRs homology model on page 53	152

Figure 4.32 CHEMBL2377740, interaction map of the most potent trisubstituted ureas (IC ₅₀ = 20 nM) docked to CaSRs homology model on page 53	154
Figure 4.33 CHEMBL3093422, interaction map of the most potent urea-based calcimimetics (IC ₅₀ = 2 nM) docked to CaSRs homology model on page 53...	155
Figure 4.34 CHEMBL3753306, interaction map of the most potent calindol derivates (IC ₅₀ = 20 nM) docked to CaSRs homology model on page 53	156
Figure 4.35 CHEMBL4100132, interaction map of the most potent Evocalcet derivates (IC ₅₀ = 15.85 nM) docked to CaSRs homology model on page 53 ..	157
Figure 4.36 CHEMBL4126450, interaction map of the most potent 3-aminopyrrolidine (IC ₅₀ = 17 nM) docked to CaSRs homology model on page 53	158
Figure 4.37 CHEMBL398092, interaction map of the most potent alkylamine derivatives; (IC ₅₀ = 3 nM) docked to CaSRs homology model on page 53.....	159
Figure 4.38 CHEMBL3974015, interaction map of the most potent alkylamine derivatives (IC ₅₀ = 39 nM) docked to CaSRs homology model on page 53.....	159
Figure 4.39 CHEMBL3895461, interaction map of the most potent alkylamine derivatives (IC ₅₀ = 4 nM) docked to CaSRs homology model on page 53.....	160
Figure 4.40 histogram of residue interactions of CaSR transmembrane with the top-scoring positive allosteric modulators from 16 different scaffolds, docking done with AutoDock VINA, model was used from page 46, interaction profile; CHEMBL2092942; CHEMBL292376; CHEMBL569182; CHEMBL180672; CHEMBL1224424; CHEMBL1200776; CHEMBL1256368; CHEMBL2346788; CHEMBL2377740; CHEMBL3093422; CHEMBL3753306; CHEMBL4100132; CHEMBL4126450; CHEMBL3980928; CHEMBL3974015; CHEMBL3895461 .	161
Figure 4.41 IP1 standard curve with IP1 standards (CisBio).....	165

9 Appendix:

appx. 9.1 Txt files of the generated pharmacophore models

Pharmacophore A.txt file (CaSR model active, undocked):

```
#
# This file is created by LigBuilder/Pocket
# It lists the pharmacophore features detected
# Creation time: Wed Jun 26 15:35:18 2019
#
# The 1st column: feature ID number
# The 2nd column: feature property
# The 3rd column: feature symbol
# The 4th column: X coordinate
# The 5th column: Y coordinate
# The 6th column: Z coordinate
# The 7th column: average score
# The 8th column: neighbouring grids around this feature
# The 9th column: Depth of the feature
#
<Start
<Feature_description
1 H-bond_donor_site D -87.00 72.00 19.50 66.58 106 8
2 H-bond_donor_site D -87.50 70.00 15.00 48.54 79 12
3 H-bond_donor_site D -88.50 68.00 20.00 80.39 132 12
4 H-bond_donor_site D -96.00 70.00 21.00 52.06 86 23
5 H-bond_donor_site D -85.00 68.50 20.50 36.17 60 6
6 H-bond_donor_site D -85.00 72.00 16.50 105.01 175 6
7 H-bond_donor_site D -94.50 73.00 20.00 41.84 73 23
8 Hydrophobic_site H -94.50 68.50 24.00 114.92 249 23
9 Hydrophobic_site H -90.50 71.00 20.00 32.68 73 15
10 H-bond_acceptor_site A -87.00 66.50 15.00 29.53 66 14
11 Hydrophobic_site H -97.00 71.00 24.50 21.36 48 25
<Feature_internal_distance
1 2 3 4 5 6 7 8 9 10 11
1 ----- 4.95 4.30 9.34 4.15 3.61 7.58 9.42 3.67 7.11 11.22
2 4.95 ----- 5.48 10.40 6.22 3.54 9.11 11.50 5.92 3.54 13.47
3 4.30 5.48 ----- 7.83 3.57 6.36 7.81 7.23 3.61 5.43 10.07
4 9.34 10.40 7.83 ----- 11.11 12.05 3.50 3.67 5.68 11.37 3.77
5 4.15 6.22 3.57 11.11 ----- 5.32 10.52 10.12 6.06 6.18 12.89
6 3.61 3.54 6.36 12.05 5.32 ----- 10.17 12.60 6.60 6.04 14.46
7 7.58 9.11 7.81 3.50 10.52 10.17 ----- 6.02 4.47 11.11 5.52
8 9.42 11.50 7.23 3.67 10.12 12.60 6.02 ----- 6.18 11.88 3.57
9 3.67 5.92 3.61 5.68 6.06 6.60 4.47 6.18 ----- 7.58 7.91
10 7.11 3.54 5.43 11.37 6.18 6.04 11.11 11.88 7.58 ----- 14.51
11 11.22 13.47 10.07 3.77 12.89 14.46 5.52 3.57 7.91 14.51 -----
<End
```

Pharmacophore B.txt file (CaSR model inactive, undocked):

```
#
# This file is created by LigBuilder/Pocket
# It lists the pharmacophore features detected
# Creation time: Thu Jun 27 12:00:24 2019
#
# The 1st column: feature ID number
# The 2nd column: feature property
# The 3rd column: feature symbol
# The 4th column: X coordinate
# The 5th column: Y coordinate
# The 6th column: Z coordinate
# The 7th column: average score
```



```

# The 8th column: neighbouring grids around this feature
# The 9th column: Depth of the feature
#
<Start
<Feature_description
  1 H-bond_donor_site      D   -32.50   -9.00    7.50   106.21  160  5
  2 H-bond_donor_site      D   -25.50   -7.50    3.00    67.48  106  2
  3 H-bond_donor_site      D   -28.50   -7.00    9.50    33.10  56  7
  4 H-bond_donor_site      D   -29.00   -9.50    6.50    41.81  73  3
  5 H-bond_acceptor_site   A   -32.50   -7.00    4.50    28.52  56  3
  6 H-bond_acceptor_site   A   -33.00  -12.50    7.50    37.16  77  1
  7 Hydrophobic_site       H   -29.00   -7.00    4.00    40.50  97  3
<Feature_internal_distance
      1      2      3      4      5      6      7
  1 -----  8.46  4.90  3.67  3.61  3.54  5.34
  2  8.46 -----  7.18  5.34  7.18 10.07  3.67
  3  4.90  7.18 -----  3.94  6.40  7.38  5.52
  4  3.67  5.34  3.94 -----  4.74  5.10  3.54
  5  3.61  7.18  6.40  4.74 -----  6.28  3.54
  6  3.54 10.07  7.38  5.10  6.28 -----  7.65
  7  5.34  3.67  5.52  3.54  3.54  7.65 -----
<End

```

Pharmacophore C.txt file (CaSR model active, pre-docked):

```

#
# This file is created by LigBuilder/Pocket
# It lists the pharmacophore features detected
# Creation time: Wed Jun 26 15:49:19 2019
#
# The 1st column: feature ID number
# The 2nd column: feature property
# The 3rd column: feature symbol
# The 4th column: X coordinate
# The 5th column: Y coordinate
# The 6th column: Z coordinate
# The 7th column: average score
# The 8th column: neighbouring grids around this feature
# The 9th column: Depth of the feature
#
<Start
<Feature_description
  1 H-bond_donor_site      D   -86.00   66.00   19.50   83.17  128  13
  2 H-bond_donor_site      D   -83.00   68.50   19.50  104.69  167  9
  3 H-bond_donor_site      D   -83.00   65.00   21.00   38.05  61  8
  4 H-bond_donor_site      D   -84.00   66.50   16.00   26.24  43  12
  5 H-bond_donor_site      D   -92.50   70.50   17.50   52.18  86  23
  6 H-bond_donor_site      D   -80.50   67.50   17.00   86.64  144  7
  7 H-bond_donor_site      D   -89.50   72.00   16.50   32.12  56  18
  8 Hydrophobic_site       H   -80.50   65.50   23.50   85.18  184  1
  9 H-bond_acceptor_site   A   -84.00   63.00   16.50   34.07  74  14
 10 Hydrophobic_site       H   -87.00   69.00   18.00   27.04  60  15
 11 Hydrophobic_site       H   -93.50   70.00   21.00   60.43  137  27
 12 Hydrophobic_site       H   -93.50   73.50   20.50   21.94  51  28
<Feature_internal_distance
      1      2      3      4      5      6      7      8      9      10      11      12
  1 -----  3.91  3.50  4.06  8.15  6.22  7.57  6.82  4.69  3.50  8.63  10.65
  2  3.91 -----  3.81  4.15  9.91  3.67  7.97  5.59  6.34  4.30  10.71  11.67
  3  3.50  3.81 -----  5.32 11.52  5.34 10.56  3.57  5.02  6.40  11.63  13.52
  4  4.06  4.15  5.32 -----  9.51  3.77  7.79  8.34  3.54  4.39  11.29  12.63
  5  8.15  9.91 11.52  9.51 ----- 12.38  3.50 14.32 11.38  5.72  3.67  4.36
  6  6.22  3.67  5.34  3.77 12.38 ----- 10.07  6.80  5.72  6.75 13.83 14.74
  7  7.57  7.97 10.56  7.79  3.50 10.07 ----- 13.12 10.55  4.18  6.34  5.85
  8  6.82  5.59  3.57  8.34 14.32  6.80 13.12 -----  8.22  9.21 13.98 15.56
  9  4.69  6.34  5.02  3.54 11.38  5.72 10.55  8.22 -----  6.87 12.63 14.71

```

```

10  3.50  4.30  6.40  4.39  5.72  6.75  4.18  9.21  6.87  -----  7.23  8.29
11  8.63 10.71 11.63 11.29  3.67 13.83  6.34 13.98 12.63  7.23  -----  3.54
12 10.65 11.67 13.52 12.63  4.36 14.74  5.85 15.56 14.71  8.29  3.54  -----
<End

```

Pharmacophore D.txt (CaSR model inactive, pre-docked):

```

#
# This file is created by LigBuilder/Pocket
# It lists the pharmacophore features detected
# Creation time: Wed Jun 26 15:52:44 2019
#
# The 1st column: feature ID number
# The 2nd column: feature property
# The 3rd column: feature symbol
# The 4th column: X coordinate
# The 5th column: Y coordinate
# The 6th column: Z coordinate
# The 7th column: average score
# The 8th column: neighbouring grids around this feature
# The 9th column: Depth of the feature
#
<Start
<Feature_description
 1 H-bond_donor_site      D   -32.50   -9.00    7.50  106.21  160  5
 2 H-bond_donor_site      D   -25.50   -7.50    3.00   67.48  106  2
 3 H-bond_donor_site      D   -28.50   -7.00    9.50   33.10   56  7
 4 H-bond_donor_site      D   -29.00   -9.50    6.50   41.81   73  3
 5 H-bond_acceptor_site   A   -32.50   -7.00    4.50   28.52   56  3
 6 H-bond_acceptor_site   A   -33.00  -12.50    7.50   37.16   77  1
 7 Hydrophobic_site       H   -29.00   -7.00    4.00   40.50   97  3
<Feature_internal_distance
      1      2      3      4      5      6      7
 1 -----  8.46  4.90  3.67  3.61  3.54  5.34
 2  8.46 -----  7.18  5.34  7.18 10.07  3.67
 3  4.90  7.18 -----  3.94  6.40  7.38  5.52
 4  3.67  5.34  3.94 -----  4.74  5.10  3.54
 5  3.61  7.18  6.40  4.74 -----  6.28  3.54
 6  3.54 10.07  7.38  5.10  6.28 -----  7.65
 7  5.34  3.67  5.52  3.54  3.54  7.65 -----
<End

```

10 Funding

This work is part of the Marie Skłodowska-Curie “CaSR Biomedicine” project and funded by the Horizon 2020 programme of the European Union under grant agreement 675228.

11 Publication

Co-Authorship:

Chavez-Abiega S, Mos I, Centeno PP, Elajnaf T, Schlattl W, Ward DT, et al. Sensing Extracellular Calcium – An Insight into the Structure and Function of the Calcium-Sensing Receptor (CaSR). Vol. 1131, Advances in Experimental Medicine and Biology. 2020.

Curriculum Vitae

Personal data

Name Wolfgang Schlattl
Date/ Place of Birth June 21st 1985 in Amstetten
Nationality Austria
LinkedIn www.linkedin.com/in/wolfgang-schlattl-research-in-drug-discovery

Education

10/2016 - present Early Stage Researcher and PhD student in Complex Systems in Life Science at S.A.F.AN. – Bioinformatics and University of Turin

02/2014 Graduation with honours in pharmacy University of Vienna, Austria
Thesis title: Protein Folding Mechanism – Analysing the Folding Mechanism of myPGK Utilizing Simulations and Experiments
Supervisor: Prof. Dr. Gerhard Ecker
(http://othes.univie.ac.at/31461/1/2014-02-01_0402261.pdf)

Work experience

01/2020 – present Pharmacist at Apotheke zum heiligen Josef, Stadtplatz 4, 2130 Laa an der Thaya (Austria)
05/2014 – 08/2015 Training for the pharmacist profession at the pharmacy Apotheke zur Madonna, Langobardenstraße 16, 1220 Vienna

Personal skills

Native language German
Foreign languages English (written and spoken), Italian (basics)

Technical skills MS- Office (Word, Powerpoint, Excel)
Molecular dynamics simulation with GROMACS, Desmond and YASARA
Visualization with VMD, YASARA
Homology modelling with Modeller, YASARA
Molecular modelling with MOE and YASARA
Pharmacophore modelling with LigBuilder2.0
OS: MS Windows, Linux, Mac, Debian

Research Areas

Homology modelling, Docking, Pharmacophore modelling, Molecular dynamics simulations, Virtual screening, Gq -IP1 assays, qualitative Gibbs free energy calculations