
# Finding novel TRPV4 inhibitors for treatment of pain and inflammation – a CADD approach

Master's thesis by

Linda Wilkman

Supervisor: Dr. Outi Salo-Ahen
Bioscience, Specialization in Pharmacy
Faculty of Science and Engineering
Åbo Akademi University
Åbo, Finland 2021

**Acknowledgements**

I would like to thank my supervisor Dr. Outi Salo-Ahen for all the help, knowledge and patience during this process.

I would also like to thank my family and friends for continued love, support and encouragement.

**Abstract**

The human TRPV4 receptor is an intriguing target for drug development. Currently, there are no commercially available medicines targeting TRPV4. The goal for this thesis was to discover novel TRPV4 antagonists for the treatment of inflammation and pain. The experiments were performed using various computer-aided drug design methods.

A comparative model of the human TRPV4 was built by using the experimentally solved western clawed frog TRPV4 structure as the template structure.

Various methods were used for locating the putative antagonist binding site in human TRPV4. A cavity, corresponding to the antagonist binding site discovered in TRPV1 and TRPV5, called the vanilloid binding site, was selected for the docking studies.

The Enamine Ion Channel molecule library was used for virtual screening. The molecule library was docked into the putative binding cavity and binding free energy calculations were performed for the best ranked ligands from both methods. The ten best ranking molecules were submitted for 300-ns molecular dynamics simulations to analyze the stability of their binding interactions at the receptor binding pocket.

The same library was screened also using a pharmacophore model that was based on the structures of two known TRPV4 inhibitors, GSK205 and HC-067047.

Pharmacokinetic and toxicity properties of the ligands with the best binding qualities from the molecular dynamics simulations and the top-ranked hits from the pharmacophore-based screening were evaluated with SwissADME and ProTox II.

The ligands Z1213735368, Z1157726398 and Z221487176 from the docking study and the ligands Z443243482 and Z1728742868 from the pharmacophore study were chosen for further biological evaluation via our collaborators, to determine if the ligands act as human TRPV4 antagonists and if they reduce pain and inflammation in the body.

Keywords: Transient receptor potential vanilloid 4, TRPV4, CADD, virtual screening, molecular dynamics simulations, pharmacophore modeling.

**Table of contents**

## 1. Introduction

Many receptors of the transient receptor potential (TRP) channel family are targets of interest for drug development research. This project focuses on the transient receptor potential vanilloid 4 (TRPV4) receptor and its role in mediating pain and inflammation. Especially GlaxoSmithKline has strived to develop TRPV4 antagonists, but currently the molecules have not passed the clinical study phases. The aim of this thesis work was to find novel TRPV4 antagonists using a range of computer-aided drug design (CADD) methods. The selected ligands will then be purchased and sent to our collaborators for further evaluation.

With CADD, the aim is to determine ligands that are protein specific and are bound with high affinity. The interaction between the ligand and the protein must be energetically favored (Henrich et al., 2010). CADD methods are divided into structure-based drug design (SBDD) and ligand-based drug design (LBDD) methods (Yu & Mackerell, 2017). The thesis applies methods from both groups.

SBDD methods are used when the structure of the target, for example, a protein is known. It would be especially important to have access to the experimentally solved structures of membrane proteins since the group stands for most of the drug target proteins (Arinaminpathy, Khurana, Engelman, & Gerstein, 2009). However, it is generally demanding to determine crystal structures of membrane proteins, and thus, non-membrane proteins are often easier to work with. Nonetheless, based on available related structures, models of membrane proteins can be built by comparative (homology) modeling.

Virtual screening and docking approaches used in this thesis belong to the SBDD approaches as do the molecular dynamics (MD) simulations and the methods used for locating the putative binding site(s). Well known examples of medicines discovered with structure-based drug design methods include the HIV medicines raltitrexed and amprenavir and the osteoarthritis and rheumatoid arthritis medicine flurbiprofen (Batool, Ahmad, & Choi, 2019).

LBDD methods are used when the structure of the target protein or the binding site is unknown. Instead, the important features of known ligands are identified and then used

to build or identify new putative ligands (Yu & Mackerell, 2017). The LBDD method used in this thesis is the pharmacophore modeling approach.

It is essential to examine the ADMET (absorption, distribution, metabolism, excretion and toxicity) properties of a drug candidate. By predicting the ADMET properties in an early drug discovery phase, the drug candidate failure, related to pharmacokinetic properties in the later phases of drug discovery, is reduced drastically (Daina, Michielin, & Zoete, 2017). In this thesis, the ADMET properties of the most promising virtual hit compounds are predicted by using two online tools.

## 2. Literature review

### 2.1 TRP channels, pain and inflammation

TRP channels work as ion channels in the body. In mammals, the protein family is arranged into six subfamilies: ankyrin (TRPA), canonical (TRPC), melastatin (TRPM), mucoliptin (TRPML), polycyctin (TRPP) and vanilloid (TRPV). In total, 33 TRP proteins have been identified in mammals (Nikolaev et al., 2019). This thesis concentrates on one specific TRP channel, the TRPV4 channel. TRPV4 is one of the six channels that belong to the TRPV subfamily (Baylie & Brayden, 2011).

The TRP channels link the nervous system with the environment. Various TRP channels convert stimuli that could be caused by mechanical, chemical (both innate and external) or thermal irritation into electrical signals. This is an important initial step in achieving the feeling of pain or hot and cold (Wang & Woolf, 2005).

The sensitivity towards pain is increased after tissue injury or inflammation. This leads to less contact with the damaged area, so that the damage can be healed with minimal distraction. TRP receptors in the nociceptor terminals increase the sensitivity towards pain after inflammation (Wang & Woolf, 2005).

Many TRP receptors have been a source of interest for drug design. The proteins are not only involved in mediating pain or inflammation, but in many diverse processes in the body. Thus, it is challenging to design medicines that target TRP receptors, due to the adverse effects that can occur when the receptor is activated or blocked at the wrong site in the body (Kaneko & Szallasi, 2014). TRPV1 is the most studied TRP receptor as it is considered an interesting target in the drug development community (Samanta, Hughes, & Moiseenkova, 2018).

It has been discovered that pain can also be mediated by more than one TRP receptor at a certain system or site. TRPV4 and TRPA1 are located at the same places in the body. For example, they both mediate temporo-mandibular joint (TMJ) pain, pancreas pain and inflammation and pain of the large intestine. It appears that both TRPA1 and TRPV4 are also involved in osteoarthritis (Kanju et al., 2016).

## 2.2 The TRPV4 channel

### 2.2.1 The role of the channel

TRPV4 receptors can be activated directly by warm temperature (27–35°C), by changes in osmolarity and by external and internal chemical stimulation (Grace, Bonvini, Belvisi, & McIntyre, 2017).

TRPV4 receptors play an important role as mechanosensors and in signaling and sensing pain. The receptors can be found in afferent (sensory) neurons in the spinal ganglion and in the trigeminal ganglion (Kanju et al., 2016). Poole et al. (2013) found protease-activated receptor 2 (PAR-2), which is a significant actor in mediating pain and inflammation, to activate TRPV4 through a mechanism that is not yet fully understood. The complicated mechanism is believed also to include mediators such as arachidonic acid, phospholipase A2, epoxyeicosatrienoic acid and an unknown tyrosine kinase. This is one example of indirect activation of the receptor. In mice, TRPV4 has been identified to be a part of channeling pain caused by UVB radiation (Moore et al., 2013).

Besides being a pain mediator, TRPV4 is also involved with various respiratory diseases such as asthma, chronic obstructive pulmonary disease (COPD) and cough. The ATP levels have been discovered to be elevated in these conditions and it has been suggested that TRPV4 plays a role in the cascade which leads to ATP being released (Grace et al., 2017).

In addition, TRPV4 activation is involved with conditions such as pulmonary oedema, brain oedema, itch (caused by serotonin) and colitis. TRPV4 is expressed in cells at the epithelia and the endothelia and in the gastrointestinal tract tissues (Grace et al., 2017).

### 2.2.2 The structure of TRPV4

The human TRPV4 receptor is a homotetramer and every subunit of the human protein is built up from 871 amino acids (White et al., 2016). Both the N-terminal and the C-terminal of the protein are located at the intracellular region. The three-dimensional (3D) TRPV4 structure of the western clawed frog (*Xenopus tropicalis*) has been solved with cryo-electron microscopy (cryo-EM) to a resolution of 3.8 Å (Protein Data Bank (PDB) ID 6BBJ) (Fig. 1) and since the sequence identity between the human and the frog TRPV4 is 78% (Deng et al., 2018), one would also expect the structure to be similar between these two

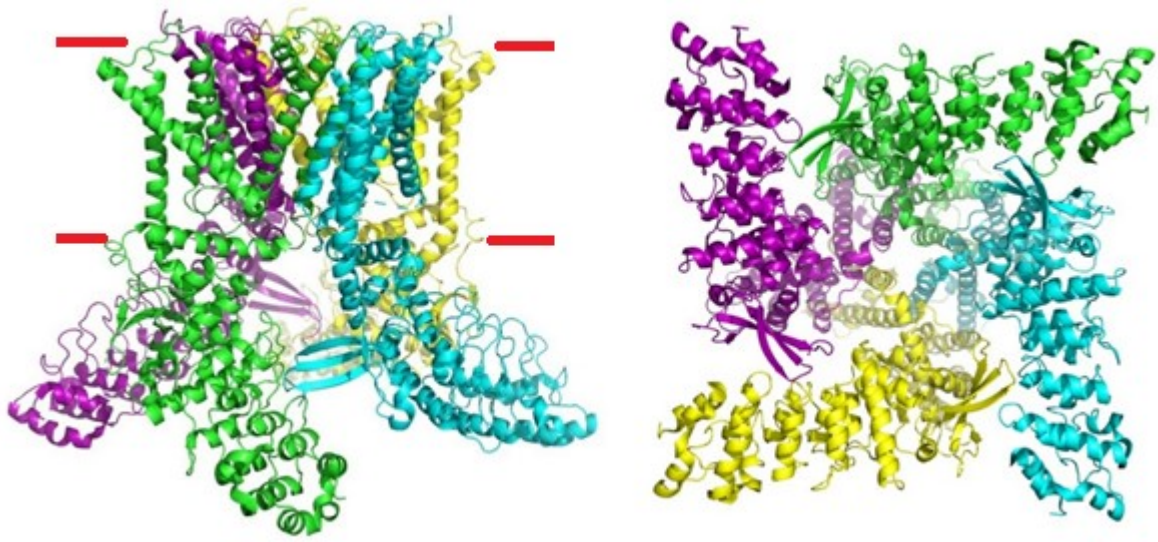

**Figure 1.** The structure of the frog TRPV4 (PDB ID 6BBJ). The four different subunits are colored with different colors. The two pictures show the structure from two different angles, left: the side view; right: the bottom view. The red lines mark the approximate location of the cell membrane.

proteins. The domains the frog model shows are, starting at the N-terminal end, first comes ankyrin repeat domain (ARD), which includes six ankyrin repeats (Deng et al., 2018). Ankyrin repeats are important for the protein function, and they are often engaged in protein-protein interactions. One ankyrin repeat is composed of 33 residues and the residues form two alpha helices which are connected by one loop (White et al., 2016). Next comes the linker domain which is a three-stranded beta sheet that includes two beta strands from the N-terminal and one beta strand from the C-terminal region. A helix-turn-helix motif and an alpha helix called the pre-S1 helix are situated in the intracellular region. These are also parts of the linker domain. The S1-S4 domain and the pore domain are transmembrane domains. The S1-S4 domain is built out of four alpha helices and the loops that connect them. The pore domain is built out of two bigger alpha helices, S5 and S6, and one smaller alpha helix called the pore helix. The composition of the S1-S4 domain and the pore domain are domain-swapped in the structure (Fig. 2), which signifies that the S1-S4 domain from one subunit interacts with a pore domain from another subunit. A TRP domain is found between the intracellular and the transmembrane area (Deng et al., 2018). The TRP domain amino acid sequence is conserved amidst the TRPV proteins. The TRP domain in TRPV1 has been proposed to be important for channel activity and channel stabilization (García-Sanz et al., 2007). Figure 3 shows the structure of one TRPV4 subunit.

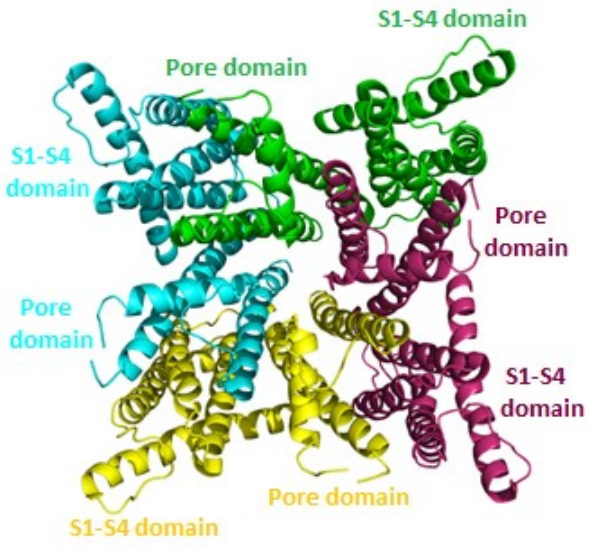


**Figure 2.** The domain-swapped structure. The different subunits are seen in different colors. Only the S1-S4 domains and the pore domains are shown, the domains are labeled separately for each subunit.

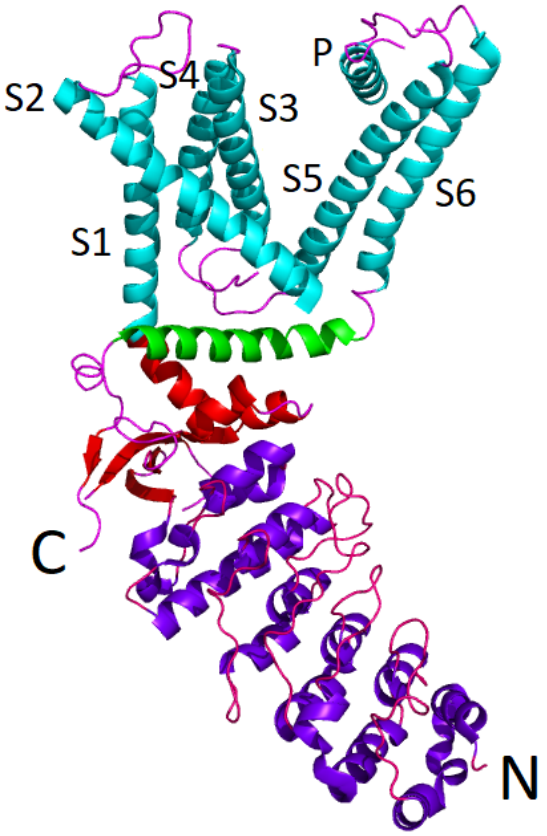


**Figure 3.** The structure of one frog TRPV4 subunit. The ankyrin repeat domain in purple-blue and hot pink, the linker domain in red, S1-S4 domain and pore domain in cyan and TRP domain in green. The S1-S6 helices and pore (P) helix are labeled. Other than ankyrin repeat domain loops are seen in magenta. The N- and C-terminals are labelled.

### 2.2.3 TRPV4 ion pore

TRPV4 receptors are more permeable for ions that are bivalent, for example Ca2+ and Mg2+ ions, but in general TRPV4 is a nonselective channel and other non-bivalent ions may also permeate the ion pore. The western clawed frog TRPV4 does not have an upper gate at the selectivity filter, which all the other known TRPV receptors have (Deng et al., 2018). The narrowest point at the selectivity filter in the frog TRPV4 is 10.6 Å and defined by the residue Gly 675 (corresponds to Gly 679 in human TRPV4). This diameter is too wide to be working as a gate, if compared to the corresponding structure in TRPV1, which has a diameter of 7.6 Å in the open channel conformation. The closed conformation of the TRPV1 upper gate is 4.8 Å in diameter. The frog TRPV4 has only one lower gate in the intracellular S6 pore helix. The important residue for this gate was determined to be Met 714 (corresponds to Met 718 in human TRPV4). The side chain of Met 714 made the narrowest point of the channel only 5.3 Å wide. Thus, the resolved structure is not an open channel structure, as a diameter of 5.3 Å is too small for ions to pass. The amino acids constructing the lower gates in various TRPV receptors

are diverse, despite the fact that the amino acid sequences of S6 helices are conserved among the TRPV receptors (Deng et al., 2018).

Deng et al. (2018) discovered that all the various cations bind nearly to the same place in the ion pore. In the study, Cs+, Ba2+ and Gd3+ were all discovered to bind near the oxygen of the carbonyl group in Gly 675. The distance between the two backbone oxygen atoms of these glycine residues in opposite domains was found to be around 10.6 Å. This is large enough to fit a hydrated ion. Because TRPV4 has only one ion-binding site, the ions can freely move to both sides of the ion-binding site. The only exception was found with Gd3+ that is bound with very high affinity to TRPV4 and, therefore, blocks the ion pore.

Densities that were believed to belong to bound lipids were also observed in the selectivity filter area in TRPV4. Lipids are believed to be important for the stability of the protein and in TRPV4 they seem to be an essential part of the selectivity filter, filling up the unoccupied space. If the lipid findings are true, it is a unique feature, since similar bindings in the pore area have not been reported in other channels in similar tetrameric proteins (Deng et al., 2018).

The corresponding residues of the human TRPV4 mentioned in this section were obtained by multiple sequence alignment of various TRPV4 proteins (see section 4.1).

### 2.2.4 Mutations in the TRPV4 protein

Mutations in the TRPV4 amino acid sequence cause both disability and fatality (Fig. 4). Mutations in TRPV4 are found, for example, in the ankyrin repeat domain at the finger 3 loop. Another location is the S4-S5 linker and at the residues forming contacts with the TRP helix. Some mutations are found in the S5-S6 pore domain or in other areas that could affect the gating properties. Many mutations are found in the outer edges of the intracellular N-terminal domains of the channel. Mutations that cause a certain disease are usually found in one or a couple of the above-mentioned locations. This makes TRPV4 a possible target for illness-specific medicine development (Deng et al., 2018).

### 2.3 3D structures and drug design

Widely known problems with the drug development industry are not only the time consuming and extremely costly development steps, but also the fact that many promising drug candidates fail in the clinical trials after many years of development. By solving 3D structures of important target proteins and further by locating possible binding sites in the proteins, the scientists are making drug design faster, cheaper, and more accurate in finding specific drug candidates (Batool et al., 2019).

When there is no experimentally solved 3D structure of the target protein available, comparative modeling is one of the commonly used methods for building a theoretical model structure. The method can be used when the sequence identity between the protein used as the template structure and the protein to be modeled is generally greater than 40% (Batool et al., 2019).

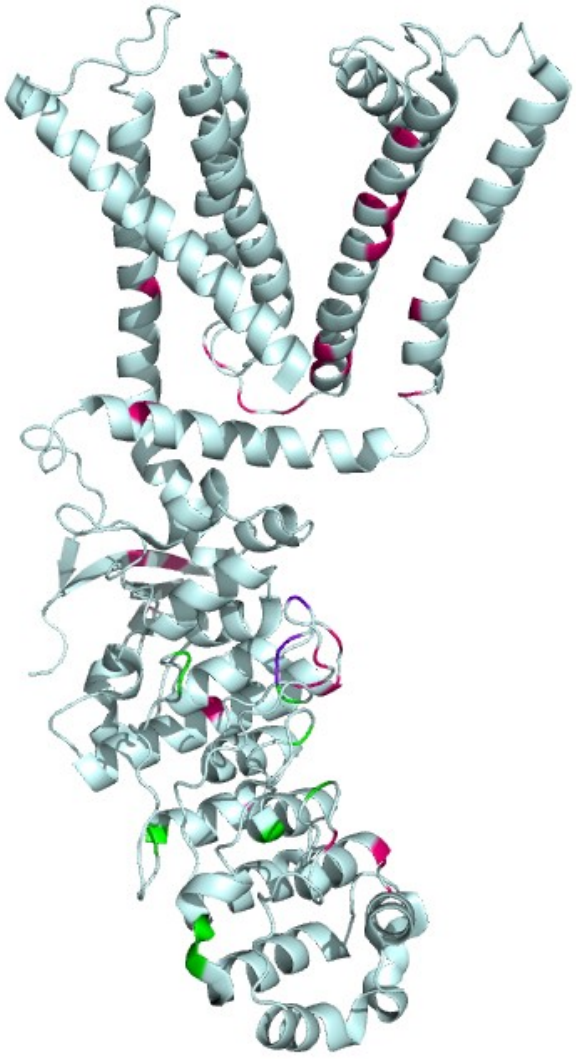

**Figure 4**. The known TRPV4 mutations. The figure is visualizing one chain in pale cyan. The areas colored with hot pink cause skeletal dysplasia, green areas cause peripheral neuropathy, and the purple-blue areas cause osteoarthropathy. The figure was modified from the figure in the study from Deng et al. (2018).

Three previous studies that describe the putative structure of the human TRPV4 have used the cryo-EM rat TRPV1 model as a template structure. The sequence identity between these two proteins is 41% (Berna-Erro et al., 2017, Garcia-Elias et al., 2015; Teng, Loukin, Anishkin, & Kung, 2015).

A human TRPV4 homology model by Cao et al. (2018) used rat TRPV2 as a template structure for almost the whole human TRPV4 sequence. An exception was a short C-

terminal part, which is less conserved in evolution. This part was modeled with several servers used for structure prediction. The overall sequence identity between rat TRPV2 and human TRPV4 is 37.8% (own alignment with Clustal Omega; Sievers et al., 2011).

### 2.4 Locating TRPV4 antagonist binding site(s)

To be able to find molecules that would inhibit human TRPV4 it was necessary to determine at least one possible antagonist binding site. During the experimental phases of this thesis, there were no antagonist binding sites of TRPV4 described in the literature.

One study describes the binding site for 5',6'-epoxyeicosatrienoic acid (5',6'-EET), which is a natural TRPV4 agonist (Berna-Erro et al., 2017). The study applied comparative homology modeling (with human TRPV1 as the template structure), molecular docking and MD simulations, microscale thermophoresis (MST), intracellular calcium imagining and electrophysiological techniques as research methods. The binding site was predicted to be in a pocket formed by the S1-S4 helices and their linkers. Important residues for binding were determined to be Lys 535, Phe 549, Gln 550, Tyr 591 and Arg 594. The study found these to be conserved residues. The residue Lys 535 was discovered to be critical for 5',6'-epoxyeicosatrienoic acid binding; when the lysine residue was mutated to alanine, the binding interaction was lost according to the MST analysis and the channel activation by the natural agonist (but not by heat or GSK1016790A agonist).

A binding site for intracellular ATP has been located at the ARD, residues from both finger 1 and finger 2 are involved in the binding (Li, Kao, & Chang, 2020). ATP sensitizes the TRPV4 receptor (Phelps, Wang, Choo, & Gaudet, 2010).

The Cryo-EM and nanodisc structure of the rat TRPV1 with bound antagonist capsazepine (Gao, Cao, Julius, & Cheng, 2016) and cryo-EM structure of rabbit TRPV5 with bound antagonist econazole (Hughes et al., 2018) (Fig. 5) are currently the only available experimentally solved TRPV structures with bound antagonists. These structures were used to visually compare the binding sites with the human TRPV4 model. The ligands are bound in a cavity that is located at a corresponding site in these two proteins. The site is called the vanilloid binding site (VBS). In TRPV1, the VBS is located between the S3 and S4 helices and the S4-S5 linker from one subunit and the S6 helix from an adjacent subunit.

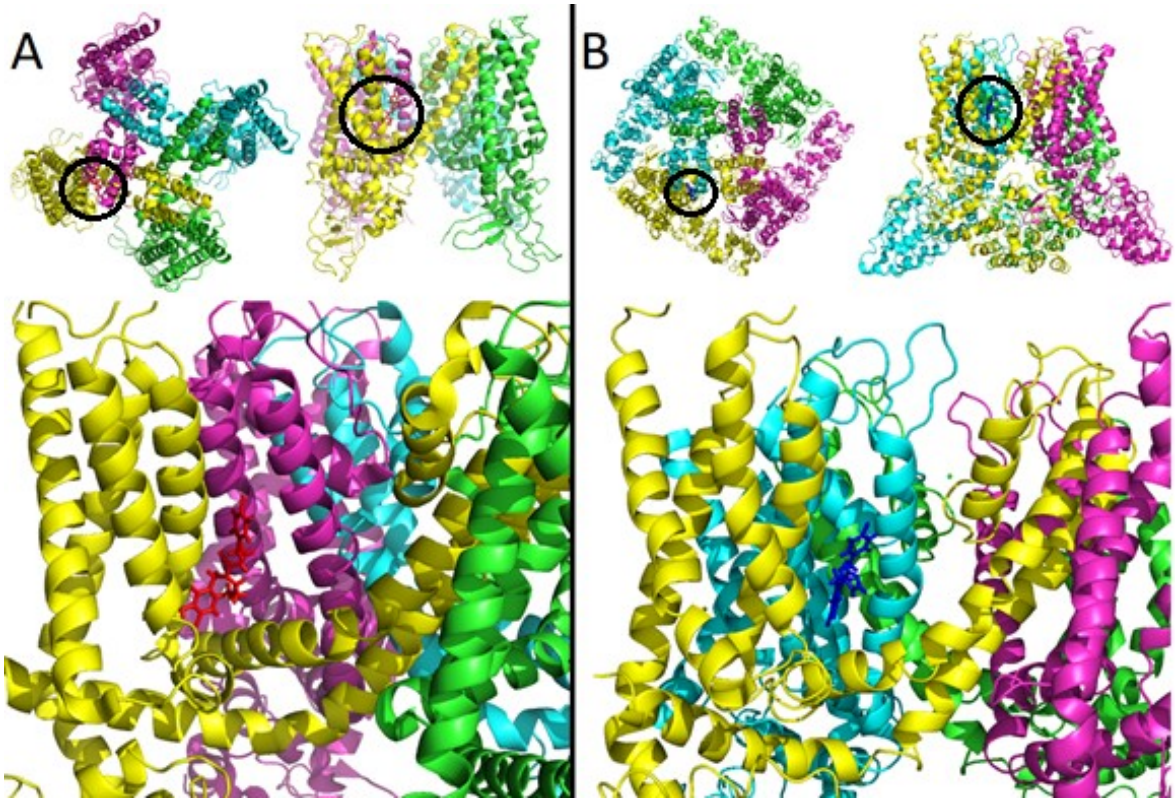


**Figure 5.** A) Structure of rat TRPV1 with bound capsazepine (red) (PDB ID 5IS0). B) Structure of rabbit TRPV5 with bound econazole (blue) (PDB ID 6B5V). The smaller figures visualize the protein shapes from different angles and the larger figures visualize the vanilloid binding site. The different subunits are colored in different colors. The black circles indicate the locations of the ligands.

Vanilloids are a group of molecules that bind into TRPV1 and act as agonists of the protein. The activation of TRPV1 by a vanilloid compound leads to pain sensation in the body. An example of a vanilloid compound is capsaicin. The antagonists of TRPV1 also bind into the VBS (Gao et al., 2016).

Gao et al. (2016) discovered that crucial residues for agonist (resiniferatoxin) binding in rat TRPV1 were Tyr 511, Ser 512, Leu 515, Val 518, Met 547, Thr 550, Arg 557, Ile 573 and Leu 669 from an adjacent subunit. These residues could form hydrophobic interactions or hydrogen bonds with the agonist. Based on a multiple sequence alignment (see section 5.2), the corresponding residues in human TRPV4 are Ser 548, Phe 549, Leu 552, Ile 555, Leu 584, Met 587, Arg 594, Leu 610 and Leu 705. Phe 549 and Arg 594 were also found to be important residues at the TRPV4 agonist binding site (Berna-Erro et al., 2017). These residues could give some direction when distinguishing if the ligand is an agonist or antagonist. Gao et al. (2016) showed that the agonist resiniferatoxin and antagonist capsazepine bind into the same location, but capsazepine did not interact with the residues

Arg 557 and Glu 570. Crucial residues for antagonist binding were not presented in the study.

Hughes et al. (2018) discovered that the crucial residue for econazole binding in rabbit TRPV5 was Phe 425. The corresponding residue in human TRPV4 (obtained with multiple sequence alignment, see section 5.2) was Phe 549 (it was also found to be an important residue at the TRPV4 agonist binding site; Berna-Erro et al., 2017).

Computational tools for detecting binding pockets can also be used. Finding the most probable binding pocket in such a big protein can be challenging, since the number of possible cavities can be large. Computational tools calculate, for example, interaction energies between the target protein and such functional groups that are known to be important for drug binding (Batool et al., 2019), thus recognizing druggable binding sites. A druggable binding site is also hydrophobic, adequate in size (not too small) and not too shallow (Halgren, 2008).

Some clues for a binding site can also be found by examining the amino acid sequence. It has been discovered that tryptophan is located more frequently in ligand binding sites than in other parts of the protein. Also, charged residues act more often as catalytic residues compared to hydrophobic or polar residues (Henrich et al., 2010).

### 2.5 Examples of TRPV4 antagonists

Some known TRPV4 antagonists (Fig. 6) are for example ruthenium red, RN-1734 (Vincent et al., 2009) and HC-067047 (Zhang et al., 2015). GlaxoSmithKline has developed many TRPV4 antagonists, and they have started clinical trials with at least one of their molecules, GSK2798745. There are at least three clinical studies involving GSK2798745. A phase II clinical study evaluating the effect and safety of GSK2798745 when treating pulmonary edema has been successfully finished (Goyal et al., 2019; ClinicalTrials.gov, 2021a). A phase I clinical study assessing the safety, pharmacodynamics, tolerability and pharmacodynamics of GSK2798745 when treating diabetic macular edema is ongoing (ClinicalTrials.gov, 2021b). A phase II clinical study, evaluating the effect and side effects when treating chronic cough with GSK2798745 was aborted due to the lack of efficacy (Ludbrook et al., 2019; ClinicalTrials.gov, 2021c). Kanju et al. (2016) reported two

derivatives of GSK205, another TRPV4 antagonist, to act as TRPV4/TRPA1 dual inhibitors. The molecules reduced trigeminal pain provoked by formalin.

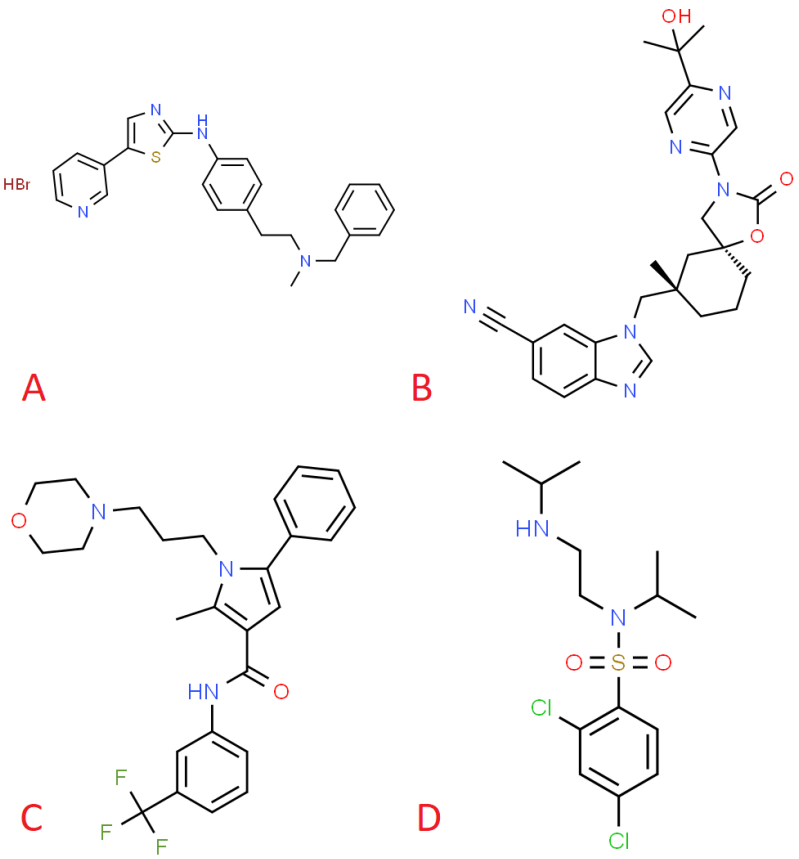


**Figure 6.** The structures of some known TRPV4 antagonists. A) GSK205, B) GSK2798745, C) HC-067047 and D) RN-1754. The structures were obtained from ChemSpider (http://www.chemspider.com/).

### 2.6 Virtual screening by molecular docking

There are many powerful computational tools available for virtual screening of molecular libraries and for MD simulations of proteins and protein-ligand complexes.

A common approach for structure-based virtual screening is molecular docking. By using energy scoring functions and search algorithms, the aim is that the molecular docking method predicts accurately the native binding mode of the ligand in the binding cavity. There are still challenges to overcome with the docking methods, for example the flexibility of the target protein and the ligands can contribute to bias in the prediction results (Guedes, de Magalhães, & Dardenne, 2014).

Glide (Friesner et al., 2006) is a commonly utilized docking software used for virtual screening. Three modes of docking and scoring functions can be used with Glide: virtual high-throughput screening (vHTS), standard precision (SP) and currently the most optimized function, the extra precision (XP) docking mode. The aim with these scoring functions is to calculate the estimated binding affinity for a ligand and a protein. vHTS

generates the roughest estimate and is thus the fastest, SP mode being the second fastest and XP the most accurate and slowest.

The scoring function for XP Glide is the sum of van der Waals interaction energies, Coulomb interaction energies and factors that both support and prevent ligand binding. XP Glide finds the best pose by dividing the ligand in portions and then, one portion at the time, finding the best pose and avoiding penalties by repositioning the different structural moieties of the ligand. Although the XP Glide is superior compared to the other Glide scoring functions, it still has its limitations, especially in distinguishing the compounds that are active from the compounds that are inactive (Friesner et al., 2006).

GOLD (Verdonk et al., 2003) is another common commercial docking software used for virtual screening. GOLD applies fitting points when docking the ligand into the binding cavity. A genetic algorithm (GA) is employed when GOLD examines potential binding modes. The GA performs modification and optimization of various ligand parameters such as the geometries of the ring structures.

### 2.6.1 Binding free energy estimation and MD simulations

Since the scoring functions of docking programs are fast and not so accurate, more accurate methods are used for re-ranking the top candidates and for studying the stability of the docked complexes.

MD-based molecular mechanics-generalized Born surface area (MM-GBSA) method is used to calculate the binding free energy of ligand-protein complexes. MM-GBSA is useful when the aim is to rank ligands according to the relative binding affinity instead of trying to calculate the absolute binding affinity between a protein and a ligand. The binding free energy is calculated from the changes in molecular mechanics energy in the gas phase, the free energy of the solvation and the conformational entropy at the time of binding. Tingjun, Junmei, Youyong and Wei (2011) investigated with a set of ligands and proteins if the length of the MD simulation was critical for the MM-GBSA calculations. The investigated simulation lengths varied from 400 – 4800 ps. They discovered that longer simulation times do not necessary mean more accurate results. Instead, the study emphasized the importance of using a good force field. A force field is used for calculating the predicted

force for every atom in the system, in other words, it predicts the potential energy of the simulation system (Bowers et al., 2006).

MM-GBSA calculations can be performed with for example Prime (Jacobson et al., 2004). The default solvation model in Prime is the VSGB 2.0 and according to Li et al. (2011) the solvation model is valid for SBDD. The force fields available are the OPLS_2005 (Kaminski, Friesner, & Jorgensen, 2001) or the OPLS3e (Roos et al., 2019). The output from the Prime MM-GBSA calculations gives properties such as the Coulomb energy, Van der Waals energy and Generalized Born electrostatic solvation energy. Prime binding energy (DG bind) is calculated according to the equation below (Prime 4.0 User Manual, 2015):

$$\text{DG bind} = \text{E\_complex(minimized)} - \text{E\_ligand(minimized)} - \text{E\_receptor(minimized)}$$

MD simulations are also inherently suitable for studying the time-dependent movements (dynamics) of a protein, or the movements and interactions of a protein with one or multiple bound ligands. When performing a MD simulation, a solvation box surrounding the protein or protein-ligand complex is first added to the simulation system. Water is commonly used as the solvent. A suitable algorithm and a force field is required for the simulation. Newton's second law is applied for portraying the particle movements. The analysis of the simulation is performed using the trajectories that include the simulation information (for example atomic coordinates and energies) saved at set time steps during the simulation. For now, simulation times longer than milliseconds are still challenging to execute because of the requirement of extremely powerful computers. However, the computational advancements are increasing the capacity continuously (Salo-Ahen et al., 2021).

In drug design, the stability of ligand binding is frequently studied by MD simulations. When performing MD simulations with membrane proteins, it is crucial to add the membrane with the lipids to the system. If the membrane is not added, there can be important interactions between the protein and the lipids that are not identified (Arinaminpathy et al., 2009).

The MD simulations in this thesis were performed with Desmond (Bowers et al., 2006). Desmond is developed by D. E. Shaw Research and the purpose of the software is to

execute fast MD simulations of biomolecular systems (D E Shaw Research, 2021). Other examples of commonly used MD simulation software are AMBER (Case et al., 2018) and GROMACS (Abraham et al., 2015).

### 2.7 Pharmacophore-based virtual screening

A pharmacophore includes the key structural properties that a ligand must possess to be able to bind into a specific binding site (Blass, 2015). With ligand-based pharmacophore modeling the properties of known ligands that are important for protein-ligand binding can be specified. The pharmacophore model can be used for finding new ligands with similar activity but different scaffold, compared to known compounds. Pharmacophore modeling could also give indirect information on the properties a favorable binding site should have and could be used to facilitate the identification of putative binding sites. An example of a medicine discovered with pharmacophore modeling is the tuberculosis medicine isoniazid (Batool et al, 2019).

PHASE (Dixon et al., 2006) as implemented in Maestro was used in this thesis for pharmacophore modeling and ligand screening. The algorithm in PHASE determines first the ligand alignments and then identifies the pharmacophores. The created pharmacophores are ranked with the scoring function PhaseHypoScore (Phase, 2021). In PHASE, the pharmacophore model can be built by using one or multiple ligands. The ligands can be active or inactive. The pharmacophore model indicates the spatial locations of the included features. These are hydrogen bond acceptors and donors, aromatic rings, hydrophobes and the positive and negative ionizable features. The directionalities of the features are presented where relevant. In addition, it is possible to add other desired features to the created pharmacophore. This is performed by adding the SMARTS patterns of the desired features to the pharmacophore (Dixon et al., 2006).

### 2.8 Properties of druggable ligands and ADMET prediction

The “Lipinski’s rule of five” is a classic guideline when creating and discovering new ligands. According to the rule, an orally active drug candidate should have a molecular weight of maximum 500 Daltons, Log P should be equal or less than 5 and the molecule should have equal or less than 10 hydrogen bond acceptors and 5 hydrogen bond donors (Lipinski, Lombardo, Dominy, & Freeney, 1997).

Prediction of the pharmacokinetic and toxicologic properties of drug candidates gives information about how the molecule would behave in the body. It is important to determine the ADMET properties of the molecule as soon as possible in the drug development process.

Water solubility is important for parenteral and orally administered medicines. A molecule with good water solubility also requires less effort, when preparing the pharmaceutical formulation, compared to less soluble molecules (Daina et al., 2017). This correlates with the “Lipinski’s rule of five” which states that the ideal lipophilicity value (log P) for a good drug candidate should be low (Lipinski et al., 1997).

In drug discovery, it is important to know which cytochrome P450 (CYP) isoforms are inhibited or induced by a new drug candidate or if the molecule interacts with the P-glycoprotein (P-gp). This leads to a better understanding of the possible interactions and adverse effects that may occur because of the interaction (Daina et al., 2017).

Toxicity assessments for chemicals are critical in the pharmaceutical industry (Banerjee, Eckert, Schrey, & Preissner, 2018). Since the amount of chemicals exposed to humans, animals and the environment increases, it is important to evaluate the toxicity profile for both single chemicals and chemical mixtures. The *in silico* toxicity prediction methods complete the *in vivo* methods. With toxicity prediction, the cost, time and need of animal testing are decreased.

Important predictions for toxicity include the lethal dose 50 (LD50) value. It is also of interest to predict if the drug candidate is hepatotoxic, cardiotoxic or cytotoxic, acts as a carcinogen or induces mutagenicity (Banerjee et al., 2018).

## 3. Aims

The main aim for this thesis work was to identify novel putative TRPV4 antagonists for the treatment of pain and inflammation. The specific aims were as follows:

- To build a comparative model of human TRPV4 and test the stability of the model with molecular simulations.

- To locate antagonist binding site(s) in human TRPV4 by using structural data, information from literature and by computational methods.

- To identify putative TRPV4 antagonists by vHTS of a molecule library using molecular docking with the docking programs GOLD and Glide.

- To test the best ranked ligands for their binding stability at TRPV4 with MD simulations.

- To create a pharmacophore model of known TRPV4 inhibitors and carry out a pharmacophore-based virtual screening of a molecular library with the program PHASE.

- To perform *in silico* ADMET prediction with SwissADME and ProTox II for the best scoring ligands from both the docking and the pharmacophore-based virtual screening results.

- To choose the virtual hits with the best qualities to be sent to our collaborators for further activity testing.

## 4. Material and methods

### 4.1 Building a stable comparative model of human TRPV4

The 3D structure of the human TRPV4 receptor has not yet been solved experimentally. Therefore, to be able to build a 3D model of the protein using comparative modeling, a homologous template structure was needed. The TRPV4 structure of the western clawed frog (*Xenopus tropicalis*) was used as a template structure, since it had been solved with cryo-EM to a resolution of 3.8 Å. As mentioned before, the reported amino acid sequence identity was 78% between the human and frog TRPV4 receptor (Deng et al., 2018).

The amino acid sequences of both human and the western clawed frog TRPV4 were downloaded from the UniProt Knowledgebase (https://www.uniprot.org/) and the western clawed frog TRPV4 3D structure was downloaded from the PDB database (Berman et al., 2000, rcsb.org) (PDB ID 6BBJ). The frog cryo-EM structure contains the residues 144 to 784 (Deng et al., 2018) from each subunit. A few loop structures were not experimentally modeled. These included the large extracellular loop, and a couple of small intracellular loops. The experimentally solved structure was used as a template to model the corresponding residues of the human TRPV4 sequence. A multiple sequence alignment of TRPV4 proteins from different organisms was performed to learn which residues correspond in the two structures. The sequences used were the human TRPV4 (Uniprot ID Q9HBA0), mouse TRPV4 (Uniprot ID Q9EPK8), rat TRPV4 (Uniprot ID Q9ERZ8), chicken TRPV4 (Uniprot ID A0A1D5PXA5), two western clawed frog TRPV4 sequences (Uniprot ID F7BWY7 and F6YG73) and the rhesus macaque TRPV4 (Uniprot ID A0A1D5R8Y1). The sequences were downloaded on April 12, 2018.

The software MODELLER (Šali & Blundell, 1993) was used for homology modeling. The residues 148 – 788 from each chain of the human TRPV4 sequence were modeled. Ten alternative 3D models were created and the best model with the most negative DOPE score (Shen & Sali, 2006) was chosen for further studies. In addition, Ramachandran Plot of the model, which shows the stereochemical properties of the structure (Batool et al., 2019), was examined and the model was visually inspected with PyMOL (DeLano, 2002).

To investigate the stability of the model, MD simulations were performed with Desmond (Bowers et al., 2006). Since the TRPV4 is a membrane protein, the initial step was to set up

the membrane with Desmond System Builder. The residues used to set up the membrane were the residues 470 – 492, 506 – 636 and 665 – 719 in all four subunits, these residues correspond to the S1 – S6 helices without the extracellular loops (Fig. 7A) and are the transmembrane parts of the protein (Deng et al., 2018). The solvent model was set to TIP3P (Jorgensen, Chandrasekhar & Madura, 1983). The periodic boundary conditions were used; the simulation box shape was orthorhombic; the box size calculation method was buffer and the distances around the solute were 10 Å x 10 Å x 10 Å. The box volume was 3613897 $Å^3$ and the force field used was OPLS3e (Roos et al., 2019). The system was neutralized with 20 Cl- ions since the charge of the protein was +20. NaCl was added to the system with a physiological concentration of 0.15 M.

The model system (Fig. 7B) was then uploaded to the Desmond Molecular Dynamics tool. The total simulation time was 300 ns, but due to technical limitations, the simulations had to be run in 100-ns parts. During the simulation from 0 to 100 ns the trajectory was recorded every 100 ps, from 100 to 200 ns every 200 ps and from 200 ns to 300 ns every 300 ps. The approximate number of frames was 1000. The ensemble class was NPT, temperature 300 K and pressure 1.01325 bar. The model system was relaxed before the simulation.

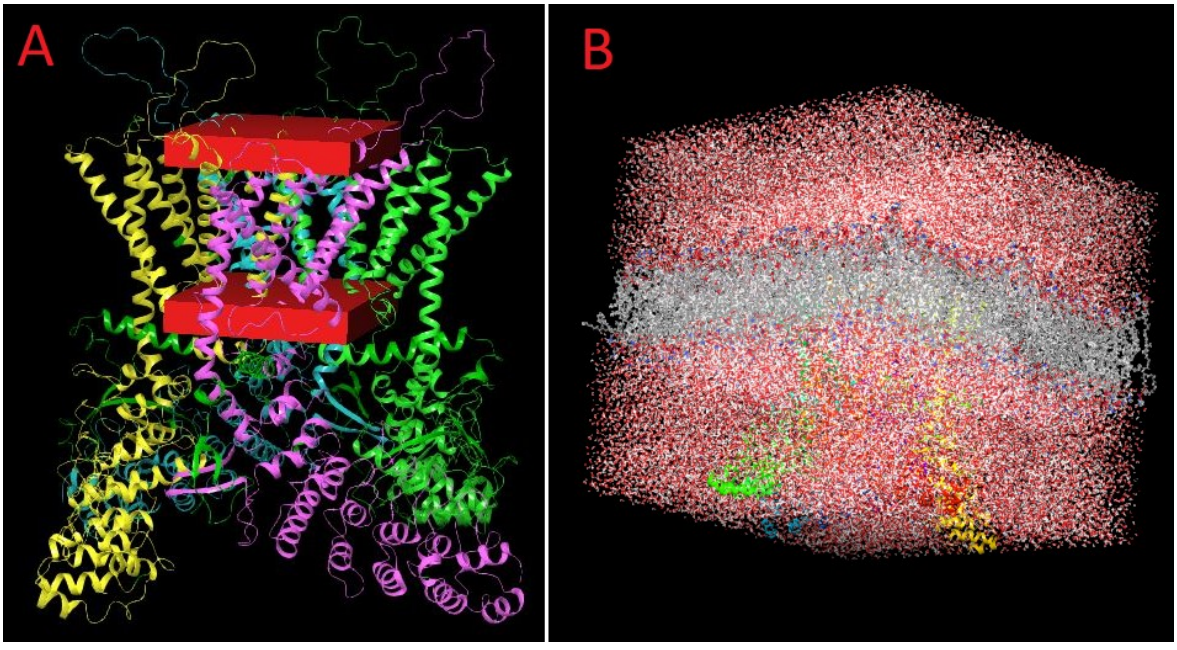


**Figure 7.** A) The cell membrane is visualized as the area between the red rectangles. B) The simulation system with the solvent box, the membrane is seen in grey. The figures were prepared with Maestro.

Before merging the three 100 ns trajectories, the files had to be reduced in size. This was done by decreasing the number of frames. Each trajectory initially had 1001 frames. Only every tenth frame was saved and so the remaining trajectory file had 101 frames. The three trajectory files were merged, and the final trajectory had 303 frames.

The trajectories of the simulation were visualized with the Desmond trajectory panel as implemented in Maestro. A video of the simulation was also prepared with the same tool.

The protein root mean square deviation (RMSD) and the root mean square fluctuation (RMSF) plots of the MD simulation were analyzed to see if the protein remained stable during the simulation.

Pictures of the protein were prepared with PyMOL (DeLano, 2002). Also, throughout the thesis, pictures were prepared with PyMOL, if not stated otherwise.

### 4.2 Locating the putative antagonist binding site

Putative antagonist binding pockets were located by using the available literature and the TRPV structural data in the PDB, but also with computational tools such as ConSurf (Ashkenazy et al., 2016), SiteMap (Halgren, 2009) and Clustal Omega (Sievers et al., 2011).

Since the binding sites are often conserved in evolution (Henrich et al., 2010), ConSurf server (Ashkenazy et al., 2016) aims to find the parts of the (model) structure or sequence that are resembling known ligand binding sites in other proteins and the residues that are important for protein function. Chain A of the TRPV4 model was evaluated. All the 149 sequences approved for the multiple sequence alignment had the sequence identity percentage between 35 and 95% compared to the human TRPV4 sequence.

A search in the PDB database (Berman et al., 2000, rcsb.org) was performed to be able to investigate the experimentally solved 3D structures of other TRPV proteins with bound antagonists. These structures were used to visually compare the binding sites with the pockets found in the human TRPV4 model. The structures previously mentioned were used; the rat TRPV1 with the bound antagonist capsazepine, the resolution of the structure is 3.43 Å (PDB ID 5IS0) and the rabbit TRPV5 with the bound antagonist econazole, the resolution of the structure is 4.80 Å (PDB ID 6B5V). The structures were aligned with the

human TRPV4 model in PyMOL (DeLano, 2002). The purpose was to see if cavities could be seen in the corresponding sites in the protein model.

Multiple sequence alignment of all six subtypes of TRPV proteins was performed with Clustal Omega (Sievers et al., 2011) to be able to locate the conserved residues. All but human TRPV4 have experimentally solved 3D structures. The human TRPV4 (Uniprot Q9HBA0), rat TRPV1 (Uniprot O35433), rabbit TRPV2 (Uniprot G1SNM3), human TRPV3 (Uniprot Q8NET8), rabbit TRPV5 (Uniprot Q9XSM3) and human TRPV6 (Uniprot Q9H1D0) were used.

The locations of tryptophan and charged residues were identified in the human TRPV4 model to determine if the results would support the putative vanilloid binding site.

SiteMap (Halgren, 2009) as implemented in Maestro was used to locate putative binding cavities. It was particularly interesting to see if SiteMap would recognize the cavity that corresponds to the rat TRPV1 and rabbit TRPV5 antagonist binding cavity. The scoring functions in SiteMap yield important information regarding the cavity's theoretical suitability for drug binding. SiteMap uses so called site points for identifying potential binding cavities. An algorithm is used to recognize site points that are most likely involved in favorable binding interactions. SiteMap then connects multiple site points, and these points form a putative binding site (Halgren, 2007). The default settings in SiteMap were used; at least 15 site points were required per reported site.

### 4.3 Virtual screening by molecular docking

#### 4.3.1 With Glide

Once the putative binding pocket was selected, a virtual screening with the docking program Glide (Friesner et al., 2006), as implemented in Maestro, was performed by docking a compound library downloaded from Enamine (https://enamine.net/) to the selected binding site. The library was called the Ion Channel Library and it contained 36 800 molecules. The molecules were not prefiltered before docking since the compounds had been selected specifically for ion channels. The molecules were prepared for docking with the LigPrep tool of Maestro (LigPrep, Schrödinger, LLC, New York, NY, 2020). Force field OPLS3e was used and Epik (Epik, Schrödinger, LLC, New York, NY, 2021) was used for predicting the side chain protonation states. Other settings were left as default; at most 32

stereoisomers were created per ligand. Ligand preparation yielded 172 217 ligands in total, this includes the stereoisomers and tautomers generated for the library compounds. Seven structures were dropped in the preparation process.

The human TRPV4 model was prepared for docking with the Protein Preparation Wizard in Maestro (Sastry, Adzhigirey, Day, Annabhimoju, & Sherman, 2013). The protein was pre-processed with the default options; the hydrogens were added and possible disulfide bonds were created. The hydrogen-bond network was optimized at pH 7.0 and a restrained minimization in the OPLS3e force field was first run for only the hydrogens and then for the heavy atoms using the convergence criteria of RMSD 0.30 Å.

The docking site was defined with Glide's Receptor Grid Generation tool. The center of the grid was determined to be located at the coordinates X=154.79, Y=147.11 and Z=166.45. The length of the ligand was set to be less or equally 18 Å. The binding site residues with thiol or hydroxyl groups were allowed to be rotatable; for example, Ser 548, Tyr 553, Tyr 556, Ser 557 and Tyr 591 in A chain and Tyr 628, Thr 701 and Thr 706 in C chain.

The library of the prepared ligands was then screened using the Virtual Screening Workflow tool in Maestro. With this tool, it is convenient to setup the whole docking pipeline as one job. It is also possible to include ligand preparation and filtering in the pipeline when necessary. The first screening was executed with the robust vHTS mode. Roughly the best 10% of the compounds were selected for the next screening step. Then the docking was executed again with these selected compounds using the SP mode and again roughly the best 10% of the compounds were chosen for the next screening step. The last screening round that was executed with the XP docking mode yielded the best 10 % of the compounds (approximately 170 compounds).

Molecules that had the Glide XP score lower than -9 kcal/mol were chosen for the Prime/MM-GBSA binding free energy calculations. The MM-GBSA score suggests the affinity between the ligand and the protein; the lower the score, the stronger the affinity. The top compounds were ranked according to their predicted binding free energy.

Protein-ligand interaction analysis was visually performed with PyMOL (DeLano, 2002).

### 4.3.2 With GOLD

Two rounds of virtual screening with the Enamine Ion Channel library were performed with the docking program GOLD (Jones, Willett, Glen, Leach, & Taylor, 1997). The binding site center was defined with the same coordinates previously used with Glide docking. The docking site around the center was set to 18 Å x 18 Å x 18 Å. The scoring function CHEMPLP was used. The first screening had the genetic algorithm (GA) setting at 10%, which can be used for screening large libraries. This screening mode is the fastest, but also the least accurate. At most five poses were produced per ligand per docking run. The first screening was set to save the 500 best-ranking molecules. The binding mode of these 500 molecules was visually inspected and the molecules that were docked into the right cavity were selected. The selected 184 molecules were then redocked using the GA setting 100%, the "Default" mode. This mode performs a much more accurate screening; the software tries to find the most favorable pose for every ligand (Cambridge Crystallographic Data Centre, 2019). The screening was set to save the 50 best ranking molecules, according to the GOLD fitness score. The higher the GOLD fitness score, the higher the ranking.

MM-GBSA binding free energy calculations were then performed for the 50 best ranked molecules with Prime/MM-GBSA method of Maestro. The estimated free energies of binding from MM-GBSA helped re-score the poses and compare the results from GOLD and Glide.

## 4.4 Evaluation of the ligand-protein complex stability

Protein-ligand complexes with the best MM-GBSA scores from both GOLD and Glide screenings were visually analyzed and 10 complexes with structurally differing ligands were chosen as the most promising candidates for the MD analysis. The water box and membrane were first set up for the protein-ligand complexes with Desmond's System builder. These 10 protein-ligand complexes were then submitted for 300-ns MD simulation each with Desmond (Bowers et al., 2006) as implemented in Maestro. The goal was to find the protein-ligand complexes that would remain stable during the simulation, thus supporting the possible good binding affinity. The parameters used for both the System builder and MD in Desmond were the same as used in the simulation of the protein model

alone. The box volume altered slightly between the complexes as did the amount of Cl- atoms used to neutralize the charge of the simulation systems.

A reference simulation with complex 1 was performed without the lipid membrane. The purpose of the simulation was to compare the results with the complex 1 simulation that includes the membrane and to conclude whether the lipid membrane affects the results. The RMSD and protein-ligand interactions plots were inspected from the reference simulation. Simulation parameters were otherwise the same.

### 4.5 Analysis of the MD simulations

The trajectories of the simulations were visualized with the Desmond trajectory panel as implemented in Maestro. Videos of the simulations were also prepared with the same tool.

Based on the stability of the complexes, as evaluated with the energy, RMSD/RMSF plots and the protein-ligand interaction plots, ligands with the best scores and features were chosen for ADMET prediction.

### 4.6 ADMET prediction

ADMET properties were predicted with two online tools. ADME properties with SwissADME (Daina et al., 2017) and toxicity with ProTox II (Banerjee et al., 2018). The tools give a large amount of information, but for this thesis the most interesting predictions from ProTox II were the LD50 value, the toxicity class, and the predictions if the compound would act as a hepatotoxic, immunotoxic, mutagenic, carcinogenic or cytotoxic agent.

From SwissADME the most interesting predicted values were: if the molecule (i) is absorbed from the gastrointestinal (GI) tract, (ii) can permeate the blood-brain-barrier (BBB), (iii) acts as a P-glycoprotein (P-gp) substrate, (iv) acts as an inhibitor for various CYP enzymes and (v) how well the molecule can permeate the skin. SwissADME also predicts the synthetic accessibility, meaning how easy or difficult the molecule is to synthesize (Daina etl al., 2017). However, this is not so relevant for this project since the molecules can be purchased from the EnamineStore website.

Solubilities of the Enamine compounds were obtained from the EnamineStore website (https://www.enaminestore.com/). These predicted values were also compared with the SwissADME predictions. SwissADME predicts the solubility with three different

approaches. In this thesis the two topological methods were observed. The first method is the ESOL method by Delaney (2004). This method uses nine parameters related to the structure of the ligand to predict the solubility. These parameters include molecular weight, the number of hydrogen bond acceptors and donors, clogP, and the number of rotatable bonds. The second method by Ali et al. (2012) applies quantitative structure–property relationships (QSPR) calculations in the predictions.

A chart of the complete workflow for the docking project is visualized in Figure 8.

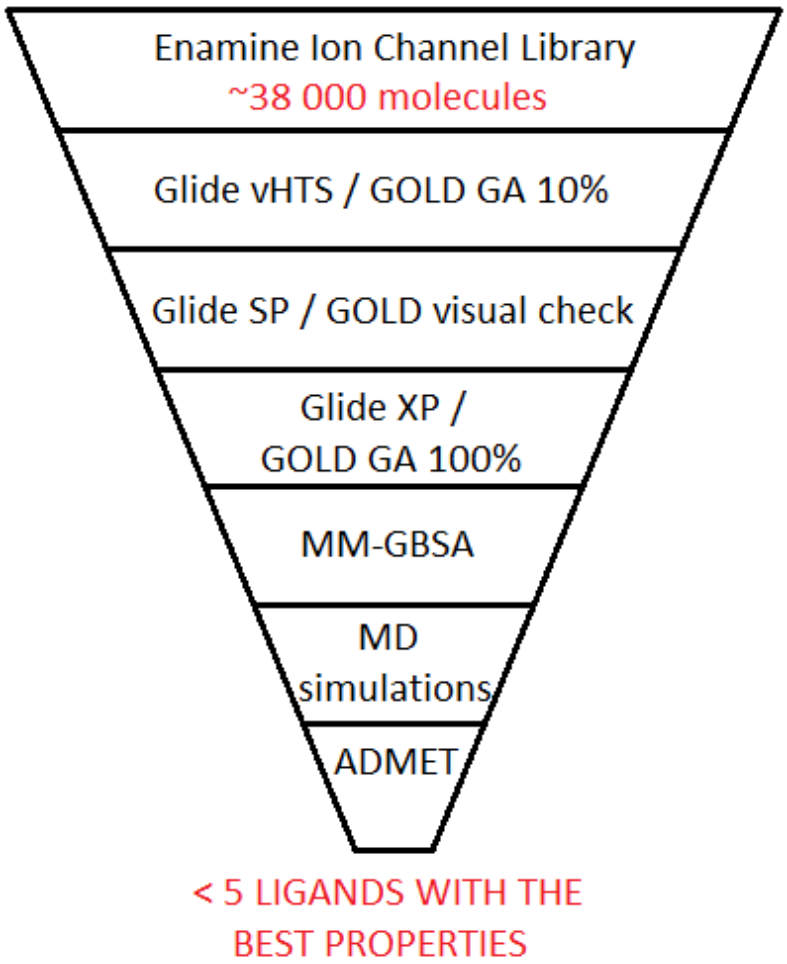


**Figure 8.** The workflow for the docking project. At the end of the virtual screening process, less than 5 of the best scoring ligands were chosen for biological testing.

### 4.7 Pharmacophore-based virtual screening

Pharmacophore modeling of two known human TRPV4 antagonists was used to identify new antagonist candidates.

PHASE (Dixon et al., 2006) as implemented in Maestro was used for pharmacophore modeling. The TRPV4 antagonists GSK205 and HC-067047 were used as the known active molecules for the model building. GSK205 is frequently used as a tool compound when designing new TRPV4 antagonists (Kanju et al., 2016). HC-067047 has been found to reduce pain caused by, for example, inflammation and neuropathy (Dias et al., 2019). The ligands were preprocessed with LigPrep. Since the two ligands were not aligned, the “Find best alignment and common features” method was

used. The hypothesis criterion was set as follows: the hypothesis shall match at least 50% of the active molecules, the number of features in the hypothesis shall range between 4 and 7 and the amount would preferably be at least 5. The criterion for the hypothesis difference was set to 0.50. The Phase Hypo Score was chosen as the scoring function. Additionally, the number of hydrogen acceptors was set to range from 0 to 9 and the number of hydrogen donors set to range from 0 to 4.

The resulting pharmacophore hypothesis was used for the PHASE ligand screening. The default settings in PHASE were used and the number of hits to report was set to 150. The output conformers were minimized.

As previously in the docking project, ADMET predictions were performed with the SwissADME and ProTox II tools for the best ranked hits.

To evaluate if the two known antagonists bind to the same vanilloid binding pocket, they were docked at that site in the TRPV4 model, and the binding free energies of the complexes were calculated.

## 5. Results

### 5.1 The human TRPV4 model

The DOPE score for the chosen human TRPV4 model was -319982.78125. The Ramachandran plot (Fig. 9) for the model showed 93.4% of the residues in most favored regions, 6.3% in additional allowed regions, 0.2% in generously allowed regions and 0.2% in disallowed regions. The residues in the disallowed regions were Ser 189 in A chain (Ser 42 in the plot), Tyr 657 in B chain (Tyr 1151 in the plot), Thr 656 in C chain (Thr 1791 in the plot) and Ser 758 in D chain (Ser 2534 in the plot). These residues were all located in flexible loop structures and were therefore not considered affecting the overall fold of the protein or the binding pocket studied in this work. The residues in the generously allowed regions were also located in flexible loop areas. These were Arg 775 in B chain (Arg 1269 in the plot), Asn 228 (Asn 2004 in the plot), Cys 652 (Cys 2428 in the plot) and Arg 774 (Arg 2550 in the plot) in D chain. Of these residues, Cys 652, Thr 656 and Tyr 657 were located at the extracellular loops that had not been solved in the frog TRPV4 cryo-EM structure. All other residues mentioned had the corresponding residues in the template structure. The Ramachandran Plot for the frog TRPV4 cryo-EM structure had no residues in the generously allowed and disallowed regions.

When visually analyzing the protein with PyMOL, the model appeared highly similar to the template (Fig. 10). Aligning the human TRPV4 model with the frog TRPV4 structure in PyMOL (Fig. 11) the root mean square deviation (RMSD) value was 0.649 Å. This is a good fit considering a perfect fit would have the value 0. 2136 C-alpha atoms from each protein were aligned and 296 C-alpha atoms were rejected from the alignment.

The results from the MD simulation revealed that the protein remained stable during the simulation and did not undergo any big conformational changes. The RMSD plot of the simulation is visualized in Figure 12A. The plot indicates that the protein reached the equilibrium state at around 160 ns and after that the fluctuation of the protein remained at around 4.3 Å. Fluctuations of this size indicate that the protein is moving in a “normal” manner. If the difference in the values of the fluctuation is high, it often implies that the protein is going through changes in its structure, for example unfolding. Also, if the RMSD value is not flattened at the end of the simulation, it signifies that the duration of the simulation is too short, and that the system has not stabilized.

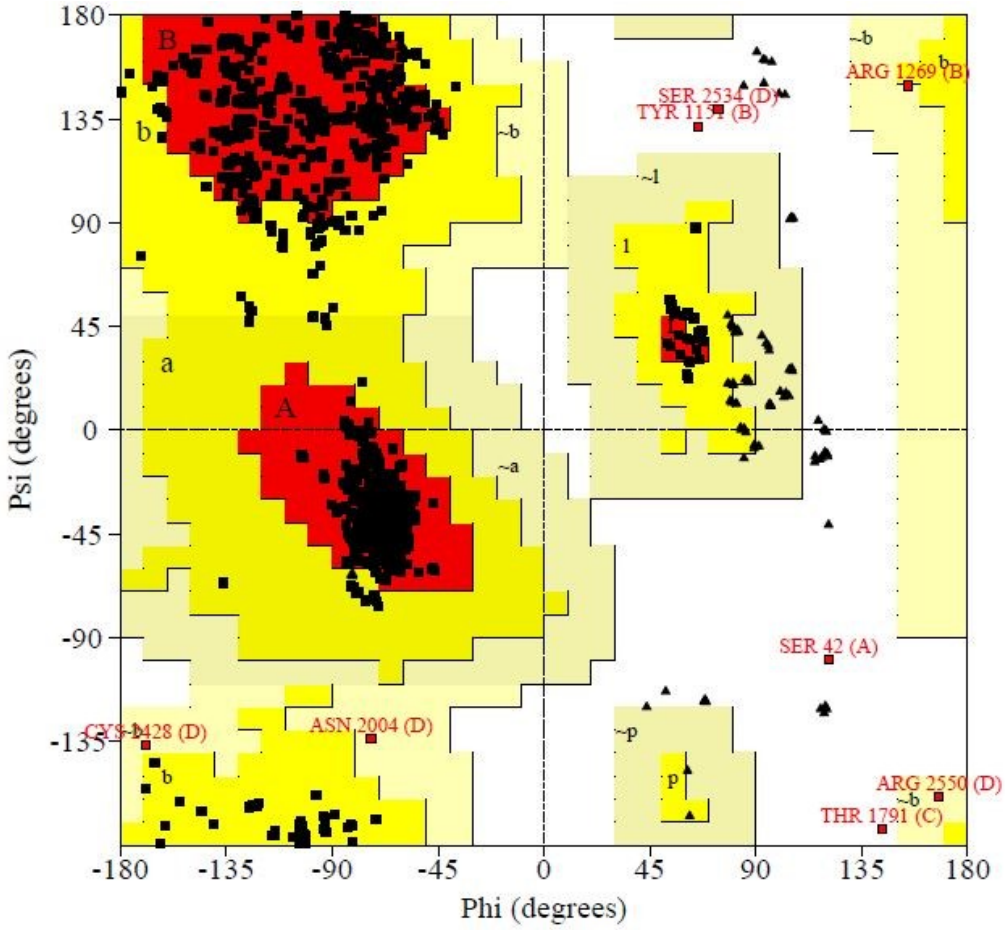


**Figure 9.** The Ramachandran plot for the chosen human TRPV4 model. The red areas are allowed regions and the bright yellow areas are additionally allowed regions. The pale-yellow areas are generously allowed regions and the white areas are disallowed regions. The triangles are glycine residues. As glycine does not have a side chain, it can be located at the disallowed region.

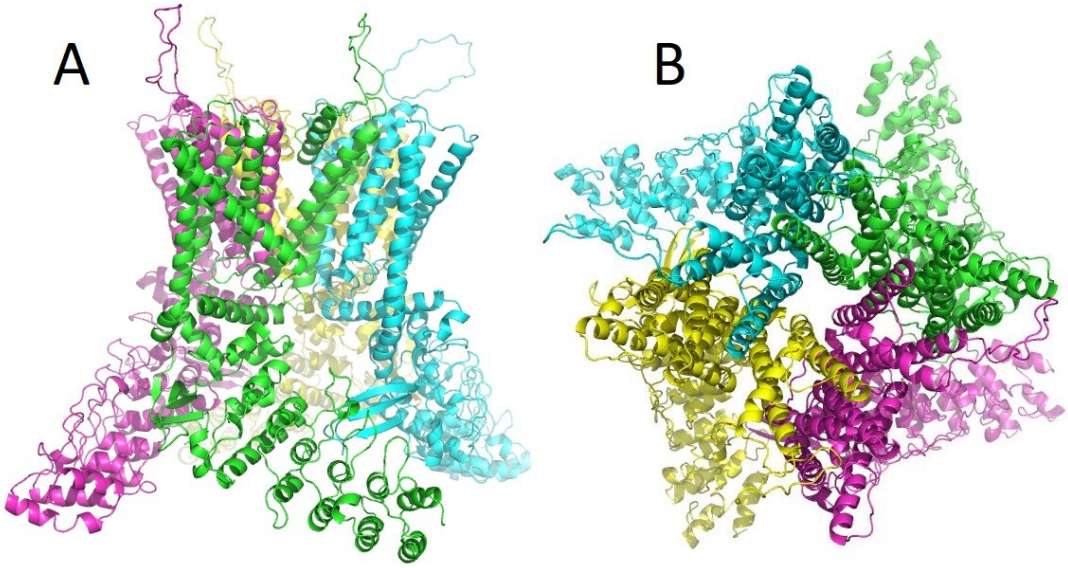


**Figure 10.** The human TRPV4 model obtained from homology modeling. Figures A and B show the protein from different angles. All four subunits are colored differently.

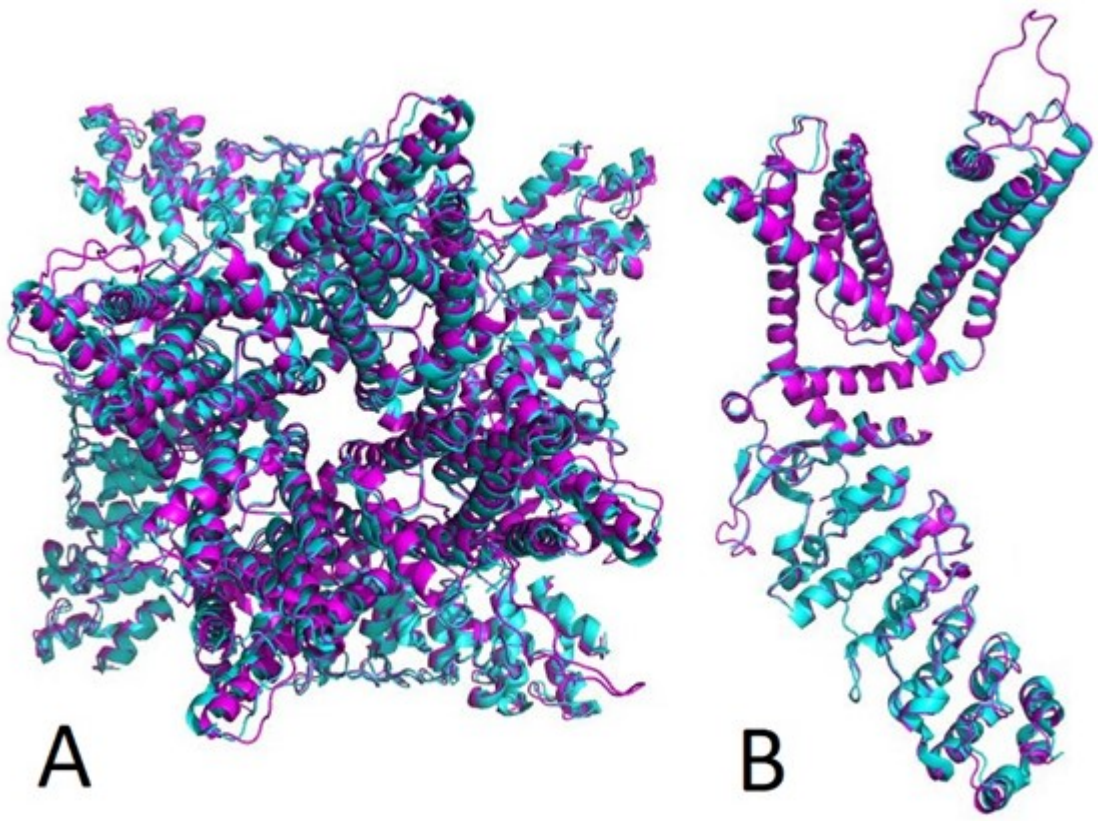


**Figure 11.** The aligned human TRPV4 model and the frog TRPV4 cryo-EM structure (PDB ID 6BBJ). The human protein can be seen in magenta and the frog protein in cyan. Figure A shows all four subunits and figure B shows one subunit from a different angle.

The RMSF plot visualizes how local parts of the protein alter during the simulation. Loop structures and N- and C-terminal ends are the structures which usually peak on RMSF plots. Alpha helices and beta strands should not peak since secondary structures are stiffer than unstructured sections of the protein.

The protein RMSF plot is presented in figure 12B. The four tallest peaks represent the large extracellular loop structures, and the other taller peaks represent the beginnings and ends of the four subunits. The A subunit starts at residue 1, subunit B begins at residue 642, subunit C begins at residue 1283 and subunit D begins at residue 1924. A peak can be seen in all subunit terminals.

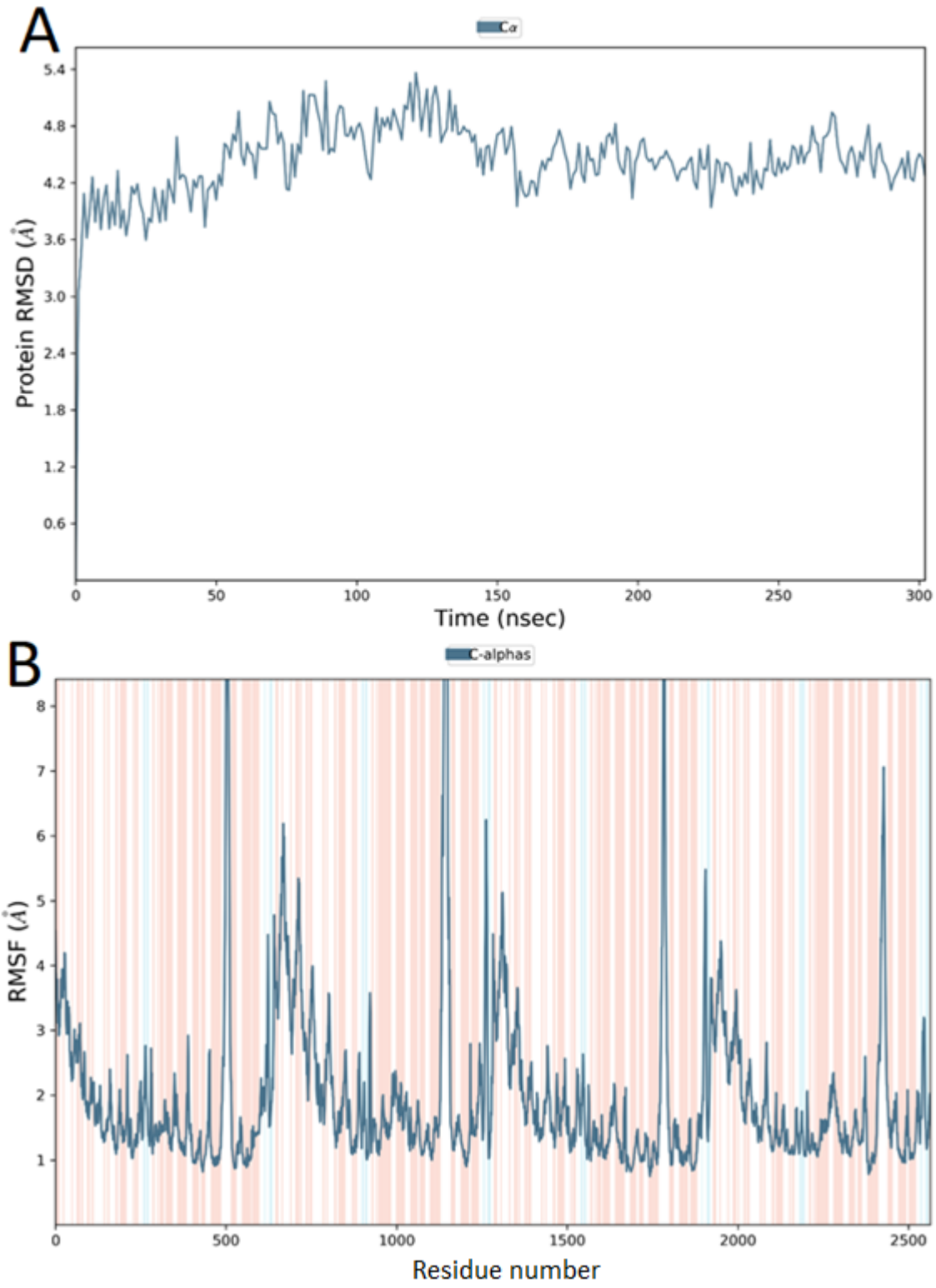


**Figure 12.** Stability of human TRPV4 model during a 300-ns MD simulation. A) RMSD during the simulation. The x-axis shows the simulation time and the y-axis shows the protein RMSD value; B) RMSF during the simulation. The x-axis shows the residue number and y-axis shows the RMSF value. Note, that the residue number 1 in the plot corresponds to the residue number 148 because the plot does not show the original N- or C-terminal of the human TRPV4 since these parts were not modeled. The white areas show the unstructured sections, the light red areas show the alpha helical sections, and the light blue areas show the beta strand sections of the protein.

### 5.2 The putative antagonist binding site

The results from ConSurf can be seen in Figure 13. The alpha helices S4, S5, S6 and the TRP domain are mostly very conserved, and the ARD is both highly conserved and non-conserved. Helices S2 and S3 seem to have most of non-conserved residues in TRPV4 when the loops are excluded. Especially helices S1 and S2 have non-conserved residues facing outwards from the protein center but the residues facing inwards are at least somewhat conserved. The putative vanilloid binding site in the TRPV4 protein is built up from both highly and less conserved parts.

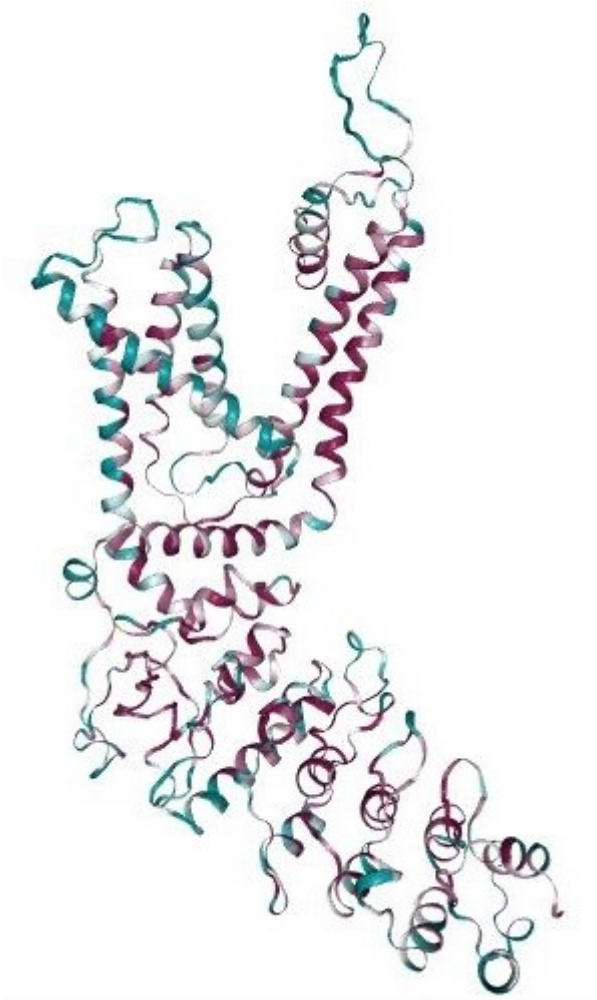

**Figure 13.** Results from ConSurf, chain A from the human TRPV4 model was evaluated. The dark purple indicates very conserved parts, and the dark turquoise shows the non-conserved areas.

There are eleven tryptophan residues in one human TRPV4 chain/subunit. They occupy the positions 122, 409, 463, 586, 733, 737, 776, 785, 788, 822 and 864. Residues 122, 822 and 864 were not included in the human TRPV4 model. According to the model, residues 409 and 776 are located at the three-stranded beta sheet structure and residues 785 and 788 at the loop that follows the C-terminal end beta strand. Residue 463 is located at the Pre-S1 helix. Residue 586 is located at the S2 helix and this residue is found at the vanilloid binding site. Residues 733 and 736 are located at the TRP domain. Figure 14 visualizes the locations of the tryptophan residues and the charged residues (histidine, aspartate, glutamate, arginine, and lysine) on one subunit of the human TRPV4 model.

The results from the multiple sequence alignment of the various TRPV proteins yielded 96 identical residues and 179 similar residues. The least conserved areas were located at the terminal ends of the chains. The vanilloid binding pocket area was discovered to contain many conserved residues. Figure 15 presents the multiple sequence alignment for the vanilloid binding pocket in TRPV proteins and Figure 16 visualizes the location of the conserved residues in the human TRPV4 model.

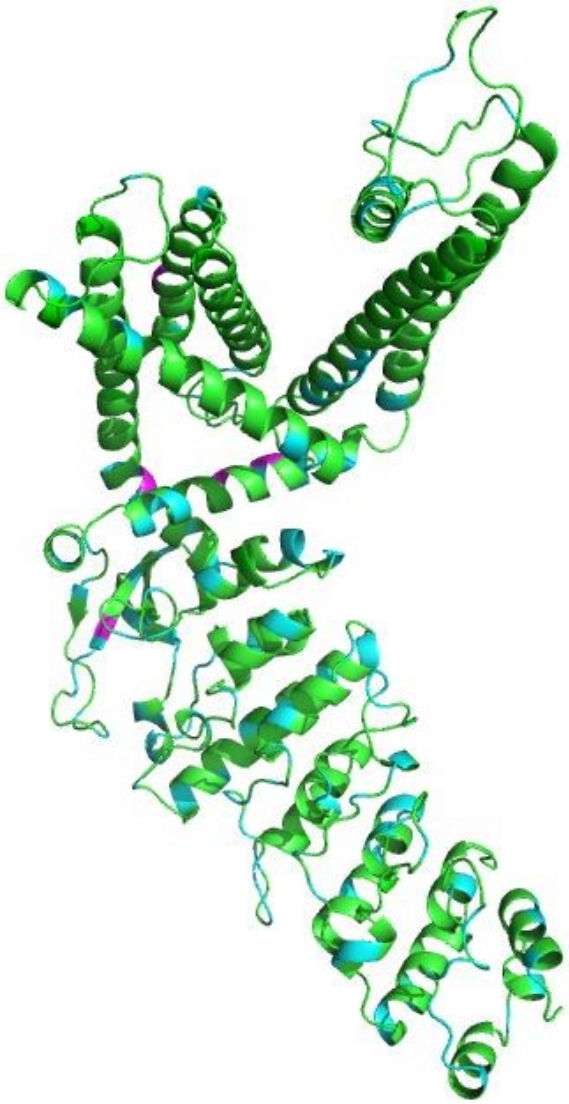

**Figure 14.** One subunit of human TRPV4 in green, the tryptophan residues are shown in magenta and the cyan colored are charged residues (histidine, aspartate, glutamate, arginine and lysine).

As expected, SiteMap located many cavities, one reason being the large size of the protein. SiteMap also identified the cavity that corresponds to the rat TRPV1 and rabbit TRPV5 VBS cavity. The properties of the cavity were explored further. The pocket map shows the shape of the cavity and the hydrogen-bond acceptor and donor regions, as well as the hydrophobic regions. The map indicated that the binding pocket is mostly hydrophobic with small hydrogen-bond donor areas at the edges of the binding pocket (Fig. 17). SiteMap also calculates the different scores of the cavity properties. These scores are meant to aid in locating the cavities that are large enough for binding and have the properties a druggable binding site usually needs. The scores also help to locate the best cavities in a protein, especially in situations when the binding site is not known, and when the protein investigated is large. SiteScore considers various calculations such as the size of the cavity, how enclosed the cavity is and the

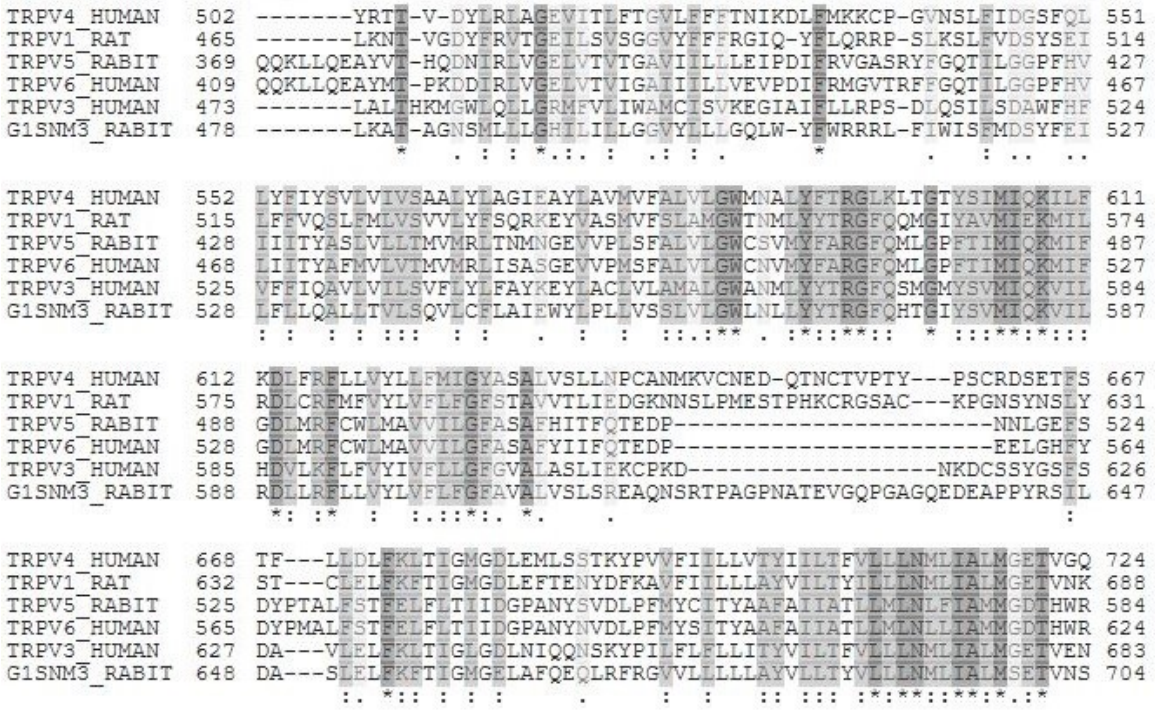


**Figure 15.** Multiple sequence alignment of the vanilloid binding area of TRPV proteins. The identical residues can be seen in dark grey, and the similar residues in two shapes of lighter grey. The multiple sequence alignment was performed with Clustal Omega.

hydrophilicity of the cavity. When SiteScore is over 1, the site is considered as a very potential target (Halgren, 2009). The druggability score (Dscore) is calculated similarly as SiteScore, but the hydrophobicity score is taken more into account in Dscore. Dscore separates the druggable cavities from cavities that are suited less or not at all for drug binding (Halgren, 2009). When Halgren (2009) calculated the Dscore value for several known binding sites in known proteins, the average value for druggable binding sites was 1.108. The SiteScore for the modeled human TRPV4 cavity was 1.271 and Dscore 1.400. These values indicate that the cavity is suitable for ligand binding, but also for drug binding. This site was also the best ranked site, of 30 different sites evaluated with SiteScore.

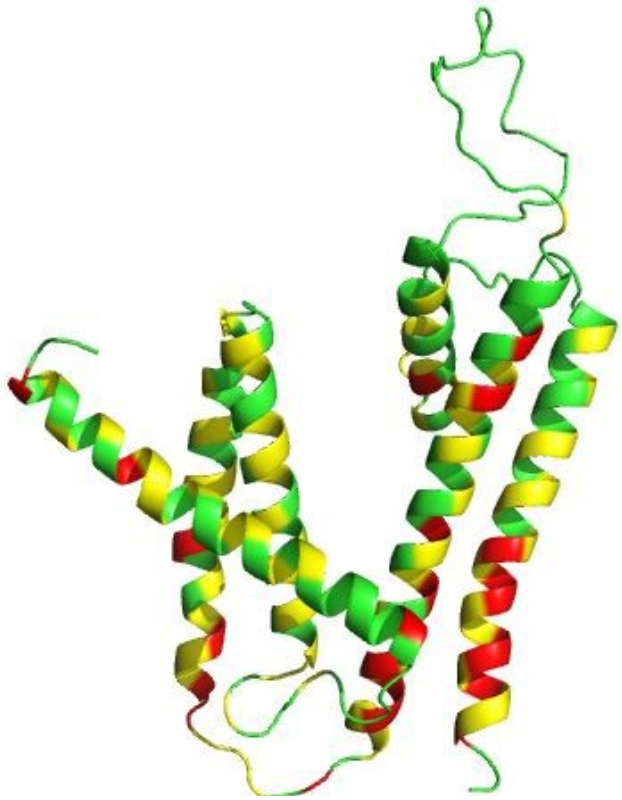

**Figure 16.** The conserved residues in human TRPV4 according to the TRPV protein alignment. Residues 502 – 724 from one subunit are shown, the residues correspond to S2 through S6. The protein is shown in green; the red area shows the location of the evolutionary identical residues and yellow area shows the locations of the similar residues.

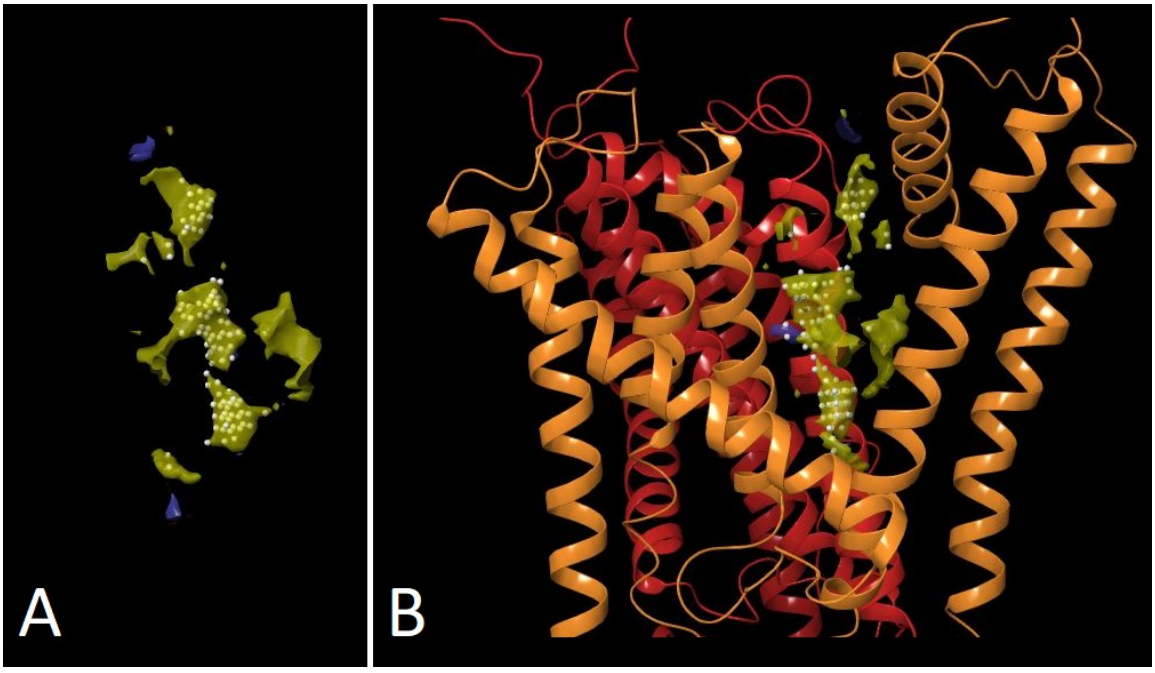


**Figure 17.** The VBS as identified with SiteMap. Map color code: yellow represents the hydrophobic regions, blue represents the hydrogen-bond donor regions and the white spheres represent the site points. A) the pocket map for the binding cavity; B) the pocket map inside the protein, two protein subunits are colored with orange and red. The figures were prepared with Maestro.

### 5.3 Virtual screening

#### 5.3.1 With Glide and GOLD

Twenty-eight molecules from the Glide docking and 50 molecules from the GOLD screening were included in the Prime MM-GBSA calculations. The results of the Glide and GOLD docking and Prime MM-GBSA calculations can be seen in Tables 1 and 2, respectively. Eight best scoring ligands from both docking groups were selected. The tables show the Enamine ID, MM-GBSA score, XP score or GOLD fitness score, molecular weight, calculated log P, the water solubility as log S and the number of hydrogen bond donors (HBD) of the eight best ranked molecules. In both Glide and GOLD docking groups, the ligand that had the best MM-GBSA score also had the best Glide XP score or GOLD fitness score. The table also lists which ligands were selected for the MD simulations. The criteria for the selection are described in section 5.3.3. The docked ligands are visualized in Figures 18 and 19. Superimposed figures of all the best scoring ligands without showing the pockets are also presented (Fig. 18I and 19I). The MM-GBSA values are in general higher in the GOLD docking group compared to the Glide docking group.

**Table 1.** Molecular properties and docking scores for the eight best scoring ligands in the Glide docking group

| Enamine ID | MW (g/mol) | CLogP | Log S | HBD | XP score | MM-GBSA | Selected |
|---|---|---|---|---|---|---|---|
| Z1053001754 | 480.6 | 4.838 | -6.464 | 2 | -9.889 | -65.87 | ✓ |
| Z118339292 | 401.5 | 2.615 | -6.362 | 2 | -9.690 | -52.68 | |
| Z1436055452 | 370.4 | 2.586 | -4.487 | 2 | -9.377 | -54.73 | |
| Z169988240 | 471.6 | 4.891 | -6.23 | 2 | -9.783 | -63.77 | ✓ |
| Z31055919 | 382.4 | 4.208 | -7.21 | 1 | -9.160 | -54.26 | ✓ |
| Z18691736 | 450.5 | 4.029 | -6.09 | 1 | -9.090 | -53.21 | |
| Z1213735368 | 429.4 | 4.613 | -6.56 | 2 | -9.491 | -56.91 | ✓ |
| Z32325702 | 376.4 | 4.407 | -6.25 | 1 | -9.062 | -53.69 | |

**Table 2.** Molecular properties and docking scores for the eight best scoring ligands in the GOLD docking group

| Enamine ID | MW (g/mol) | CLog P | Log S | HBD | GOLD fitness | MM-GBSA | Selected |
|---|---|---|---|---|---|---|---|
| Z30995330 | 375.5 | 2.884 | -3.486 | 0 | 82.083 | -61.80 | ✓ |
| Z146317224 | 397.5 | 4.447 | -5.197 | 1 | 80.400 | -59.74 | |
| Z1624971550 | 490.0 | 5.652 | -6.7 | 1 | 75.060 | -60.95 | |
| Z229153032 | 457.6 | 4.506 | -5.5 | 2 | 76.829 | -63.45 | ✓ |
| Z1157726398 | 501.6 | 3.357 | -7.04 | 1 | 77.728 | -61.29 | ✓ |
| Z29940381 | 466.5 | 5.699 | -7.85 | 2 | 83.858 | -70.22 | ✓ |
| Z51867077 | 470.9 | 3.238 | -6.891 | 1 | 82.111 | -67.91 | ✓ |
| Z221487176 | 448.9 | 4.079 | -6.067 | 2 | 82.300 | -60.48 | ✓ |

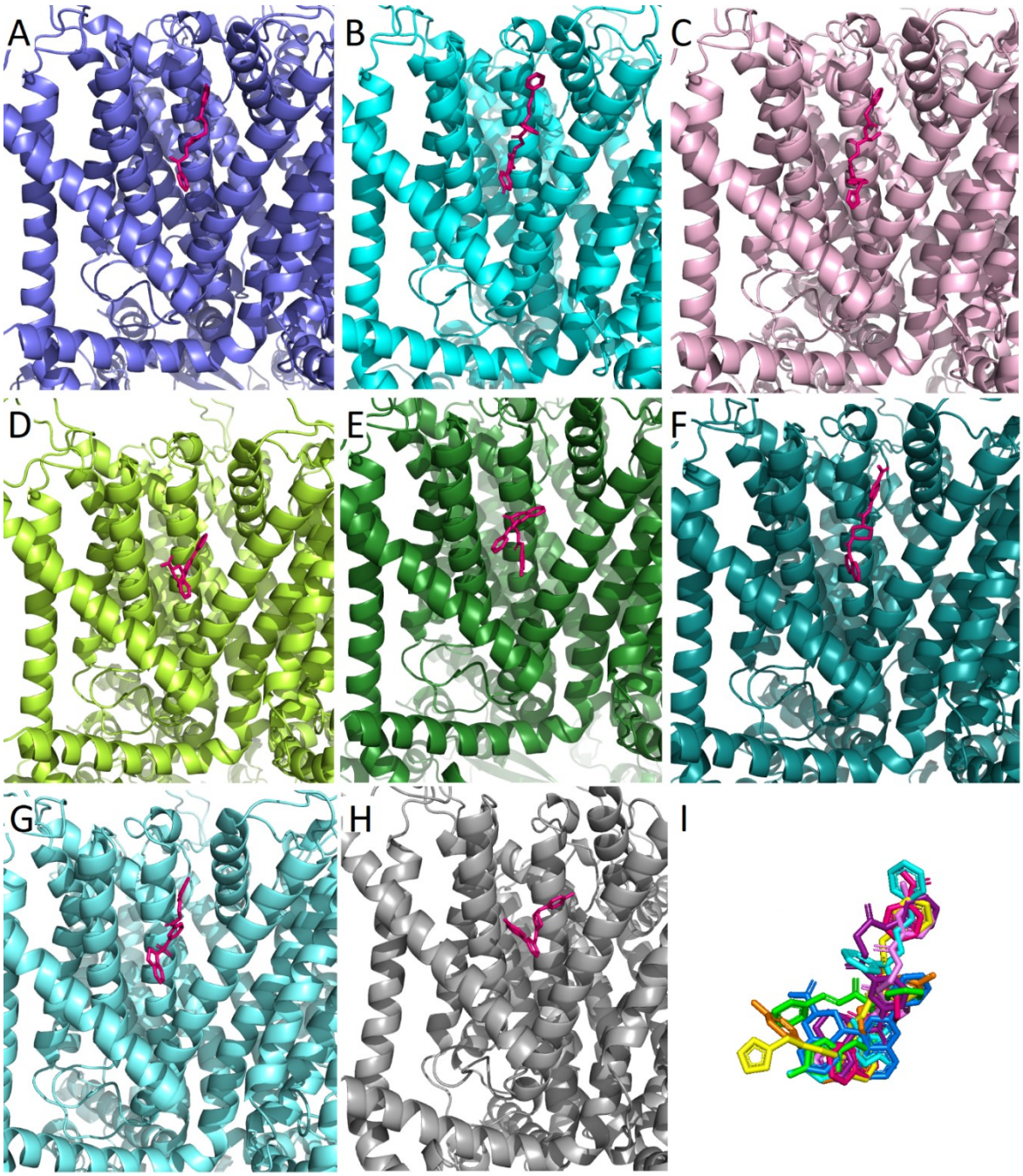


**Figure 18**. The docked ligands from the Glide screening. The ligands are visualized in hot pink and the human TRPV4 protein in various colors. A) Z1436055452, B) Z169988240, C) Z1053001754, D) Z31055919, E) Z1213735368, F) Z118339292, G) Z18691736, H) Z32325702. I) all the docked ligands superimposed and seen from side.

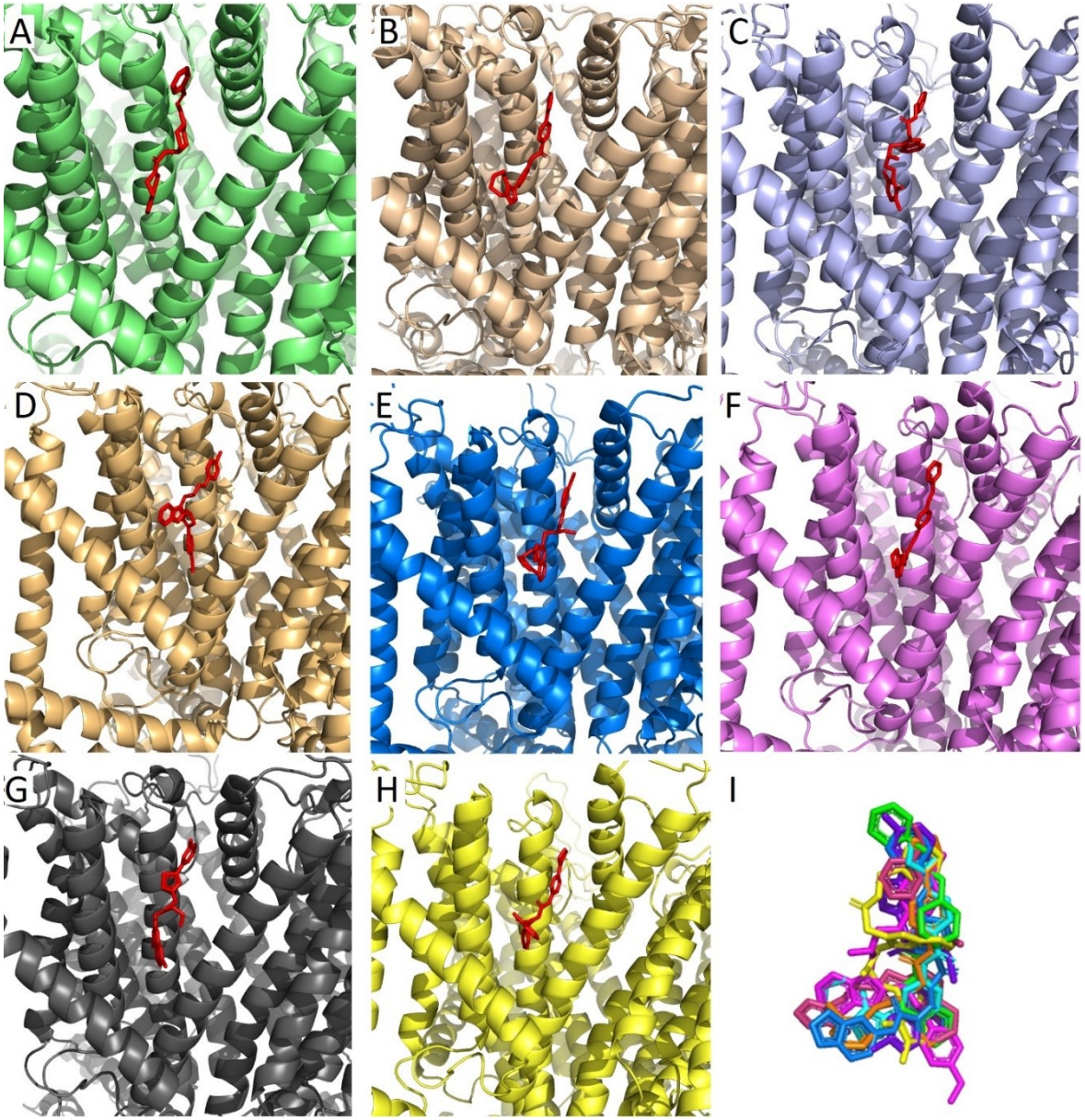


**Figure 19.** The docked ligands from the GOLD screening. The ligands are visualized in red and the human TRPV4 protein in various colors. A) Z30995330, B) Z146317224, C) Z229153032, D) Z1624971550, E) Z1157726398, F) Z29940381, G) Z51867077 and H) Z221487176. I) all the docked ligands superimposed and seen from side.

### 5.3.2 Selection of the ten most promising molecules for MD simulations

The selection of the ten most promising antagonist candidates was made foremostly by comparing the MM-GBSA results from both GLIDE and GOLD screenings, but also by choosing structurally different molecules. Four molecules from the GLIDE screening and six molecules from the GOLD screening were selected.

The 2D structures of the selected molecules from the GLIDE screening are visualized in Figure 20 and from the GOLD screening in Figure 21. The protein and molecule complexes were renamed by giving each complex a number between 1 and 10. The protein and molecules from the GLIDE screening were renamed as following; Z1053001754 = 1, Z1213735368 = 2, Z169988240 = 3, Z31055919 = 4 and the protein and molecules from GOLD screening were renamed as following; Z30995330 = 5, Z229153032 = 6, Z1157726398 = 7, Z29940381 = 8, Z51867077 = 9, Z221487176 = 10.

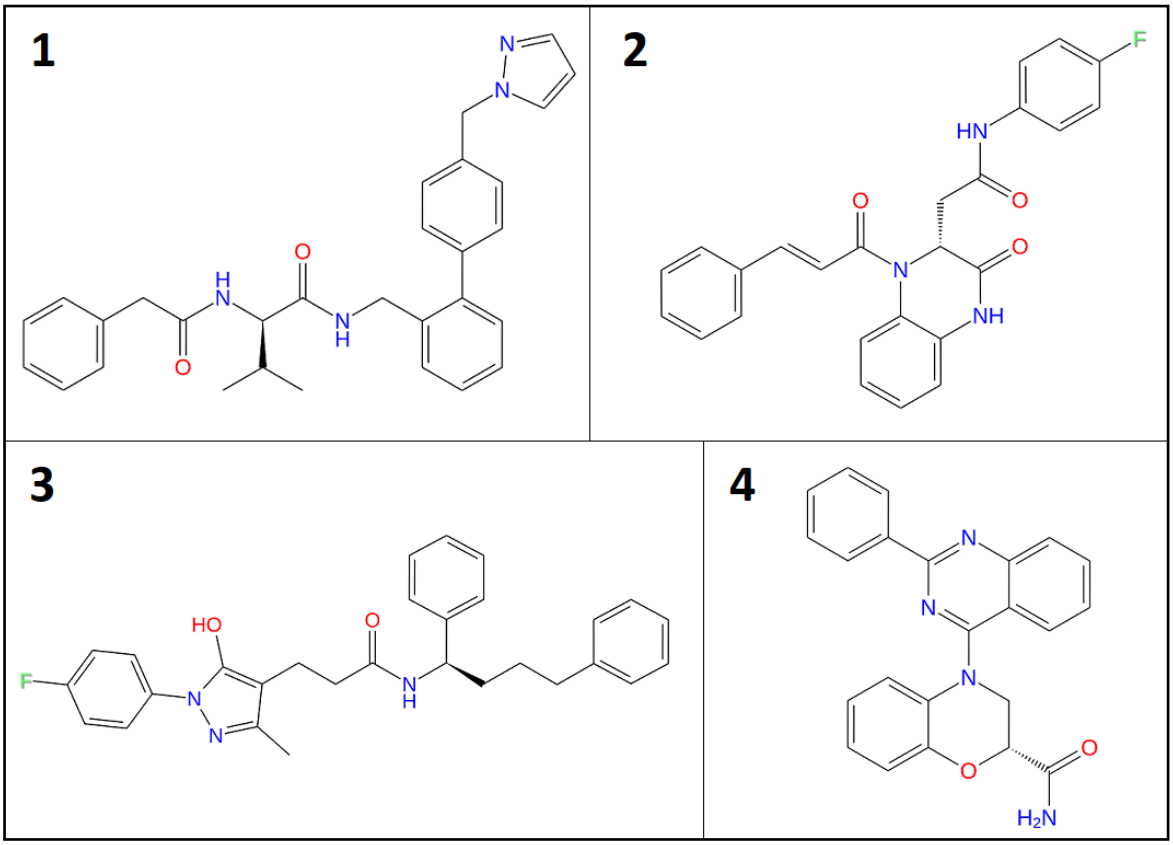


**Figure 20.** 2D structures of the selected molecules from the GLIDE screening. The ligands represented are following; 1 = Z1053001754, 2 = Z1213735368, 3 = Z169988240 and 4 = Z31055919. The pictures were obtained from the Desmond simulation output file.

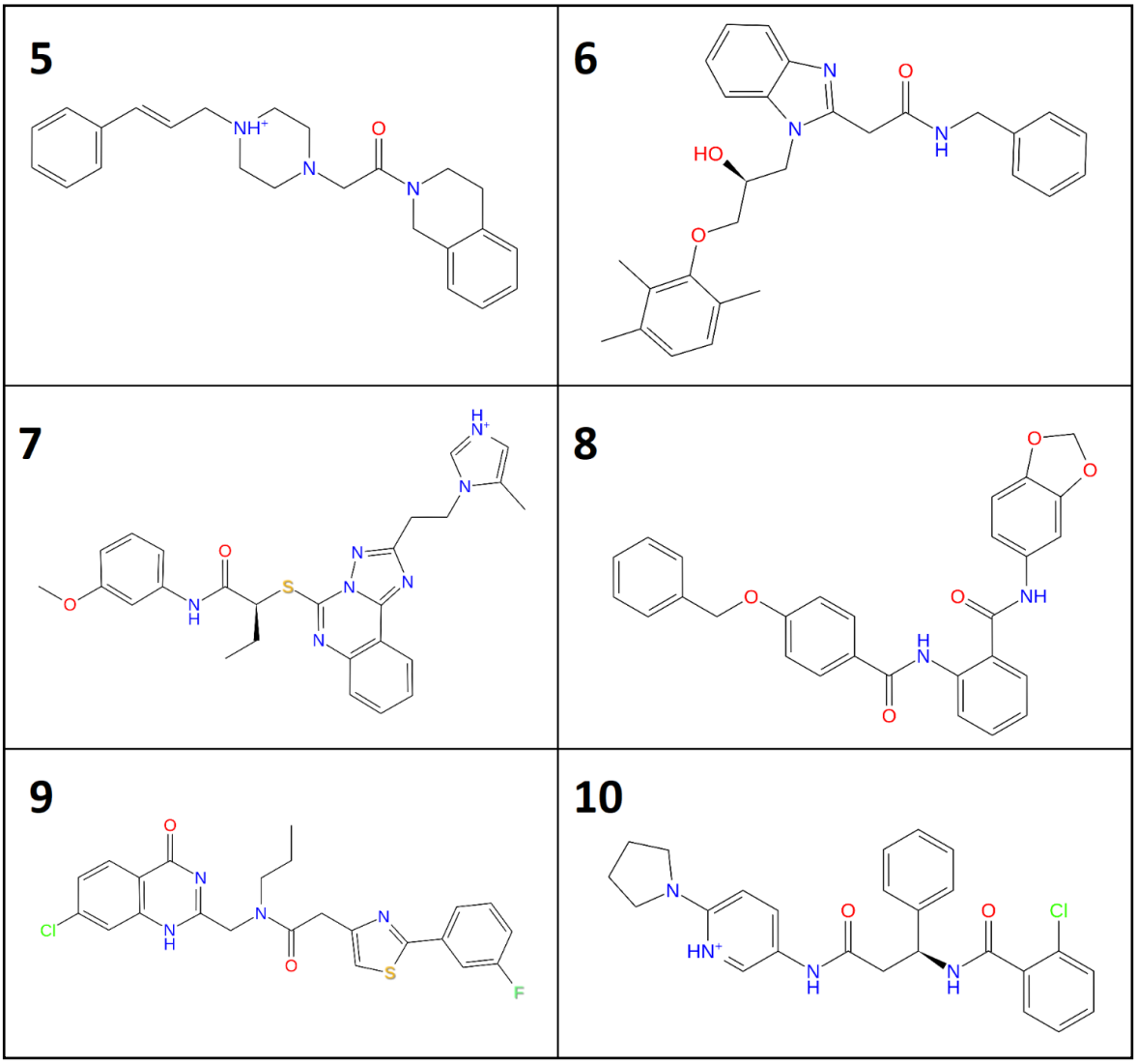


**Figure 21.** 2D structures of the selected molecules from the GOLD screening. The ligands represented are following; 5 = Z30995330, 6 = Z229153032, 7 = Z1157726398, 8 = Z29940381, 9 = Z51867077 and 10 = Z221487176. The pictures were obtained from the Desmond simulation output file.

### 5.4 Molecular dynamics simulations

#### 5.4.1 RMSF analysis

The MD simulation RMSF plots of the protein-ligand complexes are presented in Figure 22 (complexes 1 to 4), Figure 23 (complexes 5 to 8) and Figure 24 (complexes 9 to 10). All plots displayed similar patterns, but the protein fluctuation patterns were not identical between the different protein-ligand complexes. The largest peaks were mostly observed in the extracellular loop regions and at the subunit N- and C-terminals. In addition, larger peaks could be observed in other unstructured loop regions in complexes 1 and 8. In general,

there appeared to be more fluctuation in the protein-ligand complexes when compared to the protein model alone.

### 5.4.2 RMSD and simulation video analysis

The RMSD plots are presented in Figure 25 (complexes 1 and 2), Figure 26 (complexes 3 and 4), Figure 27 (complexes 5 and 6), Figure 28 (complexes 7 and 8) and Figure 29 (complexes 9 and 10). The ligand RMSD value provides information about the stability of the ligand in the binding pocket. A stable ligand does not have a significantly higher RMSD value compared to the protein RMSD value. A considerably higher ligand RMSD value indicates weak or no binding interactions with the protein. In addition, the ligand migrating away from the binding site during the simulation is probable (Desmond 4.2 User Manual, 2015).

The RMSD plot for complex 1 predicted a stable protein and a quite stable binding of the ligand. The protein equilibrated at around 20 ns and for the rest of the simulation the fluctuation remained at around 5.6 Å. The ligand RMSD value kept below the protein RMSD value during most of the simulation and the simulation video confirmed that the ligand remained in the binding cavity. The simulation video visualized that the middle part of the ligand stayed quite fixed but the aromatic rings at the ligand ends rotated intensely.

The RMSD plot for complex 2 revealed a stable protein and a ligand which had for most of the simulation a slightly higher RMSD value compared to the protein RMSD. Towards the end of the simulation, the RMSD values for both the ligand and the protein were quite similar. At the end of the simulation, the protein RMSD was still increasing, so it is possible that the system never equilibrated during the simulation. The simulation video showed that the ligand remained in the binding cavity and kept a quite similar position throughout the simulation

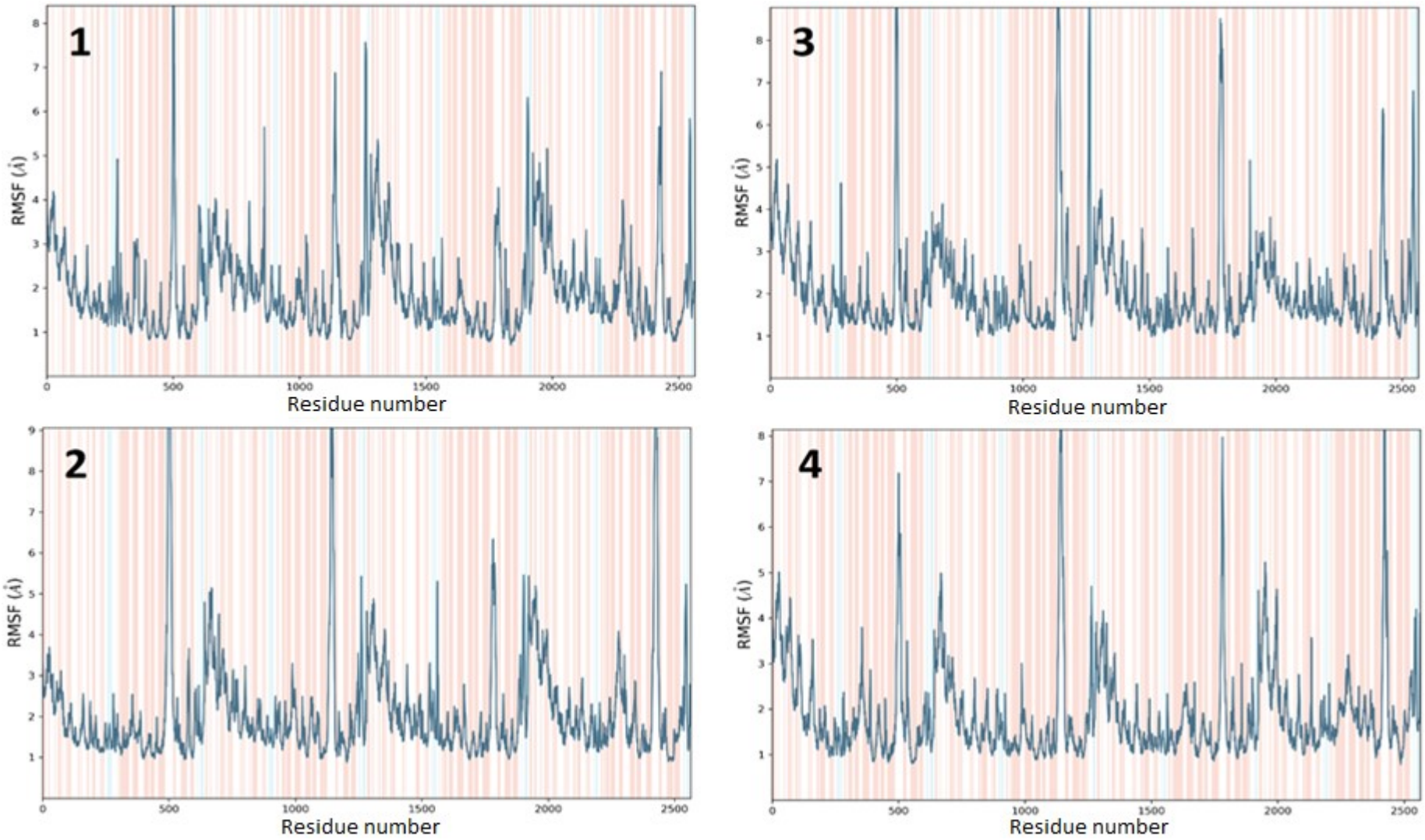


**Figure 22.** The RMSF plots for complexes 1-4 (labeled). Note, that the residue number 1 in the plot corresponds to the residue number 148 because the plot does not show the original N- or C-terminal of the human TRPV4 since these parts were not modeled. The white areas show the unstructured sections, the light red areas show the alpha helical sections, and the light blue areas show the beta strand sections of the protein.

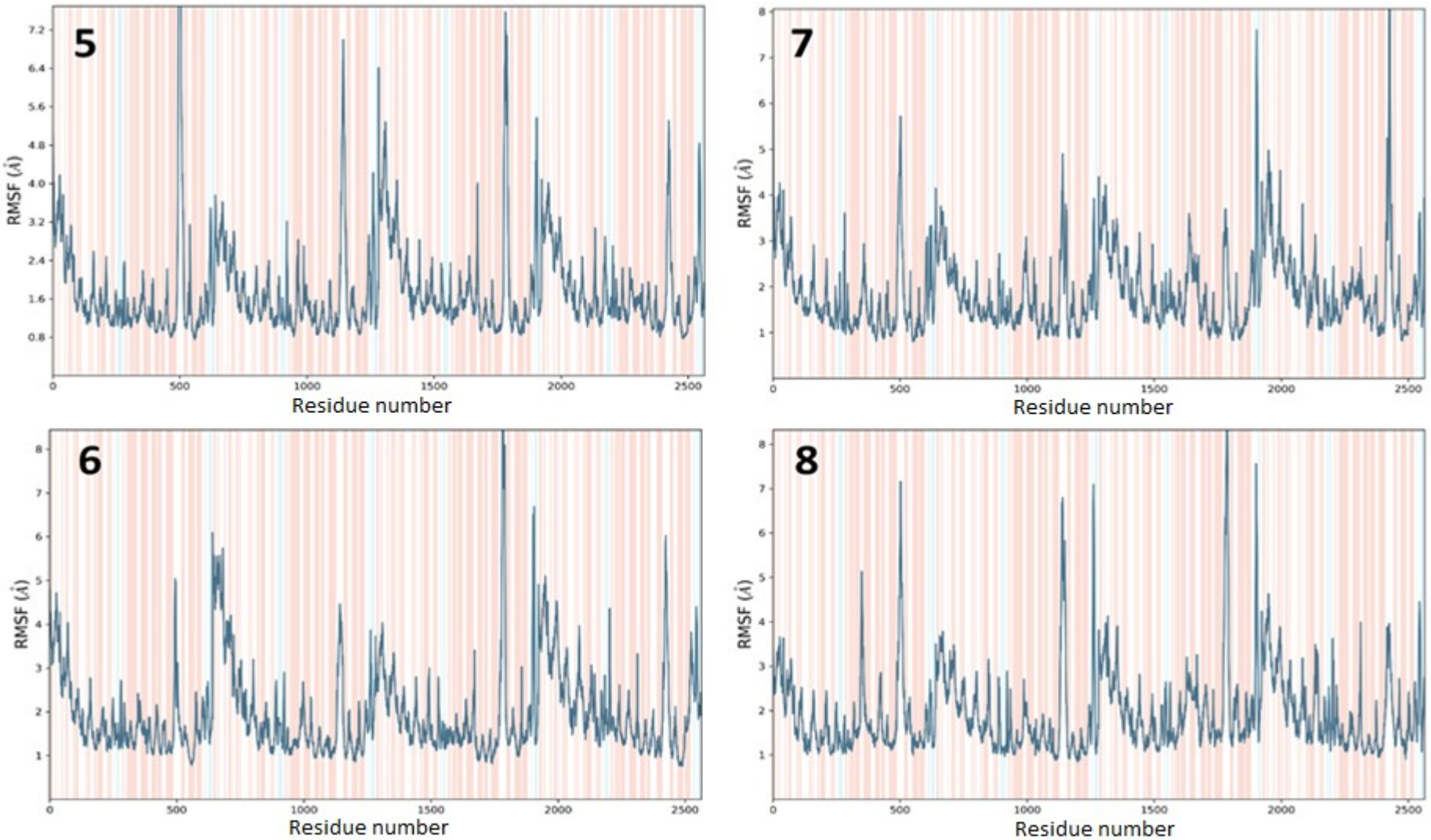


**Figure 23.** The RMSF plots for complexes 5-8 (labeled). Note, that the residue number 1 in the plot corresponds to the residue number 148 because the plot does not show the original N- or C-terminal of the human TRPV4 since these parts were not modeled. The white areas show the unstructured sections, the light red areas show the alpha helical sections, and the light blue areas show the beta strand sections of the protein.

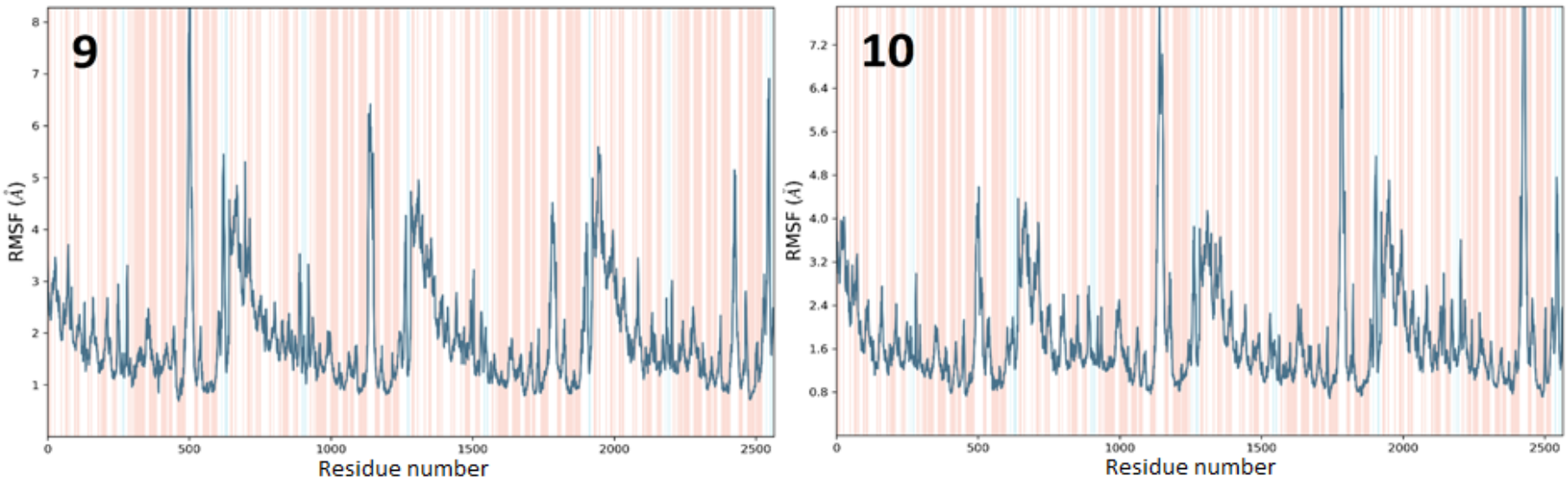


**Figure 24**. The RMSF plots for complexes 9-10 (labeled). Note, that the residue number 1 in the plot corresponds to the residue number 148 because the plot does not show the original N- or C-terminal of the human TRPV4 since these parts were not modeled. The white areas show the unstructured sections, the light red areas show the alpha helical sections, and the light blue areas show the beta strand sections of the protein.

The RMSD plot for complex 3 predicted a stable protein, but the ligand was not very stable in the binding cavity. The ligand RMSD altered somewhat and the ligand RMSD value was higher than the corresponding protein RMSD value during some parts of the simulation. The simulation video visualized that the ligand stayed in the binding cavity, but parts of the ligand were rotating considerably. The ligand underwent bigger changes in its pose a few times.

The RMSD plot for complex 4 was quite like the plot for complex 3. The ligand RMSD was somewhat larger than the protein RMSD throughout most of the simulation. The protein equilibrated at around 150 ns and for the rest of the simulation the fluctuation remained at around 5.4 Å. When comparing the simulation videos, it was noticeable that towards the end of the simulation the ligand in complex 4 was not altering its position as much as the ligand in complex 3.

The worst RMSD result was obtained with complex 5. The ligand RMSD value was significantly higher that the protein RMSD value throughout the simulation. The simulation video confirmed that the ligand did not bind into the cavity. Instead, it slid away from the cavity towards the extracellular parts of the protein.

The results from complex 6 RMSD plot and video were quite like the results obtained from the simulation involving complex 4. Both protein and ligand stabilized towards the end of the simulation. Though, it is possible that the protein did not fully equilibrate during the simulation since the protein RMSD was still increasing at the end of the simulation.

The simulation video of complex 7 revealed that the ligand underwent a large structural change at the beginning of the simulation before stabilizing. The structural change is seen in the RMSD plot by the big leap of the ligand RMSD value at around 90 ns. The protein equilibrated at around 100 ns and for the rest of the simulation the fluctuation remained at around 4.6 Å.

The RMSD plot for complex 8 indicated and the simulation video confirmed that the ligand structure altered and moved in the binding cavity. The same applied to complex 9, but here the difference between the corresponding RMSD values between the protein and ligand were higher compared to complex 8. Thus, the ligand binding in complex 9 was less stable. In complex 8 the protein equilibrated at around 50 ns and in complex 9 at around 100 ns. The fluctuation remained at around 4.8 Å for both proteins during the rest of the simulation.

The simulation video and the RMSD plot for complex 10 revealed large conformational changes for the ligand at 70 ns. The ligand stabilized after that. The protein was stable during the simulation, equilibrating at around 50 ns. The fluctuation remained at around 5.2 Å during the rest of the simulation.

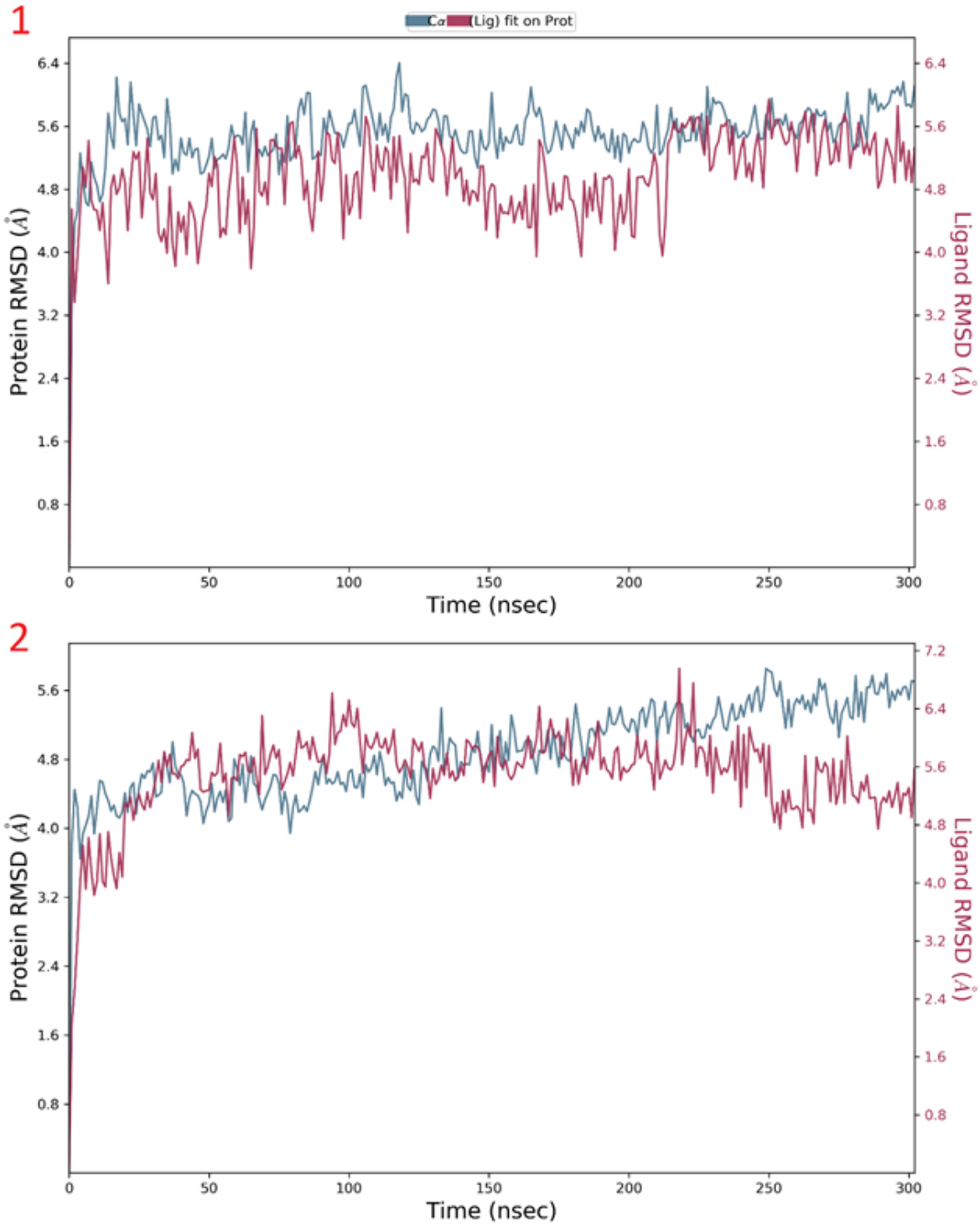


**Figure 25.** The RMSD plot for complex 1 (above) and complex 2 (below). X-axis shows the simulation time in ns. The RMSD fluctuation for the protein is presented in blue and the value seen in the left Y-axis. The RMSD fluctuation for the ligand is presented in magenta and the value seen in the right Y-axis.

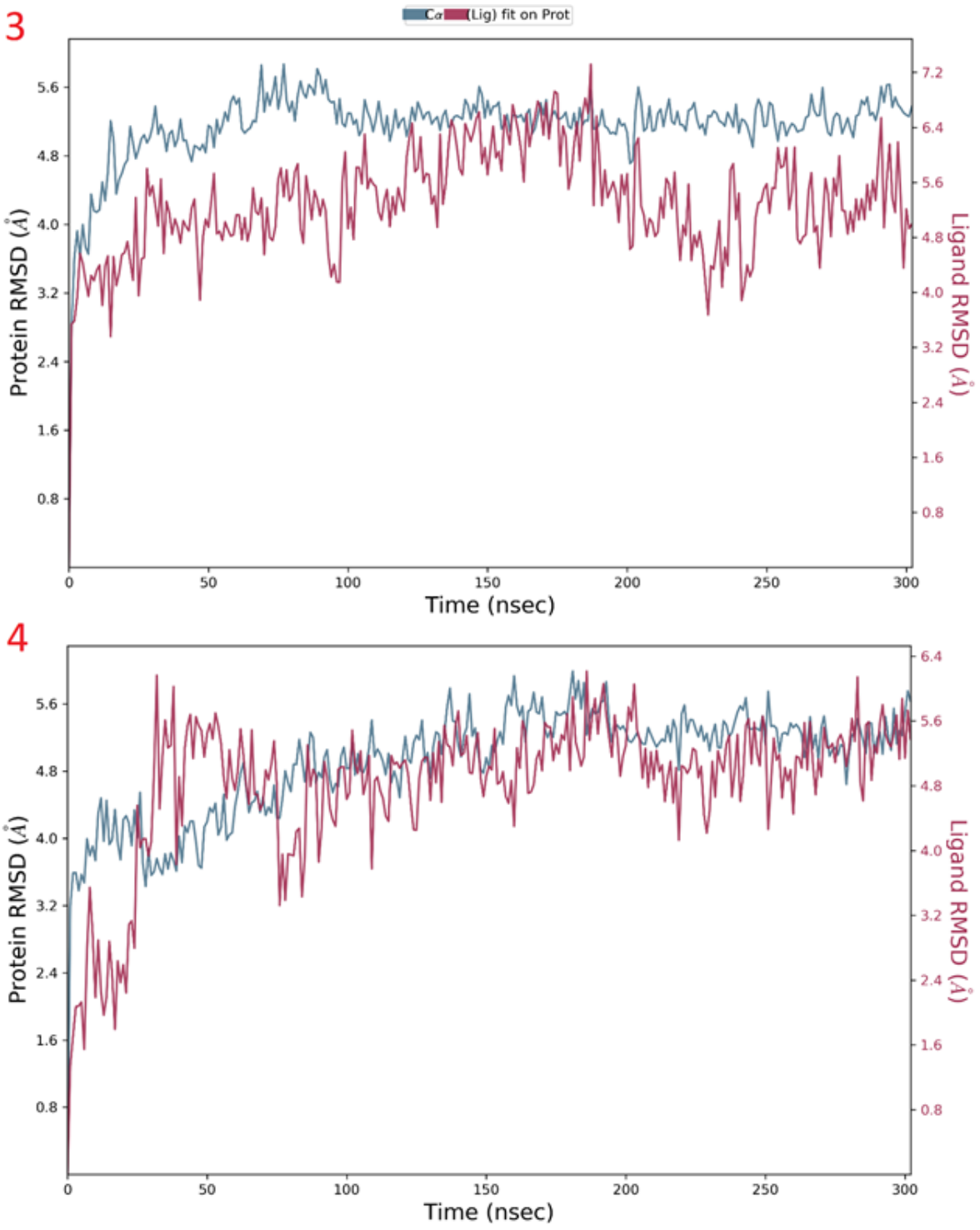


**Figure 26.** The RMSD plot for complex 3 (above) and complex 4 (below). X-axis shows the simulation time in ns. The RMSD fluctuation for the protein is presented in blue and the value seen in the left Y-axis. The RMSD fluctuation for the ligand is presented in magenta and the value seen in the right Y-axis.

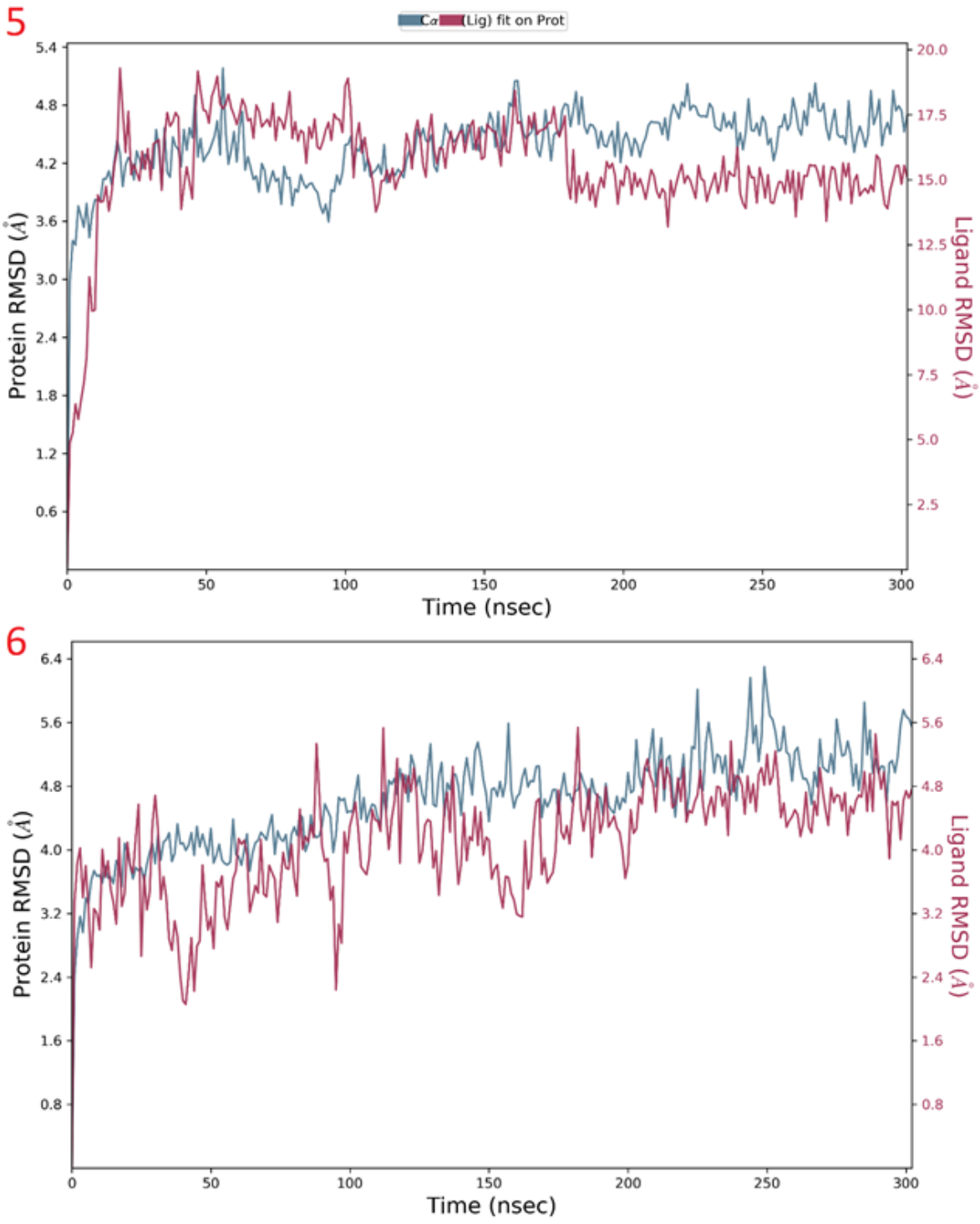


**Figure 27.** The RMSD plot for complex 5 (above) and complex 6 (below). X-axis shows the simulation time in ns. The RMSD fluctuation for the protein is presented in blue and the value seen in the left Y-axis. The RMSD fluctuation for the ligand is presented in magenta and seen the value in the right Y-axis.

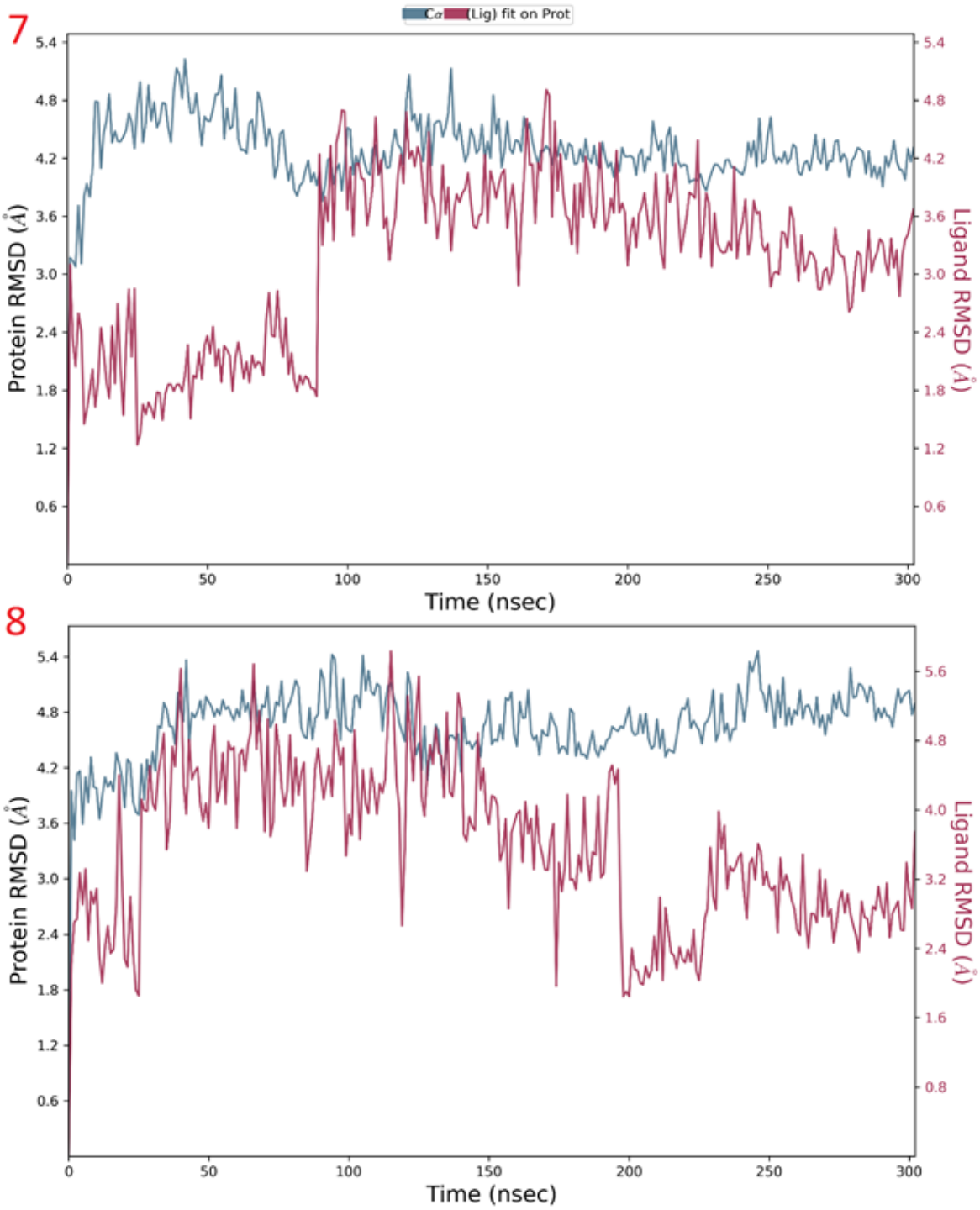


**Figure 28.** The RMSD plot for complex 7 (above) and complex 8 (below). X-axis shows the simulation time in ns. The RMSD fluctuation for the protein is presented in blue and the value seen in the left Y-axis. The RMSD fluctuation for the ligand is presented in magenta and the value seen in the right Y-axis.

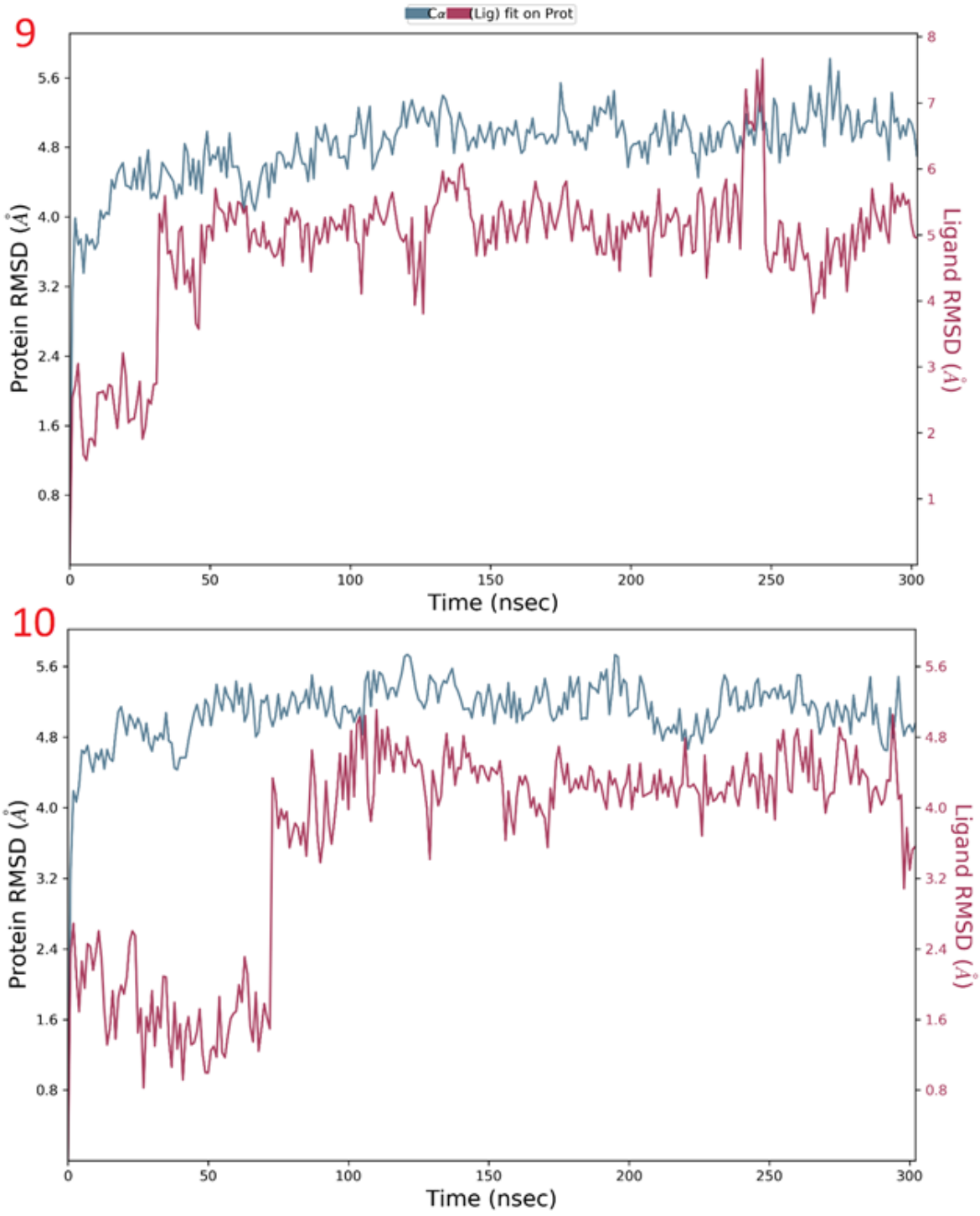


**Figure 29.** The RMSD plot for complex 9 (above) and complex 10 (below). X-axis shows the simulation time in ns. The RMSD fluctuation for the protein is presented in blue and the value seen in the left Y-axis. The RMSD fluctuation for the ligand is presented in magenta and the value seen in the right Y-axis.

### 5.4.3 Protein-ligand interactions

The protein-ligand interaction plot presents the contacts the ligand and the protein form on residue level during the simulation. The value of the interactions fraction can easily be converted to percentage, for example the interactions fraction 0.4 equals 40%. The percentage gives information on how long a certain residue interacts with the ligand during the simulation. It is also specified if the contacts are hydrophobic or ionic, or if the binding partners form a hydrogen bond or a water bridge.

The protein-ligand interaction plots for complexes 1-10 are presented in Figures 30-34. Majority of the contacts observed during the simulations were hydrophobic, this was expected since the binding pocket was hydrophobic according to the results from SiteMap. All 10 ligands had contacts between multiple residues, but most of the contacts were short-lived, only about 10% or less of the simulation time. In complex 8, 28 residues interacted with the ligand.

Ligands with mainly hydrophobic contacts were found in complexes 1, 3, 4, 6, 8, 9 and 10. Ligands in complexes 3 and 9 had both low interaction fraction values for the contacts, indicating low affinity. Ligands in complexes 1 and 4 formed long-term interactions with Phe 554 and ligands in complexes 6, 8 and 10 formed long-term interactions with Phe 674. Here, long-term interactions mean interactions that occur for more than 80% of the simulation time. If the interaction fraction value is higher than 1 (100%), then more than one interaction is formed between the ligand and the residue. The interaction types can be similar or different. Phe 674 and Phe 554 are both conserved residues. Phe 674 is an identically and Phe 554 is a similarly conserved residue according to the multiple sequence alignment (see Figure 15 in section 5.2).

Only two complexes showed long-term hydrogen bonds between the ligand and a residue: ligand in complex 7 with Val 560 and ligand in complex 2 with Phe 615. The ligand in complex 7 further had a long-term hydrophobic interaction with Phe 525. Hydrogen bonds between the ligand and the protein are considered important when determining the ligand-protein specificity (Wade & Goodford, 1989). Val 560, Phe 615 and Phe 525 residues were not conserved residues according to the multiple sequence alignment (see Figure 15 in section 5.2). These residues could possibly contribute to ligand selectivity for TRPV4.

Complex 5 showed all four types of interactions, but the interactions were all short-termed. This was expected since the molecule fluctuated away from the binding site during the simulation. No complexes showed long-termed water bridges or ionic contacts. None of the long-term contacts were formed with the corresponding VBS residues presented in section 2.4. The most interesting complexes were 1, 2, 7 and 10. Complexes 1 and 10 showed the longest lasting hydrophobic contacts and complexes 2 and 7 the longest lasting hydrogen bond contacts.

Since the protein-ligand interaction plots do not show when during the simulation the interactions occur, the plot that visualizes the simulation timeline with interactions and contacts were inspected for complexes 1, 2, 7 and 10. This plot also revealed if there were multiple contacts or interactions occurring at the same time between the ligand and one amino acid. The plots are visualized in Figures 35 and 36.

The interactions between Phe 554 and the ligand in complex 1 are not persistent and there are gaps during the simulation (the white area), meaning that there are no interactions between the ligand and the residue at that moment. In addition, there are parts where multiple contacts with the ligand have been formed, this is the reason why the interactions fraction value is over 100%.

In complex 10 the interaction pattern between Phe 527 and the ligand is quite like the pattern in complex 1, except for the gaps that are shorter, meaning the ligand is in contact with the residue more compared to the ligand in complex 1.

In complex 2, there are only a few gaps where the ligand and Phe 615 do not form a contact. Thus, the ligand forms multiple contacts with the residue during the simulation, which can also be seen in the interactions fraction figure.

In complex 7, a continuous interaction (H-bond) is formed at around 80 ns between the ligand and Val 560. There were no multiple interactions observed. The interactions between the ligand and Phe 525 were more frequent in the beginning of the simulation, but the gaps grew larger in size and number towards the end of the simulation.

The locations of the important residues are visualized in Figure 37. Phe 525 is located at the S2 helix, Phe 554 and Val 560 are located at the S3 helix, Phe 615 is located at the S5 helix and Phe 674 is located at the pore helix.

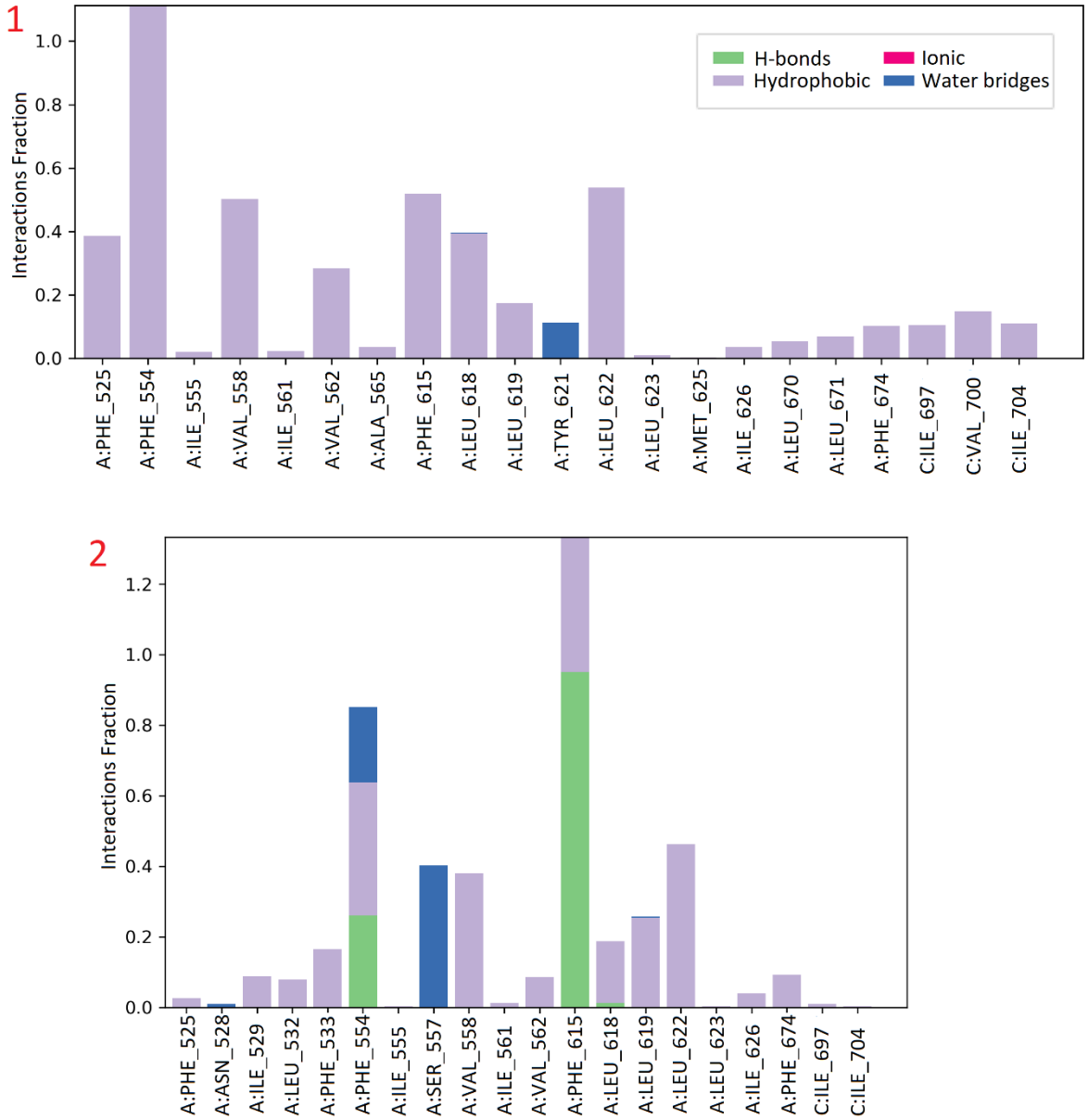


**Figure 30.** Protein-ligand interaction plots for complex 1 (above) and complex 2 (below). Interactions are shown as fractions from the total MD simulation time (300 ns). The residue numbers are presented on the X-axis and the interactions fraction value is presented on the Y-axis. The residue numbers correspond to the correct human TRPV4 residue numbers, the letter in front of the number (A-D) indicates the chain where the residues are located at.

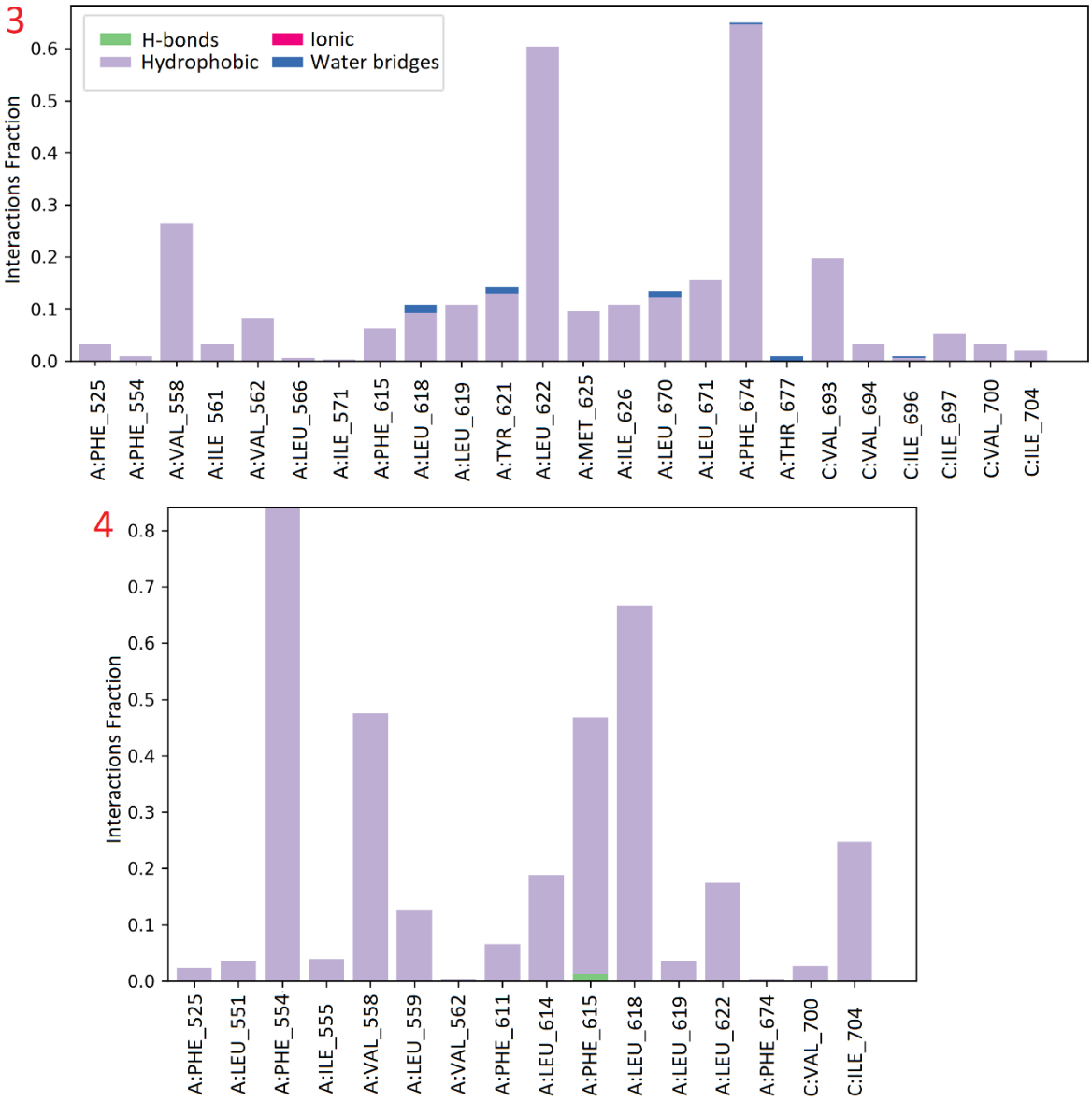


**Figure 31.** Protein-ligand interaction plots for complex 3 (above) and complex 4 (below). Interactions are shown as fractions from the total MD simulation time (300 ns). The residue numbers are presented on the X-axis and the interactions fraction value is presented on the Y-axis. The residue numbers correspond to the correct human TRPV4 residue numbers, the letter in front of the number (A-D) indicates the chain where the residues are located at.

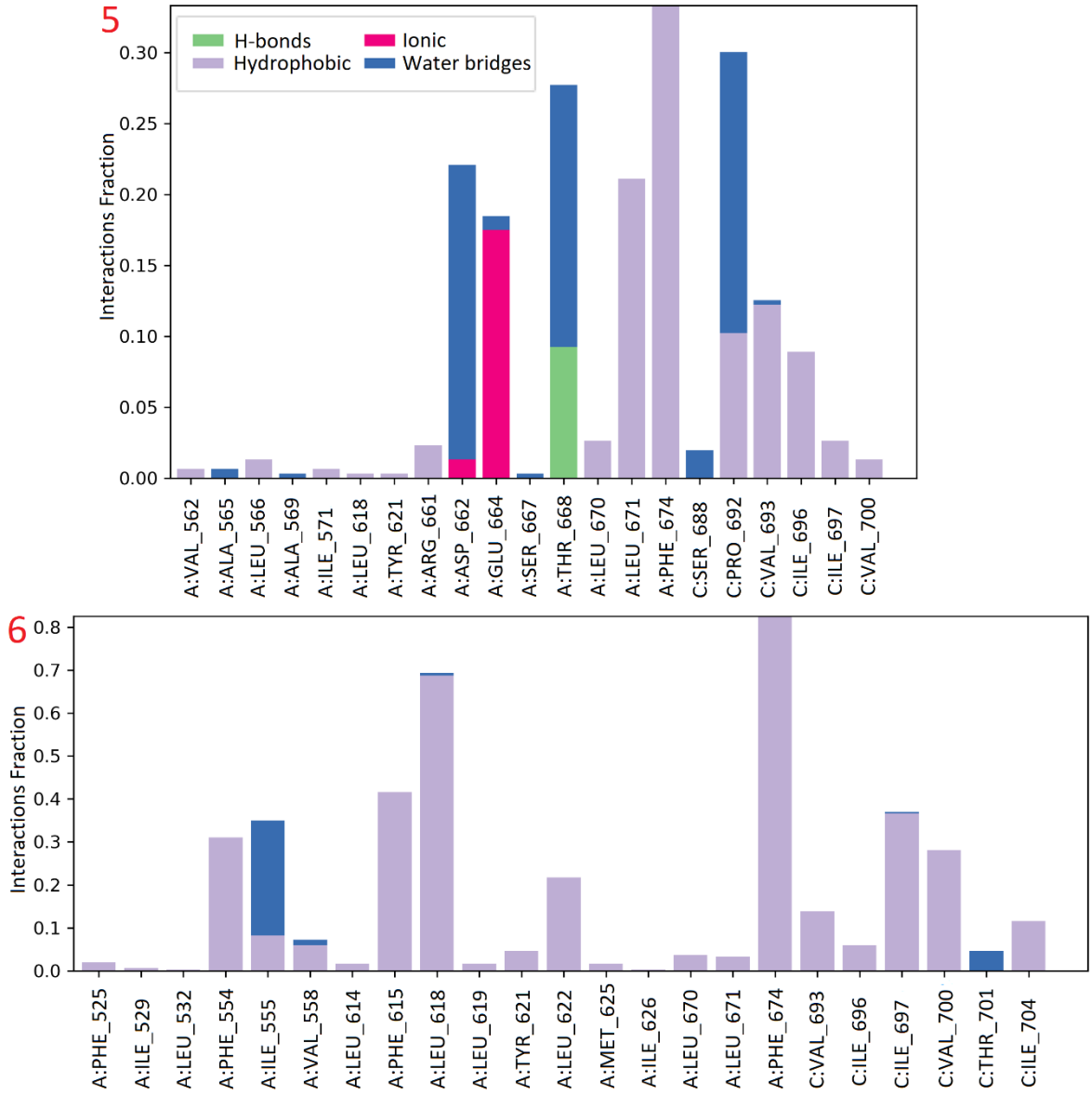


**Figure 32.** Protein-ligand interaction plots for complex 5 (above) and complex 6 (below). Interactions are shown as fractions from the total MD simulation time (300 ns). The residue numbers are presented on the X-axis and the interactions fraction value is presented on the Y-axis. The residue numbers correspond to the correct human TRPV4 residue numbers, the letter in front of the number (A-D) indicates the chain where the residues are located at.

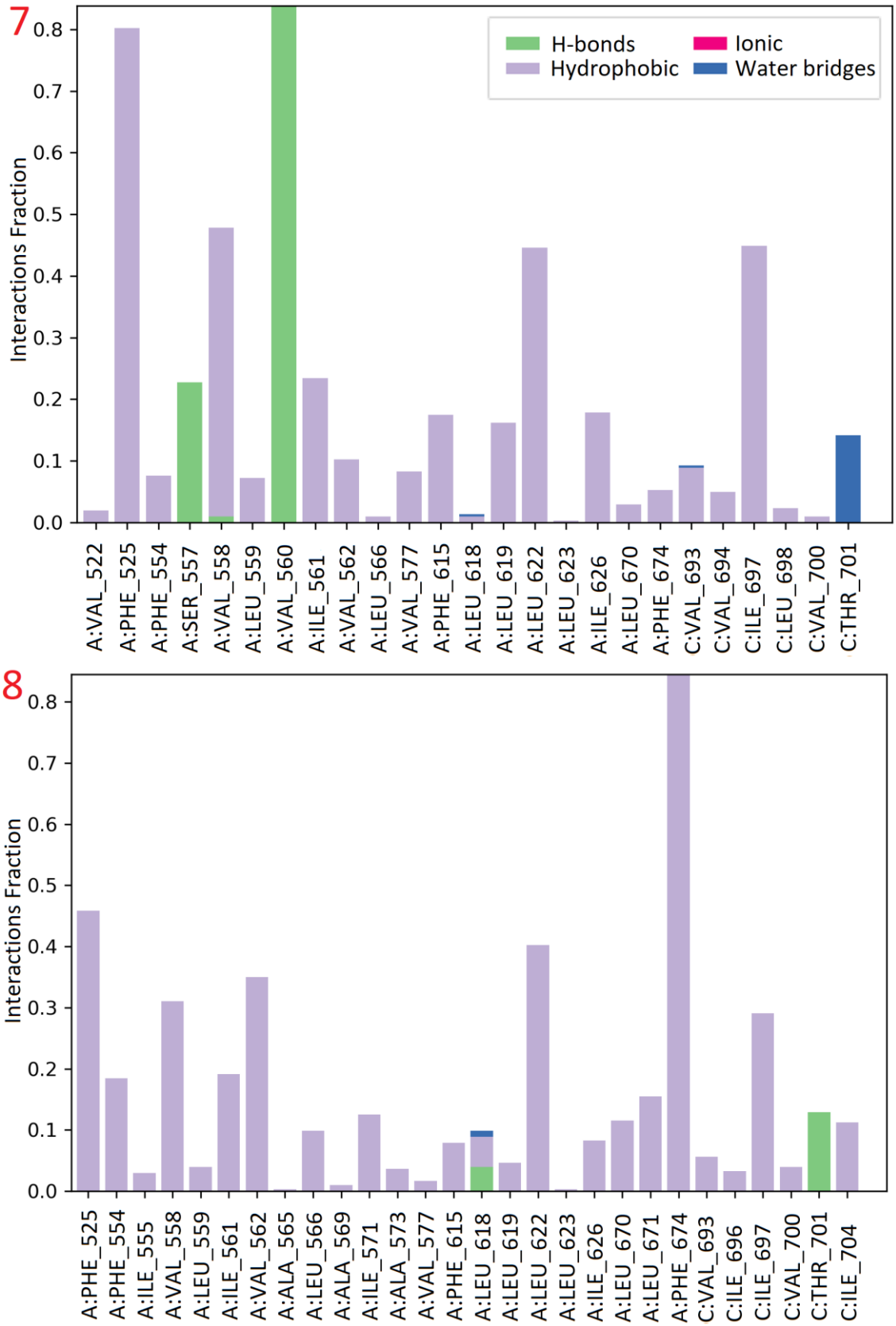


**Figure 33.** Protein-ligand interaction plots for complex 7 (above) and complex 8 (below). Interactions are shown as fractions from the total MD simulation time (300 ns). The residue numbers are presented on the X-axis and the interactions fraction value is presented on the Y-axis. The residue numbers correspond to the correct human TRPV4 residue numbers, the letter in front of the number (A-D) indicates the chain where the residues are located at.

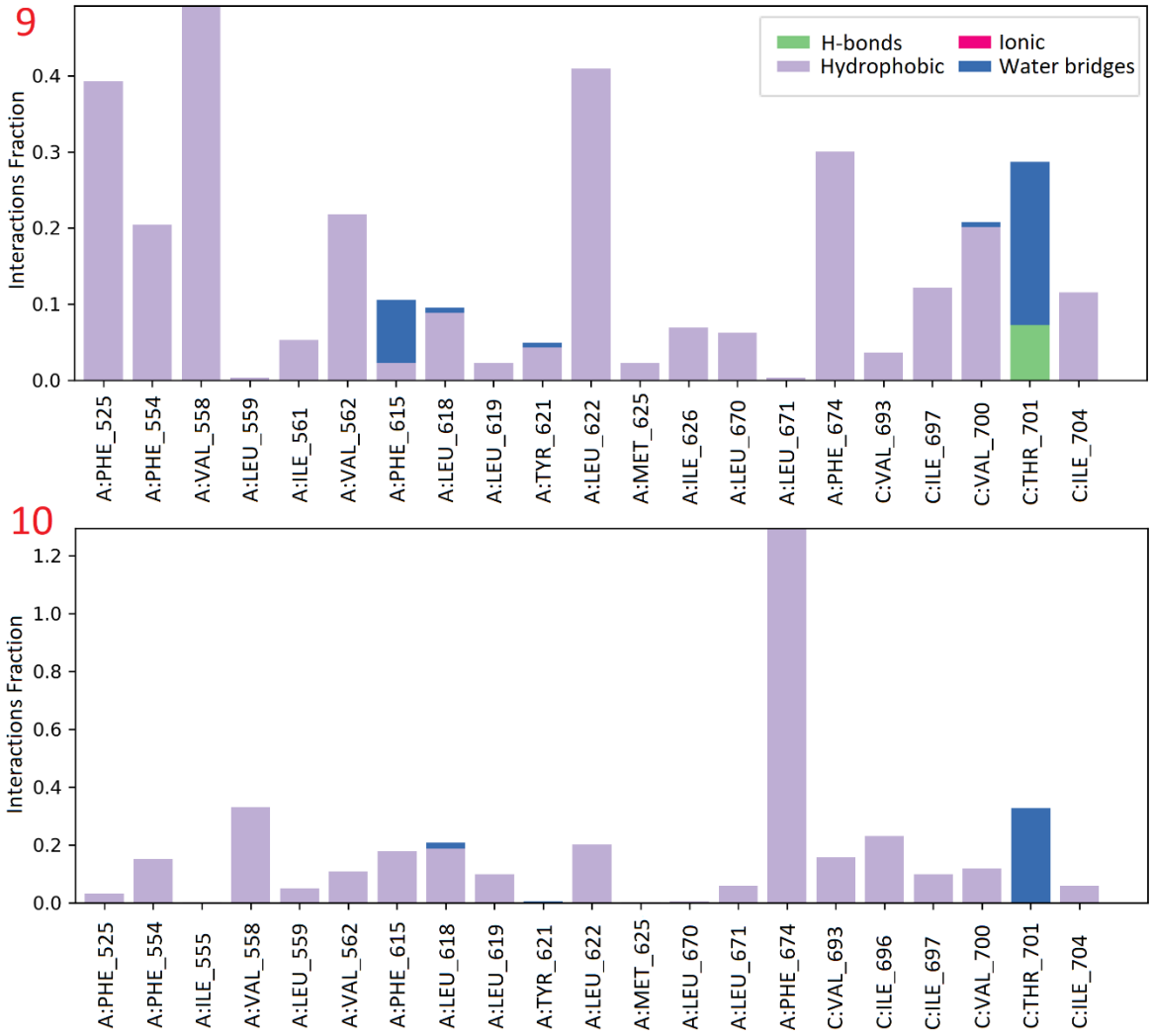


**Figure 34.** Protein-ligand interaction plots for complex 9 (above) and complex 10 (below). Interactions are shown as fractions from the total MD simulation time (300 ns). The residue numbers are presented on the X-axis and the interactions fraction value is presented on the Y-axis. The residue numbers correspond to the correct human TRPV4 residue numbers, the letter in front of the number (A-D) indicates the chain where the residues are located at.

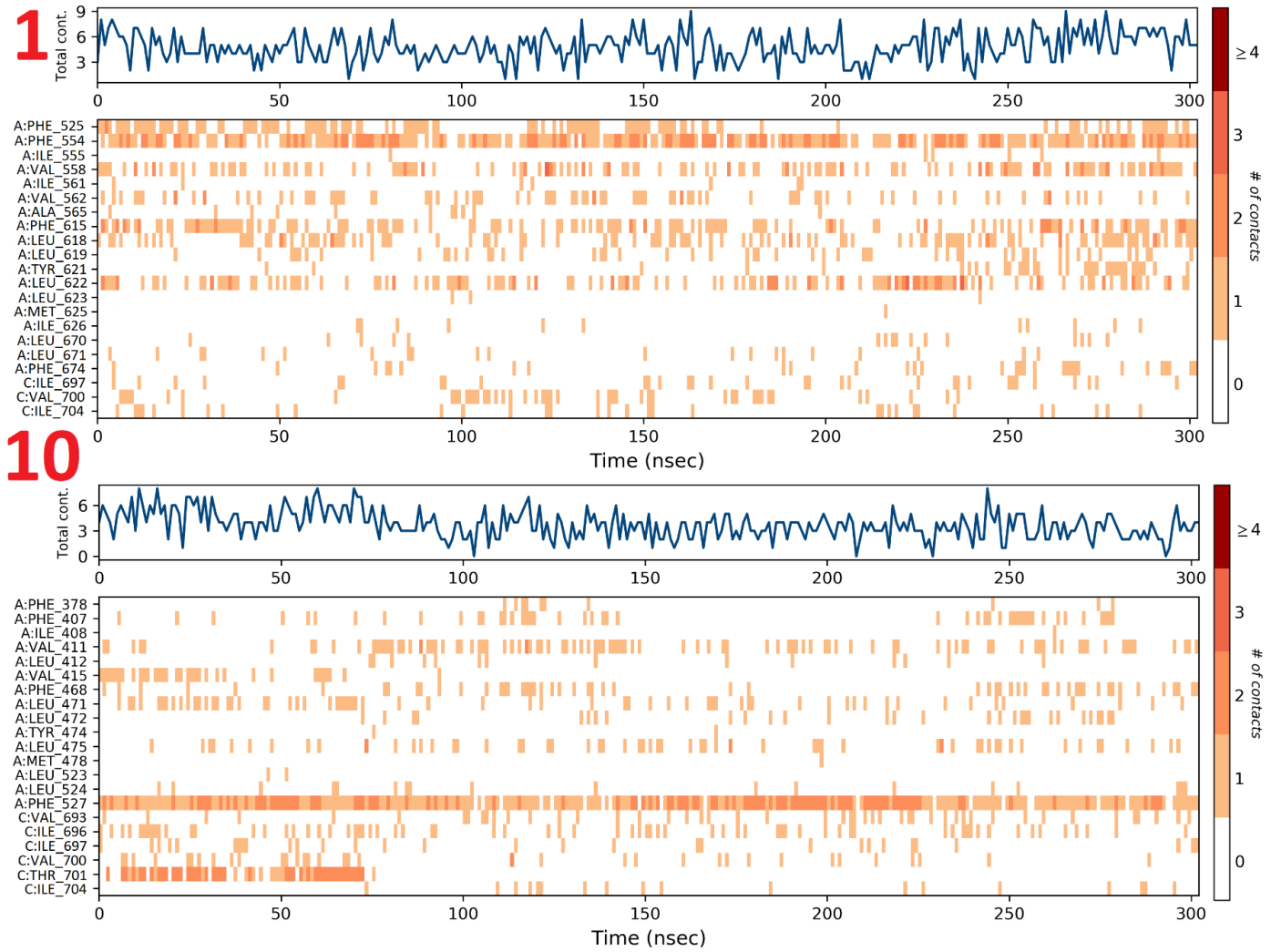


**Figure 35.** The number of protein-ligand interactions versus the simulation time for complexes 1 and 10 (labeled). The upper plot reveals how many contacts or interactions the ligand forms in total with the protein during a certain simulation time and the bottom plot reveals the number of contacts or interactions formed by a specific amino acid during a certain simulation time. The residue numbers correspond to the correct human TRPV4 residue numbers, the letter in front of the number (A-D) indicates the chain where the residues are located at.

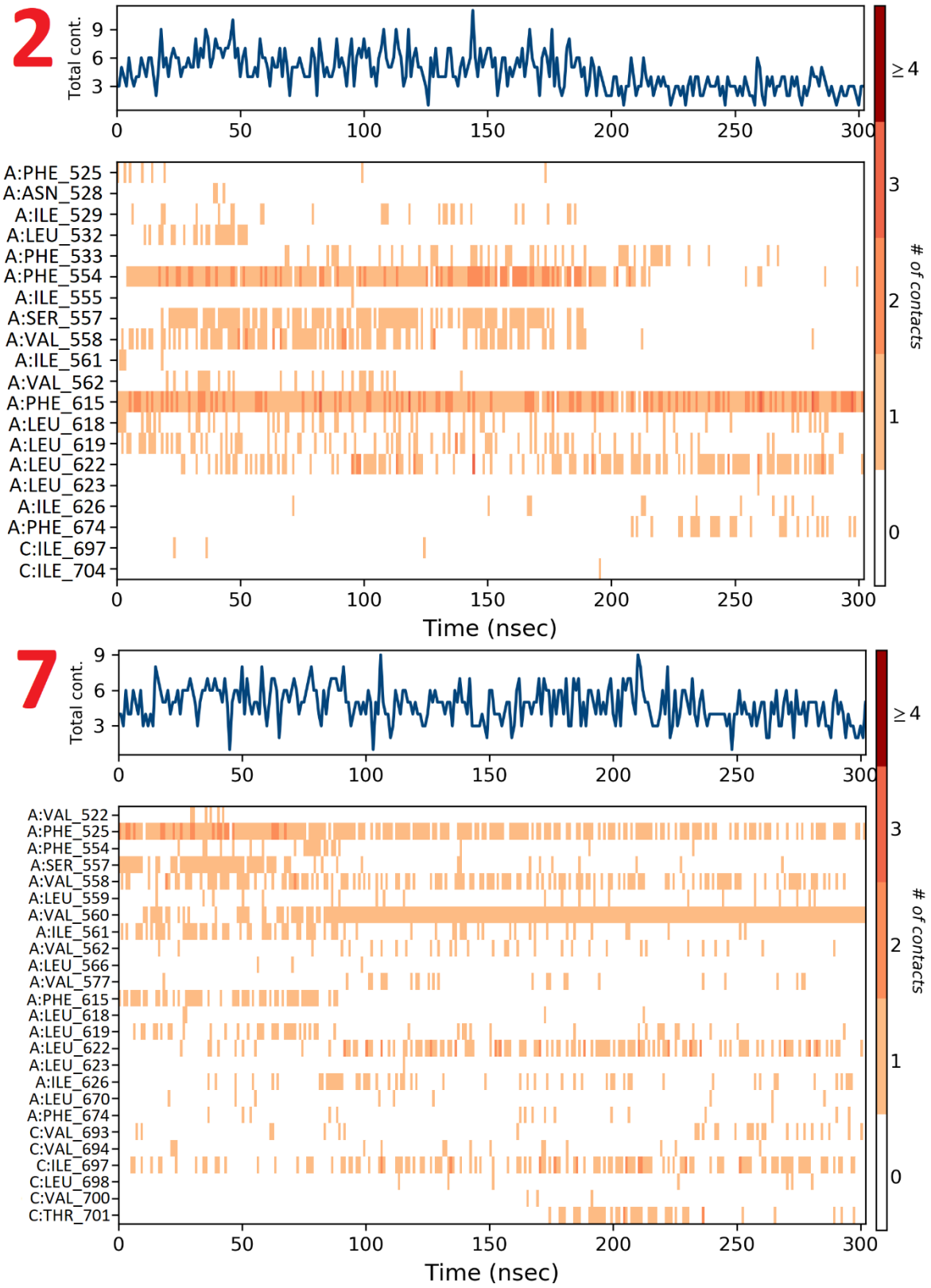


**Figure 36.** The number of protein-ligand interactions versus the simulation time for complexes 2 and 7 (labeled). The upper plot reveals how many contacts or interactions the ligand forms in total with the protein during a certain simulation time and the bottom plot reveals the number of contacts or interactions formed by a specific amino acid during a certain simulation time. The residue numbers correspond to the correct human TRPV4 residue numbers, the letter in front of the number (A-D) indicates the chain where the residues are located at.

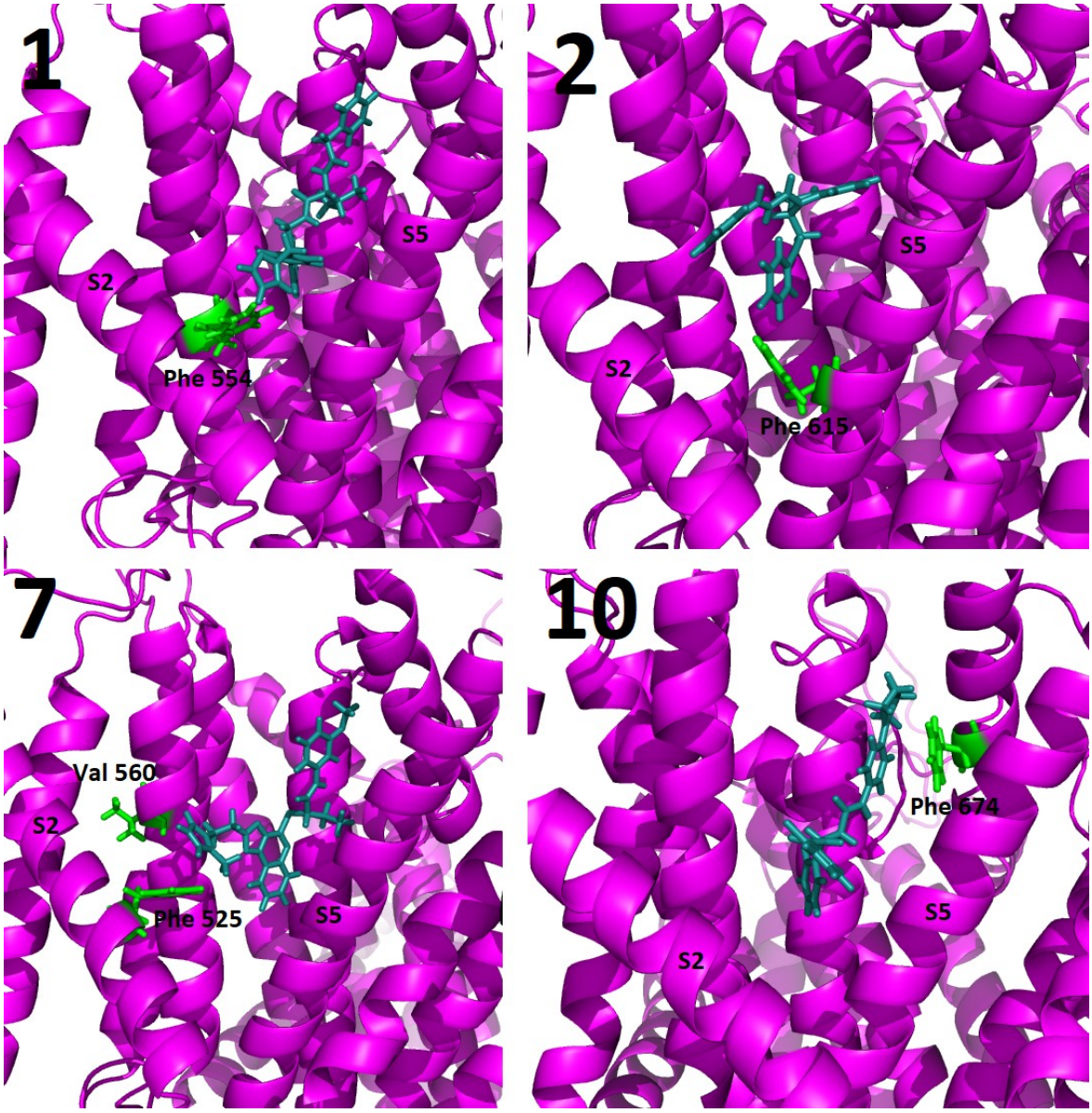


**Figure 37.** The locations of the residues forming the longest lasting interactions with the ligands. The complex numbers are labeled. The protein is seen in magenta and the ligands in teal. The important residues are visualized in green, and they are labeled and shown as sticks. Alpha helices S2 and S5 are also labeled. Val 560 and Phe 615 formed hydrogen bonds with the specific ligands and Phe 525, Phe 554 and Phe 674 formed hydrophobic contacts.

### 5.4.4 Reference simulation without the lipid membrane

When comparing the protein-ligand RMSD plots from the two different complex 1 simulations a clear difference could be observed. The ligand was fluctuating considerably more when the lipid membrane was absent, this was confirmed by the simulation video. The protein-ligand interaction plot indicated short-term interactions between the residues and the ligand, this information supports the results obtained from the RMSD plot and the simulation video. The longest occurring interaction was formed between Ile 697 in C chain and the ligand, this hydrophobic interaction was observed only during 40% of the simulation. As previously mentioned, complex 1 with the lipid membrane formed hydrophobic interactions between Phe 554 and the ligand for over 100% of the simulation time (more than one interaction was formed since the value was over 100%). Overall, when comparing the protein-ligand interaction plots, it is evident that the results differ enormously. Corresponding residues did not form similarly lasting interactions and the residues that form the longest lasting interactions are completely different in the two simulations. The RMSD plot and the protein-ligand interaction plot for the complex 1 without the lipid membrane are visualized in Figure 38.

From the differing results between the two simulations, it is evident that to get accurate results, it is crucial to build up a lipid membrane to the system when performing MD simulations with membrane proteins.

The true difference between the results can only be achieved by performing many parallel simulations of the systems.

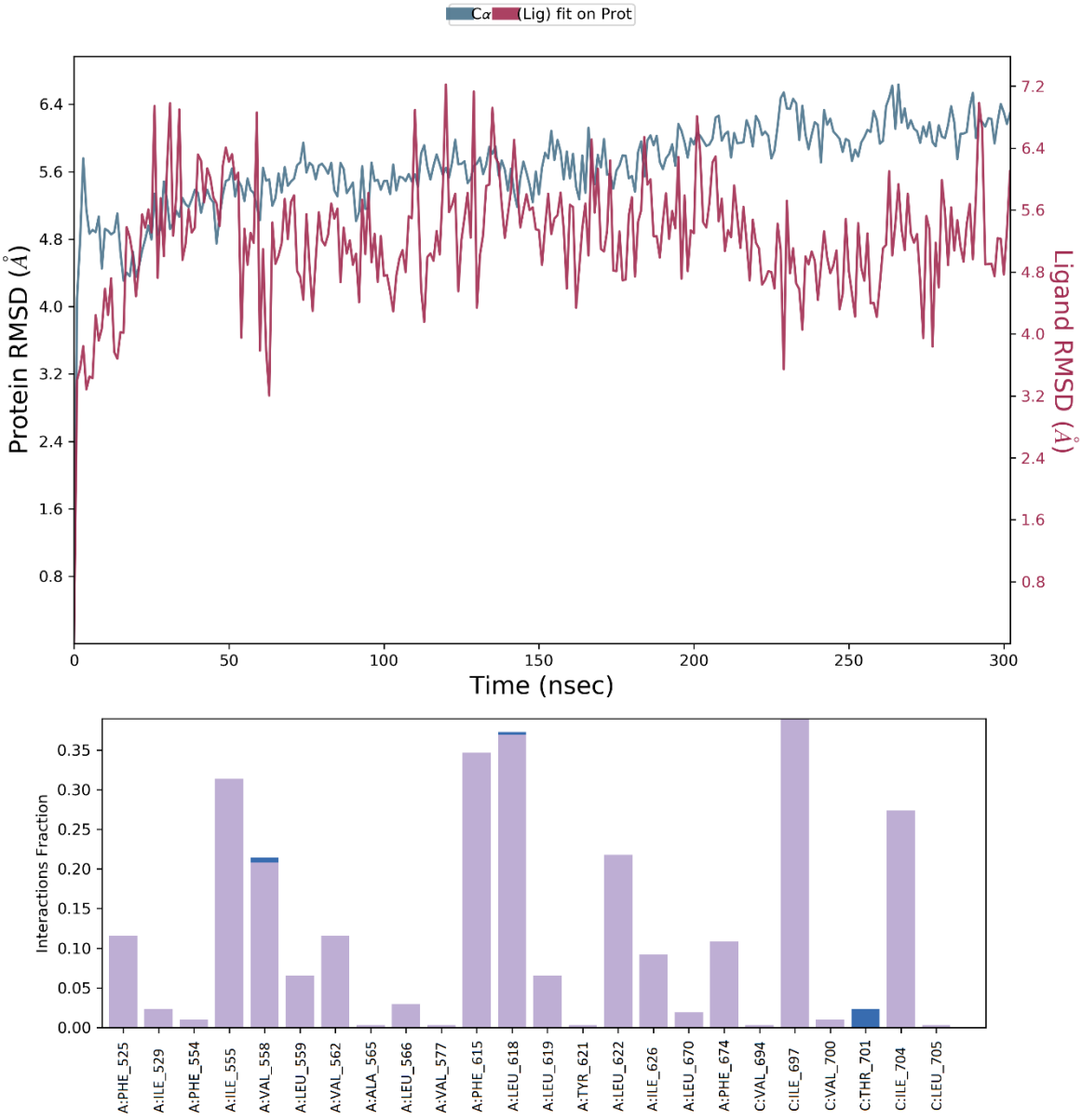


**Figure 38.** The RMSD (above) and protein-ligand interaction (bottom) plot for the reference simulation for complex 1 without the lipid membrane. The RMSD fluctuation for the protein is presented in blue and the value seen in the left Y-axis. The RMSD fluctuation for the ligand is presented in magenta and the value seen in the right Y-axis. Interactions are shown as fractions from the total MD simulation time (300 ns). The residue numbers are presented on the X-axis and the interactions fraction value is presented on the Y-axis. The residue numbers correspond to the correct human TRPV4 residue numbers, the letter in front of the number (A-D) indicates the chain where the residues are located at.

### 5.5 Prediction of ADMET properties and the selection of the best molecules

The results from the ADMET prediction are presented in Table 3. The prediction accuracy for the ProTox II results ranged between 54 and 68%. The toxicity class in ProTox II is divided into six categories and is directly connected to the LD50 value when the molecule is taken orally. Class 1 molecules are categorized fatal and class 6 molecules non-toxic. All the compounds tested in this thesis were either class 5 (may be harmful) or class 4 (harmful). The hepatotoxicity, immunotoxicity, cytotoxicity, carcinogenicity and mutagenicity predictions are presented as active or inactive. If a ligand is predicted active, it means that the ligand is toxic whereas an inactive result predicts a non-toxic ligand. Only a few of the targets could be predicted with high accuracy. Immunotoxicity was predicted with high accuracy for all ligands, two of the ligands were predicted to be inactive and two were active. Mutagenicity was predicted as inactive for one ligand and cytotoxicity was predicted as inactive for two ligands. Hepatotoxicity and carcinogenicity could not be predicted with high accuracy for any ligand.

Water solubility is presented as log S. The molecule is predicted to be highly soluble when the value is over 0, very soluble when the value is between 0 and -2, soluble when the value is between -2 and -4, moderately soluble when the value is between -4 and -6, poorly soluble when the value is between -6 and -10 and insoluble when the value is < -10. The ligands Z1213735368 and Z221487176 were predicted to be moderately soluble with both methods. The ligands Z1053001754 and Z1157726398 were predicted to be moderately soluble with the ESOL method and poorly soluble with the QSPR method.

SwissADME predicted that all molecules would have high GI absorption. A couple of molecules would permeate the BBB and three molecules would function as P-gp substrates. Only one molecule was predicted to act as a CYP1A2 inhibitor, the rest of the CYP enzymes were predicted to be inhibited by all the molecules.

The synthetic accessibility value is a score between 1 and 10, the score 1 means the molecule is very easy to synthesize and the score 10 means it is very difficult to synthesize. SwissADME predicted that all the molecules are relatively easy to synthesize.

The skin permeation value was included because of the possibility that the administration route of the medicine could be transdermal. The less negative the score is, the better the

skin permeation. For example, a molecule with the value around -10 cm/s does not permeate the skin well and a molecule with the value around -5 cm/s permeates the skin well (SwissADME – FAQ, 2021). The molecules in this thesis were predicted to permeate the skin quite well.

The ligands Z1213735368, Z1157726398 and Z221487176 were selected to be sent to our collaborators for further experimental evaluation. Molecules Z1213735368 and Z1157726398 were selected because they formed hydrogen bonds with the protein. Molecule Z221487176 was selected because the hydrophobic connection it formed with the protein was more consistent compared to the hydrophobic connection formed with molecule Z1053001754. The ADMET prediction results did not find any drastic differences between the ligands.

**Table 3**. The ADMET prediction results performed with SwissADME and ProTox II for the ligands in complexes 1, 2, 7 and 10. The values for hepatotoxicity, carcinogenicity, immunotoxicity, mutagenicity and cytotoxicity that could be predicted with minimum 70% accuracy are presented with bright green or red background. Values that were below 70% but higher than 50% in accuracy are presented in light green and light red. If a feature is labeled "inactive" it means that the ligands is not predicted to be for example mutagenic and the opposite if the feature is labeled "active". With the different colors from the toxicity class, GI absorption, BBB permeability, P-gp substrate and CYP inhibition predictions, the purpose was to make it easier to distinguish between differing results.

| | **Z1053001754 (Complex 1)** | **Z1213735368 (Complex 2)** | **Z1157726398 (Complex 7)** | **Z221487176 (Complex 10)** |
|---|---|---|---|---|
| **LD50 (mg/kg)** | 2290 | 2500 | 500 | 1800 |
| **Toxicity Class** | 5 | 5 | 4 | 4 |
| **Hepatotoxicity** | Inactive | Inactive | Inactive | Inactive |
| **Carcinogenicity** | Active | Inactive | Active | Inactive |
| **Immunotoxicity** | Inactive | Active | Active | Inactive |
| **Mutagenicity** | Inactive | Inactive | Active | Inactive |
| **Cytotoxicity** | Inactive | Inactive | Inactive | Inactive |
| **Solubility (log S) ESOL/QSPR** | -5.57 / -6.20 | -4.53 / -4.61 | -5.62 / -6.86 | -5.00 / -5.33 |
| **GI absorption** | High | High | High | High |
| **BBB permeant** | No | Yes | No | Yes |
| **P-gp substrate** | Yes | No | Yes | Yes |
| **CYP1A2 inhibitor** | No | No | No | Yes |
| **CYP2C19 inhibitor** | Yes | Yes | Yes | Yes |
| **CYP2C9 inhibitor** | Yes | Yes | Yes | Yes |
| **CYP2D6 inhibitor** | Yes | Yes | Yes | Yes |
| **CYP3A4 inhibitor** | Yes | Yes | Yes | Yes |
| **Synthetic accessibility** | 3.78 | 3.72 | 4.41 | 3.53 |
| **Skin permeation Log $K_p$ (cm/s)** | -5.77 | -6.58 | -6.17 | -6.18 |

### 5.6 Pharmacophore modeling

#### 5.6.1 Modeling and ADMET prediction

The resulting pharmacophore hypothesis is visualized in Figure 39. Figure 39A presents the aligned antagonists GSK 205 and HC-067047 and how the pharmacophore is built from them. Figure 39B presents only the pharmacophore hypothesis which had four features, the orange rings correspond to aromatic rings and the blue ball with an arrow corresponds to a hydrogen bond donor and its direction.

From the PHASE screening, the screened ligands were ranked according to the PHASE fitness score. The fitness score has a value between -1 and 3 where 3 equals the perfect fit to the pharmacophore. Six ligands with the highest fitness scores were chosen. All chosen molecules had the fitness score over 2. The 2D structures and the pharmacophore fit for the chosen molecules are visualized in Figure 40.

The hit ligands were also submitted for ADME properties prediction with SwissADME and for toxicity prediction with ProTox II. The results are visualized in Table 4. The prediction accuracy for the ProTox II results ranged between 54 and 67%.

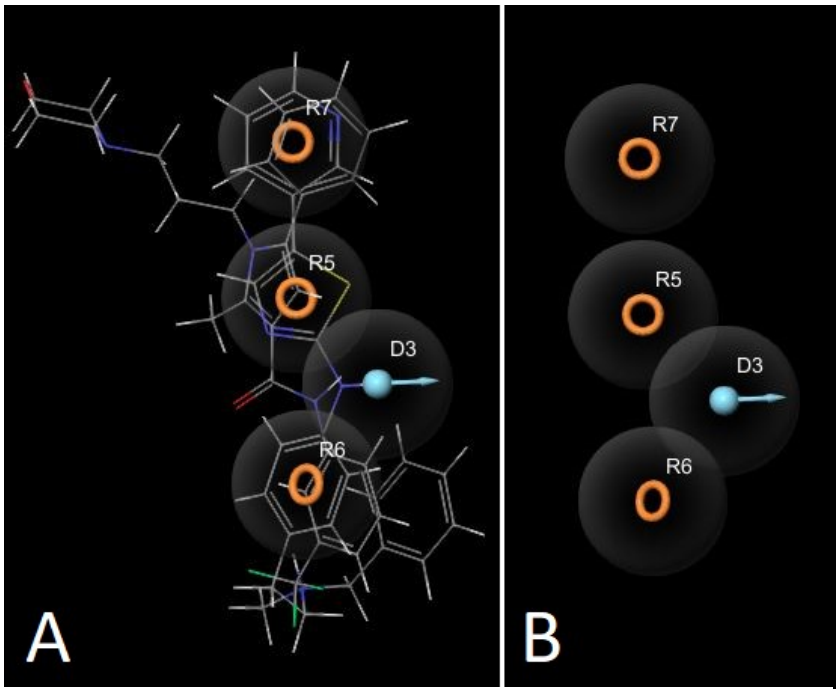


**Figure 39**. A) The pharmacophore model based on the aligned known TRPV4 antagonists GSK205 and HC-067047; B) The pharmacophore model without the ligand structures. The orange rings correspond to aromatic rings and the blue ball with an arrow to a hydrogen bond donor.

The ligands were predicted to belong to the toxicity classes 4 or 5 and the LD50 values ranged between 350 and 5000 mg/kg. The skin permeation value ranged between -5.79 and -6.40 cm/s. This meant that the ligands would probably permeate the skin quite well. The ligands were ranked as moderately

water soluble. The synthetic accessibility scores predicted that all the ligands are easy to synthesize.

Ligands 2, 3, 4, 5 and 6 were predicted to have high GI absorption. Ligand 5 was predicted to function as a P-gp substrate and ligands 5 and 6 were predicted to permeate the BBB. All ligands worked as CYP inhibitors. However, ligands 2 and 3 were not predicted to inhibit CYP2D6.

It could be predicted with high accuracy that ligands 2, 5 and 6 would not be immunotoxic, ligands 3 and 4 would not be mutagenic, and ligand 6 would not be cytotoxic. Hepatotoxicity and carcinogenicity could not be predicted with high accuracy for any of the ligands.

The ligands 5 and 6 were almost identical to the structure and pharmacophore fit. The difference is one chlorine substituent in the structure of ligand 6. The chlorine substituent makes a great impact on the predicted LD50 value when comparing these two ligands.

The ligands Z443243482 (ligand 4) and Z1728742868 (ligand 2) were chosen for further biological evaluation via our collaborators. Z443243482 had the highest LD50 value and good solubility values. Although Z1728742868 had the lowest LD50 value (350 mg/kg), it was not predicted to be hepatotoxic, cytotoxic, carcinogenic, mutagenic or immunotoxic. Immunotoxicity was the only one predicted with high accuracy. Z1728742868 was also the molecule that had the highest PHASE fitness score amongst the molecules that also had high GI absorption.

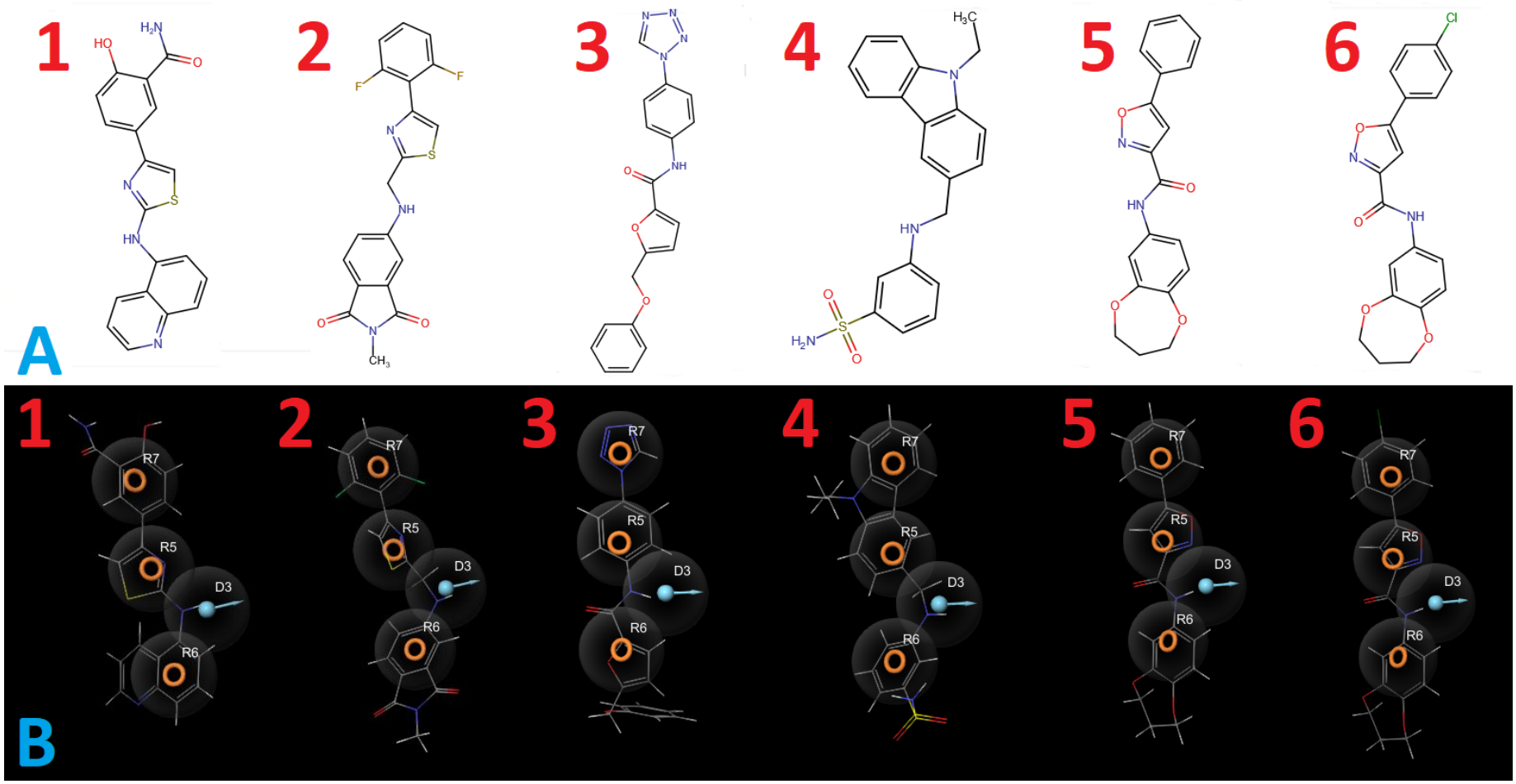


**Figure 40.** A) The 2D structures (1A-6A); B) Pharmacophore fit (1B-6B) of the chosen ligands. The ligand structures were obtained from the EnamineStore website (https://www.enaminestore.com/). The pharmacophore fit figures were prepared with Maestro. The Enamine IDs for the molecules are: 1 = Z916071674, 2 = Z1728742868, 3 = Z26551809, 4 = Z443243482, 5 = Z164852238 and 6 = Z229634356.

**Table 4.** PHASE fitness score, molecular properties and the ADMET prediction results performed with SwissADME and ProTox II for the best scoring ligands. The values for hepatotoxicity, carcinogenicity, immunotoxicity, mutagenicity and cytotoxicity that could be predicted with minimum 70% accuracy are presented with bright green or red background. Values that were below 70% but higher than 50% in accuracy are presented in light green and light red. If a feature is labeled "inactive" it means that the ligands is not predicted to be for example mutagenic and the opposite if the feature is labeled "active". With the different colors from the toxicity class, GI absorption, BBB permeability, P-gp substrate and CYP inhibition predictions, the purpose was to make it easier to distinguish between differing results.

| Enamine ID | Z916071674 | Z1728742868 | Z26551809 | Z443243482 | Z164852238 | Z229634356 |
|---|---|---|---|---|---|---|
| Phase Fitness Score | 2.148908 | 2.055982 | 2.052274 | 2.037829 | 2.028276 | 2.024096 |
| MW (g/mol) | 362.4 | 385.4 | 361.4 | 379.5 | 336.3 | 370.8 |
| CLogP | 4.528 | 3.749 | 3.204 | 3.999 | 3.524 | 4.238 |
| Log S | -5.445 | -5.575 | -5.04 | -5.005 | -5.12 | -5.83 |
| HBD | 3 | 1 | 1 | 2 | 1 | 1 |
| LD50 (mg/kg) | 1000 | 350 | 2000 | 5000 | 4000 | 500 |
| Toxicity Class | 4 | 4 | 4 | 5 | 5 | 4 |
| Hepatotoxicity | Active | Inactive | Active | Inactive | Active | Active |
| Carcinogenicity | Active | Inactive | Active | Active | Active | Inactive |
| Immunotoxicity | Active | Inactive | Inactive | Inactive | Inactive | Inactive |
| Mutagenicity | Inactive | Inactive | Inactive | Inactive | Active | Active |
| Cytotoxicity | Inactive | Inactive | Inactive | Inactive | Inactive | Inactive |
| Solubility (log S) ESOL/QSPR | -4.83/-6.24 | -4.52/-4.89 | -4.09/ -4.62 | -4.70/-5.17 | -4.22/-4.49 | -4.82/-5.14 |
| GI absorption | Low | High | High | High | High | High |
| BBB permeant | No | No | No | No | Yes | Yes |
| P-gp substrate | No | No | No | No | Yes | No |
| CYP1A2 inh. | Yes | Yes | Yes | Yes | Yes | Yes |
| CYP2C19 inh. | Yes | Yes | Yes | Yes | Yes | Yes |
| CYP2C9 inh. | Yes | Yes | Yes | Yes | Yes | Yes |
| CYP2D6 inh. | Yes | No | No | Yes | Yes | Yes |
| CYP3A4 inh. | Yes | Yes | Yes | Yes | Yes | Yes |
| Synthetic accessibility | 2.94 | 2.94 | 3.05 | 2.73 | 3.16 | 3.15 |
| Skin permeation Log $K_p$ (cm/s) | -5.79 | -6.30 | -6.40 | -6.00 | -6.03 | -5.79 |

### 5.6.2 Testing the pharmacophore hypothesis of the known antagonists

If both the known antagonists bind into different antagonist binding sites in the protein, the pharmacophore model based on their superposed structures may be wrong. On the other hand, if the two antagonists bind into the same binding site, the results obtained from the pharmacophore modeling could be valid. However, it is of course not certain if either of the compounds binds specifically to the putative VBS that we used in this study.

The results from the test docking are visualized in Figure 41. Both GSK 205 and HC-067047 were successfully docked into the vanilloid binding pocket. However, the docking scores were poor. The XP score for GSK 205 was -4.557 and for HC-067047 -2.605. The MM-GBSA score for GSK 205 was -40.82 and for HC-067047 -37.61. Both XP score and MM-GBSA score are low for these ligands when compared to the MM-GBSA scores obtained from the docking studies of the best Enamine library compounds. Especially the Glide docking results suggest that these two antagonists may bind into some other binding pocket than the VBS. On the other hand, the accuracy of the model can also affect the docking results.

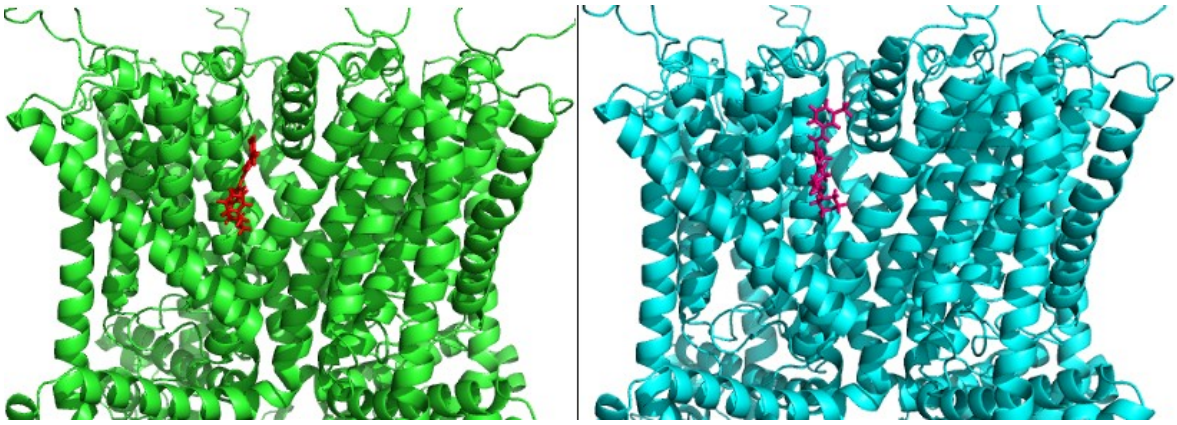

**Figure 41.** The docked antagonists GSK 205 in red and human TRPV4 in green (left panel) and HC-067047 in magenta and human TRPV4 in cyan (right panel). Only the transmembrane parts of the proteins are visualized.

## 6. Discussion

Despite of huge advances in the computer-aided drug design methods, there are still many problems to be solved. For example, the restrictions in estimating the ADMET properties are yet to be conquered, the flexibility of the target protein and the ligand are not always considered and the knowledge in choosing the right scoring function should be enhanced (Batool et al., 2019).

TRPV4 receptors are found in various locations in the body and because they are involved in so many different functions, it has been suggested that the administration route for the possible future inhibitors could be through local injections, for example in the joint (intra-articular) or in the spinal cord (intrathecal). Other non-systemic administration routes could be via inhalation or transdermal administration. Hence, the predicted skin permeation abilities were presented in this thesis for the best scoring ligands. Via local administration it would be easier to reach the specific area and cells meant to be targeted and hopefully decrease the possibility for serious adverse effects (Kanju et al. 2016).

It has been suggested that another way to avoid adverse effects would be to inhibit a part of the signaling pathway leading to TRPV4 activation, instead of directly blocking the ion channel. Grace et al. (2014) discovered that bafetinib, which is a tyrosine kinase inhibitor, blocked the signaling pathway between PAR2 and TRPV4. Bafetinib is believed to inhibit tyrosine kinases to activate TRPV4. Bafetinib did not affect PAR2 activation.

After the experimental phases of this thesis were completed, a study describing an antagonist binding site was discovered (Doñate-Macian et al., 2020). The study applied SiteMap for locating the binding pocket and Desmond for MD simulations. The western clawed frog TRPV4 3D structure was used for the CADD approaches. The study discovered the binding pocket for TRPV4 antagonist HC-067047. The pocket is formed by S2-S3 linker and the alpha-helices S4 and S5 (Fig. 42). The pocket occupied partly the same areas that the agonist binding site for agonist 5',6'-EET occupy (see section 2.4). The study also applied human TRPV4 transfected HeLa cells for the validation of important residues. Mutation studies confirmed the residue Asp 546 to be crucial for inhibition intermediated by HC-067047. Tyr 591 was discovered to be another important residue. The study also

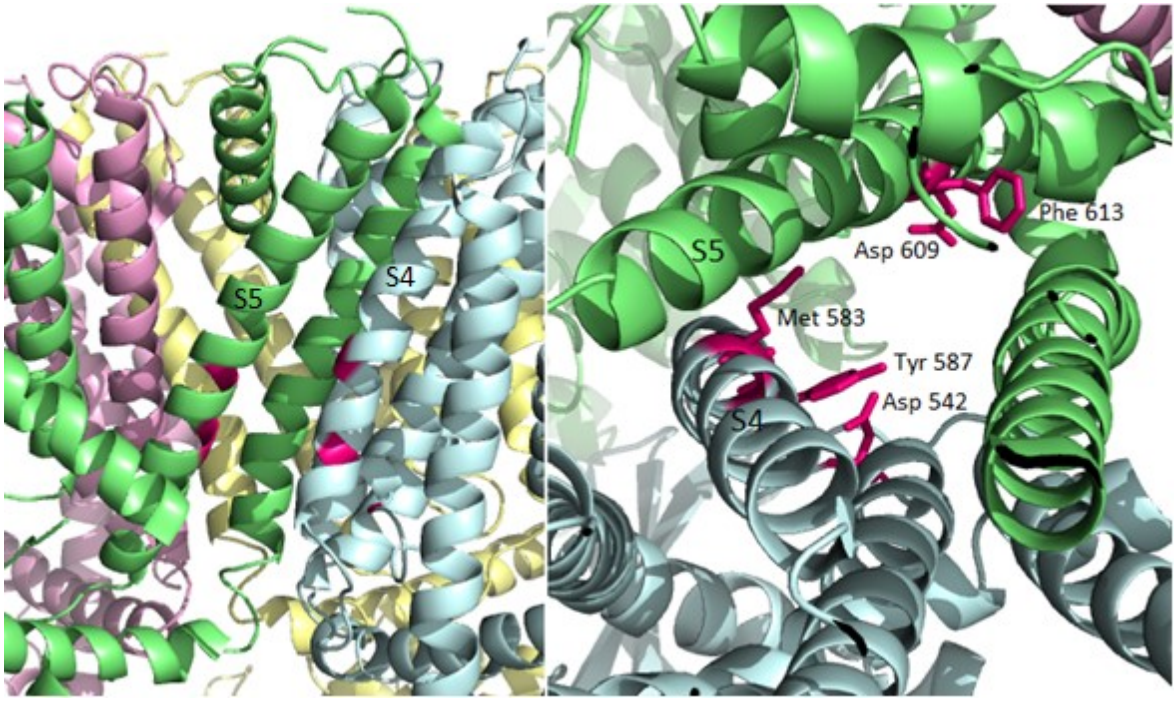


**Figure 42.** The antagonist binding site located by Doñate-Macian et al. (2020). The left panel visualizes partly the VBS (on the left of the labeled S5) and the transmembrane part next to the VBS of the frog TRPV4 (PDB ID 6BBJ), the magenta-colored parts are the residues forming the binding cavity, the chains of the protein are colored with different colors. The figure on the right visualizes the binding site from another angle. The important residues are shown as sticks and labeled. The S5 and the S4 from adjacent chains are also labeled.

aimed to discover novel TRPV4 antagonists against Zika virus. Since the antagonist in the study was the same as one of the antagonists used in this thesis for creating the pharmacophore model, there is a possibility that the ligands from the pharmacophore modeling study would rather bind to this pocket. This would be consistent with the poor docking results. But since the binding pocket for the other antagonist used in pharmacophore modeling, GSK205, is still unknown, the accuracy of the pharmacophore model and the resulting hit ligands needs further experimental evaluation. It is yet to be discovered if this binding pocket presented in the study by Doñate-Macian et al. (2020) also serves in mediating pain and inflammation.

AlphaFold (Jumper et al., 2021) is an artificial intelligence system that has recently predicted the 3D structures of the human proteome. The prediction of one human TRPV4 subunit was found (with UniProt entry code Q9HBA0) in AlphaFold (Fig. 43A). According to AlphaFold's predicted local-distance difference test (pLDDT), a large part of the protein structure was predicted with a confidence score of over 70 (on a scale from 0 to 100). The

AlphaFold model was aligned with the human TRPV4 model that was used in this thesis in PyMOL (Fig. 43B). The RMSD was 2.171 Å, 446 C-alpha atoms from each subunit were aligned and 195 C-alpha atoms were rejected from the alignment. The two models were almost identical at the ARD, but the transmembrane domains, where the binding sites have been located at, were not identical, although similar. Perhaps the fact that the AlphaFold structure only consists of one subunit could affect the overall structure of the model?

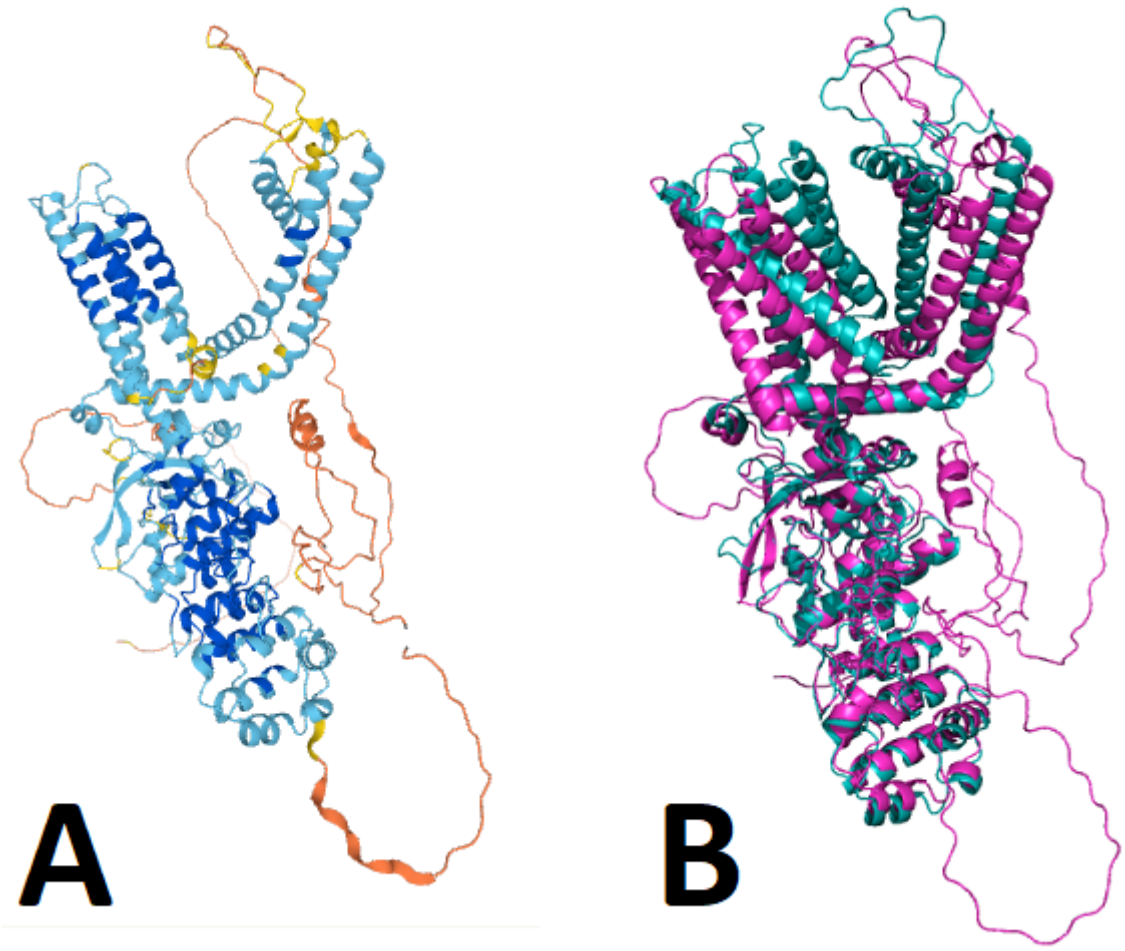


**Figure 43.** A) The AlphaFold human TRPV4 subunit model. The coloring indicates the model confidence; dark blue means that the pLDDT score > 90 (very high confidence), light blue means that the pLDDT score is between 70 and 90, yellow means that the pLDDT score is between 50 and 70 and orange means that the pLDDT score is < 50 (very low confidence). B) The aligned human TRPV4 subunit model structures, the AlphaFold model can be seen in magenta and the model used in this thesis work in teal.

The selected ligands were not predicted to be easily soluble. If necessary, the solubility may be increased with various methods. For example, the pH can be adjusted, a salt form can be added to the molecule or various dispersion methods can be used. The poor solubility of new drug candidates remains a challenge for the drug development community (Savjani, Gajjar, & Savjani, 2012).

Since the binding site used in this thesis is the corresponding site to the vanilloid binding site in TRPV1, it could be of interest to dock and simulate the best scoring ligands with TRPV1 to determine if the ligands could act as dual antagonists (or agonists).

## 7. Conclusions

In this study, new putative TRPV4 antagonists were searched for the treatment of pain and inflammation. The human TRPV4 model was built with the program MODELLER by using the frog TRPV4 structure as the template structure. The putative antagonist binding site was discovered by using the program SiteMap. The discovered binding site corresponded to the VBS found in TRPV1. The Enamine Ion Channel molecule library was docked into the putative binding site with the docking programs GOLD and Glide. A 300-ns MD simulation was performed with Desmond for the 10 most promising protein-ligand complexes. Finally, ADMET predictions were performed with SwissADME and ProTox II, to be able to choose the most promising compounds with good predicted pharmacokinetic properties. A pharmacophore model was created by using two known TRPV4 antagonists with the program PHASE. The Enamine Ion Channel library was then screened with the model and six most interesting ligands were chosen for ADMET prediction. The final selected ligands from the molecular docking study were Z1213735368, Z1157726398 and Z221487176 and from the pharmacophore study Z443243482 and Z1728742868.

The lack of experimentally solved good quality membrane protein structures is a factor that makes computer-aided drug design methods slower and more unreliable. The solution to the challenges lies in the development of even more powerful and accurate tools for computer-aided drug discovery, such as artificial intelligence. The development relies in the future resources and investments in the science.

Because the putative binding site used in this thesis serves as a binding site in other TRPV proteins, it is possible that the location also acts as a binding site in TRPV4. If it would be determined that the site actually serves as a binding site, the next thing to do would be to identify the properties and functions that are mediated through the binding site.

Hence, further experimental testing and knowledge of the TRPV4 antagonist binding site(s) and the critical residues involved are required, to be able to conclude if the ligands discovered in this thesis would act as TRPV4 antagonists against pain and inflammation.

## 8. Summary in Swedish –Svensk sammanfattning

### Upptäckt av nya TRPV4 inhibitorer för behandling av smärta och inflammation med hjälp av datorstödd läkemedelsdesign

Målsättningen med pro gradu-avhandlingen var att upptäcka några nya potentiella antagonister till proteinet TRPV4 och sedan skicka dessa molekyler till vårt samarbetslaboratorium för fortsatta undersökningar. Arbetet utfördes genom att använda olika metoder av datorstödd läkemedelsdesign.

TRP-proteiner fungerar som jonkanaler (membranproteiner) i kroppen. Trettiotre olika TRP-proteiner har upptäckts hos däggdjur och dessa indelas vidare i åtta undergrupper (Nikolaev et al., 2019). Denna avhandling koncentrerade sig på det specifika TRP-proteinet TRPV4. TRPV4 har en viktig roll vid smärtsignalering och då smärtkänslan uppstår. Proteinet fungerar som mekanoreceptor i kroppen och påträffas i de inåtledande nerverna (sensoriska nerverna) i spinalganglierna och i trigeminusganglierna (Kanju et al., 2016). Den exakta mekanismen hur TRPV4 fungerar och vilka kemiska ämnen som medverkar i aktiveringen och inaktiveringen är ännu till stor del okända. För tillfället finns inga läkemedel på marknaden som skulle fungera som TRPV4-antagonister eller -agonister. Antagonisten GSK2798745 har man lyckats få till den kliniska prövningsfasen, en studie är för tillfället i fas II kliniska studier (ClinicalTrials.gov., 2021a) och en annan studie är i fas I kliniska studier (ClinicalTrials.gov., 2021b).

För tillfället är den västafrikanska klogrodans (*Xenopus tropicalis*) TRPV4-struktur den enda experimentellt lösta TRPV4-strukturen som finns tillgänglig för allmänheten (Deng et al., 2018). Denna struktur användes som mall i avhandlingen, då den teoretiska modellen av människans TRPV4 byggdes upp. Metoden kallas för jämförande modellering och kan tillämpas då aminosyrasekvensidentiteten mellan två proteiner är minst 40 %. Grodans och människans TRPV4-sekvensidentitet är 78 % (Deng et al., 2018). Multipel linjering utfördes med några olika arters TRPV4-aminosyrasekvenser med programmet Clustal Omega (Sievers et al., 2011). Körningen gav de motsvarande aminosyrorna för proteinerna och denna information användes då jämförande modellering genomfördes med programmet MODELLER (Šali & Blundell, 1993). Modelleringen resulterade i tio modeller och den bästa modellen valdes genom att undersöka DOPE-poängen, Ramachandran-graferna och

genom visuell granskning av proteinerna med programmet PyMOL (DeLano, 2002). Med programmet Desmond (Bowers et al., 2006) utfördes molekyldynamisk simulering för att undersöka stabiliteten av den bästa modellen. Simuleringstiden var 300 ns och lipidmembranen tillsattes till proteinet för att åstadkomma ett så realistiskt resultat i simuleringen som möjligt. Ramachandran-grafen visade för den bästa modellen att 93,4 % av aminosyrorna hittades i de mest gynnsamma regionerna, 6,3 % hittades i de ytterligare tillåtna regionerna, 0,2 % hittades i de generellt tillåtna regionerna och 0,2 % hittades i de förbjudna regionerna. Aminosyrorna i de generellt tillåtna och förbjudna regionerna var lokaliserade i de flexibla loopstrukturerna och ansågs därför inte skadliga för proteinstrukturen. DOPE-poängen för den bästa modellen var -319982,78125 och resultaten av simuleringen visade att proteinet var stabilt.

Information om antagonistbindningsställen fanns inte tillgänglig i litteraturen då dockningsstudien utfördes. Därför undersöktes om positionerna för antagonistbindningsfickor hade hittats hos andra TRP-proteiner. Proteiner inom samma huvudgrupp är släkt med varandra och är därför ofta liknande till struktur och aminosyrasekvens. Två experimentellt lösta TRPV-strukturer med bundna antagonister upptäcktes, TRPV1 med antagonisten capsazepin (Gao et al. 2016) och TRPV5 med antagonisten ekonazol (Hughes et al., 2018). Bindningsställena hos dessa två proteiner fanns på motsvarande positioner. TRPV4-modellen kördes med SiteMap (Halgren, 2009) programmet som lokaliserade och rangordnade möjliga bindningsfickor i proteinet. Fickan som gav bästa poängen motsvarade fickan i TRPV1 och TRPV5. Denna förmodade bindningsficka valdes för dockningsstudien.

Nästa steg var att docka ett molekylbibliotek i bindningsfickan. Enamine Ion Channel Library (https://enamine.net/) valdes som molekylbibliotek. Molekylbiblioteket bestod av cirka 38 000 molekyler ämnade för att dockas i jonkanaler. Programmen Glide (Friesner et al., 2006) och GOLD (Verdonk et al., 2003) användes för dockning och LigPrep användes för att förbereda molekylerna för dockning. Glide-dockning utfördes tre gånger med metoderna vHTS, SP och XP. Efter varje dockningsomgång valdes de bästa 10 % av resultaten för nästa dockningssteg. XP-dockningen resulterade i cirka 170 molekyler. MM-GBSA-beräkningar utfördes för 28 molekyler med de bästa XP-poängen. MM-GBSA-poängen uppskattar affiniteten mellan proteinet och liganden (Tingjun et al. 2011).

GOLD-dockning utfördes två gånger, först med metoden GA 10 % och sedan med GA 100 %. Positionerna i bindningsfickan för de resulterande 500 molekylerna från GA 10 %-körningen granskades visuellt och cirka 180 molekyler valdes för dockning med GA 100 %-metoden. MM-GBSA-beräkningar utfördes för de resulterande 50 molekylerna.

Genom att jämföra MM-GBSA-resultaten, valdes fyra molekyler från Glide-gruppen och sex molekyler från GOLD-gruppen för molekylärdynamisk simulering. Simuleringarna kördes med programmet Desmond, simuleringstiden var 300 ns och lipidmembranen tillsattes till varje protein-ligandkomplex.

Av simuleringsresultaten granskades RMSD, RMSF och protein–ligand-interaktionsdiagrammen. Videon av simuleringarna granskades visuellt. Fyra ligander valdes som mest intressanta. Två av dessa bildade långvariga vätebindningar med proteinet och de två andra bildade långvariga hydrofobiska interaktioner med proteinet. Prognos av ADMET egenskaper (absorption, distribution, metabolism, exkretion och toxicitet) för de fyra liganderna utfördes med online-verktygen SwissADME (Daina et al., 2017) och ProTox II (Banerjee et al., 2018). ADMET-prognosen visade inga drastiska skillnader mellan liganderna. Liganderna Z1213735368, Z1157726398 och Z221487176 valdes för fortsatta studier hos samarbetslaboratoriet.

En kontrollsimulering utfördes med ett protein-ligandkomplex utan lipidmembran. Målet var att få en inblick i betydelsen av lipidmembranen för simuleringen. Resultaten visade att det är kritiskt att tillsätta lipidmembranen till membranproteiner inför molekyldynamisk simulering. Liganden skiftade kraftigt strukturen och bindningsläget och bildade inga långvariga interaktioner med proteinet då lipidmembranen saknades.

Farmakoformodellering användes också för att hitta nya möjliga antagonister. För körningen användes programmet PHASE (Dixon et al., 2006) och farmakoforhypotesen byggdes upp av två kända TRPV4-antagonister, GSK205 (Kanju et al. 2016) och HC-067047 (Zhang et al. 2015). Hypotesen, som bestod av fyra aromatiska ringar och en vätebindningsdonator, screenades med samma molekylbibliotek som användes i dockningsstudien. Sex molekyler med de bästa PHASE-poängen valdes för prognos av ADMET-egenskaper. Efter en evaluering av alla resultaten valdes liganderna Z443243482 och Z1728742868 för fortsatta studier hos samarbetslaboratoriet.

Kontrolldockning av GSK205 och HC-0607047 i den förmodade bindningsfickan utfördes, för att få en inblick över bindningsförmågan av dessa antagonister, i denna ficka. Dockningen lyckades för båda molekylerna, men docknings- och MM-GBSA-poängen tydde inte på starka interaktioner mellan proteinet och antagonisterna. Detta kunde betyda att bindningsfickan som användes i denna avhandling möjligtvis inte binder dessa två antagonister.

Studien av Doñate-Macian et al. (2020) hittades efter att den experimentella delen av avhandlingen var slutförd. Denna studie beskrev ett bindningsställe för TRPV4, som binder antagonisten HC-0607047. Studien efterforskade nya ligander mot zikavirus. Bindningsstället i studien motsvarade inte bindningsstället som användes i denna avhandling. Detta betyder att de molekyler som uppfunnits med hjälp av farmakoforstudien möjligtvis kunde binda i bindningsstället som studien av Doñate-Macian et al. (2020) beskrivit. Dock vet man inte om GSK205 binds i denna ficka, eller om fickan deltar i förmedlandet av smärta och inflammation.

För att undvika allvarliga biverkningar som kan uppkomma då TRPV4 blockeras i alltför stor utsträckning i kroppen, har det föreslagits att de möjliga framtida läkemedlen skulle påverka TRPV4 receptorer endast lokalt. Olika läkemedelstillförselvägar kunde vara till exempel administration direkt i leden genom intraartikulär injektion, via lungorna genom inhalation eller via huden genom transdermal administration (Kanju et al. 2016).

Mera experimentellt erhållen kunskap och kunskap om TRPV4-antagonistbindningsställen och de viktiga aminosyrorna som medverkar i protein-ligandbindningen krävs, för att kunna bestämma om de ligander som presenteras i avhandlingen fungerar som TRPV4-antagonister.

## 9. References


Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1–2, 19–25.

Ali, J., Camilleri, P., Brown, M., B., Hutt, A., J., & Kirton, S., B. (2012). *In silico* prediction of aqueous solubility using simple QSPR models: The importance of phenol and phenol-like moieties. *Journal of Chemical Information and Modeling*, 52 (11), 2950-2957. DOI: 10.1021/ci300447c

Arinaminpathy, Y., Khurana, E., Engelman, D. M., & Gerstein, M. B. (2009). Computational analysis of membrane proteins: the largest class of drug targets. *Drug Discovery Today*, 14, 1130-5. DOI: 10.1016/j.drudis.2009.08.006

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., & Ben-Tal, N. (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic acids research*, 44(W1), W344–W350. DOI: https://doi.org/10.1093/nar/gkw408

Banerjee, P., Eckert, A, O., Schrey, A. K., & Preissner, R. (2018). ProTox-II: a webserver for the prediction of toxicity of chemicals. *Nucleic Acids Research*, 46, W1, W257–W263. DOI: https://doi.org/10.1093/nar/gky318

Batool, M., Ahmad, B., & Choi, S. (2019). A structure-based drug discovery paradigm. *International journal of molecular sciences,* 20(11), 2783. DOI: https://doi.org/10.3390/ijms20112783

Baylie, R. L., & Brayden J. E. (2011). TRPV channels and vascular function. *Acta Physiologica (Oxford, England)*, 203, 99-116. DOI: 10.1111/j.1748-1716.2010.02217.x

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., … Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research*, 28, 235-242.

Berna-Erro, A., Izquierdo-Serra, M., Sepúlveda, R. V., Rubio-Moscardo, F., Doñate-Macián, P., … Valverde, M. A. (2017) Structural determinants of 5’,6’-epoxyeicosatrienoic acid binding to and activation of TRPV4 channel. *Scientific Reports*, 7, 10522.

Blass, B. E. (2015). *Basic principles of drug discovery and development* (pp. 218-220). Philadelphia, PA: Academic press.

Bowers, K. J., Chow, E., Xu, H., Dror, R. O., Eastwood, M. P., Gregersen, B. A., … Shaw, D. E. (2006, November 11-17). *Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters.* Proceedings of the ACM/IEEE Conference on Supercomputing (SC06), Tampa, Florida.

Cambridge Crystallographic Data Centre. (2019). *GOLD User Guide*. Retrieved 27.7.2021 from: https://www.ccdc.cam.ac.uk/support-and-resources/ccdcresources/GOLD_User_Guide.pdf

Cao, S., Anishkin, A., Zinkevich, N. S., Nishijima, Y., Korishettar, A., Wang, Z., Fang, J., Wilcox, D.A., & Zhang, D. X. (2018). Transient receptor potential vanilloid 4 (TRPV4) activation by arachidonic acid requires protein kinase A-mediated phosphorylation. *The Journal of Biological Chemistry*, 6, 5307-5322. DOI: 10.1074/jbc.M117.811075

Case, D. A., Ben-Shalom, I. Y., Brozell S. R., Cerutti, D. S., Cheatham, T. E. III, Cruzeiro, V. W., … Kollman, P. A. (2018). AMBER 2018 [Computer software], University of California, San Francisco.

ClinicalTrials.gov. (2021a). *A first time in human study to evaluate the safety, tolerability, pharmacokinetics, and pharmacodynamics of GSK2798745 in healthy subjects and stable heart failure patients*. Retrieved 24.10.2021 from: https://clinicaltrials.gov/ct2/show/NCT02119260

ClinicalTrials.gov. (2021b). Study 212669: *A phase I study to evaluate the safety, tolerability, pharmacokinetics and pharmacodynamics of GSK2798745 in participants with diabetic macular edema*. Retrieved 24.10.2021 from: https://clinicaltrials.gov/ct2/show/NCT04292912

ClinicalTrials.gov. (2021c). *A study to assess the effectiveness and side effects of GSK2798745 in participants with chronic cough.* Retrieved 24.10.2021 from: https://clinicaltrials.gov/ct2/show/NCT03372603

D E Shaw Research. (2021). *Desmond*. Retrieved 13.11.2021 from: https://www.deshawresearch.com/resources_desmond.html

Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, 7, W257-W263. DOI: https://doi.org/10.1038/srep42717

Delaney, J. S. (2004). ESOL: Estimating aqueous solubility directly from molecular structure. *Journal of Chemical Information and Computer Sciences*, 44 (3), 1000-1005. DOI: 10.1021/ci034243x

DeLano, W. L. (2002). PyMOL. DeLano Scientific, San Carlos, CA, 700.

Deng, Z., Paknejad, N., Maksaev, G., Sala-Rabanal, M., Nichols, C. G., Hite, R. K., & Yuan, P. (2018). Cryo-EM and X-ray structures of TRPV4 reveal insight into ion permeation and gating mechanisms. *Nature Structural & Molecular Biology*, 25(3), 252-260.

Desmond 4.2 User Manual. (2015). Retrieved 27.9.2021 from: http://gohom.win/ManualHom/Schrodinger/Schrodinger_2015-2_docs/desmond/desmond_user_manual.pdf

Dias, F. C., Alves, V. S., Matias, D. O., Figueiredo, C. P., Miranda, A. L. P., Passos, G.F., & Costa R. (2019). The selective TRPV4 channel antagonist HC-067047 attenuates mechanical allodynia in diabetic mice. *European Journal of Pharmacology*, 856:172408. DOI: 10.1016/j.ejphar.2019.172408

Dixon, S. L., Smondyrev, A. M., Knoll, E. H., Rao, S. N., Shaw, D. E., & Friesner, R. A. (2006). PHASE: A new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening. 1. Methodology and preliminary results. *Journal of Computer-Aided Molecular Design*, 20, 647-671.

Doñate-Macian, P., Duarte, Y., Rubio-Moscardo, F., Pérez-Vilaró, G., Canan, J., Díez, J., González-Nilo, F., & Valverde, M. A. (2020). Structural determinants of TRPV4 inhibition and identification of new antagonists with antiviral activity. *British Journal of Pharmacology*. DOI: 10.1111/bph.15267

Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., Sanschagrin, P. C., & Mainz, D. T. (2006). Extra precision Glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of Medicinal Chemistry*, 49, 6177–6196.

Gao, Y., Cao, E., Julius, D., & Cheng, Y. (2016). TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. *Nature*, 534, 347-51. DOI: 10.1038/nature17964

Garcia-Elias, A., Berna-Erro, A., Rubio-Moscardo, F., Pardo-Pastor, C., Mrkonjić, S., Sepúlveda, R. V., Vicente, R., González-Nilo, F., & Valverde, M. A. (2015). Interaction between the linker, Pre-S1, and TRP domains determines folding, assembly, and trafficking of TRPV channels. *Structure*, 23(8), 1404-1413.

García-Sanz, N., Valente, P., Gomis, A., Fernández-Carvajal, A., Fernández-Ballester, G., Viana, F., Belmonte, C., & Ferrer-Montiel A. (2007). A role of the transient receptor potential domain of vanilloid receptor I in channel gating. *The Journal of Neuroscience*, 27(43), 11641–11650.

Goyal, N., Skrdla, P., Schroyer, R., Kumar, S., Fernando, D., Oughton, A., … Cheriyan, J. (2019). Clinical Pharmacokinetics, Safety, and Tolerability of a Novel, First-in-Class TRPV4 Ion Channel Inhibitor, GSK2798745, in Healthy and Heart Failure Subjects. *American Journal of Cardiovascular Drugs*, 19 (3): 335-342. DOI: 10.1007/s40256-018-00320-6

Grace, M. S., Lieu, T., Darby, B., Abogadie, F. C., Veldhuis, N., Bunnett, N. W., & McIntyre, P. (2014). The tyrosine kinase inhibitor bafetinib inhibits PAR2-induced activation of TRPV4 channels in vitro and pain in vivo. *British Journal of Pharmacology*, 171 (16), 3881-94. DOI: 10.1111/bph.12750

Grace, M. S., Bonvini, S. J., Belvisi, M. G., & McIntyre, P. (2017). Modulation of the TRPV4 ion channel as a therapeutic target for disease. *Pharmacology & Therapeutics*, 177, 9-22. DOI: 10.1016/j.pharmthera.2017.02.019

Guedes I. A., de Magalhães, C. S., & Dardenne, L. E. (2014). Receptor-ligand molecular docking. *Biophysical Reviews*, 6 (1): 75-87. DOI: 10.1007/s12551-013-0130-2

Halgren T. (2007). New method for fast and accurate binding-site identification and analysis. *Chemical biology & drug design*, 69 (2): 146-8. DOI: 10.1111/j.1747-0285.2007.00483.x

Halgren, T. A. (2009). Identifying and characterizing binding sites and assessing druggability, *Journal of Chemical Information and Modeling*, 49 (2), 377-389. DOI: 10.1021/ci800324m

Henrich, S., Salo-Ahen, O. M., Huang, B., Rippmann, F., Cruciani, G., & Wade, R. C. (2010). Computational approaches to identifying and characterizing protein binding sites for ligand design. *Journal of Molecular Recognition*, 23 (2), 209-219.

Hughes, T.E., Lodowski, D.T., Huynh, K.W., Yazici, A., Del Rosario, J., … Moiseenkova-Bell, V.Y. (2018). Structural basis of TRPV5 channel inhibition by econazole revealed by cryo-EM. *Nature Structural & Molecular Biology*, 25, 53-60. DOI: 10.1038/s41594-017-0009-1

Jacobson, M. P., Pincus, D. L., Rapp, C. S., Day, T. J., Honig, B., Shaw, D. E., & Friesner, R. A. (2004). A hierarchical approach to all-atom protein loop prediction. *Proteins: Structure, Function and Bioinformatics*, 55, 351-367.

Jones, G., Willett, P., Glen, R. C., Leach, A. R., & Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology*, 267, 727-748. DOI: 10.1006/jmbi.1996.0897

Jorgensen, W. L., Chandrasekhar, J., & Madura, J. D. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79, 926. DOI: 10.1063/1.445869

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596, 583–589. DOI: 10.1038/s41586-021-03819-2

Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., & Jorgensen W. L. (2001). Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *The Journal of Physical Chemistry B*, 105 (28), 6474-6487. DOI: 10.1021/jp003919d

Kaneko, Y., & Szallasi, A. (2014). Transient receptor potential (TRP) channels: a clinical perspective. *British journal of pharmacology*, 171 (10), 2474–2507. DOI: 10.1111/bph.12414

Kanju, P., Chen, Y., Lee, W., Yeo, M., Lee, S. H., Romac, J., ... Liedtke, W. B. (2016). Small molecule dual-inhibitors of TRPV4 and TRPA1 for attenuation of inflammation and pain. *Scientific Reports*, 6, article number 26894.

Li, D., Kao, T.-H., & Chang, S.-W. (2020) The structural changes of the mutated ankyrin repeat domain of the human TRPV4 channel alter its ATP binding ability. *Journal of the Mechanical Behavior of Biomedical Materials*, 101, 103407. DOI: /10.1016/j.jmbbm.2019.103407

LigPrep. Schrödinger Release 2020-3: LigPrep, Schrödinger, LLC, New York, NY, 2020.

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Freeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, 23, 3-25.

Ludbrook, V., Hanrott, K. E., Marks-Konczalik, J., Kreindler, J. L., Bird, N. P., ... Smith, J. (2019). S27 A placebo-controlled, double-blind, randomised, crossover study to assess the efficacy, safety and tolerability of TRPV4 inhibitor GSK2798745 in participants with chronic cough. *Thorax*, 74, A18.

Lupyan, D., Mezei, M., Logothetis, D. E., & Osman, R. A. (2010). Molecular dynamics investigation of lipid bilayer perturbation by PIP2. *Biophysical Journal*, 98(2):240-7. DOI: 10.1016/j.bpj.2009.09.063

Moore, C., Cevikbas, F., Pasolli, H. A., Chen, Y., Kong, W., … Liedtke, W.B. (2013). UVB radiation generates sunburn pain and affects skin by activating epidermal TRPV4 ion channels and triggering endothelin-1 signaling. *Proceedings of the National Academy of Sciences of the United States of America*, 110(34):E3225-34. DOI: 10.1073/pnas.1312933110

Nikolaev, Y. A., Cox, C. D., Ridone, P., Rohde, P. R., Cordero-Morales, J. F., … Martinac, B. (2019). Mammalian TRP ion channels are insensitive to membrane stretch. *Journal of Cell Science*, 132 (23): jcs238360.

Phase [Computer software]. (2021). Retrieved 25.10.2021 from: https://www.schrodinger.com/products/phase

Phelps, C. B., Wang, R. R., Choo, S. S., & Gaudet, R. (2010). Differential regulation of TRPV1, TRPV3, and TRPV4 sensitivity through a conserved binding site on the ankyrin repeat domain. *The Journal of Biological Chemistry*, 285 (1): 731-40. DOI: 10.1074/jbc.M109.052548

Poole, D. P., Amadesi, S., Veldhuis, N. A., Abogadie, F. C., Lieu, T., Darby, W., … Bunnett, N. W. (2013). Protease-activated receptor 2 (PAR2) protein and transient receptor potential vanilloid 4 (TRPV4) protein coupling is required for sustained inflammatory signaling. *The Journal of Biological Chemistry*, 288(8), 5790-802.

Prime 4.0 User Manual. (2015). Retrieved 24.10.2021 from: http://gohom.win/ManualHom/Schrodinger/Schrodinger_2015-2_docs/prime/prime_user_manual.pdf

Roos, K., Wu, C., Damm, W., Reboul, M., Stevenson, J. M., Lu, C., … Harder, E. D. (2019). OPLS3e: Extending force field coverage for drug-like small molecules. *Journal of Chemical Theory and Computation*, 15 (3), 1863-1874. DOI: 10.1021/acs.jctc.8b01026

Šali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, 234 ,779–815.

Salo-Ahen, O. M. H., Alanko, I., Bhadane, R., Bonvin, A. M. J. J., Honorato, R. V., Hossain, S., … Vanmeert, M. (2021). Molecular dynamics simulations in drug discovery and pharmaceutical development. *Processes*, 9 (1): 71. https://doi.org/10.3390/pr9010071

Samanta, A., Hughes, T., & Moiseenkova-Bell, V. Y. (2018). Transient receptor potential (TRP) channels. *Sub-cellular Biochemistry*, 87, 141–165. https://doi.org/10.1007/978-981-10-7757-9_6

Savjani, K. T., Gajjar, A. K., & Savjani, J. K. (2012). Drug solubility: importance and enhancement techniques. *ISRN pharmaceutics*, 195727. DOI: /10.5402/2012/195727

Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R., & Sherman, W. (2013). Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *Journal of Computer-Aided Molecular Design*, 27(3), 221-234. DOI: https://doi.org/10.1007/s10822-013-9644-8

Shen, M. & Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. *Protein Science*, 15, 2507–2524.

Sievers, F., Wilm, A., Dineen, D. G., Gibson, T. J., Karplus, K., … Higgins, D.G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* 7, 539. DOI:10.1038/msb.2011.75

SwissADME – FAQ. Retrieved 15.8.2021 from: http://www.swissadme.ch/faq.php

Teng, J., Loukin, S. H., Anishkin, A., & Kung, C. (2015). L596-W733 bond between the start of the S4-S5 linker and the TRP box stabilizes the closed state of TRPV4 channel. *Proceedings of the National Academy of Sciences of the United States of America*, 112(11), 3386-3391.

Tingjun H., Junmei W., Youyong L., & Wei W. (2011). Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *Journal of Chemical Information and Modeling*, 51 (1), 69-82. DOI: 10.1021/ci100275a

Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W., & Taylor, R. D. (2003). Improved protein-ligand docking using GOLD. *Proteins*, 52 (4), 609-23. DOI: 10.1002/prot.10465

Vincent, F., Acevedo, A., Nguyen, M. T., Dourado, M., DeFalco, J., … Duncton, M. A. (2009). Identification and characterization of novel TRPV4 modulators. Biochemical and Biophysical Research Communications, 389, 490-494. DOI: 10.1016/j.bbrc.2009.09.007

Wade, R. C., & Goodford, P. J. (1989). The role of hydrogen-bonds in drug binding. *Progress in Clinical and Biological Research*, 289, 433-44.

Wang, H., & Woolf, C. J. (2005). Pain TRPs. *Neuron*, 46, 9-12.

White, J. P., Cibelli, M., Urban, L., Nilius, B., McGeown, J. G., & Nagy, I. (2016). TRPV4: Molecular conductor of a diverse orchestra. *Physiological Reviews*, 96, 911-973. DOI: 10.1152/physrev.00016.2015

Yu, W., & MacKerell, A. D., Jr. (2017). Computer-Aided Drug Design Methods. *Methods in Molecular Biology (Clifton, N.J.)*, *1520*, 85–106. DOI: 10.1007/978-1-4939-6634-9_5

Zhang, L. P., Kline, R. H. 4th, Deevska, G., Ma, F., Nikolova-Karakashian, M., & Westlund, K. N. (2015) Alcohol and high fat induced chronic pancreatitis: TRPV4 antagonist reduces hypersensitivity. *Neuroscience*, 311, 166-179. DOI: 10.1016/j.neuroscience.2015.10.028.