UNIVERSITÀ DEGLI STUDI DI MILANO

FACOLTÀ DI SCIENZE DEL FARMACO

Department of Pharmaceutical Sciences

DOCTORATE SCHOOL IN CHEMICAL SCIENCES AND TECHNOLOGIES

Curriculum
Pharmaceutical Sciences (XXVII Cycle)
CHIM/08

In silico screening of Taste Receptors: an integrate modeling approach

Matteo Lo Monte

R09667

Tutor: Prof. Giulio VISTOLI

Coordinator: Prof. Ermanno VALOTI

Academic year 2013/2014

Index

Chapter 4 - Computational studies concerning the hTRPM5 ion channel

Chapter 5 - Computational studies concerning the hTRPV4 ion channel

Chapter 6 - Computational studies concerning the hTRPV1 ion channel

Conclusions

References

1. Introduction

"Taste is more than a sensual experience: it is a signal of nutritional value or danger. It has evolved as a vital survival mechanism in mammals and driven epic periods of human history — it was, after all, the quest for spices that helped launch the age of exploration[1,2]. Yet our understanding of how taste works has lagged behind the other senses."

With this fascinating sentence begins the editorial of Herbert Brody for the recent issue of *Nature* dedicated to taste. It highlights the centrality of this experience which finds countless implications in history, art, science, food industry, and obviously cooking, and which can also have beneficial therapeutic implications with relevant outcomes in drug discovery.

The way by which food is experienced is not limited to the mouth — odour, vision, hearing and even touch can profoundly change the taste thus affecting food preference. All starts before birth, since amniotic fluid contains glucose, fructose, fatty acids and amino acids, and newborns prefer the sweet taste of mother milk to other flavours.

The basic tastes include sweet, bitter, sour, and salty, even though this exclusive club might admit new members. After umami, the glutamate-dependent tastefulness, other recognized tastes include carbonation, metallic and fattiness.

Notably, taste is a whole-body experience since taste receptors, besides being located in the taste buds, are also found in non-sensory tissues, like the gut or the airways, playing still not completely known roles, for example, in glucose metabolism as well as in energy homeostasis.

Synthetic ligands for "*taste*" receptors can found several potential applications as both nutritional and therapeutic agents. With regard to medicinal perspectives, one may hypothesize that metabolic syndrome and diabetes can be treated by artificially activating the taste receptors so as to trick the body, which, believing it had eaten, would initiate the release of hormones from the gut, pancreas and fat tissues. Not to mentioned that the ability to modulate the taste receptors of fatty acids should reduce the natural preference for fatty foods, resulting in new anti-obesity strategies. From a nutritional standpoint, taste receptors have received great interest in designing new artificial sweeteners, taste enhancers, synthetic flavours and, more in general, a detailed knowledge of taste receptors should allow the rational design of new ingredients (ranging from foods stabilizers to drug excipients) with an enhanced palatability avoiding (or masking) bitter and sour tastes.

1.1 Taste sense.

Taste is the sensory impression of food or other substances on the tongue and is one of the five senses. It is produced when a substance interacts with taste receptor cells (TRC) located on taste buds. Along with smell (olfaction) and trigeminal nerve stimulation (experiencing texture, pain, and temperature), it determines flavours.

In humans, taste receptors are located on taste buds (gustatory calyculi) and on other areas including the upper surface of the tongue and the epiglottis.[3] The tongue is covered by thousands of small and clearly visible bumps, called papillae. Within each papilla, there are hundreds of taste buds, except for the filiform papillae that do not contain taste buds. Moreover, thousands of (between 2000 and 5000) are located on the back and front of the tongue. Other taste buds are located on the roof, sides and back of the mouth, and in the throat. Each taste bud contains from 50 to 100 taste receptor cells. Taste buds are able to differentiate among the five basic tastes (i.e., sweetness, sourness, saltiness, bitterness, and umami as deeply described later) by detecting interaction with different molecules or ions. Relying on this capacity of distinction, the German scientist D.P. Hanig proposed in 1901 the so-called "tongue map" hypothesis, according which there are different area on the tongue assigned to detect a unique taste. Only first in 1942, with Edwin Boring, and then in 1974, with Virginia Collings, this theory have been refuted and replaced by the fact that all tastes are equally detected across areas of the tongue.

From a physiological point of view, a functional distinction can be based on the nature of the receptors involved in taste signaling. Sweet, umami, and bitter tastes are triggered by G protein-coupled receptors (GPCR), on the cell membranes of taste buds, while saltiness and sourness are perceived through ion-channels which are activated by alkali metal or hydrogen ions, respectively.[4]

1.1.1 Sweetness.

Sweetness is regarded as a pleasant experience. Foods rich in carbohydrates are commonly associated with sweetness, although there are other natural and synthetic compounds which induce a sweet sensation at very low concentrations, and as such they can be exploited as non-caloric surrogates. Examples of non-sugar sweet substitutes include saccharin, aspartame, acesulfame potassium, sucralose, xylitol, erythritol, and stevia. Other compounds, such as miraculin, may alter

perception of sweetness itself. The molecular basis for detecting sweetness varies between both individuals and species and has only begun to be understood since the late 20th century.

Studies indicate that sensitivity to sweetness has a very ancient evolutionary beginning, since it is present as chemotaxis even in motile bacteria, such as E. coli[5]. Newborns also prefer high sugar concentrations and solutions including sugars sweeter than lactose, the sugar found in breast milk[6,7]. The perception of sweetness is based on the multipoint attachment theory, which involves multiple binding sites between a sweetness receptor and a sweet substance. Sweetness shows the highest taste recognition threshold, as it is detectable at around 1 part in 200 of sucrose in solution. Bitterness has the lowest detection threshold, about 1 part in 2 million of quinine in solution. During the evolution, a sweet food was related to an energy source, while bitterness indicates toxicity. The high detection threshold for sweetness and low detection threshold for bitterness would have predisposed our ancestors to prefer energy-rich foods while completely avoiding bitter-tasting foods. Even leaf-eating primates prefer immature leaves, which are higher in proteins and sugars and lower in fibers and (potentially toxic) secondary metabolites than mature leaves. Thus the *'sweet tooth'* has an ancient evolutionary origin and, while food processing has changed consumption habits, human physiology remains essentially unchanged.

About the mechanism of sweet sensing and among the involved GPCRs, T1Rs (taste receptor type 1) is the most important sub-family, followed by T2Rs one. They belong to the C type GPCRs, which are characterized by a large extracellular Venus flytrap domain (VFT) harboring the orthosteric binding site and which are organized in homo- and hetero-dimers.

T1R family consists of three members: T1R1 and T1R2 (identified in 1999[8]) and T1R3 (discovered in 2001[9]), that dimerize to reach a functional structure. The T1R2-T1R3 heterodimer is the actor of the sweet taste sensing[10], while the T1R2-T1R3 is responsible for the umami taste[11]. They are expressed in the TRC type II cells as well as in the testis and in small intestine[12].

To prove that the T1R2-T1R3 heterodimer functions as a sweet taste receptor, studies were conducted on knockout mice without the genes for T1R2 and/or T1R3. Mice lacking one of the genes (T1R2 or T1R3) showed a reduced response to sugars, sweeteners, and D-amino acids, thus confirming that the T1R2-T1R3 heterodimer is the main sweet taste receptor *in vivo*.

However, knockout mice showed very low, but still measurable responses to very high concentrations of natural sugars (>300 mM). This suggests that there may be additional sweet taste receptors the activation pathway of which is independent of T1R, or that T1R2 and T1R3 can function as sweet receptors also in the absence of their heteromeric partners. In knockout mice

lacking both genes, i.e., T1R2 and T1R3, the response to sweeteners has been completely eliminated[13].

T1R2-T1R3 heterodimer recognizes a wide variety of sweet substances ranging from natural sugars (sucrose, fructose, glucose, maltose), sweet amino acids (D-tryptophan, D-phenylalanine, D-serine), the sweeteners (aspartame, cyclamate, saccharin) and sweet proteins (monellin, brazzein, thaumatin)[14]. Humans and mice have different ability to identify sweeteners and sweet proteins. For example, mice do not respond to aspartame or monellin, even though the introduction of human T1R2 gene into mouse genome assimilated their taste preferences to those seen in humans.

The sweet signal transduction occurs via two distinct pathways. The first one involves cAMP and the second one involves inositol 1,4,5-trisphosphate (IP_3). It was shown that sugars cause an increase of cAMP, while sweeteners stimulate the production of IP_3[15].

Briefly, sugars bind the T1R receptor and activate the adenylate cyclase (AC) through Gs protein thus converting ATP to cAMP. cAMP can directly cause the influx of cations through cyclic nucleotide-dependent ion channels, or it can activate protein kinase A (PKA) that causes them to shut down through phosphorylation of the K^+ channels of the basal membrane. This leads to TRC cell depolarization resulting in the Ca^{++} influx through voltage-gated calcium channels and neurotransmitter release[16]. Expression of cyclic nucleotide gated ion channel in TRC cells may also contribute to membrane depolarization and influx of calcium as induced by an increase of cAMP[17].

Differently, sweetener bind to the T1R receptor and activate Gq proteins. Dissociated subunits $Gq\alpha$ or $Gq\beta\gamma$ stimulate phospholipase C $\beta2$ (PLC $\beta2$) that hydrolyzes phosphatidylinositol 4,5-bisphosphate (PIP_2) to diacylglycerol (DAG) and IP_3[18]. Elevated IP_3 levels cause the release of calcium ions from intracellular compartments leading to the depolarization of the TRC cell membrane and release of neurotransmitter[19]. Different signaling pathways to sugars and sweeteners were found in rat model. Notably, the cAMP level in rat cells after stimulation with sweeteners is significantly lower than after sucrose, while the level of IP_3 is higher in response to sweeteners than in response to sucrose. In response to sugars and sweeteners, TRC cells showed elevated Ca^{++} level, although with high sweetener concentrations the Ca^{2+} level is independent of the extracellular Ca^{2+} concentration and ruled by the Ca^{2+} release from intracellular compartments[20].

1.1.2 Bitterness.

Bitterness is the most sensitive taste and is usually perceived as unpleasant and sharp, even though it can be sometimes desirable and intentionally added by various bittering ingredients. Common bitter foods and beverages include coffee, unsweetened cocoa, South American mate, bitter gourd, olives, citrus peel, many plants in the *Brassicaceae* family, dandelion greens, wild chicory, and escarole. The ethanol confers a bitter taste to alcoholic beverages[21], which also contain other bitter ingredients such as hops in beer and orange in bitters. Quinine, a typical ingredient of tonic water, is also known for its bitter taste.

Bitterness is particularly relevant to study evolution and has clear toxicological implications since many natural bitter compounds are known to be toxic. The ability to detect bitter and toxic compounds at low thresholds provides an important protective function[22]. Various food processing techniques able to detoxify otherwise inedible foods, rendering them safe and tasty[23]. Furthermore, use of fire, changes in diet, and avoidance of toxins has led to evolution in human bitter tasting resulting in a reduced sensitivity when compared to other species[24].

Several studies have shown that T2Rs, such as T2R38 which is coupled to the G protein gustducin, are responsible for the human ability to taste bitter substances[25]. They were identified by their ability to detect known bitter ligands as well as by their morphology since they are surface bound and monomeric receptors[26]. The human T2R family comprises about 25 different taste receptors, some of which can recognize a wide variety of bitter compounds. Over 550 bitter molecules have been identified, and for about 100 ligands one or more specific receptors have been identified. T2Rs are located in the TRC cells of tongue and palate as well as in extra-oral tissues, including the airways and upper respiratory tract, where they cause bronchodilation and influence host susceptibility to infection. They are also expressed in the gastrointestinal tract, where they play a role in the emetic reflex, thus emphasizing their potential therapeutic relevance, although their exact non-sensing functions remain partially unknowns[27].

Recent pharmacological studies on T2Rs led to the identification of constitutively active mutants (CAMs)[28], useful for the identification of new bitter taste blockers, as well as for elucidating the mechanisms of T2R activation and their physiological pathways.

In contrast to the T1Rs, T2R receptors have shorter extracellular N-terminal domain[29]. Likewise T1Rs, these receptors should be able to form dimers, even though there are no clear evidence supporting this hypothesis[30]. T2Rs are coupled with gustducin and are characterized by both highly conserved motifs in cytoplasmatic loops and transmembrane domains and variable extracellular

regions, structural features which are required for the recognition of structurally heterogeneous ligands[31]. The binding site of the tastants corresponds to the first and the second extracellular loop and transmembrane domains TM 1, 2, 3, and 7[32].

Although TRC cells express either T1R or T2R receptors in the same taste buds, the T1R and T2R receptors are not co-expressed in the same cell[33]. Most T2Rs are co-expressed in the same subtype of TRC cells, which recognize a huge variety of bitter substances, but without distinguishing them. Initially, it was thought that bitter substances stimulate all cells expressing T2R receptors, even though successive studies have shown that each TRC cell recognizes one or more, but not all bitter ligands. These results are explained by Behrens and co-workers, who revealed that human TRC cells sensitive to bitter substances express only some T2Rs, and not the entire T2R family[34].

Ligands for T2R receptors are amides, such as denatonium (phenylmethyl-[2-[(2,6-dimethylphenyl)amino]-2- oxoethyl]-diethylammonium benzoate), alkaloids, such as strychnine, caffeine, and quinine, amino acids, urea, fatty acids, phenols, amines, and even potassium, magnesium, and calcium salts. Depending on their substrate specificity, T2R receptors can be activated by many heterogeneous bitter substances or by a few homogeneous ligands only[35]. The T2R family includes many orphan receptors, for which ligands have not yet been identified.

1.1.3 Umami.

The Japanese term umami was proposed by Kikunae Ikeda and can be translated as "pleasant savoury taste". Umami taste is detected by receptors for glutamate, which is a common food additive added in its salt form as the monosodium glutamate (MSG)[36]. Hence, umami has to be considered as distinct from saltiness, even though there was a long debate whether umami is a basic taste since its existence was proposed in 1908. Finally, umami is now recognized as the scientific term to describe the taste of glutamates and nucleotides such as guanosine monophosphate (GMP) and inosine monophosphate (IMP)[37]. It is a pleasant long lasting and mouth-watering taste giving a coating sensation over the tongue. Umami taste is perceived only in presence of other aroma and is acribable to the detection of the glutamate carboxylate anion by specific cells present on the tongue[38]. Its effect is to balance the overall flavour of a dish by enhancing the palatability of many foods. Glutamic acid imparts little umami taste, whereas its salts, known as glutamates, give the characteristic umami taste. GMP and IMP amplify the effects of glutamate[39]. Glutamate has a long

history in cooking and foods rich in glutamate were used widely in ancient Rome, in medieval Byzantine and Arab cuisine, and in China.

Many foods are rich in ingredients conferring umami taste. Glutamates are found in meats and vegetables, whereas inosinate comes mostly from meats and guanylate from vegetables. Thus, foods rich in umami taste include fish, shellfish, cured meats, mushrooms, vegetables (e.g., ripe tomatoes, Chinese cabbage, spinach, celery, etc.) or green tea, and fermented and aged products (e.g., cheeses, shrimp pastes, soy sauce, etc.). The first food rich in umami taste is the breast milk, which contains roughly the same umami taste as broths.

Taste buds on the tongue and other regions of the mouth can detect umami taste, regardless of their location. Biochemical studies revealed that the taste receptors responsible for the sense of umami are modified forms of mGluR4, mGluR1 and of taste receptor type 1 (T1R1 + T1R3), which are expressed in all regions of the tongue bearing taste buds[40].

The mGluR1 and mGluR4 receptors specifically recognize glutamates, whereas T1R1 + T1R3 are responsible for the mentioned synergism, although the specific role of each receptor in taste buds remains unclear. These GPCRs involve similar signaling pathways includin G proteins beta-gamma, PLC β2 and PI$_3$ which mediate release of calcium from intracellular stores[41]. In turn, Ca^{2+} activates the selective cation channel transient receptor potential melastatin 5 (TRPM5, see Chapter 4) that leads to membrane depolarization and the consequent release of ATP and secretion of neurotransmitters including serotonin[42].

Cells responding to umami taste do not possess typical synapses, but ATP conveys taste signals to gustatory nerves and in turn to the brain that interprets and identifies the taste quality[43]. mGluRs are activated are not sensitive to nucleotides, while T1R1+T1R3 are activated by a broad range of amino acids and shows a strongly potentiated response in the presence of nucleotides. Loss of the T1R1 or T1R3 genes depresses but not eliminates responses to umami. When mammalian taste buds are apically stimulated by umami tastants, their functional responses do not resemble to that of a single umami receptor. Thus, umami taste may involve multiple receptors expressed in different taste cells. This receptor diversity may emphasize the complex perception of umami taste, with different mixtures of amino acids, peptides, and nucleotides producing subtly distinct tastes[44].

1.1.4 Sourness.

Sourness is the ability to taste acidity. The sourness of a given compound is rated by using an index by which all compounds are compared to dilute hydrochloric acid, which stays as reference with a sourness index of 1.

The most common sour foods are fruits, such as lemon, grape, orange, tamarind, and sometimes melon. Wine also usually has a sour flavour, and milk, if not kept correctly, can generate a sour taste. Children usually show a greater preference for sour foods than adults, as exemplified by very popular sour candies, mostly containing citric acid.

Various mechanisms have been proposed to detect sour taste. These include acid-sensing ion channels (ASICs), hyperpolarization-activated cyclic nucleotide-gated (HCN) channels, and two pore domain K^+ channels (K2P). Sour taste is detected by a small subset of cells distributed across all taste buds, which can be identified by the expression of PKD2L1, Polycystic kidney disease 2-like 1 protein, (transient receptor potential polycystic 3, TRPP3), although this protein is not involved in sour tasting.

Protons that are abundant in sour foods can directly enter into the taste cells, thus triggering an electrical response. Also weak acids, such as acetic acid, which are not fully dissociated at physiological pH, can have a similar effect penetrating taste cells and eliciting an electrical response. After their intracellular penetration, hydrogen ions inhibit K^+ channels, which normally hyperpolarize the cell. The combination of these two depolarizing effects fires action potentials and releases neurotransmitter within the taste cell.

1.1.5 Saltiness.

The saltiness is mostly produced by the sodium ions, even though other metal ions group can also confer a taste salty, depending on their similarity to Na^+ in terms of size and charge. Similarly to what was described for sourness, also the saltiness of substances is described by an index by which each ingredient is compared to sodium chloride (NaCl), which has an reference index value of 1[45].

The apical amiloride-sensitive epithelial sodium channel (ENaC) is the major player involved in salty sensation as also confirmed by animal experiments. Deletion of *ENaC* gene causes animals to die early, as this channel plays a critical role in the kidneys, lungs, and other organs and this hampers the studies concerning the ENaC role in saltiness. To tackle this problem, Zuker and co-

workers developed an animal model lacking functioning ENaC receptors only in their taste receptor cells[46]. The so engineered mice responded normally to sweet, umami, sour, and bitter tastants with no taste for sodium chloride. Even when depleted of salt, mice without ENaC in their taste cells showed little interest in saline solutions. "*While this result was largely expected, it's very satisfying to validate the theory*," Zuker says.

The influx of Na^+ ions through the EnaC channel directly depolarizes the taste cell and releases neurotransmitter onto the afferent nerve fibers (10, 12). ENaC in expressed in TRCs as well as in other epithelia where it is essential for salt and water balance. Regardless of expressing tissue, it shows some common features including small conductance (~5 pS), $Na^+ \sim Li^+ >> K^+$ ion selectivity, regulation by extracellular Na^+ (self-inhibition) and intracellular Na^+ (feedback inhibition), and it is modulated by hormones (e.g., aldosterone, vasopressin, atrial natriuretic peptide)[47,48]. In detail, aldosterone and arginine-vasopressin (AVP) increase ENaC-mediated currents, while atrial natriuretic peptide and oxytocin reduce these currents. Clearly, ENaC plays key roles in regulating blood pressure and extracellular fluid volume.

1.1.6 Fattiness.

Fattiness is the last discovered basic taste and is involved in the natural preference for caloric end nutrient foods, thus supporting a primordial instinct by which humans and animals tend to eat foods that can improve their survival chances.

Interestingly, fatty foods send signals to the brain not only through receptors on taste buds but also, once they are ingested and absorbed in the gastrointestinal tract, at a systemic level though unknown mechanisms. Signals from both taste buds and peripheral tissues are integrated by the brain and further reinforce the preference for fatty foods via a reward system.

Fat tasting is mediated by the free fatty acid receptors (FFAR), a GPCR type A subfamily activated by free fatty acids (FFAs), which play important roles not only as essential nutritional components but also as signaling molecules in numerous physiological processes. In the last decade, FFARs have been derived from the human genome database and identified by the GPCR deorphanization strategy. To date, several FFARs have been identified and characterized as critical components in various physiological processes. FFARs are classified according to the chain length of their ligands. FFA2 and FFA3 are activated by short chain FFAs, GPR84 is activated by medium-chain FFAs, whereas FFA1 and GPR120 are activated by medium- or long-chain FFAs. FFARs act as

physiological sensors for FFAs contained in food and derived by digestion in the gastrointestinal tract. Moreover, they are involved in the regulation of energy metabolism as mediated by insulin and incretin as well as by sympathetic system, taste preferences, and inflammatory responses related to insulin resistance. Therefore and because FFARs play key roles in relevant pathophysiological processes, it comes as no surprise that FFARs represent promising therapeutic targets for the treatment of metabolic disorders including type 2 diabetes and metabolic syndrome.

1.2 Transient Potential Receptors (TRPs): a general introduction.

Besides the above described main tastes and their signaling pathways, a lot of additional sensations (such as pungency, spiciness, coolness, numbness, astringency, metallicness), mediated by many, much less known mechanisms, cooperate to give a 360° taste experience.

Among the proteins involved in perception of these additional tastes, the Transient receptor potential ion channel (TRPs) represent the most involved and physio-pharmacologically interesting family. The interest for such ion-channels also derives from the observation that the really few is known about their tridimensional structure as well as their mechanism of action and tissue-dependent roles[49,50].

They are mostly non-selective cation channels, characterized by a low voltage dependence. TRP channels are polymodal receptors since they respond to a variety of activation and modulation stimuli and are involved in diverse functions such as thermosensation, pheromone reception, and regulation of blood flow and cellular osmotic pressure.

The first TRP channel was cloned from the fly drosophila. To date, more than fifty diverse genes have been individuated among different species, twenty-eight of which have been cloned in humans. They are commonly classified in six subfamililies, according to their aminoacidic sequences similarity: (i) the "canonical" TRPC (TRPC1-7), (ii) "vanilloid" TRPV (TRPV1-6), (iii) "melastatin-related" TRPM (TRPM1-8), (iv) "polycystin-related" TRPP, (v) "mucolipin-related" TRPML and (vi) "ankyrin" TRPA[51].

All members have in common the same three-dimensional architecture made of six transmembrane segments (S1–S6) , and cytoplasmic N- and C-termini. The six membrane spanning helices can be grouped into two functionally different blocks: the first, made of S1-S4, that acts as *sensor module* and hosts the ligand binding domain, and the second, the *pore module*, made of S5-S6 and the relative linker loop, that brings the residues involved in both cations attraction and selectivity.

Similarly to the highly homologous Kv channels, a functional TRPs channel appears to be in a tetrameric conformation. In detail, a highly conserved segment of 25 amino acids, containing the *TRP box* motif, located in the C-terminus domain, immediately following the last transmembrane segment, is thought to stabilize the entire multimeric protein[52].

As anticipate above, TRP channels are characterized by a polymodal behavior which involves a variety of activation and regulatory mechanisms, some of which are still not completely known. Among the different physical stimuli, one may cite light, temperature, osmolality, mechanical force, while chemical factors include pH and endogenous and/or exogenous agonists. These stimuli may directly activate the channels or may open them through a indirectly-mediated pathway often involving PLC-coupled receptors. Moreover, some TRP channels are constitutively active because they are not activated by an identified stimulus.

Moreover, the polymodal profile is further complicated by the wide localization of TRP channels among all the human body. TRP are found not only in sensory tissues, where they play crucial roles in sensing pathways, but in non-sensory tissues as well. They are expressed in gastro-enteric systems, airways, bladder, pancreas, brain or even genitals, where they are involved in many physiologic mechanisms such as homeostasis, glucose metabolism and energy balance or synaptic activity, pain feelings and tissues suffering.

On these grounds, they clearly appears not only as interesting target for sensing modulation, but also as potential multi-therapeutic target for the treatment of many pathologic conditions.

Aim of the here reported Ph.D studies

As described in the Introduction, taste perception plays a central role and may have relevant (and industrial) implications in food-related researches but also in medicinal chemistry and nutraceutical science. In these contests, *in silico* studies can have manifold roles. They can be exploited both in rational drug design to optimize already existing ligands and in virtual screening campaigns to find new ligands possibly exhibiting new chemotypes. The so designed ligands can found applications as synthetic tastants but also they can have a medicinal role since the number of therapeutic implications of the taste receptors is exponentially growing. Moreover, computational studies can also have a mechanistic role allowing a better understanding of the receptor recognition processes and consequently of the activation mechanism at a molecular level. Hence, the numerous computational studies recently reported in literature concerning taste receptors come as no surprise.

Notably, the hitherto published studies involve almost exclusively GPCR proteins involved in taste signaling. For example, the bitter receptors have been extensively studied in silico for the relevance of predicting the bitter taste of excipients and food additives. Most studies involved homology modeling and docking simulations as clearly exemplified by the study of Goddard and co-workers, who generated the 3D structure of TAS2R38 bitter receptor and modeled its binding to well-known reference agonists[53]. In some studies, the computed complexes are then validated by targeted MD simulations and the so refined models were utilized to suggest key residues whose role was then verified by mutagenesis[54].

By considering the richness of already available studies on taste GPCR proteins, we decided to focus the here described studies on a completely different family of transmembrane receptors variously involved in taste signaling, namely the TRP ion channels. This choice has a three key reasons. First, TRP channels play relevant roles in taste perception and in particular TRPM5 (see Chapter 4) acts a general door for taste transduction in the gustatory system of bitter, sweet and umami tastes, thus suggesting that TRPM5 ligands can have a modulating role for all taste perception.

Second, TRP ligands have notable medicinal roles in analgesia as well as in metabolic related diseases such as obesity, diabetes and metabolic syndrome. Moreover, TRPM8 ligands can have a relevant therapeutic application in treatment of prostate cancer and recently studies revealed that testosterone is a potent endogenous agonist for TRPM8 which thus can play notable roles in many testosterone induced effects including psychosomatic traits.

Third, a reliable homology model was already generated using a fragmental approach by the group where the here reported studies have been conducted. Such a model has already been validated by docking simulations and MD simulations which allowed the TRPM8 mechanism of activation to be hypothesized at an atomic level.

On one hand, this model represents a satisfactory starting point for new studies on TRPM8 including rational design of novel agonists and virtual screening campaigns aimed at finding novel potent and selective antagonists (see Chapter 3). On the other hand, the encouraging results already obtained by the TRPM8 homology model emphasize that the exploited fragmental approach may be conveniently utilized to model other relevant TRP members. Thus, the present thesis describes the generation by fragments and the following validation by docking simulations for the TRPM5 (see Chapter 4) and TRPV4 (see Chapter 5). Notice that the latter shows a clearly poorer homology degree with the TRPM8 and thus its modeling represents a more challenging task compared to that of TRPM5 for which the computational strategy already used for TRPM8 could be closely used. Lastly, the thesis describes the modeling of TRPV1 as performed using a simple global approach based on the use of previously generated TRPV4 model as the template (see Chapter 6). Such a simplified strategy has the primary objective to verify whether less computationally demanding procedures can be exploited to obtain satisfactory results at least when the two involved models show a high degree of similarity (as in the case of TRPV1 and TRPV4). Nevertheless and considering the unavoidable approximations introduced by such a simplified approach, more TRPV1 homology models were generated and a tailored filtering strategy was applied.

As a preamble, the here reported studies deserve two major considerations. The first concerns their industrial profile which is reflected in the kind of objectives pursued and in the major computational strategies adopted. Starting from the key collaborations with Axxam s.p.a. and Dompé s.p.a. which inspired and contributed to the computational studies on TRPM8, the same general approach was maintained in the following reported studies which were primarily aimed at validating homology models and computational protocols to be exploited in virtual screening campaigns. The second more technical considerations concern the organization by which the results are here described which follows a chronological order. Such an approach might perhaps mix methods and results but gives a clear idea of how the here reported studies were developed during the Ph.D period.

2. Computational strategies

At the time I started my doctoral activity, among the TRPs family, only the three-dimensional structure of the TRPM8 member was already been modelled, by homology techniques[55], and validated, through molecular docking and molecular dynamics studies[56], in the group where I spent my Ph.D. course. Hence, the here described TRPM8 studies exploited such a homology model, while the studies concerning all others TRP members started from the generation of reliable three-dimensional models. Indeed and as trans-membrane proteins, the 3D structures of such receptors are not easily resolved by crystallographic methods and, thus, their *in silico* construction was performed by homology modelling.

2.1 Homology modelling.

Known also as comparative modelling, it is a computational technique which allows the generation of the three-dimensional structure for a given protein (*target*) by exploiting its amino acid sequence similarity with one or more proteins (*templates*), the structures of which have already been experimentally resolved. This method is based on the assumption that two proteins with high sequence similarity have similar secondary and tertiary structure. This hypothesis is supported by experimental evidence in which denatured proteins, placed in renaturating conditions, quite similar from the physiological point of view of pH and ionic strength, assume always the original folding. Furthermore, chaperones, proteins capable of favouring certain forms of hydrophobic folding, are able to recognize specific amino acid sequences, inducing particular arrangements of the protein chains. From these lines of evidence, it is clear that the information about the secondary and tertiary structure are intrinsically encoded in the primary one (the so-called *Anfinsen's principle*).

Another aspect to be considered is that, often, proteins with high sequence similarity are homologous, namely evolutionarily derived from the same ancestral gene: homologous proteins often retain the same physiological functions of the ancestor protein. All these considerations further corroborate the basic assumption leading to the comparative modelling approaches.

In daily practice, a choice can be made in terms of which homology modelling program to use, since there are many available as online or free downloadable tools. One of the driving reasons to

prefer one program rather than another one is, often, the availability of reliable templates and their degree of homology with the target protein.

Highly homologous templates allow the whole protein to be easily modelled. Contrarily, the lower is the homology between target and templates, the greater is the need to resort to alternative approaches. A convenient example of these different methods involve the subdivision of the primary sequence into structural segments that are then aligned and modelled separately. This strategy is known as "fragmental approach" and allow to valorise the local homology in order to obtain more reliable models.

Concerning the here reported studies, four different homology modelling programs have been utilized depending on the homology level of the available templates. In the case of low homology, I-TASSER and FUGUE were used, while in the case of high homology the choice fell on SWISS-MODEL or MODELLER. The specific protocols utilized will be detailed in the following charters, here a general description of the mentioned methods will be briefly given.

In detail, I-TASSER[57] server is an integrated platform for automated protein structure and function prediction based on the "*sequence-to-structure-to-function*" paradigm. Starting from an amino acid sequence, it generates three-dimensional atomic models from multiple threading alignments and iterative structural assembly simulations. I-TASSER guarantees a reliable product, quite complete and optimized, that needs just a final refinement. On the other hand, among the programs mentioned above, it is the most demanding in terms of time of calculation (days or weeks) and doesn't allow to queue more than one request; this makes it a powerful tool to be cautiously used only if opportune.

Beside I-TASSER, FUGUE[58] as well represents a valid alternative for modelling studies depending on templates with medium-low homology. In particular, considering its quick feedback, it is useful specially for the fragment approach, since it allows to model many small fragments in a short time. It can be used as stand-alone solution or, and this is the way I often used it, as optimizing tool for a protein modelled by a more detailed and time demanding program, such as I-TASSER, in order to improve not correctly modelled local regions. In detail, FUGUE has three key features: (1) improved environment-specific substitution tables. Substitutions of an amino acid in a protein structure are constrained by its local structural environment, which can be defined in terms of secondary structure, solvent accessibility, and hydrogen bonding status. The environment-specific substitution tables have been derived from structural alignments in the HOMSTRAD[59] database. (2) Automatic selection of alignment algorithm with detailed structure-dependent gap penalties. FUGUE uses the global-local algorithm to align a sequence-structure pair when they greatly differ in length and uses the global algorithm in other cases. The gap penalty at each position of the

structure is determined according to its solvent accessibility, its position relative to the secondary structure elements (SSEs) and the conservation of the SSEs. (3) Combined information from both multiple sequences and multiple structures. FUGUE is designed to align multiple sequences against multiple structures to enrich the conservation/variation information. The combination of these three key features implemented in FUGUE improves both homology recognition performance and alignment accuracy.

This in case of low homology templates, while in the opposite situation, when the target protein do align with high homology a template, SWISS-MODEL[60] or MODELLER[61] were the preferred tools.

The first is a good compromise between the required time and reliability of the model and indeed it is widely used in standard modelling studies. The program compares the target sequence to a single template, that can be suggested by the user as well as chosen by the program, and this clearly justifies the need of highly homologous templates.

Lastly, when highly homologous three-dimensional structures are available, the homology modelling can be performed by using Modeller, that aligns the sequences and build the models by satisfying spatial restraints. The restraints can operate on distances, angles, dihedral angles, pairs of dihedral angles and some other spatial features defined by atoms or pseudo atoms. Presently, MODELLER automatically derives the restraints only from the known related structures and their alignment with the target sequence thus resulting in generated models which are very similar to the template structures as long as their sequences are conserved. The so generated models are then optimized by exploiting molecular probability density functions (pdf). The molecular pdf for comparative modelling is optimized with the variable target function procedure in Cartesian space by conjugate gradients minimizations.

Regardless of the utilized method, homology models are often incomplete because of possible local reduced homologies by which it is not possible to generate reliable three-dimensional structures. If these local unpredicted regions are not particularly large, simplified *ab initio* methods can be used to remedy the problem. Otherwise, a second homology modelling run, using the same or different approaches, should be performed in order to model the missing regions and so to complete the predicted structure.

2.2 Model Validation.

Once the complete model is generated, it is subjected to a set of refinement processes in order to optimize its structure, by minimizing intra-molecular repulsions, which may to arise from the substitution of amino acids different from those of the template. The refinement of the highly conserved segments, such as the transmembrane helices, involved energy minimizations often keeping the backbone atoms fixed to preserve the predicted folding, while the optimization of loop segments was markedly more difficult and required specific procedures depending on their length and complexity as well as on the availability of reliable templates. To this end, MD runs of the protein embedded in suitable environments were utilized to gain a better relaxation and a more correct arrangement of whole models. The specific protocols utilized in these refinement phases will be detailed in the following chapters.

Lacking reference structures with which to compare the obtained models, the so generated models require a careful validation and their reliability was assessed by two kinds of analysis: (1) a structural validation which involved a set of checks through which the model was evaluated by using well-known structural parameters (see below) and (2) a functional validation which is focused on the binding site(s) and was performed by suitably designed docking simulations with a view to assess the predictive power of the analyzed model.

With regards to structural validation, a first preliminary check had the objective to avoid unphysical occurrences such as *cis* peptide bonds, wrong configurations, improper bond lengths, non-planar aromatic rings or colliding side-chains. Then the model underwent to a set of conformational analyses which afford an assessment of the overall quality of the model as compared with well-refined resolved structures and which also highlight regions that may need further refinement. By using specific programs, such as Procheck, this analysis involved the calculation of the percentage of residues falling in the allowed regions of the Ramachandran Plot, of the chi-space concerning the correct conformation of the side-chains, of the optimal planarity of the aromatic rings. These checks involved also the analysis of all non-bonded interactions between different pairs of residues as well as of all main-chain bond lengths and bond angles in the model. The analyses produced detailed residue-by-residue results thus revealing the protein regions requiring further refinement.

2.3 Docking set-up.

Since the ligand binding domain (LBD) of the studied receptors was known as were the key residues involved in the ligand recognition, the generated models were also validated by analysing the binding modes and the interaction features of the complexes obtained by molecular docking.

As the studies were focused on the competitive antagonism, molecules with antagonist activity were collected from literature and patents, following each time appropriates criteria of selection, but at the same time maintaining a common thread. In order to ensure statistically significant results and considerations, several tens of ligands (from 30 up to 150 molecules) per each receptors have been collected with the wider as well as equally represented range of potency as possible. Chemical diversity were also take into account, providing challenging test sets.

Once selected, all test set libraries were prepared by using an automatic script of the VEGAZZ[62] suite of programs which performs the following tasks for each compound: (i) generating the 3D structure; (ii) adding the hydrogen atoms; (iii) assigning the atom types and the Gasteiger's atomic charges; (iv) for ionizable molecules, selecting the predominant form at physiological pH; (v) for undefined chiral molecules, generating all possible stereoisomers; (vi) minimizing the so obtained molecules combining steepest descent and conjugate gradient algorithms. So optimized molecules are then ready for molecular docking calculations.

2.4 Molecular Docking.

Among so many available docking programs, the present studies mostly involved three programs, namely Autodock, PLANTS and LiGen, that represent conceptually different approaches and, as such, allow the ligand-protein interaction to be explored from different points of view.

Autodock[63] is a well-known example of stochastic docking program, based on genetic algorithm. Relying on grids of interaction energies, pre-calculated in batch mode by the internal tool AutoGrid, and based on the macromolecular target and on each type of atom in the ligand, including dispersion/repulsion forces and hydrogen bonding energies, AutoDock performs the docking calculation exploring the space driven mostly by polar interactions, according to the AMBER force field. It allows the flexibility of the ligand, while maintaining the protein atoms fixed. It was considered for long time as a landmark in the field of molecular docking, even though it has been recently outclassed by a new generation of docking programs, which are characterized be less time-

demanding calculations, parallel architectures, flexibility of both ligand and protein target and more refined scoring functions, which are able to equally account for polar and hydrophobic interactions that suitably contribute to the total energy of the so generated complexes.

Differently, PLANTS[64] is a "knowledge-based" molecular modelling program, since it places and scores the ligand-protein interactions by comparing them to an internal pattern of interactions and evaluating the recurrence. It introduces an innovative approach, based on ACO technique, for sampling the search space. in detail, the heuristic search is based on the MAX-MIN Ant System (MMAS) algorithm, through which it is possible to calculate the probability of each "ant" to choose an appropriate value of the degree of freedom relying on both the heuristics information derived from the calculation of potential through the force field, and the influence of "pheromones" previously deposited. Since the MMAS algorithm requires that only one solution for iteration can be used to deposit a pheromone, by convention, PLANTS uses only the best pose. The update of a generic pheromone depends on the score which is calculated for the best solution found for a specific iteration: only in cases in which the score has a negative value and that corresponds to a situation of high affinity between ligand and receptor, the pheromone is deposited.

Plants includes two types of docking scores. The first one involves the so-called Piecewise Linear Potential (PLP). In this score function, atoms are classified into the following four types: H-bond donor, H-bond acceptor, H-bond donor/acceptor, and nonpolar. Each pair of interacting atoms is then assigned one of the three interaction types: a H-bond interaction between donors and acceptors, a repulsive interaction for donor-donor and acceptor-acceptor contacts, and a dispersion for other contacts. The energy of each interaction type is represented by a piecewise linear function, with different parameters for different types of interaction. Total PLP Energy is then the sum of corresponding energy of each pair of ligand and protein heavy atoms.

The second Plants score function (CHEMPLP) consists of two parts: the first uses a modified version of the PLP function, which includes the estimation of the steric complementarity between ligand and receptor, while the second part implements the function CHEMSCORE of the program GOLD, that is particularly optimized in estimating the contribution of hydrogen bonds. Differently from AUTODOCK score, these scoring account for both electrostatic and hydrophobic interaction, equally.

Since PLANTS can generate multiple solutions very similar following different paths, a specific algorithm based on the calculation of the RMSD is able to identify the redundant poses and to keep only the most significant one, thus greatly simplifying the analysis of the obtained results.

Third, LiGen[65] is a "pharmacophore-based" molecular modelling program. It is less known than the previous ones since it was released very recently. It consists of a set of modules that work sequentially or as standalone tools, according to the user's need: *LiGenPass* for binding site recognition, *LiGenPocket* for binding site analysis and structure-based pharmacophore definition, *LiGenDock* for docking and virtual screening, and *LiGenBuilder* for de novo design. Thus, LiGenPocket and LiGenDock are the two modules that constitute the docking engine of LiGen.

First, LiGenPocket computes volume, shape, and physicochemical properties (donor, acceptor, hydrophobic,etc.) of the binding pocket and proposes a pharmacophore model based on these characteristics. Then, LiGenDock uses such pharmacophore scheme as the driving force for the docking procedure. It strongly depends on the quality of the pharmacophore, but it can perform even better than the other utilized programs and it requests much less time for the calculation, especially if compared to AutoDock, that requires more than thirty times the time taken by PLANTS.

LiGen scoring function, LiGebScore, for the evaluation of binding free energy in receptor-ligand complexes, derives from the well known X-SCORE[66] function. It is an empirical scoring function that computes the binding free energy as the sum of the van der Waals interactions, hydrogen bonding, deformation penalty, and hydrophobic effects between receptor and ligand. Moreover, it is able to provide a clear physical meaning of the suitable features for the ligand recognition, such as number of hydrogen bonds, hydrophobic zones, or frozen rotatable bonds during the binding process.

2.5 QSAR and Virtual screening.

Scoring structural and physicochemical properties of so obtained complexes allow us to validate the reliability of our homology model in different ways. First, certainly, the canonical QSAR studies. By exploiting tens of docking scores, as well as extended sets of ligand-dependent and complex-related descriptors, it is possible to develop correlative equations with the experimental activities that, besides permitting to predict the putative bioactivity of unknown ligands in an hit search context, act as benchmark to assess the goodness of the involved model, supporting its reliability with their statistical parameters.

Beside the QSAR analyses, the virtual screening campaigns are raising as effective validation tools to assess the reliability of the generated homology models. Virtual screening involves databases

made of active compounds dispersed among thousands of "decoy" molecules, performed in order to verify the capacity of the receptor to "select" for its ligands. Such a validation strategy appears to be particularly effective in the here reported studies which have as primary objective the screening of proprietary libraries.

Virtual screening (VS) is a computational technique used in drug discovery to search libraries of small molecules in order to identify those structures which are most likely to bind to a drug target, typically a receptor or an enzyme.

It has been defined as the "automatically evaluating very large libraries of compounds" using computer programs. As this definition suggests, VS has largely been a numbers game focusing on how the enormous chemical space of over conceivable compounds can be filtered to a manageable number of molecules that can be synthesized or purchased, and tested. Although searching the entire chemical universe may be a theoretically interesting problem, more practical VS scenarios focus on designing and optimizing targeted combinatorial libraries and enriching libraries of available compounds from in-house compound repositories or vendor offerings. As the accuracy of the method has increased, virtual screening has become an integral part of the drug discovery process[67].There are two broad categories of screening techniques: ligand-based and structure-based approaches.

In the former, given a set of structurally diverse ligands that binds to a receptor, a pharmacophore model of the receptor can be built by exploiting the collective information contained in such set of known ligands. A candidate ligand can then be compared to the pharmacophore model to determine whether it is compatible with it and therefore likely to bind.

A different strategy is to develop logic-based rules describing features of substructures and chemical properties related to bioactivity using support vector inductive logic programming. These logic-based features provide insights into bioactivity which can be understood by medicinal chemists. Support vector machine integrate the features to yield QSAR models which are then used to screen molecular databases. This approach is well suited for scaffold hopping to identify novel active molecules and is implemented in the package INDDEx. Another approach to ligand-based virtual screening utilizes 2D chemical similarity analysis methods to scan molecular databases against one or more active ligand structures. A popular approach to ligand-based virtual screening is based on shape similarity analysis by searching molecules with shape similar to that of known actives, considering that such molecules will fit the target's binding site and hence will be likely to bind the target. The interested reader can find several prospective reviews of this class of approaches in literature. Ligand-based methods typically require a fraction of a second for a single

structure comparison operation. A single CPU is enough to perform a large screening within hours, even though several comparisons can be made in parallel in order to accelerate the processing of very large databases of compounds.

Contrarily, the structure-based virtual screening involves docking simulations of candidate ligands into a target protein followed by a scoring function to evaluate the likelihood that the ligand will bind to the protein with high affinity. The size of the task requires a parallel computing infrastructure, such as a cluster of Linux systems, running a batch queue processor to handle the work, such as Sun Grid Engine or Torque PBS. A means of handling the input from large compound libraries is needed. This requires a form of compound database that can be queried by the parallel cluster, delivering compounds in parallel to the various computer nodes. Commercial database engines may be too ponderous, and a high speed indexing engine may be a better choice. Furthermore, it may not be efficient to run one docking simulation per job, because the ramp up time of the cluster nodes could easily outstrip the amount of useful work. To solve this problem, it is necessary to process batches of compounds in each cluster job, aggregating the results into a log file. A secondary process, to analyse the log files and extract high scoring candidates, can then be run after the whole calculation has been run.

Virtual screening (VS) techniques are well established tools in the modern drug discovery process and an almost unmanageable number of different VS techniques are available today[68]. Along with that, a plethora of comparative performance assessments has been published in recent years in order to support computational chemists in finding the best approach, particularly in the field of structure-based VS. However, the specific needs and aims of the VS campaigns differ considerably[69], and so do the fields of application of the different algorithmic approaches. Therefore, comparing different VS algorithms is a challenging task, particularly in case of those approaches involving protein–ligand molecular docking.

The performance of VS approaches is mostly measured in terms of their ability to discriminate between active and inactive compounds. Actives and inactives are thereby injected to the VS workflow either as one mixed library or as two dedicated collections. In the latter case, the resulting hit lists are combined before the actual analysis of the results. The collection and preparation of these compounds at the very beginning of an investigation represents the Achilles' heel of any evaluation of VS protocols.

About actives, a crucial aspect during the selection for our VS studies is the common mechanism of binding of the ligands to the target protein. Since the pharmacophore models in general are able to represent only one specific binding mode, one has to make sure that either all compounds of the

active set share the identical binding mode or that all possible binding modes are accounted for by specific pharmacophore models in a parallel screening. Moreover, the performance assessments of the scoring functions and of their ability to predict the protein–ligand affinity require reliable bioactivity information. In the best case, all data should be obtained by the same bioassay and preferably also by the same laboratory.

For what concern the inactive compounds, it is common practice to collect the so-called decoys (i.e. molecules presumably inactive against the examined target) from structural pools or to create them by using virtual combinatorial library generators such as, for example, Ilib Diverse[70] and SmiLib [71]. It has been confirmed by several studies that the characteristics of the known inactive compounds or of the decoys molecules chosen for VS assessments have a significant impact on the performances of the utilized VS approaches. Both known inactive molecules and decoys are required to meet some essential prerequisites in order to achieve meaningful results. The most important need for decoys is the similarity of their physicochemical properties to the set of the active compounds.

Probably one of the best examples for the hidden impact of decoy characteristics on the performances of VS techniques is the direct dependence of docking scores on the molecular weight of the ligands. It is known that the bigger is the ligand the higher is the computed score of the complex, since the number of functional groups and the related interactions grows proportionally.

In the here reported studies at least three physicochemical properties were utilized in order to design unbiased decoy sets. Mass, volume and charge are always taken into account because their combination well frame the key features of the simulated ligands, along with descriptors such as virtual logP[72], number of rotors, and Lipole (encoding for the separation between hydrophobic and hydrophilic functional groups within a given molecule), which are checked to better define the physicochemical properties of the set.

Last but not least, an adequate ratio of actives and decoys is necessary for both statistical reasons and to be compatible with real-life scenarios[73]. Only evaluations performed using 'high-quality haystacks' allow robust conclusions on the performance of the VS methods.

About this, all collected decoy sets have from 10000 up to 15000 molecules, with a percentage of actives toward the target of interest between 1 and 1.5%, a range of values normally used by the scientific community since it approximates the random value of actives in non-focused large libraries.

2.6 Virtual screening validation tools.

Many metrics are currently used to evaluate the performance of ranking methods in virtual screening (VS), for instance, the area under the receiver operating characteristic curve (ROC), the area under the accumulation curve (AUAC), the average rank of actives, the enrichment factor (EF), and the robust initial enhancement (RIE).

ROC, AUAC, and all averaged rank metrics show the same drawback, called "early recognition problem", since they are unable to evaluate whether the active compounds are focused on the first positions of the ordered list and thus they appear to be poor metrics for analysing VS campaigns whose purpose is to rank actives very early. Even though their use has often been questioned for the arbitrariness of the defined thresholds, the enrichment factors appear to be meaningful when comparing virtual screening campaigns performed by using the same ligand dataset, the same receptor and different computational approaches (as here performed). The major concern in using the enrichment factors in comparative studies is that they are focused on the top-ranking molecules and unavoidably ignore what happens in the remaining part of the ranking.

The goal of virtual screening method is to detect more true positives from a compounds database respect to those that would be selected using a random approach. In detail, as schematized in Figure 2.1, a VS can be seen as a strategy to select a group of molecules (n) among a starting set (N) that would contain both active compounds (true positive, TP) and inactive compounds (also called decoys, false positive, FP); among the non selected molecules there will be as well active compounds (false negative, FN) and decoys (true negative, TN). In order to evaluate the performances of such a VS method, by assessing the enrichment of actives from a mixed database, several descriptors have been developed.

Two important values are usually considered in developing these enrichment descriptors: (i) the sensitivity[74], accounting for the true positive rate (Eq. 1) and (ii) the specificity, that accounts for the ratio of discarded inactives respect to the entire decoys population (Eq. 2).
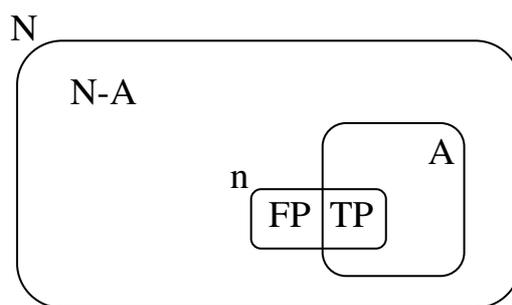
Figure 2.1 Selection of n molecules

from a database containing N entriesby a VS protocol.

Eq. 1 $$Se = \frac{N \text{ selected actives}}{N \text{ total actives}} = \frac{TP}{TP + FN}$$

Eq. 2 $$Sp = \frac{N \text{ discarded inactives}}{N \text{ total inactives}} = \frac{TN}{TN + FP}$$

Most enrichment descriptors do not include a weight for the rank that is assigned to the active molecules by the VS algorithm, considering only the fraction of selected and not the relative positions within the ranking, thus giving rise to the so-called 'early recognition problem'.

The relevance of this problem depends on the percentage of the molecules retrieved by a VS campaign which will be investigated, e.g., by biological testing. Clearly, this problem becomes a crucial issue when only a very restricted fraction of the top-ranked compounds (e.g. 0.1 % or less) will be tested and then one needs that these small fraction includes some active compounds. Therefore, it is not only important to have a VS protocol that performs well in discriminating actives from decoys, but also to have a VS workflow that is able to rank the active compounds at the beginning of a rank-ordered list.

Selectivity and specificity can be utilized to plot the receiver operating characteristic (ROC) curve, the area under which is a widely reported metric to evaluate VS performances, since it possesses suitable statistical characteristics. For example, it appears to be independent of the ratio of active compounds and decoys, even though this aspect was questioned by recent studies; it can be

interpreted as the probability that an active compound will be ranked before an inactive molecule; it has a value between 0 (worst performance attainable) and 1 (best performance) with 0.5 which corresponds to a random picking.

However, recent works have shown that the area under the ROC curve is clearly completely unable to tackle the early recognition problem because it does not depend on the specific ranking of the active compounds. Hence, some metrics have been recently proposed which are conceptually related to ROC curves but introduce weighting functions to account for the position of the active compounds. As an example, the Boltzmann-enhanced discrimination of receiver operating characteristic (BEDROC), which is derived as a generalization of ROC, but which addresses the "early recognition" problem by incorporating an exponential function to weight the contribution of the rank to the final score.

Another frequently used metric is the enrichment factor EF (Eq. 3). This descriptor takes into account the improvement of the hit rate by a VS protocol compared to a random selection[75].

$$Eq.\ 3 \qquad\qquad EF = \frac{TP/n}{A/n}$$

A flaw of the EF descriptor is its marked dependency on the ratio of active compounds contained in the considered set, thus limiting its use in comparing more VS methods only when dealing with the same databases of active compounds and decoys. Another disadvantage is that all active compounds contribute equally to the measured value, not considering their position within the analyzed subset. Thus, also the EF descriptor do not consider the ''*early recognition problem*''.

In relation to EF, the statistical significance of the enrichment (Eq. 4) is used to assess VS performance. It describes the probability that a random selection of molecules contains an equal or higher number of active compounds than a molecule selection derived by a VS protocol.

$$Eq.\ 4 \qquad\qquad S = \sum_{k=TP}^{A} \frac{\binom{A}{k}\binom{N-A}{n-k}}{\binom{N}{n}}$$

To grasp a complete picture of all the distribution, avoiding to take blunder considering what happens in the best ranked cluster only, in the here reported VS campaigns the results of an enrichment study was and represented by an histogram plot as obtained by subdividing the ranking into 100 bins. Ideally, the active compounds should be included in the first bins, decrease gradually in the following bins and be absent from the last bins. By contrast, a virtual screening which evenly distributes the active compounds throughout the ranking should be considered as ineffective regardless of how many active ligands are placed in the first bins.
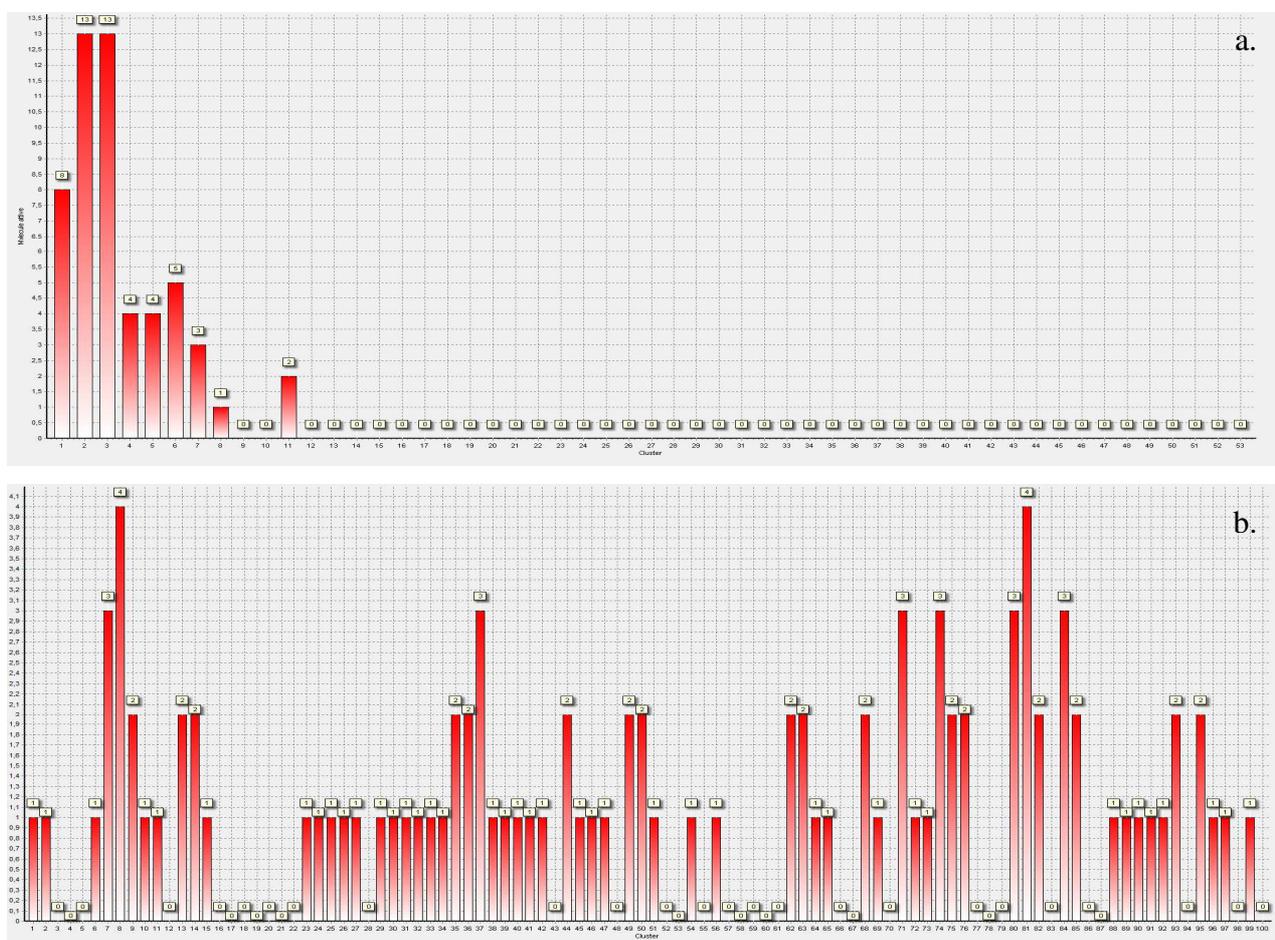
Figure 2.2 Exemplificative distribution plot derived from (a) a satisfactory and (b) an unsatisfactory virtual screening, respectively.

As examples, the first graph (Figure 2.2a) reports a satisfactory screening, since all active compounds take place on the left of the plot, falling in the best-ranked bins. Conversely, the second one (Figure 2.2b) describes a totally unsuitable screening, since the active compounds fall evenly in all bins, including the worst-ranked ones.

Besides a graphical support for the analysis of virtual screening analyses, these distributions can also be exploited to derive a novel parameter which quantitatively describes how much the active compounds are focused in the first bins and thus which can be exploited to tackle the "*early recognition problem*" . In detail, this parameter corresponds to the skewness of the distribution which is a measure of its asymmetry and is computed by Eq. 5 as the third standardized moment[76]. In a suitable distribution the parameter has to be the largest possible positive number.

$$Eq.\ 5 \qquad Skeweness = \frac{\dfrac{\sum_{i=1}^{n}(x_i - \overline{x})^3}{n}}{\dfrac{\left(\sum_{i=1}^{n}(x_i - \overline{x})^2\right)^{3/2}}{n}}$$

3. Computational studies concerning the hTRPM8 ion channel

3.1 hTRPM8 and its mechanism of action: an introduction

3.1.1 TRPM8 structural features.

Located on region 2q37.1 of chromosome 2, the human Trpm8 gene spans 102,124 bases, contains 25 exons and codes for a protein of 1104 amino acids. Several shorter isoforms have been described, but they are poorly characterized.

Structural details about the TRPM8 protein are still quite limited. Like the rest of the TRP channels, it has an overall topology similar to that of voltage-gated potassium (Kv) channels, with six transmembrane segments (S1–S6) flanked by large amino- and carboxy-termini located intracellularly. Charge-neutralizing mutations along segment S4 and S4–S5 linker region have been proved to reduce the channel gating, suggesting they are part of the voltage sensor[77]. The putative pore loop is formed by the segment between helices S5 and S6, while within the segments S1-S4, referred to as "*sensor module*", stays the orthosteric ligand binding domain.

TRPM8 subunits are thought to tetramerize into functional channels[78,79]. Oligomerization process of TRPM8 is still cloudy, it has been related to a coiled-coil domain at the C-terminal end of the protein, even if mutants missing the entire C-terminus have been proved to be able to tetramerize as well and to localize to the plasma membrane. The TRP domain, a 25 amino acid motif containing the TRP box (EWKFAR), also located within the C-terminal region (within residues 990–1025), is known to be important in PIP2 regulation and in the energetics of channel opening, i.e., translating drug binding into channel opening[80].

The C-terminus of TRPM8, as well as the TRPV1 one, have proved to contains structural elements involved in temperature-dependent gating within several functional studies on chimeric channels[81]. Moreover, the C-terminus accommodates critical sites for the activation of TRPM8 by phosphatidylinositol 4,5-bisphosphate (PIP2)[82]. Actually, little is known about the function of the long intra-cytoplasmic N-terminus of TRPM8. It contains four conserved regions, TRPM family homology sequences (MHR)[83]. Several studies showed so far that removing till the first 39 amino acids of the protein does not affect channel activity, but larger deletions do give rise to non-functional channels[84].

Bioinformatic tools predict the presence of five putative phosphorylation sites on the N-terminal domain (Ser9, Thr17, Thr32, Ser121, and Ser367), some of which are confirmed protein kinase A (PKA) phosphorylation sites[85]. A short, leucine-rich conserved sequence among TRPM channels,

known as the ICF region (indispensable for channel function), has been characterized in some detail. TRPM8 proteins lacking this region are still expressed in significant, even if reduced, amount at the cell surface but do not constitute functional channels themselves[86].

As discussed later, the main binding sites for chemical agonists and antagonists have been mapped within the *sensor module* of the monomer. In particular, residueTyr745 in the S2 segment is involved in channel activation by menthol and also mediates inhibition by SKF96365[87]. In contrast, the agonist effects of icilin require specific residues within the helix S3.

TRPM8 channels is also subject to post-translational modifications able to affect its function. For example, the channel is glycosylated on an asparagine residue in position 934, near the putative vestibule of the ion channel pore. This change is not critical for channel assembly or function but it modulates the channel activity[88]. Moreover, glycosylation of TRPM8 affects its association to lipid rafts[89]. Flanking this glycosylation site, two cysteine residues, namely Cys929 and Cys940, are proved to be essential for channel function with their disulfide bond[90]. A recently observed modification of TRPM8 consists in the covalent linkage of the protein to repeated units of R-3-hydroxybutyrate, giving rise to the attachment of large polymeric chains of poly-(R)-3-hydroxybutyrate (PHB)[91]. Mutants lacking specific PHB-binding sites showed highly reduced response to menthol or cooling.

3.1.2 TRPM8 functional features.

TRPM8 is a non-selective cation channel, with modest calcium permeability. The relative permeability for Ca^{2+} versus Na^+ (P_{Ca}/P_{Na}) has been estimated at 1–3.3[92]. At the single-channel level, the current versus voltage relationship of TRPM8 channels is approximated linear with an estimated conductance of ~60–70 pS at 20°C, increasing to ~95 pS at 30°C[93].

In sensory neurons and their terminals, TRPM8 is thought to monitor fine changes in ambient temperature. In agreement with this physiological role, the gating of TRPM8 shows steep temperature dependence, with Q_{10} estimates around 30–40[94].

Cold activation is accompanied by wide changes in both enthalpy and entropy but leads to relatively small changes in free energy[95]. The molecular mechanism(s) underlying TRPM8 thermal sensitivity is still quite unknown. About that, experimental evidences that TRPM8 channels purified from bacterial expression system and then reconstituted into planar lipid bilayers can be gated by temperature[96], in addition to evidences about the thermally sensitive modules identified in TRPM8

structure, as well as in other TRP members[97], clearly support an intrinsic thermal sensitivity of the channel[98].

A functional feature shared by TRPM8 with other TRP channels is the polymodal activation, that makes them signal integrators. Beside chemical stimuli (e.g., menthol), TRPM8 is also gated by voltage. In order to better understand the function of the channel it has to be explained how the different sensors within the protein communicate with the pore to induce its opening and closing. Both low temperature and menthol increase probability of TRPM8 opening[99]. Moreover, menthol sensitizes the activation of TRPM8 by temperature and vice versa.

The voltage dependence of TRPM8 shows a characteristic outward rectification at depolarized transmembrane potentials and a rapid and potential-dependent closure at negative membrane potentials. TRPM8 gating charge is reasonably small compared to that of classical voltage-gated potassium channels (about 0.6–0.8 e). As a result, small changes in the Gibbs free energy of the channel can cause large shifts (>100 mV) in the voltage-activation curve[100]; while, the shallow voltage dependence curve implies a wide gating window, enabling a precise fine-tuning of the extent of channel activation rather than an all/none type of response. Charge-neutralizing mutations in the TRPM8 domains S4 and the S4–S5 linker, namely on Arg842 and Lys856, have proved to reduce the gating charge, suggesting that this region is part of the voltage-sensing domain, as happens in voltage-dependent potassium channels (K_v).

A detailed characterization of both voltage- and temperature-dependent gating of TRPM8 channels at the single-channel level showed an intricate gating mechanism. The observed increased channel open probability, caused by cooling or membrane depolarization, appeared to be mainly due to a decrease in closure intervals duration time, showing only a modest increase in mean open times. This behaviour would fit much more to a multi-state gating model, in which would exist at least two open and five closed states. Coherently, expression of short TRPM8 isoforms in prostate downregulated the channel activated by cold or menthol, but did not lead to any change if activated by icilin[101].

At the single channel level, the effect of short isoforms co-expression mimicked the effects of high temperature (i.e., long closures). In native channels, the apparent temperature activation curve of TRPM8 is shifted to temperatures warmer then the ones related to recombinant TRPM8 channels expressed in hippocampal neurons[102]. This shift is in line with an apparent shift in voltage-dependent activation toward more negative potentials. Altogether, this would explain the important difference in thermal threshold observed between cold thermo-receptor fibers expressing TRPM8

and recombinant channels. In line with this behaviour, static warm ambient temperatures have proved to shift the temperature threshold of TRPM8 to warmer values[103] .

The TRPM8 polymodal activation has stimulated the development of several models aiming to describe the relation between chemicals binding and channel gating. In 2004, Nilius and colleagues proposed that cooling and menthol activate TRPM8 through a shift in the voltage dependence of activation ($V_{1/2}$). In more detail, both agonists would shift the TRPM8 voltage-activation curve towards more negative potentials, so increasing the probability of channel openings and encouraging inward currents at physiological membrane potentials. A two-state gating model, with a direct coupling of temperature and voltage, reproduced the macroscopic *G/V* curves correctly and could adequately predict the temperature dependence of TRPM8 whole-cell currents at different potentials. However, the complexities of the gating of single TRPM8 channel indicate that multi-state models would better explain the gating mechanisms[104].

Differently, Latorre and collegues proposed an allosteric model in which TRPM8 gating activation by voltage and temperature lead to augment the open probability of the channel both separately and in concert. In this view, sensors for voltage, chemical agonists, and temperature would be located in different protein domains. In this direction, Matta and Ahern found that very high agonist concentrations are able to activate TRPM8 in a totally independently from voltage, while voltage alone cannot completely activate TRPM8[105], which would be the case if voltage was the final gate in channel activation (two-state model). While the allosteric model would seem to reproduce experimental data, the independence of the many fit parameters is not obvious at all.

As highlighted by Zhu[106], it is a challenging issue to understand the key differences between the allosteric model and the two-state model, since this mechanisms can be influenced by experimental factors such as desensitization or voltage clamp errors. In a more recent publication, Voets etcelleagues generated a Monod–Wyman–Changeux (MWC)[107] model explaining the combined effects of voltage, temperature, and menthol on TRPM8 gating. The model was proved to faithfully reproduce experimental data, predicting the menthol-related shifts in both the voltage-activation curves and the temperature activation threshold. Practically, without the ligand, the MWC channel is reduced to a two-state model. In practice, the MWC model appears to be able to predict voltage- and time-dependent kinetics, while being more simple than the allosteric one. Nevertheless, this model cannot accommodate data which point to a menthol-induced increase in the maximum open probability.

Finally, Yang et collegues, working on the TRPM8 thermodynamics, reformulated the allosteric model idea, proposing that activation of TRPM8 by temperature, voltage, and menthol constitutes

three stand-alone mechanisms. It remains to be clarify how this view can be consolidated with the existing data of agonist-induced effects on TRPM8 voltage dependence. In a recent publication, Jara-Oseguera and Islas introduced a novel hypothesis about the allosteric modulation of the temperature-dependent TRP channels activation. In a MWC-type model, they showed that a large decreasing of the coupling constant between temperature sensor activation and channel gating could reverse the sign of the temperature sensitivity[108]. Practically, TRPM8 channel would bring a heat sensor negatively coupled to the gate that during cooling would cause the opening of the channel.

3.1.3 TRPM8 expression.

TRPM8 channel was originally cloned from prostate tissue[109] and named Trp-p8, before being detected in a subpopulation of small-diameter cold-sensitive peripheral sensory neurons[110]. Following investigations using knockout mice lacking functional TRPM8 led to attest that it plays a crucial role in detecting cold temperatures in the environment[111].

Moreover, TRPM8 was proved to be strongly upregulated in several tumoral conditions, originally in testis and prostate, but recently even in many other cancerous tissues[112]. In prostate, healthy or cancerous as well, TRPM8 gene expression is regulated by androgen receptors, but the precise subcellular localization of TRPM8 in this tissue is somewhat controversial, apparently depending on the oncogenic status.

3.1.4 TRPM8 physiology.

Relying on the detection of TRPM8 in a subclass of thermo-sensitive peripheral sensory neurons, early studies investigated its putative role in temperature perception and pain.

Diverse sensations can follow to cold temperature sensing, going from pleasant to painful as temperature decreases. These sensations are thought to derive from activity in sensory fibers with different thermal thresholds.

Exposing the skin to mild cold temperatures excites a subset of peripheral somatosensory neurons. In the same fibers, menthol appears to sensitize the response to low temperatures, behaviour known as "cooling mimetic action". Both cooling and menthol are able to generate inward depolarizing currents which in turn depolarizes the cell and releases the chemical signal[113]. Evidences from

several experimental studies run on cold-sensitive neurons showed how the non-selective cation current (I_{cold}) reports physical/chemical properties in line with TRPM8-dependent currents in transfected cells[114]. Furthermore, it was evidenced how several peripheral cold-sensitive neurons does not express TRPM8, which on contrary was proved to be expressed within the most part of such cold sensing tissues[115], suggesting the need of an additional cold sensor.

The role of TRPM8 in detecting noxious cold stimuli have been increasingly studied during the recent period, approaching it in diverse physio/pathological contest, such as, just to mention one, the hypersensitivity to cold (known as cold allodynia) and its treatment with TRPM8 antagonists[116]. Notably, the variable presence of potassium channels has proved to modulate the sensory fibers sensitivity to temperature changing, further demonstrating how variable the operating range of cold receptors could be.

Receptors sensitive to cold temperature in peripheral neurons are known to be constitutively active at neutral skin temperatures as well, generating a steady neural signal that reaches the brain and maintains constant all related processes. When exposed to a lower temperature, cold receptors activity increases leading to spontaneous thermoregulatory responses, among which include decrease of superficial vessels volume as well as increase in thermogenesis from fat, by directly increasing the expression of thermogenin (uncoupling protein 1).

In line with this physiological behaviour, cooling agents acting on TPRM8 lead to response in body temperature, which increases in defence of the organism; while, contrarily, TRPM8 blockers have been proved to inhibit this reflex, leading the body temperature to decrease of about 1 °C [117].

3.1.5 TRPM8 related pathologies.

To date, TRPM8 expression profile variations as well as its incorrect functionality have been related to several pathologic conditions, most of which, further supported by the relatively confined presence of TRPM8 in sane organs, have been identified as cancerous diseases, often prostate related. This clearly makes it a promising target for both diagnosis and, hopefully, for therapeutic treatment too, reason why the pharmacological interest in it keeps growing.

As an example of a diagnostic use, TRPM8 expression profile can be analysed to determine the prostate tumor progression, since directly proportional to the androgen hormone. Furthermore, TRPM8 mRNA has recently been reported to be tracked in blood and urine as well[118]. Here, it can

be used as metastasis prognosis detector since its level was proved to be much higher in this condition than in both subjects sane and patients with tumor still localized at organ level.

Unfortunately, from the pharmacological targeting point of view things get more complicated, since TRPM8 role in prostate cancer is still unknown.

Beside prostate tumor, TRPM8 over-expression have been detected in other different kind of cancers, in which it passes from being virtually undetected in physiological conditions to be considerably increased in tumoral environment.

To mention just few examples, cooling agent TRPM8 triggering in melanoma cells would reflect into sustained calcium influx and reduced cell viability[119]; while, in breast adenocarcinoma, it is mainly detected with abnormal frequency in small, level I cancers, suggesting how it would be related to low proliferative diseases.

Cancer is not the only big issue with which TRPM8 has been related so far.

The neural mechanisms underlying the cold-related pain are quite intricate, involving both activation of nociceptors via low temperature and several interdependent afferent inputs in the central nervous system, where finally the sensory experience takes place[120].

In order to better understand these mechanisms, several studies have been conducted on TRPM8 antagonism in animal models of neuropathic pain. For example, not only cold allodynia was proved to be sensibly alleviated by treating the animal with a TRPM8 antagonist, the Capsazepine[121], but also TRPM8 knockout models showed no response to temperature variation in neuropathic cold pain conditions.

On the other hand, cold is also known to have analgesic effects on pain and so TRPM8 agonists as well are object of interest in pain alleviation, specially when inflammatory conditions are existing, in line with the common practice of cooling the sore skin to receive a lenitive sensation. The intricate paths underlying this analgesic effect are still far from being completely clears.

3.1.6 TRPM8 ligands.

For convenience, it is usual to refer to TRPM8 ligands subdividing them into two main categories depending on their chemical structure: (i) the first one hosts the so called para-menthane based compounds, while (ii) the second one contains the non-para-menthane based ones. Among the first one, clearly united by the shared isopropyl-methylcyclohexan structure, both agonists and antagonists take place; while, contrarily, except for icilin, an agonist twenty time more potent then

menthol, the second class mostly accounts for antagonists only. To date, several TRPM8 ligands with diverse chemical structures have been individuated, mostly discovered through HTS of proprietary in-house chemical libraries in several pharmaceutical companies.

3.1.6.1 Para-menthane-based TRPM8 ligands: agonists.

l-(–)-Menthol is the classic menthane-based TRPM8 agonist. It is a common and popular flavoring and cooling additive in household remedies (topical pain gels, throat lozenges, toothpaste) and foods (gum, candies, teas), and is known to produce its cooling effects by activating the TRPM8 ion channel. Menthol binds to TRPM8 with an EC_{50} of 66.7±3.3μM in Xenopusoocytes or 30μM in transfected CHO cells, however; this monoterpenoid also non-selectively activates TRPV3 (EC_{50}20 mM), inhibits mouse TRPA1 (IC_{50}68μM) but activates human TRPA1.

Menthol is also known to have anesthetic, analgesic and counter-irritant properties, most of which are now linked to its TRPM8 agonist activity. During a wide mutagenesis study, Tyr745, placed in the middle of segment S2, was identified as crucial for the menthol sensitivity of mouse TRPM8[122].

Residues Tyr1005 and Leu1009 situated in the TRP box of the intracellular C-terminus were also found to participate in channel activation by menthol, but were postulated to be involved in events downstream of menthol binding. Using tandem tetramers of mutant TRPM8 channels with low affinity for menthol, Janssens and Voets estimated that each channel can bind up to four menthol molecules[123].

Starting from the individuation of both TRPM8 gene and mRNA in normal and malignant prostate cells, TRPM8 ligands have been investigated for efficacy in prostate cancer and benign prostatic hyperplasia (BPH), a non-cancerous enlargement of the prostate gland. Among these agonists, for exemple, the benzimidizalone D3263[124] was advanced to Phase I clinical trials in diseased patients and appeared to be well-tolerated, showing preliminary efficacy for disease stabilization in three out of eight male patients with advanced prostate cancer.

The expression of TRPM8 in prostatic tissue is regulated by androgen and is upregulated in androgen-responsive tumors[125] till its lost when reaching the androgen independence condition, making TRPM8 expression a valid prognostic marker for cancer progression.

Clearly, TRPM8 agonists may have therapeutic potential for treatment of early-stage prostate cancer and prevention of tumor progression.

3.1.6.2 Para-menthane-based TRPM8 ligands: antagonists.

Among para-menthane-based TRPM8 antagonists, Ortar and colleagues have reported a series of (−)-menthylamine-based molecules which TRPM8 affinity and selectivity was obtained by installing substituents found in TRPV1 ligands such as BCTC onto the TRPM8 (−)-menthylamine scaffold[126].

3.1.6.3 Non para-menthane-based TRPM8 ligands: agonists.

Among non-menthane-based ligands for the TRPM8 ion channel, Icilin is a well characterized potent agonist and *"potent cooling agent"*, acting as TRPM8 agonist with 200-fold greater efficacy (EC_{50}0.36μM) than menthol[127]. To date, icilin seems to be the only one non-p-menthane based TRPM8 agonist. The structural dissimilarity between TRPM8 agonists icilin and menthol seems to point to different binding sites and/or different activation mechanisms, further supported by the fact that activation of TRPM8 by icilin and cold (17°C) is modulated by intracellular pH, whereas the same is not true for menthol.
Icilin activation of TRPM8 seems to require the concomitant elevation of intracellular calcium, with TRPM8 acting as a coincidence detector of two different stimuli. A comparative analysis of rat TRPM8 and icilin-insensitive avian TRPM8 based on the analysis of chimeras and mutant channels revealed the critical residues for icilin sensitivity to be Asn799, Asp802, and Gly805 in the transmembrane domain[128].

3.1.6.4 Non para-menthane-based TRPM8 ligands: antagonists.

There has been a significant effort by the pharmaceutical industry in discovering TPRM8 antagonists for various therapeutic applications, particularly pain. By and large, most TRPM8 antagonists reported in the journal literature and patents are non-p-menthane based antagonists, obtained by optimizing high-throughput screening hits. Little commonality other than the presence of perhaps two phenyl rings, one of which usually contains a fluoro- or trifluoro-substitution, is

seen in these compounds. Some well-characterized TRPM8 antagonists have been used as tools by other laboratories to understand the pharmacology of TRPM8 in pain models.

Among this class, 4-benzyloxyphenylmethylamides, 4-benzyloxybenzoic acid amides, 2-benzyloxy-benzoic acid amides (AMTB), benzyloxy-phenylmethylcarbamates (PBMC), 3-Chloro-2-pyridinyl-N-[4-(1,1-dimethylethyl)phenyl]-1-piperazinecarboxamide (BCTC) and benzyloxyphenylmethylureas have been found to possess nanomolar activity.

High affinity antagonists, such as AMTB and PBMC have been used by both academia and industry as selective pharmacological probes of TRPM8 to explore the range of therapeutic applications resulting from blockade of this receptor, such as the regulation of deep body temperature (Tb)[129], treatment of urological disorders[130] and cold hypersensitivity associated with inflammatory and neuropathic conditions[131].

A focused high-throughput screen of in-house chemical libraries, followed by further structural elaborations, leads Calvo et al. to identify 2-vinyl cycloalkyl-substituted benzimidazoles as moderately potent TRPM8 antagonists[132].

As well, Parks et al. run into a series of heteroatom-substituted spiro[4,5]dec-2-enes substituted at the 2-position of a benzimidazole core and optimizing studies culminated in the identification of a trifluoro-substitute, revealed an orally bioavailable, potent, and selective TRPM8 antagonist exhibiting antiallodynic properties in *in vivo* pain models[133].

Janssen has also developed an extensive series of benzothiophenephosphonate esters[134] and pyrrolidines[135], including JNJ-39267631, active in various models of allodynia associated with neuropathic pain and hyperalgesia associated with inflammatory pain. Compound JNJ-39267631 is also active in the WDS model, has potent antagonist activity against icilin, menthol and cold-induced TRPM8 responses, and is selective versus TRPV1, TRPV2 and TRPA1 channels.

The actions of TRPM8 blockers on native thermo-receptors are still relatively unexplored, with only a few studies addressing the antagonism of TRPM8 in sensory neurons[136,137,138] and sensory nerve terminals[139].

Malkia and colleagues showed that several antagonist compounds (e.g., BCTC and SKF96365) act as negative allosteric modulators of channel gating, shifting the voltage activation of TRPM8 towards more positive potentials, suppressing the depolarizing effects of temperature and chemical agonists[140]. They showed that co-application of different concentrations of an agonists and an antagonist could be used to titrate the voltage-activation curve of TRPM8 to any desired value, leading to changes in thermal activation threshold. Tyrosine residue 745 at the menthol binding site is critical for inhibition mediated by the compound SKF96365. In contrast, the inhibition by BCTC

is unaffected suggesting that at least one other binding site exists on the TRPM8 channel from where this drug exerts its negative allosteric modulation.

3.1.7 TRPM8 computational background.

As mentioned before, at the time I started my PhD course, a homology model of the hTRPM8 was already obtained and validated by the group[141]. Briefly, the construction of the monomeric structure of TRPM8 channel was achieved in three steps, here resumed. The first part involved the modeling of the vast cytoplasmic N-terminus by the classical homology techniques, basing on the structure of importin-β (PDB: 1QGR). The second step concerned the generation of the TM bundle by fragments using the structure of Shaker channel Kv1.2 (PDB: 2A79) to guide the final assembly. Lastly, the C-terminus was modelled using the experimental structure of HCN2J pacemaker (PDB: 1Q43). Once assembled, structurally optimized and energetically minimized, the obtained monomer was used to build the corresponding quaternary structure through protein–protein automatic docking using the EscherNG[142] program, via a two sequential steps strategy. In the first step, two TRPM8 monomers were docked in order to obtain the corresponding homo-dimer, and then two dimmers were docked together to generate the final TRPM8 tetramer.

Then, several docking calculations of representatives agonists and antagonists were performed on both monomeric and tetrameric protein, allowing to validate the reliability of the model and, in particular, of its ligand binding domain.

Relying on several studies emphasizing the key role of S4 helix and S4–S5 linker which act as a voltage sensor integrating thermal and chemical stimuli in TRPM8[143], and, considering the recently demonstrated critical role of S4 helix in a voltage-dependent phosphatase[144], in which a conformational shift during the channel activation range from the standard α-helix to a more elongated 3–10 helix, our group hypothesized a three steps activation mechanism.

Moreover, mutagenesis studies revealed that the neutralization of the positively charged residues, which characterize the S4 and S4-S5 linker, alters voltage dependence and thermal sensitivity and affects menthol binding and activation, supporting a general activation mechanism based on conformational shifts happening in this region. The proposed role of S4 is also in line with the evidence that opening of the pore does not require conformational changes in S5 or S6 but is based on a rigid-body movement of the entire module[145].

3.1.8 Proposed activation mechanism for TRPM8.

Analysing the geometric relations between menthol binding cavity, the segment S4 and S4–S5 linker, one can note that Arg842 is placed between Glu802 and Asp835 and probably it can interact with such negative residues changing the conformation of its side chain. Moreover, the ion-pair between Arg842 and Glu802 may be influenced by Tyr745 and then by the interaction between this latter and ligands. The particular arrangement of these residues seems to suggest that Arg842 can move between Glu802 and Asp835 depending on ligand binding and/or voltage and such a transition may influence the S4 structure. Considering our model, one can suppose that the ligand binding breaks the interaction between Tyr745 and Glu802 favoring that between Arg842 and Glu802, which implies a down movement of Arg842 and the shift of S4 segment to 3–10 helix. This may mean that the menthol binding induces a local change in S3 which approaches Asp802 to S4 promoting the contact between Arg842 and Asp802. Globally, these considerations may suggest that: (i) the close state is characterized by S4 in canonical αhelix, while the open state requires its elongation to 3–10 helix; (ii) in the close state Arg842 interacts with Asp835 stabilizing the αhelix conformation in the S4 helix; (iii) in the close state Glu802 is turned towards S2 and interacts with Tyr745; this interaction is broken when the ligand interacts with Tyr745; (iv) the open state requires a transition of Glu802 towards S4, so that it can attract Arg842 which, in turn, leaves Asp835; (v) such a shift of Arg842 unblocks the S4 conformation, which assumes the elongated 3–10 helix. Such observations propose that the open state involves an elongation of S4 which induces a joint movement of S4–S5 linker and, more importantly, the opening of S5–S6 pore. The described mechanism could be triggered also by voltage, which can influence the arrangement of charged residues and/or by pH of milieu which can change the ionization state of considered charged residues.

3.1.9 Molecular Dynamics study.

With a view to substantiating the hypothesized mechanism, a molecular dynamics (MD) simulations study involved the TRPM8 monomer alone and in complex with selected ligands. To avoid too long simulations, the MD runs were intentionally targeted so as to focus the attention on the conformational shifts directly induced by ligand binding (namely, the approaching of S3 towards S4 and the shifts of the latter).

In detail, the study involved four MD simulations considering the TRPM8 alone and in complex with menthol, icilin and AMTB, a known TRPM8 antagonist. While the choice of the representative agonists (namely, menthol and icilin) was quite obvious involving the most known natural and synthetic compounds, AMTB was chosen considering that a recent mutational analysis suggested that some hydrophobic inhibitors (such as capsazepine and BCTC) interact in a lateral binding cavity which only marginally overlaps with the orthosteric binding site[146], whereas the more polar AMTB should occupy the orthosteric cavity in a competitive manner.

The MD simulations lasted 5 ns and involved TRPM8 alone and bound to menthol, icilin and AMTB. To simplify the simulation space, the MD runs involved a single monomer and all atoms were kept fixed apart from those within a sphere of 15 Å radius around the bound ligand. This sphere was chosen to include the lower region of the transmembrane voltage sensor (S1–S4) plus the S4–S5 linker.

Figure 3.1.1 reports the dynamic profiles for the distance between Tyr745 and Asp802 showing that the starting conformation of all simulated complexes is similarly stabilized by a close contact between Tyr745 and Asp802. Nevertheless, the four simulations show quite different behaviours. Indeed, while the free TRPM8 (Figures 3.1.1-3, blue plot) and the channel bound to AMTB (Figures 3.1.1-3, yellow plot) conserve the contact during the whole simulation, the binding with menthol and icilin disrupts this interaction and the monitored distance progressively increases during the simulation. Moreover, Figure 3.1.1 shows that menthol (Figures 3.1.1-3, red plot) induces a slower but more pronounced conformational shift compared to that exerted by icilin (Figures 3.1.1-3, green plot). The monitored differences could be somewhat explained by a visual comparison of the first and last frames of the monitored simulations. During the simulation time, both agonists reinforce their interaction with Tyr745 through the hydroxyl function for menthol and through the nitro group and the pyrimidinone ring for icilin. Due to its small size and hydrophobic nature, menthol appears to be wedged in the binding site physically distancing Tyr745 from Asp802 and explaining the significant increases as seen in Figure 3.1.1. Despite its polarity, AMTB, conversely, engages the TRPM8 binding site mainly stabilizing p-p stacking and charge transfer interactions and it never contacts Tyr745 or Asp802 during all the monitored simulation. Taken together, these results seem to confirm that TRPM8 agonists are characterized by ability to strongly engage Tyr745, thus liberating Asp802 which in turn can approach Arg842, while the antagonists, irrespective of their polarity, occupy the binding site or block the entrance, as recently proposed by Malkia[147], without contacting Tyr745 nor influencing its interaction with Asp802. This result shows

that the performed simulations can clearly discriminate between agonists and antagonists and affords an interesting confirmation for the first triggering step of the proposed mechanism.
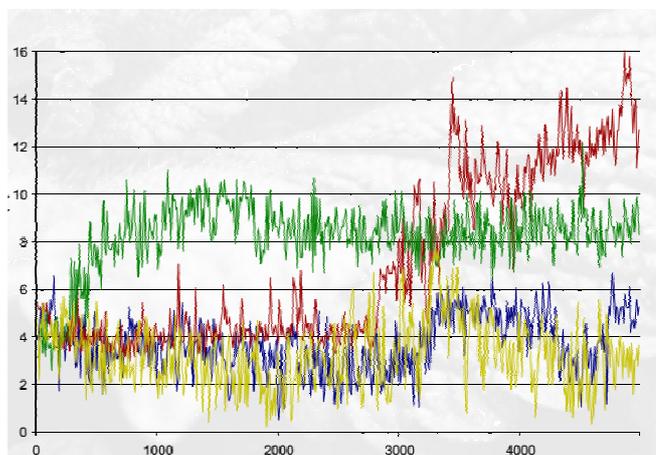
Figure 3.1.1 Dynamic profile of the distance between Tyr745 and Asp802
within a 5ns MD simulation.

Concerning the approaching of Asp802 towards Arg842, Figure 3.1.2 reports the distance between these residues as monitored during the MD runs. One may note that all complexes are characterized by an initial distance of about 10–12 Å which indicates a lack of significant interactions between them. As already seen in Figure 3.1.1, also Figure 3.1.2 shows marked differences between the performed simulations. Indeed, while the distance between Asp802 and Arg842 remains constantly above 10 Å during the entire simulations of the TRPM8 free and bound to the AMTB antagonist, the simulated agonists induce a clear approaching characterized by a distance between Asp802 and Arg842 of about 5 Å thus indicating a relevant ionic interaction between them. Remarkably, the timing of the approaching of Asp802 towards Arg842 as seen for the agonist complexes is in line with that of moving away between Tyr745 and Asp802 as reported by Figure 3.1.1. This observation seems to confirm that the two monitored structural changes can be seen as related steps of a concerted mechanism activated by agonists and leading to the channel opening.
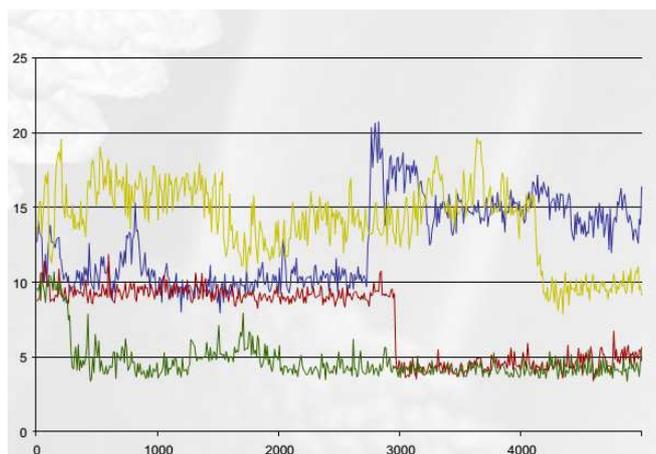
Figure 3.1.2 Dynamic profile of the distance between Asp802 and Arg842
within a 5ns MD simulation.

The last analysis involved the conformational shifts which happen in the S4 segment due to the agonist binding. To simplify the analysis and considering that the imposed constraints could alter the exact movements of this segment, the attention was simply focused on the atomic fluctuations as described by RMSD values computed taking into account only the S4 atoms and irrespective of the conformational shifts underlying such fluctuations. Figure 3.1.3 depicts the dynamic profile of the RMSD values showing interesting differences between the monitored simulations. Indeed, after a common starting equilibration phase, the S4 segment evidences a clearly greater mobility when the channel is bound to agonists compared to TRPM8 alone or bound to AMTB. Yet again, there is an evident synchronization between the RMSD variations induced by the simulated agonists and the conformational shifts as analyzed by Figures 3.1.1 and 3.1.2. Indeed, the RMSD increase occurs almost immediately with icilin and after about 3 ns with menthol. This further confirms that the monitored structural variations can be seen as sequential steps of the same activation mechanism selectively induced by the agonist binding.
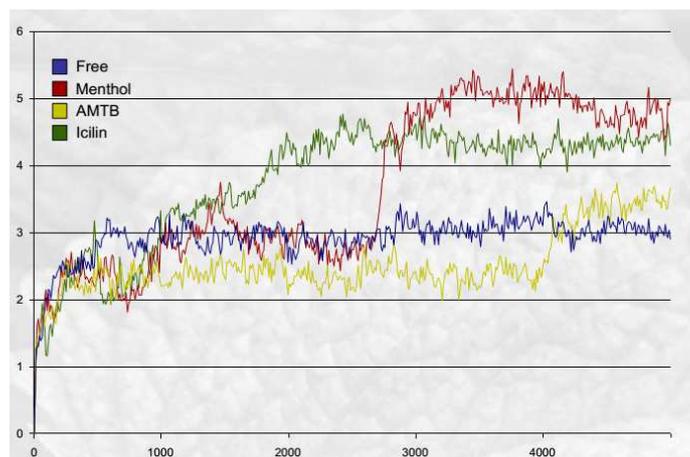
Figure 3.1.3 RMSD fluctuations of S4 helix
within a 5ns MD simulation.

The comparison of the three plots for menthol complex evidences how all main transitions occur at the same time (black circle) thus suggesting that they are concatenate steps of an unique concertate mechanism (Figure 3.1.4).
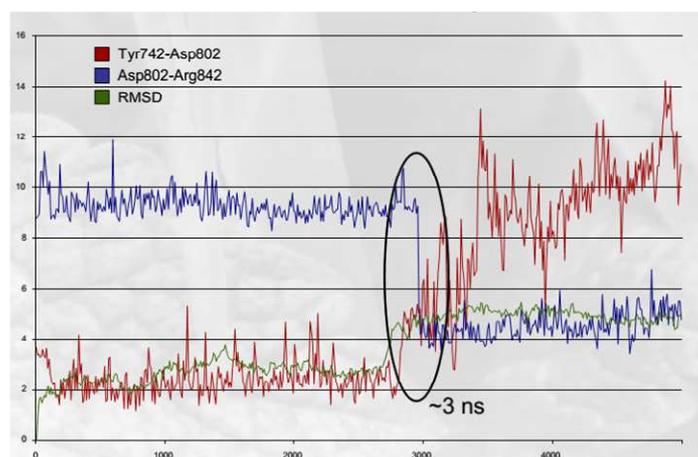
Figure 3.1.4 Summary of the three steps mechanism for the TRPM8-menthol complex.

3.1.10 Here presented computational studies.

As anticipated above, the studies focused on TRPM8 and carried out during my Ph.D activity were thus based on this reasonably reliable homology model, the validation of which culminated in the above described MD study to which I contributed before beginning my Ph.D course. These studies can be roughly subdivided in four parts as schematized below:

1. Virtual screening study performed combining ligand-based and structure based approaches on a rather large proprietary database, which led to the identification of potent and selective naphthyl based antagonists.

2. Docking simulations on a set of isoxazolylamine agonists which allowed a more precise analysis of the TPM8 binding site revealing which residues are responsible of ligand potency and which residues of ligand affinity.

3. Blind docking simulations performed on the whole TRPM8 tetramer with a view to identify additional allosteric binding sites whose reliability was next assessed by mutagenesis.

4. Selectivity studies by considering the binding sites of all members of TRPM family and utilizing the same database already used in the first study.

Apart from the second study which was performed in collaboration with the Department of Pharmacy of the University of Naples Federico II, all other three studies were conducted within a collaboration with Dompé s.p.a. and Axxam s.p.a.

3.2 TRPM8 Virtual Screening study

3.2.1 Introduction.

Nowadays, virtual screening plays a key role in the discovery of innovative leads and represents an efficient method to reduce the costs and to increase the hit rates compared to wet-lab high-throughput screening (HTS). Notably, also virtual screening techniques can be subdivided into ligand-based and structure-based approaches, the former mainly based on the similarity analyses and the latter mostly involving docking simulations. Recent studies suggested the fruitful opportunity of combining both methods even though robust comparisons of such methods as well as of their possible combinations suffer from a lack of reliable datasets large enough to allow statistically significant analyses[148].

Based on these premises and considering that TRPM8 is well suited to high throughput screening (HTS) campaign using a calcium sensitive fluorescence dye approach[149,150], the present study started from a HTS campaign on a highly diverse corporate library of 124,107 compounds, described later. With the HTS results in hand, a second part of the study exploited such an invaluable wealth of biological data to compare ligand-based and structure-based virtual screening approaches. In detail, the primary objective of the computational study involved the evaluation of the predictive power of the SMARTS-based pharmacophore mapping compared to well-established docking simulations which are permitted by the previously described development of a TRPM8 homology model.

3.2.2 Library generation.

The compounds collection used for the first HTS campaign is composed by 124,107 synthetic small molecules and is a joined Dompé S.p.A. and Axxam S.p.A. corporate collections. For the assembly of the screening collection several computational tools were used: millions of commercially available compounds were screened *in silico*, applying a number of scientifically recognized filters to select a set of compounds that adhere to strict drug-like criteria: almost 90% of the library has zero violations of Lipinski's Rule of Five[151] for oral drug-likeness and more than 85% has zero violations of Oprea's parameters for lead-likeness[152]. Then, several other filters were applied to

minimize the number of known aggregators[153,154], false positives[155], toxicophores[156], and compounds containing reactive functional groups[157]. Particular attention was given to the choice of compounds with calculated favorable physicochemical and pharmacokinetic (ADMET) parameters, resulting in almost 80% of the library having predicted oral drug-like properties and no major liabilities[158,159,160,161]. The selection of compounds with a relatively high degree of novelty was performed by application of a modified version of the framework analysis proposed by Murcko[162]. Following a hierarchical fragmentation schema, progressive levels of structure abstractions were used to classify compounds according to their scaffolds or frameworks. Compounds were then clustered by framework or scaffold and compared to similarly clustered compounds currently at various stages of development, including launched drugs. The library was assembled in such a way that frameworks and scaffolds not represented in development compounds make up 63-71% of the library, respectively. Cluster analysis of these structural descriptors was also applied to ensure high chemical diversity of the screening collection.

3.2.3 Initial High Throughput Screening.

The first part of the study involved the HTS analysis of a highly diverse corporate library performed by Axxam s.p.a. and aimed to identify novel TRPM8 antagonists. In Primary Screening, all the 124,107 compounds were tested once at 10 μM. Valid results were produced for all of the 376 compound plates.

The screening was evaluated for each plate both in compound addition phase (CA, when the compound is injected onto the cells) and subsequent target activation phase (TA, when the reference agonist is added at EC_{80} concentration to open the channel). Compounds active in CA (potential agonists) and compounds which exhibited an unusual and fast rise of the fluorescent signal upon addition (potentially auto-fluorescent) were flagged and excluded from further characterization. The threshold for the primary screening hit selection was calculated according to the distribution of the percentage inhibition of all the compounds that was centered around 8.9% with a standard deviation of 14.8%. The threshold was then set at 53% corresponding to [percent inhibition mean]+[3*percent inhibition standard deviation].

Such a primary screening allowed the identification of 2916 compounds, corresponding to a hit rate equal to 2.4 %. The selected compounds underwent Hit Confirmation phase in which they were tested at 3 concentrations (10μM, 3μM, 1.0μM), with triplicate data points.

Hence, based upon the performance at 1μM, 3μM and 10μM, a response score was computed (i.e. a weighted score averaging the percentage inhibition information for a compound at the three tested concentrations) for each compound which was used to rank the Hit Confirmation data. In this way, 1739 compounds were selected for Activity Determination (IC_{50} calculation), as they showed a response score > 50. The percentage of the here selected compounds (1.40 %) thus represents the random percentage, namely the percentage of confirmed hits found in the library. This random value will be used in the following virtual screening analyses to evaluate their enrichment factors. During Activity Determination, a full dose-response curve was performed at 8 concentrations in triplicates. Remarkably, 49 compounds were identified as strong blockers, showing sub-micromolar IC_{50} values.

While avoiding a systematic analysis of all confirmed hits (apart from the naphthyl analogues described below), Figure 3.2.1 compares the distribution for some key physicochemical properties for confirmed hits (blue bars) and for the complete library (red bars). The first two analyzed properties suggest that, on average, the confirmed hits are larger molecules as demonstrated by the fact that about 50% of the confirmed hits have a molecular weight greater than 350 gr/mol. This is reflected in the number of rotatable bonds even though the highest percentage corresponds in both sets to molecules with 4 rotors. The last two compared properties show that the confirmed hits are on average apolar molecules and this trend is clearly evidenced by both polar areas where the most populated bin corresponds to molecules with PSA around 50 Å^2 and lipophilicity values where the highest percentage corresponds to molecules with log P around 4. The precise balance between an intermediate PSA value and a good lipophilicity emphasizes the key role of some H-bonds (as confirmed by the pharmacophore model, see below) while accounting for the transmembrane position the TRPM8 binding site.
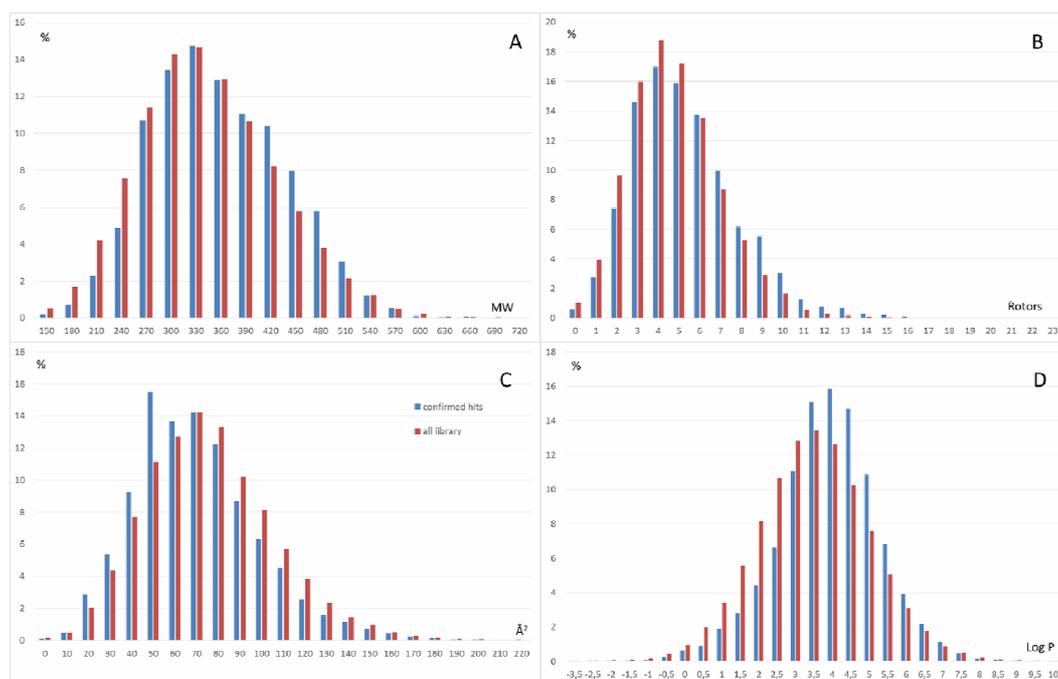
Figure 3.2.1 Comparison for some key ligand descriptors between the confirmed hits (blue bars) and all screened library (red bars). In detail, the comparison involves: (A) Molecular Weight; (B) Number of rotatable bonds; (C) Polar surface area computed as described in 13; and (D) logP (14).

3.2.4 Naphthyl derivatives.

Among the TRPM8 antagonists identified by this HTS campaign, a set of five naphthyl derivatives (**1-5**, as compiled in Table 3.2.1) were selected and pharmacologically characterized as interesting for drug-likeness, chemical tractability and potency since it includes the most potent blocker (**1**). In particular, the monitored set includes two naphthyl derivatives the confirmed inhibition percentage of which is greater than 50% (**1-2**) plus three inactive compounds (**3-5**) included in this analysis to offer more complete structure-activity relationships. With a view to better rationalizing their potency, the naphthyl compounds underwent docking simulations on the hTRPM8 model. In detail, Figure 3.2.2 shows the putative complex for compound **1** within the TRPM8 binding site evidencing the key role played by the ionic interaction between the ligand's ammonium head and Asp802, a role which is further substantiated by the observation that all reported amido derivatives were found to be markedly less potent. The naphthyl moiety is seen to engage the pocket normally

occupied by menthol where it can elicit $\pi-\pi$ stacking with Phe809 and Tyr745 which in turn stabilizes a H-bond with the ligand ether function. The heteroaromatic ring stabilizes H-bonds plus $\pi-\pi$ stacking with the amido group of Asn799, a kind of interaction already described for a recently reported set of isoxazole agonists[163]. Taken together, docking results emphasize that an antagonist has to engage (and constrain) at the same time Tyr745 and Asp802 thus interfering with the activation mechanism triggered by these two residues.
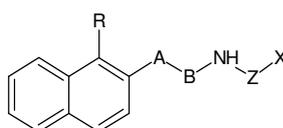
Cpd	R	A	B	Z	X	% inh at 10 μM	pIC$_{50}$
1	Cl	-O-CH$_2$	CH$_2$	CH$_2$	2-furanyl	101.11	7.15
2	F	-O-CH$_2$	CH$_2$	CH$_2$	2-furanyl	98.18	6.54
3	H	CH$_2$-O-	CH$_2$	C=O	3-pyridyl	7.92	< 5
4	H	CH$_2$-O-	CH$_2$	C=O	5-bromine 3-pyridyl	0	< 5
5	H	-O-CH$_2$	C=O	CH$_2$	3-pyridyl	0	< 5

Table 3.2.1. Structure and TRPM8 inhibition values for the naphthyl derivatives identified by the HTS campaign.
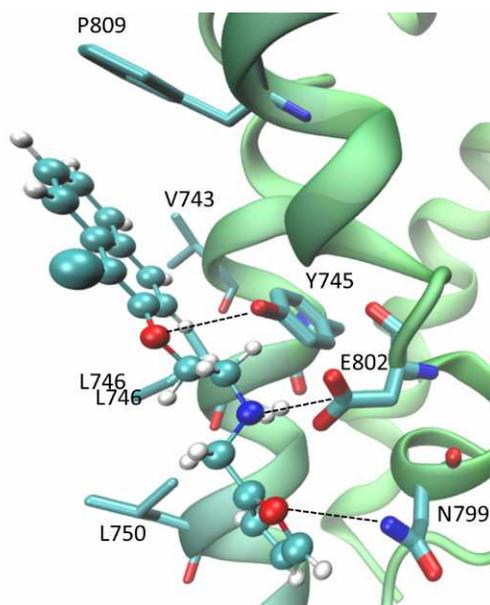
*Figure 3.2.2 Main interaction stabilizing the putative complex between the hTRPM8 homology model and the here reported compound **1**.*

3.2.5 Pharmacological characterization of compound 1.

Among the above described naphthyl analogues, compound 1, the most potent representative of the class, showed a good potency (middle nM range) regardless of the agonist used (pIC_{50} value of 7.l5 with Cooling Agent 10, 7.17 with Icilin, 6.7 with WS12). Moreover and interestingly, compound 1 showed an excellent selectivity profile towards other close members of the TRP channels family, since it is completely inactive on TRPA1, TRPV1 and TRPV4.

To further validate its TRPM8 blocking behavior, 1 was tested in the presence of a TRPM8 physiological stimulus, e.g. by giving cold temperatures to the cells (Figure 3.2.3). After incubation with increasing concentrations of 1, HEK-293/hTRPM8 cells were exposed to a temperature drop by using the ABI Prism® 7900HT instrument (from 25°C to 14°C) and the Ca^{2+} influx was recorded by a Ca^{2+}-sensitive fluorescent dye. Calculated pIC_{50} was 6.76, in the same range as the ones obtained with TRPM8 agonists (Figure 3.2.3a).

As second orthogonal assay, TRPM8 inhibition was evaluated by manual patch-clamp. Outward currents were elicited upon addition of Cooling Agent 10 (50 µM) at +40 mV, in the presence and in the absence of 1 (300 nM). When the compound was perfused, outward current was almost fully

blocked, suggesting a decrease of hTRPM8 activity (Figure 3.2.3b). After compound 1 washout, hTRPM8 activity was restored.

In order to investigate its binding mode, TRPM8 blockade was finally assessed on a construct carrying the mutation Asp802Ala. Asp802 has been identified as a key residue required for activation by icilin in mammalian TRPM8 while this amino acid is not fundamental for stimulation by menthol[164]. As shown in Figure 3.2.3c compound 1 pIC_{50} was decreased by the mutation in the icilin binding site (5.2 instead of 7.2 on wt TRPM8). This finding confirms what was evidenced by docking results, namely that Asp802 is important for compound 1 binding and blocking behavior and gives a hint as to its mechanisms of action: compound 1 could block TRPM8 activation by competing with agonists such as menthol in disrupting the interaction between Tyr745 and Asp802.
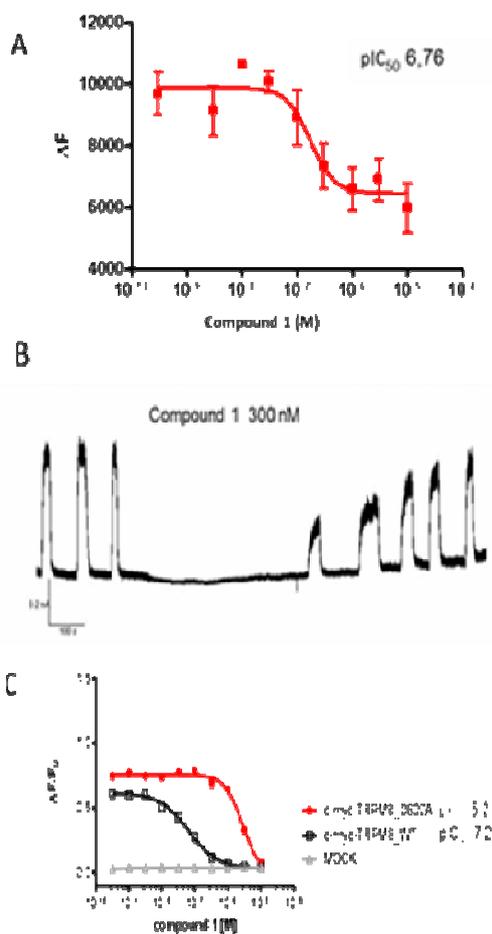
Figure 3.2.3. Validation of TRPM8 antagonist compound 1

by using cold temperatures as stimulus (3A), by manual patch clamp (3B)

and by site-directed mutagenesis (3C).

3.2.6 *In silico* virtual screening: setting the scene.

The second part of the study had the primary objective of comparing a docking-based approach with SMARTS-based pharmacophore filters. While commonly used in chemical data mining, SMARTS strings have been scantly exploited to encode pharmacophore models although the combined use of more strings should allow the selection of chemically diverse molecules which however satisfy well-defined pharmacophoric rules[165].

Hence, the computational studies began with the development of a reliable pharmacophore model based on a set of patented TRPM8 antagonists. The key elements were transformed into a set of SMARTS strings to filter the already experimentally screened corporate library. To further emphasize the potentialities of SMARTS-based filtering, the same set of patented TRPM8 antagonists was used to screen the corporate library based on 2D molecular similarity analysis as expressed by ECFP6 descriptors and Tanimoto distances. Finally, a set of known TRPM8 antagonists was used to optimize docking procedures which were then applied to the screening of the same corporate library. The two utilized approaches were finally compared including also their mutual combination. The here reported ligand-based studies were performed by the industrial partners.

3.2.7 Pharmacophore development, SMARTS-based screening and similarity analyses.

As recently reviewed, several TRPM8 antagonists have been reported mostly in patent literature. While showing a marked structural variability, they share some key common features (as schematized in Figure.3.2.1), since most antagonists possess two cyclic (usually but not necessarily aromatic) systems connected by a variable linker suitably decorated by H-bonding groups. Although not necessary for activity, several linkers are characterized by a central basic function. As already described in Figure 3.2.2 and confirmed by mutational analyses, docking simulations revealed that a TRPM8 antagonist should contact both Tyr745 and Asp802, hence the beneficial role of a suitably spaced hydrogen bond acceptor (HBA) which should interact with Tyr745 and a hydrogen bond donor system (HBD) which should interact with Asp802. The first cyclic system should mimic the menthol carbon skeleton, while the second aromatic ring should approach Asn799. In both cases, the interactions are strengthened by apolar contacts with the surrounding hydrophobic residues. This pharmacophore is also characteristic for some agonists such as D-3263 developed by Dendreon now in phase I clinical trial. The so obtained general scheme was then

converted into a set of ten SMARTS strings which encode for the variability in length and flexibility of the linker as well as for the size and complexity of the ring moieties (Table 3.2.2).

SMARTS	Total Ratio	SMARTS Ratio	EF	SMARTS Positive	SMARTS Negative	k value
1	0.014	0.4	28.57	2	3	0,391
2	0.014	0.132	9.43	177	1164	0,120
3	0.014	0.127	9.04	10	69	0,114
4	0.014	0.104	7.41	11	95	0,091
5	0.014	0.093	6.66	11	107	0,080
6	0.014	0.088	6.31	25	258	0,075
7	0.014	0.08	5.71	24	276	0,067
8	0.014	0.079	5.67	17	197	0,066
9	0.014	0.063	4.48	14	209	0,049
10	0.014	0.052	3.75	275	4961	0,039

Table 3.2.2 Enrichment Factors and k^{166} values for the 10 SMARTS strings related to the pharmacophoric model.

The SMARTS-based pharmacophore mapping was performed using a small set of 12 known TRPM8 antagonists (Figure 3.2.4) which allowed a general motif formed by a hydrogen bond acceptor system (HBAcc) and a hydrogen bond donor system (HBDon) to be revealed. The two H-bonding groups can be spaced by two atoms, as in the case of 12, 5, 6 and 9, or one atom in the case of 7 and 8, or they can be spaced by more flexible chains such as in 2 and 3. Considering the constrained systems in 1, it has been argued that the *cis* orientation is preferred and this is consistent with the proposed binding mode for TRPM8 inhibitors, on the basis of the derived 3D homology models of the ion channel (Figure 3.2.2). All the three classes of selected chemotypes can be

schematized by a pharmacophore model which define the relevant common chemical pattern characterizing these active compounds (Figure 3.2.5)

1 (AMGEN) Prous ID: 726494	2 (BAYER AMTB) Prous ID: 445976	3 (BAYER) Prous ID: 424573
4 (BAYER) Prous ID: 424587	5 (DENDREON) Prous ID: 399718	6 (DENDREON) Prous ID: 399722
7 (GLENMARK) Prous ID: 689868	8 (ITALIAN CNR) Prous ID: 696404	9 (JANSSEN JNJ41876666) Prous ID: 706061
10 (JANSSEN) Prous ID: 688352	11 (JANSSEN) Prous ID: 688356	12 (JANSSEN) Prous ID: 703270

Figure 3.2.4. Known TRPM8 antagonists used to develop the pharmacophore model.

The H-bond acceptor group is represented by direct and reverse sulphonamide, ether, ester, ketone and imine; the H-bond donor group is represented by amide, amine and alcohols. The spacers are represented by a carbon atom, and are restricted to some rules: between the H-bond acceptor and

the H-bond donor (SPACER 1), is allowed a linear distance of 2 or 3 atom, while more than 3 atoms are allowed only in very flexible moieties due to a possible tridimensional arrangement; between the H-bond acceptor and the last ring (SPACER 2), is allowed a linear distance of 0 or 1 atom. This pattern which condenses the structure-activity relationships among the active hits can be encoded by 10 SMARTS strings. An example is shown below:

*[*R]~[#16X4&$([SX4](=[OX1])(=[OX1])[NX3;H1]),#7X3&$([NX3][SX4](=[OX1])(=[OX1])),#8X2&$([OX2][CX3](=[OX1])),$([NX2]),$(*[NX3;H1]),#6X3&$([CX3](=[OX1])[OX2]),#6X3&$(*=O),#8,#16,#7&H0]~[#6!R]~[#6!R]~[#6&$(*-[OX2;H1]),#6&$(*=[NX2;H1]),#6&$(*-[SX2;H1]),#7X3&H1,#16X4&$([SX4](=[OX1])(=[OX1])[NX3;H1]),#7X3&$([NX3][SX4](=[OX1])(=[OX1])),#8X2&$([OX2][CX3](=[OX1])),#6X3&$([CX3](=[OX1])[OX2]),#6&$(*-[NX3;H2])]~[#6!R]~[*R]*

One compound is labelled as active if its percent inhibition value is greater than 50% .
The Enrichment Factor (EF) obtained using the strings were calculated by the following equation:

$$EF = \frac{SMARTS\ ratio}{TOTAL\ ratio}$$

Where:

$$SMARTS\ ratio = \frac{SMARTS\ actives}{SMARTS\ actives + SMARTS\ inactives}$$

And:

$$TOTAL\ ratio = \frac{TOTAL\ actives}{TOTAL\ actives + TOTAL\ inactives}$$

A second part of SMARTS-based filtering was then exploited to discard potentially unselective molecules. In detail, a collection of published compounds (MDDR Database) was analyzed in order to identify moieties directly connected to the reported mechanisms of action. With the same rationale used for the definition of the pharmacophore model, a total of 37 SMARTS strings related to the unselective moieties were defined (7 representative SMARTS can be found in Table 3.2.3). Once the optimal set of SMARTS string was defined, they were used to screen the corporate library

with a view to identify TRMP8 blockers while discarding unselective or undesired scaffolds. As reported in Table 3.2.4, this analysis allowed the selection of 6400 molecules including 479 confirmed hits (i.e. with a Response Score > 50). Hence, the percentage of confirmed hits is here equal to 7.48%, thus giving a noteworthy enrichment factor (EF) of 5.35 compared to experimentally determined random hit rate. The performance of each SMARTS string is reported in Table 3.2.2; notably, even the worst performing filter shows a EF value markedly greater than 1, and for all strings the k value is greater than 0 thus confirming that the obtained results cannot be due to a randomly correct prediction. Table 3.2.4 shows also the analysis of the chemical variability of the so selected set. Indeed, the selected set shows a ratio of decorated scaffolds per active compound which, albeit less than that observed in the full library, is in line with that obtained by docking simulations (see below) and emphasizes that suitably designed SMARTS strings can yield reasonably diverse results which are not heavily constrained by the exploited pharmacophore rules.

RING 1	HBA	SPACER 1	HBD	SPACER 2	RING 2

Figure 3.2.5 Schematic representation of the developed pharmacophore model.

SMARTS Strings	Mechanism of Action
[#7,#8](-[#6](-[#6]-[#6](-[#7,#8]-[*R,#6!R])=[#8])=[#8])-[*R,#6!R]	Antineoplastic
[c,n]1[c,n][c,n]([c,n][c,n]1)-[#6]-[#6](=[#8])-[#7]-[*R,#6!R]	Antineoplastic
[#16](=[#8])(=[#8])(-[#7]-[#6](-[#6]-[#7]-[*R,#6!R])=[#8])-[*R,#6!R]	Anticoagulant
[#7&R0](-[#6&R0](-[#6&R0](-[#7&R0]-[*R,#6!R])-[*R])=[#8])-[*R,#6!R]	Anticoagulant
[#6](-[#6](-[#7]-[*R,#6!R])=[#8])(-[#7]-[*R,#6!R])=[#8]	Antineoplastic
[#6](-[*R])(=[#8])-[#7]-[#6](-[#6](-[#6]-[#7]-[*R,#6!R])-[#8])-[#6]-[*R]	Nootropic Agent
[*R,#6!R]-[#7!R]-[#6!R]-[#6!R](-[#7!R]-[*R,#6!R])=[#8]	Anticoagulant

Table 3.2.3 A collection of published compounds (MDDR Database) was analyzed in order to identify moieties directly connected to mechanisms of action. With the same rational used for the definition of pharmacophore strings, a selection of 7 SMARTS strings were presented. Through a sub-structural search, they were used to eliminate cross-selective compounds from the library.

Method	Cpds	Actives	Hit rate	EF	Active scaffolds
Full library	124107	1739	1.40	1.00	1054 (0.61)
Similarity-based	6400	349	5.39	3.85	202 (0.57)
Ligand-based	6400	479	7.48	5.34	238 (0.49)
Structure-based	6400	677	10.57	7.55	282 (0.42)
Consensus	1075	126	11.72	8.37	37 (0.29)

Table 3.2.4 Major results for the here performed virtual screening campaigns.

The satisfactory flexibility of the utilized SMARTS string is clearly emphasized by the comparison with the results obtained by a simple 2D similarity analysis as encoded by ECFP6 descriptors and Tanimoto distances. The similarity screen was repeated considering each of the 12 known TRPM8 antagonists (Figure 3.2.4) already utilized for pharmacophore development and the compounds were ranked by using a consensus function by which a given compound is ranked according to its maximum so computed Tanimoto distance. In other words, a compound is evaluated by considering its capacity to approximate at least one of the 12 known antagonists. Table 3.2.4 reports the obtained results and reveals that such a similarity search placed among the 6400 top-ranked compounds 349 confirmed hits which corresponds to a enrichment factor markedly lower than that obtained by SMARTS strings. This result can be explained by considering that such a similarity search does not include any kind of molecular flexibility and so it can find only those compounds which are similar to the 12 known antagonists. This consideration is confirmed by the lowest number of active scaffolds as reported in Table 3.2.4.

3.2.8 Structure-based virtual screening: setting strategy.

A similar strategy based on already reported TRPM8 inhibitors allowed the best docking strategy for screening on the TRPM8 receptor structure to be designed.

Following a rationale similar to that utilized in the ligand-based approach, structure-based screening started collecting a benchmarking set composed of 50 published TRPM8 antagonists and 4950 decoy compounds taken from the inactive compounds of the screened library. With a view to identifying the most productive strategy, docking simulations involved the already published human TRPM8 homology model and were performed using AutoDock 4.0, Vina and LiGen.

In detail, both benchmarking set and screened library were prepared by using the already mentioned automatic script of the VEGAZZ suite of programs.

With regard to the AutoDock simulations, all calculations were performed by using the parallelized Gridock tool[167] and had the following characteristics: (i) the search was focused on a 10 Å radius sphere around Asp802 so as to encompass the entire binding cavity; (ii) the resolution of the grid box had a spacing of 0.375 Å; (iii) each compound was docked into the grid by the Lamarckian algorithm; (iv) the flexible bonds of the ligand were automatically recognized and left free to rotate; (v) the genetic-based algorithm ran 10 simulations per substrate with a maximum number of generations of 27,000; (vi) the crossover rate was increased to 0.8, and the number of individuals in each population to 150. Docking simulations were repeated by carrying out both rapid (100,000 energy evaluations per ligand and per generation) and exhaustive (2,000,000 energy evaluations) analyses and all other parameters were left at the AutoDock default settings[168]. Vina open-source program utilizes an iterated local search global optimizer algorithm which involves a sequence of steps consisting of a mutation and a local optimization and each step can be accepted or discarded according to the Metropolis criterion[169]. The box parameters were the same already utilized by AutoDock program. Before performing docking simulations, LiGen requires the development of a suitable pharmacophore model by using LiGenPocket. To this end, a rather large and very potent TRPM8 antagonist (a benzyloxy-phenylmethylcarbamate derivative, $IC_{50} = 0.45$ nM) not included in the benchmarking set was utilized with a view to avoiding bootstrapping situations and developing the pharmacophore as exhaustive as possible. Docking simulations were then carried out adopting the best performing set of parameters as derived by previous studies and focusing the search on a 10 Å radius sphere around Asp802. The ligand was considered as flexible and the results were ranked by the LigenScore function.

Table 3.2.5 compares the computed enrichment factors as derived by using AutoDock, Vina and LiGen and considering the top 10%. Overall, the reported EF values reveal a rather low predictive power of all considered scoring functions except for LiGenScore, the optimization of which will be described below. With regard to AutoDock results, Table 3.2.5 shows that the EF values obtained by exhaustive docking are slightly better than those derived by fast docking, a clearly expected result which however cannot compensate the hugely greater computational time required by the exhaustive search. Similarly, Vina score shows no predictive power. Conversely and even adopting the default parameters, LiGenScore reveals an encouraging predictive power ranking about one half of the active molecules[170] in the top 10%.

Based on these results, the LiGen simulations were then optimized by tuning the arrangement and the extension of the box where the docking search is performed, whereas the docking parameters of the program were not adjusted since they were already optimized by experimental design to optimize virtual screening performances and a further specific optimization can result in a unproductive overfitting. Hence, the LiGen simulations were repeated by varying the radius of the considered sphere (i.e. equal to 12, 10, 8, 6.5 Å) and its center (around Asp802 or Tyr745) so as to take into account also a possible adjacent sub-pockets. Table 3.2.6 reports the remarkable enrichment factors as obtained by the best performing LiGen simulation which involved a 8 Å radius sphere around Asp802 and such an optimized protocol was then utilized to analyze the screened library.

Score Function	Enrichment Factor Fast docking	Enrichment Factor Exhaustive docking
AD_Intermolecular	1.01	1.31
AD_Electrostatic	1.02	1.12
AD_VdW-Hbond-Desolv	1.03	1.30
Vina	1,01	--
Ligen	4.91	--

Table 3.2.5 Enrichment factors as derived by using AutoDock,

Vina and LiGen and considering the top 10%.

top % N	Cpds	Active	EF
1	50	18	36
2	100	27	27
5	250	38	15.2
10	500	41	8.2
20	1000	43	4.3

Table 3.2.6 Enrichment factors as derived by using LiGen
and focusing the search in 8 Å radius sphere around Asp802.

3.2.9 Structure-based virtual screening: results.

Such an optimized docking protocol (based on LIGEN software and using the LiGenScore consensus scoring algorithm) was exploited to virtually screen the corporate library and the top-ranked 6400 compounds (a number of compounds similar to that obtained by SMART-based filtering) were selected and analyzed. As evidenced in Table 3.2.4, docking-based screening identified 677 confirmed hits with an enrichment factor (7.55) greater than that achieved by SMARTS strings. Interestingly and contrary to what was expected, the docking-selected library shows a chemical variability as represented by a ratio of decorated scaffolds per active compound which is comparable to (and even slightly less than) that obtained using SMARTS strings (0.42 vs. 0.49).

Among the chemotypes identified by the HTS campaign, the study describes the pharmacological characterization of a set of naphthyl TRPM8 antagonists. Besides including the most active compound found within the corporate library, the naphthyl series was selected for its chemical novelty, its optimal lipophilicity profile while possessing ionizable and/or polar moieties, and because the compounds showed an activity in the nanomolar range combined with a remarkable selectivity since they are completely inactive against closely related TRP members. In fact, most of the available TRPM8 antagonists suffer from an unsuitable selectivity profile which hampers their clinical applications and indeed to date only the phosphonate analogue JNJ-39267631 has advanced

to clinical trials. The here reported naphthyl analogues might show a proper balance between potency and selectivity thus reducing unfavorable off-target effects and representing valuable candidates for clinical studies. From a mechanistic standpoint, mutational studies are in line with docking simulations and confirm the key role of Asp802. This finding is in agreement with the previously hypothesized activation mechanism and emphasizes that an antagonist has to interfere with the interaction between Tyr745 and Asp802 to induce the TRPM8 blockade.

With regard to *in silico* screening campaigns, the SMARTS-based results evidence the remarkable potentialities of this approach which gives productive results with very reduced computational costs (16 minutes for virtual screening 1 million compounds on a single CPU). Moreover, the reported method emphasized the fruitful possibility of combining pharmacophore SMARTS strings with other strings encoding for unselective (anti-targets SMARTS strings can be easily identified from annotated databases such us MDDR) or undesired scaffolds as well as strings to screen the compounds according to the Lipinski rule or other ADME/Tox filters. In this way, it is possible to collect a library of SMARTS strings which allow a proper filtering of any database. This combination of filters may reduce the number of hits but greatly enhances the quality of the retrieved compounds, because the final goal of a screening campaign is the identification of valuable hits suitable for progress to the lead optimization phase. The potentiality of suitably combined SMARTS strings is clearly documented by the comparison with the similarity analysis based on the EPCF6 descriptors which, despite being based on the same set of known TRPM8 blockers, give poorer results in terms of both enrichment factor and active scaffold variability.

Although supported by similarity analyses of known TRPM8 antagonists, the pharmacophore mapping involves a subjective definition of the pharmacophoric elements whose arbitrariness is however minimized by the possibility of exploiting many SMARTS strings so as to cover all conceivable combinations of the key elements. Although one of the major concerns with the SMARTS strings is the poor flexibility of the so performed substructure searches, our results clearly indicate that the filtered subset contains the same structural variability observed in docking results. It should be underlined that the chemical variability obtained by SMARTS strings can be modulated by varying the number of utilized strings, while that observed in the docking-based results is unavoidably constrained by the shape and size of the explored binding site and this limit is here further exacerbated by the fact that computed poses are evaluated without any complex optimization to reduce the computational costs. Nevertheless, docking-based screening provides a remarkable enrichment factor in line with the preliminary benchmarking tests. Collectively, the

remarkable results of docking-based screening represent a relevant validation of both the TRPM8 homology model and the used docking strategy.

The obtained screening results allow the positive effects of combining ligand-based and structure-based methods to be investigated. In detail, the two generated subsets share 1075 common compounds which represent about 1/6 of the entire subsets thus suggesting that SMARTS strings and docking simulations select rather different chemical spaces, a result which can be explained considering the different criteria by which the two methods collect their subsets. Among the shared compounds, there are 126 confirmed hits thus yielding an enhanced enrichment factor of 8.37. This means that the two methods, while covering rather diverse chemical space, can be successfully combined in a parallel approach which corresponds to selecting from among the compounds which fulfill the pharmacophore filters those achieving the most stable complexes. Remarkably, such an enhanced enrichment factor cannot be ascribed to the fact that it involves a restricted subset of 1,075 molecules. Indeed, the 1,075 top-ranked molecules as obtained by docking simulations include 113 confirmed hits with a percentage in line with that obtained for all 6,400 top-ranked molecules. In other words, the consensus approach allows the selection of a rather restricted number of compounds which benefits from ligand-based and structure-based approaches.

3.3 Isoxazole derivatives as potent TRPM8 agonists.

3.3.1 Introduction.

As mentioned before, both peripheral and central activation of TRPM8 induces analgesic effects that specifically reverse the sensitization of the behavioural reflexes elicited by peripheral nerve injury[171]. These effects are produced in a range of very low concentrations of topically applied TRPM8 activators, whereas high concentrations of menthol were found to cause both cold and mechanical hyperalgesia in healthy volunteers[172,173]. These findings suggest the potential relevance of TRPM8 activators as therapeutic strategy for pain treatment.

In this study[174], 12 isoxazole derivatives have been designed and tested for their pharmacological properties both in *F11* sensory neurons *in vitro* and in an *in vivo* model of cold allodynia.

The experiments revealed that, when compared to menthol, some newly-synthesized compounds were up to 200-fold more potent, though none of them showed an increased efficacy. Some isoxazole derivatives potentiated allodynic responses elicited by acetone when administered to rats subjected to sciatic nerve ligation; when compared to menthol, these compounds were efficacious at earlier (0-2 min) but not later (7-9 or 14-16 min) time points.

In order to investigate what lies under this behaviour, docking experiments have been performed on the hTRPM8 receptor model, revealing that newly-synthesized compounds might adopt two possible conformations, thereby allowing to distinguish "menthol-like" compounds (characterized by high efficacy/low potency), and "icilin-like" compounds (with high potency/low efficacy).

3.3.2 Compounds design and experimental/computational SAR analyses strategies.

Given the potential role of TRPM8 agonists for the treatment of neuropathic pain conditions, increasing efforts in the last few years have been dedicated to the design of selective and potent TRPM8 ligands. To date, the most active cooling compound, icilin (Scheme 3.3.1B), is characterized by the presence of a central tetrahydropyrimidine-2-one ring substituted by a nitrophenyl and phenol moieties. Several other analogues, based on the tetrahydropyrimidine-2-one moiety, have been described and patented for their cooling activity[175].

The need for a heterocyclic ring to elicit efficient interactions with TRPM8 is also witnessed by the isolation of the potent cooling agents alpha-ketoenamine (Scheme 3.3.1C and D) from roast malt extract[176]; however, these molecules are penalized by their great instability to oxidation, probably related to the reactive enamine functional group. A similar instability was also found for a class of aminophenols synthesized as hybrid derivatives of icilin and ketoenamines. Therefore, and in view of these stability issues, different heterocyclic moieties have been utilized as chemical scaffolds for the synthesis of new TRPM8 modulators; these include benzimidazole-based (Scheme 3.3.1E), as well as fused oxazole (Scheme 3.3.1F) and thiazole derivatives (Scheme 3.3.1G)[177,178].

In addition, aminoisoxazoles, in particular the 3-amino derivatives, are resistant to the typical degradation occurring to the isoxazole rings[179], and are suitable for simple and high-yield chemical derivatization of the amino group. Substituted aminoisoxazole derivatives seem to include the minimum structural requirements for a TRPM8 ligand, as seen for menthol or icilin, such as hydrogen bonding groups, a compact (mainly cyclic) hydrocarbon backbone, a correct hydrophobic/hydrophilic balance with a log P range between 1 and 5, and a molecular weight in the range of 150-350 g/mol. Indeed, several structural similarities can be observed between the alphaketoenamine derivatives and the aminoisoxazoles (Scheme 3.3.1H) [180].

Based on these evidence and on the known SARs for menthol (Scheme 3.3.1A) and its analogues[181], in this study 12 new isoxazole derivatives, carrying modifications on the isoxazole ring, have been synthesized. These mainly involved the introduction of aminoaliphatic or aminoaromatic chains in position 3 and/or the introduction of a methyl group in position 5 of this scaffold ring.
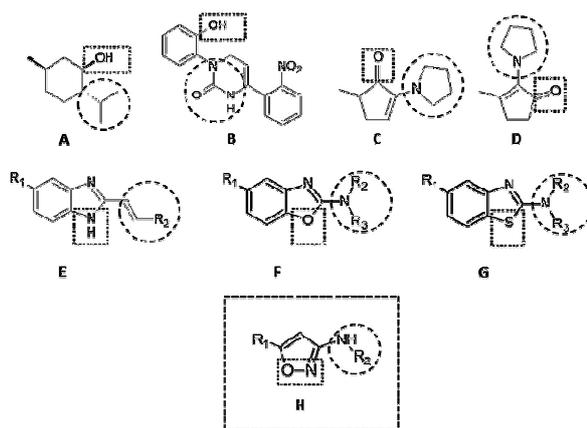
Scheme 3.3.1 (A-G) Molecular structures of main TRPM8 modulators: menthol (A), icilin (B), a-ketoenamine (C and D), benzo-fused imidazole (E), oxazole (F) and thiazole derivatives (G); generic structure of the newly synthesized compounds of the present study (H).

When tested for their ability to trigger TRPM8-induced $[Ca^{2+}]_i$ responses in sensory neurons *in vitro*, some of these derivatives showed higher potency when compared to menthol. Thereafter, the most potent compounds were tested *in vivo* in an animal model of cold allodynia, where they showed strong and rapid, although short-lasting, allodynic responses.

According to the previously resumed devices, the 12 new isoxazole derivatives synthesized bear substituents at position 3 and 5 of the heteroaromatic ring (Scheme 3.3.2).

In particular aminoaliphatic (1a-e) and aminoaromatic (1f-h) chains were introduced at position 3 for derivatives belonging to series 1, while a methyl group occupies position 5. By contrast, derivatives belonging to the series3 (3b and 3d) lack the methyl group in position 5. Finally, the derivatives 2b and 2d are isomers of 1b and 1d, due to inversion of the oxygen and nitrogen heteroatoms.

To characterize their pharmacological activity as TRPM8 agonists *in vitro*, these new compounds were tested for their ability to enhance $[Ca^{2+}]_i$ in differentiated F11 cells, as done for menthol (For procedure details see study reference). All tested molecules increased $[Ca^{2+}]_i$ in a concentration-dependent manner, although with different efficacy and potency.

None of the tested compounds showed a higher efficacy than menthol. The compounds referred as 1d and 1f showed the same efficacy than menthol; instead, the other ten tested compounds were less effective than menthol, showing the following rank order of efficacy: 1g>3b>3d>1h>1c>1e>1b>1a>2b>2d, suggesting that longer aliphatic chains or aromatic substituents at the position R2 are required to retain high efficacy.

By contrast, comparison of the potency among tested compounds was performed by normalization of the experimental data to the maximal value of fluorescence ratio prompted by each compound.

Eight of the tested compounds (i.e. 1a-e, 1h, 3b, 3d) showed a significantly lower EC_{50} when compared to menthol, suggesting that aliphatic substituents at the R2 and R1 positions are involved in conferring high potency to the tested compounds. The compound referred as 1g showed the same potency as menthol; finally, the other three tested compounds were less potent than menthol, with the following rank order of potency:2b>2d>1f, suggesting that the introduction of an aromatic ring at the position R2 or the inversion of the heteroatoms at the position X, Y (Scheme 3.3.2) have deleterious effects on compound potency.

As later detailed, the reported newly-synthesized molecules can be roughly subdivided into two groups: the first includes compounds characterized by submicromolar potency but with quite low efficacy (as seen for example in 1a-c derivatives), and the second includes compounds endowed with low potency and high efficacy (as exemplified by 1f and 1g).

The compound 1d appears to be the most promising compromise since it shows low micromolar potency combined with a very remarkable efficacy.
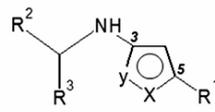
Compound	X	Y	R1	R2	R3
1a	O	N	CH₃	CH₃	CH₃
1b	O	N	CH₃	(CH₃)₂CH	H
1c	O	N	CH₃	(CH₃)₂CHCH₂	H
1d	O	N	CH₃	CH₃CH₂CH₂	H
1e	O	N	CH₃	(cyclohexyl)	H
1f	O	N	CH₃	(benzyl/phenyl)	H
1g	O	N	CH₃	(nitrophenyl)	H
1h	O	N	CH₃	(nitrophenyl)	H
2b	N	O	CH₃	(CH₃)₂CH	CH₃
2d	N	O	CH₃	CH₃CH₂CH₂	H
3b	O	N	H	(CH₃)₂CH	CH₃
3d	O	N	H	CH₃CH₂CH₂	H

Scheme 3.3.2 Representation of the synthesis procedure for all isoxazolylamine derivatives of the study and classification of the substituent groups present in each compound.

3.3.3 Computational methods.

The ligands were simulated considering their neutral forms since they are the favored ones at the physiological pH. After a preliminary optimization using the PM6 semi-empirical method, as implemented in MOPAC 2009, to discard high energy geometries and to calculate precise atomic charges, the conformational behavior of the compounds was investigated using VEGAZZ by a clustered MonteCarlo procedure which generated 1000 conformers by randomly rotating the rotors. All geometries obtained were stored and optimized to avoid high-energy rotamers. For each ligand, the so obtained lowest energy structure was then exploited in the docking simulations performed by AutoDock4.0 using the TRPM8 model recently obtained by homology techniques [23]. In detail, the calculations involved a single TRPM8 monomer and the grid box was set to include all residues within a 15 Å radius sphere around Tyr745, whose role in ligand recognition has been confirmed by mutational studies (ref), in order to comprise the entire binding cavity. The resolution of the grid was 68×77×100 points with a grid spacing of 0.450 Å. For docking simulations, the flexible bonds of the ligand were left free to rotate to account for ligand flexibility within the binding cavity. Each

substrate was docked by using the Lamarckian algorithm as implemented in AutoDock. The genetic-based algorithm ran 30 simulations per substrate with 2,000,000 energy evaluations and up to 27,000 generations. The crossover rate was increased to 0.8, and the number of individuals in each population to 150. All other parameters were left at the AutoDock default settings. The best complexes were finally minimized to favor the mutual adaptability between ligand and receptor and the optimized complexes were then used to re-calculate AutoDock docking scores, VEGA energy scores and X Score values. All mentioned minimizations were performed using the conjugated gradients algorithm as implemented in the Namd 2.51 package with the force-field CHARMm v22 and Gasteiger's atomic charges.

3.3.4 Computational results.

To identify the structural elements involved in the interactions of the here reported isoxazole derivatives with the TRPM8 channel, which might provide a plausible explanation for their different pharmacological properties, docking simulations were performed.

As shown in Figure 3.3.1A, the main interactions stabilizing the complex between TRPM8 and the most potent 1a derivative occur in a subpocket adjacent, but not coincident, with that of menthol, roughly corresponding to that occupied by the nitrophenyl moiety of icilin[182].

This result emphasizes the role of N799, which stabilizes an extended π-π stacking involving the carbamido function of the aminoacid side chain and the entire isoxazole ring. Furthermore, ligand's amino group is involved in an H-bond with the D802 carboxylate, a further key residue determining the activity of icilin, as also demonstrated by mutagenesis[183]. Finally, the interaction between ligand and the D802 seems to be essential for agonism as it destroys the H-bond between D802 and Y745, which characterizes the inactive state of TRPM8 channels, thus triggering its activation[184].

These two major contacts are further stabilized by a set of hydrophobic interactions with surrounding apolar side chains (e.g. L749, A753 and F794). Notably, the isopropyl group of the 1a derivative approaches Y745 contacting the apolar residues involved in the interactions with menthol, while the ligand's methyl group seems to act as a pivot which restrains the ligand pose within the binding site, thus maximizing its interactions (see below). Accordingly, docking results show that the desmethyl analogue, 3b, is characterized by a slightly different pose, in which the ligand slides towards F794 and loses its key H-bond with D802. Again, the isopropyl chain affords

the best potency presumably because it balances apolar contacts and steric hindrance, while bulkier groups (such as in 1b or in 1d) tend to clash against Y745 or L749. Interestingly, the isopropyl group corresponds to the substituent also inserted in a comparable pose by menthol, suggesting that this group best fits the TRPM8 cavity in that region.

By contrast, when the N-linked group is further enlarged, as seen in the1fehderivatives, the resulting ligands cannot assume the above described pose and show a completely different arrangement in which the isoxazole ring approaches Y745, while the bulky N-linked moiety approaches N799. Figure 3.3.1B displays the pose of the 1h derivative showing that the isoxazole ring is unable to elicit both significant π-π stacking and the key H-bond with Y745. The nitrophenyl moiety mimics the pose assumed by the same group of icilin stabilizing a key H-bond which involves the nitro function and the side chain of N799, while the amino group of 1hinteracts with D802. However, the inability to contact Y745 as well as the steric clashes exerted with surrounding residues could explain its poor potency, an effect which appears to be dramatically exacerbated when the ligand cannot elicit H-bonds with N799 (as seen for 1f).

Finally, docking simulations revealed that the inverted 3-methyl derivatives (2b and 2d) assume a different pose which could explain their poor potency. In detail, their isoxazole ring stabilizes an H-bond between the intra-annular oxygen atom and the N799 side chain, losing the extended π-π stacking and, more importantly, these compounds appear to be unable to elicit the pivotal contact with D802.
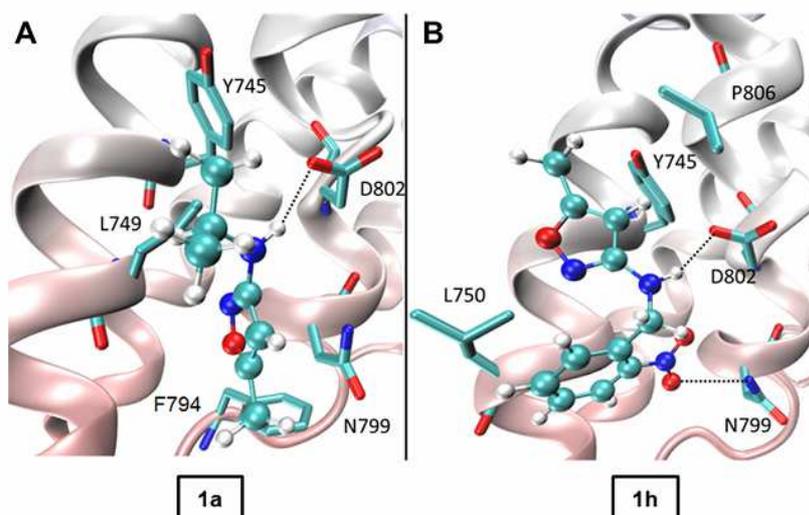
Figure 3.3.1 Magnification of the binding site in TRPM8 channel. (A, B) Main interactions stabilizing the complex between the TRPM8 binding site and 1a (A) or 1h (B). The cartoon for the transmembrane helices is colored according to their orientation within the membrane bi-layer (gray=extracellular side; pink=intracellular side).

3.3.5 Discussion.

When tested *in vitro*, the compounds containing modifications only at the 3-position, particularly a secondary amine linked to an isobutyl (3b) or a butyl (3d) group, displayed a slightly higher potency when compared to menthol, whereas their efficacy was slightly reduced. The introduction of a methyl group at the 5-position in both these compounds (generating 1b and 1d, respectively), lead to a further improvement of the potency, without altering the efficacy in the case of 1b or inducing a small, but significant efficacy increase in the case of 1d when compared to 3d.

Further modifications of the aliphatic chain connected with the secondary amine (consisting in an isopropyl or an isobutyl group in the case of 1a or 1c, respectively) prompted to similar improvement in the potency, but still not in the efficacy, when compared to the parent compounds (3b or 3d).

The increase in potency showed by the described compounds is instead prevented by the introduction of an aromatic function on the groups linked to the amine at the 3-position of the isoxazole ring. In fact, when a benzyl group was linked to the secondary amine (1f), a dramatic decrease in potency was observed; furthermore, when this high electron density group was depleted

through the introduction of the known deactivator e NO2 (at the ortho- or para-position, as in the 1h and 1g derivatives, respectively) or replaced by a cyclohexyl group (1e), the potency increased by about 20-fold.

In *in vivo* test, we observed that, similarly to menthol, the newly-synthesized 1a-d compounds also potentiated allodynic responses prompted by acetone exposure, suggesting that they also acted as TRPM8 agonists. However, when compared to menthol, 1a-d showed a distinct time-dependent profile. In fact, allodynic responses triggered by these compounds showed a shorter latency, being already evident after the first triggering stimulus was applied; in addition, 1a e 1d compounds ceased to elicit allodynic responses after the second or third acetone application. By contrast, menthol was only effective after the second acetone application, and the 1f compound failed to display allodynic effects at any time point after stimulus challenge.

These observation suggests the existence of a good correspondence between *in vitro* and *in vivo* data, with those compounds showing the highest *in vitro* potency as TRPM8 agonists being provided of faster onset and offset *in vivo* effects when compared to menthol, and the lower potency TRPM8 agonist showing no efficacy in *in vivo* testing.

About the *in silico* investigation, the results achieved in our docking experiments suggest valid structural hypothesis to rationalize the biological evidence. Indeed, the computed complexes suggest that the TRPM8 binding site can be roughly subdivided into two regions. The first region is more hydrophobic, and contains the Y745 residue playing a key role in TRPM8 activation[185]. Given the nature of the flaking residues, a ligand can occupy such a region by stabilizing a set of precise hydrophobic contacts plus key H-bonds with Y745, as seen for the prototypical example of menthol.

The second region is more polar as it is lined by N799 and D802 which can elicit reinforced H-bonds and extended π-π stacking with the more potent isoxazole derivatives. When considering these possible polar interactions, it comes as no surprise that suitable ligands which occupy this second region can afford markedly higher potencies compared to ligands which are positioned in the first sub-cavity where at most they stabilize H-bonds with Y745. Yet, the ability of molecules to interact with residues of the first region appears to be an essential requisite to achieve a good efficacy profile, regardless of the polar contacts stabilized within the second sub-pocket. Notably, icilin confirms that ligands able to suitably occupy both sub-pockets can successfully combine potency and efficacy even though such a double occupation appears to be vastly constrained by hampering steric clashes exerted by surrounding apolar residues.

Taken together, the results obtained emphasize that the newly synthesized compounds (particularly 1a-c) can induce TRPM8 activation *in vivo*. Although the TRPM8-dependent acetone test in CCI animals utilized in our study is a model of cold allodynia (indicative of an increased pain sensitivity), it should be reminded that the activation of TRPM8 channels can also mediate analgesic effects[186].

As a matter of fact, menthol itself, though acting as a reference compound for TRPM8 activators, has well-known analgesic properties[187]. Concentration-dependent effects of menthol at TRPM8 receptors might provide a plausible explanation for this apparently paradoxical behaviour; in fact, while, as indicated previously, elevated concentration of menthol trigger TRPM8-dependent nociceptive behaviour in CCI mice, lower menthol concentrations induced marked analgesia in the same rat model of neuropathic pain. However, whether additional variables such as changes in TRPM8 kinetics of activation and/or desensitization, differential treatment modalities (topical vs systemic administration), or model-dependent effects on TRPM8 expression levels at various sites contribute to the distinct pharmacological profile shown in vivo by menthol or other TRPM8 modulators remains to be investigated.

Despite all these uncertainties, the present results clearly suggest that presently-described aminoisoxazole-based derivatives display a pharmacological profile, both *in vitro* and *in vivo*, consistent with that of TRPM8 agonists. These molecules might be of considerable interest for further structural optimization and functional analysis, in order to validate their clinical utility as novel analgesics.

3.4 Accessory binding site research on TRPM8.

3.4.1 Introduction.

As reported so far, the TRPM8 homology model has shown to be structurally reliable and to perform well in terms of both ligand interaction, agonists and antagonists as well, and from a functional point of view. Anyway, the orthosteric binding site is not the only mechanism we have been interested in. The TRPM8 polymodal activation is still not completely understood as proven by the ever increasing interest by the worldwide scientific community; moreover, the interest in its allosteric modulation is becoming increasingly popular in the scientific community, as supported by the several reports appeared in literature.

For the here reported study, the rationale has been based on both literature data and in house results. As already mentioned, several antagonists, like BCTC, SKF96365 and 1,10-phenanthroline, have been proved to act as allosteric modulators, shifting the voltage curve that lies under the TRPM8 channel opening towards more positive potentials. In their mutagenesis study, Malkia and co-workers found that, substituting the key Tyr745, only the agonist activity of menthol was lost, while antagonist's desensitization was stably maintained[188]. This strongly suggests that not all the antagonists interact with the orthosteric site and thus TRPM8 should have at least one allosteric binding domain, by which antagonists like BCTC can change the physiological properties of the channel, moving the threshold for activation towards higher temperatures for native currents.

In a recent study, Taberner and co-workers investigated the allosteric modulation on TRPM8 through *in silico* calculation, first, and mutagenesis studies, later, finding the linker between the S6 of pore module and the TRP conserved box to be a target of the channel function modulation. In detail, stepwise substitutions of segments in the S6-TRP box linker of TRPM8 channel with the cognate TRPV1 channel sequences, recently resolved by cryo-electron microscopy images[189], produced functional chimeric channels, and identified Tyr981 as a central molecular determinant of the channel function. Additionally, mutations in the 986–990 region had a profound impact on channel gating by voltage and menthol, as evidenced by the modulation of the conductance-to-voltage (G-V) relationships.

Another example of allosteric antagonism came from the study of Miller and co-workers[190]. They proved the rabbit polyclonal antibody ACC-049 to be full antagonist of TRPM8 cold and icilin activation, by blocking the channel pore and avoiding the cations entrance. Species-specific

differences were found for menthol activation, supporting, again, that there are different critical residues/regions within the TRPM8 channel.

On the other hand, further motivation arose from some results obtained in my previous studies. Indeed, among the actives found through the previously described HTS study, a non negligible part of them (about 10%) were characterized by a quite hydrophilic profile, as supported by their log P values lesser than 2.5, which appears to be seemingly incompatible with the hydrophobic environment of the orthosteric intramembranous binding domain.

While considering that the difference of potency between the molecules used for setting up the screening strategy (nanomolar antagonists) and the screened ones (low micromolar antagonists) is presumably a key factor influencing the consistent decrease of the EF values as measured along the two screening studies (from around 32 to 8, one fourth), the role of a secondary binding domain, in which part of the active molecules can interact, might be a second factors to be to explain why docking simulations didn't rank most of tested antagonists in the top of the ranking.

Based on these observations, our interest has understandably been directed to the search of a secondary ligand binding domain, presumably located in the extracellular part of the channel, where hydrophilic ligands (or, more generally non Tyr745-depending antagonists) can be conveniently accommodated blocking the channel through a lid mechanism on the pore entrance.

In order to explore this possibility, a blind docking study on the whole surface of the tetrameric TRPM8 model has been performed. So obtained potential pockets have been selected and complexes analysed. The most promising area, structurally compatible with the hypothesized accessory domain, has been promoted to a mutagenesis experiment, whose results, unfortunately, did not support what was predicted *in silico*.

3.4.2 Blind docking strategy.

Although working on the single monomer would be much less time demanding, and, at a first glance, it could even be enough to explore most intra- and extra-cellular portions, especially for the sensor module, such an approximation could neglect suitable binding pockets located at the interface between two or more monomers of the channel. With a view to exploring all possible binding domains, the docking calculations were thus performed on the whole TRPM8 tetrameric structure.

To what concern the set of molecules used as probes and considering its early stage, the study was focused on known nanomolar antagonists, in order to reduce as much as possible false negative results, leaving the molecules from the HTS for the following steps.

So, a dataset of ten representative antagonists was generated, by collecting several potent AMTB derivatives, chosen due to their marked hydrophilicity, with some more hydrophobic antagonists known to be independent from Tyr745, such as BCTC, CTCP, and SKF96365.

The docking calculations were performed by using Biodock[191], an in house pure stochastic docking program, able to produce, evaluate and classify a high number of complexes between two interaction partners in a rapid and efficient way. Both the ligand and the protein were kept fixed. For each compound thousands complexes have been screened and the most stables were clustered, optimized and visually analysed.

Among the most populated and top-ranking clusters, only few had suitable features in terms of location, size and residues involved. In particular, two pockets stand out from the findings: a first one is located on the extracellular side of the sensor module while the second is located within the pore loop. So individuated regions were further characterized. Indeed, the obtained complexes were optimized by NAMD energy minimization, in order to refine the binding mode and the relative scores, and then the refined complexes were analyzed to highlight the most conserved patterns of interactions.

3.4.3 Sensor module putative pocket.

On the extra-cellular side of the sensor module domain, a calyx made by the two parallel loops (here loop 1 and loop 2), linking S1 to S2 and S3 to S4, respectively, hosted many molecules of the probe set, with remarkable interaction energies. The so arranged loops define a potential binding pocket with interesting chemical features.

Loop 1, on the right of the here reported schemes, exhibits mainly polar residues, in agreement with the hydrophilic nature of the extracellular environment, namely His720, Lys721 and Lys722. On the other flank of the calyx, loop 2 exposes several aromatic amino acids, like Tyr732, deeply in the pocket, and Phe832, more superficially. The so framed area can be seen as a triangular pyramid, oriented with the top toward the membrane, with height of about 7 Å and sides of 8, 8.5 and 9 Å, approximately. On these grounds, the maximum length of the hypothetic allosteric modulators should be not more than 14 Å, still sufficient to harbour all molecules used in this study.

As reported in the Scheme 3.4.1, the pharmacophore resulting from the pocket analyses should contain, substantially, three chemical substructures, which brings to mind those defined to develop the pharmacophore model for the TRPM8 antagonists (see paragraph 3.2.7) (*i*) an hydrophobic/aromatic moiety that should stabilize the complex through π-π stacking with the Phe832 side chain as well as through hydrophobic interaction with Leu825 aliphatic chain; (*ii*) an H-bond acceptor to form an Hydrogen bond with imidazole function of Hys720 (circled in red), flanked by protonated amines of Lysine residues, that could be able to drive partial negative head of ligands; (*iii*) and a second aromatic moiety, even better if substituted in order to form an H-bond with the hydroxyl function of the Tyr826 or the Thr732, intended to establish a second π-π stacking, so stabilizing the complex.
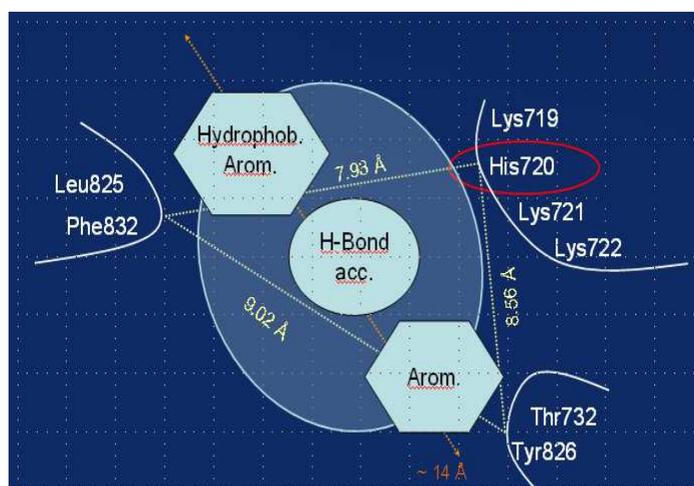
Scheme 3.4.1 Bi-dimensional representation of the putative
accessory TRPM8 ligand bind domain on sensor module
and related pharmacophore model. Putative crucial Histidine circled in red.

As example of satisfactory complexes between the putative accessory TRPM8 binding pocket and two representative antagonists, namely CTPC ((2R)-4-(3-chloro-2-pyridinyl)-2-methyl-N-[4-(trifluoromethyl)phenyl]-1-piperazinecarboxamide) and a diamine derivative taken from patent WO 2010010435 A2, are here described (Figures 3.4.1-2, respectively).

To what concern CTPC, the trifluoromethyl-phenyl moiety establishes π-π stacking with Tyr826, an interaction further stabilized by a H-bond between the hydroxyl function of Tyr826 and a fluorine atom. The piperazinecarboxamidic moiety establishes a well-conserved H-bond with His720, while the chloro-pyridyl tail stabilizes π-π stacking and apolar interactions with Phe735 and Leu825, further strengthened by a H-bond between pyridyl nitrogen atom and the Leu825 backbone group. Moreover, the complex shows a very favourable docking score which further supports how hydrophobic antagonists can be suitably accommodated whiten the hydrophilic environment of this putative pocket.

And indeed, docking results show that also the more hydrophilic benzimidazole derivatives are able to successfully interact with the accessory pocket. An example of this class is offered by the complex for diamine derivative. The ligand phenyl ring elicits π-π stacking with Tyr826 and His720, which also establishes the conserved H-bond with the ether oxygen, while the oxazol-pyridyl moiety and the charged amine establish H-bonds with Ser824 and Ser733, respectively.
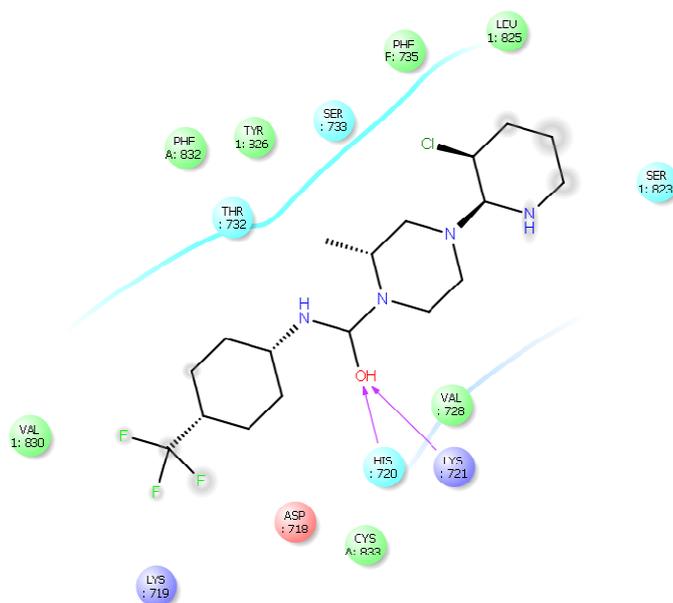
Figure 3.4.1Representative scheme of CTPC-putative sensor module binding domain interactions.
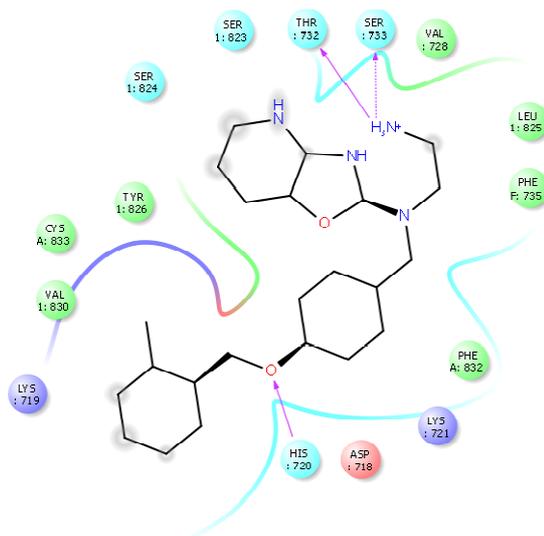
Figure 3.4.2 Representative scheme of
diamine derivative-putative sensor module binding domain interactions.

3.4.4 Pore module extracellular putative pocket.

At the entry of the channel, a second putative accessory binding domain was found (Scheme 3.4.2). Here, the four extracellular loops, linking S5 to S6 segments of each monomer, flank each other and delimit the pore. Here, the exposed residues appear to be mostly involved in ions detection and regulation.

The possibility that this region represents an allosteric modulation site is in line with many other known receptors, in which the antagonists bind the protein at the entry of the channel and physically blocks the ionic flow, acting as a cork, as proposed by Miller for the antibodies. Alternatively, the antagonist binding can lead to a refolding of the area, thus preventing the normal ionic flow.

In detail, the TRPM8 putative accessory binding site appears to be within the channel, a bit displaced in respect to the longitudinal axis of the pore. Here, three of the four monomeric intramembrane loops are mainly involved into framing the sub-pocket. Immediately deeper, going in the intracellular direction, pore module helices complete the pocket, by exposing both polar and apolar residues toward the pore. The putative site appears, then, to be straight, long and tight, allowing the ligand to enter maintaining an extended conformation.

The top of the putative binding site is mostly characterized by polar residues, in agreement with the extracellular environment. Here, Asn934, Lys937 and Pro938 of the monomer A, Thr930 and

Thr932 of monomer B, and Tyr905 and Asn934 of monomer C frame the extracellular side of this putative pocket. Deeper, monomer A and B mostly contribute to complete the site with several hydrophobic residues, such as Val941, Phe951, Ile957, Val960 and Pro907, Phe912 respectively.
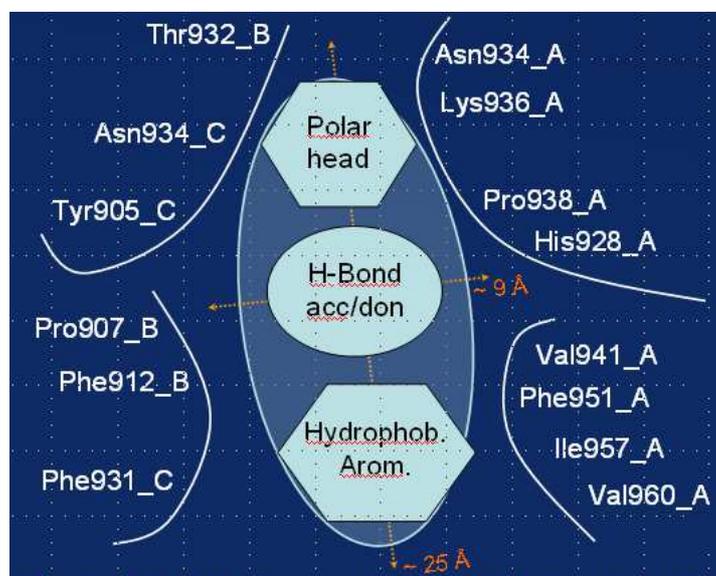
Scheme 3.4.2 Bi-dimensional representation of the putative
accessory TRPM8 ligand bind domain within the pore module
and related pharmacophoric model.

As for the previous docking calculations, the computed complexes between this putative accessory binding pocket and two representative ligands, namely BCTC and the same a diamine derivative previously shown are here reported (Figures 3.4.3-4, respectively).
BCTC enters the pocket with its t-butyl-phenyl ring, that seems to drive the molecule down to the pore attracted by the hydrophobic area at the bottom of the pocket. Here, both Phe931 and Phe951 of monomer A stabilize the complex through π-π stacking interactions with the aromatic ring, while t-butyl group elicits a set of apolar contacts with the surrounding alkyl side chains. Then, the urea moiety establishes two H-bonds with His928 on chain A and with Tyr905 on chain B, respectively. Finally, the chloro-pyridyl ring interacts with Pro938 on chain A, Ile904 on chain B and Phe931 on chain B via hydrophobic interactions, while the pose is further stabilized by a H-bond between the pyridyl nitrogen atom and the ammonium head of Lys937 on chain A.

Again, an example of a computed complex with a more polar diamine derivative is reported. Similarly to the BCTC binding mode, the hydrophobic methyl-phenyl head is inserted deeply into the pocket, attracted by hydrophobic/aromatic residues such as Phe951, Ile957, Val 960 and Met964. Next, the ether oxygen atom establishes a H-bond with His928, while the ammonium head stabilizes a ion-pair with Glu942 on chain A. On the pocket rim, the oxazol-pyridyl moiety contacts Ser902 and Lys937 on chain A, and elicits a π-π stacking interaction with Phe931 on chain C.
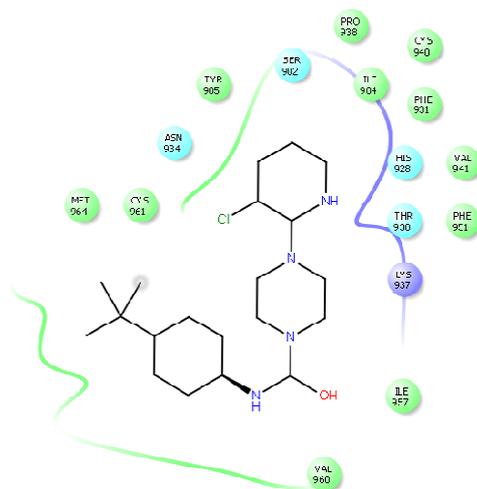
Figure 3.4.3 Representative scheme of BCTC-putative pore module binding domain interactions.
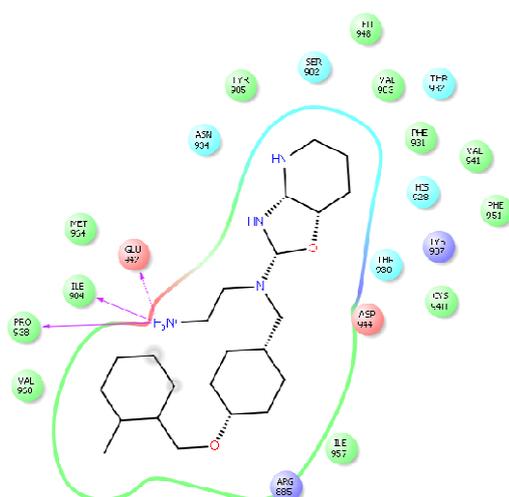
Figure 3.4.4 Representative scheme of
diamine derivative-putative pore module binding domain interactions.

3.4.5 Virtual screening validation.

In order to select the putative accessory pocket to be promoted to mutagenesis studies for an experimental validation, a virtual screening campaign has been performed on both putative binding domains: the best performing one, as evaluated in terms of enrichment factors and skewness values, will be preferred.

The database of active and decoy molecules was, then, collected by selecting 120 active molecules from the previous HTS study the log P value of which was lower than 2.5. To these molecules, 11880 decoy molecules, proved to be inactive in HTS, have been added in order to obtain a dataset containing 1% of active compounds. These decoys were selected according to their log P average value, as well as to their molecular weight and charge in order to avoid biasing conditions. Unfortunately, as happened for the entire HTS set of molecules used in the previously described virtual screening study, selected hydrophilic active molecules possess a lower potency compared to the molecules initially used as probe to detect the accessory binding pocket. This might reflect in a partial incapacity to correctly rank the 120 active compounds, but it would not represent a biasing problem since it will equally affect both putative pockets.

By considering the best performing procedure as evidenced in the previous *in silico* virtual screening study, also here docking simulations were performed by the Ligen docking program and using the previously described benzimidazole complexes as the pharmacophoric templates. Docking calculations were performed in a 8 Å radius sphere around the center of mass of the antagonists used as the templates. The LiGen score function was used to rank the resulting complexes.

To evaluate the performances of the two screening studies and besides the skewness values, the enrichment factors were calculated by considering the following arbitrarily chosen percentages of the ranking: top 1, 2, 5, 10 and 20 %. Results of the screenings on both the sensor module and the pore module putative pockets are reported in Table 3.4.1 and Table 3.4.2, and the histogram plots of the entire distributions are reported in Figure 3.4.5 and Figure 3.4.6, respectively.

n (top %N)	Cpds	Act	EF	Sk
1	120	4 (1)	4	
2	240	11 (2)	5,5	
5	600	24 (5)	4,8	2,01
10	1200	45 (10)	4,5	
20	2400	64 (20)	3,2	

Table 3.4.1 Sensor module-Virtual Screening evaluation.

"Cpds" reports the number of molecules hosted in the considered subset; "Act" stays for the actives compared to the expected one in parenthesis; EF reports the Enrichement factor values; "Sk" the Skewness value of the entire distribution.

n (top %N)	Cpds	Act	EF	Sk
1	120	1 (1)	1	
2	240	4 (2)	2	
5	600	12 (5)	2,4	0,78
10	1200	22 (10)	2,2	
20	2400	40 (20)	2	

Table 3.4.2 Pore module-Virtual Screening evaluation.

"Cpds" reports the number of molecules hosted in the considered subset; "Act" stays for the actives compared to the expected one in parenthesis; EF reports the Enrichement factor values; "Sk" the Skewness value of the entire distribution.
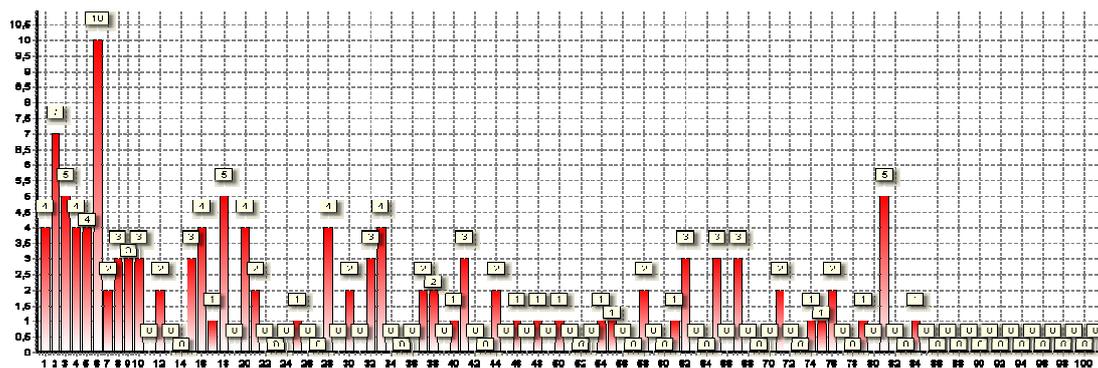
Figure 3.4.5 Sensor module-Virtual Screening distriburion.
The histogram plot reports the number of actives per cluster.

Figure 3.4.6 Pore module-Virtual Screening distribution.
The histogram plot reports the number of actives per cluster.

As clearly evincible from results in both tables and histogram plots, the best distribution was obtained for the docking calculation performed into the extracellular putative pocket within the sensor module. Especially for the lowest percentages, namely top 1 and top 2 %, the markedly higher EF values emphasize the greater reliability of the first binding pocket. However and even if it performs better than the pore module binding site, the screening results are far from the ones performed within the orthosteric binding site, where both EF and skewness values were higher and statistically more significant. On the basis of the screening results, the putative accessory pocket on the sensor module underwent to mutagenesis validation.

3.4.6 Mutagenesis study.

With these results in hands, the following step was selecting the putative key residue of the selected accessory binding site for a mutagenesis study, performed by Axxam s.p.a along with four residues within the TRPM8 orthosteric binding site, selected to investigate differences in the mechanism of action between icilin and menthol-related compounds.

His720 was then chosen, since it is involved in the most conserved H-bond in the generated complexes. It was mutated into an alanine residue, whose apolar side chain cannot preserve the mentioned polar interaction.

Aim of the study was to prove for the generated mutants the activity of icilin and menthol-related molecules; then, the capability of selected TRPM8 antagonists to inhibit icilin- or menthol-like Cooling Agent 10-mediated activation was tested.

For the first step, mutation of His720 in Ala showed no effect on both icilin-mediated and menthol-related compounds activations compared to the wild type response profile, meaning that this amino acid is not involved in agonist binding, in line with the expectations.

Contrarily, as expected for the orthosteric residues, mutations of Asn799 and Asp802 in Ala completely abolished TRPM8 activation by icilin since these aminoacids are involved in the icilin binding. The same effect was also observed for mutation of Tyr745Ala and for Ile746Ala indicating that also these amino acids play crucial roles for the icilin binding.

In the case of activation by menthol-related compounds, all the constructs behaved similarly towards the tested agonists Cooling Agent 10, WS-3 and menthol confirming a similar binding mode. Mutations of Ile746 or Tyr745 in Ala greatly reduced or completely abolished their TRPM8 activation while mutations of Asn799 and Asp802 which are located in the icilin-binding site showed any effect on neither maximal TRPM8 activation nor EC_{50}.

Then, the activity of selected TRPM8 antagonists has been evaluated on the TRPM8 mutants. Unfortunately, His720 mutation keep showing no effect on the receptor response profile, suggesting that is not involved in antagonist activity, even thought this was encouragingly supported by computational studies.

3.5 Computational selectivity prediction study on hTRPMs.

3.5.1 Introduction.

The last part of the studies on TRPM8 involved the selectivity prediction for the active compounds as obtained from the above described wet-lab HTS study. Indeed, it is well known that a major factor which hampered the clinical development of TRPM8 antagonists is their poor selectivity profile which render their use potentially unsafe.

The present study was focused on the highly homologous TRPM subfamily (TRPM1-TRPM8) and although it includes some clear approximations which will be later detailed, it led to some meaningful results. The study started from the already discussed docking results on TRPM8 and, with regard to TRPM5, it exploited the complete TRPM5 monomer which was modelled by fragments as described in Chapter 4. Hence, the study can be subdivided into two major parts. In the first part, the sensor modules (S1-S4) of the not yet modelled TRPM members were modelled by using the corresponding TRPM8 domain as the template. Then, the seconds part involved analysis and comparison of the results obtained by docking simulations performed on the so generated seven TRPM sensor modules (TRPM1-TRPM7) by utilizing a dataset including all the 1739 active molecules as found by HTS plus 10000 randomly selected inactive molecules, thus obtaining a reasonably manageable dataset.

3.5.2 Homology modeling of the sensor module.

Although utilizing a homology model as the template to generate other homology models is highly questionable, here such a choice can be justified by several key reasons. Firstly and as reported in Figure 3.5.1, there is a high degree of homology at least to be concern the S2-S3 segments which are directly involved in ligand recognition. In detail the triad of key residues involved in the binding processes (Tyr745, Asn799 and Asp802) are constantly conserved in all TRPM members and the sole minor difference is seen for TRPM6 which bears a glutamate (Glu911) instead of an aspartate. The same degree of similarity is also observed in S4 (data not shown) that, despite not being directly involved in binding, includes some key residues involved in the channel activation as above described for TRPM8. Some differences are observed in S1, especially with regard to their length

since in four members it includes a greater number of residues showing an additional peptide of five conserved residues in their N-terminal tail. However, S1 has a very limited role in binding and thus such difference should minimally impact on docking results. Secondly, the utilized TRPM8 model was extensively validated by both MD and docking simulations investigating in depth both ligand binding and activation mechanism. Hence and while considering the potential risks in using a homology model as the template, the here utilized TRPM8 template seems to be accurate enough to assure that reliable homology models can be generated at least to be concern the appropriate architecture of their binding sites. Thirdly and while avoiding to trivialize the performed calculations, the primary objective of this analysis is to gain some insights about the factors governing the subtype-selectivity with the TRPM subfamily without delving a specific TRPM member. Hence, the introduced approximations appear to be adequate to the planned objectives and do not significantly undermine the reliability of the obtained global results.

Figure 3.5.1 S1-S3 sequences alignment for all TRPM members;
key residues boxed in black.

On these grounds the primary sequences of the sensor modules were retrieved from UniProt and firstly aligned by using ClustalX as reported in Figure 3.5.1 and above discussed. The amino acid sequences were then used to generate the corresponding homology models by using Modeller and

the TRPM8 model as the template. In detail, the models were generated by Modeller9.10 using the default parameters and generating 20 structures for each run. Among the generated models, the best structures were selected according to (a) the scores computed by Modeller9.10 (i.e. DOPE and GA341) (b) the percentage of residues falling in the allowed regions of the Ramachandran Plot; (c) the agreement with the predicted secondary structure from the sequence alignment especially to be concern the TM helices; (d) the lack of not predicted gaps. Notice that the loops of the sensor modules are rather short segments and so their folding is clearly constrained by the arrangement of the TM helices. The so selected four-helix bundles were then completed by adding hydrogen atoms using VEGA. According physiological pH, Arg, Lys, Glu and Asp residues were preserved ionized, while His residues were considered neutral by default. The completed models were carefully checked to avoid unphysical occurrences such as cis peptide bonds, wrong configurations, improper bond lengths, non-planar aromatic rings or colliding side-chains. Finally, they were optimized by a final minimization made up by two phases: first a minimization without constraints until RMS = 0.1 and then a second minimization with backbone fixed until RMS = 0.01 to preserve the predicted structure.

The high degree of similarity among the modelled sensor modules is clearly documented by Figure 3.5.2 that for clarity reports the superimposition of the pivotal segments (S2 and S3) highlighting the three key residues involved in ligand binding. While considering their remarkable similarity, the considered binding sites show some significant differences which allow them to be grouped into two well defined groups. As schematized in Table 3.5.1, the TRPM's binding site can roughly be subdivided into two sides. A more internal side which is bears the three key residues and is flanked by other mostly hydrophobic residues and a more external side which is mainly characterized by hydrophobic residues which however show some relevant differences regarding the abundance of aromatic residues. In detail, Table 3.5.1 reveals that (a) TRPM3, TRPM4 and TRPM8 show only one aromatic residue in their external region; (b) TRPM1 and TRPM6 include two aromatic residues and (c) TRPM2, TRPM5 and TRPM7 include three or more aromatic residues in the more external side of their cavity. A careful analysis of the modelled binding site emphasizes that the pockets of the subtypes belonging to the third group are clearly more deepen, and while remaining substantially superficial binding site, they should be sterically more constrained for the hindrance exerted by the aromatic residues. Some slight differences are noticeable in the residues flanking the three key conserved residues. Indeed in most subtypes the key residues are surrounded by apolar side-chain but in some cases the more internal side is characterized by additional polar residues which can play a crucial role in determining the subtype selectivity. As examples, one may cite

TRPM2 which shows a second aspartate residue in that region or TRPM6 and TRPM7 which possess there an additional H-bonding threonine residue.
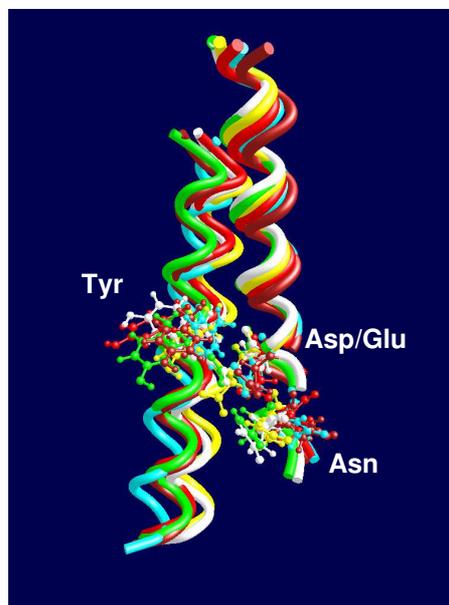
Figure 3.5.2 Superimposition of TRPM1-7 S2 and S3 helices.

Key residues are highlighted in ball and stick.

Subtype	Key residues	Internal side	External side
TRPM1	Tyr838, Asn900, Asp903	Trp899, Ile907	Tyr841, Phe845, Leu482
TRPM2	Tyr806, Asn869, Asp872	Val873, Asp866	Phe809, Leu810, Phe813, Trp868
TRPM3	Tyr905, Asn967, Asp970	Leu971, Ile974	Leu839, Tyr908, Leu911
TRPM4	Tyr790, Asn865, Asp868	Leu869	Trp864, Leu791, Leu794
TRPM5	Tyr742, Asn808, Asp811	--	Phe743, Leu746, Phe749, Trp807
TRPM6	Tyr849, Asn911, Glu914	Thr915	Leu850, Leu853, Phe856, Trp910
TRPM7	Tyr863, Asn925, Asp928	Thr829, Ile926	Phe866, Leu867, Tyr870, Phe924
TRPM8	Tyr745, Asn799, Asp802	Leu806, Phe807	Ile736, Leu750

Table 3.5.1 Residues involved in the ligand binding domain, reported for TRPM1-8.

3.5.3 Docking results.

With the above described homology models in hand, docking simulations involved a reasonably manageable dataset including the 1739 active molecules as found by HTS plus 10000 randomly selected molecules from among the inactive compounds as found by HTS. The resulting database underwent to docking simulations on all the seven remaining TRPM subtypes. TRPM8 was not considered since these ligands were already docked in the previous study and more importantly their biological activity towards TRPM8 was experimentally evaluated by HTS. Docking simulations were performed by using AutoDock. For all receptors, the search involved a 10 Å radius sphere around the key (and central) aspartate residue thus encompassing the entire cavities by applying the computational settings already described in paragraph 3.2.8. Although Ligen performed better in previous docking simulations, AutoDock was here utilized since Ligen results are particularly dependent on the availability of validated starting complexes to derive the corresponding pharmacophore models. Without these initial complexes, a stochastic program like AutoDock should be more reliable.

The so obtained results of each screening were ranked and analyzed by considering for how many subtypes the active molecules was in the top 30%, 50% and 70%. Ideally a TRPM8 selective ligand should not be in the top 70% for all other TRPM members. Contrarily, a pan-selective ligand should be in the top 30% for all TRPM members.

Figure 3.5.3 shows the obtained distributions and reveal a parabolic profile which should be indicative of a sort of binary behaviour. When considering the top 30% distribution a large part of compound are absent for all remaining seven members (the termed bin '0') and the found percentages progressively decreases in the following bins. This result is clearly understandable since it is improbable to find active molecules when considering a so challenging percentage (namely the top 30%) However and as later discussed, the distribution shows a parabolic trend since they re-increase in the bins '6' and '7'. On the other hand when considering the most generous top 70% profile most compounds appear to be understandably affinitive for all subtypes and the bin '7' is the most populated one.

A remarkable result is offered by the intermediate top 50% and here the parabolic profile is really evident since '0' and '7' are the two most populated bins while the central bins are progressively less populated. As mentioned above, these results seem to suggest that the active molecules can be roughly subdivided into selective and pan-selective compounds and molecules which putatively interact with three of four subtypes represent a marked minority. While considering the small differences between subtypes described above, this result can be interpreted considering the high degree of similarity shared by all TRPM members. Stated differently, an active molecule either has some peculiar features which render it selective for TRPM8 or it interacts with the three key residues and thus it unavoidably binds all subtypes. The obtained results suggest that the TRPM8 binding site has some distinctive structural properties which do not allow intermediate possibilities and which should allow the design of truly selective ligands.
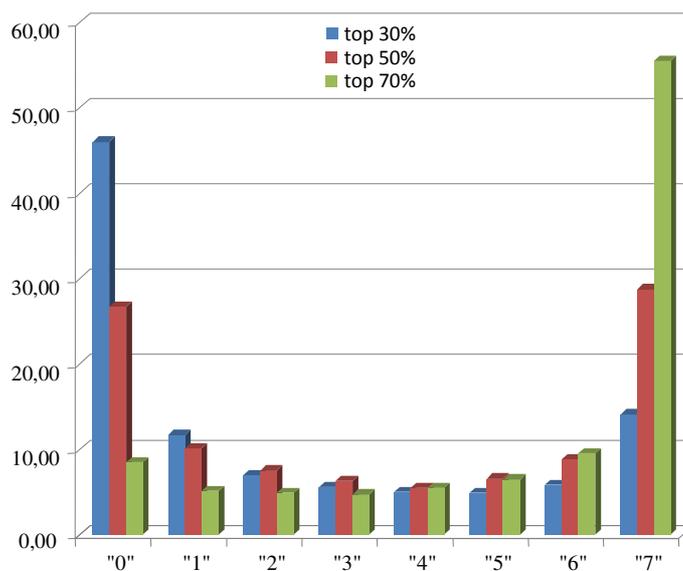
Figure 3.5.3 Graphical distribution of actives selectivity profile.

Actives are considered ranked within the top 30, 50 and 70 % of the set;

bins going from "0" to "7" account for the TRPM8 itself to all TRPM members together,

respectively, by including one more model per bin.

4. Computational studies concerning the hTRPM5 ion channel

4.1 TRPM5 structural features.

Located within the Beckwith-Wiedemann syndrome critical region-1 on chromosome 11p15.5, and identified in a study on genes correlated to such a disease, although it has been shown to be unrelated to this pathological condition, the human TRPM5 gene comprises 24 exons and predicts for a 1165 amino acids protein[192]. Within the melastatin sub-family, TRPM5 has highest homology with TRPM4, showing about the 42% of identity and 55% of similarity, while it is more distantly related to other members, such as the cold and menthol receptor TRPM8, with which the homology level is about 30% of identity and 45% of similarity.

In line with the conserved structure of almost all TRP members, TRPM5 as well is thought to spatially arrange in large intracellular amino- and carboxy-termini and in a transmembrane domain composed by six TM helices (S1-S6). It is expected to assemble in a homo-tetrameric structure since so far there is no evidence that it can co-assemble with any other TRP monomers, and no different interacting proteins have been reported as well.

The binding site for chemical agonists and antagonists has been mapped within the transmembrane portion, on the lateral surface of the first four helices block, as seen for the TRPM8. Actually, the entire site appears to be quite conserved compared to the TRPM8 one, especially for the highly conserved key residues Tyr742 on the S2 helix, and Asn808 and Asp811 on the S3 helix. On these grounds, each monomer unity is expected to be able to contact the ligand, and so four molecules would bind the tetrameric channel, causing conformational changes which lead to channel opening.

Intracellular N- and C-termini brings several site expected to be involved in protein-protein interaction, specially along the C-terminus, were coiled-coil domains have been predicted.

The pore channel is predicted to be lined by the transmembrane segments S5-S6 and relative linker, relying on the fact that mutations addressed toward aminoacids in theis area reflect into changing in the ions selectivity and permeability.

4.2 TRPM5 functional features.

Studying the TRPM5 profile in taste cells introduced the idea that it is part of a phospholipase-mediated pathway[193], specially if considering that TRPM5 currents can be generated by stimulating G_q-coupled receptors, and their consequent PLC cascade, in many heterologous cellular type[194].

As known, PLC hydrolyzes PIP_2 producing DAG and IP_3, which in particular stimulates the exit of calcium from cytoplasmic stocks, event that, in line with the TRPM5 polymodal activation profile discussed later, would directly activate the channel (Figure 4.1).

To date, this mechanism have been supported by diverse independent studies. For example, it has been proved that the previously mentioned PLC dependant TRPM5 triggering doesn't happens if IP_3 receptors are blocked or intracellular calcium is not free to move out from the stores and to reach the cellular membrane (calcium buffering[195]), or can contrarily be overstimulated, for example increasing IP_3 level (UV-uncaging[196]).

The resulting model, in which TRPM5 is activated as a consequence of the activation of other receptors (taste receptor or GPCRs) leading to increase intracellular calcium concentrations, is consistent with physiological data from taste cells and with targeted deletion of taste transduction molecules[197,198].

There are many ways in which intracellular calcium signal can be induced and, moreover, depending on the trigger, the duration time as well as the intensity may change. For example, calcium signal can rise and end very fast, like happens when induced by ryanodine receptors[199], reaching 20–30 μM concentrations, or can increase slower and stays on for several seconds.

Clearly, to better understand how TRPM5 channels can be activated by intracellular calcium, two fundamental aspects must be investigated: (i) the distance between TRPM5 and calcium source, the closer the easier to be activated, and (ii) its sensitivity to calcium.

The first question can be quickly answered considering that both TRPM5 channels and IP_3 receptor or PLCβ2 are known to locate across plasma membrane of the whole taste cell, condition that strongly supports the possibility that they are part of the same signalling path, like happens, for example, for the signaling components of fly phototransduction[200].

For what concern the second aspect, it might be a bit more complicated to be investigated, as seems to be reading the not always consistent results in literature. Among the diverse techniques for determination of a given receptor calcium sensitivity so far developed, the inside-out patches is in all probability the most accurate. Using this approach, experimental evidences attested that TRPM5 channels are activated by intracellular calcium with an EC_{50} ranging from 20 to 30 μM[201], and this value would be in line with the one measured for the activation of native channels in taste cells, that

is about 8 μM. Moreover, this sensitivity profile is not stable during the activation, but has been proved to decrease up to 80 μM, event that might be partially explained considering the decreasing stocks of PIP_2 in the environment. All these data depict a condition in which TRPM5 would be so activated only if located close to the calcium stores.

The situation changes if approaching the measurement with the whole cell recording technique, from which evidences TRPM5 appears to be much more sensitive to calcium activation[202], allowing it to respond to the stimulus even if far from its location. Several considerations can intervene to level this important difference, such as the "calcium-induced calcium release", that would allow even really low calcium concentrations to activate TRPM5 through the release of more calcium from intracellular stores, making the channel able to detect not only local but global calcium changes as well.

It is worth noting that TRPM5 and TRPM4, despite being the only two TRP channels activated by Ca^{2+}, they are not crossed by this ion. Moreover, TRPM5 channels show little discrimination among the monovalent cations too, namely Na^+, K^+ and Cs^+, that has been attributed to several residues along the putative pore, basing on mutagenesis studies in which changing them with the correspondent residues of the calcium permeable TRPV6 led to a slight but detectable calcium permeability[203]. This further confirmed that S5-S6 linking loops of each monomer act in cation detecting and filtering, but unfortunately, so far, no experimental evidence has been able to reveal the relation between residues and ion selectivity, suggesting that multiple residues or regions might cooperatively contribute to this feature.

Beside calcium-induced activation, TRPM5 was proved to respond to voltage shifts as well[204], like happens for TRPM8 and TRPM4 members, suggesting how this event could intervene as a modulator of the diversely-induced channel gating[205], as the cold regulation of TRPM8, TRPV1 and TRPM5, and the PIP_2 regulation of TRPM4 and TRPM8[206]. Unfortunately, differently from the mentioned modulating system for which a good part is now clear, voltage modulation of TRPM5 is still quite unknown.

TRPM5 activity is regulated by lipids as well. Several experimental evidences proved the phosphoinositide $PI(4,5)P_2$ to participate in the activation of many ion channels, among which TRPM5, and indeed the its hydrolysis products have been proposed to underlie rundown of many $PI(4,5)P_2$-sensitive ion channels[207]. Beside PIP_2, arachidonic acid as well, probably generated in taste cells from DAG by the actions of PLA2-IIA, is another lipid that seems to be involved in TRPM5 regulation[208].

Recently, TRPM5 was proved to respond to long-chain unsaturated free fatty acids, so being related to the fat taste sensing. Just to mention one example, linoleic acid was proved to depolarize mouse taste cells so eliciting a robust intracellular calcium rise via the TRPM5 activation[209]. The induced responses were shown to depend on G-protein-phospholipase C pathway, suggesting how G-protein-coupled receptors are involved in the transduction of fatty acids. Moreover, mice lacking TRPM5 channels exhibits no preference for and reduced sensitivity to linoleic acid.

Finally, TRPM5 was found to respond to warm temperatures, even though it is structurally more related to the cold-activated TRPM8 than to the heat-activated TRPV members[210]. Here, heat was speculated to intervene by shifting the midpoint for voltage-dependent activation toward negative voltages, contrarily to what happens for TRPM8, which goes toward more positive values. Nevertheless, heat is not sufficient to activate TRPM5, which even at warm temperatures still requires the Ca^{2+} contribute.

TRPM5 is sensitive to extracellular pH level below 7.0 and is completely blocked by pH 5.9[211]. By comparison, TRPM4 is insensitive to pH levels as low as 5.4. Two residues account for most of the pH sensitivity of TRPM5, a glutamate in the S3–S4 linker and a histidine in the pore region.
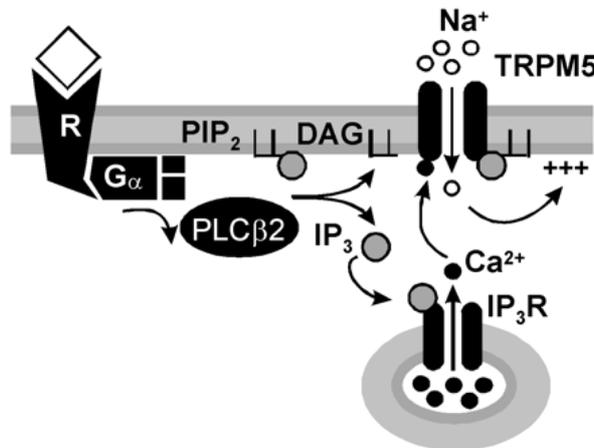
Figure 4.1 TRPM5 taste transduction mechanism.
(Picture from the work of Liu and colleagues[212])

4.3 TRPM5 expression.

TRPM5 was first reported to be expressed in mammalian taste buds, in a subset of cells that co-express receptors for bitter, sweet, and umami tastes. Subsequent studies showed a more wide-spread distribution of the channel. Using a mouse in which the TRPM5 promoter drives expression of GFP[213] or direct antibody labelling, TRPM5 expression has been detected in two distinct subsets of olfactory neurons, namely ciliated and microvillous neurons, in the vomeronasal organ, in the gastrointestinal tract, and in the respiratory system. In most of these tissues, TRPM5 is expressed in solitary cells. Within the gastrointestinal system, TRPM5 is expressed in the stoma, small intestine, and colon in sparely distributed solitary cells.

TRPM5 is also expressed in pancreatic islets, where it is localized into insulin secreting β-cells. In addition, there are a few reports describing the TRPM5 expression in the central nervous system.

4.4 TRPM5 physiology.

As already mentioned, TRPM5 was found to be highly expressed in the subset of taste receptor cells that mediate bitter, sweet, and umami taste. The corresponding taste-sensing pathways start activating the G protein gustducin and phopholipase C β2 followed by an intracellular cascade leading to membrane depolarization and ATP release. Data from electrophysiological experiments showed that a TRPM5-dependent current follows the increase of IP_3 or Ca^{2+} concentrations[214].

Tastants-mediated activation of taste receptors was found to generate TRPM5-related events as well[215], attesting how TRPM5 is crucial for normal sensing, how strongly suggested by experimental studies on knockout animals in which the lack of TRPM5 reflects in a considerable less sensitive to bitter, sweet, and umami, although preserving both sour and salty sensing[216].

Similarly to TRPM8, besides sensory tissues, TRPM5 is known to be expressed in many non sensory tissues as well. Just to mention few examples, duodenum and stomach cells express TRPM5, that it is believed to play a role in post-ingestive chemosensation[217]; the respiratory and olfactory epithelia as well present TRPM5 in solitary chemosensory cells, where it may function downstream of bitter receptors to detect inhaled irritants or pheromones[218].

An important role is played by TRPM5 in pancreatic β-cells, where it was found to be involved in insulin secretion[219]. In detail, rising blood levels of glucose induces its uptake into β-cells, where it imbalances ATP against ADP, and so blocks a specific family of ATP/ADP sensitive potassium

channel, with consequent depolarization of the membrane potential may be through the opening of a still not determined channel, probably sodium permeable. Here, voltage gated calcium permeable channels are gated allowing the intracellular calcium level to rise and directly acts on the insulin releasing from intracellular stores[220].

To date, many actors have been proposed to regulate these depolarizing currents, including TRPM4, TRPM2, and TRPM5, which in particular has been tested in knockout animal in vivo studies. In detail, in response to glucose administration, TRPM5 lacking animals showed a decrease in both insulin secretion and glucose clearance if compared with wild type mice[221], suggesting how TRPM5 would intervene in the only partially known genetic propensity to develop the pathology and, hopefully, how important it would be as target for pharmacological treatment of diabetes or metabolic diseases in general.

4.5 TRPM5 pharmacology.

The first field in which TRPM5 ligands, specially antagonists, have been explored is the taste sensing manipulation, with the aim to alterate the TRPM5-related tastes perception, may be so masking unpleasant sensations for both drugs and foods oral ingestion.

To date, really few selective molecules have been discovered. Among the non selective ones, for example, the bitter chemical quinine was found to block TRPM5 too, with an EC_{50} of 50 µmol, and the contemporary administration of quinine with a sweetener agent was proved to silence the sweet response, so supporting the mentioned interaction between sweet and bitter tastes[222].

In a HTS for specific TRPM5 antaponists, triphenylphosphine oxide (TPPO) was identified[223], reported to act with an IC_{50} value of 12 µmol. Contrarily, the related triphenylphosphine showed no effect, indicating how the oxygen atom would be crucial. TPPO appears to be specific for TRPM5, and no effect of TPPO on the activity of TRPA1, TRPV1, or TRPM4 was observed.

Nicotine as well has been proved to inhibits TRPM5 channel. In a study aimed to investigate the consequences of the oral administration of nicotine and its metabolites, such as cotinine, they were found to reduce the activity of neurons in the nucleus of the solitary tract[224].

Nicotine in particular affects TRPM5 channel, decreasing the maximal whole-cell conductance and augmenting the rate of channel closure. Besides negatively acting on TRPM5 it also affects IP_3 receptors, as well important for sensing the sweet taste. Moreover, intracellular Ca^{2+}-release

channel activity was proved to be promoted by the nicotine, while highly inhibited by quinine, and only modestly by caffeine, theobromine and theophylline.

4.6 TRPM5 Homology modelling.

As for TRPM8, at the time of the study and so up to now, the TRPM5 structure was not yet resolved by crystallographic experiments. So, the TRPM5 modelling was based on homology techniques by adopting a fragmental strategy, since highly homologous global templates were unavoidable. Such an approach tends to favour the local homologies shared with different templates thus allowing more reliable models to be generated.

Human TRPM5 primary sequence was then subdivided into structural domains, namely the N- and C-termini, the six transmembrane helices and the five relative loops. Except for the intracellular domains, modelled by using the importin beta and the HCN2 pacemaker channel as the templates, respectively (as already done for TRPM8) all fragments were submitted to FUGUE online server for their homology modelling.

As already described, FUGUE tool is able to manage more then one request per time. In absence of highly homologous templates, it conveniently produces many different solutions among which it is possible to select the most promising ones in a reasonably time. Another interesting feature is that FUGUE generates structurally reliable models even for rather short peptides thus rendering this on line tool particularly well suited for the fragmental strategy

For each segment and among the realistic models produced by FUGUE, the preferred structure was chosen considering the following major conditions: (a) the predicted secondary structure from the sequence alignment, as obtained using ClustalX; (b) the lack of not predicted gaps; (c) the prediction score as calculated by the programs; (d) the helix conformation of six transmembrane segments with characteristic slight bend of helices containing proline and glycine residues; (e) the global "U" shape for the loops in which the two ends are close enough to join to adjacent TM segments. Then, the so selected models were completed by adding side chains and hydrogen atoms using VEGA. According physiological pH, Arg, Lys, Glu and Asp residues were preserved ionized, while His residues were considered neutral by default. After a careful visual scrutiny of obtained structures to avoid unphysical conditions, the segments were minimized and assembled using the mammalian Shaker Channel Kv1.2 (PDB: 2A79) as the template, the same used to assemble the TRPM8 model. First, the transmembrane helices (S1-S6) were superimposed to the corresponding

segments of the Shaker Channel by matching the equivalent Cα atoms. Next, the extramembrane loops were manually added, as their arrangement is reasonably defined by the position of transmembrane segments. However and as described later, their conformation was further relaxed by targeted molecular dynamics simulations. The superimposed fragments were then assembled by manually adding the connecting bonds and the so obtained model underwent an initial minimization until RMS = 1 to discard high-energy interactions, followed by a local minimization until RMS = 0.05, where all atoms were kept fixed except for atoms included within a 7.5 Å sphere around the manually connected bonds (at the fragment ends). The completed model was carefully checked to avoid unphysical occurrences such as cis peptide bonds, wrong configurations, improper bond lengths, non-planar aromatic rings or colliding side-chains. Finally, the model was optimized by a final minimization made up by two phases: first a minimization without constraints until RMS = 0.1 and then a second minimization with backbone fixed until RMS = 0.01 to preserve the predicted structure.

4.7 TRPM5 optimization by MD simulation.

Differently from the TRPM8 modelling study, in which the entire tetrameric structure was generated by protein-protein docking starting from the refined monomer, in this study only the monomeric unity of TRPM5 has been considered. This is justified by considering that the orthosteric binding domain we were interested in is located on the intramembrane surface of the *sensor module* (S1-S4), in line with the other TRPM members, and as such their structure is not influenced by the presence of the other monomers.

For further improvement, the TRPM5 monomer underwent to a molecular dynamics simulations study, that allows to better refine the entire model, especially to be concern the extra- and intra-cellular loops while the transmembrane helices which were reliably modelled by homology techniques were maintained fixed. Thus, the model was inserted in a cubic box (80 Å x 90 Å x 90 Å) of water molecules and underwent a preliminary minimization to optimize the relative position of solvent molecules. During this minimization and the following MD run the transmembrane segments were kept fixed and only loops atoms are free to move. Indeed, the selected water medium is clearly unsuitable for TM segments and the MD run was aimed at refining the loop conformations. In detail, the MD run had the following characteristics: (a) Newton's equation was integrated using the r-RESPA method (every 4 fs for long-range electrostatic forces, 2 fs for short-

range non bonded forces, and 1 fs for bonded forces); (b) the simulation space was stabilized by introducing the Periodic Boundary Conditions (90 Å × 100 Å × 100 Å); (c) the long-range electrostatic potential was treated by the Particle Mesh Ewald summation method (80 × 100 × 100 grid points); (d) the temperature was maintained at 300 ± 10 K by means of the Langevin's algorithm; (e) Lennard-Jones (L-J) interactions were calculated with a cut-off of 10 Å and the pair list was updated every 20 iterations; (f) a frame was stored every 5 ps, to yield 2000 frames (g) the atoms of TM segments were kept fixed. The simulations were carried out in two phases: an initial period of heating from 0 K to 300 K over 300000 iterations (300 ps, i.e. 1 K/ps) and the monitored phase of 10 ns. During the MD run, the loops show significant beneficial rearrangements as assessed by an increase of about 10 % of residues falling in the allowed regions of the Ramachandran plot (from 75.95 % to 84.25%). The so obtained last frames was finally minimized and utilized in the following docking simulations. All minimizations in model construction and equilibration were performed using the conjugated gradients algorithm. The structural reliability of the final model was finally assessed by several parameters such as Ramachandran Plot and side chain χ-space.

4.8 TRPM5 binding site.

Before starting docking simulations, the so generated TRPM5 monomer was utilized to analyze its binding cavity comparing it with that already described of TRPM8. Similarly to the TRPM8 binding site, the TRPM5 ligand binding domain appears to be markedly apolar, in line with the transmembrane environment. Moreover, key residues involved in ligand recognition for both models appear to be exactly conserved. In particular, Tyr742, Asp811 and Asn808 (corresponding to Tyr745, Asp802 and Asn799 in TRPM5) arrange their side-chains in a very similar way, and thus they should be involved in similar inter and intramolecular interactions, as already mentioned for TRPM5.

However a slight difference can be noted between TRPM5 and TRPM8 binding sites. Indeed, while the TRPM8 pocket appears to be completely superficial and lined only by few aliphatic residues, such as Ile746 and Leu750, which surround Tyr745 and contribute to the hydrophobicity of this region, the TRPM5 cavity appears to be slightly deeper and surrounded by a greater number of aromatic residues which should restrict the cavity capaciousness rendering it more affinitive for ligands rich in aromatic moieties.

4.9 QSAR validation of the TRPM5 orthosteric site.

Relying on the high homology of the orthosteric binding site between TRPM5 and TRPM8, and considering the multiple successful studies through which TRPM8 binding domain has been validated, all docking calculations were focused on this site.

As already mentioned, TRPM5 common activating role in taste sensing clearly suggests that TRPM5 blockers could be used as innovative taste masking agents for oral pharmaceuticals with unpleasant, bitter-tasting components. On this considerations, I have been mostly interested in TRPM5 antagonism, dealing with antagonists from the initial validation stages up to the virtual screening campaigns. At the date of the work, really few selective ligands were available in literature. Among them, it is worth mentioning the compounds reported by the patent US 7674831 B2[225] from Redpoint Bio Corporation, "Heterocyclic compounds as sweetener enhancers", dated march the 9th, 2010. It describes a set of sweet enhancers developed and tested for their capability to improve the activation of TRPM5 in presence of endogenous gating inducers, such as ATP. Among the reported heterocyclic compounds, seventeen molecules showed negative percentage of activation, ranging from -139 to -16%, thus acting as antagonists.

So, these molecules have been selected as training set for docking calculations aimed to validate the reliability of the binding site of the TRPM5 model.

The ligand's structure was built using the VEGA ZZ software, and the overall geometry and the atomic charges were optimized using MOPAC 6.0. Their conformational profile was explored by a MonteCarlo procedure (as implemented in VEGA) which generated 1000 conformers by randomly rotating the rotors. All geometries so obtained were optimized and clustered according to similarity to discard redundant ones; in detail, two geometries were considered as non-redundant if they differed by more than 60° in at least one torsion angle.

As done for the TRPM8, all calculations about the orthosteric binding site were performed on the single monomer, since this region is located far from the inter-monomer interface regions, within the membrane.

Although the above reported docking simulations on TRPM8 were carried out by the recently developed LiGen program, most of the docking simulations here described were performed by using Autodock 4.2 and PLANTS programs. When comparing the results obtained by the two pieces of software, PLANTS appeared to provide better results in terms of computed poses, predictive power and requested computational time. Thus and as in this docking analysis, several

studies were initially performed by AutoDock and then repeated by using PLANTS. For the sake of clarity, only the optimized results obtained by PLANTS were here described.

In detail, PLANTS was used with default settings and without geometric constraints. The search was focused on a 10.0 Å radius sphere around a centroid defined by the three key residues Tyr742, Asp811 and Asn808 thus encompassing the entire binding cavity. For each campaign, speed 1 was used, 10 poses was generated for each ligand and scored by using the most exhaustive ChemPLP function. The so obtained best complexes were then minimized by keeping fixed the atoms outside a 10.0 Å radius sphere around the bound ligand. The refined complexes were then utilized the recalculate a set of docking scores including (a) the three PLANTS score functions, (b) the AudoDock ones, namely the predicted pKi, the internal (intermolecular) energy, the non-bond energy accounting for Van der Waals interactions plus Hydrogen bonds plus the Desolvatation energy, the Electrostatic interaction, (c) the X-SCORE scores, namely HP, HM and HS[226], (d) the AMMP non-bond interaction energy, calculated with SP4 force field, (e) the VEGA ZZ scores, such as the MLPInS score[227], accounting for the compatibility of hydrophobicity/hydrophilicity between ligand and protein, in its four distance related functions (linear, square, cubic or Fermi's), the Electrostatic energy, with and without distance dependent dielectric function, and the non-bond interaction energy calculated with CHARMM force field.

As a representative example of the obtained complexes, Figure 4.2 reports the detail of the complex between TRPM5 and compound **2** (2-(2-(2-Methoxyphenoxy)ethylthio)-1H-benzimidazole; -136% of activation). It is interesting to observe that the key Tyr742 residue stabilizes $\pi-\pi$ stacking interactions with the ligand benzimidazole which also contacts Asp811. The methoxyphenoxy moiety approaches Asn808 with which it elicits H-bonds. Finally the complex is further stabilized by a rich set of hydrophobic contacts involving both aliphatic and aromatic side chains.

Similarly the ligands extracted by the minimized complexes were utilized to compute a set of representative ligand-based descriptors. Overall, these descriptors represent the starting point for a multivariate QSAR studies by canonical regression techniques, aimed to develop a model, the predictive power of which can indirectly confirm the reliability of the obtained model thus allowing the activity on unknown antagonists to be predicted.
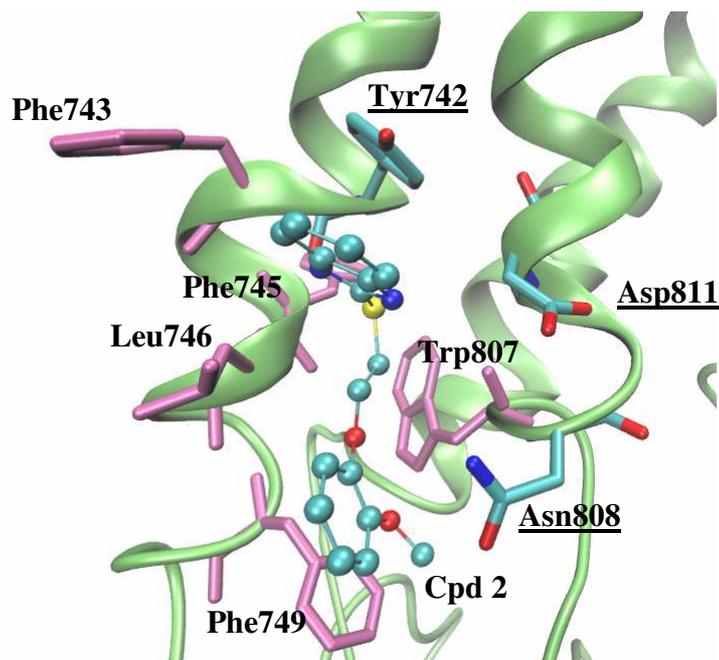
Figure 4.2 Representative scheme of the Cpd2-TRPM5 complex;
Compound 2 is visualized in ball and stick; key residues are underlined;
hydrophobic residues are reported in violet.

In this study, the QSAR analyses can be subdivided into two steps since the first part involved the development of correlative models by including only ligand-based descriptors, while, in the second step, the developed equations exploited all available descriptors. The equations generated in the first phase have two major objectives. On one hand, the comparison of the statistics for the best equations obtained in the two steps should reveal the effective contribution of the docking simulations. Ideally, the equations obtained in the second phase should show significantly better statistics; if not, docking simulations would lose a large part of their relevance. One the other hand, ligand-based equations can be seen as fruitful tools to filter huge databases so focusing docking simulations only on the most promising datasets.

Considering the small size of the dataset, the developed equations have to include at most two variables, to avoid overfitting conditions. To be concern the first ligand-based study, the best equation was obtained combining (i) the number of the flexible torsions (*FlexTors*) of the

compound and (ii) its molecular weight (*MW*). The equation and its statistical parameters are here reported, followed by the dispersion plot of the distribution (Eq.1, Figure 4.3):

Eq. 1   *pInhib =   2.6332 + 0.0954*FlexTors – 0.0038*MW*

$$r^2 = 0.49 \ \ q^2 = 0.21 \ \ SE = 0.199 \ \ F = 6.15 \ \ P = 1.32e\text{-}02$$
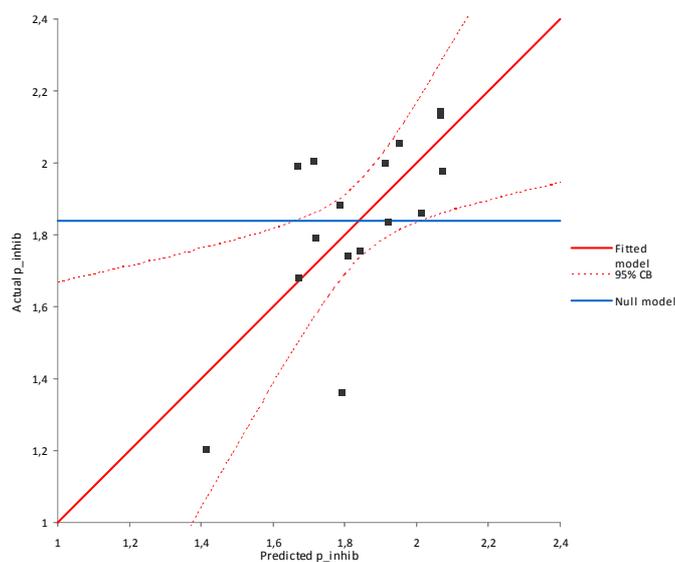
Figure 4.3 Linear regression for the pInhib predicted with the correlative equation Eq.1

As indicated by the signs in the equation, flexible torsions are predicted to positively contribute, suggesting how the ligand can preserve its flexibility within the pocket, a results easily explainable considering its superficiality. Again, the second term suggests that, despite being not so deep, the binding site has however well-defined and constraining spatial requirements since small ligands perform better.

The statistics of the Equation are not really satisfactory, a poor result which emphasizes the scarce predictive power of these ligand-based descriptors and invites to involve also docking scores, in order to improve the robustness of the model. And indeed, combining both ligand and docking descriptors led to much better predictive equations, among which the best performing one is here reported (Eq 2, Figure 4.4). It was obtained with (i) the number of the flexible torsions, thus

confirming the role of ligand flexibility, and (ii) the scoring function *HM_score*[228], which is one of the three scoring components generating by the well-known X-Score function and which accounts for steric hindrance, H-bonds, and hydrophobic interactions, the last one weighted by a value, *HM*, that accounts for the compatibility of the ligand apolar atoms with the environment of the pocket. Correlative equation and distribution plot are here reported:

Eq.2   *5.4302 - 0.8914*HM_Score + 0.0709*FlexTors*

$$r^2 = 0.75 \quad q^2 = 0.62 \quad SE = 0.137 \quad F = 19.92 \quad P = 1.10e\text{-}04$$
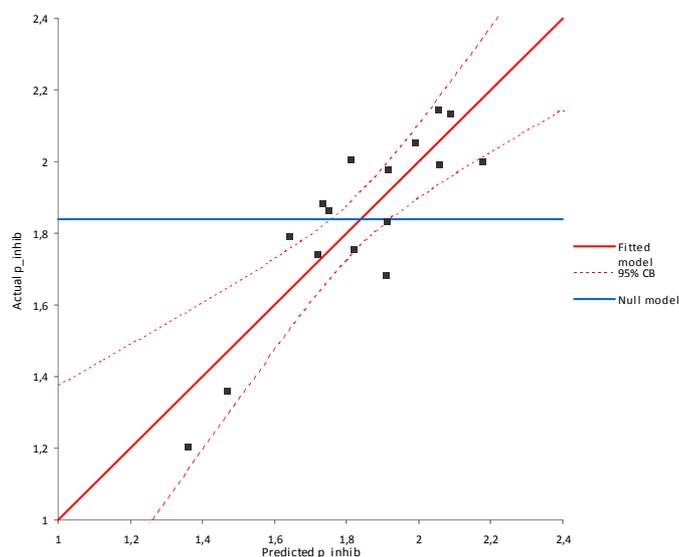
Figure 4.4 Linear regression for the pInhib predicted with the correlative equation Eq.2

In this case, the statistics of this docking-based equation are consistently higher than those of the previous model, as evidenced by their r^2 and q^2 values. This result affords an encouraging confirmation for the reliability of the here generated TRPM5 homology model, and allow us to move to the following step of the validation process, namely the virtual screening test.

4.10 TRPM5 virtual screening.

Once the reliability of the main binding site was encouragingly validated by testing its capability to correctly interact with representative ligands, the following step was to further assess the TRPM5 performances in a virtual screening study. Indeed, successful results of the so developed screening strategy would represent a further validation of the reliability of the TRPM5 homology model and clearly could be applied in screening campaigns on large databases.

To this end and similarly to what was already described for TRMP8 (see 3.2.8) a database of known antagonists was collected from patent literature. In detail and besides the 17 compound already used in QSAR analyses, antagonists came from four different patents, namely US 0259875 A1[229], US 0306030 A1[230], US 0033393 A1[231] and US 7723075 B2[232]. They included (a) 10 triaryl-substituted imidazole derivatives from the first patent, (b) 5 dimethoxyphenyl-acetohydrazide derivatives from the second, (c) 65 acetohydrazide derivatives from the third and (d) 3 heterogeneous derivatives from the fourth patent. In total, 100 active compounds were collected and jointed to 9900 putative TRPM5 inactive molecules, selected from among the compounds screened in the HTS experiment for their activity toward TRPM8 and proved to be inactive ligands in order to reach the already described percentage of 1% of active compounds. Albeit questionable, the choice of the here utilized decoys was supported by the high homology between the two receptor binding sites, and, however, it was due to usual lack of experimental published data about inactive molecules. Clearly, inactive compounds were selected with a physicochemical profile similar to that of the active ones to avoid biasing conditions.

As above mentioned all docking simulations started by using AutoDock 4.2 and then were performed by PLANTS obtaining markedly better results. Also here AutoDock 4.2 produced enrichment factors so close to the random (best obtained EF = 1.56) not to give a glimpse of room for improvement even exhaustively changing the docking parameters. By contrast and since from the first run performed with default parameters, PLANTS afforded EF values almost double compared to those obtained by AutoDock. Hence, PLANTS was definitively adopted for this screening campaign by gradually varying the available parameters especially to be concern the scoring function and the protein region on which the docking search was focused. The best results were thus obtained by using the PLP95 scoring function and focusing the search on a 8 Å radius sphere centred on the carboxylic carbon atom of Asp811. Ligands were considered as flexible while all protein residues were kept fixed. Although PLANTS allows the flexibility of some selected side chains to be considered, this would increase the required computational time so as to render the

simulations unsuitable to screen large databases. The simulations were carried out at speed 1 level, to guarantee more detailed calculations, and 1 pose was computed for each ligand.

Although PLANTS calculates the simple total Score function as well as several weighted scores (e.g. by the rotatable bonds and by the molecular weight) in order to harmonize the values with the unavoidably different physicochemical properties of the ligands within the database, the best results were produced by the simple total score.

In detail, the arbitrarily chosen enrichment factors (namely, top 1, 2 , 5, 10 and 20%) were evaluated. The Skewness factor was also considered and visualized by the histogram plot of the entire distribution (Table 4.1, Figure 4.5).

n (top %N)	Cpds	Act	EF	Sk
1	100	6 (1)	6.00	
2	200	10 (2)	5.00	
5	500	28 (5)	5.60	2.3
10	1000	52 (10)	5.20	
20	2000	65 (20)	3.25	

Table 4.1 Virtual screening evaluation, molecules ranked per PLP95 Total score.
"Cpds" reports the number of molecules hosted in the considered subset; "Act" stays for the
actives compared to the expected one in parenthesis; EF reports the Enrichement factor values;
"Sk" the Skewness value of the entire distribution.
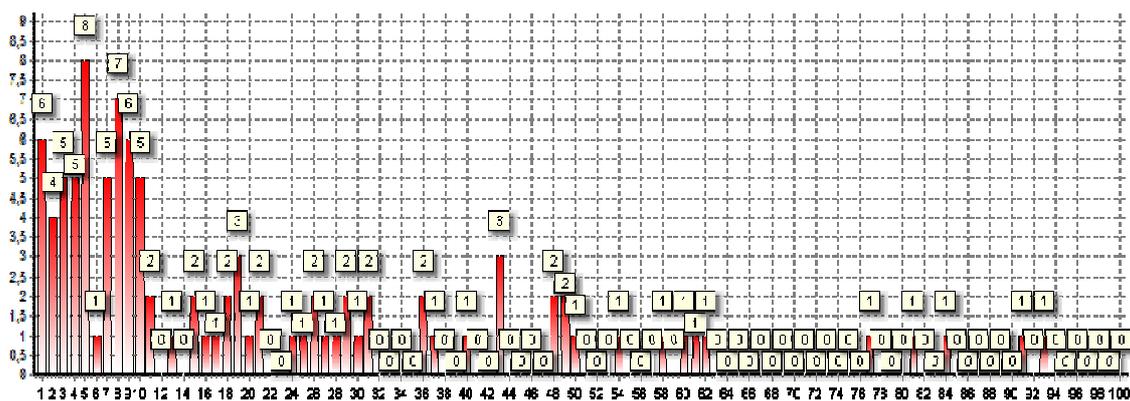
Figure 4.5 Virtual screening ranking per PLP95 Total score.
The histogram plot reports the number of actives per cluster.

Table 4.1 shows remarkable results, especially for the Top 1% enrichment factor and this value is extremely relevant because it suggests that such a screening strategy can conveniently tackle the problem of early recognition. The richest cluster is the fifth, hosting 8 active compounds, however, it rises among other rich clusters, in line with the expectations. Notably, in the top 15% of the distributions, the simulations ranked 57 active compounds out of 100, more than one half (Figure 4.5). This encouraging result clearly emphasizes the reliability of both the here generated homology model and of the utilized docking strategy, and promotes them to be used in the virtual screening studies with promising perspectives.

4.11 Correlative equations versus virtual screening.

In the previous screening analysis the ligands were ranked by simply using the PLANTS scores without any further mathematical processing. For the preliminary QSAR validation, two notable predictive models were however developed and therefore they can be interestingly exploited also to rank the screened database with a view to comparing the screening performances of these predictive equations with that of a "standard" virtual screening campaign. Moreover and since the first equation includes only ligand-based descriptors, it is possible to investigate whether they can be effective in screening analyses. Such a study should reveal the concrete role played by docking simulations and if it is possible to hugely reduce the computational costs by screening the databases by pure ligand-based methods.

Thus, the first correlative equation (Eq. 1) including only ligand descriptors was utilized to predict the activity (*pInhib*) of all 10000 molecules and the corresponding metrics were calculated. Hopefully (for docking supporters) the obtained results (Table 4.2, Figure 4.6) show that this first correlative equation not only not enrich the set, but does even worse than random.

n (top %N)	Cpds	Act	EF	Sk
1	100	1 (1)	1	
2	200	0 (2)	0	
5	500	4 (5)	0.8	0.1
10	1000	6 (10)	0.6	
20	2000	11 (20)	0.55	

Table 4.2 Virtual screening evaluation, molecules ranked per pInhib predicted with Eq. 1.
"Cpds" reports the number of molecules hosted in the considered subset; "Act" stays for the
actives compared to the expected one in parenthesis; EF reports the Enrichement factor values;
"Sk" the Skewness value of the entire distribution.
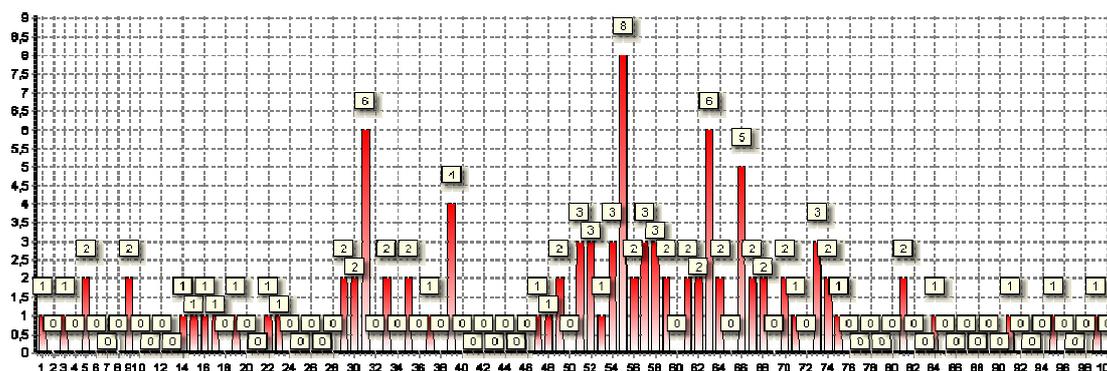
Figure 4.6 Virtual screening ranking per pInhib predicted with Eq.1.
The histogram plot reports the number of actives per cluster.

This very poor result may be due fact that also the starting equation showed rather poor statistics and thus a second ligand-based equation including three descriptors and showing better statistics was tested (Eq. 3). Even though such an equation may be questionable when considering the low number of compounds, it could yet perform better in screening campaigns. However, the distribution of actives appears clearly to be randomly dispersed among all the bins even in this second test, revealing the complete unsuitability of these ligand-based models in virtual screening (Table 4.3, Figure 4.7) Clearly, part of this failure can be due to the unavoidable differences between training set and screened database. Moreover it should be emphasized that ligand-based pharmacophore models can be really effective in virtual screening (as seen for SMARTS-based study on TRPM8) because they are able to catch the structural features characterizing the active molecules. Conversely, it is very improbable that simple physicochemical descriptors can be able to suitably discriminate between active and inactive compounds even when combined in multivariate models.

Eq. 3   *1.999 + 0.1011*Hb Acc + 0.0556*Torsions – 0.0746*Heavy Atoms*

$$R^2 = 0.69 \; Q^2 = 0.54 \; SE = 0.161 \; F = 8.87 \; P = 2.26e\text{-}03$$

n (top %N)	Cpds	Act	EF	Sk
1	100	1 (1)	1	
2	200	1 (2)	0.5	
5	500	1 (5)	0.2	0.1
10	1000	7 (10)	0.7	
20	2000	17 (20)	0.85	

Table 4.3 Virtual screening evaluation, molecules ranked per pInhib predicted with Eq. 3.
"Cpds" reports the number of molecules hosted in the considered subset; "Act" stays for the
actives compared to the expected one in parenthesis; EF reports the Enrichement factor values;
"Sk" the Skewness value of the entire distribution.
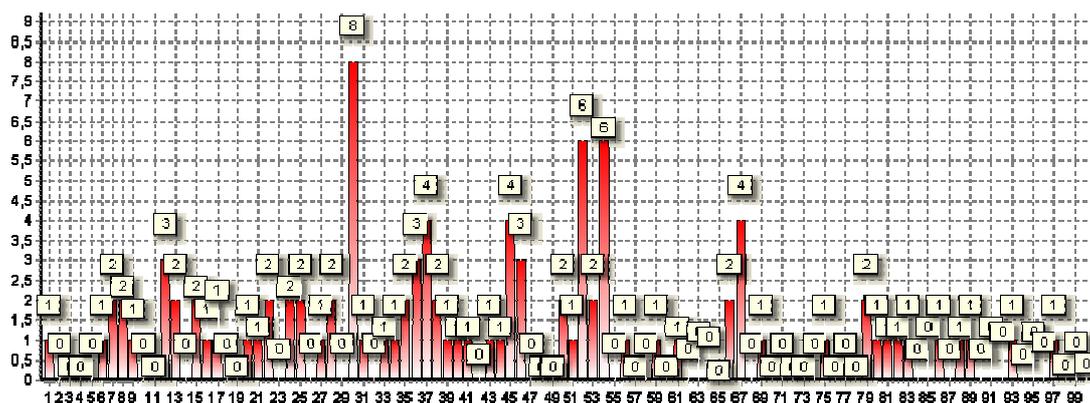
Figure 4.7 Virtual screening ranking per pInhib predicted with Eq.3.
The histogram plot reports the number of actives per cluster.

The results enhance when docking scores are involved in the study. Since the calculation of X-Score for all screened database was unfeasible, the second QSAR model was modified by including the PLP Total Score instead of the *HM_score* so obtaining a new equation with slightly worse statistics ($r^2 = 0.68$).

Table 4.4 reports the EF values of this structure-based model, showing that they are notably better than those obtained by first equations but a bit worse than those generated by canonical virtual screening strategy (Table 4.4, Figure 4.8). This result can be still explained considering that the molecules used to derive the equation have different and limited properties compared to the screened library thus emphasizing that QSAR equations have a very restricted efficacy when dealing with large and heterogeneous datasets.

n (top %N)	Cpds	Act	EF	Sk
1	100	4 (1)	4	
2	200	7 (2)	3.5	
5	500	10 (5)	2	1,42
10	1000	26 (10)	2,6	
20	2000	41 (20)	2,05	

Table 4.4 Virtual screening evaluation, molecules ranked per pInhib predicted with Eq. 2. "Cpds" reports the number of molecules hosted in the considered subset; "Act"stays for the actives compared to the expected one in parenthesis; EF reports the Enrichement factor values; "Sk" the Skewness value of the entire distribution.
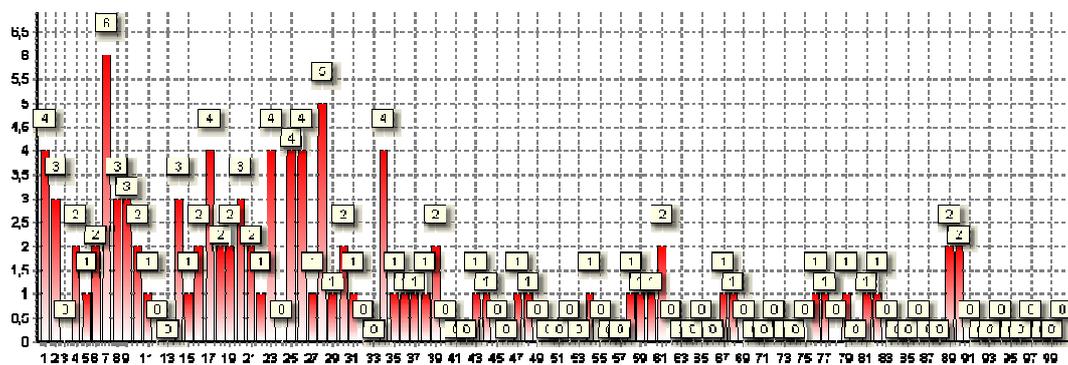
Figure 4.8 Virtual screening ranking per pInhib predicted with Eq.2.
The histogram plot reports the number of actives per cluster.

5. Computational studies concerning the hTRPV4 ion channel

5.1 TRPV4 structural features.

Ion channel TRPV4 was individuated by screening expressed sequence tag databases for sequences with similarity to TRPV1 and TRPV2, already individuated by expression cloning. It was initially indicated with different names by the different groups who cloned it, such as OTRPC4 (OSM-9-like TRP channel 4)[233], VROAC (vanilloid receptor–related osmotically activated channel)[234], TRP12[235] or VRL-2 (vanilloid receptor–like channel 2)[236].

The human TRPV4 gene is located on the long arm of chromosome 12 at position 24.1 and predicts for a 871 amino acids protein. Within the TRPV family, it shows higher homology with TRPV1 and TRPV3 members, with 40.3 % and 36.3 % of identity and 57 % and 52.7 % of sequence similarity, respectively. By contrast, the homology with other members sensibly decrease, till 25% of identity and 37.5% of similarity with TRPV2.

Like most of TRPs, TRPV4 is expected to arrange in a tetrameric functional channel, and so far, even though it shares high homology with other TRPV sequences, there is no evidence for heteromeric assemblies with other TRPV monomers. In line with the other elements of the sub-class, both N- and C-termini are located in the cell and the transmembrane portion is organized in a six membrane spanning segments (S1–S6), linked by five loops (L1-L5), the last of which, connecting S5 to S6, partially enters the membrane (Figure 5.1). About this segment, the determination of several the bacterial potassium channel structure, namely KcsA [237], KirBac1.1[238] and KirBac3.1[239], has been of great help in better understanding the general structure of these channels, confirming, inter alia, that they presents a four-fold symmetry around the central pore region.

The big intracellular N-terminus represents more than half of the entire TRPV4 sequence, and branches (i) three ankyrin repeat domains (ARD1–3), (ii) four sites predicted to receive a protein kinase C (PKC)–phosphorylation, (iii) a cAMP-dependent phosphorylation site, and (iv) two additional PKC regions located immediately after the ARD3 domain.

With regard to the channel core, the central domain is expected to consist of six transmembrane spanning helices and relatives linking loops, the last of which, between TM5 and TM6, partially re-enters the membrane and together with S5 and S6 is predicted to build up the *pore module* of the channel. As for other members of this class of receptors, TRPV4 transmembrane domain is

predicted to undergo to post-translational glycosylation[240], and a protein kinase C phosphorylation site is predicted to remain immediately after the helix 2, in the intracellular environment.

As mentioned, TRPV4 channel, as TRPVs in general, shows a three-dimensional structure very similar to the bacterial KcsA K^+ channel. Relying on structural information of the latter, the pore loop domain is expected to be made by a series of non polar residues arranged in a short alpha helix followed by few residues more which role is to manage the ion selectivity, operating as a filter. Mutagenesis studies addressed in this region found several residues to affect both selectivity and permeability. Among these residues, Asp672, Asp682 and, although it is not that clear how, Met680 have showed to sensibly weaken calcium channel crossing.

Finally, TRPV4 C-terminus is predicted to starts from around position 712 up to 871 and is thought to bring the docking site on which at least two proteins, namely calmodulin[241] and microfilament-associated protein 7[242] are expected to land on.
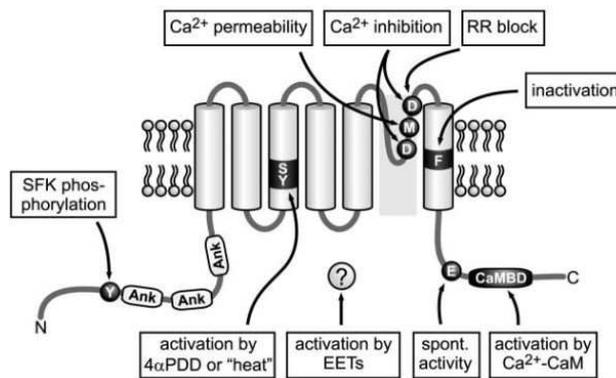
Figure 5.1 Schematization of monomeric TRPV4 structure.

Interesting point such as important residues or sites are highlighted.

(Figure takrn from Plant and Strotmann[243])

5.2 TRPV4 functional features.

Similarly to the highly homologous TRPV1 and TRPV3, TRPV4 has a specific permeability and conductance profile, clearly different from the other vanilloid members, especially concerning Ca^{2+} permeability.

The reversal potential in normal extracellular solutions is positive close to 0 mV. The outward rectification results from a block by extracellular Ca^{2+}. Lowering the extracellular calcium concentrations increases inward currents and leads to a loss of rectification in calcium-free solutions[244].

TRPV4 is a non-selective cation channel, only slightly permeable to calcium, with a preference for sodium of about two times[245]. It is also crossed by other divalent cations, such as magnesium, strontium and barium, with a ratio compared to the sodium of 2–3, 9, 0.7–7, respectively[246,247].

Lower is the discrimination level among monovalent cations, for axample potassium, cesium and litium permeability has proved to range from 0.9 up to 2, respect to the sodium[248].

Unfortunately, all these experimental evidences still have not found a completely clear correlation with the structure of the channel.

5.3 TRPV4 polymodal activation.

Considering its wide expression in many unrelated tissues and the nature of its receptor family, the variety of the physical and chemical stimuli activating TRPV4 comes as no surprise. It likely acts as a polymodal sensor protein with specified functions, depending on the cell type and on the co-expressed proteins pattern profile. Different chemicals stimuli, such as diacylglycerol and phorbol esters as well as arachidonic acid and epoxygenase metabolites, interact with the protein and gate the channel. Besides them, physical stimuli as well, such as mechanic stress, temperature and osmotic pressure, are able to activate it.

With regard to chemical stimuli, the first ligand discovered to activate TRPV4 was the synthetic phorbol ester, 4α-phorbol 12,13-didecanoate (4α-PDD)[249], found to be able to activate TRPV4 at relatively modest concentrations (EC_{50} about 0.2 μM), with good specificity, since so far it was proved not to activate any other TRPV channels.

Since 4α-PDD is a metabolically stable non-protein kinase C–activating ligand, it was speculated that it might bind itself TRPV4. Moreover, recent studies have demonstrated that other 4α-PDD analogues can activate the channel with a potency depending on their apolarity[250].

Initially, experimental studies aimed to investigate the ability of protein kinase C-activating phorbol esters to activate TRPV4 demonstrated that they actually were able to do it, but quite slightly. Only in a second time, performing the same studies at higher temperature, TRPV4 was demonstrated to be potently activated by phorbol esters as well as by mechanical stimuli[251]. In detail, at 37 °C, sub-μM concentrations of phorbol myristate acetate were proved to be enough for activating TRPV4 in a PKC-dependent manner, obtaining an effect similar to that induced by 4α-PDD, and suggesting how TRPV4 channel might be someway strictly related to signaling pathways underlying the induction of PKC.

TRPV4 main binding domain is reasonably expected to be within the four helices sensor module, in line with the other TRPs channels. Moreover, this speculation is further supported by the individuation of a YS motif similar per both nature and position to that one of the known main binding site of TRPV1 channel[252]. In detail, two residues, namely Tyr556 and Ser557, located on the third transmembrane segment, close to the intracellular side, have been so selected and studied[253]. In particular, coherent results came from mutagenesis studies in which changing on Tyr556 was proved to completely cancel the responsiveness to 4αPDD, but not to arachidonic acid, suggesting how diverse paths would underlie the chemical-mediated TRPV4 activation[254].

Beside phorbole esters, arachidonic acid (AA) and related epoxyeicosatrienoic metabolites (EET) generated via the cytochrome P450 epoxygenase pathway, namely 5,6-EET and 8,9-EET, were found to activate TRPV4[255]. This represented the first evidence for endogenous ligands.

Concerning the mechanical stimuli, TRPV4 was firstly cloned while investigating on its capacity to respond to hypo-osmotic cell swelling, behaviour that once confirmed proposed it as a potential mechano-sensitive channel[256], probably acting as controller in all volume related processes.

To support this speculation, several studies have been conducted, for example, by applying fluid shear stress across the apical surface of TRPV4-expressing cells[257], and the resulting TRPV4 activation confirmed its capacity to respond to mechanical stimuli as well, even if this process was proved to be strictly temperature-dependant.

Two principal mechanisms have been proposed to explain how TRPV4 detects and reacts to this kind of stimulation: (i) direct mechanical activation and (ii) indirect activation through other transduction pathways. Indeed, cell swelling leads to the activation of both the phospholipase C (PLC)/diacylglycerol (DAG) and the phospholipase A2 (PLA2)/arachidonic acid (AA) transduction

pathways, able to modulate TRPV4 activation. TRPV4 knockout mice demonstrated that the inhibition of PLA2 or cytochrome P450 epoxygenase activity reduced the swelling-induced activation of TRPV4, while application of the EET metabolites or inhibition of EET hydrolysis enhanced both the swelling and AA-induced activation of TRPV4.

Hypotonicity-induced nociception has also been described as a TRPV4-mediated process[258]. TRPV4 is highly sensitive to changes in extracellular osmolarity. Reductions in the extracellular osmolarity result in increases in $[Ca^{2+}]_i$ and membrane currents, whereas osmolarities above 300 mosmol/L decrease both $[Ca^{2+}]_i$ and currents in those cells displaying spontaneous activity[259].

Beside cell suffering and osmotic pressure, temperature has been related to TRPVs activation, in particular for TRPV1–4 members, which have been shown to be activated by warm to noxious temperatures. It has been demonstrated that temperature dependency may involve a shift of the channel's voltage dependency to more physiologically relevant voltage ranges. Elevating the temperature induces a left shift in the voltage sensitivity, thus activating the channel. Cold-activated TRPs show a complementary right shift in the voltage sensitivity with cooling.

5.4 TRPV4 expression.

TRPV4 is widely express in many different tissues. It is present in kidney[260], in the heart, endothelium, brain, liver, placenta, lung, trachea, and salivary glands[261]. TRPV4 mRNA is also present in airway smooth muscle[262] and in the substantia nigra pars compacta[263].

In the kidney tubules, TRPV4 expression is localized to constitutively or conditionally water impermeable (antidiuretic hormone-sensitive) segments, where it is mainly localized to the basolateral membrane[264]. Moreover, TRPV4 is highly expressed in pancreatic Islet of Langherans, as well as in testis seminiferus ducts and prostate. In other different tissues TRPV4 is still expressed albeit in lower concentrations. In female breast and reproductive system it is expressed in endometrium, ovary and placenta. Lower level are detected into blood and immune system, as in lymphatic and tonsil germinal center cells. TRPV4 has been found in skeletal and smooth muscles, as well as in skin and adipose issues, but with markedly lower abundance.

5.5 TRPV4 physiology.

Its ubiquitous expression combined with its polymodal activation makes TRPV4 an important actor in many unrelated physiological processes. Directly involved in osmoregulation, TRPV4 is expressed in the cerebral circumventricular, which is responsible for the osmotic balance controlling it by levelling water intake/excretion. TRPV4 plays an important role in regulating the osmotic balance in kidney, as well.

TRPV4 is highly present in epidermal keratinocytes, which role is to monitor temperature changes, so contributing to warmth perception. TRPV4 might also play a role in regulating thermogenesis, since it is expressed in neurons of the pre-optic and anterior hypothalamus which are involved in thermoregulation and contain specialized warm- and cool-sensitive neurons also activated by hypo-osmolarity[265].

Activation of TRPV4 after damaging the stratum corneum of the skin accelerates epidermal barrier recovery, probably by stimulating lamellar body exocytosis and intracellular lipid bilayer formation in keratinocytes of the stratum granulosum. Moreover, TRPV4 can act as humidity sensor in the skin by regulating transcutaneous water loss.

5.6 TRPV4 related pathologies.

TRPV4-related channelopathies are directly or indirectly involved in several diseases, specially skeletal dysplasies, such as mild brachyolmia or spondylomethaphyseal dysplasia (SMD). Hence, TRPV4 proved to be a promising target for the pharmacological treatment especially for liver pathologies as well as pancreatic, respiratory and cardiovascular diseases.

Certainly, one of the most interesting pharmacological implication of TRPV4 is in the treatment of the type II diabetes. Located in pancreatic β-cells, known to decrease in number in patients affected by type II diabetes as a consequence of the deposition of amyloid in the pancreatic islets[266], TRPV4 was proved to be activated by physical changes induced by islet amyloid polypeptide (hIAPP), and someway leading to β-cells apoptosis beginning.

Moreover, TRPV4 has recently been proposed as a target for pharmacological treatment of inflammation diseases, such as arthritis, inflammatory bowel diseases and interstitial cystitis, since it is co-expressed with Protease activated receptor 2 (PAR2) within a subset of primary sensory neurons. Activation of PAR2 by several inflammation mediators induces sensitization of TRPV4-

expressing cells to mechanical and osmotic stimuli[267]. Several TRPV4 knockdown animal models have shown that TRPV4 mediates such an increased sensitivity in response to PAR2 activation.

5.7 TRPV4 ligands.

Up to date, many TRPV4 agonists have been developed. This is not surprising when considering the studies performed in the last years to characterize the structure-activity relationships of the well-known TRPV4 agonists such as 4α-PDD and the arachidonic acid metabolites, further supported by exhaustive mutagenesis studies. For example, only to mention a class, several phorbol esters has been developed by Nilius and Appendino groups by exploring substitution on esterification pattern on ring moiety and varying the length of the aliphatic chains[268]. Among many findings, this work led to identify 4α-PDH, with shorter chains made by only 6 carbon atoms, which is a five time more potent agonist than 4α-PDD (EC_{50} equal to 0,07 and 0.37 μM, respectively).

Oppositely, the TRPV4 antagonists are not so abundant. To date, only a relatively limited number of TRPV4 antagonists have been developed, and most of them are relatively unspecific[269]. Thus, the well-known voltage-dependent blockers, ruthenium red (RR), Gd^{3+} and La^{3+}, which inhibit inward but not outward TRPV4 currents, show similar actions on other TRPV channels. Similarly, citral oil from lemongrass blocks transiently TRPV4 in a micromolar range, but also modulates other channels in sensory neurons, such as TRPV1-3, TRPM8 and TRPA1[270].

However, it is worth mentioning a series of molecules patented by GSK (GlaxoSmithKline) which are derivatives based on a diazabicyclo[2,2,1]hept-2-yl chemotype. These compounds show structural similarities to the TRPV4 agonist GSK 1016790A25, and can be regarded as conformationally-constrained analogues of this compound. Unfortunately, no structure-activity relationships have been reported for these TRPV4 antagonists rendering them unamenable for rational correlative studies. Nevertheless, the conserved substructure in common with several agonists suggest that they can bind in the same binding site. This means that these molecules can be used in docking simulations to assess the reliability of a homology model as performed in the here reported final virtual screening study.

5.8 Computational studies: setting the scene.

The above mentioned scarcity of well-characterized TRPV4 antagonists rendered the validation of the here generated TRPV4 homology model, and especially of its ligand binding domain, truly difficult. In particular, the scarcity of known antagonists prevented detailed and validating QSAR analyses to be carried out. Trusting the truly reliable results afforded by fragmental approach for the TRPM8 and TRPM5 members, the same approach was here employed also to generate the TRPV4 homology model which was preliminarily validated by docking a small set of ten known phorbol ester agonists. Although the dataset was not enough to perform QSAR studies, the relevance of the found stabilizing interactions and the agreement of these contacts with the available SAR studies provided a preliminary validation of the TRPV4 homology model. With these results in hand, the last part of this study involved a virtual screening campaign as performed by utilizing a dataset including 130 known antagonists taken from patent literature and 12870 suitably chosen decoy molecules.

5.9 TRPV4 homology modelling and optimization.

At the time of the study, TRPV4 (and no member of TRPV subfamily) was not experimentally resolved. Nevertheless, in march 2014, Liao and co-workers reported determined the three-dimensional structure of the rat TRPV1 as resolved by single particle electron cryo-microscopy (cryo-EM)[271]. This finding, albeit far from the exhaustiveness of the X-ray resolved structures, represents a notable starting point for a better understanding of TRPV channels, and, as described later, a good test bench for the human TRPV1 homology models. However, since the here reported study on TRPV4 was carried out prior of this findings, homology modelling was performed by fragments.

Due to the low homology degree with the available templates, fragmental multi-template approach was, indeed, the preferred modelling strategy, since it allows a more reliable model to be generated, exploiting local homologies. In order to streamline the modelling studies and as successfully done for the TRPM5 model, only the monomeric structure of the TRPV4 was generated, since the study, was focused on the main ligand binding domain, which is located on the lateral surface of the sensor module (S1-S4), in line with the other TRP channels. On these considerations, the modelling of both termini were avoided.

The precise subdivision of the TRPV4 sequence in the structural segments with in particular the identification of the key TM helices was performed by carefully aligning the TRPV4 sequence with that of the other TRPV members. With these alignments in hand, FUGUE was used to generate the starting models for each segment. Due to the already mentioned low homology with the available templates, this modelling study was a really challenging task and some long segments were subdivided into more subfragments and modelled separately. Nevertheless and since the general protocols for both homology modelling and model optimization have already being described, please refer to paragraph 4.6 and 4.7, respectively, for the computational details.

5.10 TRPV4 binding site.

Before starting docking simulations, the so generated TRPV4 monomer was utilized to characterize its binding cavity, analysing the involved residues at the light of what was reported in literature. By analogy to TRPV1 channel, for which a YS motif in the S2–S3 linker was shown to be involved in the binding with its reference agonist, the capsaicin, a highly conserved (and correspondent) YS motif on the S3 helix of TRPV4, namely Tyr556 and Ser557, was hypothesized as the core of the TRPV4 orthosteric ligand binding domain. This consideration was further corroborated by mutagenesis studies showing the Tyr556 mutations induce the loss of channel activation mediated by 4α-PDD.

The so defined region is located in the transmembrane region, close to the intracellular environment, on the opposite side respect to the pore of the channel, similarly to what was observed in the TRPM5 and TRPM8 models. The putative pocket evidently hosts both Tyr556 and Ser557, along with Tyr553 and several apolar residues, such as Phe554, Ile555, Ile561 and Leu584. Analysing the arrangement of the involved residues, two main regions can be clearly individuated. The two tyrosine residues and the serine confer a strong polar nature to a small sub-pocket, beside which, instead, apolar residues line a more hydrophobic and more extended region. The disposition of these residues brings to mind the "*internal/external*" distribution described for the TRPM members (see paragraph 3.5.2).

On these grounds, the phorbol esters are expected to accommodate its polar functions close the YS motif, while the aliphatic chains should be inserted into the hydrophobic region (Figure 5.2).
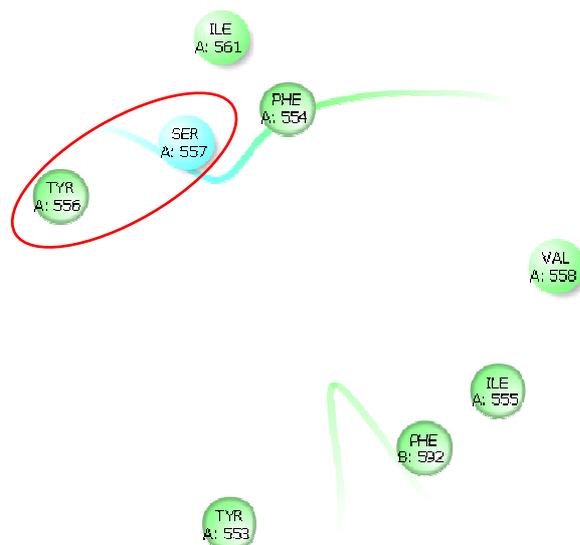
Figure 5.2 TRPV4 ligand binding domain residues disposition.
Key residues are circled in red.

5.11 Docking studies.

As mentioned before, preliminary docking calculations involved a small set of phorbol esters (Table 5.1), which were built using VEGAZZ software. Ligand's conformational analyses have been performed similarly to what was already described (see paragraph 4.9 for the computational details). Once available both protein and ligands, docking calculations were performed in a 10 Å radius sphere centred on the key Tyr556. Considering the successfully docking campaigns performed on TRPM5, the PLANTS docking program was utilized. Ten poses were generated for each ligand and scored by using the most exhaustive ChemPLP function. The so obtained complexes were analyzed considering both the computed scores and the expected contacts. The selected best complexes were then minimized keeping fixed all atoms outside a 10 Å radius sphere centred around the bound ligand and the docking scores were finally recalculated.

In line with the expectations, Figures 5.3-4 clearly show that the parent compound 4α-PDD arranges its polar functions toward the YS motif, establishing two H-bonds with Tyr556 and Tyr553 with the two hydroxyl functions carried by the seven member aliphatic ring. Besides these pivotal interactions, the carbonyl group of one ester tail forms another H-bond with Ser557, confirming the relevance of both YS residues. The complex is further stabilized by several hydrophobic

interactions between Phe554, Ile555, Val558, Ile561 and Leu584 and the two aliphatic tails of the ligand. (Figures 5.3 and 5.4).
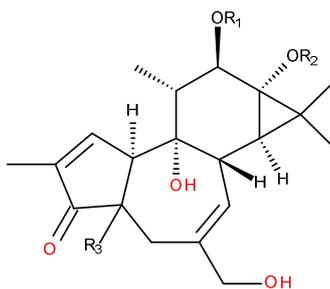
Compound	R1	R2	R3	EC50 (µM)
1 (αPDD)	n-COC$_9$H$_{19}$	n-COC$_9$H$_{19}$	αOH	0.37
2	n-COC$_3$H$_7$	n-COC$_3$H$_7$	αOH	22.47
3 (αPDH)	n-COC$_5$H$_{11}$	n-COC$_5$H$_{11}$	αOH	0.07
4	n-COC$_7$H$_{15}$	n-COC$_7$H$_{15}$	αOH	>50
5	n-COC$_8$H$_{17}$	n-COC$_8$H$_{17}$	αOH	>50
7	n-COC$_{13}$H$_{27}$	n-COC$_{13}$H$_{27}$	αOH	2.83
8	H	n-COC$_9$H$_{19}$	αOH	0.45
6	COCH$_3$	COCH$_3$	αOH	>50
9	n-COC$_9$H$_{19}$	n-COC$_9$H$_{19}$	βOH	0.40
10	n-COC$_9$H$_{19}$	n-COC$_9$H$_{19}$	H	>50

Table 5.1 Chemical structure of the docked 4α-PDD derivatives.
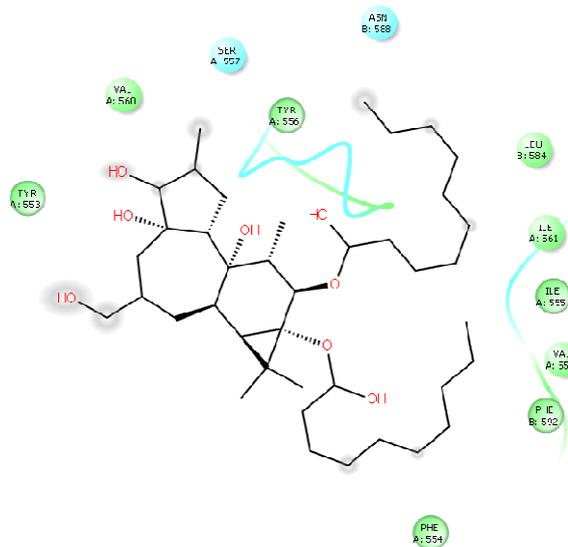
Figure 5.3 Representative scheme of the TRPV4-4α-PDD complex.
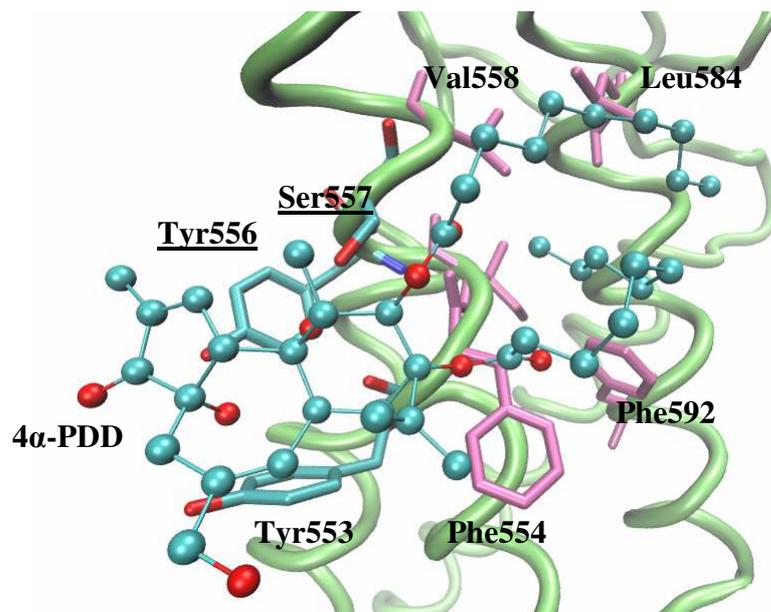
Figure 5.4 TRPV4-4α-PDD complex.

Key Tyrosine and Serine residues are underlined and reported in light blue;

hydrophobic residues are reported in violet.

The docking results for the 4α-PDD derivatives are in clear agreement with the reported activity. In detail, the simulations allow for two main considerations. First, docking results emphasize the key role played by the hydroxyl group in 4 which is involved by two key H-bonds thus explaining why the deoxygenated analogue is inactive. As an aside, docking results confirm that the 4β-PDD epimer is able to rearrange and to retain the key H-bonds also supported by the close carbonyl function. Second, the good activity of compound 8 which bears only one ester tail is clearly understandable considering that docking results suggest that the key role of this ester group is to elicit H-bond with Ser557, a role which is clearly maintained by the simple hydroxyl function. Finally, it is more difficult to rationalize the fluctuating dependence of activity on chain length which appears to be due to steric hindrance and chain fitting and in which the apolar contacts with the mentioned hydrophobic residues play a key role.

Beside agonists, a preliminary docking study has been also conducted with some representative antagonists. In detail, this study involved a set of sulphonamide derivates, among which a dichlorophenyl derivative, RN-1734[272] (IC$_{50}$ 2.3 μM), is here reported in complex with TRPV4 (Figure 5.5) as well as capsazepine, well-known TRPV1 antagonist recently proved to be also a TRPV4 inhibitor (Figure 5.6), and the nanomolar HC-067047 derivative[273] (Figure 5.7).

When compared to above described complexes for agonists, all analyzed antagonists are characterized by their ability to interact only partially with the YS motif, establishing an unique H-bond with only one of the two key residues. More in detail, RN-1734 and HC-067047 contact only Ser557, capsazepine interacts with only Tyr556 while all ligands stabilize π−π stacking with Tyr556. Moreover, all computed complexes are further stabilized by a set of hydrophobic interactions that ligand's hydrophobic groups elicit with aliphatic side chains within the TRPV4 binding site. Analysing the complexes and considering the differences in potency (micromolar for RN-1734 and capsazepine, and nanomolar for HC-067047), one might speculate that the polar and π−π interactions are mandatory for the activity which in turn increases with the number of hydrophobic interactions, which can be reasonably related the size and lipophilicity of the compounds.
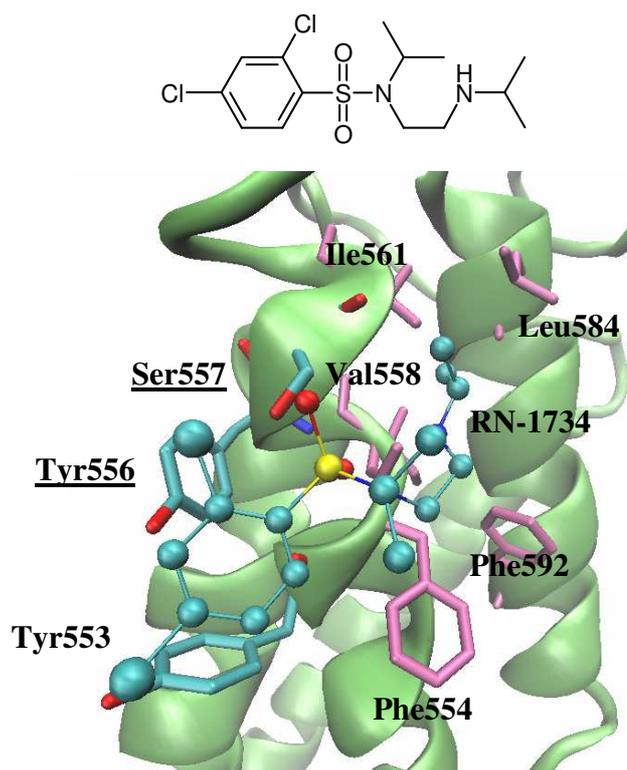
Figure 5.5 RN-1734 structure (upper) and TRPV4- RN-1734 complex (below).
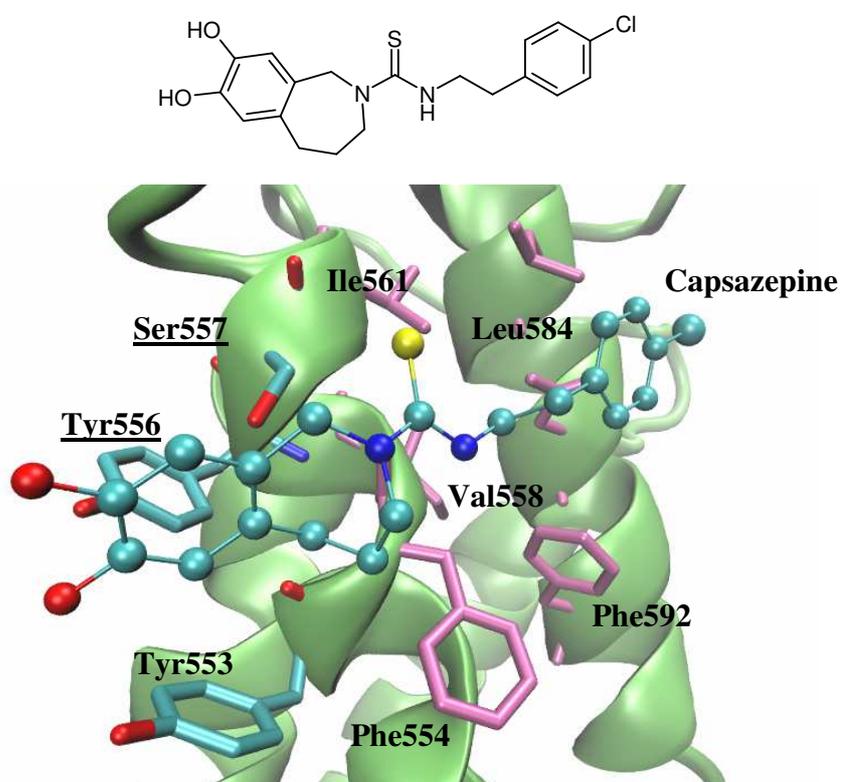
Figure 5.6 Capsazepine structure (upper) and TRPV4-Capsazepine complex (below).
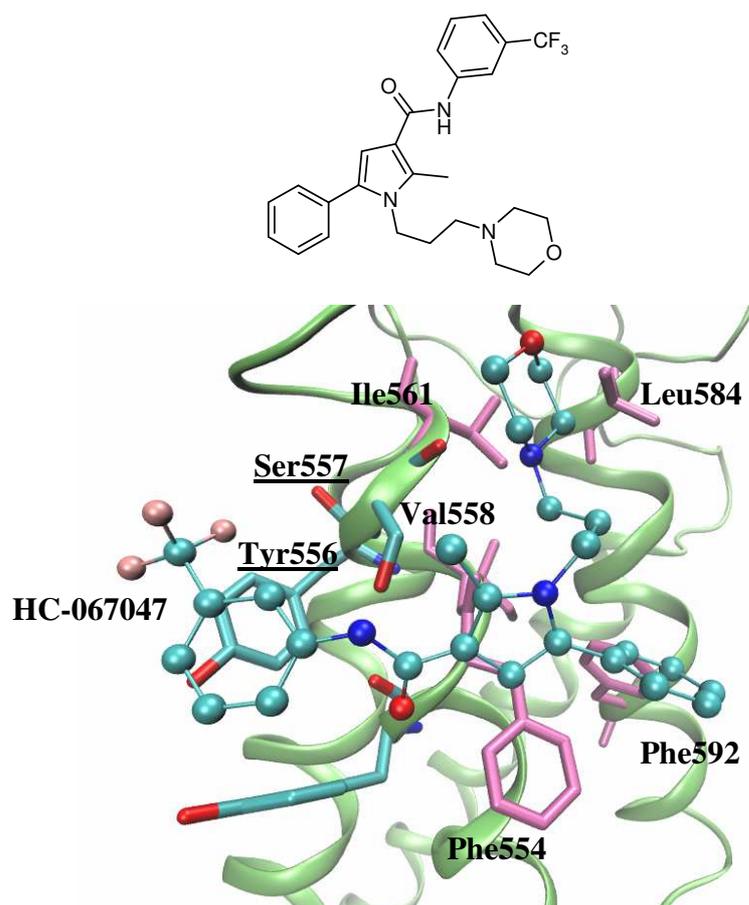
Figure 5.7 HC-067047 structure (upper) and TRPV4- HC-067047 complex (below).

5.12 Virtual screening studies.

Once the TRPV4 ligand binding domain has been analyzed by docking simulations and similarly to what was already done for the previous TRPM members, the successive computational study involved a virtual screening campaign, performed in order to further validate the reliability of the model, by testing its capacity to discriminate between active and inactive compounds.

As already detailed in paragraph 3.2.8, to which the reader is invited to refer for all the computational details, a library of 130 structurally different TRPV4 antagonists was collected by selecting compounds from several patents, mostly acyclic 1,3- and 1,4-diamine derivatives from US0259965A1[274] and US0105259A1[275], as well as amino azepine derivatives from US0293477A1[276] and US0124619A1[277].

As already mentioned, the virtual screening strategy requires that these 130 actives are dispersed among inactive molecules, which also in this case come from the inactive ligands of the previously described HTS study on TRPM8. In order not to bias the results, putative decoys were selected if they fulfilled several physicochemical property ranges, which were determined by analysing the active compounds.

Once again, we had to speculate that TRPM8 decoys would also be inactive toward TRPV4, a choice that is indeed justified by the lack of available data about TRPV4 inactive compounds. In order the guarantee the statistical significance of the study, the 130 active ligands were jointed to 12870 decoys, obtaining a 13000 molecule dataset in which the percentage of active molecules correspond to 1%.

Docking calculations were, then, performed using PLANTS, testing all three scoring functions, in order to maximize its performances. Calculations were focused on the previously described binding site, namely in a 10 Å radius sphere, centred on Tyr556 and had the same docking characteristics already described in paragraph 4.9.

For every simulations, the generated complexes were scored according to different functions and ranked from best to worse. Indeed and besides the Total Score value, three normalized values of each scoring functions were considered. In detail, the total scores were normalized per (i) the rotatable bonds, (ii) the cubic root of molecular weight, (iii) the cubic root of heavy atoms. The cubic root function is applied to obtain a smoothed normalization thus avoiding an excessive overestimation of the smallest molecules, while the normalization per rotatable bonds is an indirect way to account for entropic factors by penalizing too flexible ligands. Virtual screening performances were, than, evaluated by calculating the EF and the Skewness values.

The results for the all three scoring function are reported in Table 5.2, which shows the EF values as calculated in the top 1, 2, 5, 10 and 20 % of the set, by ranking molecules according to all three Total Scores and their mentioned normalized values. While the normalized scores reveal worst EF values, the obtained results afford an overall encouraging validation of the predictive power of the here modelled TRPV4 model. Best performing score is the PLP simple Total score, which performances are reported in detail in Table 5.2. Clearly appreciable even by the histogram plot itself, the screening produced really remarkable results (Table 5.3, Figure 5.8). In more detail, the best performing results are especially obtained within the smallest top ranked clusters, namely 1, 2 and 5% of set. For example in the Top 5% the screening places more than a half of the total actives, reaching the impressive EF values of almost 13.5 in the top 2%. Even if only a very small part of the screened molecules are effectively promoted to the experimental stages, the results obtained in

the Top 1% of the distribution mean that it is possible to decuple the odds to find an active compounds respect to the random. The goodness of the screening is further supported by the remarkable Skewness value of 3.5, which emphasize the predictive power of both homology model and docking protocol as assessed in the entire distribution.

With these gratifying results in hand, the model as well as the entire screening strategy could be reasonably exploited to screen large databases, looking for new hits.

ChemPLP	Total score	Score$_{Rotors}$	Score$_{HA}$	Score$_{MW}$
1	3.84	2.30	0.76	0.76
2	4.23	1.53	0.38	0.76
5	4.15	2.76	1.07	1.23
10	2.76	2.53	1.69	2.15
20	2.2	1.96	1.46	1.42
PLP	Total score	Score$_{Rotors}$	Score$_{HA}$	Score$_{MW}$
1	10.76	8.46	3.07	3.84
2	13.46	9.23	2.69	3.07
5	10.3	7.53	3.84	2.15
10	6.76	4.6	2	2.3
20	3.8	2.92	1.46	1.69
PLP95	Total score	Score$_{Rotors}$	Score$_{HA}$	Score$_{MW}$
1	6.92	2.30	1.53	2.30
2	8.07	2.69	1.92	3.07
5	7.38	3.23	2.76	2.61
10	6.38	3.00	2.84	3.53
20	3.76	2.38	2.30	2.73

Table 5.2 Virtual screening evaluation for ChemPLP, PLP, PLP95 scoring function and their wheightings. EF reported are calculated in the top 1, 2, 5, 10 and 20 % of the set.

n (top %N)	Cpds	Act	EF	Sk
1	130	14 (1.3)	10.76	
2	260	35 (2.6)	13.46	
5	650	67 (6.5)	10.3	3.5
10	1300	88 (13)	6.76	
20	2600	99 (26)	3.80	

Table 5.3 Virtual screening evaluation, molecules ranked per PLP Total score.
"Cpds" reports the number of molecules hosted in the considered subset; "Act" stays for the
actives compared to the expected one in parenthesis; EF reports the Enrichement factor values;
"Sk" the Skewness value of the entire distribution.
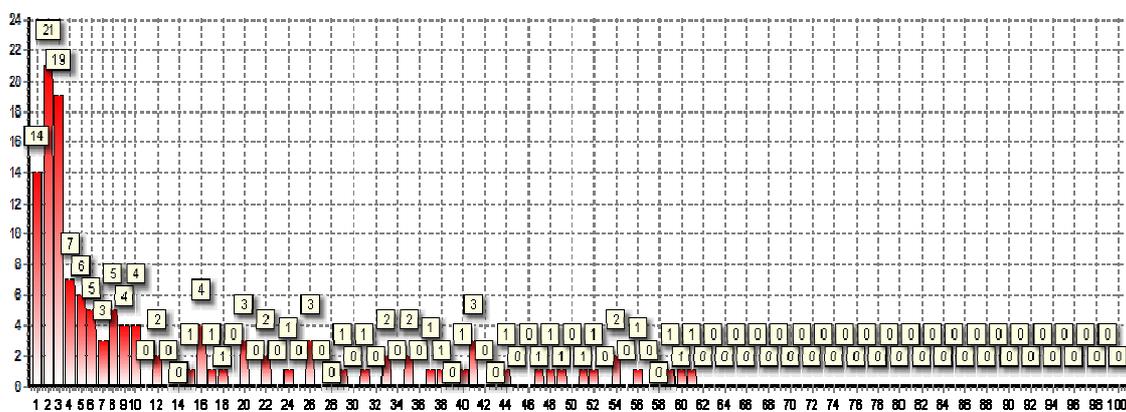
Figure 5.8 Virtual screening ranking per PLP Total score.
The histogram plot reports the number of actives per cluster.

6. Computational studies concerning the hTRPV1 ion channel

6.1 TRPV1 structural features.

TRPV1 was cloned in 1997 by Caterina and co-workers within a research project on somatic sensory and pain[278]. This finding represented a significant improvement for the understanding of molecular mechanisms involved in the transduction of noxious thermal and chemical stimuli by sensory neurons. Originally described as a receptor for capsaicin and related natural irritants (the so called vanilloids), TRPV1 is indeed involved in detecting different noxious stimuli, such as heat, acids, negatively charged pollutants and endogenous pro-inflammatory molecules.

Located on chromosome 17p13, hTRPV1 gene encodes a 839 amino acids protein. Beside the complete channel, two splice variants lacking residues in the N-terminus domain are known, namely TRPV1b and VR.5'sv[279]. The former protein shows no sensitivity to heat, low pH or capsaicin and was proved to interact with TRPV1 channel, so acting as a down-regulator[280]. Similarly, the second variant does not respond to TRPV1 agonists but interacts with TRPV1 and definitively inhibits the gating[281].

In line with other TRP channels, TRPV1 is homologous to the voltage gated potassium channels (Kv), with which it is expected to share the three-dimensional arrangement, made by N- and C-termini located in the cellular environment, and a transmembrane domain composed of six segments (S1-S6), with a pore-loop defined by S5-S6 and relative linking loop. TRPV1 as well assembles in a functionally active homotetrameric channel.

Like other TRPV channels, the N-terminus contains six ankyrin repeats, whose structure has been determined via high resolution x-ray crystallography[282]. Similarly to what already described for TRPV4, the pore loop is expected to arrange in a short helix and a selectivity filter, partially entering the plasma membrane[283]. Here, several residues have been individuated as important for the response to diverse stimuli, as discussed later; phosphorylation sites as well have been found along the intracellular loops as well as in the S5-6 segments.

A motif common in all TRPs channels, namely the TRP box, is predicted to locate in the C-terminus region, where is thought to arrange in an α-helix parallel and close to the membrane, on the intracellular side.

As already mentioned in the previous chapter, an important improvement has been recently achieved in the knowledge of TRPV structure, in general, and more specifically for the TRPV1

member, with the resolution of the rat TRPV1, via electron cryo-microscopy technique with a 3.4 Å resolution [284]. While confirming many predicted structural elements, such as the TRP box location and arrangement or the spatial architecture of the first four ankyrin repeats, this study provided novel information, such as, for example, that the last two ankyrin domain are actually involved in channel assembly, or that several sub-region of the N-terminus are involved in connecting the ankyrin repeats to the segment S1, which in turn contacts the last two ankyrin on the following monomer, further stabilizing the tetramer.

6.2 TRPV1 functional features.

TRPV1 channel shows no selectivity among monovalent cations, resulting almost equally permeable to sodium, potassium, cesium, lithium and rubidium, while its permeability to divalent cations appears to be much higher, as reported by Ahern and colleagues studies, in which capsaicin mediated TRPV1 gating showed for calcium and magnesium values five and ten times bigger than sodium permeability, respectively; similar ratio was reported for heat-mediated channel activation[285].

Protons were proved to enter the cell easily passing through TRPV1 channel[286], that it is also permeable to polyvalent cations as well, albeit a little less, via a "pore dilation" mechanism[10].

So far, diverse mutagenesis studies attested the involvement of several amino acids in the process of ion selection. Clearly, most of them are located in the pore-forming loop, but others are surprisingly located in the transmembrane helices, specially in the last segment.

About the pore region, there is a well conserved sequence motif "TIGMG" which is typical of TRPV channels and is highly similar even to that of the bacterial KcsA channel, that is "TVGYG"[287]. Beside this sequence, other amino acids are involved in regulating both cations passing and channel gating. As a general rule, neutralization of acid residues or replacement of neutral residues with basic ones induce a reduction of permeability as seen in D646N, E648A or Y671S, the last one leading to a loss of calcium-mediated desensitisation[288].

Deeper in the membrane, Ile679 plays a crucial role in permeation, as proposed by electron cryo-microscopy, which showed that it represents the narrowest point along the ion path. Comparisons of the position of closed and open states reveals that Ile679 forms a hydrophobic seal able to impede the permeation of hydrated ions when the channel is close, but, at the same time, rearrangements occurring in the activated state remove this cap, allowing the ion permeation.

6.3 TRPV1 polymodal activation.

TRPV1 receptors are activated by vanilloids like capsaicin, inducing the influx of calcium and sodium, and thus depolarizing the cell. At very ineffective low concentrations, capsaicin can however sensitize TRPV1 receptors to protons and heat. Similarly, protons sensitize TRPV1 to capsaicin and heat.

Vanilloids binding site is expected to be in transmembrane sensor module, far from the pore channel, and this is supported by the fact that, when added extracellularly, membrane-nonpermeant derivatives of capsaicin are inactive. Moreover, mutagenesis studies revealed that three residues, namely Tyr511, Ser512 and Tyr550, are crucial for the capsaicin activation. In detail, mutations of one of the first two residues abolished capsaicin response, but not activation by heat or protons, while mutating Tyr550 only reduces the capsaicin sensitivity of the channel[289].

TRPV1 is a major heat sensor in peripheral sensory neurons that innervate the body. Thermal activation of these neurons shows a temperature threshold of about 42 °C although this can be lower during inflammation. The TRPV1 regions responsible for heat activation have not been yet completely established. Mutational studies involving pore region and C- or N-termini show that all these regions influence the temperature responses of TRPV1.

TRPV1 is known to act not only in thermal-related pain conditions, but it maintains a low activity at physiological temperature as well. This role is fundamental for body thermal regulation, as pointed out in clinical studies in which diverse TRPV1 blockers were related to high fever side effect[290]. Moreover, at room temperature and physiological pH, TRPV1 acts as a voltage-gated outwardly rectifying channel, resulting activated by depolarizing voltage itself[291].

In addition, TRPV1 sensitivity to heat and capsaicin is modulated by variations in proton concentrations within the range of pH 6–9[292,293]. The potentiating effect of pH on capsaicin activation is related to both an increased affinity for capsaicin and an enhanced channel gating. Generally speaking, the effect of protons on TRPV1 can be explained considering the neutralization effect of the acidic residues lining the pore. Thus, Glu648, located between the selectivity filter and the S6 mediates the agonist effect of protons, without affecting pH modulation[27].

6.4 TRPV1 expression.

TRPV1-expressing tissues can be divided into two categories: neuronal and non neuronal tissues.

For what concern the first group, TRPV1 has been detected in the most part of both somatic and visceral sensory neurons and only in small/medium size fibers of dorsal root, trigeminal and vagal ganglia[1]. These neurons form non- or thinly-myelinated branches, named C or Aδ fibres, directed toward almost all organs in the human body.

TRPV1 is also present within the CNS, particularly expressed in the preoptic area of the hypothalamus, where it acts as body temperature downregulator[294]. Moreover, TRPV1 also intervenes in generating postsynaptic long-term depression when activated by endogenous anandamide, specially in the nucleus accumbens[295].

Besides neurons, TRPV1 is additionally expressed in diverse non-neuronal cells, where its activation induces a rapid rise in intracellular Ca^{2+} levels. Among these tissues, one may mention the endothelium, the immune cells (lymphocytes, dendritic cells and mast cells), keratinocytes, the smooth muscle cells and the urothelium, thymocytes and macrophages.

In more detail, TRPV1 activation in perivascular sensory fibers was attested to reflect into neuropeptide CGRP- mediated vasorelaxation of resistance arteries[296]. TRPV1 expression was also detected in arteriolar smooth muscle in skeletal and thermoregulatory tissues[30], where it is speculated to manage the blood flow with a still not completely understood mechanism.

Taste buds are surrounded by TRPV1 containing peptidergic nociceptors from the trigeminal nerve that are responsible for the burning sensation typically induced by capsaicin containing foods by a calcium mediated transmitter release. Once reached the brain, this information can be interpreted as reward or threat and so the substance can be ingested or rejected. TRPV1 expression has also been reported in bladder urothelium, where the it regulates bladder contractions by mediating urothelial ATP release in response to stretch[297]. Finally, several studies have reported expression of functional TRPV1 in human keratinocytes[298].

6.5 TRPV1 physiology.

One of the most important implications of TRPV1 channel in physiological processes is, surely, the detection of temperature increasing. In sensory peripheral neurons, it is known to transduce not only

thermal nociception due to noxious temperature (reasonably intended major than 42 °C) but physicochemical stimuli as well, generating the perception of local 'burning' heat.

Moreover, TRPV1 has a remarkable physiopathological role in inflamed tissues, in which even lower temperatures were proved to elicit a painful sensation (so called thermal hyperalgesia), as demonstrated in a knockout animals based study, in which mice lacking TRPV1 did not exhibit the expected sensitization to warm temperatures exposure of injured or inflamed tissues[299].

Beside noxious heat sensing, both central and peripheral TRPV1 expressing neuronal tissues have recently been implicated in controlling the physiological body temperature too.

Pain sensation is augmented by the acidic extracellular pH during ischemia or inflammation. On these grounds, it has been found that protons only activate the channels when they are added from the extracellular solution, suggesting that the proton-active gate is on the extracellular part of the molecule[15]. Lowering the pH causes sigmoidal increases in the current. This current has a reversal potential close to zero millivolts, suggesting that, at neutral pH, protons do not carry much of the current in buffers containing physiological concentrations of sodium and calcium.

Finally, TRPV1 is thought to intervene in the transduction process of osmotic information in the CNS. Hyperosmotic-induced cell volume diminution leading to the augmentation of the membrane conductance was proved to be balanced by ruthenium red in functional TRPV1 expressing mice; while, contrarily, it was not detected in knockout animals[300].

In addition, in TRPV1 knockout mice, the consume of water under systemic hyperosmotic challenge was attested to be much less than in the normal animals, so introducing TRPV1 in thirst induction process too[301].

6.6 TRPV1 ligands.

Besides CAPS.like molecules, TRPV1 was proved to be both stimulated or silenced by diverse amphiphilic compounds. Among these, can be mentioned the resiniferatoxin, a highly irritant diterpene twenty time more potent than CAPS[302], quinazolinone[303], evodiamine[304] and 17-β-estradiol[305], as well as long acyl and amide chains bringing chemicals, such as anandamide[306], olvanil and omega-3 polyunsaturated fatty acids are known to gate TRPV1[307]. Other natural TRPV1 agonists are 12-hydroperoxyeicosatetraenoic acid (12-HPETE) and N-arachidonoyl dopamine (NADA)[308].

Piperine from black pepper, eugenol from loves and zingerone from ginger have also been shown to activate TRPV1[309]. Moreover, gingerols from raw ginger, and shogaols, possessing a vanillyl moiety, can also activate the channel.

Considering the various anatomical/physiological contests in which TRPV1 channel is implicated, the possibility of making it a target for the pharmaceutical treatment of several important pathological conditions clearly appears as a great goal.

To date, two opposite approaches are mainly walked for antagonize TRPV1 channel in diseases: (i) an indirect approach, which involves TRPV1 agonists and aims to the activation/desensitization of the channel, and (ii) a classical direct approach involving TRPV1 antagonist.

The first method lies of the fact that, somewhat surprisingly, capsaicin is able to attenuate painful sensations, long desensitizing TRPV1 after the brief activation. According to this, it is used, for example, for treating osteoarthritis and rheumatoid arthritis as well peripheral neuropathy pain.

About the second method, capsazepine, setup by modifying the chemical structure of the agonist capsaicin, is the first reported TRPV1 antagonist, even if later was proved to be not selective.

As a competitive antagonist, it locates in the capsaicin-binding site and thus impedes the channel activation. Unfortunately, capsazepine was not considered a good candidate for therapeutic use, because not selective, metabolically instable and affected by a bad pharmacokinetic profile.

Nevertheless, the SAR of capsazepine was used as starting point for developing new generation of TRPV1 blockers, that can be broadly divided in (i) competitive or (ii) non-competitive antagonists.

The first class is the most studied and populated one. It accounts for compounds that binds to the agonist binding site and locks the TRPV1 channel in a closed state. The most part of competitive antagonists developed so far can be divided into two sub-classes, (a) classic and (b) non-classic antagonists. The first class hosts molecules which have in common a carbonyl group that acts as H-bonding donor or acceptor, brought as thiourea, urea, cinnamide, ester or amide. This class of molecules we have mostly been interested in, in particular, we explored the pattern of interactions with TRPV1 binding site of a series of thiadiazole derivatives as well as several propanamides derivatives, reported to effect with nanomolar activity (see later).

On the other side, non-classic TRPV1 antagonists brings a carbonyl group which is either present as a part of heterocyclic ring or is unrecognizable. Among them the quinazoline or benzimidazole analogues appear structurally different from capsazepine but still retain the key binding elements.

Finally, non-competitive antagonists have been individuated as well. They are pore blockers binding allosteric sites. To date, this class is much more attractive from the therapeutical point of view, since these molecules would act specially against pathologically-over-activated TRPV1

channels, minimizing the rise of side effects. About these, so far, really few molecules have been improved in terms of selectivity, aspect that afflicts the most part of the purified toxins from which drug development of this class of molecules usually starts.

6.7 TRPV1 homology modeling.

At the time I started working on human TRPV1 channel, immediately following the previously described TRPV4 study, the mentioned experimental resolution of the TRPV1 structure by cryo-EM technique was not yet disclosed. So, once again, homology techniques represented the only available approach to model the TRPV1 three-dimensional structure.

However, here a different strategy was applied relying on the high sequence homology between hTRPV1 and hTRPV4, namely 40.3% of identity and 57.0% of similarity, the latter even higher when considering only the transmembrane region (67% of similarity). Indeed the previously generated homology models (TRPM5 and TRPV4) were constructed by using a fragmental approach which gave rather encouraging results (as discussed in the previous chapters) but required a significant amount of computational effort which renders such a strategy unsuitable when the study would require the generation of many receptor models. On the other hand, the study performed on TRPM8 ligand's selectivity indicated that the use of a homology model as the template to generate other homology models, while being questionable on a theoretical basis can afford encouraging results at least when there is a notable homology between the two involved proteins. Clearly, the study performed on TRPM8 was a first attempt which deserves additional confirmations. In this context, the TRPV1 appears to be an interesting case study, considering the really encouraging results obtained by the here presented TRPV4 homology model which should allow a proper modeling of the TRPV1. Hence and while considering that using a homology model as the template may generate an output model which is affected by an excessive and risky similarity with the template inheriting its structural weakness and uncertainties, an overall homology approach using TRPV4 as the template was adopted

Therefore, MODELLER 9.14 was here utilized, since, as described in chapter 2, it allows reliable models to be generated when used in robust homology conditions. In more detail, the transmembrane portion of the TRPV1 primary sequence has been modelled. Alignment first, then model generation were performed with default settings and ten output homology models were produced, in order to consider different solutions among which the most reliable one can be chosen.

As expected, the output models were rather similar, showing the predicted overall folding in agreement with secondary structure predicted by sequence alignment, a reliable arrangement of the sensor module and in particular of the key residues without relevant unpredicted gaps. Based on these remarkable preliminary evaluations, the choice of the most reliable models was carried out by considering structural parameters as well as MODELLER score functions. Indeed and for each generated output model MODELLER computes two scoring functions, namely DOPE and GA341. In detail, DOPE assessment score parameterizes the model conformational energy as calculated by the program during the model refinement and thus the lower the better. It is a not absolute score, since it can only be used to rank models calculated from the same alignment. GA341 assessment score goes from 0 (worst) to 1 (native-like), and so the higher the better. It is a transferable score which can be used to compare models obtained by different calculations. For this reason, GA341 should be preferred to evaluate the model reliability considering its general meaning.

Besides MODELLER scores, the percentage of residues falling in the allowed regions of the Ramachandran Plot and of the side chain χ-space were evaluated to describe the correct conformation of both the backbone and the side chains. For all models, χ-space percentage values were higher than 95% (Table 6.1), and thus all models can be seen reasonably satisfying to be concern the conformation of the side-chains.

With these data in hands, a careful model evaluation was performed in order to select the structures with which docking simulations can be performed. Indeed and considering the potential weakness above described for the here exploited homology modeling, focusing all docking simulations on only one generated model (even though the seemingly most reliable one) would represent a tremendous approximation able to undermine the reliability of all following simulations. On the other hand and at the light of the general similarity among the output models, running docking calculations on all the ten models would be uselessly onerous and probably redundant. On these grounds, a three step selection strategy aimed at progressively reducing the number of the filtered models was setup. In detail, (i) the first filter was based on structural quality and MODELLER scores and allows five models to be selected, then (ii) the second step was based according to similarity analysis among the before selected models in order to individuate the three most representative structures, as evaluated by a pair-wise RMSD analysis (Table 6.2), and finally (iii) the last step involved docking calculations on a representative set of 30 known antagonists and was based on the correlations between the resulting docking scores as computed for each generated complex and the experimentally activity values. This last step allows the selection of the best

performing TRPV1 homology model on which the following docking simulations will be performed.

As described above for χ-space percentage values, also Ramachandran plots, while showing a greater variability, reveal rather similar and satisfying results for all generated models, as seen in Table 6.1. Hence, the first filter was based on the MODELLER scores and in particular the selection was based on GA341 score choosing the five models with highest GA341 values (namely model 1, 3, 5, 8, 10). Gratifyingly the selected models revealed also very low DOPE score values and remarkable Ramachandran plots. Thus, models 1, 3, 5, 8, 10 have been selected and compared in terms of similarity by a pair-wise RMS analysis as performed by matching backbone atoms only (see table 6.2). By using this similarity analysis, three models were selected by considering (a) the structure with the lowest RMSD mean (model 5) which should represent a sort of average structure accounting for most structural diversity exhibited by the compared models; (b) the two structures with the highest RMSD means which should represent two so different structures as not to be properly represented by the previous average model (models 3 and 10). Fortunately, the comparison of models 3 and 10 shows a very high RMSD value, thus suggesting that they effectively should represent two relevant (and not redundant) extreme structures in this comparison.

The so selected three TRPV1 homology models were completed by adding hydrogen atoms and carefully checked to avoid (and fix) unphysical situations. Next, they were refined by an the optimization phase, that was pursuit by the same three steps strategy already used for the previous models, combining energy minimizations with suitably targeted MD simulations. Please refer to paragraphs 4.6 and 4.7 for all the computational details.

Output model	DOPE score	GA341 score	Ram. Plot	χ space
1	-26894.78	0.78807	80.39	98.04
2	-27244.13	0.56092	81.18	95.29
3	-27002.02	0.55980	81.96	97.82
4	-27066.26	0.52961	83.53	97.82
5	-26898.42	0.74333	80.39	96.71
6	-27113.99	0.51953	78.40	96.20
7	-27290.32	0.49000	77.90	98.66
8	-27667.22	0.69363	81.40	96.20
9	-26843.68	0.47999	82.24	95.87
10	-26982.94	0.67185	79.97	97.90

Table 6.1 Homology models structural analyses.

	Model 1	Model 3	Model 5	Model 8	Model 10
Model 1	--	--	--	--	--
Model 3	2.71	--	--	--	--
Model 5	1.45	1.78	--	--	--
Model 8	1.37	1.98	2.25	--	--
Model 10	2.23	2.45	1.78	2.53	--
Means	1.94	2.23	1.82	2.03	2.25

Table 6.2 Pair-wise RMSD analyses; values are expressed in Armstrong.

6.8 Preliminary docking simulations.

The so refined three homology models were finally tested and filtered by assessing the capability of their ligand binding domains to correctly interact with a set of known antagonists.

To this purpose, a set of thirty CAPS-like antagonists was generated by collecting known molecules from literature and patents, including 23 heterocyclic molecules from the patent EP 2604599 A1[310], and 7 1,3,4-thiadiazole heterocyclic derivatives from the recent study of Rebolledo and co-workers[311], all having IC_{50} values within the micromolar range. The ligand's conformational profile was explored by using the VEGA ZZ software, as described in paragraph 4.9.

Relying oh the remarkable results obtained in the previous docking studies and considering the high homology between TPRV4 and TRPV1, docking calculations were also here performed by PLANTS, using a computational protocol similar to that applied for the previous models and detailed in paragraph 4.9. In detail, calculations were focused on a 12 Å radius sphere centred around the centroid defined by the Cα atoms of the key residues Tyr511 and Ser512. So, one pose per compound was generated and scored by all three implemented functions. The obtained complexes were evaluated without any post-docking refinement, in line with the idea to optimize a docking strategy to be applied in virtual screening campaigns, in which approaches involving post-docking selection/optimization/rescoring procedures are clearly unfeasible. Besides the three PLANTS total scores also the already mentioned normalized values were considered, in order to harmonize the scores to the set of molecules in use.

With these data in hand, a comparison of the correlations between the experimental IC50 values and the computed scores for the three selected models was performed. All resulting correlative values (expressed by the Pearson coefficient, r) are reported in table 6.3.

ChemPLP	Total score	Score$_{Rotors}$	Score$_{HA}$	Score$_{MW}$
Model 3	-0.22	0.11	-0.46	-0.24
Model 5	-0.33	0.09	-0.53	-0.33
Model 10	-0.05	0.14	-0.34	-0.15
PLP	Total score	Score$_{Rotors}$	Score$_{HA}$	Score$_{MW}$
Model 3	-0.17	0.12	-0.44	-0.22
Model 5	-0.62	0.06	-0.62	-0.44
Model 10	-0.12	0.13	-0.34	-0.17
PLP95	Total score	Score$_{Rotors}$	Score$_{HA}$	Score$_{MW}$
Model 3	-0.44	0.08	-0.55	-0.36
Model 5	-0.49	0.07	-0.62	-0.43
Model 10	-0.24	0.13	-0.45	-0.25

Table 6.3 Pearson coefficient r values reported are calculated for model 3,8 and 10, correlating compound activity with ChemPLP, PLP, PLP95 scoring function and their weightings.

Table 6.3 clearly shows that model 5 affords the best results reaching quite remarkable r values, almost independently from the score used, and giving its best correlation with PLP Total score. Apart from the normalized Score$_{Rotors}$ values which show unsatisfactory results for all score functions and all models, there are no significant differences between the correlations produced by total and normalized scores (which are even better in some cases) thus indicating that the reported results are not markedly biased by structural differences within the considered dataset. It is worth mentioning that the model affording the best docking results is the model chosen in the similarity analysis since it showed the lowest RMSD mean value and thus it should represent a sort of average of the compared models. This results may have a general meaning suggesting that the model with lowest RMSD mean effectively represents a structural centroid able to account for most structural variability and as such able to properly accommodate most ligands. The obtained results are particularly interesting especially when considering both the number of simulated compounds as

well as the roughness of the unrefined complexes, for which it is reasonable to expect notable improvements if refined by post-docking optimization strategies. On these grounds, the Model 5 was selected and promoted to the following docking simulations.

6.9 Detailed docking simulations.

With a definitive model in hand, a deeper docking analysis was performed by considering visually inspected and refined complexes in order to collect a set of docking-based descriptors to be involved in multiparametric QSAR studies. The so derived correlative equation would provide both a further validation of the reliability of the homology model and a useful tool for unknown molecules in provisional studies.

In detail, the mentioned visual filtering is based on a detailed analysis of the ligand domain, whose main features come from mutational studied reported in literature. The key YS motif, namely Tyr511 and Ser512, is located on the segment S3, decentred toward the intracellular environment. It exposes the side chains on the external surface of the S1-S4 bundle, toward a flank of the sensor module. As for all other studied TRPs, the binding region is actually more an exposed surface than a real pocket, on which the ligands would land rather than fit. Besides the YS motif, several other residues characterize the binding site, and are arranged parallel to the membrane giving it a "C" shape, with the open side toward the center of the channel. On the rim, three polar residues, namely Tyr487, Ser510 and Glu513, delimit the left side of the binding region, conferring it a rather polar character. In the other regions, the binding site is mostly characterized by hydrophobic residues such as Met514, Phe507 and Ile564 and appears to be markedly extended especially in the right side.

Interestingly, the Thr550 residue which proved to be important, albeit not crucial, for the activity of capsaicin by the mutagenesis studies, is located on the S4 segment, a bit far from the YS motif, with a distance (about 14 Å between both hydroxyl functions) which does not completely exclude it from the interacting site. In detail, it is possible to delineate a region along the S4 helix and lined by Tyr554 and Tyr555, which connects the YS motif to Thr550, a region in which ligands can be stabilized by intermolecular stacking interactions.

In order to deeper investigate these structural features, a pilot docking was performed using the reference agonist, Capsaicin. Fifty different poses were generated in order to exhaustively explore all binding modes permitted by such a binding site.

Docking calculations were performed by PLANTS, focused on the same region already selected for the previous docking simulations, and the same settings were applied. All generated poses were clustered and the so identified different binding modes were visually inspected.

As expected, in most poses capsaicin fitted a binding site perpendicularly to the membrane axis, orienting its hydroxy-3-methoxyphenylic moiety toward the Ser512, with which hydroxyl group establishes a first H-bond, while the amide function elicits a second H-bond with the hydroxyl group of Tyr511 which in turn stabilizes π–π stacking with the ligand aromatic ring along with Phe507. Finally the carbon chain stabilizes apolar contacts with Met 514 as well as with the carbon skeleton of the side chain of Arg447, while the ligand double bond can elicit weak π–π stacking with the aromatic moiety of Tyr554 and Tyr555 which also contribute with apolar contacts (Figure 6.1).

Interestingly, a second binding pose, almost equally scored and significantly represented among the computed poses, was individuated. In more detail, Capsaicin is arranged parallel to the membrane axis and, despite being differently harboured, the YS motif are engaged in interactions almost superimposable to that already seen for the previous binding mode. The aliphatic chain is inserted between Tyr554 and Tyr555 with which it can stabilize the already described π–π stacking, and approaches Thr550, without in fact contacting it (Figure 6.2).

In order to further investigate the described dual binding modes, the already used thirty CAPS-like antagonists set was re-docked in the ligand binding domain, producing 20 poses per ligand by still using the best performing PLP score. The so generated complexes were then visually analysed and, interestingly, the two major binding modes already seen for capsaicin were maintained for all the tested antagonists, thus suggesting that this dual binding mode involves all ligands regardless of their intrinsic activity. Hence, the best complexes of each binding mode (henceforth named BM1 and BM2) were selected for each compound, generating two sets of 30 complexes. As already done in the previous filtering step, correlative studies were performed to compare the two binding modes, thus selecting that offering the better performances at least for the considered classes of molecules.

Table 6.4 reports the Pearson coefficient r values obtained correlating the PLP docking scores with the experimental activity IC50 values, for the two monitored binding mode.
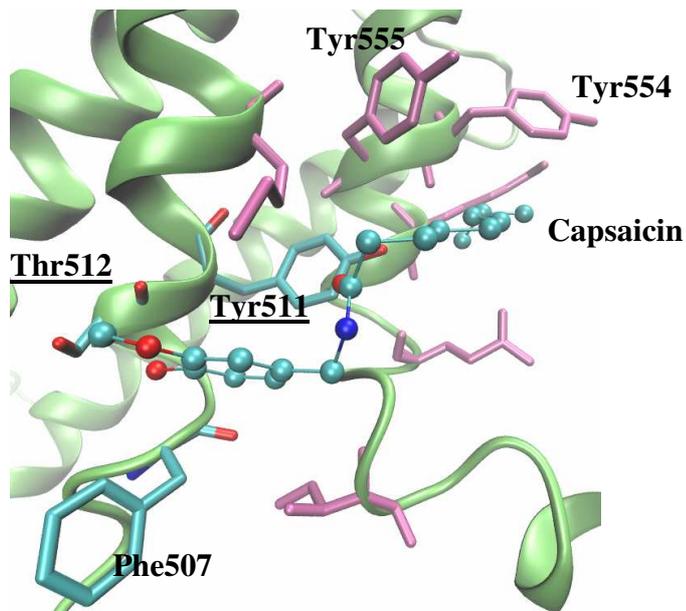
Figure 6.1 Capsaicin-TRPV1 complex, binding mode 1.
Capsaicin is reported in ball and stick; key YS residues are underlined;
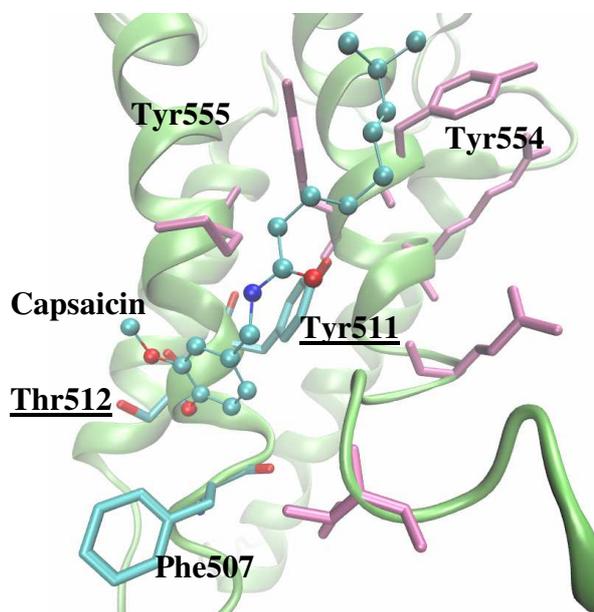
Figure 6.2 Capsaicin-TRPV1 complex, binding mode 2.
Capsaicin is reported in ball and stick; key YS residues are underlined;

PLP	Total score	Score$_{Rotors}$	Score$_{HA}$	Score$_{MW}$
BM1	-0.69	-0.11	-0.40	-0.61
BM2	-0.23	-0.04	-0.12	-0.22

Table 6.4 Pearson coefficient r values calculated per BM1 and BM2, correlating compound activity with PLP scoring function and their normalized values.

Table 6.4 clearly shows that BM1 performs markedly better than to BM2, in all computed score values. As seen in previous analysis, the best result was obtained by the PLP simple Total score, with a *r* value equal to -0.69. As expected, the correlation obtained using refined (and rescored complexes) performs better than those obtained by unrefined results (as compiled in Table6.3). However, the differences between refined and unrefined complexes are not so significant even considering that here all results show a coherent binding mode. Therefore one may suppose that refining the complexes and discriminating between BM1 and BM2 might not markedly enhance the poor correlations as derived by Models 3 and 10. Also for this reason, the following simulations were kept focused on Model 5.

Next, BM1 selected complexes were involved in a wider multi-parametric QSAR analysis, with a view to deriving correlative equations including both docking-based and ligand-based descriptors. Considering the relatively low number of considered molecules, it would be statistically appropriate not to involve more than three independent variables to derive robust correlative equations, avoiding overfitting conditions.

The best equation derived (Eq.1) involved the PLP Total score, and suggests that a good TRPV1 antagonists should be flexible and of limited size.

Eq. 1   3.5319 – 0.0558 PLP_TotalScore - 0.1008 HeavyAtoms + 0.1049 FlexTorsions

$$r^2 = 0.82 \quad q^2 = 0.74 \quad SE = 0.160 \quad F = 37.55 \quad P = 2.06e\text{-}09$$

As done within the study on the TRPM5 channel, a correlative equation was derived using only ligand-based descriptors. This equation has the key objective to reveal the effective role of docking simulations; not to mention its possible use in a pre-docking screening study.

Best performing ligand-based equation (Eq. 2) is in line with the previous one confirming the positive role of flexibility and the detrimental effect of molecular size (here described by radiuses of gyration). Remarkably, the docking score is here replaced by the total number of H-bonding groups, an easily understandable result considering the pivotal role of the H-bonds interactions stabilized with the YS motif. However, Eq. 2 shows markedly worse statistics and this result emphasizes the key role of docking simulations and indirectly affords a positive validation for the here generated and selected TRPV1 homology model.

Eq. 2   *8.9328 – 0.5282*Gyr_rad + 0.0910*FlexTorsions + 0.3261*HbTot*

$$r^2 = 0.58 \;\; q^2 = 0.44 \;\; SE = 0.243 \;\; F = 11.90 \;\; P = 4.32e\text{-}005$$

6.10 TRPV1 virtual screening.

Once the predictive power of TRPV1 ligand binding domain for the selected TRPV1 model has been positively assessed by docking simulations and the binding mode BM1 was proved to be more predictive at least for the considered antagonists, the following model validation involved a virtual screening campaign, aimed at further validating the reliability and the predictive power of the model, by testing its capacity to discriminate between active and inactive compounds.

By applying the same computational protocols as detailed in paragraph 3.2.8, a library of 130 antagonists was made up by collecting potent compounds (nM activity) in literature. In detail, 74 sulfonylaminophenyl derivatives coming from the work of Kim and co-workers[312] were jointed to 56 2-thio pyridine propanamides analogues from the work of Ha and co-workers [313].

The resulting 130 molecules were dispersed among 12870 inactive molecules, which also in this case were taken from the inactive ligands of the previously described HTS study on TRPM8, obtaining a 13000 compounds set with the percentage of active molecules corresponding to 1 %. Once again, in order not to bias the results, putative decoys were selected if they fulfilled several physicochemical property ranges, which were determined by analysing the active compounds.

For the sake of completeness and while best results in previous docking simulations were obtained when using PLP score function, docking calculations were performed by PLANTS program, involving all three implemented scoring functions and their related normalized values.

Moreover, in order to focus the obtained poses only on the best performing BM1, docking calculations were addressed in a 8Å radius sphere around the centroid defined by the Cα atoms of the key residues Tyr511 and Ser512. Such a sphere is large enough to encompass the entire BM1 region but not to involve the upper region around Thr550.

Table 6.5 reports all computed EF values and clearly shows that PLP scoring function, and in particular, the Total score affords the best result (Table 6.6). More importantly, the performed virtual screening campaign proved successful as clearly evincible from the histogram plot (Figure 6.3) as well as from the very high EF values as obtained in the lowest thresholds (Top 1% and 2%) which indicate that such a screening strategy can be conveniently utilized also in the really challenging early recognition.

CHEMPLP	Total score	Score$_{Rotors}$	Score$_{HA}$	Score$_{MW}$
1	6.15	3.07	1.53	0.76
2	5.38	1.53	1.53	0.76
5	4.30	2.92	2.61	1.38
10	2.92	2.69	2.00	2.15
20	2.46	2.19	1.76	1.50

PLP	Total score	Score$_{Rotors}$	Score$_{HA}$	Score$_{MW}$
1	12.30	8.46	3.84	9.23
2	9.61	5.76	4.23	6.53
5	6.30	4.00	3.84	4.61
10	5.69	4.15	3.23	5.00
20	3.69	3.07	2.15	3.26

PLP95	Total score	Score$_{Rotors}$	Score$_{HA}$	Score$_{MW}$
1	10.7	7.69	3.07	3.84
2	8.07	5.38	2.30	2.69
5	7.23	5.35	4.15	2.92
10	3.84	5.00	3.15	2.84
20	3.42	2.73	2.73	2.30

Table 6.5 Virtual screening evaluation of model 8,molecules ranked per ChemPLP, PLP, PLP95 scoring function and their normalized values.

n (top %N)	Cpds	Act	EF	Sk
1	130	16 (1.3)	12.30	
2	260	25 (2.6)	9.61	
5	650	41 (6.5)	6.30	3.2
10	1300	74 (13)	5.69	
20	2600	96 (26)	3.69	

Table 6.6 Detail of the virtual screening evaluation on model 8,

molecules are ranked per PLP Total score.

"Cpds" reports the number of molecules hosted in the considered subset; "Act" stays for the

actives compared to the expected one in parenthesis; EF reports the Enrichment factor values;

"Sk" the Skewness value of the entire distribution.
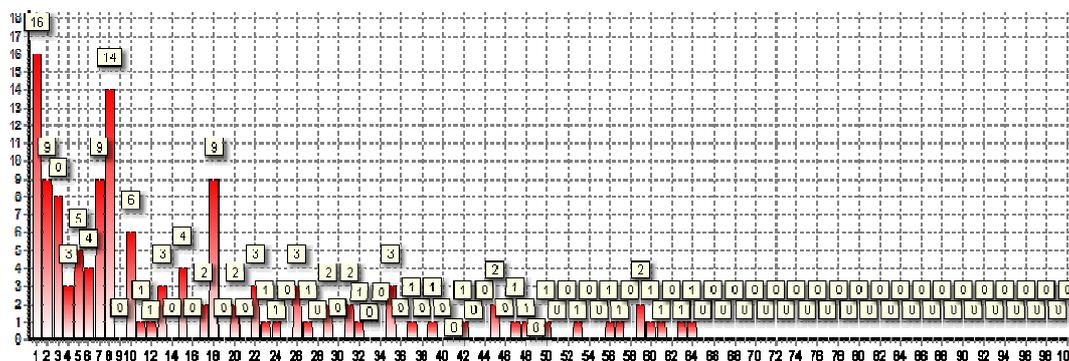
Figure 6.3 Virtual screening on model 8, molecules ranked per PLP Total score.

The histogram plot reports the number of actives per cluster, each cluster contains 130 cpds.

Finally and to further support the here adopted filtering strategy to select the homology model by which the docking simulations are carried out, virtual screening calculations were also performed with the same settings on model 3 and 10 as well. The simulations were performed applying the same computational procedures as previously described for the selected model even though only the best performing PLP scoring function was utilized.

Table 6.7 compiles the EF values obtained with total scores and relative normalized values and reveals that Model10 affords relevant results even though however worse than those provided by the best model, and once again best EFs values are generated by the Total score. Differently, Model 3, while still enriching the set, never reached EF values higher than 4. Overall, these last results emphasize that the selected best model is the best performing one, thus confirming the reliability of the here described selection strategy.

Model 3	Total score	Score$_{Rotors}$	Score$_{HA}$	Score$_{MW}$
1	3.07	1.53	0.76	1.53
2	2.60	0.76	0.76	0.76
5	2.61	2.46	1.38	1.07
10	2.92	1.76	1.46	2.84
20	2.23	1.84	1.65	2.57
Model 10	Total score	Score$_{Rotors}$	Score$_{HA}$	Score$_{MW}$
1	7.69	3.07	1.53	3.07
2	6.15	1.53	2.30	2.69
5	4.46	2.76	2.00	3.23
10	4.46	2.92	2.84	3.23
20	3.04	2.23	1.69	2.30

Table 6.7 Virtual screening evaluation of model 1 and 3, molecules ranked per PLP scoring function and its weightings. EF reported are calculated in the top 1, 2, 5, 10 and 20 % of the set.

Conclusions

Even though the present Thesis involves four similar but rather unconnected case studies and some specific conclusions are already reported at the end of each chapter, the reported results allow for some relevant general considerations. Firstly, the homology modeling by fragments, which was initially developed to support the generation of "*problematic*" GPCR models, has now leaved the GPCR world, where the reasonable richness of resolved structures allows reliable models to be generated for most GPCRs, and offers its notable potentialities to the modeling of clearly more complicated ion-channels. Specifically, the here reported results afford convincing confirmations that such an approach can be conveniently exploited to model complex transmembrane proteins provided that an even far protein with the same topology has been resolved to assist the final assembly. By this way, it is possible to properly account for the local homologies while fulfilling the overall folding imposed by the final template and while avoiding to generate models which are forcedly constrained to comply to the structure of the available template. The here reported results provide further remarkable confirmations for the reliability of the already modeled (and exploited) TRPM8 model, while the here generated TRPM5 and TRPV4 models afford results (despite obtained in a validating preliminary phase) in line with those of TRPM8 further emphasizing the reliability of the fragmental approach. Not to mention that the described targeted strategy to model TRPV1 suggests that previously generated homology models can be then exploited to assist the modeling of highly homologous proteins still obtaining encouraging results but with a significant saving of the required computational efforts.

The second consideration involves the character of the described computational studies. While avoiding dangerous oversimplified scenarios, the computational studies can be roughly subdivided into two groups. In the first ones, the simulations involve a restricted number of compounds which are deeply analyzed by using really sophisticated and time demanding computational strategy which are able to produce a rich arsenal of information concerning not only the predicted activity but also the kinetic profile and the mechanism of action. Such a wealth of data can be used to rationally optimize new ligands and to delve the specific factors influencing the biological profile. These studies are rarely amenable for large libraries and typically performed within academic environments. The second group includes industrial driven projects which usually involve very huge databases of compounds which are analyzed often using ligand-based chemoinformatics methods. Such *in silico* approaches imply a set of reasonable and necessary approximations which

permits to conveniently deals with the large data involved. Clearly, the results which can be derived by these simplified strategies are less informative and basically allow the discrimination between active and inactive compounds, a kind of result which represent the major goal of a virtual screening campaign. While conducted in an academic environment, the here presented studies substantially belong to the second group since they were aimed at developing homology models and computational strategy amenable for virtual screening analyses of huge libraries of compounds. Nevertheless, they were designed and performed with a view to pursuing an acceptable compromise in which introduced approximations and exhaustiveness of structure-based simulations are modulated to yield reliable results with a still bearable computational effort. While requiring further confirmations the results obtained for TRPM5, TRPV1 and TRPV4 emphasize that it is possible to reach convenient compromises by which the unavoidably introduced approximations does not undermine the predictive power of the obtained models.

In this regard, one may clearly understand the attempt, here performed in some cases, to compare and combine structure-based and ligand-based strategies for virtual screening campaigns. As detailed in Chapter 3, virtual screening studies are assisting to the historic duel between structure-based and ligand-based approaches even though more recent trends tend to synergistically combine the two methods. The results obtained by the TRPM8 simulations emphasize the remarkable potentialities in combining more different methods to generate improved enrichment factors.

References

[1] Brody H, *"Taste"*, Nature 486, S1 (2012)

[2] Cipolla Carlo M, *"Allegro ma non troppo"*, Il Mulino.

[3] *Human biology*, 201-464, Daniel D. Chiras. Jones & Bartlett Learning, 2005.

[4] *Human Physiology: An integrated approach* 5th Edition -Silverthorn, Chapter 10, 2010.

[5] Blass EM, *"Opioids, sweets and a mechanism for positive affect: Broad motivational implications"*, Dobbing 115–124, 1987.

[6] Desor JA, Maller O, *"Taste acceptance of sugars by human infants"*, Journal of Comparative and Physiological Psychology, 84 (3): 496–501, 1973.

[7] Schiffman S, *"Taste and smell in disease"*. The New England Journal of Medicine, 308 (22): 1337–43,1983.

[8] Hoon MA, Adler E, *"Putative mammalian taste receptors: a class of taste-specific GPCRs with distinct topographic selectivity"*, Cell, 96:541–51, 2999.

[9] Nelson G, Hoon M A, *"Mammalian sweet taste receptors"*, Cell, 106:381–90, 2001.

[10] Ohkuri T, Yasumatsu K, *"Multiple sweet receptors and transduction pathways revealed in knockout mice by emperature dependence and gurmarin sensitivity"*, Am. J. Physiol. Regul. Integr. Comp. Physiol. 296:R960–R971, 2009.

[11] Damak S, Rong M, *"Detection of sweet and umami taste in the absence of taste receptor T1r3"*, Science, 301:850–853, 2003.

[12] Ishimaru Y, Matsunami H, *"Transient receptor potential (TRP)channels and taste sensation"*. J. Dent. Res. 88:212–218, 2009

[13] Zhao GQ, Zhang Y, *"The receptors for mammalian sweet and umami taste"*, Cell, 115:255–266, 2003.

[14] Jiang P, Cui M, *"Identification of the cyclamate interaction site within the transmembrane domain of the human weet taste receptor subunit T1R3"*, J. Biol. Chem. 280:34296–34305, 2005.

[15] Varkevisser B, Kinnamon SC, *"Sweet taste transduction in hamster: Role of protein kinases"*, J. Neurophysiol. 83:2526–2532, 2000.

[16] Margolskee RF, *"Molecular mechanisms of bitter and sweet taste transduction"*, J. Biol. Chem. 277:1–4, 2002.

[17] Lindemann B, *"Receptors and transduction in taste"*, Nature, 413:219–225, 2001.

[18] Margolskee RF, *"Molecular mechanisms of bitter and sweet taste transduction"*, J. Biol. Chem. 277:1–4, 2002.

[19] Ohkuri T, Yasumatsu K, *"Multiple sweet receptors and transduction pathways revealed in knockout mice by emperature dependence and gurmarin sensitivity"*, Am. J. Physiol. Regul. Integr. Comp. Physiol. 296:R960–R971, 2009.

[20] Lindemann B, *"chemoreception: Tasting the sweet and the bitter"*, Curr. Biol. 6:1234–1237, 1996.

[21] Scinska A, Koros E, *"Bitter and sweet components of ethanol taste in humans"*. Drug and Alcohol Dependence 60 (2): 199–206, 2000.

[22] Glendinning JI, *"Is the bitter rejection response always adaptive?"*, Physiol Behav 56 (6): 1217–1227, 1994.

[23] Johns T, *"With Bitter Herbs They Shall Eat It: Chemical ecology and the origins of human diet and medicine"*, Tucson: University of Arizona Press, 1990.

[24] Wang X, *"Relaxation Of Selective Constraint And Loss Of Function In The Evolution Of Human Bitter Taste Receptor Genes"*, Human Molecular Genetics, 13(21), 2671-2678, 2004.

[25] Maehashi K, Matano H, *"Bitter peptides activate hTAS2Rs, the human bitter receptors"*, Biochem Biophys Res Commun 365 (4): 851–855, 2008.

[26] Lindemann B, *"Receptors and transduction in taste"*, Nature, 413 (6852): 219–225, 2001.

[27] Wu SV, Rozengurt N, *"Expression of bitter taste receptors of the T2R family in the gastrointestinal tract and enteroendocrine STC-1 cells"*, Proc. Natl. Acad. Sci. 99:2392–2397, 2002.

[28] Sai PP, Rajinder PB, *"Chapter Ten – Constitutive Activity of Bitter Taste Receptors (T2Rs)"*, Advances in Pharmacology, Volume 70, 2014, Pages 303–326, 2014.

[29] Maehashi K, Huang L, *"Bitter peptides and bitter taste receptors"*, Cell. Mol. Life Sci. 66:1661–1671, 2009.

[30] Roper SD, *"Signal transduction and information processing in mammalian taste buds"*, Pfl¨ugers Arch. 454:759–776, 2007.

[31] Maehashi K, Huang L, *"Bitter peptides and bitter taste receptors"*, Cell. Mol. Life Sci. 66:1661–1671, 2009.

[32] Upadhyaya J,Pydi SP, *"Bitter taste receptor T2R1 is activated by dipeptides and tripeptides"*. Biochem. Biophys. Res. Commun. 398:331–335, 2010.

[33] Sugita M, *"Taste perception and coding in the periphery"*. Cell Mol. Life Sci. 63:2000–2015, 2006.

[34] Behrens M, Foerster S, *"Gustatory expression profiles of the full complement of human bitter taste receptor genes"*, Chem. Senses. 31(8):E72, 2006.

[35] Chandrashekar J, Mueller KL, *"T2Rs Function as Bitter Taste Receptors"* Cell, Vol. 100, 703–711, 2000.

[36] *"Umami taste receptor identified"*. Nature, 2013.

[37] Yamaguchi S, Kumiko N, *"Umami and Food Palatability"*. Journal of Nutrition 130 (4): 921S–26S, 2000.

[38] Chandrashekar J, Hoon MA, *"The receptors and cells for mammalian taste"*. Nature 444 (7117): 288–94, 2006.

[39] Yasuo T, Kusuhara Y, *"Multiple receptor systems for glutamate detection in the taste organ"*. Biological & Pharmaceutical Bulletin 31 (10): 1833–7, 2008.

[40] San Gabriel A, Uneyama H, *"Cloning and characterization of a novel mGluR1 variant from vallate papillae that functions as a receptor for L-glutamate stimuli"*. Chem Senses 30, 2005.

[41] Kinnamon SC, *"Taste receptor signaling — from tongues to lungs"*, Acta Physiol: 204(2):1748-17161, 2011.

[42] Roper SD, *"Signal transduction and information processing in mammalian taste buds"*. Pflügers Archiv 454 (5): 759–76, 2007.

[43] Iwatsuki K, Ichikawa R, *"Identification of the vesicular nucleotide transporter (VNUT) in taste cells"*. Biochem Bhiphys Res Commun 388 (1): 1–5, 2009.

[44] Chaudhari N, Pereira E, *"Taste receptors for umami: the case for multiple receptors"*, Am J Clin Nutr, 90(suppl):738S–42S, 2009.

[45] McLaughlin S, Margolskee RF, *"The Sense of Taste"*, American Scientist 82 (6): 538–545, 1994.

[46] Chandrashekar J, Zuker CS, *"The cells and peripheral representation of sodium taste in mice"*, Nature, Mar 11;464(7286):297-301, 2010.

[47] Shigemra N, Ohkuri T, *"Amiloride-sensitive NaCl taste responses are associated with genetic variation of ENaC alpha-subunit in mice"*, Am J Physiol Regul Integr Comp Physiol 294: R66–R75, 2008.

[48] Kretz O, Barbry P, *"Differential expression of RNA and protein of the three pore-forming subunits of the amiloride-sensitive epithelial sodium channel in taste buds of the rat"*, J Histochem Cytochem 47: 51–64, 1999.

[49] Clapham DE, *"TRP channels as cellular sensors"*, Nature, 426:517–524, 2003

[50] Montell C, Birnbaumer L, *"The TRP channels, a remarkably functional family"*, Cell,108:595–598, 2002.

[51] Moran MM, Xu H, *"TRP ion channels in the nervous system"*, Curr Opin Neurobiol, 14:362–369, 2004.

[52] Parekh AB, Putney JW Jr, *"Store-operated calcium channels"*, Physiol Rev, 85:757–810, 2005.

[53] Tan J, Abrol R, *"3D structure prediction of TAS2R38 bitter receptors bound to agonists phenylthiocarbamide (PTC) and 6-n-propylthiouracil (PROP)"*, J Chem Inf Model, 23, 52(7):1875-85, 2012.

[54] Marchiori A, Capece L, *"Coarse-grained/molecular mechanics of the TAS2R38 bitter taste receptor: experimentally-validated detailed structural prediction of agonist binding"*, PLoS One, 31, 8(5):e64675, 2013.

[55] Pedretti A, Marconi C, *"Comparative modeling of the quaternary structure for the human TRPM8 channel and analysis of its binding features"*, Biochim Biophys Acta, 1788(5):973-82, 2009.

[56] Pedretti A, Labozzetta A, *"Exploring the activation mechanism of TRPM8 channel by targeted MD simulations"*, Biochem Biophys Res Commun, 414(1):14-9, 2011.

[57] http://zhanglab.ccmb.med.umich.edu/I-TASSER

[58] http://tardis.nibio.go.jp/fugue/

[59] http://www-cryst.bioc.cam.ac.uk/homstrad/

[60] http://swissmodel.expasy.org

[61] https://salilab.org/modeller

[62] http://nova.disfarm.unimi.it/cms/index.php?Software_projects:VEGA_ZZ

[63] http://autodock.scripps.edu/

[64] www.tcd.uni-konstanz.de/research/plants

[65] Beato C, Beccari AR, *"Use of experimental design to optimize docking performance: the case of LiGenDock, the docking module of LiGen, a new de novo design program"*, J Chem Inf Model, Jun 24;53(6):1503-17, 2013.

[66] Wang R, Lai L, *"Further development and validation of empirical scoring functions for structure-based binding affinity prediction"*, J. Comput.-Aided Mol, 16,11-26, 2002.

[67] Kirchmair J, Markt P, *"Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection—What can we learn from earlier mistakes?"*, J Comput Aided Mol Des 22:213–228, 2008.

[68] Klebe G, *"Virtual ligand screening: strategies, perspectives and limitations"*. Drug Discovery Today 11:580–594, 2006.

[69] Chen H, Lyne PD, *"On evaluating molecular-docking methods for pose prediction and enrichment factors"*. J Chem Inf Model 46:401–415, 2006.

[70] Wolber G, Langer T, 13th European symposium on quantitative structure-activity relationships, Duesseldorf, Germany, Aug 27 - Sept 1, 2000

[71] Schueller A, Haehnke V, *"SmiLib v2.0: A Java-Based Tool for Rapid Combinatorial Library Enumeration"*, QSAR Comb Sci 26:407–410, 2007.

[72] P. Gaillard, P.A. Carrupt, B,*"Molecular Lipophilicity Potential, a tool in 3D QSAR: Method and applications"*, Journal of Computer-Aided Molecular Design, 8(2), 83-96, 1994.

[73] Edwards BS, Bologa C, *"Integration of virtual screening with high-throughput flow cytometry to identify novel small molecule formylpeptide receptor antagonists"*, Mol Pharmacol 68:1301–1310, 2005.

[74] Triballeau N, Acher F, *"Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4"*, J Med Chem 48:2534–2547, 2005.

[75] Triballeau N, Acher F, *"Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4"*, J Med Chem, Apr 7;48(7):2534-47, 2005.

[76] Groeneveld RA, Meeden G,*"Measuring Skewness and Kurtosis"*. The Statistician 33, 391–99, 1984.

[77] Voets T, Owsianik G, *"TRPM8 voltage sensor mutants reveal a mechanism for integrating thermal and chemical stimuli"*. Nat Chem Biol 3(3):174–82, 2007.

[78] Dragoni I, Guida E, *"The cold and menthol receptor TRPM8 contains a functionally important double cysteine motif"*, J Biol Chem 281:37353–60, 2006.

[79] Phelps CB, Gaudet R, *"The role of the N terminus and transmembrane domain of TRPM8 in channel localization and tetramerization"*, J Biol Chem 282(50):36474–80, 2007.

[80] Bandell M, Dubin AE, *"Highthroughput random mutagenesis screen reveals TRPM8 residues specifically required for activation by menthol"*, Nat Neurosci 9:493–50, 2006.

[81] Brauchi S, Orta G, *"A hot-sensing cold receptor: C-terminal domain determines thermosensation in transient receptor potential channels"*, J Neurosci 26(18):4835–40, 2006

[82] Rohacs T, Lopes CM, *"PI(4,5)P2 regulates the activation and desensitization of TRPM8 channels through the TRP domain"*, Nat Neurosci 8(5):626–34, 2005.

[83] Fleig A, Penner R, *"The TRPM ion channel subfamily: molecular, biophysical and functional features"*, Trends Pharmacol Sci 25(12):633–39, 2004.

[84] Phelps CB, Gaudet R, *"The role of the N terminus and transmembrane domain of TRPM8 in channel localization and tetramerization"*, J Biol Chem 282(50):36474–80, 2007.

[85] Bavencoffe A, Gkika D, *"The transient receptor potential channel TRPM8 is inhibited via the alpha2 Adrenoreceptor signaling pathway"*, J Biol Chem 285(13):9410–19, 2010.

[86] Fruhwald J, Camacho LJ, *"Alternative splicing of a protein domain indispensable for function of transient receptor potential melastatin 3 (TRPM3) ion channels"*, J Biol Chem 287 (44):36663–72, 2012.

[87] Bandell M, Dubin AE, *"Highthroughput random mutagenesis screen reveals TRPM8 residues specifically required for activation by menthol"*, Nat Neurosci 9:493–500, 2006.

[88] Pertusa M, Madrid R, *"The N-glycosylation of TRPM8 channels modulates the temperature sensitivity of cold-thermoreceptor neurons"*, J Biol Chem 287:18218–229, 2012.

[89] Morenilla-Palao C, Pertusa M, "Lipid raft segregation modulates TRPM8 channel activity. J Biol Chem 84(14):9215–24, 2009.

[90] Dragoni I, Guida E, *"The cold and menthol receptor TRPM8 contains a functionally important double cysteine motif"*, J Biol Chem 281:37353–60, 2006.

[91] Cao C, Yudin Y, *"Polyester modification of the mammalian TRPM8 channel protein: implications for structure and function"*. Cell Rep 4(2):302–315, 2013.

[92] McKemy DD, Neuhausser WM, *"Identification of a cold receptor reveals a general role for TRP channels in thermosensation"*, Nature 416(6876):52–58, 2002.

[93] Zakharian E, Cao C, *"Gating of transient receptor potential melastatin 8 (TRPM8) channels activated by cold and chemical agonists in planar lipid bilayers"*, J Neurosci 30 (37):12526–34, 2010.

[94] Voets T, Droogmans G, *"The principle of temperature-dependent gating in cold- and heat-sensitive TRP channels"*, Nature 430 (7001):748–54, 2004.

[95] Brauchi S, Orio P, *"Clues to understanding cold sensation: thermodynamics and electrophysiological analysis of the cold receptor TRPM8"*, Proc Natl Acad Sci USA 101 (43):15494–99, 2004.

[96] Zakharian E, Cao C, *"Gating of transient receptor potential melastatin 8 (TRPM8) channels activated by cold and chemical agonists in planar lipid bilayers"*, J Neurosci 30 (37):12526–34, 2010.

[97] Yao J, Liu B,*"Modular thermal sensors in temperature-gated transient receptor potential (TRP) channels"*, Proc Natl Acad Sci USA 108(27):11109–14, 2011.

[98] Latorre R, Zaelzer C, *"Structure-functional intimacies of transient receptor potential channels"*, Q Rev Biophys 42(3):201–246, 2009.

[99] McKemy DD, Neuhausser WM, *"Identification of a cold receptor reveals a general role for TRP channels in thermosensation"*, Nature 416(6876):52–58, 2002.

[100] Nilius B, Talavera K, *"Gating of TRP channels: a voltage connection?"*, J Physiol 567(Pt 1):35–44, 2005.

[101] Bidaux G, Beck B, *"Regulation of activity of transient receptor potential melastatin 8 (TRPM8) channel by its short isoforms"*, J Biol Chem 287(5):2948–62, 2012.

[102] Malkia A, Cabedo H, *"The contribution of TRPM8 channels to cold sensing in mammalian neurones"*, J Physiol 567(Pt 2):415–26, 2005

[103] Fujita F, Uchida K, *"Ambient temperature affects the temperature threshold for TRPM8 activation through interaction of phosphatidylinositol 4,5-bisphosphate"*, J Neurosci 33(14):6154–59, 2013.

[104] Fernandez JA, Skryma R, *"Voltage- and cold-dependent gating of single TRPM8 ion channels"*, J Gen Physiol 137(2):173–195, 2011.

[105] Matta JA, Ahern GP, *"Voltage is a partial activator of rat thermosensitive TRP channels"*, J Physiol 585(Pt 2):469–82, 2007.

[106] Zhu MX, *"Understanding the role of voltage gating of polymodal TRP channels"*, J Physiol 585(Pt 2):321–22, 2007.

[107] Voets T, Owsianik G, *"TRPM8 voltage sensor mutants reveal a mechanism for integrating thermal and chemical stimuli"*, Nat Chem Biol 3(3):174–82, 2007.

[108] Jara-Oseguera A, Islas LD, *"The role of allosteric coupling on thermal activation of thermoTRP channels"*, Biophys J 104(10):2160–69, 2013.

[109] Tsavaler L, Shapero MH, *"Trp-p8, a novel prostate-specific gene, is up-regulated in prostate cancer and other malignancies and shares high homology with transient receptor potential calcium channel proteins"*, Cancer Res 61(9):3760–69, 2001.

[110] McKemy DD, Neuhausser WM, *"Identification of a cold receptor reveals a general role for TRP channels in thermosensation"*, Nature 416(6876):52–58, 2002.

[111] Dhaka A, Murray AN, *" TRPM8 is required for cold sensation in mice"*, Neuron 54(3):371–378, 2007.

[112] Yee NS, Zhou W, *"Transient receptor potential channel TRPM8 is over-expressed and required for cellular proliferation in pancreatic adenocarcinoma"*, Cancer Lett 297(1):49–55, 2010.

[113] Viana F, De la Pena E, *"Specificity of cold thermotransduction is determined by differential ionic channel expression"*, Nat Neurosci 5(3):254–60, 2002.

[114] McKemy DD, Neuhausser WM, *"Identification of a cold receptor reveals a general role for TRP channels in thermosensation"*, Nature 416(6876):52–58, 2002.

[115] Takashima Y, Daniels RL, *"Diversity in the neural circuitry of cold sensing revealed by genetic axonal labeling of transient receptor potential melastatin 8 neurons"*, J Neurosci 27(51):14147–57, 2007.

[116] Knowlton WM, Daniels RL, *"Pharmacological blockade of TRPM8 ion channels alters cold and cold pain responses in mice"*, PLoS One 6 (9):e25894, 2011.

[117] Knowlton WM, Daniels RL, *"Pharmacological blockade of TRPM8 ion channels alters cold and cold pain responses in mice"*, PLoS One 6 (9):e25894, 2011.

[118] Bai VU, Murthy S, *"Androgen regulated TRPM8 expression: a potential mRNA marker for metastatic prostate cancer detection in body fluids"*, Int J Oncol 36(2):443–50, 2010.

[119] Yamamura H, Ugawa S, *"TRPM8 activation suppresses cellular viability in human melanoma"*, Am J Physiol Cell Physiol 295(2):C296–C301, 2008.

[120] Fruhstorfer H, *"Thermal sensibility changes during ischemic nerve block"*, Pain 20 (4):355–361, 1984.

[121] Xing H, Chen M, *"TRPM8 mechanism of cold allodynia after chronic nerve injury"*, J Neurosci 27(50):13680–90, 2007.

[122] Bandell M, Dubin AE, *"Highthroughput random mutagenesis screen reveals TRPM8 residues specifically required for activation by menthol"*, Nat Neurosci 9:493–500, 2006.

[123] Janssens A, Voets T, *"Ligand stoichiometry of the cold- and menthol-activated channel TRPM8"*, J Physiol 589(Pt 20):4827–35, 2011.

[124] Duncan DF, Provost N, *"Preclinical validation of the TRPM8 ion channel as a cancer/benign prostate hyperplasia target"*, Eur J Cancer Suppl, 6(12):109, 2008.

[125] Zhang L, Barritt GJ, *"Evidence that TRPM8 is an androgen-dependent Ca^{2+} channel required for the survival of prostate cancer cells"*, Cancer Res, 64(22):8365–73, 2004.

[126] Ortar G, De Petrocellis L, *"Menthylamine derivatives as potent and selective antagonists of transient receptor potential melastatin type-8 (TRPM8) channels"*, Bioorg Med Chem Lett, 20(9), 2010. 2729–32

[127] McKemy DD, Neuhausser WM, *"Identification of a cold receptor reveals a general role for TRP channels in thermosensation"*, Nature, 416(6876):52–8, 2002.

[128] Chuang HH, Neuhausser WM, *"The super-cooling agent icilin reveals a mechanism of coincidence detection by a temperature-sensitive TRP channel"*, Neuron 43(6):859–69, 2004.

[129] Almeida MC, Hew-Butler T, *"Pharmacological blockade of the cold receptor TRPM8 attenuates autonomic and behavioral cold defenses and decreases deep body temperature"*, J Neurosci, 32(6):2086–99, 2012.

[130] Lashinger ES, Steiginga MS, *"AMTB, a TRPM8 channel blocker: evidence in rats for activity in overactive bladder and painful bladder syndrome"*, Am J Physiol Renal Physiol, 295(3):F803–10, 2008.

[131] Knowlton WM, Daniels RL, *"Pharmacological blockade of TRPM8 ion channels alters cold and cold pain responses in mice"*, PLoS One, 6(9):e25894, 2011.

[132] Raul R. Calvo, Sanath K, *"Discovery of vinylcycloalkyl-substituted benzimidazole TRPM8 antagonists effective in the treatment of cold allodynia"*, Bioorganic & Medicinal Chemistry Letters, 22:1903–07, 2012.

[133] Daniel J. Parks, William H, *"Design and Optimization of Benzimidazole-Containing Transient Receptor Potential Melastatin 8 (TRPM8)Antagonists"*, J. Med. Chem, 54,233–47, 2011.

[134] Colburn RW, Scott DL, *"Cold menthol receptor-1 antagonists"*, US 20120053347, 2012.

[135] Colburn RW, Dax SL, *"N-Acylheterocyclic compounds as cold menthol receptor antagonists and their preparation, pharmaceutical compositions and use in the treatment of diseases"*, WO 2010021878 A1, 2010.

[136] Reid G, Babes A, *"A cold- and menthol-activated current in rat dorsal root ganglion neurones: properties and role in cold transduction"*, J Physiol 545(Pt 2):595–614, 2002.

[137] Xing H, Chen M, *"TRPM8 mechanism of cold allodynia after chronic nerve injury"*, J Neurosci 27(50):13680–90, 2008.

[138] Knowlton WM, McKemy DD, *"TRPM8: from cold to cancer, peppermint to pain"*, Curr Pharm Biotechnol 12(1):68–77, 2011.

[139] Hirata H, Oshinsky ML, *"Ocular dryness excites two classes of corneal afferent neurons implicated in basal tearing in rats: involvement of transient receptor potential channels"*, J Neurophysiol 107(4):1199–209, 2012.

[140] Malkia A, Madrid R, *"Bidirectional shifts of TRPM8 channel gating by temperature and chemical agents modulate the cold sensitivity of mammalian thermoreceptors"*, J Physiol 581(Pt 1):155–174, 2007.

[141] Alessandro Pedretti, Cristina Marconi, *"Comparative modeling of the quaternary structure for the human TRPM8 channel and analysis of its binding features"*, Biochimica et Biophysica Acta 1788, 973–82,2009.

[142] G. Ausiello, G. Cesareni, *"ESCHER: a new docking procedure applied to the reconstruction of protein tertiary structure"*, Proteins, 28: 556-67, 1997.

[143] Voets T, Owsianik G, *"TRPM8 voltage sensor mutants reveal a mechanism for integrating thermal and chemical stimuli"*, Nat Chem Biol, 3: 174–182, 2007.

[144] Villalba-Galea CA, Sandtner W, *"S4-based voltage sensors have three major conformations"*, Proc. Natl. Acad. Sci. USA, 105: 17600–07, 2008.

[145] Bocquet N, Nury H, *"X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation"*, Nature 457: 111–114, 2009.

[146] Malkia A, Pertusa M, *"Differential role of the menthol-binding residue Y745 in the antagonism of thermally gated TRPM8 channels"*, Mol Pain 5: 62, 2009.

[147] Malkia A, Pertusa M, *"Differential role of the menthol-binding residue Y745 in the antagonism of thermally gated TRPM8 channels"*, Mol Pain 5: 62, 2009.

[148] Wilson GL, Lill MA, *"Integrating structure-based and ligand-based approaches for computational drug design"*, Future Med Chem, 3: 735-50, 2011.

[149] Bianchi BR, Moreland RB, *"Application of large-scale transiently transfected cells to functional assays of ion channels: different targets and assay formats"*, Assay Drug Dev Technol, 5: 417-24, 2007.

[150] Beck EJ, Hutchinson TL, *"Development and validation of a secondary screening assay for TRPM8 antagonists using QPatch HT"*, Assay Drug Dev Technol, 8: 63-72, 2010.

[151] Lipinski CA, Lombardo F, *"Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings"*, Adv Drug Del Rev, 23: 3-25, 1997.

[152] Oprea TI, *"Property distribution of drug-related chemical databases"*, J Comp Aid Mol Des,14: 251, 2000.

[153] Feng BY, Shelat A, *"High-throughput assays for promiscuous inhibitors"*, Nat Chem Biol, 1: 146-8, 2005.

[154] Feng BY, Simeonov A, *"A high-throughput screen for aggregation-based inhibition in a large compound library"*, J Med Chem, 50: 2385-90, 2007.

[155] Pearce BC, Sofia MJ, *"An empirical process for the design of high-throughput screening deck filters"*, J Chem Inf Model, 46: 1060-8, 2006.

[156] Kazius J, McGuire R, *"Derivation and validation of toxicophores for mutagenicity prediction"*, J Med Chem, 48: 312-320, 2005.

[157] Metz JT, Huth JR, *"Enhancement of chemical rules for predicting compound reactivity towards protein thiol groups"*, J Comput Aided Mol Des, 21: 139-44, 2007.

[158] Muegge I, *"Selection criteria for drug-like compounds"*, Med Res Rev, 23: 302-21, 2003.

[159] Olah M, Bologa CG, *"Strategies for compound selection"*, Curr Drug Discov Technol, 1: 211-20, 2004.

[160] Walters WP, Murcko MA, *"Prediction of 'drug-likeness'"*, Adv Drug Deliv Rev, 54: 255-71, 2001.

[161] Bemis GW, Murcko MA, *"The properties of known drugs. 1. Molecular frameworks"*, J Med Chem, 39: 2887-93.1996.

[162] *SMARTS Teory Manual*, Daylight Inc. http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html

[163] Ostacolo C, Ambrosino P, *"Isoxazole derivatives as potent transient receptor potential melastatin type 8 (TRPM8) agonists"*, Eur J Med Chem, 69: 659-69, 2013.

[164] Chuang HH, Neuhausser WM, *"The super-cooling agent icilin reveals a mechanism of coincidence detection by a temperature-sensitive TRP channel"*, Neuron, 43: 859-69, 2004.

[165] Sperandio O, Andrieu O, *"MED-SuMoLig: a new ligand-based screening tool for efficient scaffold hopping"*, J Chem Inf Model, 47: 1097-110, 2007.

[166] Carletta J, *"Assessing agreement on classification tasks: The kappa statistic"*, Comp Ling, 22: 249–54, 1996.

[167] Ghemtio L, Perez-Nueno VI, *"Recent trends and applications in 3D virtual screening"*, Comb Chem High Throughput Screen, 15: 749-69, 2012.

[168] Morris GM, Goodsell DS, *"Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function"*, J. Comput Chem, 19: 1639-62, 1998.

[169] Trott O, Olson AJ, *"AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading"*, J Comput Chem, 31: 455-61, 2010.

[170] Beccari AR, Cavazzoni C, *"Li-Gen: a high performance workflow for chemistry driven de novo design"*, J Chem Inf Model, 53: 1518-27, 2013.

[171] Proudfoot CJ, Garry EM, *"Analgesia mediated by the TRPM8 cold receptor in chronic neuropathic pain"*, Curr Biol 16 (16): 1591-605, 2006.

[172] Wasner G, Schattschneider J, *"Topical menthol - a human model for cold pain by activation and sensitization of C nociceptors"*, Brain, 127: 1159-71, 2007.

[173] Namer B, Seifert F, *"TRPA1 and TRPM8 activation in humans: effects of cinnamaldehyde and menthol"*, Neuroreport, 16 (9): 955-59, 2009.

[174] Ostacolo C, Ambrosino P, *"Isoxazole derivatives as potent transient receptor potential melastatin type 8 (TRPM8) agonists"*, European Journal of Medicinal Chemistry, 69: 659-69, 2013.

[175] Foster AJ, Van der Logt CPE, *"Tetrahydropyrimidine-2-one derivatives and their uses"*, Unilever plc, WO 2004026840, 2004.

[176] Bassoli A, Borgonovo G, *"Synthesis of a new family of Naryl lactams active on chemesthesis and taste"*, Eur J Org Chem, 7: 1656-63, 2006.

[177] Calvo RR, Meegalla SK, *"Benzimidazole derivatives useful as TRPM8 receptor modulators"*, Jansenn Pharmaceutica, NV, WO 2010132247, 2010.

[178] IrlapatiI NR, Thomas A, *"Fused oxazole and thiazole derivatives as TRPMs modulators"*, Glen-Mark Pharmaceuticals, SA, WO 2010010435, 2010.

[179] Murugesan N, Gu Z, *"Biphenylsulfonamide endothelin receptor antagonists. 4. Discovery ofN-[[2'-[[(4,5-dimethyl-3-isoxazolyl)amino]sulfonyl]-4-(2-oxazolyl)[1,1'-biphenyl]-2-yl]methyl]-N,3,3-trimethylbutanamide (BMS-207940), a highly potent and orally active ET(A) selective antagonist"*, J Med Chem, 46 (1): 125-37, 2003.

[180] Amano A, *"Cooling and pungent agents"*, J Jap Soc Cut Health, 17: 77-86, 1984.

[181] Watson HR, Hems R, *"New compounds with the menthol cooling effect"*, J Soc Cosmet Chem, 9: 185-200, 1978.

[182] Pedretti A, Marconi C, *"Comparative modeling of the quaternary structure for the human TRPM8 channel and analysis of its binding features"*, Biochim Biophys Acta, (5): 973-82, 2009.

[183] Chuang HH, Neuhausser WM, *"The super-cooling agent icilin reveals a mechanism of coincidence detection by a temperature-sensitive TRP channel"*, Neuron, 43 (6): 859-69, 2004.

[184] Pedretti A, Labozzetta A, *"Exploring the activation mechanism of TRPM8 channel by targeted MD simulations"*, Biochem Biophys Res Commun, 414 (1): 14-19, 2011.

[185] Bandell M, Dubin AE, *"High-throughput random mutagenesis screen reveals TRPM8 residues specifically required for activation by menthol"*, Nat Neurosci, 9 (4): 493-500, 2006.

[186] Mandadi S, Roufogalis BD, *"ThermoTRP channels in nociceptors: taking a lead from capsaicin receptor TRPV1"*, Curr Neuropharmacol, 6 (1): 21-38, 2008.

[187] Pan R, Tian Y, *"Central mechanisms of menthol-induced analgesia"*, J. Pharmacol Exp Ther, 343 (3): 661-72, 2012.

[188] Malkia A, Madrid R, *"Bidirectional shifts of TRPM8 channel gating by temperature and chemical agents modulate the cold sensitivity of mammalian thermoreceptors"*, J Physiol, 581.1: 155–74, 2007.

[189] Liao M, Cao E, *"Single particle electron cryo-microscopy of a mammalian ion channel"*, Curr Opin Struct Biol, 27:1-7, 2014.

[190] Miller S, Rao S, *"Antibodies to the extracellular pore loop of TRPM8 act as antagonists of channel activation"*, PLoS One, Sep 9;9(9):e107151, 2014.

[191] Perdetti A, *"Nuovo metodo per il docking automatico di ligandi con macromolecole a struttura 3D nota"*, Thesis degree, Università degli studi di Milano, 1995.

[192] Prawitt D, Enklaar T, *"Identification and characterization of MTR1, a novel gene with homology to melastatin (MLSN1) and the trp gene family located in the BWS-WT2 critical region on chromosome 11p15.5 and showing allele-specific expression"*, Hum Mol Genet 9 (2):203–16, 2000.

[193] Perez CA, Huang L, *"A transient receptor potential channel expressed in taste receptor cells"*, Nat Neurosci 5(11):1169–76, 2002.

[194] Hofmann T, Chubanov V, *"TRPM5 is a voltage-modulated and Ca^{2+}-activated monovalent selective cation channel"*, Curr Biol 13(13):1153–58, 2003.

[195] Liu D, Liman ER, *"Intracellular Ca^{2+} and the phospholipid PIP2 regulate the taste transduction ion channel TRPM5"*, Proc Natl Acad Sci USA 100(25):15160–65, 2003.

[196] Zhang Z, Zhao Z, *"The transduction channel TRPM5 is gated by intracellular calcium in taste cells"*, J Neurosci 27(21):5777–86, 2007.

[197] Akabas MH, Dodd J, *"A bitter substance induces a rise in intracellular calcium in a subpopulation of rat taste cells"*, Science 242(4881):1047–50, 1988.

[198] Hwang PM, Verma A, *"Localization of phosphatidylinositol signalling components in rat taste cells: role in bitter taste transduction"*, Proc Natl Acad Sci USA 87(19):7395–99, 1990.

[199] Wellman GC, Nelson MT, *"Signaling between SR and plasmalemma in smooth muscle: sparks and the activation of Ca^{2+}-sensitive ion channels"*, Cell Calcium 34(3):211–229, 2003.

[200] Montell C, Birnbaumer L, *"The TRP channels, a remarkably functional family"*, Cell 108(5):595–98, 2002.

[201] Ullrich ND, Voets T, *"Comparison of functional properties of the Ca 2+ -activated cation channels TRPM4 and TRPM5 from mice"*, Cell Calcium 37(3):267–78, 2005.

[202] Ullrich ND, Voets T, *"Comparison of functional properties of the Ca 2+ -activated cation channels TRPM4 and TRPM5 from mice"*, Cell Calcium 37(3):267–78, 2005.

[203] Nilius B, Prenen J, *"The selectivity filter of the cation channel TRPM4"*, J Biol Chem 280(24):22899–906, 2005.

[204] Talavera K, Yasumatsu K, *"Heat activation of TRPM5 underlies thermal sensitivity of sweet taste"*, Nature 438(7070):1022–25, 2005.

[205] Nilius B, Talavera K, *"Gating of TRP channels: a voltage connection?"*, J Physiol 567(Pt 1):35–44, 2005.

[206] Rohacs T, Lopes CM, *"PI(4,5)P(2) regulates the activation and desensitization of TRPM8 channels through the TRP domain"*, Nat Neurosci 8(5):626–34, 2005.

[207] Suh BC, Hille B, *"Regulation of ion channels by phosphatidylinositol 4,5-bisphosphate"*, Curr Opin Neurobiol 15(3):370–78, 2005.

[208] Oike H, Wakamori M, *"Arachidonic acid can function as a signaling modulator by activating the TRPM5 cation channel in taste receptor cells"*, Biochim Biophys Acta 1761(9):1078–84, 2006.

[209] Liu P, Shah P, *"Transient Receptor Potential Channel Type M5 is essential for fat taste"*, The Journal of Neuroscience, 31(23): 8634-42, 2011.

[210] Talavera K, Yasumatsu K, *"Heat activation of TRPM5 underlies thermal sensitivity of sweet taste"*, Nature 438(7070):1022–25, 2005.

[211] Liu D, Zhang Z, *"Extracellular acid block and acid-enhanced inactivation of the Ca^{2+}-activated cation channel TRPM5 involve residues in the S3-S4 and S5-S6 extracellular domains"*, J Biol Chem 280(21):20691-99, 2005.

[212] Liu D1, Liman ER, *"Intracellular Ca2+ and the phospholipid PIP2 regulate the taste transduction ion channel TRPM5"*, Proc Natl Acad Sci U S A, 100 (25):15160-5, 2003

[213] Clapp TR, Medler KF, *"Mouse taste cells with G protein-coupled taste receptors lack voltage-gated calcium channels and SNAP-25"*, BMC Biol 4:7, 2006.

[214] Zhang Z, Zhao Z, *"The transduction channel TRPM5 is gated by intracellular calcium in taste cells"*, J Neurosci 27(21):5777–86, 2007.

[215] Huang YA, Roper SD, *"Intracellular Ca 2+ and TRPM5-mediated membrane depolarization produce ATP secretion from taste receptor cells"*, J Physiol 588(Pt 13):2343–50, 2010.

[216] Zhang Y, Hoon MA, *"Coding of sweet, bitter, and umami tastes: different receptor cells sharing similar signaling Pathways"*, Cell 112(3):293–301, 2003.

[217] Kokrashvili Z, Rodriguez D, *"Release of endogenous opioids from duodenal enteroendocrine cells requires Trpm5"*, Gastroenterology 137(2):598–606, 2009.

[218] Lin W, Ezekwe EA Jr, *"TRPM5-expressing microvillous cells in the main olfactory epithelium"*, BMC Neurosci 9:114, 2008.

[219] Prawitt D, Monteilh-Zoller MK, *"TRPM5 is a transient Ca^{2+}-activated cation channel responding to rapid changes in $[Ca^{2+}]_i$"*, Proc Natl Acad Sci USA 100(25):15166–71, 2003.

[220] Ashcroft FM, Harrison DE, *"Glucose induces closure of single potassium channels in isolated rat pancreatic beta-cells"*, Nature 312(5993):446–48, 1984.

[221] Brixel LR, Monteilh-Zoller MK, *"TRPM5 regulates glucose-stimulated insulin secretion"*, Pflugers Arch 460(1):69–76, 2010.

[222] Talavera K, Yasumatsu K, *"The taste transduction channel TRPM5 is a locus for bitter-sweet taste interactions"*, FASEB J 22 (5):1343–55, 2008.

[223] Palmer RK, Atwal K, *"Triphenylphosphine oxide is a potent and selective inhibitor of the transient receptor potential melastatin-5 ion channel"*, Assay Drug Dev Technol 8(6):703–13, 2010.

[224] Gees M, Yeranddy A, *"Differential effects of bitter compounds on the taste transduction thannels TRPM5 and IP3 receptor type 3"*, Chem. Senses 39: 295–311, 2014.

[225] Karnail S, Robert W, *"Heterocyclic compounds as sweetener enhancers", US 7674831 B2, 2010.*

[226] Wang R, Lai L, "Further development and validation of empirical scoring functions for structure-based binding affinity prediction", J Comput-Aided Mol Des, 16:11-26, 2002.

[227] Vistoli G, Pedretti A, *"In silico prediction of human carboxylesterase-1 (hCES1) metabolism combining docking analyses and MD simulations"*, Bioorg Med Chem, 18:320–29, 2010.

[228] Renxiao W,Luhua L, *"Further development and validation of empirical scoring functions for structure-based binding affinity prediction"*, Journal of Computer-Aided Molecular Design, 16: 11–26, 2002.

[229] Atwal K, Atwal A, *"Triaryl substituted imidazole derivatives and taste-inhibiting uses thereof"*, US0259875A1, 2007.

[230] Lee SP, Zhou P, *"USe of a TRPM5 inhibitor to regulate insulin and GLP-1 release"*, US0306030A1, 2008.

[231] Stein P, Daines R, *"Hydrazone compounds and their use"*, US0033393A1, 2011.

[232] Penner R, Flieg A, *"Methods of screening for TPM5 modulators"*, US7723075B2, 2010.

[233] Strotmann R, Harteneck C, *"OTRPC4, a nonselective cation channel that confers sensitivity to extracellular osmolarity"*, Nat Cell Biol, 2:695–702, 2002.

[234] Liedtke W, Choe Y, *"Vanilloid receptor-related osmotically activated channel (VR-OAC), a candidate vertebrate osmoreceptor"*, Cell, 103:525–35, 2000.

[235] Wissenbach U, Bödding M, *"Trp12, a novel Trp-related protein from kidney"*, FEBS Lett , 485:127–34, 2000.

[236] Delany NS, Hurle M, *"Identification and characterization of a novel human vanilloid receptor-like protein, VRL-2"*, Physiol Genomics, 4:165–74, 2004

[237] Zhou Y, Morais-Cabral JH, *"Chemistry of ion coordination and hydration revealed by a K^+ channel-Fab complex at 2.0 A resolution"*, Nature, 414(6859):43–48, 2001.

[238] Kuo A, Gulbis JM, *"Crystal structure of the potassium channel KirBac1.1 in the closed state"*, Science, 300(5627):1922–26, 2003.

[239] Kuo A, Domene C, *"Two different conformational states of the KirBac3.1 potassium channel revealed by electron crystallography"*, Structure (Camb), 13(10):1463–72, 2005.

[240] Arniges M, Fernandez-Fernandez JM, *"Human TRPV4 channel splice variants revealed a key role of ankyrin domains in multimerization and trafficking"*, J Biol Chem, 281:8996, 2006.

[241] Strotmann R, Schultz G, *"Ca^{2+}-dependent potentiation of the non-selective cation channel TRPV4 is mediated by a C-terminal calmodulin-binding site"*, J Biol Chem, 278(29):26541–49, 2003.

[242] Suzuki M, Hirao A, *"Microtubule-associated [corrected] protein 7 increases the membrane expression of transient receptor potential vanilloid 4 (TRPV4)"*, J Biol Chem, 278(51):51448–53, 2003.

[243] Plant TD, Strotmann R, *"TRPV4: A Multifunctional Nonselective Cation Channel with Complex Regulation"*, TRP Ion Channel Function in Sensory Transduction and Cellular Signaling Cascades, Ch.9, 2007.

[244] Voets T, Prenen J, *"Molecular determinants of permeation through the cation channel TRPV4"*, J Biol Chem, 277:33704–10, 2002.

[245] Strotmann R, Harteneck C, *"OTRPC4, a nonselective cation channel that confers sensitivity to extracellular osmolarity"*, Nat Cell Biol, 2:695–702,2000.

[246] Watanabe H, Davis JB, *"Activation of TRPV4 channels (hVRL-2/mTRP12) by phorbol derivatives"*, J Biol Chem, 277:13569–77, 2002.

[247] Voets T, Prenen J, *"Molecular determinants of permeation through the cation channel TRPV4"*, J Biol Chem, 277:33704–10, 2002.

[248] Nilius B, Prenen J, *"Differential activation of the volume-sensitive cation channel TRP12 (OTRPC4) and volume-regulated anion currents in HEK-293 cells"*, Pflügers Arch, 443:227–33, 2001.

[249] Watanabe H, Davis JB, *"Activation of TRPV4 channels (hVRL-2/mTRP12) by phorbol derivatives"*, J Biol Chem, 277(16):13569–77, 2002.

[250] Xu F, Satoh E, *"Protein kinase C–mediated Ca^{2+} entry in HEK 293 cells transiently expressing human TRPV4"*, Br J Pharmacol, 140(2):413–21, 2003.

[251] Gao X, Wu L, *"Temperature-modulated diversity of TRPV4 channel gating: activation by physical stresses and phorbol ester derivatives through protein kinase C–dependent and –independent pathways"*, J Biol Chem, 278(29):27129–37, 2003.

[252] Jordt SE, Julius D, *"Molecular basis for species-specific sensitivity to ''hot'' chili peppers"*, Cell, 108:421–30, 2002.

[253] Vriens J, Watanabe H, *"Cell swelling, heat, and chemical agonists use distinct pathways for the activation of the cation channel TRPV4"*, Proc Natl Acad Sci USA, 101:396–401, 2004.

[254] Vriens J, Watanabe H, *"Cell swelling, heat, and chemical agonists use distinct pathways for the activation of the cation channel TRPV4"*, Proc Natl Acad Sci USA, 101(1):396–401, 2004.

[255] Vriens J, Owsianik G, *"Modulation of the Ca^{2+} permeable cation channel TRPV4 by cytochrome P450 epoxygenases in vascular endothelium"*, Circ Res, 97(9):908–15, 2005.

[256] Liedtke W, Choe Y, *"Vanilloid receptor–related osmotically activated channel (VR-OAC), a candidate vertebrate osmoreceptor"*, Cell, 103(3):525–35, 2002.

[257] Gao X, Wu L, *"Temperature-modulated diversity of TRPV4 channel gating: activation by physical stresses and phorbol ester derivatives through protein kinase C–dependent and –independent pathways"*, J Biol Chem, 278(29):27129–37, 2003.

[258] Alessandri-Haber N, Yeh JJ, *"Hypotonicity induces TRPV4-mediated nociception in rat"*, Neuron, 39(3):497–511, 2003.

[259] Strotmann R, Harteneck C, *"OTRPC4, a nonselective cation channel that confers sensitivity to extracellular osmolarity"*, Nat Cell Biol, 2:695–702, 2000.

[260] Delany NS, Hurle M, *"Identification and characterization of a novel human vanilloid receptor-like protein, VRL-2"*, Physiol Genomics, 4:165–74, 2001.

[261] Delany NS, Hurle M, *"Identification and characterization of a novel human vanilloid receptor-like protein, VRL-2"*, Physiol Genomics, 4:165–74, 2001.

[262] Jia Y, Wang X, *"Functional TRPV4 channels are expressed in human airway smooth muscle cells"*, Am J Physiol Lung Cell Mol Physiol, 28:7, L272–78, 2004.

[263] Guatteo E, Chung KK, *"Temperature sensitivity of dopaminergic neurons of the substantia nigra pars compacta: involvement of TRP channels"*, J Neurophysiol, 94:3069–80, 2005.

[264] Tian W, Salanova M, *"Renal expression of osmotically responsive cation channel TRPV4 is restricted to water-impermeant nephron segments"*, Am J Physiol Renal Physiol, 28:7, F17–24, 2004.

[265] Guler A, Lee H, *"Heat-evoked activation of TRPV4 (VR-OAC)"*, J NeuroScience 22, 6408-14, 2002.

[266] Casas S, Novials A, *"Calcium elevation in mouse pancreatic beta cells evoked by extracellular human islet amyloid polypeptide involves activation of the mechanosensitive ion channel TRPV4"*, Diabetologia 51, 2252-62, 2008.

[267] Alessandri-Haber N, Dina OA, *"A transient receptor potential vanilloid 4-dependent mechanism of hyperalgesia is engaged by concerted action of inflammatory mediators"*, J Neurosci 26, 3864-74, 2006.

[268] Klausen TK, Pagani A, *"Modulation of the Transient Receptor Potential Vanilloid Channel TRPV4 by 4alpha-Phorbol Esters: A Structure-Activity Study"*, J Med Chem, 52, 2933-39, 2009.

[269] Surprenant A, *"Pain TRP-ed up by PARs"*, J Physiol 578, 631, 2007.

[270] Stotz SC, Vriens J, *"Citral sensing by transient receptor potential channels in dorsal root ganglion neurons"*, PLoS ONE 3, e2082, 2008.

[271] Liao M, Cao E, *"Structure of the TRPV1 ion channel determined by electron cryo-microscopy"*, Nature, 504(7478):107–12, 2013.

[272] Stotz, S. C.; Vriens, *"Citral sensing by Transient receptor potential channels in dorsal root ganglion neurons"*, PLoS One, 3 e2082, 2008.

[273] Everaerts W, Zhen X,*"Inhibition of the cation channel TRPV4 improves bladder function in mice and rats with cyclophosphamide-induced cystitis"*, Proc Natl Acad Sci USA,107 19084-89, 2010.

[274] Casillas LN, Jeong JU, *"Acyclic 1,3-diamines and uses thereof"*, US20070259965 A1, 2007.

[275] Jeong JU, Marquis RW, *"Acyclic 1,4-diamines and uses thereof"*, US20090105259 A1, 2009.

[276] Casillas LN, Jamieson C, *"Novel Compounds"*, US20070293477 A1, 2005.

[277] Cichy-Knight M, *"Novel Compounds"*, US20090124619 A1, 2009.

[278] Caterina MJ, Schumacher MA, *"The capsaicin receptor: a heat-activated ion channel in the pain pathway"*, Nature, 389 (6653):816–24, 1997.

[279] Xue Q, Yu Y, *"The genomic organization of the gene encoding the vanilloid receptor: evidence for multiple splice variants"*, Genomics, 76(1–3):14–20, 2001.

[280] Vos MH, Neelands TR, *"TRPV1b overexpression negatively regulates TRPV1 responsiveness to capsaicin, heat and low pH in HEK293 cells"*, J Neurochem, 99(4):1088–102, 2006.

[281] Eilers H, Lee SY, *"The rat vanilloid receptor splice variant VR.5'sv blocks TRPV1 activation"*, Neuroreport, 18(10):969–73, 2007.

[282] Lishko PV, Procko E, *"The ankyrin repeats of TRPV1 bind multiple ligands and modulate channel sensitivity"*, Neuron, 54(6):905–18, 2007.

[283] Venkatachalam K, Montell C, *"TRP channels"*, Annu Rev Biochem, 76:387–417, 2007.

[284] Liao M, Cao E, *"Structure of the TRPV1 ion channel determined by electron cryo-microscopy"*, Nature, 504(7478):107–12, 2013.

[285] Tominaga M, Caterina MJ, *"The cloned capsaicin receptor integrates multiple pain-producing stimul"*, Neuron, 21(3):531–43, 1998.

[286] Hellwig N, Plant TD, *"TRPV1 acts as proton channel to induce acidification in nociceptive neurons"*, J Biol Chem, 279 (33):34553–61, 2004.

[287] Owsianik G, Talavera K, *"Permeation and selectivity of TRP channels"*, Annu Rev Physiol 68:685–717, 2006.

[288] Mohapatra DP, Nau C, *"Desensitization of capsaicin-activated currents in the vanilloid receptor TRPV1 is decreased by the cyclic AMP-dependent protein kinase pathway"*, J Biol Chem, 278(50):50080–90, 2003.

[289] Jordt SE, Julius D, *"Molecular basis for species-specific sensitivity to ''hot'' chili peppers"*, Cell, 108:421–430, 2002.

[290] Gavva NR, Treanor JJ, *"Pharmacological blockade of the vanilloid receptor TRPV1 elicits marked hyperthermia in humans"*, Pain, 136:202–10, 2008.

[291] Voets T, Droogmans G, *"The principle of temperature-dependent gating in cold- and heat-sensitive TRP channels"*, Nature, 430:748–54, 2004.

[292] McLatchie LM, Bevan S, *"The effects of pH on the interaction between capsaicin and the vanilloid receptor in rat dorsal root ganglia neurons"*, Br J Pharmacol, 132(4):899–908, 2001.

[293] Jordt SE, Tominaga M, *"Acid potentiation of the capsaicin receptor determined by a key extracellular site"*, Proc Natl Acad Sci USA, 97(14):8134–39, 2000.

[294] Szolcsanyi J, Joo F, *"Mitochondrial changes in preoptic neurons after capsaicin desensitization of the hypothalamic thermodetectors in rats"*, Nature, 229(5280):116–17, 1971.

[295] Chavez AE, Chiu CQ, *"TRPV1 activation by endogenous anandamide triggers postsynaptic long-term depression in dentate gyrus"*, Nat Neurosci, 13(12):1511–18, 2010.

[296] Zygmunt PM, Petersson J, *"Vanilloid receptors on sensory nerves mediate the vasodilator action of anandamide"*, Nature, 400(6743):452–57, 1999.

[297] Birder LA, Nakamura Y, *"Altered urinary bladder function in mice lacking the vanilloid receptor TRPV1"*, Nat Neurosci, 5(9):856–60, 2002.

[298] Inoue K, Koizumi S, *"Functional vanilloid receptors in cultured normal human epidermal keratinocytes"*, Biochem Biophys Res Commun, 291(1):124–29, 2002.

[299] Bolcskei K, Helyes Z, *"Investigation of the role of TRPV1 receptors in acute and chronic nociceptive processes using gene-deficient mice"*, Pain, 117(3):368–76, 2005.

[300] Sharif Naeini R, Witty MF, *"An N-terminal variant of Trpv1 channel is required for osmosensory transduction"*, Nat Neurosci, 9(1):93–98, 2006.

[301] Ciura S, Bourque CW, *"Transient receptor potential vanilloid 1 is required for intrinsic osmoreception in organum vasculosum lamina terminalis neurons and for normal thirst responses to systemic hyperosmolality"*, J Neurosci, 26(35):9069–75, 2006.

[302] Latorre R, Brauchi S, *"ThermoTRP channels as modular proteins with allosteric gating"*, Cell Calcium, 42:427–38, 2007.

[303] Culshaw AJ, Bevan S, *"Identification and biological characterization of 6-aryl-7-isopropylquinazolinones as novel TRPV1 antagonists that are effective in models of chronic pain"*, J Med Chem, 49:471–74, 2006.

[304] Pearce LV, Petukhov PA, *"Evodiamine functions as an agonist for the vanilloid receptor TRPV1"*, Org Biomol Chem, 2:2281–86, 2004.

[305] Xu S, Cheng Y, *"17beta-estradiol activates estrogen receptor beta-signalling and inhibits transient receptor potential vanilloid receptor 1 activation by capsaicin in adult rat nociceptor neurons"*, Endocrinology, 149:5540–48, 2008.

[306] De Petrocellis L, Bisogno T, *"Overlap between the ligand recognition properties of the anandamide transporter and the VR1 vanilloid receptor: inhibitors of anandamide uptake with negligible capsaicin-like activity"*, FEBS Lett, 483:52–56, 2000.

[307] Jara-Oseguera A, Simon SA, *"TRPV1: on the road to pain relief"*, Curr Mol Pharmacol, 1:255–69, 2008.

[308] Hwang SW, Cho H, *"Activation of capsaicin receptors by products of lipoxygenases: endogenous capsaicin-like substances"*, Proc Natl Acad Sci USA, 97:6155–60, 2000.

[309] Liu L, Simon SA, *"Similarities and differences in the currents activated by capsaicin, piperine, and zingerone in rat trigeminal ganglion cells"*, J Neurophysiol, 76:1858–69, 1996.

[310] Tafesse L, Kurose N, *"TRPV1 antagonists and uses thereof"*, EP2604599 A1, 2013.

[311] Rebolledo CL, Sotelo-Hitschfeld P, *"Design and synthesis of conformationally restricted capsaicin analogues based in the 1, 3, 4-thiadiazole heterocycle reveal a novel family of transient receptor potential vanilloid 1 (TRPV1) antagonists"*, Eur J Med Chem, 66:193-203, 2013.

[312] Kim MS, Ryu HC, *"2-(3-Fluoro-4-methylsulfonylaminophenyl) Propanamides as Potent Transient Receptor Potential Vanilloid 1 (TRPV1) Antagonists: Structure Activity Relationships of 2-Amino Derivatives in the N-(6-trifluoromethyl-pyridin-3-ylmethyl) Cregion"*, J Med Chem, 55(19): 8392–008, 2012.

[313] Ha TH, Ryu HC, *"TRPV1 antagonist with high analgesic efficacy: 2-Thio pyridine C-region analogues of 2-(3-fluoro-4-methylsulfonylaminophenyl) propanamides"*, Bioor & Med Chem, 21: 6657–66, 2013.