universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Chemoinformatics for chemical senses:

Bitter taste and chemesthetic compounds“

verfasst von / submitted by

Sebastian Bayer

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2022 / Vienna 2022

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 606
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Drug Discovery and Development
Betreut von / Supervisor:	Univ.-Prof. Mag. Dr. Veronika Somoza

Table of contents

Acknowledgements

I would like to thank Univ.-Prof. Mag. Dr. Veronika Somoza for giving me the opportunity to perform my master thesis at the Leibniz Institute for Food Systems Biology at the Technical University of Munich.

Special thank goes to my supervisor Dr. Antonella Di Pizio for all the help during my time in Munich and afterwards Thank you for your patience, support and guidance on this journey!

I would also like to thank my co-workers Alessandro Nicoli and Verena Weber for their comradery and help along the way!

Last but not least I would also like to thank my roommates in Freising and in Vienna for the moral support and friendship during our time together.

Abstract

Humans have been fascinated by how we perceive our environment for millennia. Senses have already been studied by Aristoteles who classified them into five types. But the machinations behind them stayed hidden to him. He could not know what receptors are or how they work, like we do nowadays. The response by an organism to a sensory active chemical compound (i.e., chemosensory perception), such as the sweet taste of sugar, the burning sensation of chili, or the eye watering fumes of onions is indeed mediated by chemosensory receptors. A plethora of such chemosensory receptors and thousands of molecules which can interact with them are now known. In parallel to the advancements in knowledge of the biology underlying the chemical senses, we are witnessing increased use of computational studies for research in all chemical fields.

This thesis focuses on the chemoinformatics and Computer-Aided Drug Discovery (CADD) investigations in chemical senses. Specifically, data sets of natural ligands for two chemosensory receptor classes have been collected and analysed.

The first investigated data set is made of natural, food-derived agonists for human bitter taste receptors (TAS2Rs). Bitter compounds lead to a natural aversive reaction, but many of them hold health benefits, and often also influence specific food typical taste, what would a coffee be without its bitterness? We collected TAS2R ligands and their food sources and investigated their chemical space and their relation to the 25 human TAS2R receptors.

The second analysed data set is made of natural agonists for human transient receptor potential (TRP) channels, responsible for chemesthesis perception. The discovery of TRP channels was rewarded with the Nobel Prize in Physiology or Medicine in 2021.[1] We focused the work of this thesis on ligands of TRP channels involved in chemoreception, i.e. TRPA1, TRPV1, TRPV3 and TRPM8. We looked into molecular features important for receptor binding and we developed in-silico models for correlating the structure of these molecules with their activities, which can be used for finding future novel TRP activators.

With this work we hope to help unravel the relationship between chemoreceptors and their natural ligands, which we encounter daily in our environment. Part of the thesis

(i.e., the analyses done on bitter compounds) has already been acknowledged and published earlier this year in the Journal of Agricultural and Food Chemistry.[2]

1. Introduction: Human taste perception

Human taste is vital for nutrient and toxin detection in consumption of food and beverages. Five taste modalities are known: sweet, salty, sour, umami and bitter.[3,4] Sweet indicates carbohydrates, which serve as energy source. Salty indicates the presence of salts, like NaCl, which are vital for the bodies osmotic balance. Umami, also called savoury, indicates protein content, and is triggered by L-amino acids, mainly L-glutamate. Sour indicates low pH and is naturally avoided as it can indicate dangerous contents or spoiled food. Bitter is also an inherently aversive taste modality and thought to guard against toxins. However, not all bitter compounds are toxic.[5] For example, coffee, one of the most abundant beverages today is known for its bitter tones, and traditional medicines are often associated with a bitter taste despite being healthy. Furthermore, in recent years a rise in consumer interest in mildly bitter and sour products has been shown.[6]

Taste sensation in humans occurs on the outer region of the ~5000 taste buds which are located on the surface of the tongue, the palate, and the epiglottis.[7] The active apical tips are in direct contact with the environment and undergo constant renewal.[8] Each taste bud is a conglomerate that consists of about 100 cells, which are classified into three types. Type I cells are the most numerous. They are electron-rich, with an irregular shaped nucleus and their functions are not fully understood yet. They contain ion channels for maintenance of K^+ ion concentration, secret proteins to eliminate extracellular neurotransmitters and are thought to be glial-like supporting cells.[9] Type II cells are larger, with a spherical nucleus and contain taste receptor G-protein coupled receptors (GPCRs) for the detection of either bitter, sweet or umami compounds. Each taste bud contains different type II cells and can therefore respond to different taste stimuli.[10] Type III cells are the rarest and their number varies depending on the location of the taste bud. The elongated cells contain ion channels for the detection of sour and salty taste. [11]

Taste cells are separated from blood supply and the surrounding epithelium by a barrier, which regulates the flow of small molecules from the outside. Furthermore, taste buds contain receptors for cell-cell and cell-neuron communication.[12] Figure 1 shows the schematic profile of a taste bud.

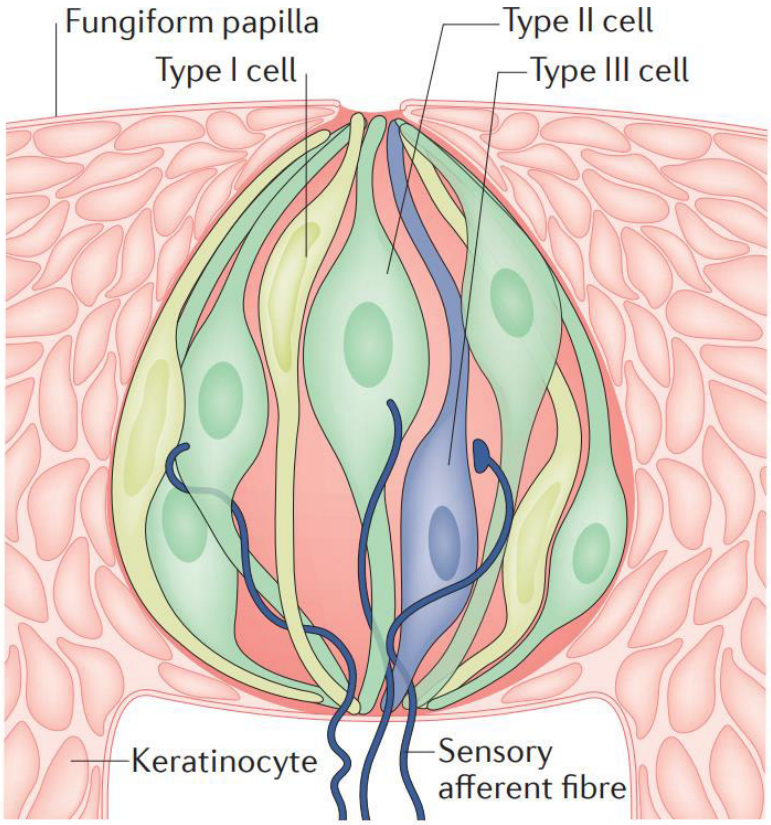


Figure 1 Schematic view of a taste bud. Type I, II and III cells are highlighted with the surrounding epidermis cells (keratinocytes), the upper surface of the tongue (fungiform papilla) and the linked sensory afferent fibre. Picture taken from Roper et al. 2017. [12]

Taste modalities are not the only characteristics by which we describe food and beverages. Chemesthesis is the chemical sensibility of the skin and mucous membranes.[13] It enables us to feel temperature, pressure and helps detect irritating compounds[14] and highly impacts the way we perceive food.[15] Many of those sensations are guided by transient receptor potential (TRP) channels. For instance, TRPV1 detects heat and capsaicin[16] and TRPM8 is activated by lower temperatures and menthol.[17] Those sensations go hand in hand with taste in the oral cavity and are considered to play an important role in the detection of food properties and compounds.[18]

In this thesis, we have investigated natural ligands of bitter taste receptors and TRP channels, which are described in more detail in the following paragraphs.

1.1 Bitter taste receptors

Bitter taste receptors (named TAS2Rs, also referred to as T2Rs) were identified 20 years ago by screening human genome data for g-coupled protein receptor (GPCR) regions in association with bitter taste.[19] 25 human bitter taste receptors are known today.[20] Their role has been proven in multiple experiments such as via calcium imaging, expression of human bitter receptors in mice and primary gastric cells.[21–23] Individual variation in bitter taste perception can be learned but are also governed by genetic differences. The most studied case is the polymorphism of TAS2R38, which can lead to a complete insensitivity of the bitter compounds phenylthiocarbamide and propylthiouracil.[24]

The molecular understanding of TAS2R structure and activation is hampered by the lack of TAS2R experimental structures. Furthermore, bitter taste receptors share little sequence similarity with other GPCRs and their classification into one of the GPCR families is still of some debate.[25] TAS2Rs have short N- and C-termini in the extracellular and intracellular environment, respectively, and 7 transmembrane (TM) helices. The ligand binding site is located in the extracellular TM core.[26] Because of their architecture and orthosteric binding site location, TAS2Rs are often ascribed towards class A GPCRs, despite their low mutual sequence similarity (~20-30%) and sequence identity (~10%).[27] Without crystal structures, only computational structure prediction can be used to gain structural insights into TAS2Rs and it has been performed in recent years to increase our knowledge of TAS2R structures and their interactions with bitter compounds.[28,29]

In comparison to other GCPR ligands, TAS2R ligands are relatively weak, with mid nanomolar to low millimolar ranges for agonist activity.[25] Indeed, higher sensitivity might lead to aversion towards food and thus a decrease in nutrient consumption and an evolutionary disadvantage. While TAS2R19 and -R60 are still orphan, TAS2R14, -R10 and -R46 are activated by one third of the known bitter compounds and to one half together.[2] The majority of bitter taste receptors have intermediate tuning breaths.[30] While sweet receptors work as dimers, TAS2Rs function as monomers. It was found

that TAS2Rs are able to form hetero and homo dimers but the dimeric function has not been demonstrated yet.[31]

Upon agonist binding, the TAS2R undergo a conformational change, and then the G protein α_{gust} separates into Gγ and Gβ subunits. Latter one activates phospholipase Cβ2 (PLCβ2), which leads to the production of the second messenger inositol 1,4,5-trisphosphate (IP3) and diacylgylcerol from the signaling precursor phosphatidylinositol 4,5-bisphosphate (PIP2), a phosphoinositide.[32,33] IP3 then binds to the membrane receptor IP3R3, thus opening it and allowing the influx of Ca^{2+} ions from the endoplasmic reticulum into the cell.[34] Following, TRPM5 in the cell membrane opens, leading to an influx of Na^{+} ions into the cell, hence leading to cellular depolarization and neurotransmitter secretion.[35] A schematic version is shown in Figure 2.

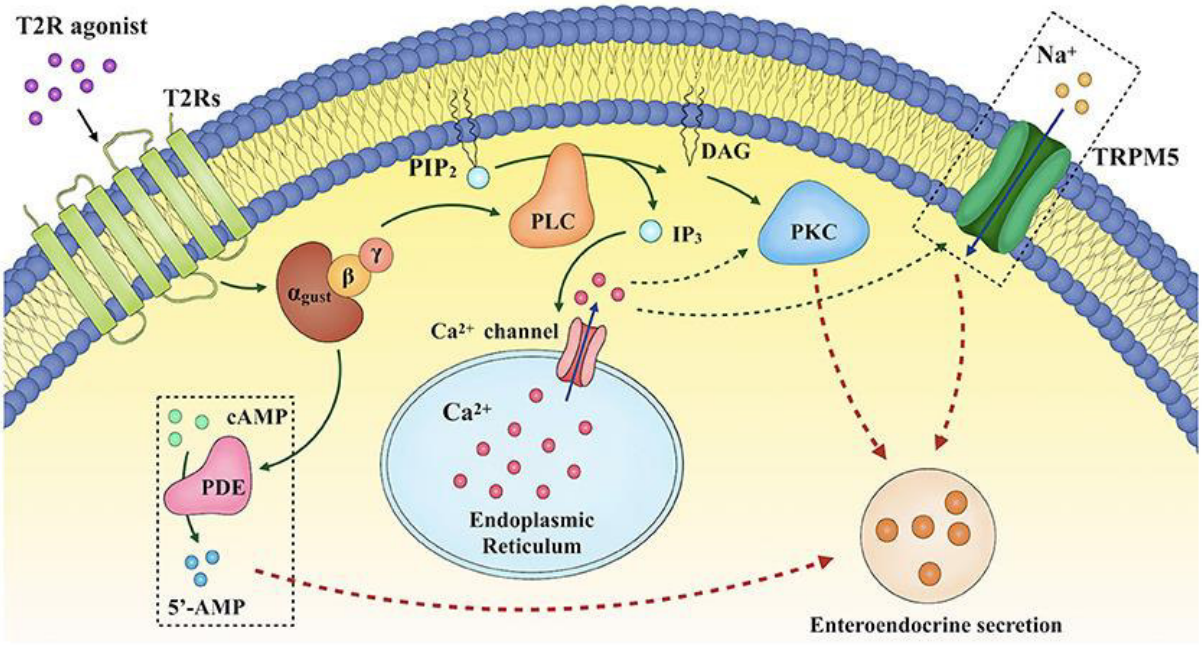


Figure 2 Signalling cascade following TAS2R ligand binding: α_{gust} protein dimer disintegrates. β subunit actives PLC (phospholipase) Cβ2, which triggers IP3 (inositol 1,4,5-trisphosphate) release. IP3 binds to membrane receptor IP3R3 (not shown), allowing the efflux of Ca^{2+} into the cell. TRPM5 activates and depolarizes the cell, leading to neurotransmitter secretion. Picture taken from Xie et al. 2018. [35]

While bitter taste receptors are primarily known for their taste activity they are involved in a multitude of different functions, when expressed in extraoral tissues in the body, such as the innate immune system, the digestive tract, and the respiratory system.[36] Therefore, TAS2Rs are now suggested as promising potential therapeutic targets and have been of significant scientific interest in the past years. For instance, in the upper airways TAS2Rs activation leads to the release of antibacterial agents,[37] marking a possible treatment for airway infections.[38] They are also discussed as a target for the treatment of asthma and similar lung diseases as they are expressed in the airways smooth muscle cells where they evoke bronchodilation.[39] A TAS2R1 agonist, KDT501, is now in clinical phase for its antidiabetic properties (ID number: NCT02444910).[40]

Moreover, TAS2R ligands tend to be “drug-like” and share similar chemical space with known drugs.[41] In fact, 10% of the compounds in the Bitterdb,[42] the largest data base for bitter ligands with over 1000 compounds, are drugs.[42] Still, the space for bitter ligands, especially natural ones hold place for many more compounds to be found.

1.2 Transient receptor potential (TRP) ion channels

In 1998, Julius et al used capsaicin, a compound that induces a burning sensation, to first identify a sensor for heat in human skin: TRPV1.[43] He was awarded with the 2021 Nobel prize in Physiology or Medicine for his groundwork on the function of transient receptor potential channels and how we perceive and sense our environment.[1]

Thanks to his work and that of many others we know today, that TRP channels are involved in the sensation and regulation of pain, temperature, pressure, vision, pH and many others.[44,45] Followingly they have been found in various different organs, such as the brain, heart, kidney, testis, lung, liver, spleen, ovary, intestine, prostate, placenta, uterus and vascular tissue.[29] They perform via the regulation of Ca^{2+} and Mg^{2+} levels in pathways that are not completely understood yet.[46,47]

28 different TRP channels have been identified in mammals. These are separated into 6 subfamilies.[48] Most important of which are TRPM (melastatin), TRPA (ankyrin) and TRPV (vanilloid). The others are: TRPC (canonical), which received the name due to its close relation with the Drosophila TRP, where the channels were first discovered.[49] It is associated with respiratory diseases.[50] TRPML (mucolipin) received its name for its involvement in the neurodevelopmental disorder mucolipidosis IV and is still largely uncharacterised.[51] The last class: TRPP is also named after an associated illness: polycystic kidney disease.[52]

A lot of research has been done on the different TRP channel associated diseases and interest in drugs which can treat them has risen in past years. Initially, drug discovery efforts have been focused on pain medication, but have expanded into respiratory disorders, neurological and psychiatric diseases, diabetes and cancer.[53] However, the fact that TRP channels are expressed in so many different tissues and organs is also a draw back as it makes specificity troublesome, and it has been a burden for drug discovery in the field.[53]

Structural data of TRP channels has long been unobtainable, however this changed with the technical advancements in cryo-EM and now all major TRP subtypes have multiple experimental structures available.[54] All TRP channels assemble as tetramers with a 4-fold symmetry alongside the ion pore. The pore is located in the transmembrane core, which consists of 6 identical helices, folded into two subunits. The first four, which are connected to the N-terminus on the one side, form a voltage sensor like domain (VSLD) and connect with the other two units which build the cation permeable pore between them and connect to the C-terminal ending. A schematic view is shown in Figure 3. The intracellular termini vary between the subtypes and give them unique regulatory functions (see paragraphs 1.2.1-4 for more details).[55,56]

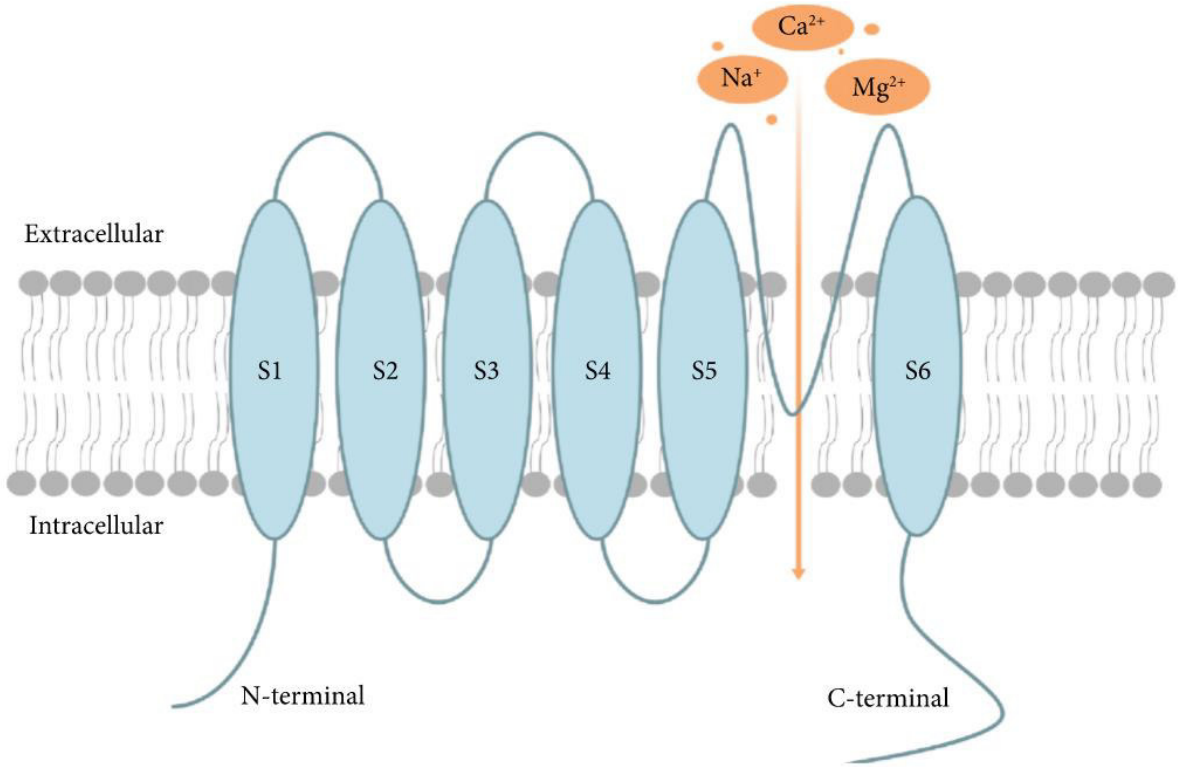


Figure 3 General representation of the transmembrane region of TRP channels. 6 identical helices form the voltage sensor like domain (S1-S4) and the ion pore (S5-S6). Picture taken from Poletini et al 2015. [56]

Depending on the type, TRP channels can be activated in up to three ways. From other receptors such as GCPRs or tyrosine kinases via e.g.: hydrolysis of PIP2, which in terms modulates the TRP channel.[57] Second, by physical activation via change in temperature or pressure, either directly or again through secondary messengers.[16] Third, by ligand activation, either exogenous molecules as Julius used in his initial study of TRPV1 or endogenous structures such as lipids.[43,58]

TRP channels have multiple important roles. Some hidden in downstream effects, as with TRPM5 that regulates the influx of Na^+ ions following the activation of bitter taste receptors,[35] as mentioned above, but also important effects due to their direct activation.

Temperature is the key word here. It is an important factor in our enjoyment of food. We expect and associate different tastes and dishes with certain temperatures and even small changes can have a significant impact. Coffee for instance is commonly served at higher temperatures and studies have shown a difference in flavour at 50, 60 and 70 °C.[59] Furthermore TRPM5 is activated by higher temperatures, leading to an increased sensation of sweetness.[60] Hot tea for example tastes sweeter than cold ice cream despite containing the same amount of sugar.[61]

The perception of hot and cold is altered by chemicals targeting TRP channels. Molecules which we commonly refer to as spices. The aforementioned capsaicin triggers a burning sensation via TRPV1 activation. So does piperine, the active ingredient of black pepper and allicin from garlic. Last one also activates TRPA1, so do isothiocyanates from mustard and curcumin, which is found in turmeric. Similarly, but with an opposite effect, menthol is a TRPM8 agonist that triggers a unique cooling sensation.[62]

TRP channels that interact on our food perception are expressed in epithelial cells of the tongue and in the oral mucosa[63,64] and in the airway system, including the nose and the trachea.[65]

1.2.1 TRPA1

Transient receptor potential channel ankyrin 1 (TRPA1) is the only member of its family in mammals. It can be activated in all three major ways: through ligands, other receptors, and mechanical stimuli. TRPA1 is expressed through many tissues in our body, and is associated with gastrointestinal tract diseases, cardiovascular diseases, itching, dental /oral pain, migraine, and a plethora of other unpleasantries. This has made it a target of great interest and it is the most studied TRP channel today, with several hundred articles published each year.[66]

The 1119 amino acid spanning the structure of TRPA1 fold into a homo tetrameric channel, which has a uniquely large intracellular terminus with 14-18 ankyrin repeat domains. Other than in TRPV1 and TRPM8, the electrophilic binding site is in the cytoplasmic region.[67] Its "clam-like" shape with a fixed lower side and a freely rotatable upper side allows for accommodation of highly diverse ligands.[68]

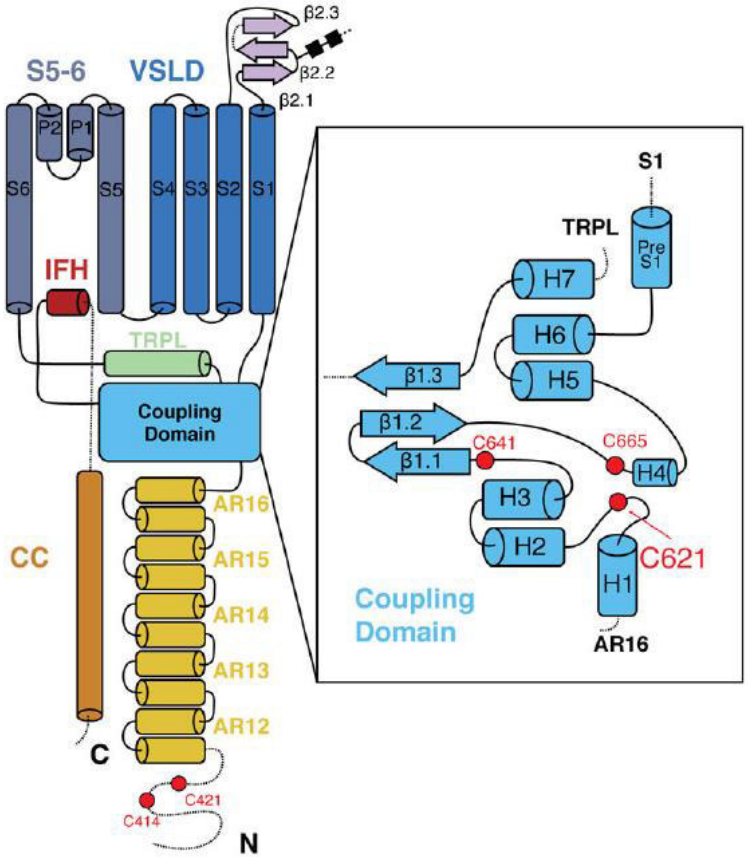


Figure 4 TRPA1 structure with the for TRP channel typical transmembrane region, a coupling domain featuring the main binding site with C621 and a large ankyrin repeat domain. Picture taken from Suo et al. 2020.[68]

It features an extra-ordinary reactive cysteine group (C621) that plays a key role in electrophilic ligand binding.[69–71] The high reactivity is caused by two factors: a lowered pKA value via interaction with the N-cap and a reduced entropy of the thiol group caused by interaction with the aromatic ring of an asparagine (F612), placing it in just the right conformation for electrophilic interaction.[68]

Electrophilic agonists of TRPA1 covalently bind to cysteine residues, foremost C621. Either via reversible reactions, such as thiol-Michael adduct formation or, via non-reversible reactions, mainly nucleophilic SN2 reactions and thiol-α,β-unsaturated reactions.[68,69] Examples are shown in Figure 6. Allyl isothiocyanate, a well-known natural compound found in e.g. wasabi,[72] forms a reversible thiol-Michael adduct (Figure 6a). JO10, a synthetic compound designed for TRPA1 mechanistic studies,[73] engages in a non-reversible nucleophlic-SN2 reaction (Figure 6b).

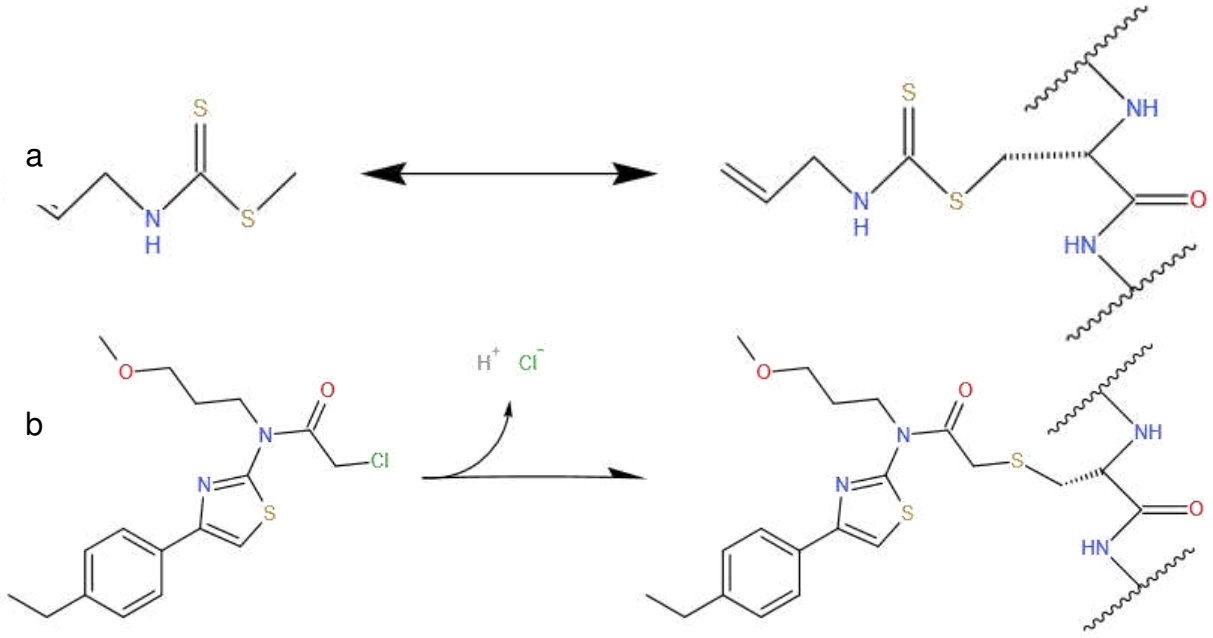


Figure 5 Different binding mechanisms of electrophilic agonists targeting cysteine C621 in TRPA1. Allyl isothiocyanate forms a reversible thiol-Michael adduct (a). JO10 engages in a non-reversible nucleophlic-SN2 reaction (b).

TRPA1 also accommodates non-electrophilic ligands such as menthol.[74] They tend to have a bimodal function, with agonist activity in lower concentrations and antagonist activity in higher concentrations.[75] Several of these ligands have been discovered with in-silico studies, mainly through ligand-based approaches.[76,77] Those place the binding pocket in the S5-S6 pore region of the receptor.[76,78] However, this has not been confirmed by experimental structures yet.

1.2.2 TRPM8

Also known as the cold receptor, TRPM8 is the main sensor for lower temperatures in humans and other mammals.[79,17,80] The calcium permeable ion channel is activated by voltage, lower temperatures, ranging from 8 to 28°C, and a handful of compounds such as menthol and the synthetic compound ilicin that can cause artificial cooling sensations. This ligand activity is conserved across multiple species.[81]

TRPM8 was discovered during a screening against a cDNA library in search for prostate cancer targets and it was later confirmed that it is a vital channel for prostate cancer cells.[82] A plethora of other therapeutic applications have since been proposed, such as urological disorders, asthma, hypersensitivity to cold air, rhinitis and colitis, as it is expressed in a large variety of tissue and organs, including the pancreas, breast, colon, brain, bladder and skin.[62,83,84] Over 1700 journal publications have been dedicated to TRPM8, however, its function and mode of action are still not fully understood.

No human TRPM8 structure is available. Thankfully, avian versions are, and we know from them that TRPM8 features the TRP channels typical transmembrane region, with the binding domain between S2 and S3. The cytoplasmic region consists of four melastatin homology regions (MHR) on the N-terminal side which are important for channel assembly and trafficking. The pore domain relates to the C-terminal domain helixes (CTDH)1 followed by CTDH 2 and a tetrameric coiled coil region.[85]

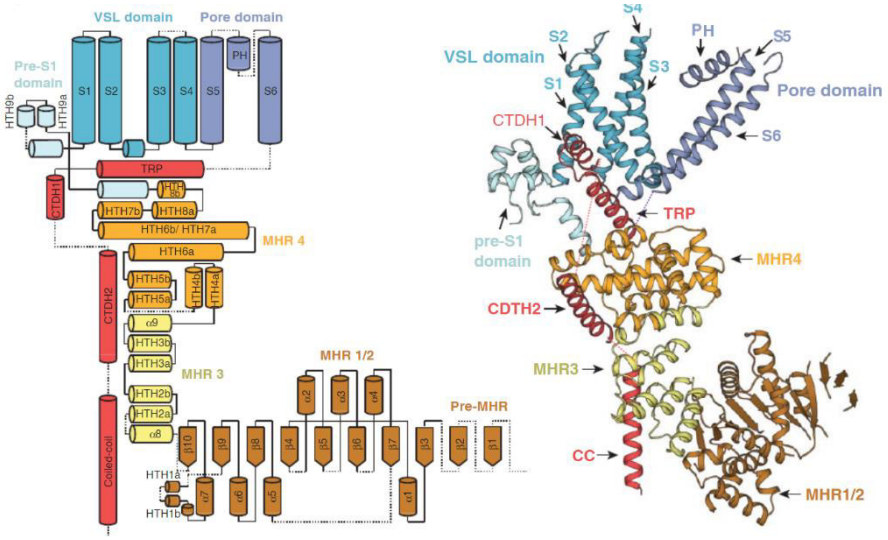


Figure 6 TRPM8 structure, featuring the transmembrane region with VSL and pore domain and the cytosolic region consisting of four melastatin homology regions (MHR) and the C-terminal domain helixes. Picture taken from Yi et al.[85]

Activation of TRPM8 requires the presence of PIP2 an endogenous lipid. It binds allosterically and can act as an agonist itself in high concentrations.[86] Menthol and icilin are the best-studied modulators. Although they, and all other known ligands, are thought to bind in the same pocket, they bind in different ways, which is highlighted by the following 3 differences. First, menthol activates orthologs across all species, while icilin does not. Secondly, only icilin requires Ca^{2+} for activation and lastly, icilin-binding is also dependent on the intracellular pH, while menthol-binding is not. An avian crystal structure with Ilicin bound has been published recently and gives some insight into the binding pocket.[79]

1.2.3 TRPV1

TRPV1 was discovered by Julius et al., as mentioned above.[43] It responds to nociceptive stimuli such as heat above 45°, low pH and noxious ligands.[43] Thus it has been of great interest for pain medication.[87] Several small molecule drugs have been developed and 9 are approved drugs as of now.[88] The majority of known modulators are agonists, and only few antagonists are known. This might seem contradictory, as activation of the channel leads to pain. However, longer activation leads to a decrease in activity, also called desensitization.[89]

TRPV1 channels consist of four subunits that fold around the ion pore with a conserved transmembrane region. The cytoplasmic region consists of the C-terminal TRP domain and the 6 N-terminal ankyrin repeat domains. The so called “vanilloid”-binding pocket lies between S2 and S3, with supporting residues in S4. TRPV1 contains phosphorylation sites that can be activated by kinases such as protein kinase a and c leading to channel sensitization.[90] Furthermore, PIP2 is needed for activation and has a binding site between the S4-S5 linker and the TRP domain.[91] Its production requires ATP, which also competitively binds with calmodulin (CaM),[92,93] a regulatory protein that desensitizes the channel and gets activated by the influx of Ca^{2+} ions during the open state. Therefore, creating a negative feedback loop.

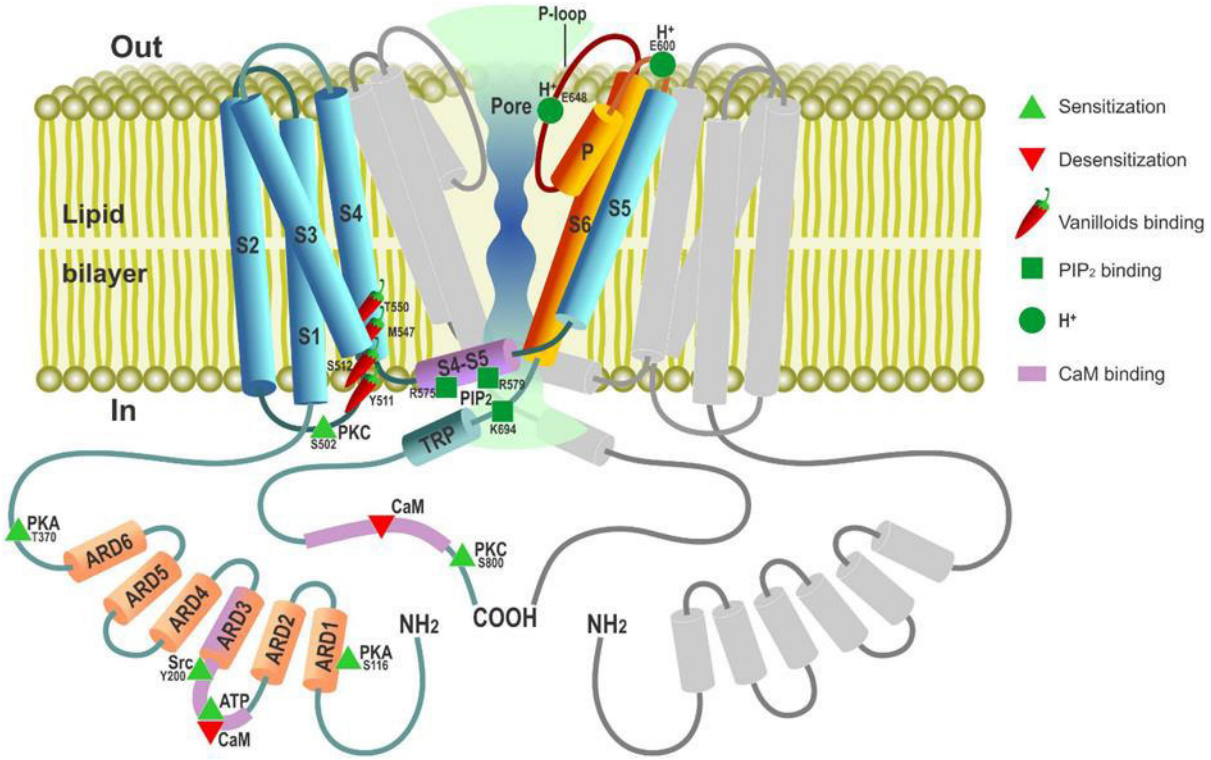


Figure 7 TRPV1 schematic representation. TRPV1 features the transmembrane region with VSL and pore domain and the cytosolic region consisting of the ankyrin repeat domain and the TRP domain. Highlighted are different binding sites for exo- and endogenous ligands and their effect via sensitization or desensitization. Picture taken from Shuba 2021.[93]

1.2.4 TRPV3

TRPV3 is expressed in the oral mucosa and the olfactory epithelium,[94] where it is activated by mild to warm temperatures (31-39°) and chemical stimuli.[95]

Its activation leads to the release of proinflammatory agents and in contrast to TRPV1 it sensitizes upon repeated activation leading to a decrease in the energy barrier of its activation.[96]

It also differs vastly in its sequence, sharing only 43% with its better-studied family brother. Its overall structure, however, shares the same features as in other TRP channels: it features the VLSD and pore domain in the transmembrane region built up by S1-S6, an N-terminal 6 ankyrin repeat domain and the C-terminal TRP domain.[97]

No endogenous modulators are known and only few exogenous ones. Many of those it shares with other TRP channels, such as such as icilin and menthol, which also activate TRPM8 and capsaicin which activates TRPV1.[98–100]

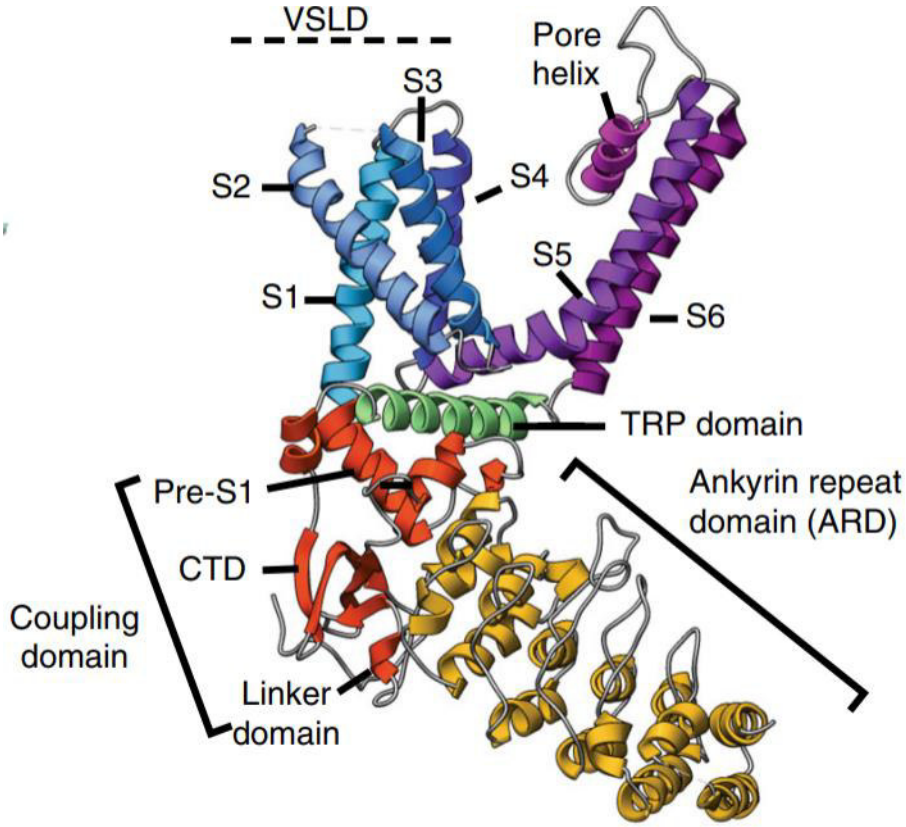


Figure 8 3D representation of TRPV3, featuring the transmembrane region with VSL and pore domain and the cytosolic region consisting of the ankyrin repeat domain and the coupling domain. Picture taken from Nilius et al. [98]

2. Results and discussion

2.1 Cheminformatics analyses of food-derived bitter compounds

Our goal was to assess the current knowledge of TAS2R natural agonists and use chemoinformatic tools to unravel the complex interactions between ligands and receptors. We started by collecting information on known natural TAS2R agonists, with a focus on food compounds. They were classified according to their chemistry and receptor association. We used this data to investigate which class of compounds activate which receptors and the role of compound and receptor promiscuity. For this we used chemoinformatics tools, such as scaffold decomposition and chemical space analysis. This work is published in the journal of agricultural and food chemistry.[2]

2.1.2 Food-derived Bitter compounds

Starting from the database of bitter compounds collected in Mayer 2019,[101] we refined a database of 247 natural TAS2R agonists, of which 133 are food-related TAS2R agonists, of those 83 were tested against all TAS2R receptors (activity value range: 1-10000 μM), here referred to as "food all-TAS2R-tested agonists dataset" (Figure 9). The compounds of this set activate 17 bitter taste receptors.

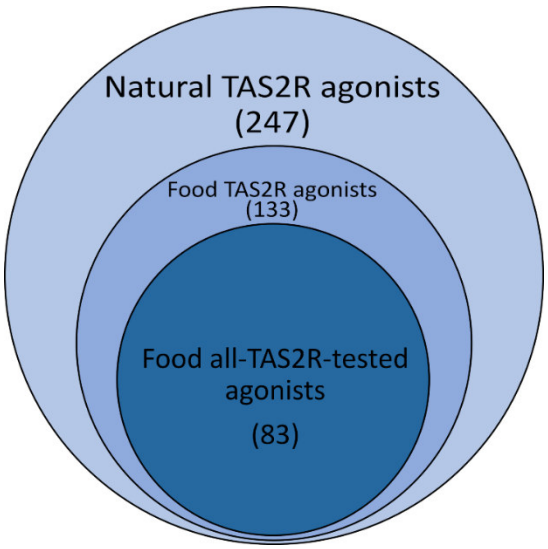


Figure 9 Visual representation of the relation between the natural TAS2R agonist set, the food TAS2R agonist and food all-TAS2R-tested agonist set. Picture from Bayer et al.[2]

2.1.3 Chemical classification

ClassyFire, a structure-based chemical taxonomy tool, which consists of up to 11 different classification levels, was used to classify and group collected TAS2R agonists.[102]

The kingdom is the highest level of classification, followed by superclass, class and subclass. Looking at the food TAS2R agonists set, 132 compounds belong to the kingdom of organic compounds. Only calcium chloride and magnesium chloride belong to a different kingdom, as they are inorganic compounds. Following 9 different groups were identified in the superclass level. The class level, then, shows 24 groups and the subsequent subclass 42 groups. As too many or too few groups make limited sense, especially considering the size of our data set, the superclass classification was used in the subsequent analyses.

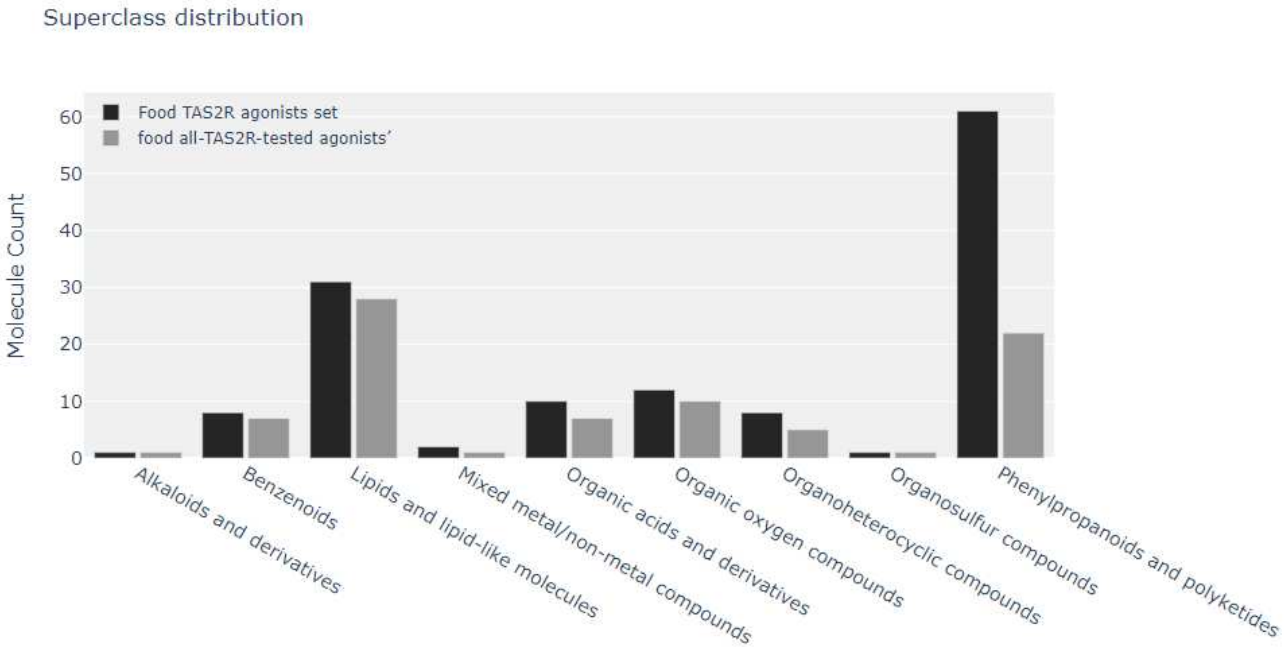


Figure 10 Superclass distribution of the food TAS2R agonist set (black) and the food all-TAS2R-tested agonist set (grey).

The food TAS2R agonists' dataset consists of phenlypropanoids and polyketides (61 compounds), lipids and lipid-like molecules (31), organic oxygen compounds (12), organic acids (10), organoheterocyclic compounds (8), benzenoids (8), mixed metal/non-metal compounds (2), organosulfur compounds (1) and alkaloids (1). A

similar distribution is seen in the all-TAS2R-tested agonists' dataset: lipids and lipid-like molecules (33), organic oxygen compounds (10), benzenoids (7), organic acids (7), organoheterocycles (5), mixed metal/non-metal compounds (2), organosulfur compounds (1) and alkaloids (1). Only the phenylpropanoids and polyketides share changes significantly, as about three third of the compounds are sorted out (22).

2.1.4 Receptor classification

Bitter taste receptors have a wide range of promiscuity. While some receptors are still orphan, like TAS2R60, others have over a hundred identified agonists.[103] Among the receptors activated by the food TAS2R agonists' set, TAS2R14 has the highest number of associated ligands, a total of 78. TAS2R39, -R1 and -R46 have significantly less associated, but are still broadly tuned with 47, 25 and 20 natural agonists, respectively. TAS2R4, -R10, -R7, -R5, -R16, -R40, -R43 and -R38, with 10±5 associated agonists, have intermediate selectivity. TAS2R30, - R31 and -R8 are more selective with 4, 2 and 2 natural agonists, respectively. Lastly, TAS2R20 and -R50 are the most selective, both with only a single associated molecule.

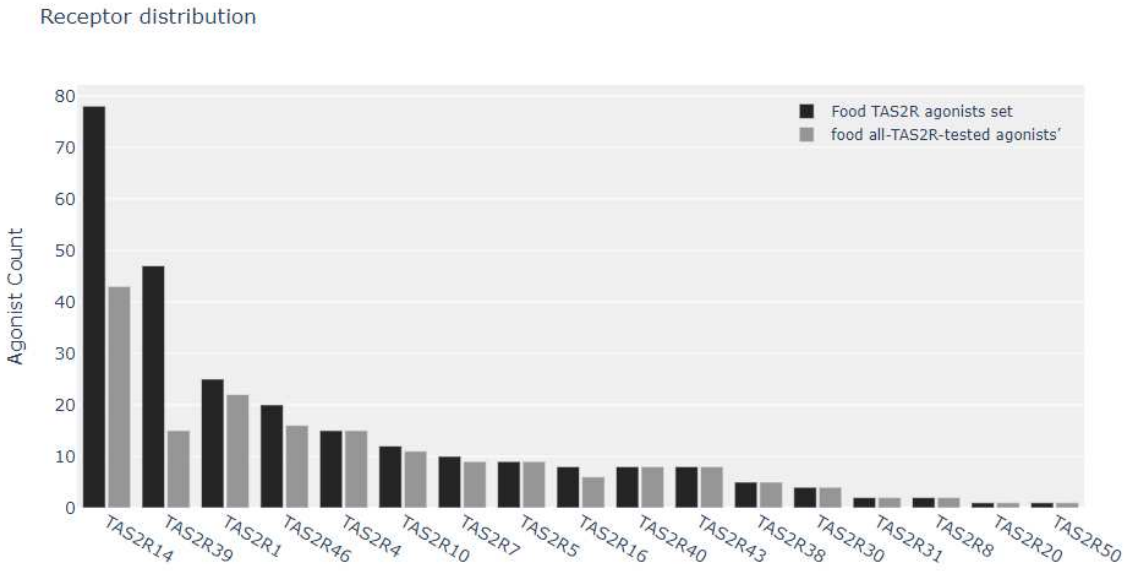


Figure 11 Agonist count per TAS2R receptor for both Food TAS2R agonist and all-TAS2R-tested agonist set.

Only the pure number of molecules can be misleading for promiscuity, as it does not consider structural diversity. To further investigate this, we performed a scaffold decomposition of the food all-TAS2R-tested agonist set and then calculated the number of scaffolds / unique scaffolds per agonist per receptor. We compared this

unique scaffold per receptor ratio with the ratio of agonists per receptor over the total number of agonists (Figure 12).

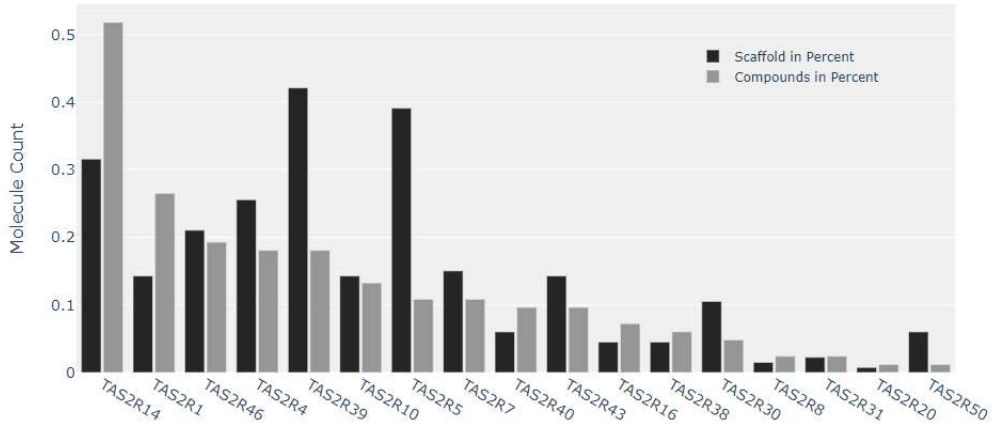


Figure 12 Unique Scaffolds per TAS2R receptor (black) and agonist count per TAS2R receptor (grey), both in proportion to the total number in the set.

Although TAS2R39 and TAS2R5 are activated only by half as many compounds as TAS2R14, they feature a higher number of unique scaffolds. TAS2R5 agonists also have the highest mean molecular weight (661). Indeed, the high number of unique scaffolds in this case relates to the complexity of the molecules. Instead, the molecular weight of TAS2R14 and TAS2R39 ligands in around 300 Dalton. Therefore, in these cases, the high number of unique scaffolds relates to structural diversity.

Table 1 Mean molecular weight of the agonists targeting different TAS2Rs.

Receptor	Mean Molecular weight
TAS2R5	661
TAS2R50	586
TAS2R4	559
TAS2R7	545
TAS2R30	423
TAS2R10	390
TAS2R14	376
TAS2R46	359
TAS2R16	352
TAS2R40	341

TAS2R39	322
TAS2R1	314
TAS2R43	314
TAS2R31	272
TAS2R38	247
TAS2R20	204
TAS2R8	160

The comparison of the promiscuity of our receptors in relation to the superclasses and their associations are shown in Figure 13. The aim of this analysis was to investigate a possible relation between chemical classes and promiscuity. Therefore, the food all-TAS2R-tested agonist set was used for this analysis.

TAS2R14 is the only receptor with ligands from 7 different superclasses. Next, TAS2R1, -R4 and -R43 come with agonists from 6 superclasses. TAS2R46, -R39, -R10, -R38, -R7 and -R40 follow with five and four superclasses. While TAS2R38 has only five agonists in the set, all of them come from different superclasses. Similarly, TAS2R30 has four agonists and associated superclasses. TAS2R31 is the only receptor with two superclasses, while there are five with one. TAS2R8 is activated by 2 organic acids. TAS2R50 and -R20 are both activated by only a single ligand, amarogentin and L-tryptophan, respectively.

TAS2R16 and -R5 are also activated by only one superclass, however, they have multiple ligands within that respective group. TAS2R16 agonists are 8 organic oxygen compounds. This selectivity is well known as the receptor is specialized in the recognition of “bitter sugars”, such as glucopyranosides and glucosinolates.[104–106] Similarly, TAS2R5 is activated by 9 compounds from the phenylpropanoids and polyketides superclass. Such polyphenols have a positive impact on the human gut,[107,108] and as TAS2R5 is expressed in the intestines, it may play a role in in their recognition.[109]

The right-hand side of the plot focuses on the compound promiscuity. It represents how many bitter taste receptors each compound can activate (five classes were set: 1, 2, 3, 4 and >4 targeted receptors). Their respective superclass is shown as well. Most compounds target one or two receptors. The mixed metal/ non-metals as well as

the organosulfur compounds are the most selective groups, only targeting a single receptor. They are also the only groups that do not target TAS2R14. Furthermore, both groups are limited in size, with only a few molecules associated.

Even smaller is the alkaloids and derivates group, consisting of only a single molecule, quinine, which is one of the most promiscuous compounds in the set. It targets a total of 8 receptors. This highlights that the size of a superclass is not the only driver for its promiscuity, as the whole set of 44 lipids and lipid-like molecules only activate one more receptor (9) than quinine. The majority of superclasses activate between six and nine receptors. The phenylpropanoids target the most, followed by the organoheterocyclic compounds and benzenoids with 11 and 10 superclasses, respectively.

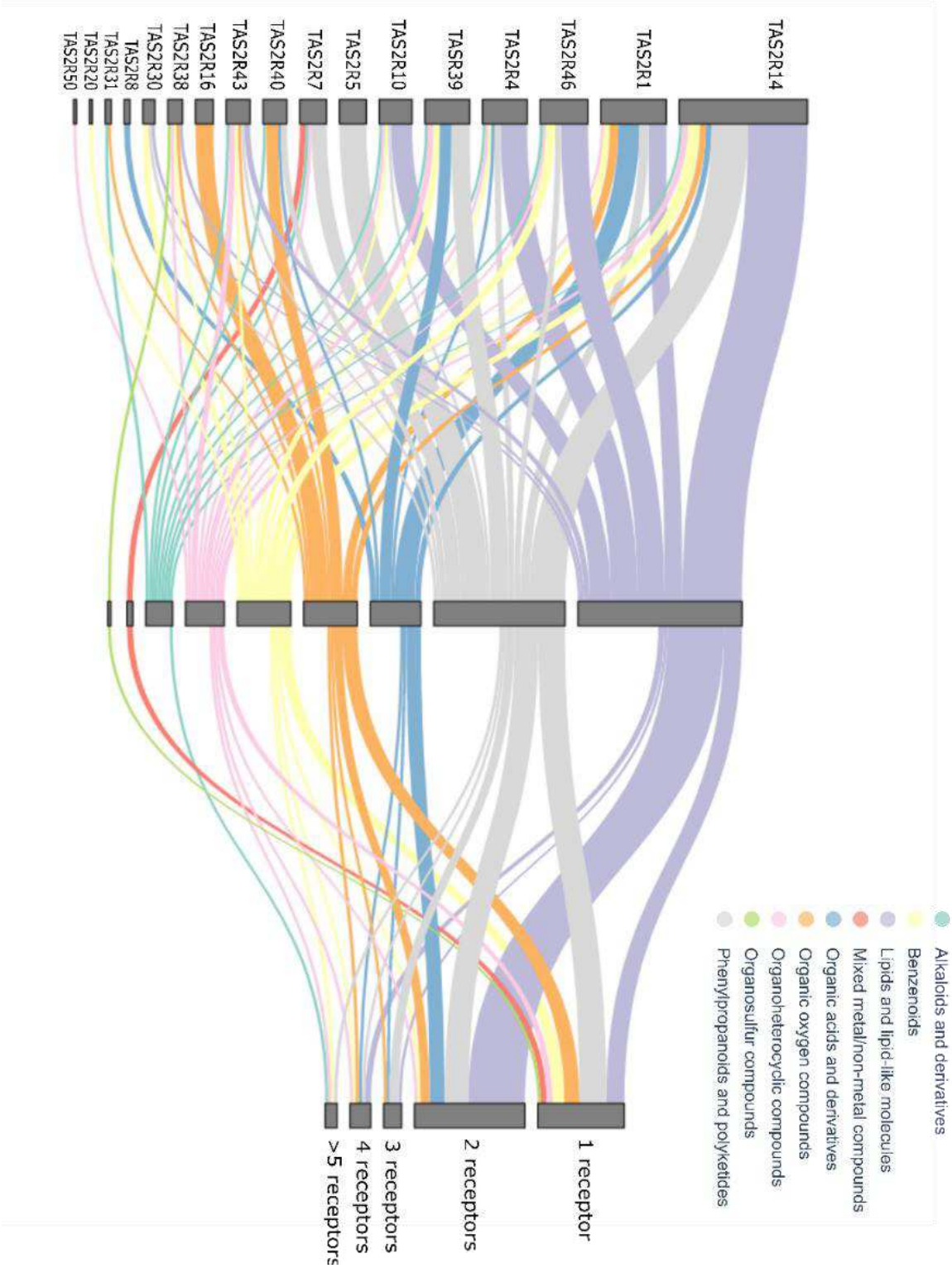


Figure 13 Alluvial plot highlighting the complexity between the different TAS2R receptors, their promiscuity (left) and the molecules targeting them (coloured lines). The superclass of the different compounds is highlighted in different colours and converge in the middle of the plot. On the right-hand side, the compound promiscuity is displayed. Picture from Bayer et al.[2]

The grouping of our compound data set into 9 superclasses in concert with the receptor associations gives an insight into the complexity of their relation. However, classification always comes with generalisation. Especially when looking at highly complex systems like molecules. For instance, quinine is classified as alkaloid, but caffeine is with the heterocycles. Thus, an accurate representation of their chemistry is limited. To overcome those issues, a t-distributed stochastic neighbour embedding (t-SNE) chemical space analysis was performed. The goal of the chemical space analysis was to include an additional structural characterization and comparison of the compounds. Fingerprints are used are molecular descriptors, and the distribution in the 2D space define the structural similarity (molecules close in space are similar, while molecules far in space are dissimilar). In Figure 14, compounds in the t-SNE plot are coloured by superclasses. The whole food TAS2R data set is plotted, with the food all-TAS2R-tested agonists highlighted using an increased dot size. Here, we see that superclasses from both sets populate the same chemical space. This is in line with the approach of many experimental screenings for compounds towards specific TAS2Rs, where they focus on chemically similar molecules such as which were previously screened towards all receptors.[110,111] This also means that we introduce no bias by only looking at one of those data sets.

Also, superclasses tend to form cluster in subregions of the chemical space. The most abundant group, the phenylpropanoids and polyketides (grey), conglomerate on the negative space. Reversely, lipids and lipid-like molecules (violet) are found on the positive side of the y-axis. The main cluster of benzoids, located close to the centre, consists of smaller one-ring structures (phenylethyl isothiocyanate, protocatechuic acid, pyrocatechin, salicylic acid, and vanillic acid), however the larger molecules amarogentin, 11β,13-dihydrolactucopicrin and lactucopicrin spread to the bottom and left side, respectively. Calcium and magnesium chloride, the two mixed metal/non-metal compounds cluster at approx. 500/0. The organoheterocyclic compounds mostly stay close to the origin of x axis but spread all around the y axis. While they are all organoherocycles they differ vastly in the number and kind of heteroatoms they incorporate.

The organic oxygen compounds (orange) build two clusters at opposite sides of the plot, one on the lower right side and a second one on the upper left. The second cluster consists of humulones (humulone, adhumulone and cohumulone) and is very close to

the three organic acids (blue) adlupulone, colupulone and lupulone. Those six compounds are closely related to each other and different from the other organic acids which are nearer the x axis (all di- and tripeptides). Similarly, allyl isothiocyanate (green) is very close to phenylethyl isothiocyanate (yellow dot), and both share the defining SCN group. Those examples highlight the ability of t-SNE to identify structurally similar molecules assigned to different ClassyFire superclasses.

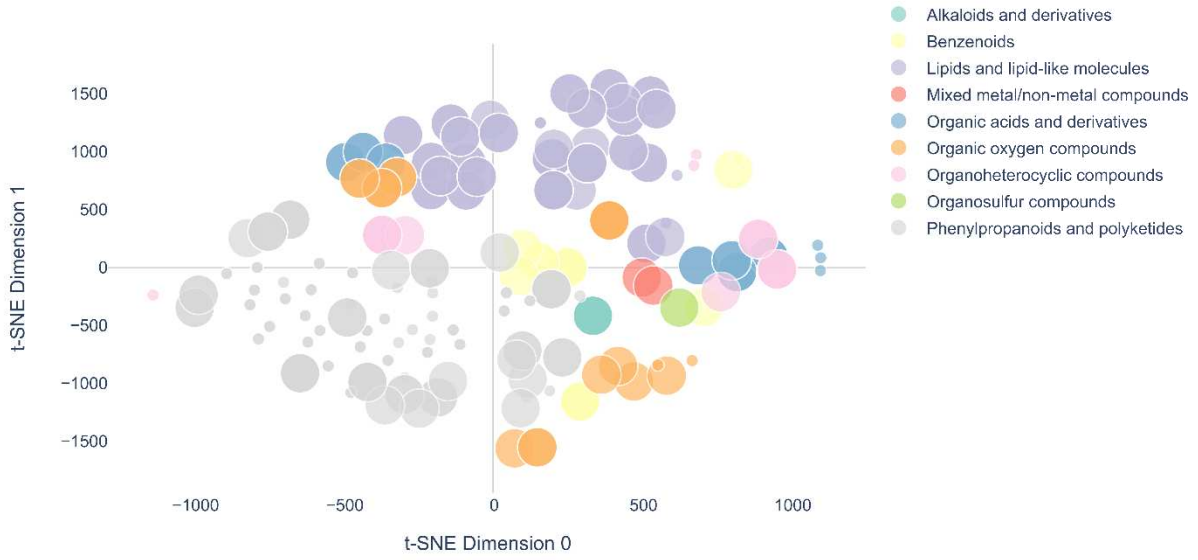


Figure 14 t-SNE plot of the food TAS2R agonist set, coloured by superclass and with the subset of food-all-TAS2R-tested agonists highlighted as larger dots. Picture from Bayer et al.[2]

In Figure 15, the food all-TAS2R-tested dataset is plotted in grey, and the size of the dots is defined by the number of receptors targeted. Most compounds are selective, meaning they target only one or two receptors. 12 compounds are labelled as intermediate, meaning they target three or four receptors and only four compounds target more receptors. Those are caffeine, quinine, amarogentine and epigallocatechin gallate (EGCG). Quinine activates 9 receptors (TAS2R4, -R7, -R10, -R14, -R31, -R39, -R40, -R43, and -R46), amarogentin 7 receptors (TAS2R1, -R4, -R10, -R30, -R39, -R43, and -R46), EGCG 6 receptors (TAS2R4, -R5, -R14, -R30, -R39, and -R43), and caffeine 5 receptors, (TAS2R5, -R9, -R14, -R43 and -R46). Interestingly, TAS2R43 is the only receptor targeted by all four compounds.

As it was previously seen in the Sankey plot (Figure 13), the superclass is not a predictor for compound promiscuity. Here we see that the position in the chemical space does not seem to predict this property either. Promiscuous, intermediate, and selective compounds are spread all around the map. One striking example of this is caffeine versus theobromine. While they are structurally alike, in the same group and close in space in the t-SNE plot, they are opposed in selectivity. Caffeine targets five receptors, while theobromine activates only TAS2R14.

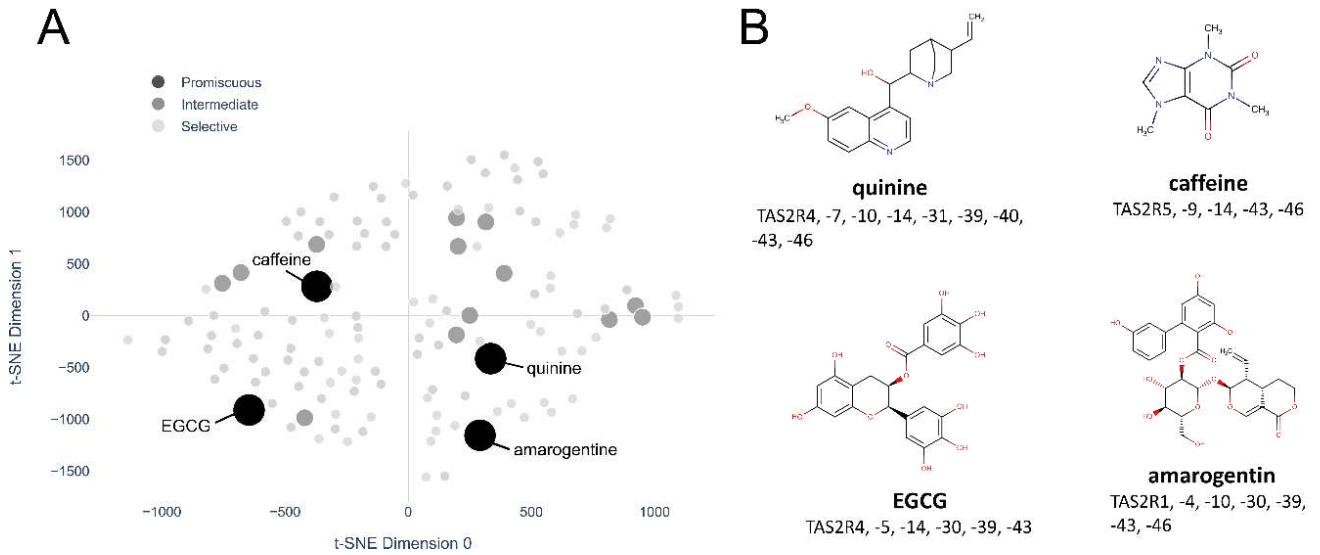


Figure 15 t-SNE plot of the food-all-TAS2R-tested set, colour, and size dependent on the promiscuity of the compound (A). Structures of the most promiscuous compounds are shown on the right-hand side (B). Picture from Bayer et al.[2]

One goal of our study is to investigate the receptor space covered by food-derived bitter compounds. This can be done by coloring the t-SNE plot by the receptors targeted. This analysis is quite complex because we have more receptors than superclasses, 17 instead of 9. Moreover, many compounds target multiple channels, thus making visualization troublesome. In Figure 16, only the agonists selective towards a single receptor are plotted and coloured according to the associated receptor.

Interestingly, we can see distinct clusters for many receptors: TAS2R4 in the upper right hand side, TAS2R5 around the y-axis in the lower side of the plot, TAS2R16 lower right hand side. Interestingly, TAS2R14 and -R39, the two receptors with the highest agonist count, build overlapping clusters. In the Supplementary information of our published paper[2] distinct plots for each receptor can be consulted.

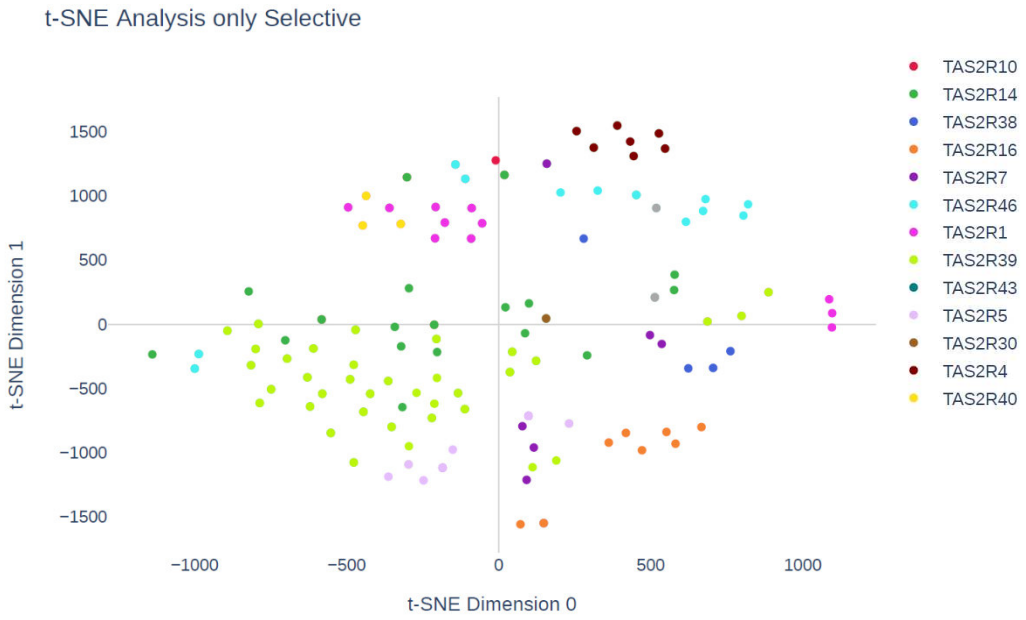


Figure 16 t-SNE plot featuring only the selective ligands, coloured by receptor targeted. Picture from Bayer et al.[2]

2.2 Chemesthetic TRP ligands: ligand- and structure-based models

Our goal was to gather currently known natural agonists of chemosensory TRP channels and use this data to create computer aided drug discovery (CADD) models to better understand the ligand-receptor interactions and help in the search for future ligands.

We used ligand- and structure-based approaches. Ligand-based pharmacophores find and display the shared molecular recognition features between a group of ligands, which bind the same receptor site. According to the IUPAC definition,[112] a pharmacophore model is 'an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response.

Structure-based approaches, on the other hand use the binding site of a macromolecule to extract information on how ligands can bind. Molecular docking is the most common structure-based tool, and samples possible conformations of ligands within the target.

The dataset used to build our models consists of 133 naturally occurring TRP modulators. It contains 91 TRPA1, 7 TRPM8, 34 TRPV1, and 38 TRPV3 agonists. An in-depth analysis is provided with each receptor. Furthermore, compounds which were tested against TRP channels and shown to be inactive were collected as well. The whole dataset is available at https://github.com/dipizio/natural_TRP_modulators.

2.2.1 TRPA1 ligands

TRPA1 has the largest agonist set and features ligands with vastly different binding mechanisms. As explained in the introduction, agonists can either act via classical non-covalent interactions or via covalent bonding through an electrophilic group. To create fitting models, we split the molecules into electrophilic (such as α,β, unsaturated ketones) and non-electrophilic. Some exceptions were made, as molecules with an electrophilic group do not necessarily engage via covalent binding. For instance, capsiate, a capsaicin-like molecule has a necessary electrophilic group, but experiments have shown that its activity is not diminished in C621 mutants.[113] Therefore, it and other electrophiles, where such information was available were grouped with the non-electrophilic. Furthermore, electrophiles with an isothiocyanate group received their own subset as they have different features.

2.2.1.1 Isothiocyanates

20 natural isothiocyanates (ITCs) that activate TRPA1 were found from literature. Most of them origin from horseradish, mustard, or wasabi. They are small molecules ranging mostly from 100-200 Dalton, with a mean of 150. All molecules and physicochemical descriptors are listed in https://github.com/dipizio/natural_TRP_modulators. Apart from the characterizing S=C=N group, they do not have specific chemical moieties, the average of HBDs (H-bond donors) is zero and the average of HBA (H-bond acceptor) is 1. With 311 Dalton, moringin is the biggest molecule of the group, it has alcohols and ketone groups, and distinguishing itself notably from the other compounds in this set.

The small size, the lack of important functional groups and the sulphur-nitrogen group set the ITCs apart from other TRPA1 agonists. However, it also makes it hard to define further boundaries for models, as the main determining factor seems to be the presence of the S=C=N group. A recent publication in the journal of Natural Products came to a similar conclusion by testing a smaller set of ITCs in vitro. Most showed similar activity and neither the carbon chain length, nor the different substituents showed a significant impact on the activity.[114]

We approached the issue by pharmacophore hypothesis development. The pharmacophore modelling can identify specific features and distances even among small and simple molecules. In a pharmacophore model, indeed, the shared important features, such as hydrogen bond donors (HBDs) and hydrogen bond acceptors (HBAs), and their geometrical orientation are distilled from a group of actives and inactives. In our case, we used Phase [115] as software to develop the modes, the EC50 values of ITCs as activity value to rank the active molecules, and a decoy set was generated for the set of inactive molecules (see paragraph 3.2.2.2.1 for more details). The outcome is a set of hypotheses ranked according to their power to discriminate actives vs. inactives, i.e., by the false positive vs false negative rate via a ROC curve, which features the external validation of our decoy/active set and by the phase hypo score which uses a set of geometric and heuristic criteria in combination with an internal test set provided to classify the models.[115] For both a higher score is desired.

We used two different feature set ups, first we used the standard features (i) and then added a custom S=C=N group for the second run (ii). For the first try we found that the best scoring pharmacophores all featured two HBA features and one HBD group.

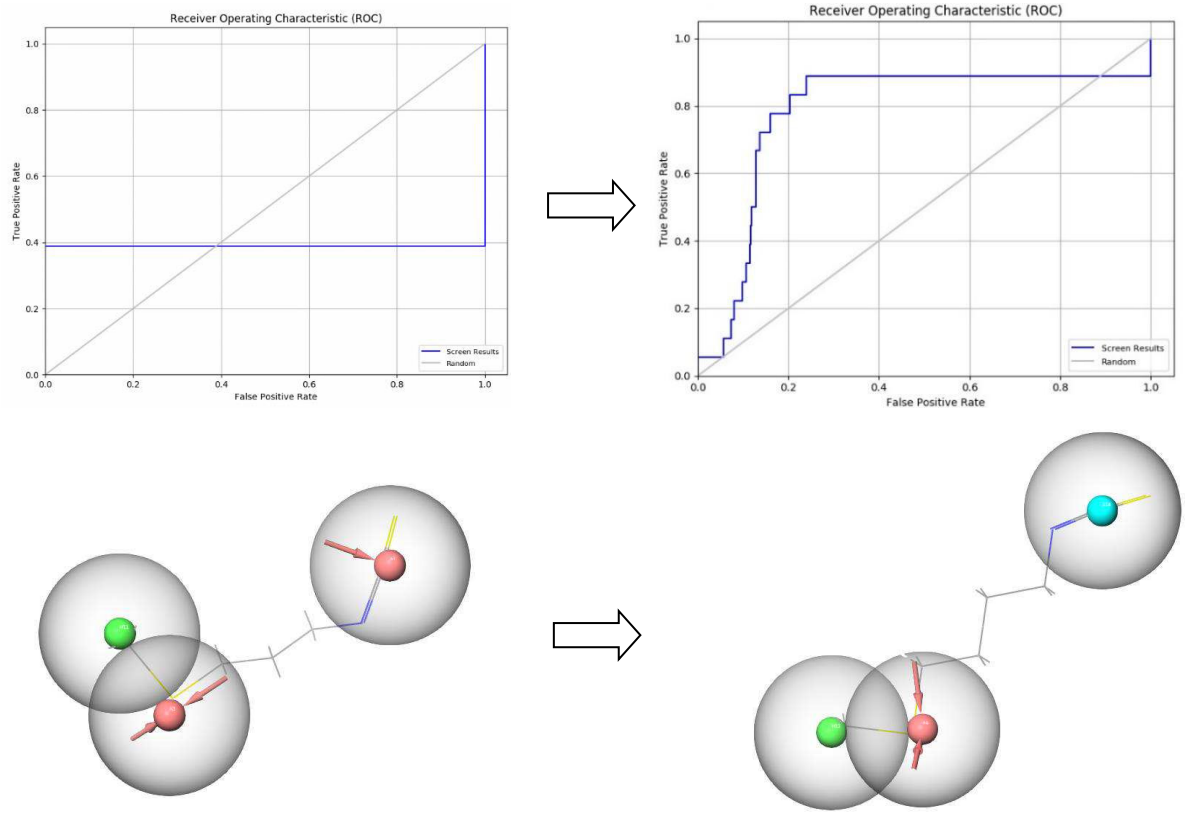


Figure 17 ROC curves and pharmacophore hypotheses for the TRPA1 ITC set with default features (i) and with ITC custom feature (ii).

The models with the S=C=N inclusion yielded similar results. Here however one acceptor group was replaced with the ITC group. The program gave both options similar phase hypo score. But, interestingly, the validation with the decoys, the ROC curve of the pharmacophores with S=C=N feature outranked the other group by far (Table 2). This is probably due to the more precise orientation of the ITC feature that increase the performance of the model increased selectivity in comparison to the standard acceptor feature. It allows to further discriminate between generic hydrogen acceptor groups and the electrophilic ITC group, ruling out false positives. Furthermore, the restricted geometric orientation implied in the HBA feature is not present in the SCN feature, which could allow for less false negatives.

Table 2 Best scoring TRPA1 ITC agonists pharmacophore hypotheses, with default features (i) and after the inclusion of the custom ITC group (ii). Phase Hypo scores and validation scores (ROC AUC) values are reported.

Feature settings	Model name	Reference molecule	Phase Hypo Score	ROC AUC
Default (i)	AAH_4	3-methylthiopropyl isothiocyanate	0.68	0.39
	AAH_5	3-methylthiopropyl isothiocyanate	0.63	0.39
With custom ITC feature (ii)	AHX_3	6-(methylthio)hexyl isothiocyanate	0.68	0.78
	AHX_4	6-(methylthio)hexyl isothiocyanate	0.65	0.60

2.2.1.2 Non-ITC Electrophilic compounds

In comparison to the ITCs, the other electrophilic ligands tend to be bigger, with a mean molecular weight of 220 and an average of three HBAs and zero HBDs (see the dataset with calculated physicochemical properties at https://github.com/dipizio/natural_TRP_modulators). Like the ITCs, their activity is mainly defined by their covalent-bond-forming group. As those are more complex and diverse, finding a fitting pharmacophore hypothesis proved troublesome. After multiple different tries the best result still only yielded a ROC curve with AUC value of 0.51.

Therefore, a Quantitative Structure-Activity Relationship (QSAR) model was used in this case. In this approach, a plethora of molecular properties, ranging from topological to physicochemical descriptors are calculated for a set of actives. Statistical models are then used to correlate activity values and molecule structures (i.e. calculated descriptors). All built models can then be validated with test sets.

The best model used 20 descriptors (Table 3) and had S.D. and R^2 values of 0.36, and 0.91, respectively, on the training set, and RMSE and Q^2 of 0.21 and 0.96, respectively, on the internal test set (Figure 18). Many of the descriptors used for our model (i.e., maximal electrotopological negative variation, mean topological charge index, Michael acceptors and Michael donors) can be associated with the electrophilicity of a molecule and reflect the importance of the electrophilic group; others refer to pharmacophoric properties of the compounds (i.e., HBDs, ring count).

Furthermore aldehyde, carbonyl and neutral carbonyls also describe functional groups which are associated with an increase / decrease in dipole moment.

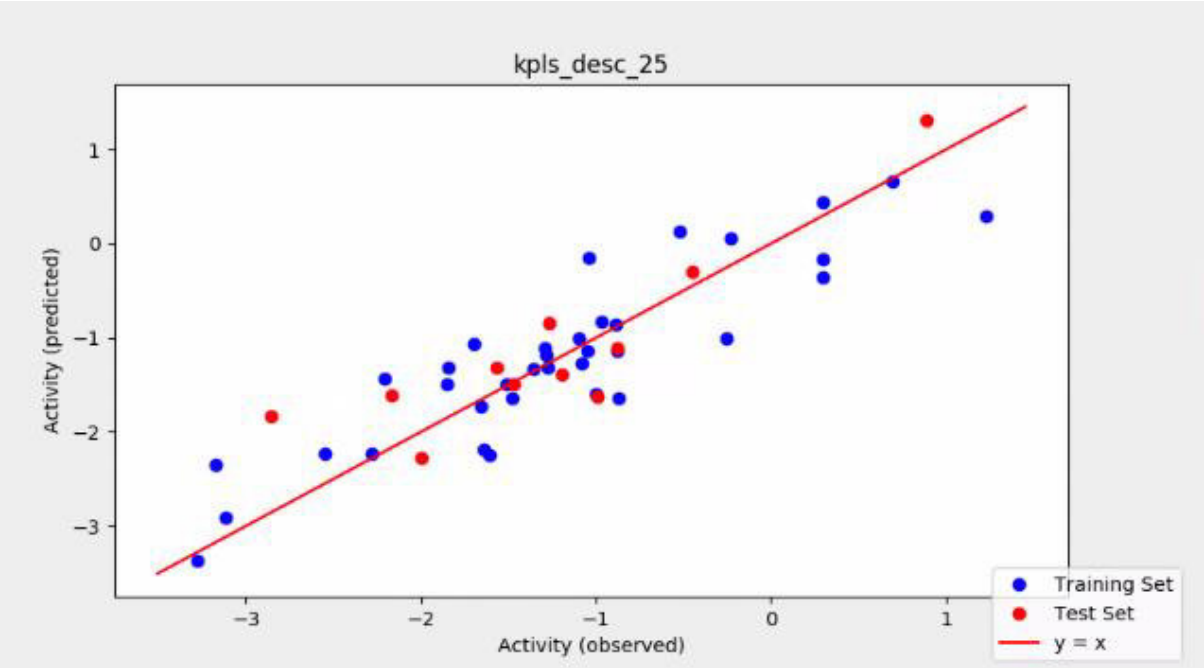


Figure 18 Visualisation of the top scoring QSAR model for the TRPA1 electrophile set.

Table 3 Descriptors of the best scoring QSAR set for TRPA1 electrophilic ligands.

1	i canvas HBD	11	i ligfilter Michael acceptors
2	i canvas dCH2 Cnt	12	i ligfilter Michael donors
3	i canvas dssC Cnt	13	i ligfilter N
4	i canvas sCH3 Cnt	14	i ligfilter Neutral carbonyls
5	i canvas ssssC Cnt	15	r desc 1-path Kier alpha-modified shape index
6	i desc Ring Count 5	16	r desc ALOGP4
7	i desc Ring Count 6	17	r desc ALOGP5
8	i desc Sum of topological distances between N. O	18	r desc Maximal electrotopological negative variation
9	i ligfilter Aldehydes	19	r desc Mean topological charge index of order 4
10	i ligfilter Carbonyls	20	r desc Mean topological charge index of order 6

2.2.1.3 Non-electrophilic binders

21 TRPA1 agonists with no strong electrophilic group or that have been shown to not interact with the electrophilic binding site were collected (https://github.com/dipizio/natural_TRP_modulators). The electrophilic agonists are more chemically diverse than the ITC group and the non-electrophiles. Their molecular weight ranges from 136 to 488, the number of HBAs from 0 to 5 and the HBDs from 0 to 7. The two largest molecules are leucettamol A and B, which are lipid like compounds identified from a marine sponge. Endogenous lipids are known to be involved in TRPA1 modulation, and their presence is important for their natural function.[116,117] The smallest agonist is limonene, a monoterpenoid. While capsaicin itself does not activate TRPA1, this set does feature some capsaicin-like compounds, namely capsiate and nordihydrocapsiate. Nicotine and Δ9-Tetrahydrocannabinol, the main active ingredients of tobacco and cannabis respectively, are also in this set.

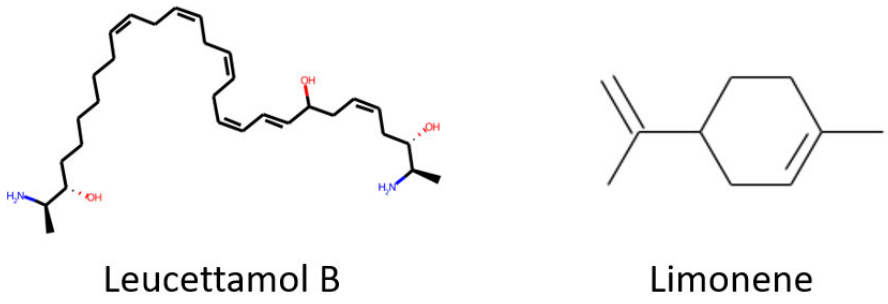


Figure 19 The largest non-electrophilic TRPA1 agonist in this set (left) and the smallest (right).

A ligand-based pharmacophore modelling approach was used for investigation of this set of compounds. The developed pharmacophore models share similar feature combinations. The two top scoring models, AHH_2 and AHR_2, feature one hydrogen bond acceptor and two hydrophobic features and one hydrogen bond acceptor, one hydrophobic and one aromatic ring feature, respectively. (Figure 20).

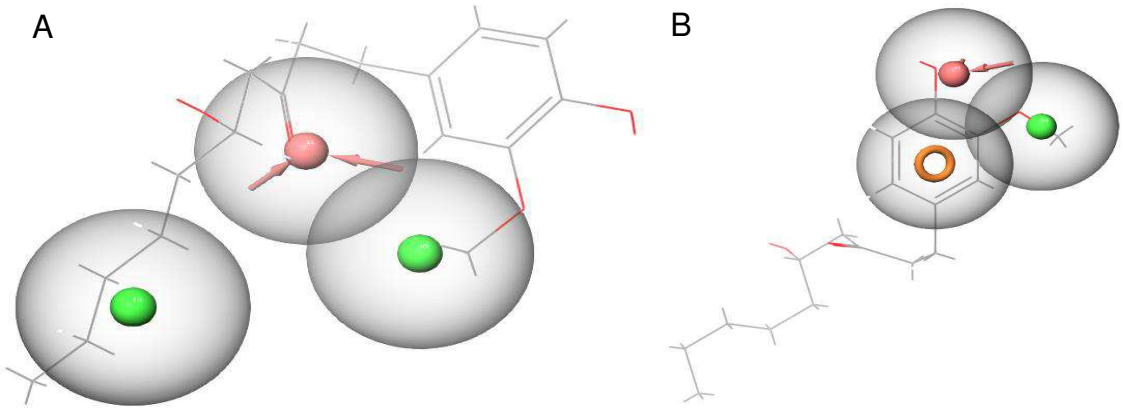


Figure 20 Top scoring pharmacophores for TRPA1 non-electrophilic agonists. On the left (A), the model AHH_2 is shown and, on the right (B), model AHR_2.

Table 4 Statistics for two top scoring pharmacophores for TRPA1 non-electrophilic agonists.

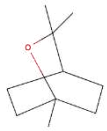
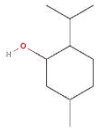
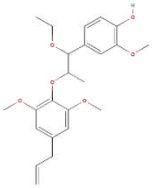
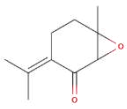
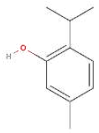
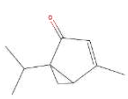
Models	Reference molecule	Phase Hypo Score	ROC
AHH_2	6-Paradol	0.73	0.74
AHR_2	6-Gingerol	0.67	0.71

Statistics for the top scoring models after validation are shown in Table 4. Interestingly, a deeper look revealed that the models had problems to match leucettamol A and B. The binding pocket for those ligands is not confirmed yet and lipid binding is also not fully understood. They might bind in different pockets and/or interact with a different binding mode. Structural investigation into the binding mode of these compounds within TRPA1 will in future clarify the correctness of our model in identifying leucettamol A and B as outliers of this set.

2.2.2 TRPM8 ligands

Few TRPM8 ligands are known and most of them are antagonists.[118] In total only 7 natural agonists were found in this study. Those are shown in Table 5 below.

Table 5 TRPM8 natural agonist data set.

Compound	Structure	Source	Reference
1,8-Cineole		Eucalyptus	119
Borneol		Thyme	120
Menthol		Mint	121
Erythro-/Threo-Δ8′-7-ethoxy-4-hydroxy-3,3′,5′-trimethoxy-8-O-4′-neolignan		Nutmeg	122
Rotundifolone		Mentha L.	123
Thymol		Thyme	124
Umbellulone		California bay laure	125

Menthol is the most renowned TRPM8 agonist. Thanks to its cooling & refreshing taste it has been used to flavour gum, sweets, tea, chocolate, sodas, and even cigarettes.[126] It is also used in pharmaceuticals to increase dermal penetration and as an active ingredient itself for its antifungal, antibacterial and analgesic affects. Furthermore, it has been proposed as an anti-cancer agent.[127]

1,8-Cineole, also called eucalyptol, has strong anti-inflammatory properties.[128] This compound activates the human TRPM8 about eight times stronger than the mouse version (EC50 of 120 µM and 925 µM, respectively). Making it an interesting compound for studying the differences of those channels.[128]

Rotundifolone, another monoterpene from Mentha x villosa Hudson leaves, has been shown in Ca^{2+} imaging approaches to activate TRPM8 more selectively than menthol in respect to other TRP channels.[123]

Borneol has been used in traditional Chinese medicine for treating dry eye disease and its effect has then been shown due to its activation of the TRPM8 channel.[120] In contrast to menthol, its activation is temperature-dependent and abolished at temperatures over 35°.[120]

Umbellulone, identified from the leaves of the so-called “Headache tree”, has been reported as a “weak” TRPM8 agonist.

Thymol is the newest discovered agonist, as its activity has been found less than a year ago by Wang et al. They tested the monoterpene on dorsal root ganglion neurons in mice. They identified a Ca^{2+} response that was not found in TRPM8 deficient models and blocked after the addition of a TRPM8 antagonist.[124]

The TRPM8 agonists mentioned thus far are relatively similar in structure and are proposed to activate the channel in a similar fashion or at least in the same binding site. In 2017, Shirai et al. identified two neolignane agonists, erythro- and threo-Δ8′-7-ethoxy-4-hydroxy-3,3′,5′-trimethoxy-8-O-4′-neolignan, which appear to act differently. While they activate the channel on their own, they also increase the potency of menthol binding. Hence, they seem to bind in a different region. SAR studies of those neolignane compounds revealed that the 4-hydroxy group is essential, while longer or bulkier alkoxy groups at C7 are also important for activity.[122]

In addition to the seven TRPM8 agonists in this set, 18 natural compounds where collected, which were proposed as TRPM8 agonists but have been shown in cell experiments to not interact with the channel. Those will here be referred to as inactives and were on average smaller, with a lower number of hydrogen bond donors and acceptors. Plus, several contained a S=C=N group.

Table 6 Difference in descriptors between TRPM8 Actives and Inactives

	Average HBA	Average HBD	Average MW
Inactives	1,39	0,33	116
Actives	1,86	0,57	190

2.2.2.1 Ligand-based pharmacophore modelling

A ligand-based pharmacophore model was developed using the actives and inactives in the dataset. As the aim of the pharmacophore model is to abstract the ligand features interacting with the receptor by a ligand superimposition the neolignan compounds were excluded from the analysis, as they do not share the same binding site as the other TRPM8 agonists.

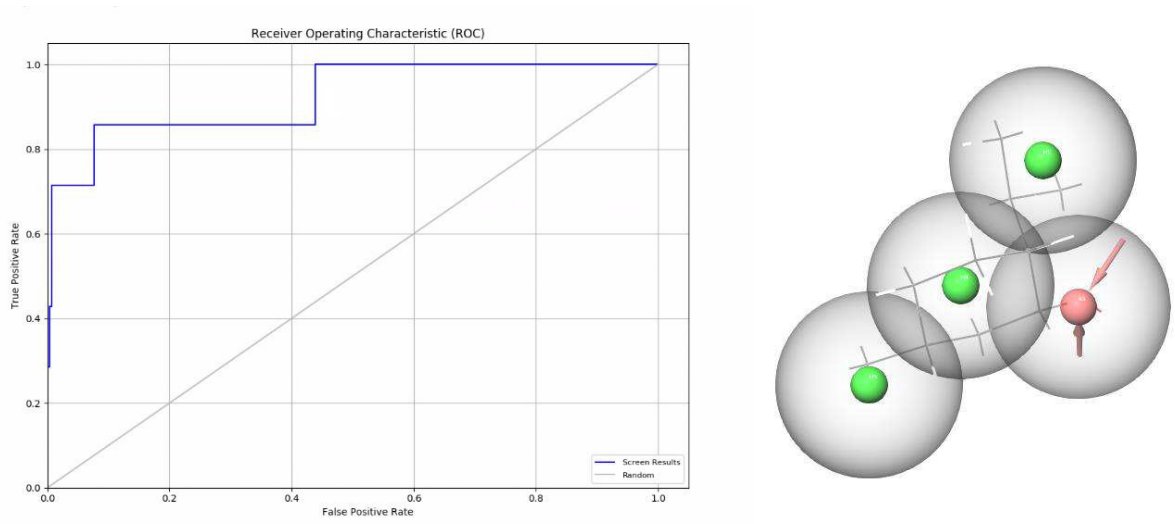


Figure 21 ROC curve (left) for the top scoring pharmacophore model (right) for the TRPM8 agonist set.

The best scoring models featured one hydrogen bond acceptor and two to three hydrophobic features (Figure 21). Indeed, the analysed agonists are all quite similar, with a central ring structure, two hydrophobic groups attached in para position and a

hydrogen bond acceptor group in between, in the form of an oxygen. The statistics for those initial models looked very promising, with high phase hypo score and ROC-curves. However, as there are so few ligands known we were not able to properly validate those models.

Table 7 Top scoring pharmacophore hypothesis models for TRPM8 and their statistics.

Models	Reference molecule	Phase Hypo Score	ROC AUC
AHHH_4	Menthol	1.15	0.92
AHH_3	1,8-Cineole	1.16	0.91
AHHH_1	Menthol	1.13	0.90
AHHHH_4	Menthol	1.11	0.90

2.2.2.2 Structure-based modelling

Although a lot of research has been done on TRPM8 in the past decades, only avian crystal structures are available.

In such cases, a structure prediction approach is pursued to build the structure of interest. A recent review on computational techniques used for protein structure prediction, advantages, limitations and current development was published in a book chapter I have been contributing during the internship time.[129] When the experimental structures of similar structures are available, homology modelling is commonly used.[130]

Very recently, Xu et al have built an accurate homology model of TRPM8 to explore the molecular mechanisms underlying the binding of menthol.[131] They found two key interactions. First the hydroxyl creates a hydrogen bond with R842 and secondly, the isopropyl group forms van der wall interactions with foremost L843 and I846.[131] The 3D coordinates of the model were made accessible through the Supporting Information of this publication.

Following these recent findings, we docked our TRPM8 agonist set in the binding site of their homology model. The output conformation of menthol after the docking matched the position and interactions observed by them, as can be seen in Figure 22. The hydroxyl group points towards R842 and the isopropyl group in the direction of L843 and I846.

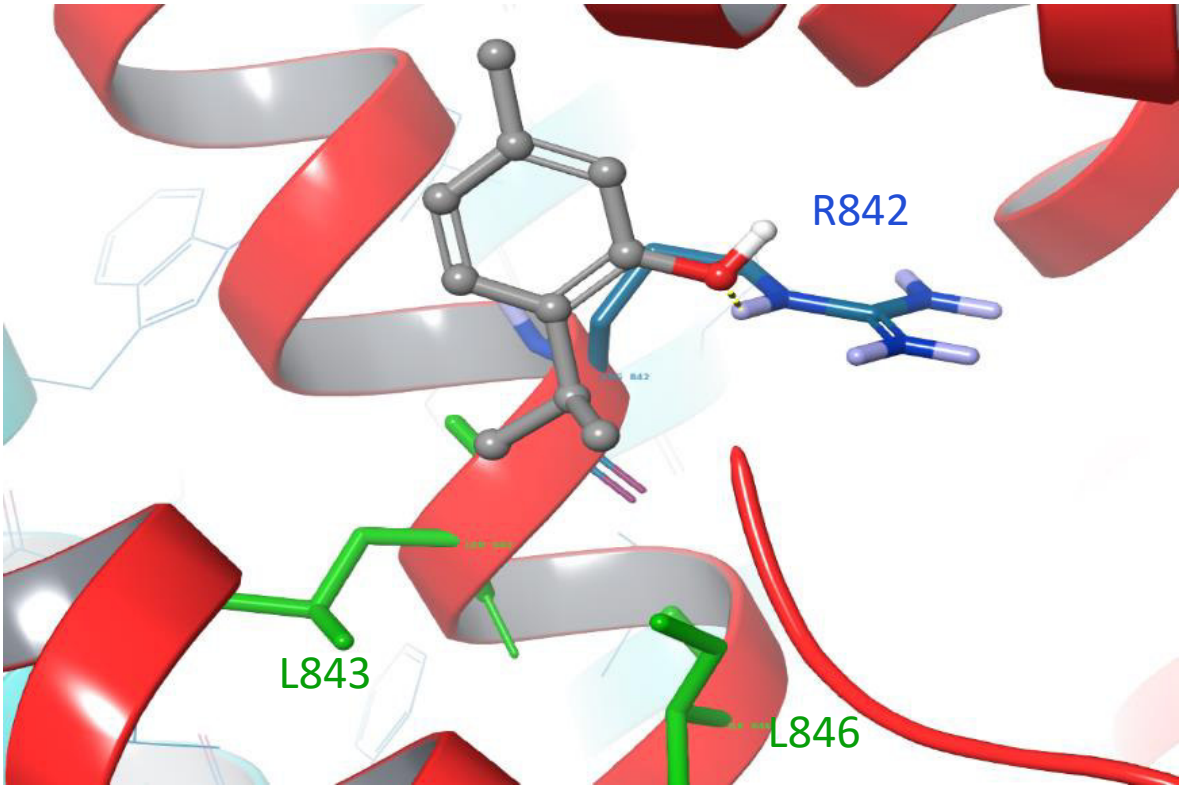


Figure 22 Predicted binding pose of menthol (grey stick) within the TRPM8. The isopropyl groups face towards L843 and L846 (hydrophobic residues shown as green sticks) and the hydroxyl group creates a hydrogen bond with R842 (shown in blue).

The docking outcome of the other ligands is shown in Figure 23 and a comparison of the docking scores is found in Table 8. For thymol, umbellone and rotundifolone we found the isopropyl group facing toward L843 and I846 and an H-bond interaction with R842, indicating very similar interactions as for menthol.

1,8-Cineole and borneol showed less similarity. The isopropyl groups are facing in the same direction as menthol, but the H-bond with R842 is not formed. This lack of interaction was also reflected in the docking score. Both showing a lower value than the other docked ligands.

Table 8 Docking Scores obtained for the investigated TRPM8 ligands

Ligand	Glide Score
Thymol	6.662
Menthol	6.471
Umbellulone	6.339
Rotundifolone	6.281
1,8-Cineol	5.904
Borneol	5.686

The best docking score was obtained for thymol, even slightly higher than menthol. The main difference in those molecules is the higher planarity and rigidness of thymol's aromatic ring. Our results indicate that thymol might bind even better than menthol, however no literature value is available for the strength of thymol binding nor a comparison to menthol.

Umbellulone and rotundifolone bulky side groups yield "bad" contacts and might thus be the reason for the lower score. Furthermore, umbellulone's higher flatness of the ring caused by the cyclopropane structure might play a role as well. Conversely, derivates with an alcohol substituted for the ketone show increased activity. Bulkier groups, however, like an acetyl, dimmish the activity, indicating sterical hindrance.[125]

Taken together these results leave only a narrow space for future natural TRPM8 ligands. However, as Shirai et al. demonstrated with novel neolignane agonists from nutmeg[122] there is room for more unexpected, allosteric compounds.

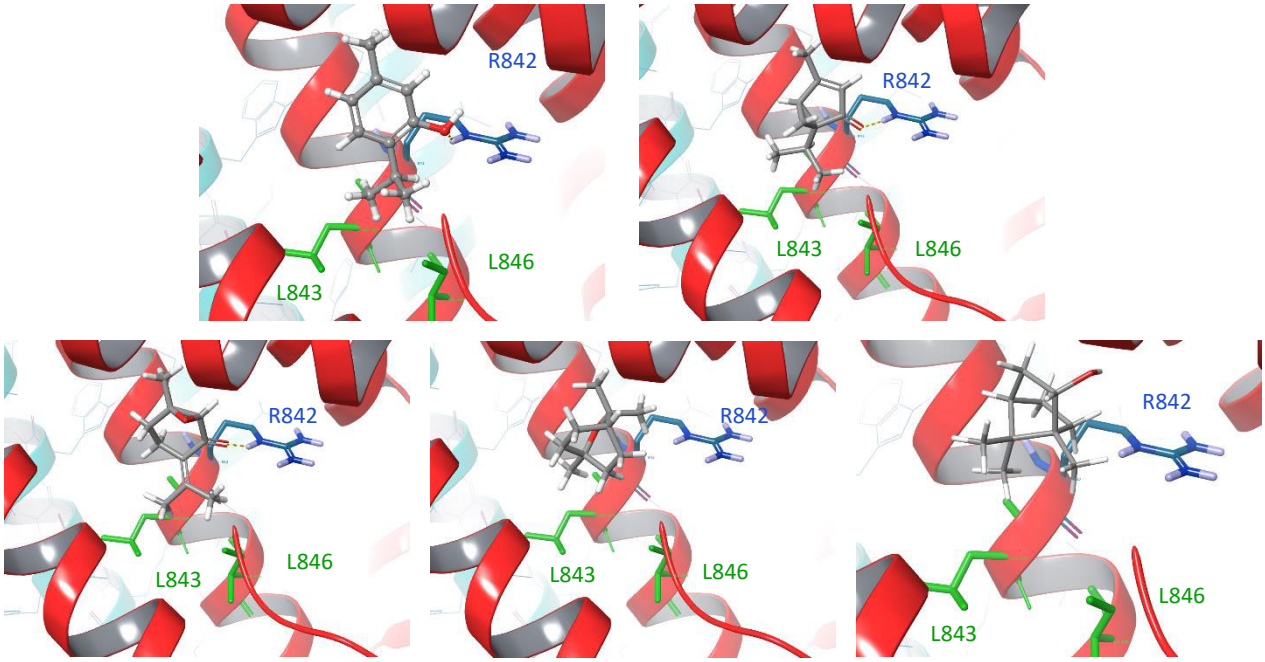


Figure 23 Docked TRPM8 ligands from left to right and top to bottom: Thymol, Umbellulone, Rotundifolone, 1,8-Cineole and Borneol. Docked structures are shown in grey, important residues in green and blue.

2.2.3 TRPV3 ligands

We collected a total of 38 natural TRPV3 agonists (mostly monoterpenoids) and 5 known inactives. All structures and activities are available at https://github.com/dipizio/natural_TRP_modulators.

2.2.3. 1 Ligand-based pharmacophore modelling

A study on the structure activity relationship of natural TRPV3 agonists has been done in 2007. Activation was compared in percent to camphor. Key findings were that a secondary hydroxyl group is a key requirement for activation and that alcohols yield better activation than ketones. Structures with aromatic or non-aromatic rings were found to be more efficient than acyclic ones.[99]

To incorporate those results, we have developed hypothesis 1, where the same compounds with an activity value of percent activation in comparison to camphor were used. However, also EC50 values are available for many natural compounds. Using those as activity value, we developed hypothesis 2, with all compounds where a EC50 value was available.

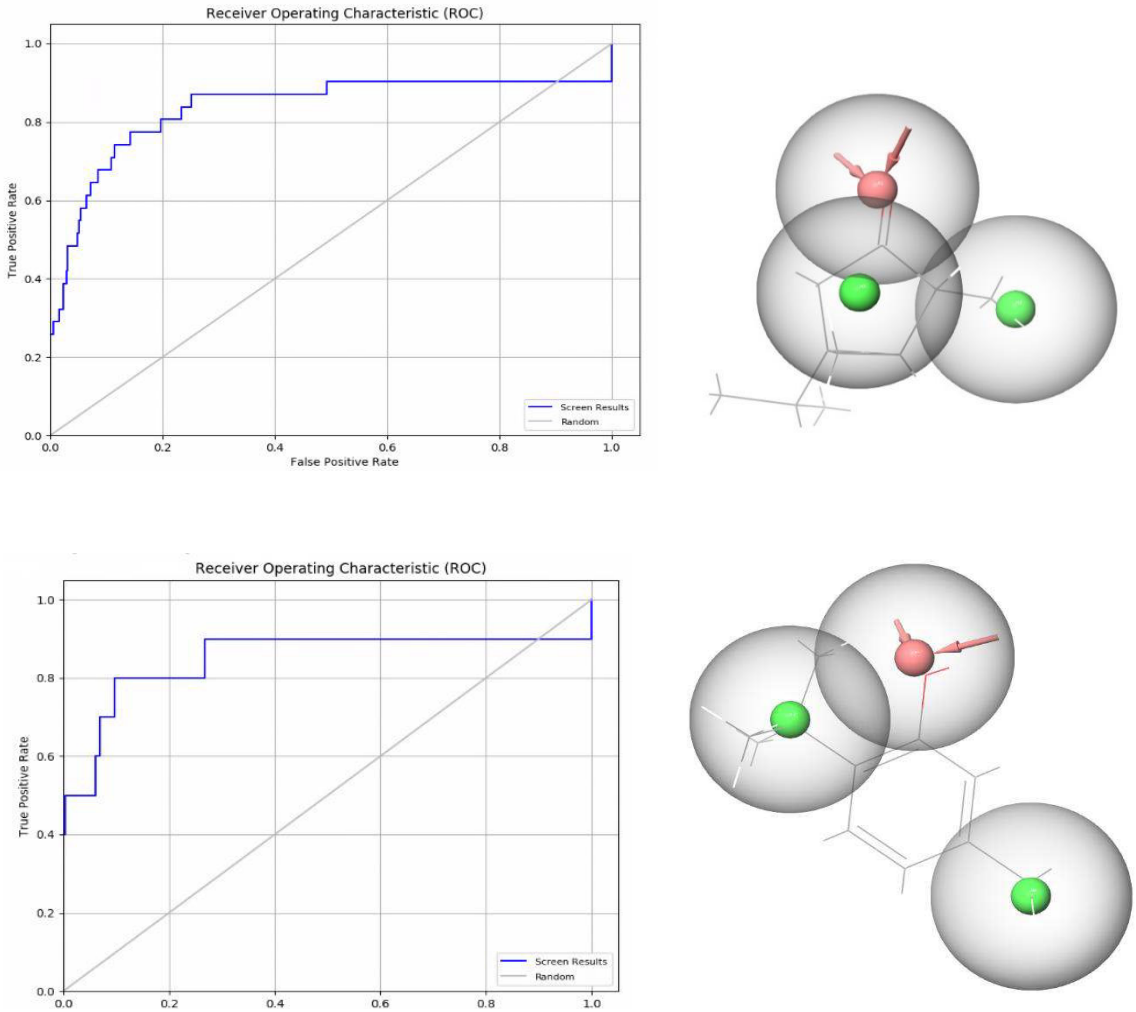


Figure 24 ROC curve and pharmacophore hypothesis for TRPV3 natural agonists. Hypothesis 1 uses activity values based on compounds with reported activity in percent activation in comparison to camphor, as reported by Vogt-Eisele et al.[99]. Hypothesis 2 uses compounds with reported EC50 values as activity value.

Table 9 Scoring values of the two best scoring pharmacophore models for TRPV3.

Hypotheses	Models	Reference molecule	Phase Hypo Score	ROC AUC
Hypothesis 1	AHH_6	Borneol	1.06	0.92
Hypothesis 2	AHH_4	Borneol	0.98	0.91

Hypothesis 1 and 2 yielded similar results. In both cases the top scoring models had one acceptor and two hydrophobic features. Statistics were also similar, although Hypothesis 2 scored slightly higher.

Taken those results together, natural data bases could be screened using the pharmacophore models, with a focus on molecules with low molecular weight, around 100-300 Dalton, featuring a ring structure and an acceptor in the form of an alcohol or ketone, to find novel TRPV3 agonists.

2.2.4 TRPV1 ligands

A total of 34 TRPV1 natural agonists were collected. Most of them origin in spice plants such as ginger, wasabi, garlic, turmeric, anis and different peppers. Their molecular weight ranges from 99 to 368, they feature between 1 to 6 HBAs and 1 to 10 rotatable bonds. Within those boundaries the molecules are evenly distributed (Figure 25), highlighting their variety. Furthermore, 27 known inactives have been collected.

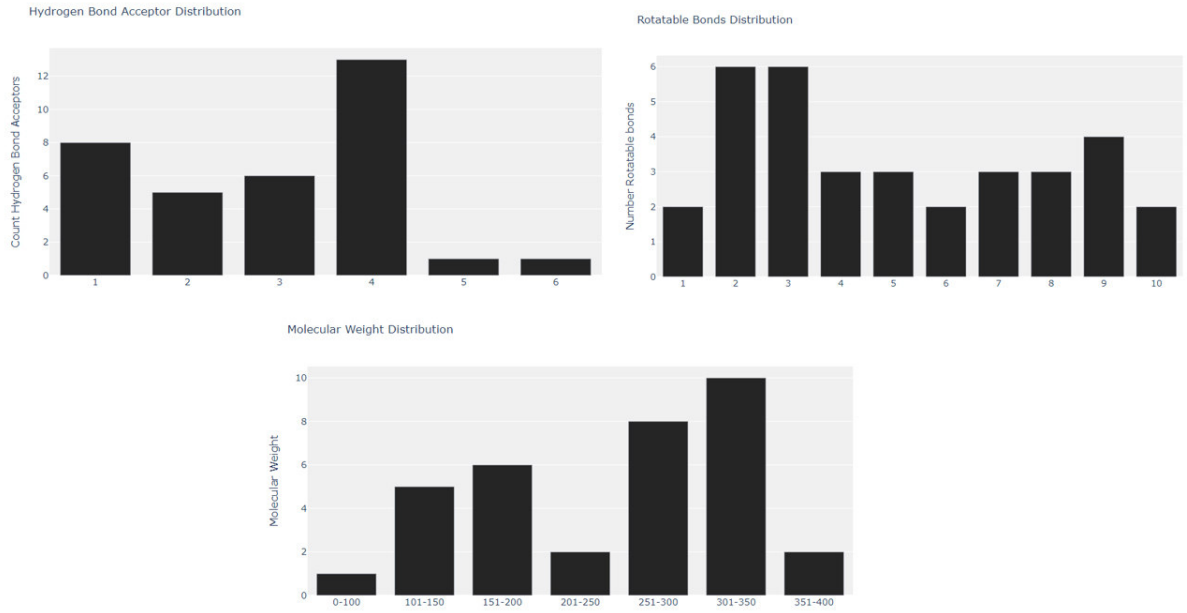


Figure 25 Bar charts highlighting the diversity of TRPV1 ligands, by showing their distribution of HBA acceptors (upper left), rotatable bonds (upper left) and weight (bottom).

2.5.1 Ligand-based pharmacophore modelling

Due to the diverse nature of TRPV1 agonists finding a fitting pharmacophore was troublesome. The best scoring model reached a ROC curve AUC slightly better than a random screening (51%) and featured two acceptor features, with a close by aromatic ring and a hydrophobic feature.

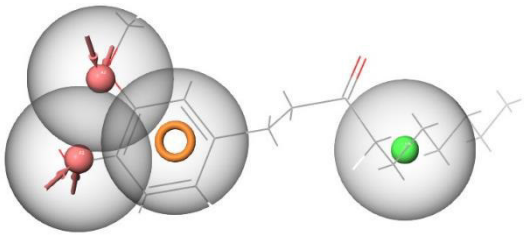


Figure 26 Best Scoring pharmacophore hypothesis for TRPV1.

TRPV1 has multiple possible binding sites, and we do not know the targeted binding site for each ligand.[132] This explains the poor performance of a pharmacophore modelling approach that uses all ligands. To increase it a structural approach could be pursued. Where the individual pockets and their features are investigated, and ligands classified accordingly to which ones they fit best.

3 Methods:

3.1 Chemoinformatics analyses of food-derived bitter compounds

3.1.1 Collection of Bitter Compounds' Set

Based on the work of Ariane Isabell Mayers Thesis "A bitter taste strategy: T2Rs and TRP receptor as molecular probes for the identification of useful nutrients in food plants".[101] The work already contained an extensive set of natural bitter molecules. The set was updated with newly published compounds and with activity data towards the individual TAS2Rs (EC50 and effective concentration values). Chemical identifiers from PubChem were manually added. The final set is available at https://github.com/dipizio/ Natural_TAS2R_agonists. Food-derived bitter compounds were used for the following analyses.

3.1.2 Chemical Classification

ClassyFire, a structure-based chemical taxonomy open source tool, which enables automated, large scale chemical classification,[102] was used to classify and group collected bitter compounds.

3.1.3 Scaffold decomposition

Unique scaffolds were calculated using the Bemis–Murcko Scaffold Decomposition tool available in Maestro (Schrödinger Release 2021-2: Maestro, Schrödinger, LLC, New York, NY, 2021).

3.1.4 Data handling and chemical space analysis in Knime

The gathered compounds and associated information (SMILES, Inchlkey, receptor association) were loaded into Knime 4.3.0.[133] Subsequently, the SMILES format was converted into the string-SMILES format, duplicates were removed, and classification results from ClassyFire were included. Three compounds not recognized by ClassyFire were discarded (sodium benzoate, Tyr-Pro-Phe-Pro-Gly-Pro-Ile-His-Asn-Ser and Leu-Val-Tyr-Pro-Phe-Pro-Gly-Pro-Ile-His-Asn). In the following two metanodes, compounds were split into separate entries for each associated receptor, e.g., coumestrol that targets TAS2R14 and -R39 was split into two entries, one for TASR14 and one for TAS2R39. Next, MACCS fingerprints were calculated using RDKIT MACCS Fingerprint. The outcome was converted into separate entries. A t-SNE was used to investigate the chemical space and was performed using the "t-SNE

(L.Jonsson node)" node using the following settings: Dimensions :2, Iterations: 5000, Theta 0.5, N. of threads:8 Seed: 1615892099957. Afterward, unnecessary columns were removed, and the rest exported into a .csv file. Built Knime workflow is shown in Figure 27.

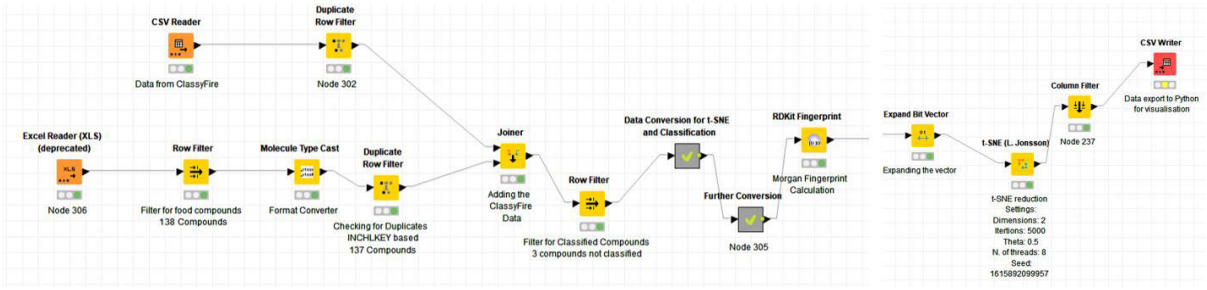


Figure 27 Knime workflow for TAS2R data set handling and t-SNE calculation.

3.1.5 Graphs

Graphs were created in jupyter notebook using Python 3. The following packages were used: numpy, pandas, matplotlib, plotly and seaborn.

The alluvial plot was generated with Plotly. Nodes settings were as followed: thickness of 20, pad to 10, and uniform color of 'grey'. TAS2Rs targeted were set as the primary source (left nodes), ClassyFire superclasses of the compounds as the primary target/secondary source (middle nodes), and the number of receptors targeted per compound was set as secondary target (left nodes). TAS2Rs were sorted by promiscuity indices (TAS2R14, -R1, -R46, -R4, -R39, -R10, -R5, -R7, -R40, -R43, -R16, -R38, -R30, -R8, -R31, -R20, -R50), superclasses by the number of compounds in each group, from largest to smallest, and compound promiscuity in increasing order (compounds that activate only one receptor, and compounds that activate 2, 3, 4 or more than 5 receptors).

3.2 CADD analyses of TRP ligands

3.2.1 Collection of TRP Compounds' Set

The TRP Compounds' Set include ligands from natural source with activity towards the human form of either TRPA1, TRPV1, TRPV3, or TRPM8. The basis for the TRP agonist set was the set published by Boonen et al in 2014.[134] It was extended and brought up to date doing an extensive literature search. The final set is available at https://github.com/dipizio/natural_TRP_modulators. EC50 values were collected when

known. The chemical identifiers, SMILES and InChI Key were derived from PubChem. Presence of functional groups, e.g. isothiocyanates, aromatic rings and ketones, was determined manually.

3.2.2 Model building and figures

All following model building has been done in the Schrödinger Suite version 12.9.123 on Linux-x_64. All associated figures (ROC curves, Pharmacophore models, Ligand/Protein depictions) were generated within Maestro.[135]

3.2.2.1 Ligand preparation

Ligands were imported in Maestro[135] and 3D structures at physiological condition were generated using the LigPrep tool[136] with the following settings: possible states at pH 7 ±2 using Epik,[136] no retainment of special chiral centres, maximum of 32 stereoisomers per ligand.

3.2.2.2 Pharmacophore development

All pharmacophore models presented in this thesis were developed using Phase[115] "Develop Pharmacophore Hypothesis" function. Details for individual created models are reported below:

3.2.2.2.1 Pharmacophore model for TRPA1 isothiocyanates

In the hypothesis settings it was defined that at least 50% of active compounds should match the hypothesis, the number of features was set from 3 to 6, hypothesis difference criterion and the return number of features per hypothesis was not changed. 7 out of 20 ITCs were used as actives for the development. As inactives 5 molecules that have been shown to not bind towards TRPA1 were used (capsaicin, dehydropipernonaline, pipernonaline, retrofractamide C and rutamarin). EC50 was defined as the activity value.

Furthermore, to appropriately consider the ITC group, custom features were developed. This was done via define features. The SMARTS pattern was set to S=C=N. The geometry was in one feature defined as group and in the other as point. The feature was once set as optional and in a second run it was forced to be included at least once in the hypothesis.

After the development of the pharmacophore, it was tested using the hypothesis validation tool. All 20 ITC agonists were used as actives and decoys were generated as described in the decoy selection.

3.2.2.2.3 Pharmacophore model for TRPA1 Non-electrophilic ligands

A sample of 7 out of the 21 agonists were used for pharmacophore development. Reported EC50 values were used as activity values. The final settings were as followed: 75% or more must match the hypothesis, 3-6 features. Hypothesis validation was done using the other 14 agonists as actives and decoy molecules as inactives.

Decoys were selected using DecoyFinder, a python-based tool for finding decoys for a given set of active molecules.[137] The chosen agonists were loaded and decoy molecules were extracted from the bitter database.[24] Decoy molecules with a similar number of rotational bonds, hydrogen bond acceptors, hydrogen bond donors, logP value and molecular weight to the active molecules were generated. Chemical similarity is defined by a MACCS fingerprints-based Tanimoto similarity. A maximum of 4 decoys were generated per active compound. Those were then exported as .sdf file.

3.2.2.2.4 Pharmacophore model for TRPM8 ligands

As the set contains only 7 TRPM8 ligands, all were used for model building. Furthermore, as for most agonists no effective concentration nor EC50 values are reported, no activity value was set. The final settings were as followed: 75% or more must match the hypothesis, 3-6 features.

3.2.2.2.5 Pharmacophore model for TRPV3 ligands

In the hypothesis settings it was defined that at least 75% of compounds should match the hypothesis and the number of features was set from 3 to 6.

For the first development ligands with a known EC50 value were used for the development and the EC50 was set as actives. In the second run ligands used by Vogt-Eisele et al. for their SAR study were used for development and the percent activation in comparison to camphor was used as activity value.[99] Both times all collected TRPV3 inactives were used as inactives.

For the validation all TRPV3 agonists were used as actives and decoys as inactives. Decoy molecules were generated using the procedure described above.[137]

3.2.2.2.6 Pharmacophore model for TRPV1

Ligands with known EC50 values were used for pharmacophore development, together with the collected inactives. Afterwards the other agonists were used for hypothesis validation with selected decoys.

Best results were yielded when 50 % or more should match the hypothesis, and number of features when the number of features was restricted from 3 to 6.

3.2.2.3 QSAR model of TRPA1 electrophilic ligands

The QSAR model was performed via the AutoQSAR tool in Schrödinger.[115,138] All 46 ligands were used. Random training set was set to 70% and 10 models were kept. EC50 value was chosen as the numerical property to fit. The models were selected according to the standard deviation of the regression (S.D.), the R^2 value for the regression, the root mean square error of the test set (RMSE) and the Q^2 for the predicted activities of the test set.

3.2.2.4 Molecular Docking

3.2.2.4.1 Ligand docking for TRPM8

Structural data for the human TRPM8 were retrieved from the Supporting Information of the paper Xu et al.[131] The model generated in this work was build using TRPM8 from Ficedula albicollis (PDB ID: 6BPQ) as template and the program Rosetta was used for modelling. The protein-ligand complex was prepared for docking using the Protein Preparation Wizard.[139] The receptor grid was generated using Glide.[140] The centroid was defined from the ligand. The docking was performed using default parameters and Glide Standard Precision as level of accuracy.

4. Conclusion

The aim of this work was to collect datasets of natural agonists for two classes of chemoreceptors, TAS2Rs and TRP channels, to computationally analyse their structures and build predictive models for their sensory activities.

For the first set, 247 TAS2R agonists where collected. They were classified according to their structure, chemistry, and receptor association. We specifically focused our analyses on food-derived compounds (133 compounds). Promiscuity versus selectivity for receptors as well as ligands was investigated. The position of agonists in chemical space was assessed and set in relation to the chemical classes, the promiscuity and to the associated receptors.

This work provides a picture of the chemical space covered by food-derived TAS2R agonists. Although the dataset is still quite small, our work is the first chemoinformatics protocol tailored on bitter compounds and can be applied to extended datasets when available.

The second data set is focused on ligands of TRP channels with a role on chemesthesis. 133 natural agonists were collected with activity data towards TRPA1, TRPV1, TRPV3 and TRPM8. The subsets were analysed individually, receptor-ligand interactions investigated, and unique models prepared accordingly.

Computational methods were accurately selected according to the different chemical composition of the TRP subsets. Pharmacophore models were generated for TRPV3, TRPM8 and TRPA1 ITC ligands. TRPV1 and TRPA1 non electrophilic compounds extend on wider chemical space. The diversity of TRPA1 electrophilics is foremost dependent on their electrophilic group, as shown by the QSAR analysis. The TRPV1 compounds' diversity is due to multiple different binding pockets.

Overall, we provide a detailed assessment of the current knowledge of natural agonists of chemosensory TRP channels. Furthermore, we build models for the classification, discovery, and design of novel TRPA1, TRPV1, TRPV3 and TRPM8 ligands.

List of abbreviations

ATP	Adenosintriphosphate
CADD	Computer-Aided Drug Discovery
CaM	calmodulin
cDNA	complementary DNA
cryo-EM	cryo-electrone microscopy
CTDH	C-terminal domain helixes
db	database
EC50	half maximal effective concentration
GPCR	g-coupled protein receptor
HBA	hydrogen bond acceptor
HBD	hydrogen bond donor
IP3	inositol 1,4,5-trisphosph
ITC	Isothiocyanate
MHR	mela-statin homology regions
PIP2	phosphatidylinositol 4,5-bisphosphate
pKA	acid dissociation constant
PLCβ2	phospholipase Cβ2
QSAR	quantitative activity relationship
RMSE	root mean square error
ROC	Receiver operating characteristic
SAR	structure activity relationship
S.D.	Standard deviation
TAS2R	bitter taste receptor
TRP	transient receptor potential
t-SNE	t-distributed stochastic neighbour embedding
VSLD	voltage sensor like domain

List of figures

List of tables

Zusammenfassung

Seit Jahrtausenden versuchen Menschen zu verstehen, wie wir unsere Umgebung wahrnehmen. Aristoteles studierte bereits unsere Sinne und teilte sie in fünf Kategorien. Wie genau diese funktionieren entzog sich ihm jedoch. Er konnte nicht wissen was Rezeptoren sind, wie sie aussehen und funktionieren, wie wir das heutzutage tun. Die Reaktion eines Organismus auf eine sensorisch aktive chemische Verbindung (chemosensorische Wahrnehmung), wie der süße Geschmack von Zucker, das Brennen von Chillies oder der tränenreizende Duft von Zwiebeln wird von chemosensorischen Rezeptoren vermittelt. Eine Vielzahl solcher Rezeptoren und tausende Moleküle, welche mit ihnen interagieren sind mittlerweile bekannt. Immer wichtiger, um die komplexen Interaktionen zu verstehen werden computergestützte Methoden und Modelle.

In dieser Arbeit wurden chemoinformatische Methoden verwendet, um chemische Sinne zu untersuchen. Zwei Datensetze an natürlichen Liganden wurden zusammengetragen und analysiert.

Das erste Set besteht aus natürlichen Agonisten für Bitterrezeptoren mit Fokus auf Lebensmittelbestandteile. Bitterstoffe führen häufig zu einer aversiven Reaktion. Trotzdem haben viele einen Nutzen für die Gesundheit und sind wichtig für das Geschmacksprofil von zum Beispiel Kaffee. Wir haben Informationen zu natürlichen Bitterstoffen gesammelt und ihre chemische Struktur im Zusammenhang mit den 25 TAS2R Rezeptoren untersucht.

Das zweite Set besteht aus natürlichen Agonisten für TRP-Kanäle, welche in Zusammenhang mit der chemischen Wahrnehmung stehen (TRPA1, TRPV1, TRPV3 und TRPM8). Die Entdeckung dieser Kanäle wurde 2021 mit dem Nobelpreis ausgezeichnet. Wir haben die molekularen Eigenschaften, welche für die Ligand-Rezeptor Interaktionen wichtig sind, untersucht und anhand dessen in-silico Modelle generiert, welche in der Zukunft helfen können, um neue Liganden zu finden.

Teile dieser Arbeit wurden bereits im Journal of Agricultural and Food Chemistry veröffentlicht.[2]

Bibliography

(1) The Nobel Prize in Physiology or Medicine 2021 https://www.nobelprize.org/prizes/medicine/2021/press-release/ (accessed 2021 -11 -25).

(2) Bayer, S.; Mayer, A. I.; Borgonovo, G.; Morini, G.; Di Pizio, A.; Bassoli, A. Chemoinformatics View on Bitter Taste Receptor Agonists in Food. *J. Agric. Food Chem.* **2021**, *69* (46), 13916–13924. https://doi.org/10.1021/acs.jafc.1c05057.

(3) On the nature of taste qualities. - PsycNET https://doi.apa.org/doiLanding?doi=10.1037%2F0033-2909.86.1.151 (accessed 2022 -01 -17).

(4) Beauchamp, G. K. Basic Taste: A Perceptual Concept. *J. Agric. Food Chem.* **2019**, *67* (50), 13860–13869. https://doi.org/10.1021/acs.jafc.9b03542.

(5) The taste of toxicity: A quantitative analysis of bitter and toxic molecules - Nissim - 2017 - IUBMB Life - Wiley Online Library https://iubmb.onlinelibrary.wiley.com/doi/full/10.1002/iub.1694 (accessed 2022 -01 -12).

(6) Europe Nutraceutical Ingredient Market | Trends, Size, Analysis, Share. *Inkwood Research*.

(7) Doty, R. L. *Handbook of Olfaction and Gustation*; John Wiley & Sons, 2015.

(8) Stone, L. M.; Finger, T. E.; Tam, P. P.; Tan, S. S. Taste Receptor Cells Arise from Local Epithelium, Not Neurogenic Ectoderm. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92* (6), 1916–1920.

(9) Kinnamon, S. C.; Finger, T. E. Recent Advances in Taste Transduction and Signaling. *F1000Research* **2019**, *8*, F1000 Faculty Rev-2117. https://doi.org/10.12688/f1000research.21099.1.

(10) Tomchik, S. M.; Berg, S.; Kim, J. W.; Chaudhari, N.; Roper, S. D. Breadth of Tuning and Taste Coding in Mammalian Taste Buds. *J. Neurosci.* **2007**, *27* (40), 10840–10848. https://doi.org/10.1523/JNEUROSCI.1863-07.2007.

(11) Huang, Y. A.; Maruyama, Y.; Stimac, R.; Roper, S. D. Presynaptic (Type III) Cells in Mouse Taste Buds Sense Sour (Acid) Taste. *J. Physiol.* **2008**, *586* (12), 2903–2912. https://doi.org/10.1113/jphysiol.2008.151233.

(12) Roper, S. D.; Chaudhari, N. Taste Buds: Cells, Signals and Synapses. *Nat. Rev. Neurosci.* **2017**, *18* (8), 485–497. https://doi.org/10.1038/nrn.2017.68.

(13) Owen, P. Chemical Senses Volume 2: Irritation: Edited by B G Green, J R Macon and M R Kare. Pp 361. Marcel Dekker, New York. 1990. $99.55: ISBN 0-8247-8323-9. *Biochem. Educ.* **1990**, *18* (3), 155–155. https://doi.org/10.1016/0307-4412(90)90239-K.

(14) Carstens, E. Overview of Chemesthesis with a Look to the Future. In *Chemesthesis*; John Wiley & Sons, Ltd, 2016; pp 268–285. https://doi.org/10.1002/9781118951620.ch14.

(15) McDonald, S. T.; Bolliet, D. A.; Hayes, J. E. *Chemesthesis: Chemical Touch in Food and Eating*; John Wiley & Sons, 2016.

(16) Ross, R. A. Anandamide and Vanilloid TRPV1 Receptors. *Br. J. Pharmacol.* **2003**, *140* (5), 790–801. https://doi.org/10.1038/sj.bjp.0705467.

(17) Dhaka, A.; Murray, A. N.; Mathur, J.; Earley, T. J.; Petrus, M. J.; Patapoutian, A. TRPM8 Is Required for Cold Sensation in Mice. *Neuron* **2007**, *54* (3), 371–378. https://doi.org/10.1016/j.neuron.2007.02.024.

(18) Roper, S. D. TRPs in Taste and Chemesthesis. *Handb. Exp. Pharmacol.* **2014**, *223*, 827–871. https://doi.org/10.1007/978-3-319-05161-1_5.

(19) Matsunami, H.; Montmayeur, J. P.; Buck, L. B. A Family of Candidate Taste Receptors in Human and Mouse. *Nature* **2000**, *404* (6778), 601–604. https://doi.org/10.1038/35007072.

(20) Go, Y.; Satta, Y.; Takenaka, O.; Takahata, N. Lineage-Specific Loss of Function of Bitter Taste Receptor Genes in Humans and Nonhuman PrimatesSequence Data from This Article Have Been Deposited with the EMBL/GenBank Data Libraries under Accession Nos. AB198983, AB199308. *Genetics* **2005**, *170* (1), 313–326. https://doi.org/10.1534/genetics.104.037523.

(21) Chandrashekar, J.; Mueller, K. L.; Hoon, M. A.; Adler, E.; Feng, L.; Guo, W.; Zuker, C. S.; Ryba, N. J. T2Rs Function as Bitter Taste Receptors. *Cell* **2000**, *100* (6), 703–711. https://doi.org/10.1016/s0092-8674(00)80706-0.

(22) Mueller, K. L.; Hoon, M. A.; Erlenbach, I.; Chandrashekar, J.; Zuker, C. S.; Ryba, N. J. P. The Receptors and Coding Logic for Bitter Taste. *Nature* **2005**, *434* (7030), 225–229. https://doi.org/10.1038/nature03352.

(23) Liszt, K. I.; Ley, J. P.; Lieder, B.; Behrens, M.; Stöger, V.; Reiner, A.; Hochkogler, C. M.; Köck, E.; Marchiori, A.; Hans, J.; Widder, S.; Krammer, G.; Sanger, G. J.; Somoza, M. M.; Meyerhof, W.; Somoza, V. Caffeine Induces Gastric Acid Secretion via Bitter Taste Signaling in Gastric Parietal Cells. *Proc. Natl. Acad. Sci.* **2017**, *114* (30), E6260–E6269. https://doi.org/10.1073/pnas.1703728114.

(24) Bufe, B.; Breslin, P. A. S.; Kuhn, C.; Reed, D. R.; Tharp, C. D.; Slack, J. P.; Kim, U.-K.; Drayna, D.; Meyerhof, W. The Molecular Basis of Individual Differences in Phenylthiocarbamide and Propylthiouracil Bitterness Perception. *Curr. Biol.* **2005**, *15* (4), 322–327. https://doi.org/10.1016/j.cub.2005.01.047.
(25) Behrens, M.; Ziegler, F. Structure-Function Analyses of Human Bitter Taste Receptors—Where Do We Stand? *Molecules* **2020**, *25* (19), 4423. https://doi.org/10.3390/molecules25194423.
(26) Behrens, M.; Meyerhof, W. Bitter Taste Receptor Research Comes of Age: From Characterization to Modulation of TAS2Rs. *Semin. Cell Dev. Biol.* **2013**, *24* (3), 215–221. https://doi.org/10.1016/j.semcdb.2012.08.006.
(27) Di Pizio, A.; Levit, A.; Slutzki, M.; Behrens, M.; Karaman, R.; Niv, M. Y. Comparing Class A GPCRs to Bitter Taste Receptors: Structural Motifs, Ligand Interactions and Agonist-to-Antagonist Ratios. *Methods Cell Biol.* **2016**, *132*, 401–427. https://doi.org/10.1016/bs.mcb.2015.10.005.
(28) Nowak, S.; Di Pizio, A.; Levit, A.; Niv, M. Y.; Meyerhof, W.; Behrens, M. Reengineering the Ligand Sensitivity of the Broadly Tuned Human Bitter Taste Receptor TAS2R14. *Biochim. Biophys. Acta BBA - Gen. Subj.* **2018**, *1862* (10), 2162–2173. https://doi.org/10.1016/j.bbagen.2018.07.009.
(29) Di Pizio, A.; Niv, M. Y. Computational Studies of Smell and Taste Receptors. *Isr. J. Chem.* **2014**, *54* (8–9), 1205–1218. https://doi.org/10.1002/ijch.201400027.
(30) Meyerhof, W.; Batram, C.; Kuhn, C.; Brockhoff, A.; Chudoba, E.; Bufe, B.; Appendino, G.; Behrens, M. The Molecular Receptive Ranges of Human TAS2R Bitter Taste Receptors. *Chem. Senses* **2010**, *35* (2), 157–170. https://doi.org/10.1093/chemse/bjp092.
(31) Kuhn, C.; Bufe, B.; Batram, C.; Meyerhof, W. Oligomerization of TAS2R Bitter Taste Receptors. *Chem. Senses* **2010**, *35* (5), 395–406. https://doi.org/10.1093/chemse/bjq027.
(32) Yu, Y.; Hao, G.; Zhang, Q.; Hua, W.; Wang, M.; Zhou, W.; Zong, S.; Huang, M.; Wen, X. Berberine Induces GLP-1 Secretion through Activation of Bitter Taste Receptor Pathways. *Biochem. Pharmacol.* **2015**, *97* (2), 173–177. https://doi.org/10.1016/j.bcp.2015.07.012.
(33) Kim, K.-S.; Egan, J. M.; Jang, H.-J. Denatonium Induces Secretion of Glucagon-like Peptide-1 through Activation of Bitter Taste Receptor Pathways. *Diabetologia* **2014**, *57* (10), 2117–2125. https://doi.org/10.1007/s00125-014-3326-5.
(34) Avau, B.; Rotondo, A.; Thijs, T.; Andrews, C. N.; Janssen, P.; Tack, J.; Depoortere, I. Targeting Extra-Oral Bitter Taste Receptors Modulates Gastrointestinal Motility with Effects on Satiation. *Sci. Rep.* **2015**, *5* (1), 15985. https://doi.org/10.1038/srep15985.
(35) Xie, C.; Wang, X.; Young, R. L.; Horowitz, M.; Rayner, C. K.; Wu, T. Role of Intestinal Bitter Sensing in Enteroendocrine Hormone Secretion and Metabolic Control. *Front. Endocrinol.* **2018**, *9*, 576. https://doi.org/10.3389/fendo.2018.00576.
(36) Lu, P.; Zhang, C.-H.; Lifshitz, L. M.; ZhuGe, R. Extraoral Bitter Taste Receptors in Health and Disease. *J. Gen. Physiol.* **2017**, *149* (2), 181–197. https://doi.org/10.1085/jgp.201611637.
(37) Lee, R. J.; Kofonow, J. M.; Rosen, P. L.; Siebert, A. P.; Chen, B.; Doghramji, L.; Xiong, G.; Adappa, N. D.; Palmer, J. N.; Kennedy, D. W.; Kreindler, J. L.; Margolskee, R. F.; Cohen, N. A. Bitter and Sweet Taste Receptors Regulate Human Upper Respiratory Innate Immunity. *J. Clin. Invest.* **2014**, *124* (3), 1393–1405. https://doi.org/10.1172/JCI72094.
(38) Lee, R. J.; Cohen, N. A. Taste Receptors in Innate Immunity. *Cell. Mol. Life Sci. CMLS* **2015**, *72* (2), 217–236. https://doi.org/10.1007/s00018-014-1736-7.
(39) Liggett, S. B. Bitter Taste Receptors in The Wrong Place: Novel Airway Smooth Muscle Targets For Treating Asthma. *Trans. Am. Clin. Climatol. Assoc.* **2014**, *125*, 64–75.
(40) KinDex Pharmaceuticals, Inc. *Effects of KDT501 on Metabolic Features in Insulin Resistant Subjects*; Clinical trial registration NCT02444910; clinicaltrials.gov, 2017.
(41) Di Pizio, A.; Behrens, M.; Krautwurst, D. Beyond the Flavour: The Potential Druggability of Chemosensory G Protein-Coupled Receptors. *Int. J. Mol. Sci.* **2019**, *20* (6), E1402. https://doi.org/10.3390/ijms20061402.
(42) Wiener, A.; Shudler, M.; Levit, A.; Niv, M. Y. BitterDB: A Database of Bitter Compounds. *Nucleic Acids Res.* **2012**, *40* (D1), D413–D419. https://doi.org/10.1093/nar/gkr755.
(43) Tominaga, M.; Caterina, M. J.; Malmberg, A. B.; Rosen, T. A.; Gilbert, H.; Skinner, K.; Raumann, B. E.; Basbaum, A. I.; Julius, D. The Cloned Capsaicin Receptor Integrates Multiple Pain-Producing Stimuli. *Neuron* **1998**, *21* (3), 531–543. https://doi.org/10.1016/s0896-6273(00)80564-4.
(44) Brauchi, S.; Orta, G.; Salazar, M.; Rosenmann, E.; Latorre, R. A Hot-Sensing Cold Receptor: C-Terminal Domain Determines Thermosensation in Transient Receptor Potential Channels. *J. Neurosci.* **2006**, *26* (18), 4835–4840. https://doi.org/10.1523/JNEUROSCI.5080-05.2006.

(45) Eccles, R. Nasal Physiology and Disease with Reference to Asthma. *Agents Actions. Suppl.* **1989**, *28*, 249–261.
(46) Vazquez, G.; Wedel, B. J.; Aziz, O.; Trebak, M.; Putney, J. W. The Mammalian TRPC Cation Channels. *Biochim. Biophys. Acta* **2004**, *1742* (1–3), 21–36. https://doi.org/10.1016/j.bbamcr.2004.08.015.
(47) Froghi, S.; Grant, C. R.; Tandon, R.; Quaglia, A.; Davidson, B.; Fuller, B. New Insights on the Role of TRP Channels in Calcium Signalling and Immunomodulation: Review of Pathways and Implications for Clinical Practice. *Clin. Rev. Allergy Immunol.* **2021**, *60* (2), 271–292. https://doi.org/10.1007/s12016-020-08824-3.
(48) Venkatachalam, K.; Montell, C. TRP Channels. *Annu. Rev. Biochem.* **2007**, *76* (1), 387–417. https://doi.org/10.1146/annurev.biochem.75.103004.142819.
(49) Cosens, D. J.; Manning, A. Abnormal Electroretinogram from a Drosophila Mutant. *Nature* **1969**, *224* (5216), 285–287. https://doi.org/10.1038/224285a0.
(50) Szallasi, A.; McAlexander, M. A. *TRP Channels as Therapeutic Targets: From Basic Science to Clinical Use*; 2015.
(51) García-Añoveros, J.; Wiwatpanit, T. TRPML2 and Mucolipin Evolution. In *Mammalian Transient Receptor Potential (TRP) Cation Channels: Volume I*; Nilius, B., Flockerzi, V., Eds.; Handbook of Experimental Pharmacology; Springer: Berlin, Heidelberg, 2014; pp 647–658. https://doi.org/10.1007/978-3-642-54215-2_25.
(52) Moran, M. M.; McAlexander, M. A.; Bíró, T.; Szallasi, A. Transient Receptor Potential Channels as Therapeutic Targets. *Nat. Rev. Drug Discov.* **2011**, *10* (8), 601–620. https://doi.org/10.1038/nrd3456.
(53) Koivisto, A.-P.; Belvisi, M. G.; Gaudet, R.; Szallasi, A. Advances in TRP Channel Drug Discovery: From Target Validation to Clinical Studies. *Nat. Rev. Drug Discov.* **2021**, 1–19. https://doi.org/10.1038/s41573-021-00268-4.
(54) Structural mechanisms of transient receptor potential ion channels | Journal of General Physiology | Rockefeller University Press https://rupress.org/jgp/article/152/3/e201811998/133640/Structural-mechanisms-of-transient-receptor (accessed 2021 -11 -25).
(55) Liao, M.; Cao, E.; Julius, D.; Cheng, Y. Structure of the TRPV1 Ion Channel Determined by Electron Cryo-Microscopy. *Nature* **2013**, *504* (7478), 107–112. https://doi.org/10.1038/nature12822.
(56) Poletini, M. O.; Moraes, M. N.; Ramos, B. C.; Jerônimo, R.; Castrucci, A. M. de L. TRP Channels: A Missing Bond in the Entrainment Mechanism of Peripheral Clocks throughout Evolution. *Temperature* **2015**, *2* (4), 522–534. https://doi.org/10.1080/23328940.2015.1115803.
(57) Clapham, D. E. Calcium Signaling. *Cell* **1995**, *80* (2), 259–268. https://doi.org/10.1016/0092-8674(95)90408-5.
(58) TRP channels interaction with lipids and its implications in disease - ScienceDirect https://www.sciencedirect.com/science/article/pii/S0005273615001017 (accessed 2021 -11 -25).
(59) Adhikari, J.; Chambers, E.; Koppel, K. Impact of Consumption Temperature on Sensory Properties of Hot Brewed Coffee. *Food Res. Int.* **2019**, *115*, 95–104. https://doi.org/10.1016/j.foodres.2018.08.014.
(60) Talavera, K.; Yasumatsu, K.; Voets, T.; Droogmans, G.; Shigemura, N.; Ninomiya, Y.; Margolskee, R. F.; Nilius, B. Heat Activation of TRPM5 Underlies Thermal Sensitivity of Sweet Taste. *Nature* **2005**, *438* (7070), 1022–1025. https://doi.org/10.1038/nature04248.
(61) Transient Receptor Potential (TRP) Channels and Taste Sensation - Y. Ishimaru, H. Matsunami, 2009 https://journals.sagepub.com/doi/10.1177/0022034508330212 (accessed 2021 -11 -26).
(62) Abe, J.; Hosokawa, H.; Okazawa, M.; Kandachi, M.; Sawada, Y.; Yamanaka, K.; Matsumura, K.; Kobayashi, S. TRPM8 Protein Localization in Trigeminal Ganglion and Taste Papillae. *Mol. Brain Res.* **2005**, *136* (1), 91–98. https://doi.org/10.1016/j.molbrainres.2005.01.013.
(63) Increased expression of TRPV1 in squamous cell carcinoma of the human tongue - Marincsák - 2009 - Oral Diseases - Wiley Online Library https://onlinelibrary.wiley.com/doi/10.1111/j.1601-0825.2009.01526.x (accessed 2021 -11 -26).
(64) Wang, B.; Danjo, A.; Kajiya, H.; Okabe, K.; Kido, M. A. Oral Epithelial Cells Are Activated via TRP Channels. *J. Dent. Res.* **2011**, *90* (2), 163–167. https://doi.org/10.1177/0022034510385459.
(65) Mickle, A. D.; Shepherd, A. J.; Mohapatra, D. P. Chapter Four - Sensory TRP Channels: The Key Transducers of Nociception and Pain. In *Progress in Molecular Biology and Translational*

Science; Price, T. J., Dussor, G., Eds.; Molecular and Cell Biology of Pain; Academic Press, 2015; Vol. 131, pp 73–118. https://doi.org/10.1016/bs.pmbts.2015.01.002.
(66) Aroke, E. N.; Powell-Roach, K. L.; Jaime-Lara, R. B.; Tesfaye, M.; Roy, A.; Jackson, P.; Joseph, P. V. Taste the Pain: The Role of TRP Channels in Pain and Taste Perception. *Int. J. Mol. Sci.* **2020**, *21* (16), 5929. https://doi.org/10.3390/ijms21165929.
(67) Paulsen, C. E.; Armache, J.-P.; Gao, Y.; Cheng, Y.; Julius, D. Structure of the TRPA1 Ion Channel Suggests Regulatory Mechanisms. *Nature* **2015**, *520* (7548), 511–517. https://doi.org/10.1038/nature14367.
(68) Suo, Y.; Wang, Z.; Zubcevic, L.; Hsu, A. L.; He, Q.; Borgnia, M. J.; Ji, R.-R.; Lee, S.-Y. Structural Insights into Electrophile Irritant Sensing by the Human TRPA1 Channel. *Neuron* **2020**, *105* (5), 882-894.e5. https://doi.org/10.1016/j.neuron.2019.11.023.
(69) Macpherson, L. J.; Dubin, A. E.; Evans, M. J.; Marr, F.; Schultz, P. G.; Cravatt, B. F.; Patapoutian, A. Noxious Compounds Activate TRPA1 Ion Channels through Covalent Modification of Cysteines. *Nature* **2007**, *445* (7127), 541–545. https://doi.org/10.1038/nature05544.
(70) Hinman, A.; Chuang, H.-H.; Bautista, D. M.; Julius, D. TRP Channel Activation by Reversible Covalent Modification. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (51), 19564–19568. https://doi.org/10.1073/pnas.0609598103.
(71) Eigenbrod, O.; Debus, K. Y.; Reznick, J.; Bennett, N. C.; Sánchez-Carranza, O.; Omerbašić, D.; Hart, D. W.; Barker, A. J.; Zhong, W.; Lutermann, H.; Katandukila, J. V.; Mgode, G.; Park, T. J.; Lewin, G. R. Rapid Molecular Evolution of Pain Insensitivity in Multiple African Rodents. *Science* **2019**, *364* (6443), 852–859. https://doi.org/10.1126/science.aau0236.
(72) Li, L.; Lee, W.; Lee, W. J.; Auh, J. H.; Kim, S. S.; Yoon, J. Extraction of Allyl Isothiocyanate from Wasabi (Wasabia Japonica Matsum) Using Supercritical Carbon Dioxide. *Food Sci. Biotechnol.* **2010**, *19* (2), 405–410. https://doi.org/10.1007/s10068-010-0057-3.
(73) Takaya, J.; Mio, K.; Shiraishi, T.; Kurokawa, T.; Otsuka, S.; Mori, Y.; Uesugi, M. A Potent and Site-Selective Agonist of TRPA1. *J. Am. Chem. Soc.* **2015**, *137* (50), 15859–15864. https://doi.org/10.1021/jacs.5b10162.
(74) Wrigley, P. J.; Jeong, H.-J.; Vaughan, C. W. Primary Afferents with TRPM8 and TRPA1 Profiles Target Distinct Subpopulations of Rat Superficial Dorsal Horn Neurones. *Br. J. Pharmacol.* **2009**, *157* (3), 371–380. https://doi.org/10.1111/j.1476-5381.2009.00167.x.
(75) Karashima, Y.; Damann, N.; Prenen, J.; Talavera, K.; Segal, A.; Voets, T.; Nilius, B. Bimodal Action of Menthol on the Transient Receptor Potential Channel TRPA1. *J. Neurosci.* **2007**, *27* (37), 9874–9884. https://doi.org/10.1523/JNEUROSCI.2221-07.2007.
(76) Araki, M.; Kanda, N.; Iwata, H.; Sagae, Y.; Masuda, K.; Okuno, Y. Identification of a New Class of Non-Electrophilic TRPA1 Agonists by a Structure-Based Virtual Screening Approach. *Bioorg. Med. Chem. Lett.* **2020**, *30* (11), 127142. https://doi.org/10.1016/j.bmcl.2020.127142.
(77) Iwata, H.; Kanda, N.; Araki, M.; Sagae, Y.; Masuda, K.; Okuno, Y. Discovery of Natural TRPA1 Activators through Pharmacophore-Based Virtual Screening and a Biological Assay. *Bioorg. Med. Chem. Lett.* **2021**, *31*, 127639. https://doi.org/10.1016/j.bmcl.2020.127639.
(78) Ton, H. T.; Phan, T. X.; Abramyan, A. M.; Shi, L.; Ahern, G. P. Identification of a Putative Binding Site Critical for General Anesthetic Activation of TRPA1. *Proc. Natl. Acad. Sci.* **2017**, *114* (14), 3762–3767. https://doi.org/10.1073/pnas.1618144114.
(79) Structural basis of cooling agent and lipid sensing by the cold-activated TRPM8 channel https://www-science-org.uaccess.univie.ac.at/doi/epdf/10.1126/science.aav9334 (accessed 2021 -12 -01). https://doi.org/10.1126/science.aav9334.
(80) Bautista, D. M.; Siemens, J.; Glazer, J. M.; Tsuruda, P. R.; Basbaum, A. I.; Stucky, C. L.; Jordt, S.-E.; Julius, D. The Menthol Receptor TRPM8 Is the Principal Detector of Environmental Cold. *Nature* **2007**, *448* (7150), 204–208. https://doi.org/10.1038/nature05910.
(81) McKemy, D. D.; Neuhausser, W. M.; Julius, D. Identification of a Cold Receptor Reveals a General Role for TRP Channels in Thermosensation. *Nature* **2002**, *416* (6876), 52–58. https://doi.org/10.1038/nature719.
(82) Zhang, L.; Barritt, G. J. Evidence That TRPM8 Is an Androgen-Dependent Ca2+ Channel Required for the Survival of Prostate Cancer Cells. *Cancer Res.* **2004**, *64* (22), 8365–8373. https://doi.org/10.1158/0008-5472.CAN-04-2146.
(83) Sabnis, A. S.; Shadid, M.; Yost, G. S.; Reilly, C. A. Human Lung Epithelial Cells Express a Functional Cold-Sensing TRPM8 Variant. *Am. J. Respir. Cell Mol. Biol.* **2008**, *39* (4), 466–474. https://doi.org/10.1165/rcmb.2007-0440OC.
(84) Stein, R. J.; Santos, S.; Nagatomi, J.; Hayashi, Y.; Minnery, B. S.; Xavier, M.; Patel, A. S.; Nelson, J. B.; Futrell, W. J.; Yoshimura, N.; Chancellor, M. B.; De, M. F. Cool (Trpm8) and Hot

(Trpv1) Receptors in the Bladder and Male Genital Tract. *J. Urol.* **2004**, *172* (3), 1175–1178. https://doi.org/10.1097/01.ju.0000134880.55119.cf.
(85) Yin, Y.; Wu, M.; Zubcevic, L.; Borschel, W. F.; Lander, G. C.; Lee, S.-Y. Structure of the Cold- and Menthol-Sensing Ion Channel TRPM8. *Science* **2018**, *359* (6372), 237–241. https://doi.org/10.1126/science.aan4325.
(86) Structure of the cold- and menthol-sensing ion channel TRPM8 https://www-science-org.uaccess.univie.ac.at/doi/10.1126/science.aan4325 (accessed 2021 -11 -29).
(87) Szallasi, A.; Sheta, M. Targeting TRPV1 for Pain Relief: Limits, Losers and Laurels. *Expert Opin. Investig. Drugs* **2012**, *21* (9), 1351–1369. https://doi.org/10.1517/13543784.2012.704021.
(88) Transient receptor potential cation channel subfamily V member 1 | DrugBank Online https://go.drugbank.com/bio_entities/BE0001023 (accessed 2021 -11 -30).
(89) Knotkova, H.; Pappagallo, M.; Szallasi, A. Capsaicin (TRPV1 Agonist) Therapy for Pain Relief: Farewell or Revival? *Clin. J. Pain* **2008**, *24* (2), 142–154. https://doi.org/10.1097/AJP.0b013e318158ed9e.
(90) Jeske, N. A.; Patwardhan, A. M.; Ruparel, N. B.; Akopian, A. N.; Shapiro, M. S.; Henry, M. A. A-Kinase Anchoring Protein 150 Controls Protein Kinase C-Mediated Phosphorylation and Sensitization of TRPV1. *PAIN* **2009**, *146* (3), 301–307. https://doi.org/10.1016/j.pain.2009.08.002.
(91) Chuang, H. H.; Prescott, E. D.; Kong, H.; Shields, S.; Jordt, S. E.; Basbaum, A. I.; Chao, M. V.; Julius, D. Bradykinin and Nerve Growth Factor Release the Capsaicin Receptor from PtdIns(4,5)P2-Mediated Inhibition. *Nature* **2001**, *411* (6840), 957–962. https://doi.org/10.1038/35082088.
(92) Lishko, P. V.; Procko, E.; Jin, X.; Phelps, C. B.; Gaudet, R. The Ankyrin Repeats of TRPV1 Bind Multiple Ligands and Modulate Channel Sensitivity. *Neuron* **2007**, *54* (6), 905–918. https://doi.org/10.1016/j.neuron.2007.05.027.
(93) Shuba, Y. M. Beyond Neuronal Heat Sensing: Diversity of TRPV1 Heat-Capsaicin Receptor-Channel Functions. *Front. Cell. Neurosci.* **2021**, *14*.
(94) Nakashimo, Y.; Takumida, M.; Fukuiri, T.; Anniko, M.; Hirakawa, K. Expression of Transient Receptor Potential Channel Vanilloid (TRPV) 1–4, Melastin (TRPM) 5 and 8, and Ankyrin (TRPA1) in the Normal and Methimazole-Treated Mouse Olfactory Epithelium. *Acta Otolaryngol. (Stockh.)* **2010**, *130* (11), 1278–1286. https://doi.org/10.3109/00016489.2010.489573.
(95) Wang, H.; Yang, P.; Lu, Y.; Wang, J.; Jeon, J.; Wang, Q.; Tian, J.-B.; Zang, B.; Yu, Y.; Zhu, M. X. Mechanisms of Proton Inhibition and Sensitization of the Cation Channel TRPV3. *J. Gen. Physiol.* **2020**, *153* (2), e202012663. https://doi.org/10.1085/jgp.202012663.
(96) Liu, B.; Yao, J.; Zhu, M. X.; Qin, F. Hysteresis of Gating Underlines Sensitization of TRPV3 Channels. *J. Gen. Physiol.* **2011**, *138* (5), 509–520. https://doi.org/10.1085/jgp.201110689.
(97) Zubcevic, L.; Herzik, M. A.; Wu, M.; Borschel, W. F.; Hirschi, M.; Song, A. S.; Lander, G. C.; Lee, S.-Y. Conformational Ensemble of the Human TRPV3 Ion Channel. *Nat. Commun.* **2018**, *9* (1), 4773. https://doi.org/10.1038/s41467-018-07117-w.
(98) Nilius, B.; Bíró, T.; Owsianik, G. TRPV3: Time to Decipher a Poorly Understood Family Member! *J. Physiol.* **2014**, *592* (2), 295–304. https://doi.org/10.1113/jphysiol.2013.255968.
(99) Vogt-Eisele, A. K.; Weber, K.; Sherkheli, M. A.; Vielhaber, G.; Panten, J.; Gisselmann, G.; Hatt, H. Monoterpenoid Agonists of TRPV3. *Br. J. Pharmacol.* **2007**, *151* (4), 530–540. https://doi.org/10.1038/sj.bjp.0707245.
(100) Sherkheli, M. A.; Gisselmann, G.; Hatt, H. Supercooling Agent Icilin Blocks a Warmth-Sensing Ion Channel TRPV3. *Sci. World J.* **2012**, *2012*, e982725. https://doi.org/10.1100/2012/982725.
(101) Mayer, A. I. A Bitter Taste Strategy: T2Rs and TRP Receptors as Molecular Probes for the Identification of Useful Nutrients in Food Plants., Universität Hohenheim, 2019.
(102) Djoumbou Feunang, Y.; Eisner, R.; Knox, C.; Chepelev, L.; Hastings, J.; Owen, G.; Fahy, E.; Steinbeck, C.; Subramanian, S.; Bolton, E.; Greiner, R.; Wishart, D. S. ClassyFire: Automated Chemical Classification with a Comprehensive, Computable Taxonomy. *J. Cheminformatics* **2016**, *8* (1), 61. https://doi.org/10.1186/s13321-016-0174-y.
(103) Di Pizio, A.; Niv, M. Y. Promiscuity and Selectivity of Bitter Molecules and Their Receptors. *Bioorg. Med. Chem.* **2015**, *23* (14), 4082–4091. https://doi.org/10.1016/j.bmc.2015.04.025.
(104) Borgonovo, G.; Zimbaldi, N.; Guarise, M.; Bedussi, F.; Winnig, M.; Vennegeerts, T.; Bassoli, A. Glucosinolates in Sisymbrium Officinale (L.) Scop.: Comparative Analysis in Cultivated and Wild Plants and in Vitro Assays with T2Rs Bitter Taste Receptors. *Molecules* **2019**, *24* (24), 4572. https://doi.org/10.3390/molecules24244572.

(105) Bufe, B.; Hofmann, T.; Krautwurst, D.; Raguse, J.-D.; Meyerhof, W. The Human TAS2R16 Receptor Mediates Bitter Taste in Response to Beta-Glucopyranosides. *Nat. Genet.* **2002**, *32* (3), 397–401. https://doi.org/10.1038/ng1014.
(106) Fierro, F.; Giorgetti, A.; Carloni, P.; Meyerhof, W.; Alfonso-Prieto, M. Dual Binding Mode of "Bitter Sugars" to Their Human Bitter Taste Receptor Target. *Sci. Rep.* **2019**, *9* (1), 8437. https://doi.org/10.1038/s41598-019-44805-z.
(107) Cardona, F.; Andrés-Lacueva, C.; Tulipani, S.; Tinahones, F. J.; Queipo-Ortuño, M. I. Benefits of Polyphenols on Gut Microbiota and Implications in Human Health. *J. Nutr. Biochem.* **2013**, *24* (8), 1415–1422. https://doi.org/10.1016/j.jnutbio.2013.05.001.
(108) Ganesan, K.; Xu, B. A Critical Review on Polyphenols and Health Benefits of Black Soybeans. *Nutrients* **2017**, *9* (5), 455. https://doi.org/10.3390/nu9050455.
(109) Wang, Q.; Liszt, K. I.; Deloose, E.; Canovai, E.; Thijs, T.; Farré, R.; Ceulemans, L. J.; Lannoo, M.; Tack, J.; Depoortere, I. Obesity Alters Adrenergic and Chemosensory Signaling Pathways That Regulate Ghrelin Secretion in the Human Gut. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **2019**, *33* (4), 4907–4920. https://doi.org/10.1096/fj.201801661RR.
(110) Roland, W. S. U.; van Buren, L.; Gruppen, H.; Driesse, M.; Gouka, R. J.; Smit, G.; Vincken, J.-P. Bitter Taste Receptor Activation by Flavonoids and Isoflavonoids: Modeled Structural Requirements for Activation of HTAS2R14 and HTAS2R39. *J. Agric. Food Chem.* **2013**, *61* (44), 10454–10466. https://doi.org/10.1021/jf403387p.
(111) Brockhoff, A.; Behrens, M.; Massarotti, A.; Appendino, G.; Meyerhof, W. Broad Tuning of the Human Bitter Taste Receptor HTAS2R46 to Various Sesquiterpene Lactones, Clerodane and Labdane Diterpenoids, Strychnine, and Denatonium. *J. Agric. Food Chem.* **2007**, *55* (15), 6236–6243. https://doi.org/10.1021/jf070503p.
(112) Wermuth, C. G.; Ganellin, C. R.; Lindberg, P.; Mitscher, L. A. Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure Appl. Chem.* **1998**, *70* (5), 1129–1143. https://doi.org/10.1351/pac199870051129.
(113) Shintaku, K.; Uchida, K.; Suzuki, Y.; Zhou, Y.; Fushiki, T.; Watanabe, T.; Yazawa, S.; Tominaga, M. Activation of Transient Receptor Potential A1 by a Non-Pungent Capsaicin-like Compound, Capsiate. *Br. J. Pharmacol.* **2012**, *165* (5), 1476–1486. https://doi.org/10.1111/j.1476-5381.2011.01634.x.
(114) Terada, Y.; Masuda, H.; Watanabe, T. Structure–Activity Relationship Study on Isothiocyanates: Comparison of TRPA1-Activating Ability between Allyl Isothiocyanate and Specific Flavor Components of Wasabi, Horseradish, and White Mustard. *J. Nat. Prod.* **2015**, *78* (8), 1937–1941. https://doi.org/10.1021/acs.jnatprod.5b00272.
(115) Dixon, S. L.; Smondyrev, A. M.; Knoll, E. H.; Rao, S. N.; Shaw, D. E.; Friesner, R. A. PHASE: A New Engine for Pharmacophore Perception, 3D QSAR Model Development, and 3D Database Screening: 1. Methodology and Preliminary Results. *J. Comput. Aided Mol. Des.* **2006**, *20* (10–11), 647–671. https://doi.org/10.1007/s10822-006-9087-6.
(116) Startek, J. B.; Talavera, K. Lipid Raft Destabilization Impairs Mouse TRPA1 Responses to Cold and Bacterial Lipopolysaccharides. *Int. J. Mol. Sci.* **2020**, *21* (11), 3826. https://doi.org/10.3390/ijms21113826.
(117) Moparthi, L.; Kichko, T. I.; Eberhardt, M.; Högestätt, E. D.; Kjellbom, P.; Johanson, U.; Reeh, P. W.; Leffler, A.; Filipovic, M. R.; Zygmunt, P. M. Human TRPA1 Is a Heat Sensor Displaying Intrinsic U-Shaped Thermosensitivity. *Sci. Rep.* **2016**, *6* (1), 28763. https://doi.org/10.1038/srep28763.
(118) Journigan, V. B.; Zaveri, N. T. TRPM8 Ion Channel Ligands for New Therapeutic Applications and as Probes to Study Menthol Pharmacology. *Life Sci.* **2013**, *92* (8), 425–437. https://doi.org/10.1016/j.lfs.2012.10.032.
(119) Caceres, A. I.; Liu, B.; Jabba, S. V.; Achanta, S.; Morris, J. B.; Jordt, S.-E. Transient Receptor Potential Cation Channel Subfamily M Member 8 Channels Mediate the Anti-Inflammatory Effects of Eucalyptol. *Br. J. Pharmacol.* **2017**, *174* (9), 867–879. https://doi.org/10.1111/bph.13760.
(120) Borneol Is a TRPM8 Agonist that Increases Ocular Surface Wetness https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0158868 (accessed 2021 -12 -12).
(121) McKemy, D. D. *TRPM8: The Cold and Menthol Receptor*; CRC Press/Taylor & Francis, 2007.
(122) Shirai, T.; Kumihashi, K.; Sakasai, M.; Kusuoku, H.; Shibuya, Y.; Ohuchi, A. Identification of a Novel TRPM8 Agonist from Nutmeg: A Promising Cooling Compound. *ACS Med. Chem. Lett.* **2017**, *8* (7), 715–719. https://doi.org/10.1021/acsmedchemlett.7b00104.

(123) TRPM8 Channel Activation Induced by Monoterpenoid Rotundifolone Underlies Mesenteric Artery Relaxation https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0143171 (accessed 2021 -12 -12).
(124) Wang, W.; Wang, H.; Zhao, Z.; Huang, X.; Xiong, H.; Mei, Z. Thymol Activates TRPM8-Mediated Ca2+ Influx for Its Antipruritic Effects and Alleviates Inflammatory Response in Imiquimod-Induced Mice. *Toxicol. Appl. Pharmacol.* **2020**, *407*, 115247. https://doi.org/10.1016/j.taap.2020.115247.
(125) Zhong, J.; Minassi, A.; Prenen, J.; Taglialatela-Scafati, O.; Appendino, G.; Nilius, B. Umbellulone Modulates TRP Channels. *Pflugers Arch.* **2011**, *462* (6), 861–870. https://doi.org/10.1007/s00424-011-1043-1.
(126) Heck, J. D. A Review and Assessment of Menthol Employed as a Cigarette Flavoring Ingredient. *Food Chem. Toxicol.* **2010**, *48*, S1–S38. https://doi.org/10.1016/j.fct.2009.11.002.
(127) Menthol: A simple monoterpene with remarkable biological properties - ScienceDirect https://www.sciencedirect.com/science/article/abs/pii/S0031942213003063?via%3Dihub (accessed 2021 -12 -12).
(128) Juergens, U. R.; Dethlefsen, U.; Steinkamp, G.; Gillissen, A.; Repges, R.; Vetter, H. Anti-Inflammatory Activity of 1.8-Cineol (Eucalyptol) in Bronchial Asthma: A Double-Blind Placebo-Controlled Trial. *Respir. Med.* **2003**, *97* (3), 250–256. https://doi.org/10.1053/rmed.2003.1432.
(129) Agamennone, M.; Nicoli, A.; Bayer, S.; Weber, V.; Borro, L.; Gupta, S.; Fantacuzzi, M.; Di Pizio, A. Chapter 13 - Protein-Protein Interactions at a Glance: Protocols for the Visualization of Biomolecular Interactions. In *Methods in Cell Biology*; Shukla, A. K., Ed.; Biomolecular Interactions Part A; Academic Press, 2021; Vol. 166, pp 271–307. https://doi.org/10.1016/bs.mcb.2021.06.012.
(130) Chothia, C.; Lesk, A. M. The Relation between the Divergence of Sequence and Structure in Proteins. *EMBO J.* **1986**, *5* (4), 823–826.
(131) Xu, L.; Han, Y.; Chen, X.; Aierken, A.; Wen, H.; Zheng, W.; Wang, H.; Lu, X.; Zhao, Z.; Ma, C.; Liang, P.; Yang, W.; Yang, S.; Yang, F. Molecular Mechanisms Underlying Menthol Binding and Activation of TRPM8 Ion Channel. *Nat. Commun.* **2020**, *11* (1), 3790. https://doi.org/10.1038/s41467-020-17582-x.
(132) Abbas, M. A. Modulation of TRPV1 Channel Function by Natural Products in the Treatment of Pain. *Chem. Biol. Interact.* **2020**, *330*, 109178. https://doi.org/10.1016/j.cbi.2020.109178.
(133) Sieb, C.; Meinl, T.; Berthold, M. R. Parallel and Distributed Data Pipelining with KNIME. *Mediterr. J. Comput. Netw.* **2007**, *3* (2), 43–51.
(134) Boonen, B.; Startek, J. B.; Talavera, K. Chemical Activation of Sensory TRP Channels. In *Taste and Smell*; Krautwurst, D., Ed.; Topics in Medicinal Chemistry; Springer International Publishing: Cham, 2017; pp 73–113. https://doi.org/10.1007/7355_2015_98.
(135) *Maestro*; Schrödinger LLC.
(136) Shelley, J. C.; Cholleti, A.; Frye, L. L.; Greenwood, J. R.; Timlin, M. R.; Uchimaya, M. Epik: A Software Program for PKaprediction and Protonation State Generation for Drug-like Molecules. *J. Comput. Aided Mol. Des.* **2007**, *21* (12), 681–691. https://doi.org/10.1007/s10822-007-9133-z.
(137) Cereto-Massagué, A.; Guasch, L.; Valls, C.; Mulero, M.; Pujadas, G.; Garcia-Vallvé, S. DecoyFinder: An Easy-to-Use Python GUI Application for Building Target-Specific Decoy Sets. *Bioinforma. Oxf. Engl.* **2012**, *28* (12), 1661–1662. https://doi.org/10.1093/bioinformatics/bts249.
(138) Dixon, S. L.; Duan, J.; Smith, E.; Von Bargen, C. D.; Sherman, W.; Repasky, M. P. AutoQSAR: An Automated Machine Learning Tool for Best-Practice Quantitative Structure–Activity Relationship Modeling. *Future Med. Chem.* **2016**, *8* (15), 1825–1839. https://doi.org/10.4155/fmc-2016-0093.
(139) Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. *J. Comput. Aided Mol. Des.* **2013**, *27* (3), 221–234. https://doi.org/10.1007/s10822-013-9644-8.
(140) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein−Ligand Complexes. *J. Med. Chem.* **2006**, *49* (21), 6177–6196. https://doi.org/10.1021/jm051256o.

Appendix

Chemoinformatics View on Bitter Taste Receptor Agonists in Food

The following publication was written by me and co-workers during the master thesis with information and results generated during the thesis:

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

pubs.acs.org/JAFC

Article

Chemoinformatics View on Bitter Taste Receptor Agonists in Food

Sebastian Bayer,[⊥] Ariane Isabell Mayer,[⊥] Gigliola Borgonovo, Gabriella Morini, Antonella Di Pizio,* and Angela Bassoli*

Cite This: *J. Agric. Food Chem.* 2021, 69, 13916−13924

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Food compounds with a bitter taste have a role in human health, both for their capability to influence food choice and preferences and for their possible systemic effect due to the modulation of extra-oral bitter taste receptors (TAS2Rs). Investigating the interaction of bitter food compounds with TAS2Rs is a key step to unravel their complex effects on health and to pave the way to rationally design new additives for food formulation or drugs. Here, we propose a collection of food bitter compounds, for which in vitro activity data against TAS2Rs are available. The patterns of TAS2R subtype-specific agonists were analyzed using scaffold decomposition and chemical space analysis, providing a detailed characterization of the associations between food bitter tastants and TAS2Rs.

KEYWORDS: *bitter molecules, food, bitter taste receptors, TAS2Rs, scaffold decomposition, chemical space*

INTRODUCTION

In humans, bitter taste is mediated by 25 bitter taste receptors (TAS2Rs), belonging to the superfamily of G protein-coupled receptors.[1−3] Bitter taste is generally considered as an aversive reaction that protects humans from ingesting toxic compounds. However, neither all compounds that are bitter are toxic nor do all toxins taste bitter.[4] Many bitter compounds, such as polyphenols, glucosinolates, and terpenes, have proven beneficial health effects,[5] and diets including higher amounts of bitter-tasting foods have been associated with better health.[6,7]

Babies have an innate preference for sweet, savory, and fatty tastes, while they reject even low levels of bitterness (and acidity). In fact, liking bitter-tasting foods is a learned behavior, and studies demonstrated that learning plays a major role in what comes to be identified as "food" and how the taste education in babies (especially improving bitter taste acceptance) could have a role on the health of the future adults.[8−10]

Interestingly, bitterness intensity is reported to have a special role in the food-medicine continuum: mild bitterness is associated with plants used as food, medium bitter taxa are seen both as food and medicine, while plants perceived to be very bitter are considered to be only medicinal.[11] This traditional knowledge and wisdom found a possible explanation on the fact that TAS2Rs have been discovered in several extra-oral tissues, including the gastrointestinal (GI), respiratory, reproduction, and urinary systems.[12−16] In these extra-oral locations, they may interact with endogenous compounds (produced by the microbiota and pathogens) and exogenous ones (such as bitter compounds in food), in both cases mediating systemic response, ranging from innate immunity[13,17] to metabolic effects.[18−21] GI bitter taste receptors are supposed to play a role (or at least to contribute) in maintaining a salutary balance among a healthy microbiome, diet, and weight.[22] Due to these recent discoveries, bitter taste receptors become interesting targets to improve human health, both through the diet (development of functional foods)[23−25] and as new drugs targets.[20,26,27]

Investigating the interaction between bitter food compounds and TAS2Rs is a key step to decode their complex effects on health and to pave the way to rationally design new additives for food formulation. More than 1000 compounds are reported to taste bitter.[28] Some of them are known for a long time when the practice of tasting pure compounds from isolation or synthesis was a quite common procedure. For instance, the taste of purified compounds—as recorded by the author—is usually described in the Beilstein Handbuch der Organischen Chemie (Handbook of Organic Chemistry), one of the oldest collections of organic compounds information edited from 1880 to 1998.[29] The compounds of particular interest in the food industry have been re-tested afterward using current sensory analysis methods.[30] Following the discovery of TAS2Rs, the associations with the cognate receptor(s) were investigated for many bitter compounds, including drugs—either from natural or synthetic origin—and food components. In vitro assays can be run on milligram scale samples and do not require preliminary toxicity tests.[31] Some of the recently identified TAS2R agonists have not even been submitted to sensory analysis.[32,33] Functional assays led to a new array of data connected to bitter compounds, and for a better understanding of the molecular recognition of bitter compounds, supporting the combinatorial coding of the bitter taste percept. Some bitter receptors are broadly tuned and respond to a wide range of structurally diverse bitter

Received: August 23, 2021
Revised: October 27, 2021
Accepted: October 29, 2021
Published: November 11, 2021

compounds, whereas others are very selective for one or a small number of compounds.[34] Interestingly, some bitter compounds can act as agonists for certain TAS2R subtypes and antagonists for others.[35] Moreover, TAS2Rs may work as heterooligomers,[36] and TAS2R subsets have been found to be co-expressed in the taste receptor cells of human circumvallate papillae, suggesting that cells are tuned for the subsets of bitter stimuli.[37] Several studies have been focusing on investigating the activation of specific TAS2Rs by the individual classes of bitter compounds.[38−40] Here, we propose a systematic chemoinformatic analysis of bitter compounds present in food for which activity data against TAS2Rs are available. The analysis of the patterns of TAS2R agonists provide a detailed characterization of associations between food bitter compounds and TAS2Rs, and a framework for chemoinformatics works on the growing number of food bitter compounds.

MATERIALS AND METHODS

A total of 247 natural compounds with TAS2R activity data were collected from the literature.[41] Canonical SMILES (simplified molecular-input line-entry system) were retrieved from PubChem. According to their presence in food sources, a set of 138 food bitter compounds was established (available at https://github.com/dipizio/Natural_TAS2R_agonists). 133 out of 138 were successfully classified using the chemical taxonomy analysis by the ClassyFire Batch compound classification tool (https://cfb.fiehnlab.ucdavis.edu/).[42] This set of 133 compounds was used for the following analyses.

Unique scaffolds were calculated using the Bemis−Murcko Scaffold decomposition tool available in Maestro (Schrödinger Release 2021-2: Maestro, Schrödinger, LLC, New York, NY, 2021).

MACCS fingerprints were calculated from the Canonical SMILES using RDKIT in KNIME 4.3.0. The t-distributed stochastic neighbor embedding (t-SNE) for chemical space visualization was obtained using the t-SNE (L. Jonsson) node in KNIME 4.3.0. The following parameters were set for the t-SNE analysis: dimensions to reduce: 2, iterations 5000, θ: 0.5, number of threads: 8, seed: 1615892099957. The t-SNE plot was then visualized with Plotly Express.

The alluvial plot was generated with Plotly Express. The nodes settings were as follows: thickness of 20, pad to 10, and uniform color of "gray". TAS2Rs targeted were set as the primary source (left nodes), ClassyFire superclasses of the compounds as the primary target/secondary source (middle nodes), and the number of receptors targeted per compound was set as the secondary target (left nodes). TAS2Rs were sorted using promiscuity indices (TAS2R14, −1, −46, −4, −39, −10, −5, −7, −40, −43, −16, −38, −30, −8, −31, −20, −50), superclasses by the number of compounds in each group, from largest to smallest, and compound promiscuity in increasing order (compounds that activate only one receptor, and compounds that activate two, three, four, or more than five receptors).

The following Python packages were used: NumPy, pandas, Matplotlib, Plotly Express, and Seaborn.

RESULTS AND DISCUSSION

TAS2R Agonists in Food. An extensive literature search led us to collect 247 bitter compounds from natural sources from which receptor activation data are available (https://github.com/dipizio/Natural_TAS2R_agonists). To the best of our knowledge, this is the largest collection of bitter natural compounds with associated TAS2Rs' activity values. Currently, the BitterDB has 58 compounds labeled as natural compounds (http://bitterdb.agri.huji.ac.il/, as of August 2021).[43] Of the 25 human TAS2Rs, 21 have been de-orphanized over the last two decades with both natural and synthetic compounds.[44,45] As reported by a number of chemical space studies, synthetic and natural compounds may chemically differ.[46−48] Therefore, we aim to focus our analyses on the bitter molecules of natural sources. Most of the collected bitter compounds are from plant origin, with the exception of a few amino acids and peptides that can be found both in plants and in animal products (like cheese). More specifically, our analyses focused on bitter molecules in food. The definition of "food" itself can be troublesome: in fact, some plants can be regarded as food only in case of lacking other sources (alimurgic plants),[49] some are used both as food and as medicines according to their ecosystem and culture.[50] In our classification, we tried to focus on bitter compounds in food commonly used in the western diet. A subset of 133 food bitter compounds with receptor information (food TAS2R agonists' dataset, available at https://github.com/dipizio/Natural_TAS2R_agonists) was built. The compounds of this set activate 17 bitter taste receptors.

The number of food TAS2R agonists associated with the receptor subtypes is reported in Table 1. For each set of TAS2R agonists, we calculated the number of unique scaffolds. Scaffold decomposition focuses on the ring systems (scaffolds), and for this reason, it is used as a measure of the chemical complexity.[51] The ratio of the number of scaffolds by the number of compounds gives rise to the "scaffold per compound": receptors with values higher than one are activated by chemically complex compounds (more scaffolds than compounds). The receptor with the highest "scaffold per compound" value is the TAS2R50. The high number of scaffolds is attributed to the size of amarogentin (MW of 586), the only TAS2R50 agonist in the food TAS2R agonists' dataset. Amarogentin is considered the most bitter natural molecule known to date[52,53] and targets eight TAS2Rs in total (https://github.com/dipizio/Natural_TAS2R_agonists). A

Table 1. TAS2Rs Activated by the Food Bitter Compounds (All-TAS2R-Tested Agonists)

receptor (alternative names)	n. of compounds	n. of unique scaffolds	scaffold per compound	median MW
TAS2R1	25	20	0,80	354
TAS2R4	14	34	2,43	586
TAS2R5	9	52	5,78	578
TAS2R7	10	26	2,60	535
TAS2R8	2	2	1,00	160
TAS2R10	12	20	1,67	410
TAS2R14	78	53	0,68	295
TAS2R16	8	6	0,75	318
TAS2R20 (TAS2R49, TAS2R56)	1	1	1,00	204
TAS2R30 (TAS2R47)	4	14	3,50	477
TAS2R31 (TAS2R44, TAS2R53)	2	3	1,50	271
TAS2R38 (TAS2R61)	5	6	1,20	163
TAS2R39 (TAS2R57)	47	66	1,40	286
TAS2R40 (TAS2R58)	8	8	1,00	354
TAS2R43 (TAS2R52)	8	19	2,38	261
TAS2R46 (TAS2R54)	20	31	1,55	335
TAS2R50 (TAS2R51)	1	9	9,00	586

https://doi.org/10.1021/acs.jafc.1c05057
J. Agric. Food Chem. 2021, 69, 13916−13924

high value of "scaffold per compound" is assigned also to TAS2R5, which is activated by complex polyphenolic structures, with a median MW of 578. Instead, in the case of TAS2R43, the "scaffold per compound" higher than 1 is associated with a rather low molecular size (median MW of 261), indicating a chemical diversity of the set of agonists. Among all receptors, TAS2R1, TAS2R14, and TAS2R16 have a "scaffold per compound" value lower than 1 (more compounds than scaffolds), pinpointing a chemical similarity of their agonists.

Broad- and Fine-Tuning of TAS2Rs Versus Food Bitter Compounds. Whereas the scaffold decomposition analysis of the entire dataset was aimed to investigate the TAS2R agonist sets, to compare the agonists of individual receptors, the exclusion of molecules tested only against certain subtypes was needed. Therefore, a smaller subset of 83 compounds tested toward all TAS2Rs was extracted (activity value range: 1−10 000 μM): we will refer to this subset as the "food all-TAS2R-tested agonists' dataset". Figure 1 shows the

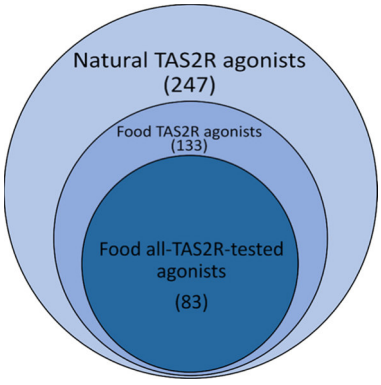


Figure 1. Venn diagram of the three datasets collected and analyzed in this work: TAS2R agonists from natural sources (in light blue), TAS2R agonists from food sources (in blue), and TAS2R agonists from food sources that were tested toward all bitter taste receptors (in dark blue). Datasets are available at https://github.com/dipizio/Natural_TAS2R_agonists

three sets of compounds used for the different analyses. To compare the receptive ranges of TAS2Rs, we defined the promiscuity index (light blue bars in Figure 2) as the number of bitter compounds that activate the individual TAS2R divided by the total number of molecules (food all-TAS2R-tested agonists' dataset).[34] To represent the diversity of the ligands set, we also calculated the number of unique scaffolds for each receptor divided by the total number of unique scaffolds, that is, 133 (gray bars in Figure 2). The height of the bars depends on the number of compounds and scaffolds of the individual receptors but also on the total numbers of molecules (i.e., 83) and scaffolds (i.e., 133); therefore, promiscuity indices per number of compounds and per scaffolds are not directly comparable. However, with this analysis we want to focus on the comparison of the different TAS2R subtypes and, therefore, on the profiles of the histograms we obtain from the analyses of the number of compounds and number of scaffolds.

TAS2R14, activated by almost half of the food all-TAS2R-tested agonists' dataset, results to be the most broadly tuned receptor. TAS2R14 is the most promiscuous receptor also when synthetic compounds are taken into account.[34] TAS2R1 follows with about half as many agonists as TAS2R14.

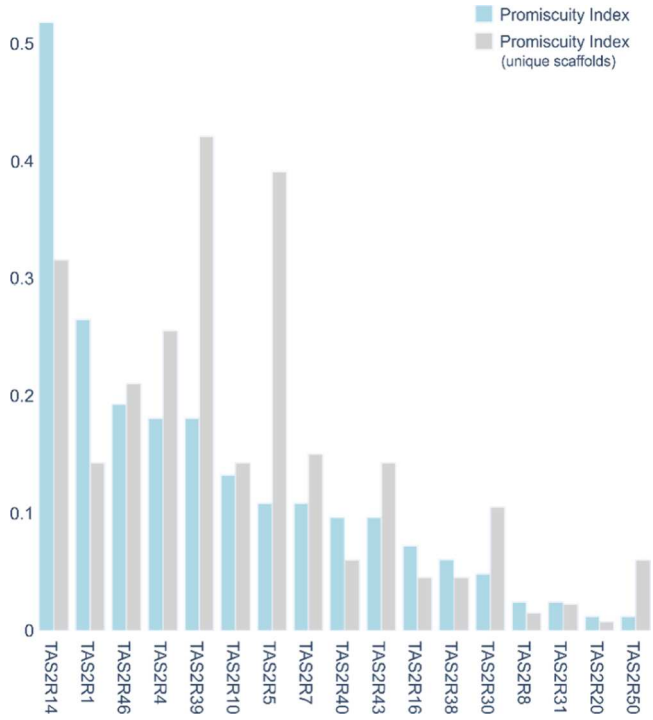


Figure 2. Receptor promiscuity indexes of bitter taste receptors toward the food all-TAS2R-tested agonists' set. Light blue bars indicate the number of compounds that activate the individual TAS2R divided by the total number of molecules; while gray bars represent the number of unique scaffolds for each receptor divided by the total number of unique scaffolds.

Intermediate-promiscuous TAS2R receptors, that is, TAS2R46, TAS2R4, and TAS2R39, follow with a promiscuity index of about 0.2. The most selective receptors are TAS2R50 and TAS2R20. Interestingly, the profile is different when looking at scaffolds. TAS2R39 and TAS2R5 are more enriched in scaffolds compared to TAS2R14. As already discussed, the high number of unique scaffolds of TAS2R5 is due to the complexity of its ligands. On the contrary, the promiscuity index calculated for the unique scaffolds of TAS2R39 agonists is higher than that of TAS2R14 agonists, although the median MW of TAS2R14 and TAS2R39 ligands is comparable.

Chemical Classes Associated with TAS2R Subtypes. The chemical classes associated with TAS2R agonists in food were assigned using the superclasses of ClassyFire, as previously done for chemical classification analyses of bitter compounds.[53,54] The ClassyFire provides an automated chemical classification based on a structure-based chemical taxonomy consisting of approximately 5000 different categories.[42] The ClassyFire chemical taxonomy consists of up to 11 different levels (kingdom, superclass, class, subclass, etc.). In our case, we decided to use the superclasses, in order to have a meaningful but simple grouping of our datasets.

The food TAS2R agonists' dataset is composed of phenylpropanoids and polyketides (with 61 compounds this superclass makes up about half of the whole set), lipids and lipid-like molecules (31), organic oxygen compounds (12), organic acids (10), organoheterocyclic compounds (8), benzenoids (8), mixed metal/nonmetal compounds (2), organosulfur compounds (1, i.e., allyl isothiocyanate), and alkaloids (1, i.e., quinine). The food all-TAS2R-tested agonists' dataset has a substantial reduction of phenylpropanoids and

https://doi.org/10.1021/acs.jafc.1c05057
J. Agric. Food Chem. 2021, 69, 13916−13924

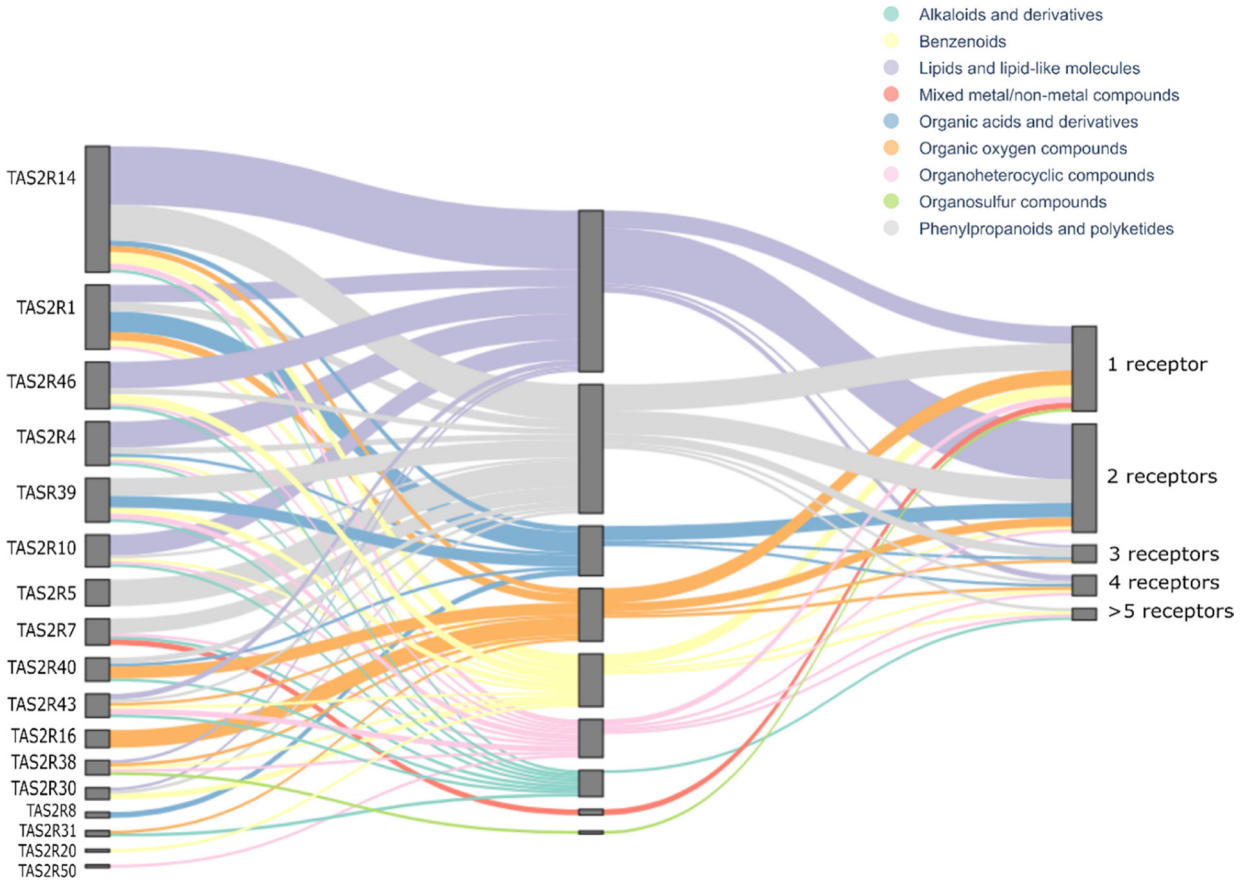


Figure 3. Alluvial plot depicting the associations of food bitter compounds with bitter taste receptors. On the left, receptors are ordered by their promiscuity (i.e., the number of compounds that activate each TAS2R); on the right, compounds (food all-TAS2R-tested agonists' set) are ordered by their promiscuity (i.e., how many receptors each compound can activate). The different colors of flows represent the superclasses (central nodes of the plot) the compounds are associated with.

polyketides to about a fourth (22), but the distribution of other superclasses is rather similar: lipids and lipid-like molecules (33), organic oxygen compounds (10), benzenoids (7), organic acids (7), organoheterocycles (5), mixed metal/nonmetal compounds (2), organosulfur compounds (1), and alkaloids (1).

To investigate the target promiscuity of TAS2Rs toward chemical superclasses, we analyzed the associations of food all-TAS2R-tested agonists' with the receptors (Figure 3). The outstanding promiscuity of TAS2R14 reflects here again, as it is the only channel that has ligands from seven superclasses. TAS2R1, TAS2R4, and TAS2R43 follow with agonists from six superclasses. In the middle, with compounds from four and five superclasses, are TAS2R46, TAS2R39, TAS2R10, TAS2R38, TAS2R7, and TAS2R40. TASR38 stands out as it only has five agonists in our set, each of them belonging to a different superclass. Among the very selective receptors, TAS2R8 and TAS2R31 are activated by two compounds, while TAS2R20 and TAS2R50 by a single agonist, L-tryptophan[55] and amarogentin,[56] respectively, which target multiple other TAS2Rs. Interestingly, we have two receptors activated by one single superclass of compounds, TAS2R16 and TAS2R5. TAS2R16 is activated by six different compounds, all organic oxygen compounds. The similarity of TAS2R16 agonists was highlighted also with the scaffolds analysis, and it is well renowned because TAS2R16 is specialized in the recognition of "bitter sugars",[57] as glucopyranosides[38] and glucosinolates.[32] Similarly, TAS2R5 shows specificity toward polyphenols: it is activated only by compounds belonging to the phenylpropanoids and polyketides superclass. This indication could be relevant for TAS2R5-targeted nutritional studies in the future because polyphenols are biomolecules with a well-established positive impact on nutrition and health.[58−60] Moreover, TAS2R5 is expressed in the small and large intestine and it may be a player in recognizing bitter polyphenols in the diet with positive effects on health.[61]

The right side of the alluvial plot in Figure 3 reports the compound promiscuity/selectivity, that is, the number of receptors each compound in the food all-TAS2R-tested agonists' set can activate. Interestingly, as observed analyzing also synthetic compounds,[34] promiscuous compounds can activate both promiscuous and selective receptors. Most of the compounds activate one or two receptors, and compound promiscuity is not driven by specific superclasses. Superclasses are distributed over all five promiscuity groups, and very selective compounds (those activating only one receptor) belong to all superclasses. Superclasses are defined in the middle nodes of the plot (Figure 3). The mixed metal/nonmetal (i.e., $CaCl_2$ and $MgCl_2$) and the organosulfur (i.e., isothiocyanate) superclasses show the highest specificity of the superclasses, targeting only TAS2R7[62,63] and TAS2R38,[39,64] respectively. The phenylpropanoids are at the other end of the spectrum activating a total of 11 receptors, followed by the organoheterocyclic compounds and benzenoids with 10 associated TAS2Rs. Most of the superclasses activate between six and nine receptors. Interestingly, the number of molecules

https://doi.org/10.1021/acs.jafc.1c05057
J. Agric. Food Chem. 2021, 69, 13916−13924

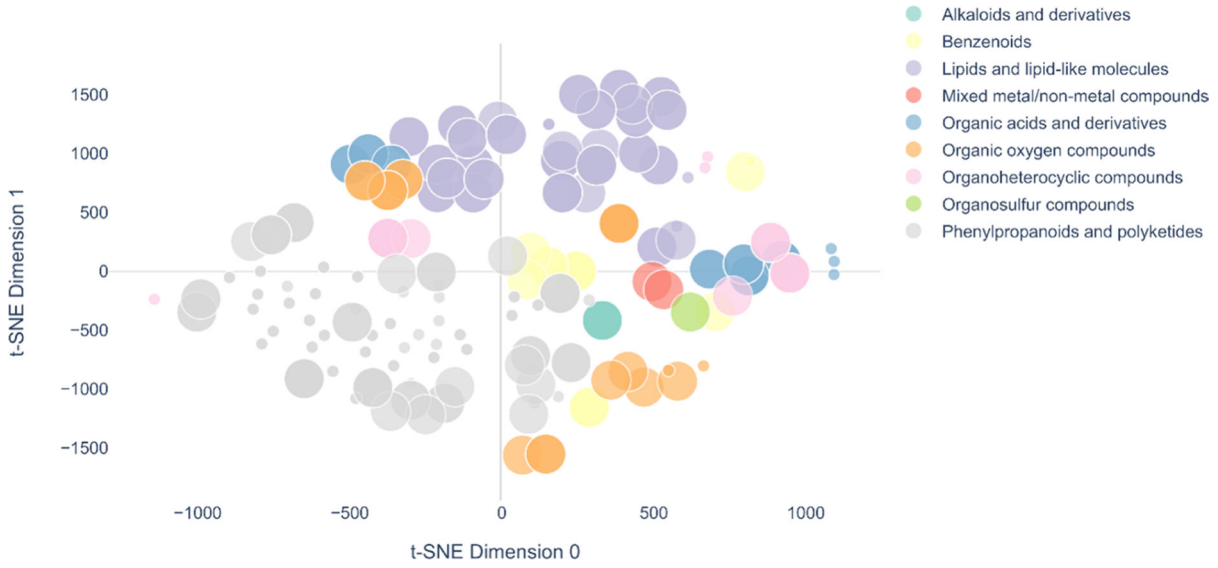


Figure 4. t-SNE plot of the food all-TAS2R-tested agonists. Compounds are colored according to the ClassyFire superclasses and sized in a bigger size when belonging to the food all-TAS2R agonists' set. t-SNE dimension 0 and 1 for all compounds reported in this plot are available at https://github.com/dipizio/Natural_TAS2R_agonists.

in a given superclass is not the only driver behind compound promiscuity, for example, 44 compounds belong to the lipids and lipid-like superclass and activate eight receptors, whereas quinine, the single compound classified as alkaloid, activates nine receptors.

Chemical Space of Bitter Compounds in Food. The classification into superclasses, which includes 26 organic and 5 inorganic generic categories,[42] necessarily leads to groups that do not completely reflect a rationale in the chemistry of compounds: for instance, only quinine is classified as alkaloid, whereas caffeine is among heterocycles; allyl isothiocyanate is in sulfur compounds, whereas phenethyl isothiocyanate is within benzenoids, and the glucosinolates sinigrin and glucoputranjivin, which also contain a sulfur atom, are in the group of oxygenated compounds. In order to better understand the chemical similarity within and between superclasses, we analyzed the chemical space of the food TAS2R agonists' set (Figure 4). Specifically, a fingerprint-based chemical space was built and visualized by t-SNE, a machine learning algorithm for data visualization developed by van der Maaten and Hinton in 2008.[65] The t-SNE performs a nonlinear projection of data points from a high-dimensional space to a low-dimensional space. Our main goal for building a chemical space was to investigate the chemical similarity of compounds in our dataset and attempt to identify (chemistry-driven and/or receptor-driven) subgroups within it. Compared to other dimensionality reduction methods, such as the principal component analysis, that concentrate on placing dissimilar data points far apart, the t-SNE tends to place similar data points close together, and we found it best fitting our aim.

In the t-SNE plot in Figure 4, the compounds are colored according to the superclasses and sized according to their belonging to the food all-TAS2R-tested (big dots) or food TAS2R (small dots) agonists. Interestingly, the food all-TAS2R-tested agonists' set has similar distribution in the chemical space as the food TAS2R agonists' set and a similar representation of superclasses. Indeed, most often the experimental screening of certain compounds toward specific TAS2Rs is focused on chemotypes that were previously screened toward all receptors.[40,66] Interestingly, with some exceptions, molecules within the same superclasses map the same regions of the chemical space. The largest superclass, phenylpropanoids and polyketides (in gray in Figure 4) span all around the lower left-hand side of the plot, while lipids and lipid-like molecules (in violet) spread on the positive side of the t-SNE-1 axis. The two mixed metal/nonmetal compounds (red dots) cluster in the right-center, very close to each other. Benzenoids (yellow dots) spread on the right-hand side of the plot, the core cluster in the center is mapped by one-ring structures (i.e., phenylethyl isothiocyanate, protocatechuic acid, pyrocatechin, salicylic acid, and vanillic acid), while the more complex structures of amarogentin, 11β,13-dihydrolactucopicrin, and lactucopicrin are the dots spreading the bottom and left side, respectively. The organosulfur compound allyl isothiocyanate (green) is very close to the structurally related phenylethyl isothiocyanate (yellow dot), which is indeed isolated from the other benzenoids. The t-SNE analysis, therefore, correctly associates compounds formally assigned to different superclasses but with the same functional groups. The organoheterocyclic compounds (eight pink dots in the plot) spread over the plot. This was expected because the heterocycle superclass includes very different oxygenated, nitrogen and other five- and six-member ring compounds. The pink dots form distinct groups: caffeine close to theobromine, lactucin close to 11β,13-dihydrolactucin, ethylpyrazine close to thiamine and L-tryptophan, and skimmianine, an alkaloid isolated from *Ruta graveolens* L., a plant having occasional uses as a food,[33] isolated on the extreme left side of the plot.

Likewise, structurally similar compounds belonging to different superclasses are associated by vicinity in the t-SNE plot. The hop-derived humulones (humulone, adhumulone, and cohumulone) are labeled as organic oxygen compounds and are in the upper left corner of the t-SNE plot (orange dots), directly opposite to three lupulones (adlupulone, colupulone, and lupulone), labeled as organic acids (blue dots). Lupulones are indeed structurally very similar to humulones and less similar to other organic acids, amino

https://doi.org/10.1021/acs.jafc.1c05057
J. Agric. Food Chem. 2021, 69, 13916−13924

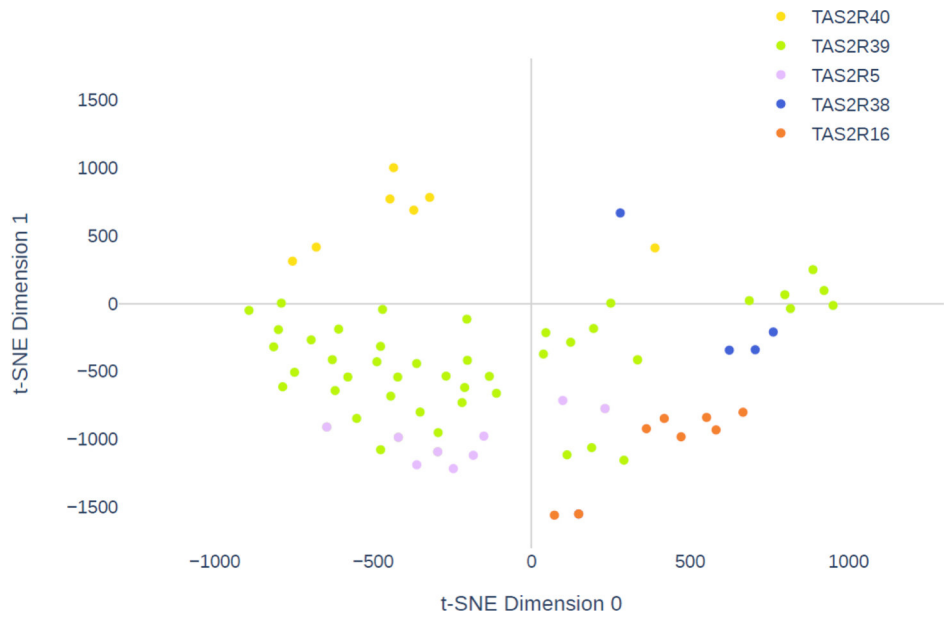


Figure 5. Receptor space of TAS2R40, TAS2R39, TAS2R5, TAS2R38, and TAS2R16 agonists.

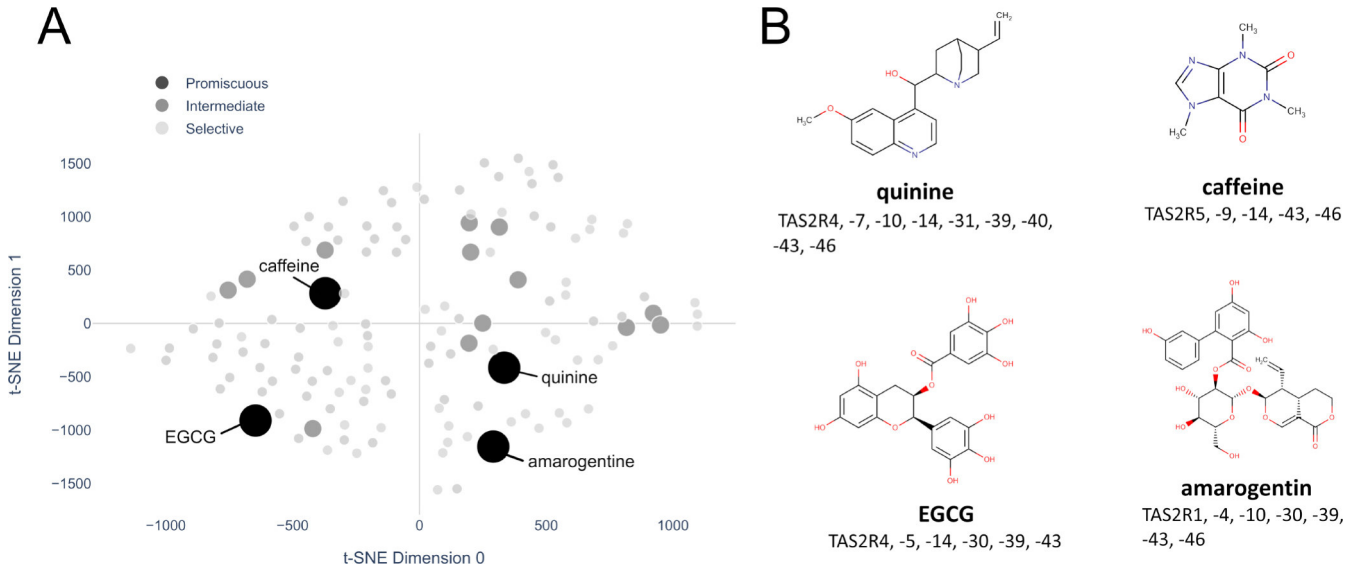


Figure 6. (A) t-SNE plot with compounds colored and sized according to their promiscuity: promiscuous compounds (activating more than five receptors) in black, intermediate (three and four receptors), and selective bitter compounds (one and two receptors). (B) Chemical structures of the most promiscuous compounds, that is, quinine, amarogentin, EGCG, and caffeine.

acids, and di/tripeptides that are placed on the right-hand side of the chemical space.

Coloring compounds in the t-SNE plot according to cognate receptors provides a view of the "receptor space" covered by food bitter compounds. In Figure 5, only TAS2R39, TAS2R5, TAS2R16, TAS2R38, and TAS2R40 ligands are plotted, for a clearer definition of the formed groups, but all individual agonists are reported in the Supporting Information Figures S1−S17. TAS2R14 agonists are the most numerous and occupy a large portion of the plot. TAS2R1, TAS2R4, TAS2R10, and TAS2R46 agonists come from many different superclasses (Figure 3) and are spread over the plot. TAS2R39 ligands mostly occupy the region of phenylpropanoids and polyketides in Figure 4. In the same area, at the bottom, TAS2R5 agonists cluster together (lilac dots). TAS2R16 agonists, all organic oxygen compounds, are also very close in space (orange dots); only two molecules, sinigrin and glucoputranjivin, are distant from the main group, reasonably because they both contain a sulfurated group. TAS2R40 ligands cluster together at the left side of the plot (yellow dots), with the exception of quinine and pantothenic acid. Three of the five TAS2R38 agonists group together closely, that is, allyl isothiocyanate, phenylethyl isothiocyanate, and ethylpyrazine, whereas limonin and sinigrin are structurally very different and distant from the core cluster (see Figure S9).

As previously shown in Figure 3, the promiscuity and selectivity of bitter compounds cannot be linked by superclasses. Indeed, annotating compounds in the t-SNE plot by their promiscuity, we do not recognize promiscuous-specific or selective-specific regions, but all types of compounds spread in the chemical space (Figure 6A). On the other hand, very similar compounds may have different promiscuity profiles:

https://doi.org/10.1021/acs.jafc.1c05057
J. Agric. Food Chem. 2021, 69, 13916−13924

caffeine and theobromine are structurally very similar, are in the same superclass and very close in space in the t-SNE plot, but one is selective for a single receptor and the other is very promiscuous.

The four most promiscuous compounds are quinine, amarogentin, epigallocatechin gallate (EGCG), and caffeine (Figure 6). Quinine activates nine receptors, that is, TAS2R4, TAS2R7, TAS2R10, TAS2R14, TAS2R31, TAS2R39, TAS2R40, TAS2R43, and TAS2R46, amarogentin seven receptors, that is, TAS2R1, TAS2R4, TAS2R10, TAS2R30, TAS2R39, TAS2R43, and TAS2R46, EGCG six receptors, that is, TAS2R4, TAS2R5, TAS2R14, TAS2R30, TAS2R39, and TAS2R43, and caffeine five receptors, that is, TAS2R5, TAS2R9, TAS2R14, TAS2R43, and TAS2R46. To be noted as TAS2R43 is a receptor shared by all four promiscuous compounds. The low number of receptor data does not allow us to drive conclusions about the roles of specific receptors for the recognition of food bitter compounds, and indeed more information is needed to see if our observations can be turned into assumptions.

In conclusion, the collection of data about TAS2R agonists in food and their analyses provide a chemical re-organization of current data and a clearer picture of the associations between food bitter compounds and their receptors. The number of compounds analyzed in this paper (133) is still quite low, but a very much higher number of TAS2R agonists is expected to be identified in food as far as the isolation of single compounds and in vitro assays on TAS2Rs will be performed. This will allow further studies and refinements on the structural classification of bitterants in relationship with their receptors. In the past, bitter compounds were identified mainly using sensory methods, which likely detect compounds with the lower recognition threshold. Further investigations aimed at identifying compounds with a mild bitter taste, difficult to be identified using sensory and taste-guided analysis, could deepen the knowledge of TAS2R agonists in food and may represent an opportunity to detect biomolecules with potential beneficial effects on health, and therefore applications in food science, nutrition, and medicine.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jafc.1c05057.

TAS2R14 agonists' space; TAS2R39 agonists' space; TAS2R5 agonists' space; TAS2R16 agonists' space; TAS2R46 agonists' space; TAS2R4 agonists' space; TAS2R10 agonists' space; TAS2R1 agonists' space; TAS2R38 agonists' space; TAS2R30 agonists' space; TAS2R40 agonists' space; TAS2R43 agonists' space; TAS2R7 agonists' space; TAS2R50 agonist's space; TAS2R31 agonists' space; TAS2R31 agonists' space; TAS2R20 agonist's space (PDF)

The database of natural TAS2R agonists and datasets used for the analyses are available at https://github.com/dipizio/Natural_TAS2R_agonists

■ AUTHOR INFORMATION

Corresponding Authors

Antonella Di Pizio − *Leibniz Institute for Food Systems Biology at the Technical University of Munich, D-85354 Freising, Germany;* orcid.org/0000-0002-8520-5165; Phone: +49(0)8161716516; Email: a.dipizio.leibniz-lsb@tum.de

Angela Bassoli − *Department of Food, Environmental and Nutritional Sciences-DeFENS, University of Milan, 20147 Milano, Italy;* orcid.org/0000-0001-5442-3808; Phone: +39(0)250316815; Email: angela.bassoli@unimi.it

Authors

Sebastian Bayer − *Leibniz Institute for Food Systems Biology at the Technical University of Munich, D-85354 Freising, Germany; Faculty of Life Sciences, University of Vienna, 1030 Vienna, Austria*

Ariane Isabell Mayer − *Department of Food, Environmental and Nutritional Sciences-DeFENS, University of Milan, 20147 Milano, Italy*

Gigliola Borgonovo − *Department of Food, Environmental and Nutritional Sciences-DeFENS, University of Milan, 20147 Milano, Italy*

Gabriella Morini − *University of Gastronomic Sciences, 12042 Pollenzo, (Bra, CN), Italy;* orcid.org/0000-0001-5494-9201

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jafc.1c05057

Author Contributions

⊥S.B. and A.I.M. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Adler, E.; Hoon, M. A.; Mueller, K. L.; Chandrashekar, J.; Ryba, N. J. P.; Zuker, C. S. A novel family of mammalian taste receptors. *Cell* **2000**, *100*, 693−702.

(2) Chandrashekar, J.; Mueller, K. L.; Hoon, M. A.; Adler, E.; Feng, L.; Guo, W.; Zuker, C. S.; Ryba, N. J. P. T2Rs function as bitter taste receptors. *Cell* **2000**, *100*, 703−711.

(3) Matsunami, H.; Montmayeur, J.-P.; Buck, L. B. A family of candidate taste receptors in human and mouse. *Nature* **2000**, *404*, 601−604.

(4) Nissim, I.; Dagan-Wiener, A.; Niv, M. Y. The taste of toxicity: A quantitative analysis of bitter and toxic molecules. *IUBMB Life* **2017**, *69*, 938−946.

(5) Bayram, B.; González-Sarrías, A.; Istas, G.; Garcia-Aloy, M.; Morand, C.; Tuohy, K.; García-Villalba, R.; Mena, P. Breakthroughs in the Health Effects of Plant Food Bioactives: A Perspective on Microbiomics, Nutri(epi)genomics, and Metabolomics. *J. Agric. Food Chem.* **2018**, *66*, 10686−10692.

(6) Lease, H.; Hendrie, G. A.; Poelman, A. A. M.; Delahunty, C.; Cox, D. N. A Sensory-Diet database: A tool to characterise the sensory qualities of diets. *Food Qual. Prefer.* **2016**, *49*, 20−32.

(7) Poelman, A. A. M.; Delahunty, C. M.; de Graaf, C. Vegetables and other core food groups: A comparison of key flavour and texture properties. *Food Qual. Prefer.* **2017**, *56*, 1−7.

(8) Mennella, J. A.; Reiter, A. R.; Daniels, L. M. Vegetable and Fruit Acceptance during Infancy: Impact of Ontogeny, Genetics, and Early Experiences. *Adv. Nutr.* **2016**, *7*, 211S−219S.

(9) Mennella, J. A. Ontogeny of taste preferences: basic biology and implications for health. *Am. J. Clin. Nutr.* **2014**, *99*, 704S−711S.

(10) Beckerman, J. P.; Alike, Q.; Lovin, E.; Tamez, M.; Mattei, J. The Development and Public Health Implications of Food Preferences in Children. *Front. Nutr.* **2017**, *4*, 66.

(11) Pieroni, A.; Nebel, S.; Quave, C.; Münz, H.; Heinrich, M. Ethnopharmacology of liakra: traditional weedy vegetables of the Arbereshe of the Vulture area in southern Italy. *J. Ethnopharmacol.* **2002**, *81*, 165−185.

(12) Shaik, F. A.; Singh, N.; Arakawa, M.; Duan, K.; Bhullar, R. P.; Chelikani, P. Bitter taste receptors: Extraoral roles in pathophysiology. *Int. J. Biochem. Cell Biol.* **2016**, *77*, 197−204.
(13) Harmon, C. P.; Deng, D.; Breslin, P. A. S. Bitter Taste Receptors (T2Rs) are Sentinels that Coordinate Metabolic and Immunological Defense Responses. *Curr. Opin. Physiol.* **2021**, *20*, 70−76.
(14) Ifrim-Chen, F. "Ectopic" Gustative and Olfactory Receptors in the Brain − New Targets for Neurodegeneration Therapy? *Farmacia* **2019**, *67*, 34−42.
(15) ZhuGe, R. H.; Roura, E.; Behrens, M. Editorial: Extra-Oral Taste Receptors: Function, Disease and Evolution. *Front. Physiol.* **2020**, *11*, 607134.
(16) Behrens, M.; Ziegler, F. Structure-Function Analyses of Human Bitter Taste Receptors-Where Do We Stand? *Molecules* **2020**, *25*, 4423.
(17) Tran, H. T. T.; Stetter, R.; Herz, C.; Spottel, J.; Krell, M.; Hanschen, F. S.; Schreiner, M.; Rohn, S.; Behrens, M.; Lamy, E. Allyl Isothiocyanate: A TAS2R38 Receptor-Dependent Immune Modulator at the Interface Between Personalized Medicine and Nutrition. *Front. Immunol.* **2021**, *12*, 669005.
(18) Dotson, C. D.; Zhang, L.; Xu, H.; Shin, Y.-K.; Vigues, S.; Ott, S. H.; Elson, A. E. T.; Choi, H. J.; Shaw, H.; Egan, J. M.; Mitchell, B. D.; Li, X.; Steinle, N. I.; Munger, S. D. Bitter taste receptors influence glucose homeostasis. *PLoS One* **2008**, *3*, No. e3974.
(19) Calvo, S. S.-C.; Egan, J. M. The endocrinology of taste receptors. *Nat. Rev. Endocrinol.* **2015**, *11*, 213−227.
(20) Behrens, M.; Somoza, V. Gastrointestinal taste receptors: could tastants become drugs? *Curr. Opin. Endocrinol. Diabetes Obes.* **2020**, *27*, 110−114.
(21) Sterneder, S.; Stoeger, V.; Dugulin, C. A.; Liszt, K. I.; Di Pizio, A.; Korntheuer, K.; Dunkel, A.; Eder, R.; Ley, J. P.; Somoza, V. Astringent Gallic Acid in Red Wine Regulates Mechanisms of Gastric Acid Secretion via Activation of Bitter Taste Sensing Receptor TAS2R4. *J. Agric. Food Chem.* **2021**, *69*, 10550−10561.
(22) Turner, A.; Veysey, M.; Keely, S.; Scarlett, C.; Lucock, M.; Beckett, E. L. Interactions between Bitter Taste, Diet and Dysbiosis: Consequences for Appetite and Obesity. *Nutrients* **2018**, *10*, 1336.
(23) Cires, M. J.; Wong, X.; Carrasco-Pozo, C.; Gotteland, M. The Gastrointestinal Tract as a Key Target Organ for the Health-Promoting Effects of Dietary Proanthocyanidins. *Front. Nutr.* **2016**, *3*, 57.
(24) Ekstrand, B.; Young, J. F.; Rasmussen, M. K. Taste receptors in the gut—A new target for health promoting properties in diet. *Food Res. Int.* **2017**, *100*, 1−8.
(25) Tiroch, J.; Sterneder, S.; Di Pizio, A.; Lieder, B.; Hoelz, K.; Holik, A.-K.; Pignitter, M.; Behrens, M.; Somoza, M.; Ley, J. P.; Somoza, V. Bitter Sensing TAS2R50 Mediates the trans-Resveratrol-Induced Anti-inflammatory Effect on Interleukin 6 Release in HGF-1 Cells in Culture. *J. Agric. Food Chem.* **2021**, DOI: 10.1021/acs.jafc.0c07058.
(26) Di Pizio, A.; Behrens, M.; Krautwurst, D. Beyond the Flavour: The Potential Druggability of Chemosensory G Protein-Coupled Receptors. *Int. J. Mol. Sci.* **2019**, *20*, 1402.
(27) Lee, S.-J.; Depoortere, I.; Hatt, H. Therapeutic potential of ectopic olfactory and taste receptors. *Nat. Rev. Drug Discov.* **2019**, *18*, 116−138.
(28) Wiener, A.; Shudler, M.; Levit, A.; Niv, M. Y. BitterDB: a database of bitter compounds. *Nucleic Acids Res.* **2012**, *40*, D413−D419.
(29) Sarkowski, H. Beilsteins,Handbuch der organischen Chemie'. In *Sehr geehrter Herr!: Autorenbriefe aus dem Springer-Archiv*; Sarkowski, H., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 1982; pp 7−9.
(30) Ryland, D.; Goldberg, E.; Aliani, M., Sensory Evaluation Techniques for Detecting and Quantifying Bitterness in Food and Beverages. *Bitterness: Perception, Chemistry and Food Processing*, 1st ed; Wiley, 2017; pp 133−160.
(31) Riedel, K.; Sombroek, D.; Fiedler, B.; Siems, K.; Krohn, M. Human cell-based taste perception—a bittersweet job for industry. *Nat. Prod. Rep.* **2017**, *34*, 484−495.
(32) Borgonovo, G.; Zimbaldi, N.; Guarise, M.; Bedussi, F.; Winnig, M.; Vennegeerts, T.; Bassoli, A. Glucosinolates in Sisymbrium officinale (L.) Scop.: Comparative Analysis in Cultivated and Wild Plants and in Vitro Assays with T2Rs Bitter Taste Receptors. *Molecules* **2019**, *24*(). DOI: 10.3390/molecules24244572
(33) Mancuso, G.; Borgonovo, G.; Scaglioni, L.; Bassoli, A. Phytochemicals from Ruta graveolens Activate TAS2R Bitter Taste Receptors and TRP Channels Involved in Gustation and Nociception. *Molecules* **2015**, *20*, 18907−18922.
(34) Di Pizio, A.; Niv, M. Y. Promiscuity and selectivity of bitter molecules and their receptors. *Bioorg. Med. Chem.* **2015**, *23*, 4082−4091.
(35) Brockhoff, A.; Behrens, M.; Roudnitzky, N.; Appendino, G.; Avonto, C.; Meyerhof, W. Receptor agonism and antagonism of dietary bitter compounds. *J. Neurosci.* **2011**, *31*, 14775−14782.
(36) Kuhn, C.; Bufe, B.; Batram, C.; Meyerhof, W. Oligomerization of TAS2R bitter taste receptors. *Chem. Senses* **2010**, *35*, 395−406.
(37) Behrens, M.; Foerster, S.; Staehler, F.; Raguse, J.-D.; Meyerhof, W. Gustatory expression pattern of the human TAS2R bitter receptor gene family reveals a heterogenous population of bitter responsive taste receptor cells. *J. Neurosci.* **2007**, *27*, 12630−12640.
(38) Bufe, B.; Hofmann, T.; Krautwurst, D.; Raguse, J.-D.; Meyerhof, W. The human TAS2R16 receptor mediates bitter taste in response to beta-glucopyranosides. *Nat. Genet.* **2002**, *32*, 397−401.
(39) Bufe, B.; Breslin, P. A. S.; Kuhn, C.; Reed, D. R.; Tharp, C. D.; Slack, J. P.; Kim, U.-K.; Drayna, D.; Meyerhof, W. The molecular basis of individual differences in phenylthiocarbamide and propylthiouracil bitterness perception. *Curr. Biol.* **2005**, *15*, 322−327.
(40) Roland, W. S. U.; van Buren, L.; Gruppen, H.; Driesse, M.; Gouka, R. J.; Smit, G.; Vincken, J.-P. Bitter Taste Receptor Activation by Flavonoids and Isoflavonoids: Modeled Structural Requirements for Activation of hTAS2R14 and hTAS2R39. *J. Agric. Food Chem.* **2013**, *61*, 10454−10466.
(41) Mayer, A. A bitter taste strategy: T2Rs and TRP receptors as molecular probes for the identification of useful nutrients in food plants. *Master of Science Molekulare Ernährungswissenschaft*; University of Hohenheim, 2019.
(42) Djoumbou Feunang, Y.; Eisner, R.; Knox, C.; Chepelev, L.; Hastings, J.; Owen, G.; Fahy, E.; Steinbeck, C.; Subramanian, S.; Bolton, E.; Greiner, R.; Wishart, D. S. ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. *J. Cheminf.* **2016**, *8*, 61.
(43) Dagan-Wiener, A.; Di Pizio, A.; Nissim, I.; Bahia, M. S.; Dubovski, N.; Margulis, E.; Niv, M. Y. BitterDB: taste ligands and receptors database in 2019. *Nucleic Acids Res.* **2019**, *47*, D1179−D1185.
(44) Meyerhof, W.; Batram, C.; Kuhn, C.; Brockhoff, A.; Chudoba, E.; Bufe, B.; Appendino, G.; Behrens, M. The molecular receptive ranges of human TAS2R bitter taste receptors. *Chem. Senses* **2010**, *35*, 157−170.
(45) Lossow, K.; Hübner, S.; Roudnitzky, N.; Slack, J. P.; Pollastro, F.; Behrens, M.; Meyerhof, W. Comprehensive Analysis of Mouse Bitter Taste Receptors Reveals Different Molecular Receptive Ranges for Orthologous Receptors in Mice and Humans. *J. Biol. Chem.* **2016**, *291*, 15358−15377.
(46) Stahura, F. L.; Godden, J. W.; Xue, L.; Bajorath, J. Distinguishing between natural products and synthetic molecules by descriptor Shannon entropy analysis and binary QSAR calculations. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 1245−1252.
(47) Skinnider, M. A.; Dejong, C. A.; Franczak, B. C.; McNicholas, P. D.; Magarvey, N. A. Comparative analysis of chemical similarity methods for modular natural products with a hypothetical structure enumeration algorithm. *J. Cheminf.* **2017**, *9*, 46.
(48) Grisoni, F.; Merk, D.; Consonni, V.; Hiss, J. A.; Tagliabue, S. G.; Todeschini, R.; Schneider, G. Scaffold hopping from natural

products to synthetic mimetics by holistic molecular similarity. *Commun. Chem.* **2018**, *1*, 44.
(49) Grauso, L.; Emrick, S.; de Falco, B.; Lanzotti, V.; Bonanomi, G. Common dandelion: a review of its botanical, phytochemical and pharmacological profiles. *Phytochem. Rev.* **2019**, *18*, 1115−1132.
(50) Pieroni, A.; Sõukand, R. Ethnic and religious affiliations affect traditional wild plant foraging in Central Azerbaijan. *Genet. Resour. Crop Evol.* **2019**, *66*, 1495−1513.
(51) Saldivar-Gonzalez, F. I.; Medina-Franco, J. L. Chemoinformatics approaches to assess chemical diversity and complexity of small molecules. *Small Molecule Drug Discovery*; Elsevier, 2019; pp 83−102.
(52) Belitz, H. D.; Wieser, H. Bitter compounds: Occurrence and structure-activity relationships. *Food Rev. Int.* **1985**, *1*, 271−354.
(53) Margulis, E.; Dagan-Wiener, A.; Ives, R. S.; Jaffari, S.; Siems, K.; Niv, M. Y. Intense bitterness of molecules: Machine learning for expediting drug discovery. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 568−576.
(54) Dunkel, A.; Hofmann, T.; Di Pizio, A. In Silico Investigation of Bitter Hop-Derived Compounds and Their Cognate Bitter Taste Receptors. *J. Agric. Food Chem.* **2020**, *68*, 10414−10423.
(55) Bassoli, A.; Borgonovo, G.; Caremoli, F.; Mancuso, G. The taste of D- and L-amino acids: In vitro binding assays with cloned human bitter (TAS2Rs) and sweet (TAS1R2/TAS1R3) receptors. *Food Chem.* **2014**, *150*, 27−33.
(56) Behrens, M.; Brockhoff, A.; Batram, C.; Kuhn, C.; Appendino, G.; Meyerhof, W. The Human Bitter Taste Receptor hTAS2R50 Is Activated by the Two Natural Bitter Terpenoids Andrographolide and Amarogentin. *J. Agric. Food Chem.* **2009**, *57*, 9860−9866.
(57) Fierro, F.; Giorgetti, A.; Carloni, P.; Meyerhof, W.; Alfonso-Prieto, M. Dual binding mode of "bitter sugars" to their human bitter taste receptor target. *Sci. Rep.* **2019**, *9*, 8437.
(58) Cardona, F.; Andrés-Lacueva, C.; Tulipani, S.; Tinahones, F. J.; Queipo-Ortuño, M. I. Benefits of polyphenols on gut microbiota and implications in human health. *J. Nutr. Biochem.* **2013**, *24*, 1415−1422.
(59) Ganesan, K.; Xu, B. A Critical Review on Polyphenols and Health Benefits of Black Soybeans. *Nutrients* **2017**, *9*, 455.
(60) Visioli, F.; Lastra, C. A. D. L.; Andres-Lacueva, C.; Aviram, M.; Calhau, C.; Cassano, A.; D'Archivio, M.; Faria, A.; Favé, G.; Fogliano, V.; Llorach, R.; Vitaglione, P.; Zoratti, M.; Edeas, M. Polyphenols and human health: a prospectus. *Crit. Rev. Food Sci. Nutr.* **2011**, *51*, 524−546.
(61) Wang, Q.; Liszt, K. I.; Deloose, E.; Canovai, E.; Thijs, T.; Farré, R.; Ceulemans, L. J.; Lannoo, M.; Tack, J.; Depoortere, I. Obesity alters adrenergic and chemosensory signaling pathways that regulate ghrelin secretion in the human gut. *Faseb. J.* **2019**, *33*, 4907−4920.
(62) Behrens, M.; Redel, U.; Blank, K.; Meyerhof, W. The human bitter taste receptor TAS2R7 facilitates the detection of bitter salts. *Biochem. Biophys. Res. Commun.* **2019**, *512*, 877−881.
(63) Wang, Y.; Zajac, A. L.; Lei, W.; Christensen, C. M.; Margolskee, R. F.; Bouysset, C.; Golebiowski, J.; Zhao, H.; Fiorucci, S.; Jiang, P. Metal Ions Activate the Human Taste Receptor TAS2R7. *Chem. Senses* **2019**, *44*, 339−347.
(64) Qin, C.; Qin, Z.; Zhao, D.; Pan, Y.; Zhuang, L.; Wan, H.; Di Pizio, A.; Malach, E.; Niv, M. Y.; Huang, L.; Hu, N.; Wang, P. A bioinspired in vitro bioelectronic tongue with human T2R38 receptor for high-specificity detection of N-C = S-containing compounds. *Talanta* **2019**, *199*, 131−139.
(65) van der Maaten, L.; Hinton, G. Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579−2605.
(66) Brockhoff, A.; Behrens, M.; Massarotti, A.; Appendino, G.; Meyerhof, W. Broad tuning of the human bitter taste receptor hTAS2R46 to various sesquiterpene lactones, clerodane and labdane diterpenoids, strychnine, and denatonium. *J. Agric. Food Chem.* **2007**, *55*, 6236−6243.

https://doi.org/10.1021/acs.jafc.1c05057
J. Agric. Food Chem. 2021, 69, 13916−13924