Published in final edited form as:

J Phys Chem Lett. 2018 May 03; 9(9): 2235–2240. doi:10.1021/acs.jpclett.8b00633.

Agonists of G–Protein-Coupled Odorant Receptors Are Predicted from Chemical Features

C. Bushdid[†,#], **C. A. de March**[‡,#], **S. Fiorucci**[†], **H. Matsunami**[*,‡,§], **J. Golebiowski**[*,†,‖]

[†]Institute of Chemistry of Nice, UMR CNRS 7272, Université Côte d'Azur, Nice, France

[‡]Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, North Carolina 27710, United States

[§]Department of Neurobiology and Duke Institute for Brain Sciences, Duke University, Durham, North Carolina 27710, United States

[‖]Department of Brain & Cognitive Sciences, DGIST, Daegu, Republic of Korea

Abstract

Predicting the activity of chemicals for a given odorant receptor is a longstanding challenge. Here the activity of 258 chemicals on the human G-protein-coupled odorant receptor (OR)51E1, also known as prostate-specific G-protein-coupled receptor 2 (PSGR2), was virtually screened by machine learning using 4884 chemical descriptors as input. A systematic control by functional in vitro assays revealed that a support vector machine algorithm accurately predicted the activity of a screened library. It allowed us to identify two novel agonists in vitro for OR51E1. The transferability of the protocol was assessed on OR1A1, OR2W1, and MOR256–3 odorant receptors, and, in each case, novel agonists were identified with a hit rate of 39–50%. We further show how ligands' efficacy is encoded into residues within OR51E1 cavity using a molecular modeling protocol. Our approach allows widening the chemical spaces associated with odorant receptors. This machine-learning protocol based on chemical features thus represents an efficient tool for screening ligands for G-protein-coupled odorant receptors that modulate non-olfactory functions or, upon combinatorial activation, give rise to our sense of smell.

Graphical Abstract

[*]**Corresponding Authors:** jerome.golebiowski@unice.fr (J.G.)., hiroaki.matsunami@duke.edu (H.M.).
[#]C.B. and C.A.d.M. contributed equally to this work.

Supporting Information

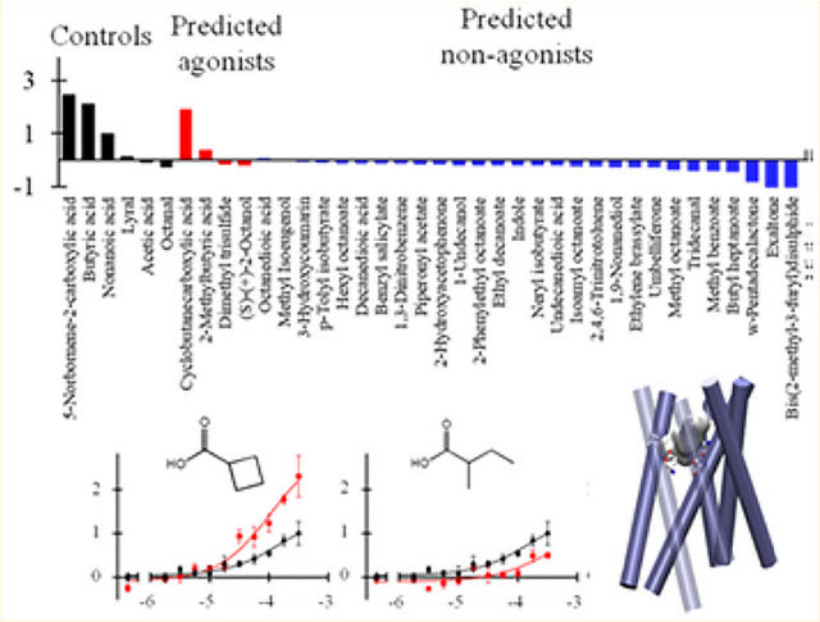
The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpclett.8b00633.

File S1. Experimental data available on OR51E1, OR1A1, OR2W1, and on MOR256–3 (XLSX)

File S2. Molecules used for the screening and category (XLSX)

Experimental methods and QSAR workflow (PDF)

The authors declare no competing financial interest.



Odorant receptor (OR) gene products represent >4% of our proteome.[1] ORs belong to the class A G-protein-coupled receptor (GPCR) family of proteins, which are responsible for transmitting signals across cell membranes. ORs play a crucial and central role in our sense of smell, endowing us with an extraordinary olfactory discrimination ability.[2] They are also ectopically expressed in various organs and modulate non-olfactory functions.[3] Therefore, understanding the chemical space that ORs respond to is not only relevant to understand olfaction but also may have a pharmacological impact.

Upon smelling, odorant compounds reach ORs present at the surface of olfactory sensory neurons. At the molecular level, odorants combinatorially activate a fraction of our 397 types of ORs,[1,4] and the pattern of activation codes for a specific olfactory percept. Thus ORs are very subtle molecular machines, and small modifications in their amino acid sequence can drastically affect their response to chemicals.[5] Accordingly, the relationship between the ligand chemical space and the receptor space is complex and subtle; however, our understanding of how OR characteristics match ligand features remains limited in the field of odor perception as well as more generally in research on GPCRs.

From a pharmacological perspective, odorants are no more than ligands for G-protein-coupled OR. GPCR ligands can be classified as agonists, non-agonists, or antagonists. However, the defining features of GPCR agonists are not well characterized. It is very likely that molecular features of candidate ligands encode their activity, and the challenge for researchers lies in identifying rules that link chemical descriptors to receptor activation. For agonists, the strength of the receptor response (efficacy) is then determined by a finely tuned interaction between the ligand and the receptor, which controls the downstream signaling.[6]

Machine learning may be a useful general approach to explore the chemical space of mammalian ORs and to identify general chemogenomic rules. Indeed, machine-learning approaches have been shown to predict olfactory perception from chemical features of

odorants.[7] Focusing on GPCR drug design, machine learning is considered sufficiently promising to increase the success rate of discovery of novel ligands in the coming years.[8] To date, machine learning mostly processed ligand similarity[9] or a few types of ligand chemical descriptors or protocols.[10–13] Most GPCR studies focused on already known data and discussed the accuracy of machine-learning models. However, the use of an external data set using novel candidates is crucial to assess the real performance of the model, as recently performed on bitter taste receptors.[14]

In this work, we designed a synergistic approach that combines machine learning, in vitro luciferase assays, and homology modeling[15] to (i) identify novel ligands for ORs and (ii) better understand the receptor features that govern agonist efficacy. We chose OR51E1 as the receptor of interest because it is a highly characterized OR. It also has ubiquitous functions in olfactory and non-olfactory tissues.[16] OR51E1 is also known as prostate-specific G-protein-coupled receptor 2 (PSGR2) and is of particular interest because it is involved not only in odor recognition but also in human prostate or lung cancer cell proliferation. Furthermore, this receptor is postulated to be a marker for neuroendocrinic carcinoma cells.[17,18]

Ligand Activity Is Predicted by a Support Vector Machine Algorithm.

All available data regarding OR51E1 were gathered (File S1). To our knowledge, 24 molecules are reported as agonists of OR51E1, while 96 are reported as non-agonists. Agonist and non-agonist spaces cannot easily be discriminated based on simple chemical descriptors (Figure 1a). The 24 known agonists of this receptor have a wide array of chemical functions and structures. About 58% of these are carboxylic acids, and the remaining 42% belong to various chemical families (aldehydes, aromatic cycles with alcohols, esters, and ethers as well as nitrogen- and sulfur-containing molecules). Of the 14 agonists within the carboxylic acid family, only one has a cyclic structure (5-norbornene-2-carboxylic acid), while the other 13 present aliphatic side chains. The carbon chains range from 4 to 14 carbon atoms (butyric acid to tetradecanoic acid). To make sense of such a complex chemical space, we used a machine-learning algorithm to capture the chemical characteristics necessary to identify a molecule as an agonist.

We further defined the odorant molecular space by computing 4884 chemical, topological, and electronic descriptors of 2577 commonly used odorants.[19] Odorants belonging to our library as well as odorants whose activity on OR51E1 has been previously reported (Files S1 and S2) were projected onto this subspace formed by the first three principal components (PCs) that account for 56.6% of variance to examine to what extent they covered the odorant space (Figure 1b and Figure S1). The first and the third molecular PCs were weighted by factors that are reasonable descriptors for the size and composition of a molecule (size of the molecule and presence/absence of oxygen, sulfur, or nitrogen atoms). The second PC was weighted by factors describing either the embranchment or the complexity of the molecule (see the Supporting Information (SI)). In this case, agonist and non-agonist spaces were found to be strongly intertwined, which highlights the need to mine them using machine learning. A quantitative structure activity relationship approach featuring a machine-learning protocol that uses a support vector machine (SVM) was performed.

As shown in Figure 1c, the initial library was made up of 258 compounds, 176 of which had never been tested before on OR51E1. A first reduction of the testable compounds was performed by assessing the applicability domain of the model. The applicability domain encompasses molecules that are considered similar to the training set. This means that the SVM model "learned" from particular structures of the training set, and only similar structures from the untested library can be reasonably screened by the model. Here we used molecular fingerprints and a Tanimoto score to compare molecules from the untested library to the training set (see the SI). In our untested library, 144 molecules that were structurally too dissimilar to those of the learning set were accordingly excluded from virtual screening. The remaining 32 compounds belonging to the applicability domain were virtually screened using the SVM algorithm.

These were split into 4 potential agonists and 28 non-agonists. The predicted agonists were 2-methyl-butyric acid, cyclobutanecarboxylic acid, dimethyl-trisulfur, and (*S*)-(+)-2-octanol. Functional luciferase in vitro assays were used to assess the activity of these 32 compounds as well as 44 compounds initially excluded from the applicability domain for quality-control reasons (Figure 1c). A total of 76 molecules were screened at 100 μM on OR51E1 and compared with a set of three positive controls and three negative controls (Figure 2a). None of the molecules outside of the applicability domain nor those predicted as non-agonists triggered a receptor response. However, of the four potential agonists, two activated the receptor and two did not. Cyclobutanecarboxylic acid and 2-methyl butyric acid exhibited an agonist dose-dependent behavior in vitro, while the two other predicted agonists (dimethyl-trisulfur and (*S*)-(+)-2-octanol) were false-positives (Figure 2b and Table S1). Thus our model accurately captured the activity of 30 compounds out of the 32 belonging to the applicability domain. It shows an in vitro hit rate of 50% (two true agonists out of four predicted agonists) and a reliability of 94% (30 correct predictions out of 32).

The transferability of the protocol was assessed by applying it to two other human receptors (OR1A1, OR2W1) and a mouse receptor (MOR256–3), for which many agonists and non-agonists are reported (see File S1). For each system, the SVM algorithm processed the part of the untested library belonging to the applicability domain. Then, in vitro assays were used to compare their ability to trigger OR response to those of control agonists. Table 1 reports in vitro hit rates for the four systems studied. Novel agonists were predicted with hit rate in the range of [39%, 50%]. For OR1A1, OR2W1, and MOR256–3, in vitro data assessed 7, 2, and 5 novel agonists, respectively (Figures S2 and S3 and File S1). This suggests that machine-learning approaches are useful to explore the wide chemical space associated with ORs and potentially other GPCRs if a sufficient number of agonists and non-agonists is known. Here the smallest database used for training a supervised model concerned OR51E1, consisting of 24 agonists and 24 non-agonists.

Focusing on OR51E1, agonists show a large range of activation strength. Compared with nonanoic acid, a known agonist of OR51E1,[20,21] the two novel agonists showed comparable potencies (see dose-response curves in Figure 2b) but different efficacies. Namely, cyclobutanecarboxylic acid triggered a response 230% higher than nonanoic acid, while 2-methyl-butyric acid had an efficacy that was 50% lower (Figure 2b). These variable strengths of the receptor responses suggest differential interactions and affinities between

ligands and the receptor cavity. We further investigated the influence of the receptor cavity features on ligand activity using a 3D model.

Agonist Efficacy Is Affected by the Receptor Cavity.

From a structural point of view, ORs, like class A GPCRs, form a bundle within the cell membrane and are composed of seven helices named TM1 to TM7. The canonical ligand binding site is located 10 Å below the extracellular side of the receptor. The largest sequence variability between ORs lies at this binding site to potentially endow our olfactory system with a large discrimination of the chemical space.[22,23] The OR51E1 3D model was built based on X-ray class A GPCR templates and showed a binding cavity with a solvent-accessible surface area of ~300 $Å^3$, ~70% of which is apolar. Consistently, the whole set of 26 agonists of OR51E1 encompasses ligands with variable bulkiness (ranging from 90 to 360 $Å^3$) and a relatively high lipophilicity (logP ranging from −0.8 to 4.7).

In the present study, nonanoic acid and cyclobutanecarboxylic acid, whose efficacies were found to differ by two orders of magnitude, shared the same binding mode. They were both predicted to bind the receptor through their carboxylic moiety at the cradle of the cavity (Figure 3a,b). The acidic function is in contact with the so-called toggle switch involved in receptor activation, $Y254^{6.48}$ (6.48 refers to the Ballesteros-Weinstein notation[24]). The ligands bind between TM3 ($S111^{3.36}$, $H108^{3.33}$), TM5 ($I206^{5.44}$), and TM6 ($Y254^{6.48}$), which is in line with typical binding modes observed in X-ray structures of agonists bound to β2-adrenergic and opsin receptors.[25,26]

More specifically, residues H108 and I206 were found to be differentially in contact with the two agonists, while S111 and Y254 were found to interact with the carboxylic acid moiety. This suggests that H108 and I206 control the bulkiness of the agonist. Mutant ORs at these positions consistently showed differential efficacy modulation in vitro (Figure 3c). The efficacy of nonanoic acid increased by more than two orders of magnitude when the receptor was reprogrammed with smaller residues at the top of its cavity (H108A and I206A in Figure 3d). These mutants both showed a nonsignificant effect on the much smaller cyclobutanecarboxylic acid.

When H108 was mutated into a ~25% bulkier and less hydrophilic residue (H108F), the response of the receptor was totally abolished for both agonists, which is likely attributed to a complete blockage of the binding cavity (Figure S4a). We show that these residues are located at the upper part of the cavity and modulate its size. They control the accessibility of agonists within the binding cavity by sensing the size of their hydrophobic moiety. This accessibility then modulates the efficacy of the receptor response.

On the contrary, mutations of residue S111 similarly modulated the response to the two different agonists; this is suggestive of a role of S111 in acid moiety recognition, which triggers the activation mechanism once the ligand is bound (Figure S4b,c).

Consistent with its low efficacy, the energy-minimized docked structure of 2-methylbutyric acid suggested that the ligand is closer to H108 and I206 compared with

cyclobutanecarboxylic acid (Figure S5). As for nonanoic acid, such tighter contacts with the top of the binding cavity would prevent these agonists from having a large efficacy.

In conclusion, we used supervised learning algorithm (SVM) to expand the chemical space of OR51E1, whereby we virtually screened a database of 176 compounds. After assessment by in vitro luciferase assays, we predicted 30 out of 32 ligand activities and revealed two novel agonists of OR51E1 in vitro, which were carboxylic acids. It is interesting to note that the structures of the acid agonists reported for this receptor are very different; some are short-chained aliphatic acids (such as butyric acid), while others exhibit cyclic structures (such as 5-norbornene-2-carboxylic acid). However, OR51E1 cannot be considered simply as a carboxylic-acid-sensing receptor. Of the 45 acids that were tested on OR51E1, 64.5% were non-agonists. Our machine-learning model was able to accurately learn from this complex set of known agonist and non-agonist structures and successfully predicted two novel ligands for this receptor. The applicability to other ORs was assessed through the use of a similar protocol on OR1A1, OR2W1, and MOR256–3. The models predicted the in vitro activity with a hit rate ranging from 39 to 50% and identified novel agonists in all cases. This suggests that machine learning is capable of predicting the activity of candidate compounds using their chemical features. Note that these machine-learning models are limited to their applicability domain, which naturally restricts the chemical space research area. Other agonists present outside this domain could be discovered using molecular modeling to predict ligand activity based on receptor activation dynamics.[27]

Furthermore, to understand the link between the strength of a response and receptor features, a 3D model of OR51E1 was built. We showed how some residues in the cavity can selectively affect efficacy. On the basis of the 3D model, we successfully reprogrammed OR51E1 *wt*, which confirmed that residues at the top of the orthosteric binding cavity control agonist efficacy. For these two mutants, nonanoic acid was specifically transformed into a highly efficacious agonist compared with the OR51E1 *wt*. This emphasizes the crucial role of some residues in modulating the strength of a receptor response. These residues belonging to TM3 and TM5 are well documented as being involved in ligand recognition in ORs.[5]

Ligand-based machine-learning approaches can be used to enhance our knowledge of the chemogenomic links between the vast space of odorants, made up of millions of molecules, and that of receptors, comprising 397 functional genes in humans. Olfactory sensing is still in its infancy, and such approaches will help widen the chemical space associated with ORs, for which few agonists and non-agonists are reported. By integrating descriptors for the receptor features, one can envision producing numerical SVM models able to reliably predict not only the activity of candidate ligands but also their efficacy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank Dr. X. Cong for critical reading of the manuscript. This research was supported by grants from the National Institute on Deafness and Other Communication Disorders, National Institute of Health grants DC014423 and DC012095, National Science Foundation (NSF) grant 1515801 (to H.M.), the Agence Nationale de la Recherche as part of NSF/NIH/ANR Collaborative Research in Computational Neuroscience, a UCA IDEX grant (to J.G.), and GIRACT, GEN foundation, and Fondation Roudnitska under the aegis of Fondation deFrance grants (to C.B.).

REFERENCES

(1). Niimura Y Olfactory receptor multigene family in vertebrates: from the viewpoint of evolutionary genomics. Curr. Genomics 2012, 13 (2), 103–14. [PubMed: 23024602]

(2). Bushdid C; Magnasco MO; Vosshall LB; Keller A Humans can discriminate more than 1 trillion olfactory stimuli. Science 2014, 343 (6177), 1370–2. [PubMed: 24653035]

(3). Kang N; Koo J Olfactory receptors in non-chemosensory tissues. BMB Rep. 2012, 45 (11), 612–22. [PubMed: 23186999]

(4). Malnic B; Hirono J; Sato T; Buck LB Combinatorial receptor codes for odors. Cell 1999, 96 (5), 713–23. [PubMed: 10089886]

(5). de March CA; Yu Y; Ni MJ; Adipietro KA; Matsunami H; Ma M; Golebiowski J Conserved Residues Control Activation of Mammalian G Protein-Coupled Odorant Receptors. J. Am. Chem. Soc 2015, 137 (26), 8611–6. [PubMed: 26090619]

(6). Yao XJ; Velez Ruiz G; Whorton MR; Rasmussen SG; DeVree BT; Deupi X; Sunahara RK; Kobilka B The effect of ligand efficacy on the formation and stability of a GPCR-G protein complex. Proc. Natl. Acad. Sci. U. S. A 2009, 106 (23), 9501–6. [PubMed: 19470481]

(7). Keller A; Gerkin RC; Guan Y; Dhurandhar A; Turu G; Szalai B; Mainland JD; Ihara Y; Yu CW; Wolfinger R; Vens C; Schietgat L; De Grave K; Norel R; Stolovitzky G; Cecchi GA; Vosshall LB; Meyer P Predicting human olfactory perception from chemical features of odor molecules. Science 2017, 355 (6327), 820–826. [PubMed: 28219971]

(8). Wacker D; Stevens RC; Roth BL How Ligands Illuminate GPCR Molecular Pharmacology. Cell 2017, 170 (3), 414–427. [PubMed: 28753422]

(9). Boyle SM; McInally S; Ray A Expanding the olfactory code by in silico decoding of odor-receptor chemical space. eLife 2013, 2, e01120. [PubMed: 24137542]

(10). Mansouri K; Judson RS In Silico Study of In Vitro GPCR Assays by QSAR Modeling. Methods Mol. Biol 2016, 1425, 361–81. [PubMed: 27311474]

(11). Shemesh R; Toporik A; Levine Z; Hecht I; Rotman G; Wool A; Dahary D; Gofer E; Kliger Y; Soffer MA; Rosenberg A; Eshel D; Cohen Y Discovery and validation of novel peptide agonists for G-protein-coupled receptors. J. Biol. Chem 2008, 283 (50), 34643–9. [PubMed: 18854305]

(12). Wu J; Zhang Q; Wu W; Pang T; Hu H; Chan WKB; Ke X; Zhang Y WDL-RF: Predicting Bioactivities of Ligand Molecules Acting with G Protein-coupled Receptors by Combining Weighted Deep Learning and Random Forest. Bioinformatics 2018, DOI: 10.1093/bioinformatics/bty070.

(13). Schmuker M; de Bruyne M; Hahnel M; Schneider G Predicting olfactory receptor neuron responses from odorant structure. Chem. Cent. J 2007, 1, 11. [PubMed: 17880742]

(14). Huang W; Shen Q; Su X; Ji M; Liu X; Chen Y; Lu S; Zhuang H; Zhang J BitterX: a tool for understanding bitter taste in humans. Sci. Rep 2016, 6, 23450. [PubMed: 27040075]

(15). Bushdid C; de March CA; Matsunami H; Golebiowski J Numerical models and in vitro assays to study Odorant Receptors. Methods Mol. Biol 2018, 1820, 10.1007/978-1-4939-8609-5.

(16). Flegel C; Manteniotis S; Osthold S; Hatt H; Gisselmann G Expression profile of ectopic olfactory receptors determined by deep sequencing. PLoS One 2013, 8 (2), e55368. [PubMed: 23405139]

(17). Leja J; Essaghir A; Essand M; Wester K; Oberg K; Totterman TH; Lloyd R; Vasmatzis G; Demoulin JB; Giandomenico V Novel markers for enterochromaffin cells and gastrointestinal neuroendocrine carcinomas. Mod. Pathol 2009, 22 (2), 261–72. [PubMed: 18953328]

(18). Weng J; Wang J; Hu X; Wang F; Ittmann M; Liu M PSGR2, a novel G-protein coupled receptor, is overexpressed in human prostate cancer. Int. J. Cancer 2006, 118 (6), 1471–80. [PubMed: 16206286]

(19). Mainland JD; Keller A; Li YR; Zhou T; Trimmer C; Snyder LL; Moberly AH; Adipietro KA; Liu WL; Zhuang H; Zhan S; Lee SS; Lin A; Matsunami H The missense of smell: functional variability in the human odorant receptor repertoire. Nat. Neurosci 2014, 17 (1), 114–20. [PubMed: 24316890]

(20). Saito H; Chi Q; Zhuang H; Matsunami H; Mainland JD Odor coding by a Mammalian receptor repertoire. Sci. Signaling 2009, 2 (60), ra9.

(21). Audouze K; Tromelin A; Le Bon AM; Belloir C; Petersen RK; Kristiansen K; Brunak S; Taboureau O Identification of odorant-receptor interactions by global mapping of the human odorome. PLoS One 2014, 9 (4), e93037. [PubMed: 24695519]

(22). Man O; Gilad Y; Lancet D Prediction of the odorant binding site of olfactory receptor proteins by human-mouse comparisons. Protein Sci. 2004, 13 (1), 240–54. [PubMed: 14691239]

(23). de March CA; Kim SK; Antonczak S; Goddard WA 3rd; Golebiowski J G protein-coupled odorant receptors: From sequence to structure. Protein Sci. 2015, 24 (9), 1543–8. [PubMed: 26044705]

(24). Ballesteros JA; Weinstein H Integrated Methods for the Construction of Three-Dimensional Models and Computational Probing of Structure-Function Relations in G Protein-Coupled Receptors. Methods Neurosci. 1995, 25, 366–428.

(25). Venkatakrishnan AJ; Deupi X; Lebon G; Tate CG; Schertler GF; Babu MM Molecular signatures of G-protein-coupled receptors. Nature 2013, 494 (7436), 185–94. [PubMed: 23407534]

(26). Park JH; Morizumi T; Li Y; Hong JE; Pai EF; Hofmann KP; Choe H-W; Ernst OP Opsin, a Structural Model for Olfactory Receptors? Angew. Chem. Int. Ed 2013, 52 (42), 11021–11024.

(27). de March CA; Topin J; Bruguera E; Novikov G; Ikegami K; Matsunami H; Golebiowski J Odorant Receptor 7D4 Activation Dynamics. Angew. Chem. Int. Ed 2018, 57 (42), 4554–4558.

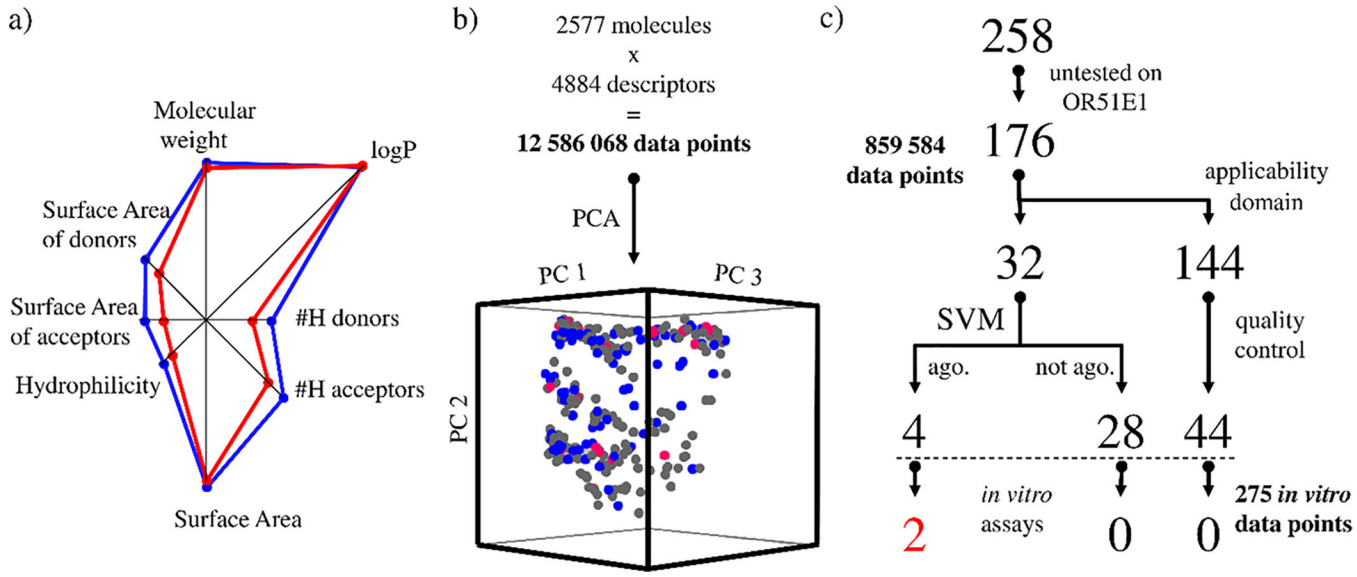


Figure 1.

Odorant molecular space and protocol used for OR51E1 virtual screening. (a) Spider plot representing the weight of eight attributes (molecular weight, logP, number of hydrogen-bond donor and acceptors, total surface area, hydrophilicity, and surface areas of acceptor and donor atoms) to define OR51E1 agonists (red) and non-agonists (blue), as identified prior to this study. (b) Projection of our library of 176 original odorants, shown in gray, into the three first principal components of the odorant space computed on the basis of chemical descriptors of 2577 odorant compounds. Known agonists and non-agonists are shown in red and blue, respectively. (c) Of our library of 258 odorants, 176 had been untested on OR51E1, and 32 belonged to the same chemical space (applicability domain) of our model. The virtual screening predicted four agonists, two of which elicited a receptor response in vitro. All predicted non-agonists were tested in vitro for quality-control purposes, as were 44 randomly selected compounds initially excluded from the library; none triggered a receptor response in vitro.

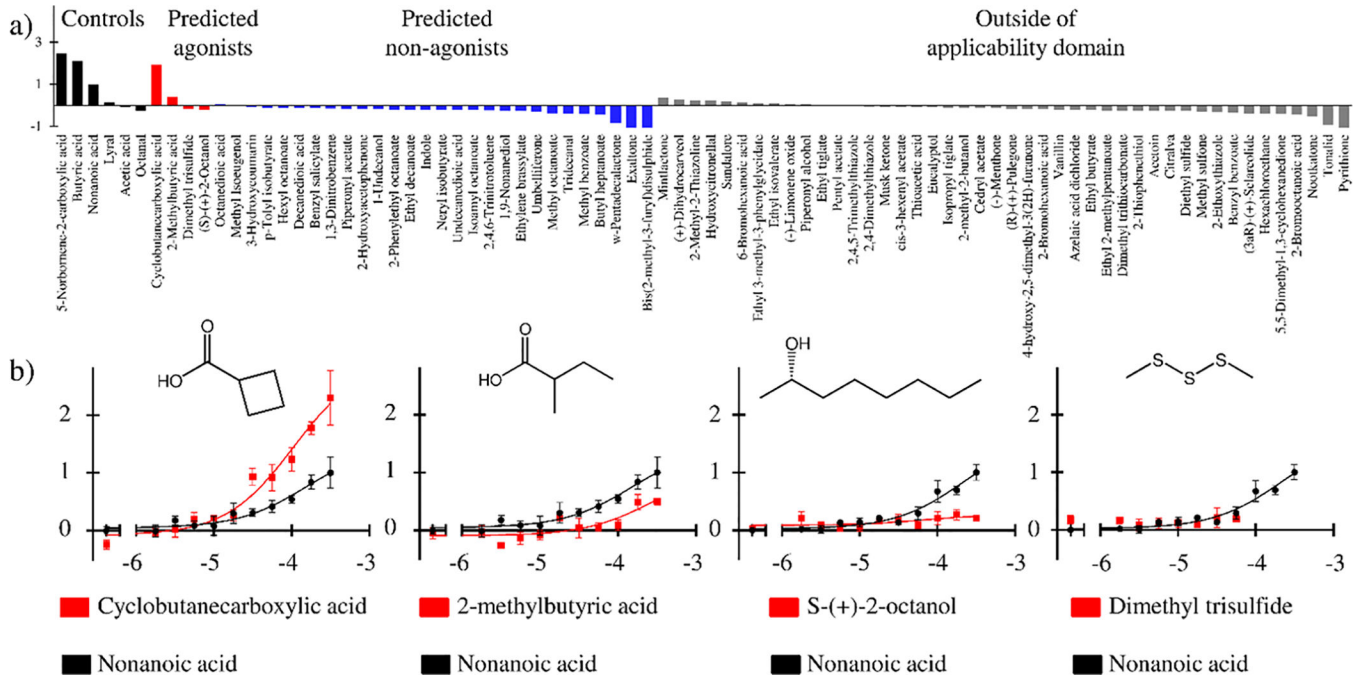


Figure 2.

In vitro screening of OR51E1. (a) Response of OR51E1 to six controls (three non-agonists and three agonists, in black), 4 predicted agonists by the SVM model (in red), 28 predicted non-agonists by the SVM model (in blue), and 44 molecules not within the applicability domain (in gray). Odorants were injected at 100 μM, and each response of OR51E1 was corrected by that of the empty vector (pCI) response and by basal activity of receptors. It should be noted that negative responses do not reflect inhibition of the response but rather cell toxicity. (b) Dose-response curves for the four potential agonists as identified by the SVM model. Cyclobutanecarboxylic acid and 2-methylbutyric acid were found to activate the receptor, while *S*-(+)-octanol and dimethyl trisulfide did not. All responses were normalized to that of nonanoic acid.

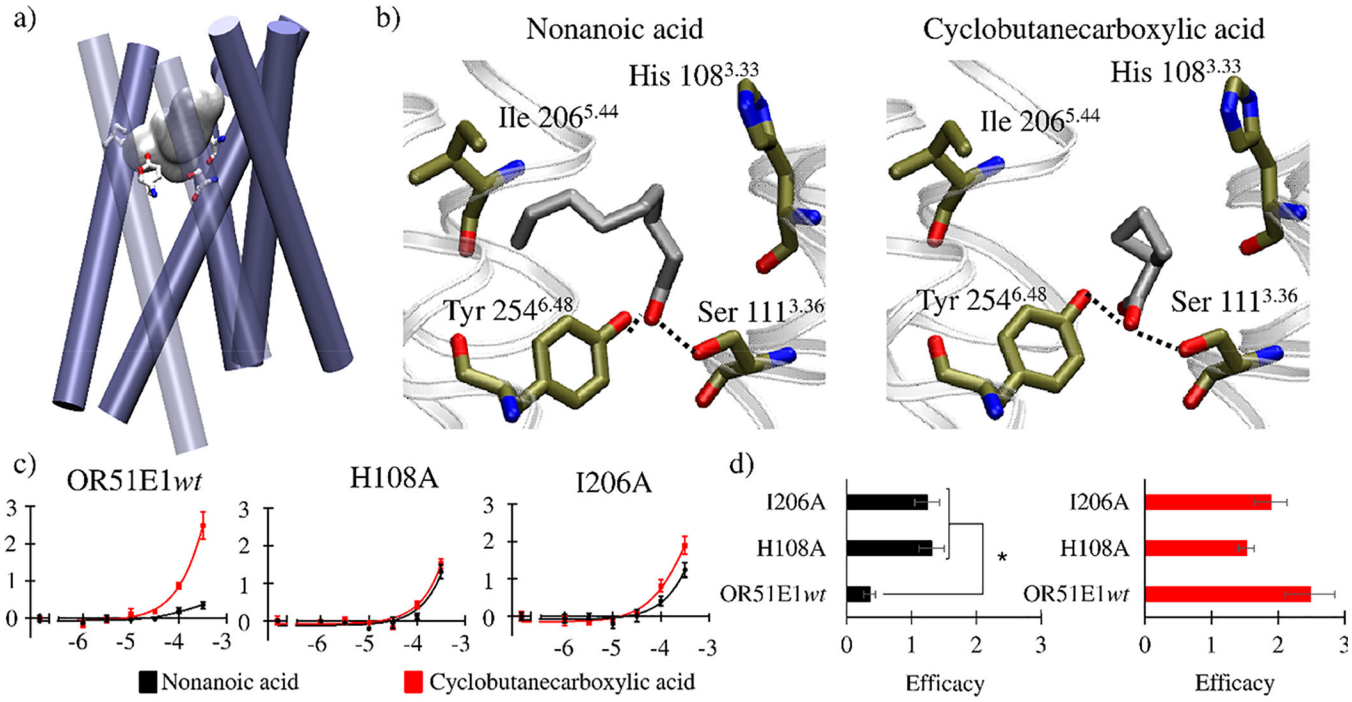


Figure 3.

Some OR51E1 cavity residues control the efficacy of bound agonists. (a) Overview of OR51E1 wild type (OR51E1 *wt*) bundle structure and binding cavity (white volume). Unfolded loops are omitted for clarity. (b) Binding mode of nonanoic acid (left) and cyclobutanecarboxylic acid (right) in the binding cradle of OR51E1 *wt*. Carbon atoms are shown in gold (or gray in the case of the ligand), oxygen atoms in red, and nitrogen atoms in blue. Hydrogen-bond contacts are shown as dashed lines. (c) Dose-response curves of OR51E1 *wt* and OR51E1 mutants to nonanoic acid (black) and cyclobutanecarboxylic acid (red). (d) Efficacies of OR51E1 *wt* and OR51E1 mutants to nonanoic acid (black) and cyclobutanecarboxylic acid (red). Error bars indicate the SEM (*$p < 0.01$).

Table 1.

Performance and Transferability of the SVM Model[a]

	OR51E1	OR1A1	OR2W1	MOR256–3
in vitro/SVM	2/4	7/18	2/5	5/10
hit rate	50%	39%	40%	50%

[a]Numbers of in-vitro-validated and SVM-predicted agonists and corresponding hit rate for the four systems studied.