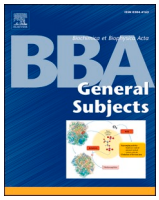
BBA - General Subjects 1869 (2025) 130825

Contents lists available at ScienceDirect

BBA - General Subjects

journal homepage: www.elsevier.com/locate/bbagen



Computational modelling of olfactory receptors

Chiemela S. Odoemelam *, Volker Steuber, Michael Schmuker

UH Biocomputation Research Group, Centre for AI and Robotics Research, University of Hertfordshire, Hatfield, United Kingdom

ARTICLE INFO

Keywords:
Olfactory receptors
Odorant molecules
Computational modelling

ABSTRACT

Olfactory receptors (ORs), the largest subfamily of G protein-coupled receptors, are essential for detecting and interpreting environmental odorants in animals. Understanding their function is crucial for deciphering olfactory perception and exploring emerging roles in non-olfactory systems. With the recent surge in available sequence data and AI-based structural predictions, computational modelling has become indispensable for investigating OR structure, ligand binding, and activation mechanisms. This review provides a comprehensive overview of computational approaches used in OR research, including homology modelling, molecular docking, molecular dynamics simulations, free energy calculations, pharmacophore modelling, virtual screening, and machine learning-based predictions. Both ligand-based and structure-based pharmacophore modelling are discussed in detail, highlighting their respective applications, strengths, and limitations. While structure-based approaches have gained prominence due to advances in receptor structure prediction tools like AlphaFold, ligand-based pharmacophore modelling remains valuable in scenarios where structural data are limited or uncertain. Case studies illustrate how these techniques have been applied to identify novel OR–ligand interactions, explore receptor dynamics, and support drug discovery. Collectively, these computational strategies offer powerful tools for decoding OR function, guiding experimental validation, and expanding our understanding of olfactory signalling in health and disease.

1. Introduction

Olfaction (or sense of smell) is one of animals' most ancient and essential sensory systems [1–4]. Animals have an olfactory system that allows them to detect, encode and interpret odour molecules within their environment [2]. Olfaction also describes the chemosensory process of sensing low amounts of airborne, volatile chemical compounds in terrestrial vertebrates and insects [3]. Fish and aquatic crustaceans, however, do not encounter airborne, volatile chemical substances. However, their olfactory systems resemble those of animals that live on land [3]. In these marine animals, the olfactory system serves critical roles in detecting food, avoiding predators, navigating environments, recognising mates and kin, and mediating social and reproductive behaviours, all through detecting waterborne chemical cues processed by specialised sensory structures [5–8]. More broadly, chemical senses play an important role in invertebrates and many vertebrates, essential for communication between individuals, detecting food and threats, and identifying dangerous substances [9–12].

The olfactory receptors (ORs) – the largest G protein-coupled receptor (GPCR) family- comprise sensory proteins that aid in recognising various environmental odorants [1,13]. About 380–400 ORs exist in humans [1,14–17], 1000 in mice [1,14,15,18,19], in cows, and in dogs, about 800–1000 OR exists [20]. In vertebrates, ORs are primarily situated in the epithelium of the nasal cavity to detect smell, and they have also been identified in other non-nasal tissues/organs such as the gut, blood vessels, airways, prostate and other organs [19,21–25]. These ORs in non-chemosensory organs are called ectopic ORs (eORs) [19,26]. Ectopic ORs might respond to compounds that are different from odorants. For instance, Cheng et al. demonstrated that a lipidated synthetic pepducin, o109-i2–2, modelled after the second intracellular loop of Olfr109, acts as an intracellular allosteric antagonist. Rather than engaging the extracellular ligand-binding site, this pepducin interferes with Olfr109-mediated signalling by modulating its interaction with G proteins and β-arrestin [27]. The study further reveals that administration of pepducin significantly ameliorates glucose metabolism disorders in high-fat diet (HFD)-induced obese mice and Akita mice, indicative of its potential as a therapeutic agent [27]. Notably, the antidiabetic effects observed were nullified in mice lacking the Olfr109 receptor, highlighting the role of this receptor in mediating the observed therapeutic outcomes [27]. These findings underscore the promising therapeutic

* Corresponding author.
E-mail address: c.odoemelam@herts.ac.uk (C.S. Odoemelam).

https://doi.org/10.1016/j.bbagen.2025.130825
Received 10 March 2025; Received in revised form 8 May 2025; Accepted 22 May 2025
Available online 27 May 2025

utility of Olfr109-derived pepducin in addressing obesity and diabetes [27,28].

Other non-odorant ligands that can be detected by ectopically expressed ORs include endogenous metabolites and synthetic compounds, mediating diverse physiological functions beyond olfaction. Notably, OR51E1, expressed in the heart and prostate, responds to medium-chain fatty acids (nonanoic acid and decanoic acid), modulating cardiac contractility, while 2-ethylhexanoic acid antagonises this effect. In acute myeloid leukaemia, OR2AT4 and OR51B5 are activated by synthetic ligands such as sandalore and isononyl alcohol, inducing apoptosis and inhibiting cell proliferation. In the liver, terpenes like (−)-carvone and (−)-citronellal activate OR1A1 and OR1A2, influencing lipid metabolism and cancer cell growth. These examples underscore the potential of extra-nasal ORs as therapeutic targets and diagnostic biomarkers in various organ systems [29]. Although the human olfactory system is weaker than many other animals, it remains exact and functionally sophisticated [30,31], including the ability to scent-track [32]. It can detect and discriminate among numerous odorants, including those at low concentrations [30,31]. This capability is attributed mainly to the diverse amino acid sequences of human olfactory receptors, which enable the recognition of structurally distinct odorants [33].

Given the structural diversity of olfactory receptors and the scarcity of experimentally resolved OR structures, computational modelling has become an indispensable tool for understanding OR function. Techniques such as homology modelling, molecular docking, molecular dynamics simulations, and, more recently, AI-based structure prediction methods like AlphaFold have been widely used to predict OR 3D structures, explore ligand-binding interactions, and investigate receptor activation mechanisms. These approaches facilitate receptor deorphanization, guide mutagenesis studies, and support drug discovery efforts targeting nasal and ectopically expressed ORs. This review explores the current landscape of computational strategies applied to OR research, highlighting their capabilities, challenges, and future directions.

2. Olfactory receptor structure and function

The structure of ORs is quite intricate and characterised by seven membrane-spanning domains (seven transmembrane domains, 7 TM) [13,19,34]. The transmembrane (TM) domains are connected by three extracellular (ECL) and intracellular (ICL) loops, respectively, an intracellular C-terminus and an extracellular N-terminus [30]. ORs have unique sequence features (Fig. 1) [30]; LHTPMY in intracellular loop 1 (IC1) [35], EF(I/L)LLG(L/F) upstream of TM1 [36], PMYFFL (TM2) [36], three C (ECL2), SY at the end of TM5 and beginning of ICL3 [35,36], KAFSTCASH at the starting of TM6 [35] and PMLNPFIY in TM7 [35,36].

Additionally, within the intracellular loop, a conserved sequence motif aspartate-arginine-tyrosine (DRY) is located at the junction of TM3 and ICL2. This motif, often referred to as $D^{3.50}$-$R.^{3.51}$-$Y^{3.52}$ in the Ballesteros-Weinstein numbering system, is a hallmark of class A GPCRs, and plays a central role in G protein coupling and activation [1,30,45–47]. The Ballesteros-Weinstein system assigns a general reference framework by denoting the most conserved residue in each helix as position x.50, where x is the helix number, allowing consistent cross-comparisons between structurally homologous receptors [48].

The olfactory sensory neurons (OSNs) within the olfactory epithelium (OE) are arranged in overlapping zones based on the expression of individual olfactory receptors (ORs) [4,49,50]. This arrangement follows a pseudostratified pattern, with basal layer progenitor cells and upper layers composed of mature neurons [4]. The mature OSN dendrites project into the nasal cavity, forming a dendritic knob at the surface of the OE [4,51]. From this knob, 5–10 long, slender cilia extend into the nasal mucus, which is approximately 60 μm thick [30,45]. These cilia serve as chemosensory structures that detect odours [30,45,52]. It is important to note that ORs are localised on the membranes of these cilia, where they interact directly with odorant molecules present in the mucus [4,51]. Mature OSNs that express the same OR gene project their axons to the olfactory bulb (OB), an extension of the brain just above the cribriform plate [4,30,45]. The axons of OSNs synapse on specific glomeruli, spherical structures in the OB dedicated to processing information from a particular odorant [4,45,51].

Each OSN expresses only one type of OR [16,53]. This implies that the activation of a specific OR solely determines the activation of an OSN [16,53]. A specific OR can bind to several odorant molecules; conversely, a single odorant can activate multiple ORs [16,30,31,54,55]. When an odorant molecule binds to an OR, it triggers a cascade of biochemical reactions inside the OSN [16,30]. This cascade of reactions ultimately causes the OSN to depolarise and generates an action potential [30]. The neuronal action potential is transmitted to the olfactory bulb, which is a relay station, forwarding the information to various brain regions for further processing [30].

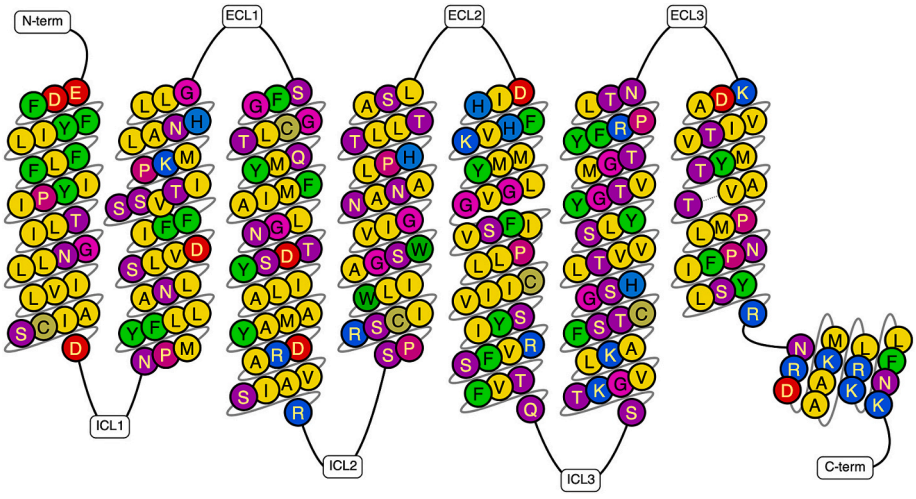


Fig. 1. Schematic showing the transmembrane topology of the human olfactory receptor OR1A1, generated using GPCRdb Tools, https://gpcrdb.org/ [37–44]. The receptor comprises seven transmembrane domains (TM1–TM7), connected by three extracellular (ECL1–ECL3) and three intracellular loops (ICL1–ICL3), with an extracellular N-terminus and intracellular C-terminus. Residues are colour-coded by physicochemical properties: green (polar uncharged), yellow (hydrophobic/aliphatic), pink (aromatic), red (acidic), blue (basic), and grey (cysteine/proline). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Computational approaches in olfactory receptor modelling

Computational methodologies employed in olfactory receptor modelling facilitate the investigation of olfactory receptors' structural, functional, and dynamic aspects. These approaches encompass diverse capabilities, including the prediction of olfactory receptor structures based on their amino acid sequences, identification of critical residues involved in ligand binding, estimation of ligand binding affinities for different olfactory receptors, and examination of the real-time dynamics of olfactory receptors. Computational approaches in modelling olfactory receptors and their response to odorant molecules that we discuss here are:

- Homology modelling, where the structure of a protein is predicted by aligning it to a closely related one where the structure is known.
- Pharmacophore modelling, where potential interaction points between ligand and receptor are abstracted to predict the binding affinity of untested ligands.
- Molecular dynamics simulations, where protein structures are modelled over time, subject to temperature-based fluctuations in their structure, governed by pre-set interactions between atoms.
- Free energy calculations estimating binding energies.
- Molecular docking, where ligand-receptor interactions are modelled by fitting ligands into binding pockets, estimating their interaction.
- Virtual screening uses computational techniques to search for potential olfactory receptor ligands in odorant databases, thus reducing large libraries to manageable sizes. Structure-based virtual screening consists of consecutive computational processes, such as target and database preparation, docking and post-docking analysis, and compound selection for experimental testing [56].
- Machine learning in human olfactory research comprises of the study of the physiology of pattern-based odour detection and recognition processes, 2) pattern recognition in olfactory phenotypes, 3) the creation of complex disease biomarkers that include olfactory features, 4) odour prediction from physico-chemical properties of volatile molecules like electronic noses, and 5) knowledge discovery in publicly available big databases [57]. Machine learning also plays an increasingly important role in the approaches above.

4. Homology modelling

Structural biology has significantly advanced our understanding of protein structure. It enables us to predict how a protein's primary sequence is arranged into secondary structures such as alpha helices and beta sheets. It identifies common motifs, domains, and folds [58,59] and provides insights into protein subunits' tertiary and quaternary organisation [58].

Protein sequences are crucial in determining their structure and function [58]. Proteins are inherently dynamic entities, exhibiting significant flexibility and structural adaptability within their drug-binding sites [60]. However, a single structural model, whether obtained experimentally or predicted, represents only a static snapshot and may not capture the full range of a protein's relevant characteristics [60].

Homology modelling, a technique used in computational biology, involves constructing a three-dimensional (3D) structure of a protein of interest based on the known 3D structure of a related protein [58,61]. This method is employed when the crystal structure of the protein of interest is unavailable [58,61]. The underlying principle of homology modelling is that evolutionarily related proteins tend to retain a conserved three-dimensional fold, particularly in their structurally stable core regions, despite substantial variation in their amino acid sequences [60]. This structural conservation arises from shared ancestry, rather than strict membership in the same protein family. Importantly, structural cores are more evolutionarily conserved than amino acid sequences, enabling reliable structural predictions even at relatively low sequence identity [62]. Homology modelling, therefore, relies on experimentally determined structures of homologous proteins (templates) to generate accurate models for related protein sequences (targets) [58,60,61].

Homology modelling yields the most accurate models when using closely related homologous structures [63]. However, suitable models can still be generated even with relatively low sequence similarity (approximately 20 %) [60]. Methods used in homology modelling are typically classified as traditional homology (template-based) or de novo modelling [60]. To assess the accuracy of three-dimensional structure predictions, computational methods such as de novo prediction and threading are compared to homology modelling using criteria such as root-mean-square deviation (RMSD) [63]. Homology modelling has been found to produce 3D structures with the highest accuracy compared to other methods [63,64]. Moreover, it is a time- and cost-efficient approach with well-defined steps [63].

Consequently, traditional homology modelling (or comparative modelling) is considered the most accurate computational method for predicting protein 3D structures [63]. As depicted in Fig. 2, homology modelling commences with selecting a suitable template protein with a known 3D structure [63]. The amino acid sequence of the target protein is subsequently aligned with the template sequence, and the structural information from the template is transferred to generate a model of the target protein [63]. The key steps involved in homology modelling are summarised in Fig. 2, while a compilation of commonly employed servers and tools for protein structure homology modelling is provided in Table 1.

The enduring challenges faced in homology modelling, such as weak sequence–structure similarities, the alignment of sequences with structures, modelling of rigid body shifts, distortions, loops, side chains, and the detection of errors in a model, persist to this day [84]. Reduced accuracy and the potential for producing incorrect models represent ongoing hurdles in this field [84]. Alignment errors remain a primary cause of deviations, necessitating meticulous manual inspection and adjustment, even when employing fully automated programs [84–87]. Choosing a suitable template, inaccurate alignments, and inefficient refinement methods remain the primary sources of errors in the homology modelling [88,89]. In the case of ORs, which are membrane-bound GPCRs, these challenges are compounded by the inherent complexity of transmembrane domain modelling, limited availability of experimentally determined templates, and the dynamic nature of these receptors. ORs are notoriously challenging to crystallise due to low expression levels and high instability, making structural data scarce. While recent advances such as AlphaFold have significantly improved predictive modelling, they often favour inactive conformations and do not fully account for the membrane environment critical to GPCR function. Consequently, hybrid approaches incorporating structurally related mouse OR templates or experimentally resolved GPCRs have emerged as a practical workaround to improve model quality and biological relevance [58,60,90].

Homology modelling of olfactory receptors presents significant challenges due to the need for experimentally solved structures available for modelling purposes. To establish a dependable homology model for Olfr73, Yuan et al. (2019) conducted a comprehensive analysis. Their approach involved comparing the sequence of Olfr73 with sequences of other class A GPCRs within the Protein Data Bank. Class A GPCRs (Rhodopsin-like receptors) are the largest and most studied GPCR subfamily. They are defined by a seven-transmembrane helical structure and conserved motifs such as the DRY motif at the cytoplasmic end of TM3, essential for receptor activation and G protein coupling [91]. Their findings revealed that the highest sequence identity achieved was 19 % compared to the beta-2-adrenergic receptor (β_2AR, PDB code: 4LDE) and 16 % compared with rhodopsin (RHO, PDB code: 4BEY). To construct their model, they used the crystal structures of both β_2AR and RHO as templates [92].

More recent work by Yu et al. (2022) further illustrates the importance of template selection and curated alignments in homology

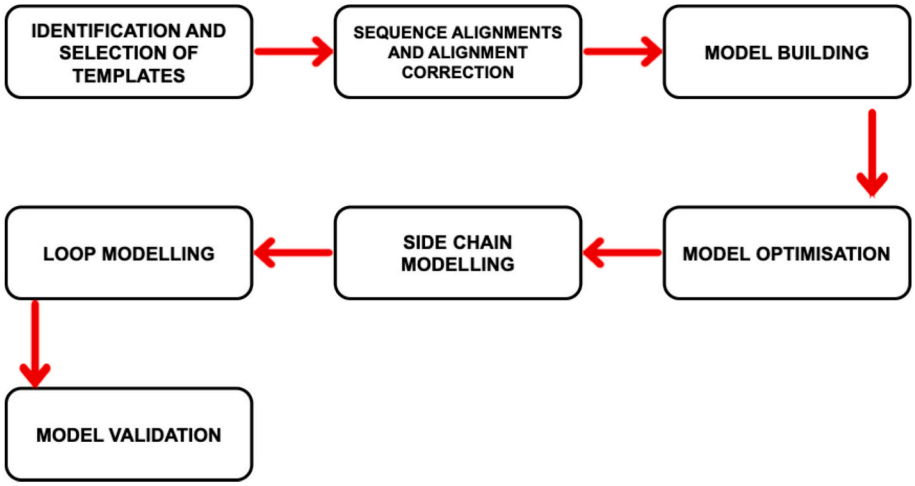


Fig. 2. Schematic representation of the typical workflow for homology modelling. The process includes template identification, sequence alignment, model building, optimisation, loop and side chain modelling, and model validation. Adopted from E. Aki-Yalcin, M.Tilahun Muhammed (2019) [63].

Table 1
Representative tools and servers for protein structure prediction and homology modelling.

Tool	URL	Ref.
MODELLER	https://www.salilab.org/modeller/	[65]
Phyre2	http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index	[66]
SWISS-MODEL	https://swissmodel.expasy.org/	[67,68]
Alphafold	https://alphafold.ebi.ac.uk/	[69]
RoseTTAFold	https://robetta.bakerlab.org/submit.php	[70,71]
RoseTTAFold Joint (RFjoint)	https://github.com/RosettaCommons/RoseTTAFold_joint	[72]
ESMFold	https://esmatlas.com/about	[73–76]
HelixFold	https://github.com/THU-KEG/HelixFold	[77]
ScaleFold	https://github.com/tensorfold/ScaleFold	[78]
I-TASSER	https://zhanggroup.org/I-TASSER/	[79–83]

modelling of olfactory receptors, particularly when investigating ligand interactions and receptor function [93]. In their study, the authors generated in-house models of mOR256–3 and mOR256–8 using MODELLER 9.21 [65], a widely used software for comparative protein structure prediction (see Table 1), guided by carefully constructed sequence alignments. Four template structures were employed: human α2AR (PDB ID: 2YDV), human CXCR1 (PDB ID: 2LNL), human CXCR4 (PDB ID: 3ODU), and bovine rhodopsin (PDB ID: 1U19). The N- and C-terminal regions were excluded from modelling. All templates represented inactive conformations, a modelling choice that warrants consideration, as GPCRs, including ORs, undergo significant conformational rearrangements during activation. These changes include the outward displacement of TM5 and TM6, the reorientation of conserved microswitches (e.g., DRY, NPxxY, PIF), and the reorganisation of intramolecular contact networks that affect ligand binding and G protein coupling [94–97]. Since OR activation mechanisms remain less well characterised than those of prototypical Class A GPCRs, relying exclusively on inactive templates may bias models toward conformations that are not functionally relevant. Future modelling efforts could incorporate activation state data or integrate inactive and active-state templates to capture the receptor's functional plasticity more accurately to address this limitation. In the Yu et al. study, sequence similarity between the templates and target ORs ranged from 31 % to 38 %, increasing to 38–44 % within the TM regions [93]. Additionally, the authors constructed three ECL2-focused chimeric models by grafting extracellular loop 2 (ECL2) regions from β_2AR (PDB ID: 2RH1), M_2R (PDB ID: 3UON), and $5HT_{2C}$R (PDB ID: 6BQH), respectively, onto the mOR256–3 scaffold. These chimaeras were used to probe the functional contribution of ECL2 to odorant binding and are distinct from the core homology models of mOR256–3 and mOR256–8 [93].

Each tool employs a distinct methodology and has different strengths, rendering the available tools well-suited for various research scenarios in protein structure prediction and homology modelling. Table 1 is not an exhaustive list of tools used in protein structure modelling, but several tools we discussed in this review.

MODELLER: Inspired by NMR techniques, MODELLER integrates diverse restraints and parameters through probability density functions. These encompass homology-derived restraints, stereochemical restraints, torsional angle, non-bonded distance parameters, and optional restraints. Optimisation involves conjugate gradient descent, molecular dynamics, and simulated annealing [84,65,98].

MODELLER is highly flexible and scriptable, making it particularly advantageous for batch processing or for fine-tuning specific modelling parameters. It is well-suited for experienced users who need control over custom restraints. However, it lacks a graphical user interface, which may make it less accessible to beginners. Model quality is typically evaluated using Discrete Optimised Protein Energy (DOPE) scores, GA341 assessment scores, and structural checks for steric clashes and geometry [65,99].

SWISS-MODEL: This model is recognised for its high degree of automation. It requires minimal user input in the form of a primary sequence. It adeptly selects templates from a comprehensive database and aligns them with the target sequence. Subsequently, it generates models for all regions except insertions and deletions, crafted using constraint space programming. Side chains are added precisely by applying a backbone-dependent rotamer library, and optimisation is carried out using the steepest descent [84,67,68,100–102]. Its main advantage lies in its ease of use and accessibility through a web-based interface, making it ideal for non-specialists and high-throughput modelling. However, its automation reduces user control over template selection and loop refinement. Model quality is evaluated using the QMEAN scoring function and global quality estimates. SWISS-MODEL predictions are also independently benchmarked via the CAMEO continuous model evaluation system [68,101].

Phyre and Phyre2: These servers leverage advanced remote homology detection methods to construct 3D models, anticipate ligand binding sites, and evaluate the impact of amino acid mutations. The alignment process employs position-specific iterated BLAST (Psi-BLAST) and secondary structure prediction algorithms, aligning the target sequence with template 3D structures. A curated dataset generated through HMM-HMM–based lightning-fast iterative sequence search (HHblits) facilitates the creation of multiple sequence alignments,

subsequently informing the prediction of secondary structures with PSI-blast-based secondary structure PREDiction (PSIPRED). Query Hidden Markov Models (HMMs) are fashioned by amalgamating alignment and secondary structure prediction data. The most favourable alignments are then employed to craft models from a database containing HMMs derived from known 3D structures. Finally, loops and side chains are modelled meticulously [84,66,103].

Phyre2 is particularly effective in detecting remote homologs, making it useful for targets with low sequence similarity to known structures. Its limitations include reduced accuracy in loop modelling and limited refinement of side chains. Confidence in Phyre2 models is typically conveyed through confidence scores based on template alignments, and additional validation can be carried out using tools like MolProbity or structural superposition for RMSD estimation [66].

I-TASSER (Iterative Threading ASSEmbly Refinement): This model is a widely used platform for protein structure prediction, particularly effective for targets with limited or no close homologs. The method combines threading, fragment assembly, and ab initio modelling to build full-length atomic models from amino acid sequences. One of its primary advantages is its ability to refine initial template-based models into more native-like structures. It has shown strong performance in both benchmark tests and large-scale studies, such as modelling over 900 human G protein-coupled receptors (GPCRs), even when sequence identity with known structures is low. The method also generates reasonable models for loop regions, which are often excluded in other GPCR-specific modelling pipelines [79,80,104].

Model evaluation in I-TASSER relies on several key metrics: Cα-RMSD (to assess structural deviation), TM-score (to evaluate overall topology), and an internal C-score, which estimates model reliability based on alignment quality and simulation convergence. TM-scores above 0.5 are generally indicative of correct global folds. I-TASSER's ability to handle low-identity targets has been validated in independent benchmarking studies, including modelling of membrane proteins like rhodopsin, where it often outperformed other homology modelling tools [105].

Nevertheless, I-TASSER has limitations. Its performance may decrease for large, multi-domain proteins or when reliable threading templates are unavailable. While it is adaptable for membrane proteins, the default implementations are generally better optimised for soluble targets. In recent years, fully end-to-end AI-based approaches like AlphaFold2 have surpassed I-TASSER in accuracy for many targets. However, I-TASSER remains a robust and accessible tool, particularly when combined with other structural bioinformatics pipelines and used alongside model validation and experimental data.

5. Artificial intelligence-based structure prediction

AlphaFold, developed by DeepMind, has revolutionised protein structure prediction by replacing traditional template-dependent approaches with deep learning-based ab initio modelling. It leverages co-evolutionary information from multiple sequence alignments (MSAs). It uses transformer-based neural networks to predict inter-residue distances and orientations, ultimately constructing full 3D structures with near-experimental accuracy [69,106].

AlphaFold2's architecture integrates MSA representations, template-based features, and pairwise residue relationships within a unified attention-based framework, trained on thousands of experimentally determined protein structures. A key innovation lies in its end-to-end differentiable pipeline, which accurately predicts backbone and side-chain geometries. AlphaFold does not rely on close sequence homologs and can infer novel folds, making it especially valuable for modelling GPCRs such as ORs, which lack high-quality structural templates [84,69,107].

The AlphaFold Protein Structure Database now includes predictions for over 200 million proteins, encompassing all annotated human proteins, including the entire human OR repertoire [106]. This database has become a vital resource for researchers exploring difficult proteins to crystallise due to low expression, poor solubility, or instability in membrane environments [84,91]. Fig. 3 illustrates the AlphaFold3-generated binding pocket of the human olfactory receptor OR1AD1. AlphaFold3 expands its modelling capability beyond proteins to include ligands, nucleic acids, and cofactors, enhancing its utility in modelling receptor–ligand interactions [108]. Fig. 3 shows the predicted binding pocket of the human olfactory receptor OR1AD1 complexed with the odorant 4-Methoxyacetophenone. The structure was generated using AlphaFold3 on the University of Hertfordshire's High-Performance Computing (UHHPC) facility, and the complex was visualised in MOE 2024.0601 (Chemical Computing Group, Montreal, Canada). The ligand (blue) is nestled between transmembrane helices TM3, TM5, and TM6 (depicted as red ribbons), forming interactions with conserved aromatic and polar residues. These contacts suggest a putative orthosteric binding site, consistent with known binding modes of class A GPCRs. This structural model supports hypotheses about odorant recognition by OR1AD1 and provides a basis for downstream docking and mutagenesis studies.

RoseTTAFold, developed by the Baker Lab, employs a three-track neural network to simultaneously integrate sequence data, pairwise distances, and 3D structural information. Initially intended for monomeric protein prediction, it has since been extended to accommodate protein complexes and nucleic acid-containing assemblies [70,109].

RoseTTAFold Joint (RFjoint) is a powerful variant that allows sequence recovery and mutation effect prediction in a unified model. It achieves strong zero-shot performance without requiring family-specific training, outperforming MSA Transformer and DeepSequence in variant impact prediction [72]. RoseTTAFold has also demonstrated robust performance in antibody modelling, particularly for challenging complementarity-determining regions such as H3 loops, even without high-quality templates [71].

Despite these strengths, RoseTTAFold has limitations, particularly in accurate side-chain placement and modelling flexible loop regions. Further refinements, such as FastRelax or domain-specific retraining, may be necessary for specialised applications [71]. Nonetheless, its accessibility, speed, and predictive versatility position it as a strong complement to AlphaFold2.

ESMFold, developed by Meta AI, eschews MSAs and templates in favour of large-scale protein language models (pLMs) trained on billions of sequences. This approach enables accurate single-sequence structure prediction, particularly valuable for orphan proteins with little or no evolutionary context [73,110].

ESMFold achieves high structural accuracy and inference speed, up to 60 times faster than AlphaFold2. This makes it suitable for proteome-scale applications and quick pre-screening in structure-based discovery [110]. It also supports applications in enzyme function prediction [74] and binding site annotation [75] and has been transformed into ESM-Design, a sequence generation tool optimised for stability and expression [76].

Comparative studies show that ESMFold performs comparably to AlphaFold2 in well-annotated protein regions (e.g., Pfam domains) but slightly underperforms in complex families where MSA-based evolutionary signals are richer [78,111].

6. Other emerging frameworks and infrastructure innovations

Several frameworks now aim to extend, optimise, or simplify AlphaFold-style modelling. HelixFold achieves fast and memory-efficient training using PaddlePaddle, reducing AlphaFold2's training time to just over 5 days [77]. ScaleFold scales AlphaFold2 training to over 2000 Graphics Processing Units (GPUs), decreasing training time to 10 h through kernel fusion and custom Compute Unified Device Architecture (CUDA) optimisation [78]. Uni-Fold, a PyTorch-based reimplementation of AlphaFold and AlphaFold-Multimer, provides faster training and enhances TM-score prediction for multimeric complexes

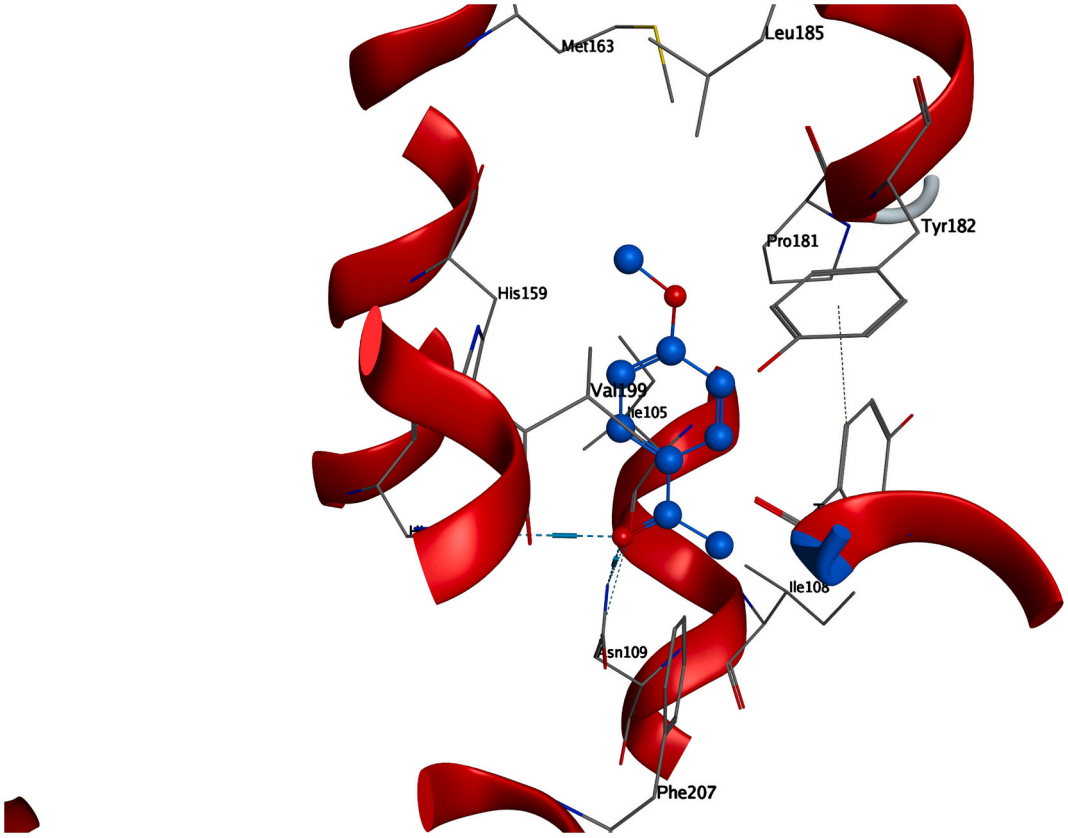


Fig. 3. A figure illustrating the binding pocket of the receptor OR1AD1 with the ligand 4-Methoxyacetophenone, generated using AlphaFold3 [108] on a High Performance Computing (HPC) system (University of Hertfordshire (UH) HPC facility) and visualised using MOE 2024.0601 (Chemical Computing Group, Montreal, Canada).

[112]. OpenComplex generalises this architecture to model RNA, protein–RNA assemblies, and non-protein targets using modular encoders [113].

Solvent, a standardised benchmarking toolkit, abstracts folding architectures into interchangeable components and facilitates reproducible model evaluation across ESMFold, OmegaFold, and others [114]. DeepFold, an AlphaFold-inspired model, integrates loss functions for side-chain accuracy and molecular energetics, ranking fourth in CASP15 [115].

Finally, BioNeMo merges AlphaFold2, OpenFold, and ESMFold predictions into a single refined structural dataset useful for structure-based ligand screening and human variant interpretation [116]. Additionally, OpenMM-Loss, a physics-aware training strategy, uses molecular dynamics gradients to enhance structural plausibility [117], while ResiRole enables functional benchmarking by quantifying how well models preserve functional sites [118].

7. Comparative insights: strengths and limitations across models

AlphaFold2 and RoseTTAFold remain the gold standards for high-resolution structure prediction across various targets. AlphaFold2 offers exceptional precision but is computationally expensive and relies on deep MSAs, making it less effective for orphan receptors without sequence homologs. RoseTTAFold is more modular and efficient, supporting multi-chain predictions with reduced computational demand; however, it lags slightly in side-chain resolution and dynamic modelling.

Conversely, Esmfold sacrifices MSA depth for speed and scalability, making it a promising option for de novo proteins, high-throughput screening, and proteins with limited evolutionary data, such as many ORs. However, its predictive power may be limited in conformationally complex or multimeric targets that contain essential co-evolutionary signals.

New platforms like HelixFold, Uni-Fold, and OpenComplex expand accessibility and the scope of structure prediction. They offer open-source frameworks that are easier to customise, retrain, and benchmark, while tools like Solvent and BioNeMo address long-standing reproducibility and integration challenges. DeepFold and OpenMM-Loss enhance physical realism and side-chain packing, marking necessary steps toward integrating structure-function prediction.

Together, these tools shift toward fast, scalable, and adaptable modelling paradigms that can infer novel folds and support functionally relevant modelling of previously inaccessible targets, such as olfactory receptors, membrane proteins, and non-canonical complexes.

8. Pharmacophore modelling

Paul Ehrlich first introduced the pharmacophore concept in 1909 [119,120]. Ehrlich defined a pharmacophore as a molecular framework that carries the essential features responsible for a drug's biological activity [119,120]. The pharmacophore concept has been expanded considerably over the past century, and it is now used to describe the essential features of a drug-target interaction [119,120]. According to the recent definition of the International Union of Pure and Applied Chemistry (IUPAC), a pharmacophore model is an ensemble of steric and electronic features necessary to ensure optimal supramolecular interactions with a specific biological target [119,120]. The pharmacophore model does not specify the drug's chemical structure but offers a framework for understanding how it interacts with its target.

A pharmacophore model can be generated using a ligand-based or structure-based technique [119]. Ligand-based pharmacophore modelling involves superimposing a set of active molecules and extracting

common chemical features essential for their bioactivity [119,121]. These features include hydrophobic areas, aromatic ring systems, hydrogen bond acceptors, hydrogen bond donors, and negatively ionisable and positively ionisable groups [120,122]. The two essential techniques in ligand-based pharmacophore modelling are handling the conformational flexibility of ligands and conducting molecular alignment [119]. These techniques are also the main challenges in this field [119]. Despite the obstacles, ligand-based pharmacophore modelling is a valuable tool for drug discovery. This technique can identify new drug candidates by identifying the common features of known ligands that bind to a particular target.

Structure-based pharmacophore modelling involves probing possible interaction points between the macromolecular target and ligands [119,121,123]. This is done by analysing the target protein's structure and identifying residues likely to interact with the ligands [119,121,123]. The pharmacophore model is then generated by identifying the common features of these interactions [119,120,123]. The Protein Data Bank (PDB) is an essential resource for the structures of protein complexes [120]. The PDB contains structures determined using various methods, including nuclear magnetic resonance (NMR) spectroscopy, X-ray diffraction, and electron microscopy. As of February 17, 2025, the PDB contains 231,356 structures from experimental methods [124] and 1,068,577 computed structure models (CSMs) [69]. The choice of pharmacophore modelling approach depends on the available data. If the structure of the target protein is known, structure-based pharmacophore modelling is the preferred approach. However, if the structure of the target protein is not known, ligand-based pharmacophore modelling can be used. Various automated pharmacophore generators have been developed in recent years. These generators are available as commercial software, such as HypoGen (Accelrys Inc., http://www.accelrys.com), DISCO [125], GALAHAD (Tripos Inc., http://www.tripos.com), PHASE (Schrödinger Inc., http://www.schrodinger.com), and MOE (Chemical Computing Group, http://www.chemcomp.com), as well as several academic programs [119]. The key steps involved in pharmacophore modelling are summarised in Fig. 4.

Pharmacophore modelling, although successful in drug design, has limitations. Challenges include modelling ligand flexibility and molecular alignment [119]. Properly selecting training set compounds, influenced by dataset size and diversity, is critical [119,126]. Overly complex pharmacophore models hinder practical use [119]. They may not reflect the quantitative structure-activity relationships (QSAR) [119]. Scoring metrics for pharmacophore-based virtual screening are lacking, and reliance on pre-computed conformation databases can miss active molecules [127–130]. Constructing pharmacophore queries lacks a standardised approach [127]. In olfaction, small ligands limit the utility of pharmacophore modelling due to the small number of pharmacophore points.

9. Molecular dynamics simulations

Molecular dynamics (MD) has emerged as a potent tool for investigating biophysical systems, benefitting from advances in computational capabilities and software availability [131–134]. MD simulations can be used to understand the structure-to-function relationships of macromolecules and the essence of protein-ligand interactions and to guide the drug discovery and design process (Fig. 5) [131]. Despite the contributions of MD to understanding complex biophysical systems, methodological difficulties remain [133]. One such difficulty is insufficient sampling, which limits the application of MD. This limitation is due to the rough energy landscapes that govern the biomolecular motion [132]. These landscapes are characterised by many local minima, separated by high-energy barriers [132,135]. As a result, it is difficult for MD simulations to escape from these local minima and explore the entire energy landscape [132].

Several methods have been developed to address the problem of insufficient sampling [132,135]. These methods include enhanced sampling techniques, such as umbrella sampling, metadynamics, and replica exchange MD [132,135]. These techniques allow MD simulations to explore a more comprehensive energy landscape range and obtain more accurate results [132,135]. In addition to sampling limitations, important methodological considerations arise when choosing between all-atom and coarse-grained (CG) simulation approaches. While all-atom MD captures detailed molecular interactions, CG approaches, such as those using the MARTINI force field, reduce complexity by grouping atoms into interaction "beads." This enables simulations over longer timescales and larger systems, making CG simulations particularly useful for membrane protein dynamics and self-assembly studies [136,137].

Another area for improvement with MD is the accuracy of the force fields used to model the interactions between atoms and molecules. Force fields are mathematical descriptions of these interactions, often

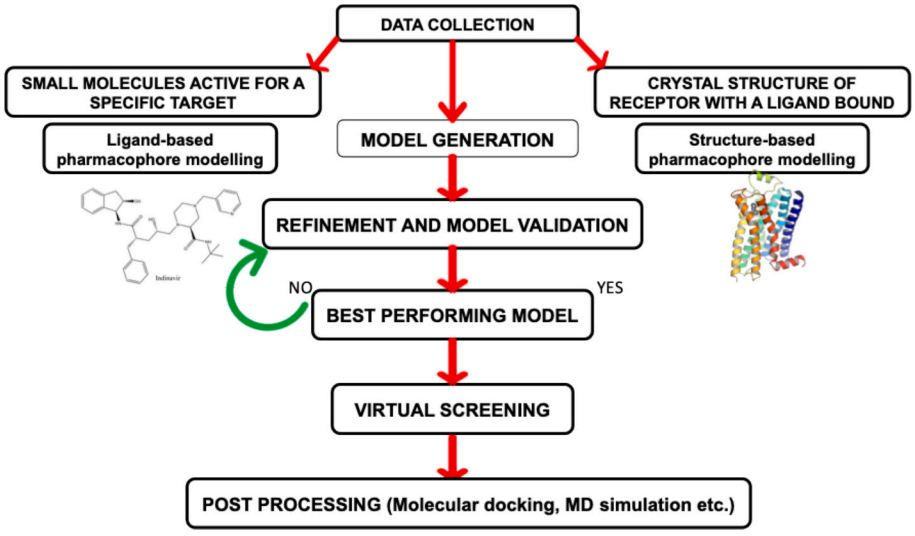


Fig. 4. Schematic showing the steps involved in pharmacophore modelling. The process starts with collecting data from either a small molecule known to be active against a specific target (ligand-based pharmacophore modelling) or crystal structures of receptor–ligand complexes (structure-based pharmacophore modelling). These inputs inform the generation of pharmacophore models, which are refined and validated. The best-performing model is chosen for virtual screening to identify new candidate compounds. Post-processing steps, such as molecular docking and molecular dynamics simulations, may be conducted to evaluate binding modes, stability, and affinity. This integrated approach facilitates the identification of potential ligands and supports structure-guided drug discovery.

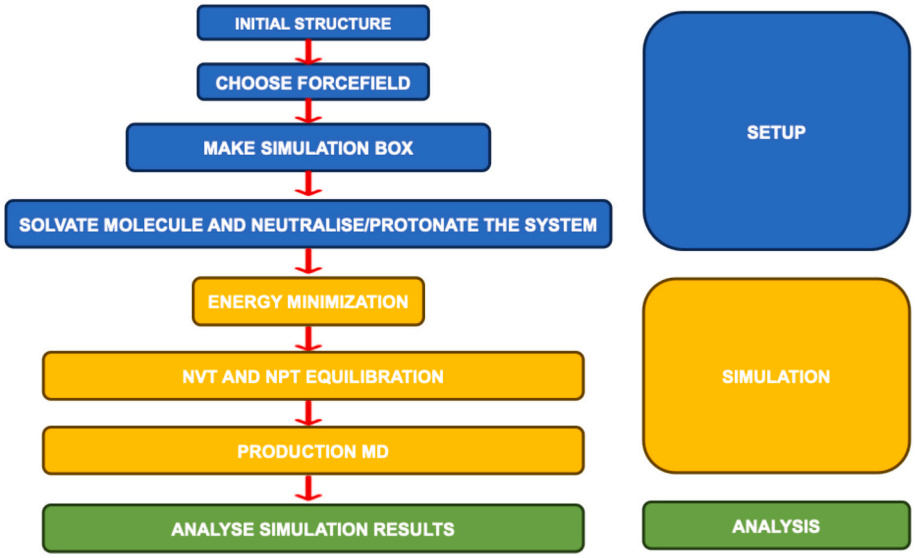


Fig. 5. Schematic representation of the key steps involved in a molecular dynamics (MD) simulation workflow. The process begins with system setup, which includes preparing the initial structure, selecting an appropriate force field, generating a simulation box, and solvating the system with suitable ion conditions. The next phase involves energy minimisation to resolve steric clashes, followed by equilibration under constant volume (NVT) and constant pressure (NPT) ensembles. Production MD simulations are then conducted to capture the system's time-dependent behaviour. The final phase involves trajectory analysis to extract structural, energetic, and dynamic information. This workflow provides a comprehensive approach for studying molecular interactions, stability, and conformational changes at the atomic level.

simplified to make the simulations computationally feasible. However, this simplification can lead to errors in the simulation results. Additionally, constructing realistic membrane-embedded protein systems is a known challenge in MD, particularly for GPCRs such as olfactory receptors. Tools like CHARMM-GUI Membrane Builder streamline this process, allowing users to embed proteins in complex lipid bilayers across several MD engines [138,139]. One emerging solution to improve realism in such systems is constant-pH molecular dynamics, which enables dynamic protonation of titratable residues during simulations. This is particularly relevant for GPCRs, where changes in protonation at conserved sites can modulate activation and ligand binding. Recent scalable implementations of constant-pH MD in GROMACS and AMBER allow this technique to be applied in complex systems [140]. Despite these difficulties, MD is a powerful tool for studying biophysical systems [133]. As computational power continues to increase and as force fields become more accurate, MD will become an even more valuable tool for understanding and predicting the behaviour of biomolecules.

MD simulations can now be restrained or guided by experimental data to address sampling and force field limitations further, enhancing model accuracy and interpretability. For example, Cryo-electron microscopy (Cryo-EM) density maps can be integrated into simulations to maintain experimentally observed conformations [141]. At the same time, SAXS/SANS data can be used to refine protein-solvent interactions and conformational ensembles [142]. Nuclear magnetic resonance (NMR) guided simulations incorporating experimental chemical shifts and Nuclear Overhauser Effects (NOEs) can enhance the validation and refinement of predicted conformational dynamics by ensuring consistency with experimental observables [143].

MD simulations are performed using computer software. Some of the most popular MD software packages include Groningen Machine for Chemical Simulations (GROMACS) (https://www.gromacs.org/index.html), Assisted Model Building with Energy Refinement (AMBER) [144], CP2K [145], Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) (https://www.lammps.org) [146].

The necessity for precisely experimentally derived protein structures in molecular docking and MD simulations is paramount [147]. However, the need for more crystal structures or high-quality protein models for olfactory receptors poses a challenge in predicting protein-ligand interactions accurately [147]. Olfactory receptors can be large and dynamically flexible, particularly those integral to intricate biological processes. Conducting MD simulations on such substantial protein systems demands considerable computational resources, limiting the scale and extent of simulations even when high-performance supercomputers are used. Correctly accounting for water molecules in simulations is crucial but often involves approximations [147].

Force fields that describe atom and molecule interactions in simulations have inherent limitations [147]. In MD simulations, these force fields may not fully capture the intricacies of protein-ligand interactions, potentially leading to inaccuracies in predicting binding affinities and protein dynamics [147]. Applications of molecular dynamics simulations to olfactory receptor systems, including studies on ligand binding, structural stability, and activation mechanisms, are discussed in later sections of this review.

10. Free energy calculations

Free energy calculations in the framework of classical molecular dynamics simulations are used in various research areas, including solvation thermodynamics, molecular recognition, and protein folding [148]. Free-energy-based simulations increasingly provide insights into protein structures, dynamics, and biological mechanisms [149]. The essential components of a free-energy calculation, a suitable model Hamiltonian, a sampling protocol, and an estimator for the free energy are independent of the specific application [148].

Binding free energy calculations based on molecular simulations provide predicted affinities for biomolecular complexes [150]. These calculations begin with a detailed system description, including its chemical composition and the interactions among its components [150]. System simulations are then used to compute thermodynamic information, such as binding affinities [150]. A primary goal of a drug discovery project is to design molecules that can bind tightly and selectively to the target protein receptor [151]. Accurate protein-ligand binding free energy prediction is essential in computational chemistry and computer-aided drug design [151]. Recent improvements in computing power, classical force field accuracy, enhanced sampling methods, and simulation setup have enabled accurate and reliable protein-ligand binding

free energy calculations [151]. This positions free energy calculations as leading in the small molecule drug discovery process [151,152].

Alchemical free energy methods represent a rigorous class of techniques involving the transformation of a molecule (e.g., a ligand) into another or a non-interacting state through a series of non-physical intermediates. A coupling parameter governs these intermediates (λ), and the resulting free energy changes are estimated using methods such as thermodynamic integration (TI), free energy perturbation (FEP), or Bennett's acceptance ratio (BAR) [153]. Unlike end-point methods, alchemical techniques explicitly sample a defined thermodynamic path, offering higher accuracy due to increased computational complexity and convergence sensitivity [154].

The Molecular Mechanics Poisson-Boltzmann Surface Area (MM/PBSA) and Molecular Mechanics Generalised Born Surface Area (MM/GBSA) methods are widely favoured for predicting binding free energies due to their superior accuracy compared to most molecular docking scoring functions, while also being less computationally intensive than alchemical free energy methods [155]. These approaches, known as end-point methods, have been extensively applied in various biomolecular studies, including protein folding, protein-ligand binding, and protein-protein interactions [155]. However, alternative approaches exist, such as Relative Binding Free Energy (RBFE) simulations, Free Energy Perturbation (FEP) for absolute binding energies, and various enhanced sampling techniques (REST2 (Replica Exchange with Solute Scaling) [151], Metadynamics (MTD) [132], Simulated Annealing [132], Adaptive biasing force (ABF) [156], and Gaussian aMD (GaMD) [156] [152]. These alternative approaches offer additional avenues for investigating binding energetics and exploring biomolecular interactions [152].

The main distinction between alchemical and end-point methods lies in the nature of the simulated transformations. End-point approaches rely on sampling only the bound and unbound states to estimate binding free energy, typically using empirical energy functions and implicit solvent models. In contrast, alchemical methods conduct transformations through intermediate λ-states between physical endpoints, enabling the calculation of absolute or relative binding energies with greater thermodynamic rigour [157].

An illustrative example of sampling difficulties in alchemical methods is the binding of toluene versus 3-iodotoluene to T4 lysozyme. The bulkier 3-iodotoluene induces a slow rearrangement of Val111, which can significantly affect free energy convergence. Suppose this conformational change is not sampled during the simulation. In that case, inaccurate results are obtained, either overly favourable or unfavourable binding predictions, depending on whether the holo or apo structure is used as the starting point [158]. This underscores the importance of careful conformational sampling and force field validation in ligand-binding free energy calculations.

Enhancing the accuracy of free energy calculations poses a complex challenge, where two crucial issues, insufficient configurational space sampling and imperfect force fields, are interlinked and hard to isolate for testing [159]. Longer simulations can naturally improve sampling and convergence, yet novel strategies may be required [159]. While advanced force fields show promise in specific scenarios, their general applicability remains unverified, especially in protein-ligand binding contexts [159–161]. Force field enhancements are expected to play a pivotal role in enhancing the accuracy of free energy calculations [159]. These improved calculations will allow for a rigorous evaluation of force field accuracy and identify areas requiring specific enhancements [159].

In practice, free energy simulations have shown their utility with homology models if input structures meet stringent quality criteria. High-resolution crystal structures, capable of precisely placing binding site residues and co-crystallized ligands from the target ligand series, are typically required for effective simulations [152,162].

Beyond sampling challenges, other issues arise in free energy methods. Perturbations that introduce formal charge changes also present known difficulties. Additionally, the treatment of explicit water molecules is a challenge. In many cases, buried waters cannot freely exchange with bulk solvent during short MD simulations, complicating the introduction of ligand modifications in these regions. Moreover, force field inaccuracies, encompassing parameterisation and inherent limitations within molecular mechanics force field equations, can constrain performance [152].

For instance, slow side chain rearrangements can be problematic when different ligands induce such rotations, as seen with toluene versus 3-iodotoluene. As Baumann et al. demonstrated, this rearrangement is often not adequately sampled in either equilibrium or nonequilibrium protocols, resulting in significant variations in free energy estimates depending on the starting structure [158]. Water sampling can also be problematic, especially when another ligand displaces a buried water molecule. In such cases, relative binding free energy calculations may converge more rapidly than absolute ones when ligands have similar shapes or displace the same water molecules. However, binding mode flips are less likely to occur with relative calculations, primarily during weak ligand interactions. Nevertheless, this issue can impact R-groups in drug-like ligands [158,163,164].

11. Molecular docking

Molecular docking is a structure-based drug design method that simulates the molecular interaction and predicts the binding mode and affinity between receptors and ligands [165,166]. The molecular docking methodology explores the behaviour of small molecules in the binding site of a target protein [166]. However, docking can also be applied in cases where the binding site is unknown, through blind docking or by combining with pocket prediction algorithms that scan the entire protein surface for potential interaction regions. Molecular docking simulates the optimal conformation of the ligand according to the complementarity and preorganisation of the binding site, which can predict and obtain the binding affinity and interaction mode between the ligand and receptor [165].

Molecular docking programs perform a search algorithm in which the ligand conformation is evaluated recursively until convergence to the minimum energy is reached [166]. Finally, an affinity scoring function ranks the candidate conformations and orientations (poses) as the sum of the electrostatic and van der Waals energies [166]. The scoring function is a mathematical function used to evaluate the binding affinity between a ligand and a receptor. Some scoring functions may also incorporate solvation, entropy, and empirical energy terms, or utilise machine learning–based methods. The choice of scoring function can significantly impact the accuracy of the docking results [165,166]. Molecular docking results can be used to identify potential drug candidates [165]. The candidates with the highest binding affinities are typically considered the most promising drug candidates. However, it is essential to note that molecular docking is a preliminary step in drug discovery. Additional experiments are normally required to confirm the candidate poses' binding affinity and assess their potential toxicity.

Over the past twenty years, more than 60 distinct docking tools and applications have been developed for academic and commercial purposes [166]. These programs employ various ligand placement strategies, which can be broadly categorised into different approaches [166]. Incremental construction techniques, such as FlexX (BioSolveIT, https://www.biosolveit.de), build the ligand from smaller fragments [166]. Shape-based algorithms, such as DOCK (UCSF DOCK, https://dock.compbio.ucsf.edu/Overview_of_DOCK/index.htm), identify potential binding poses by matching the shape of the ligand to the shape of the receptor binding site [166]. Genetic algorithms like GOLD (CCDC, https://www.ccdc.cam.ac.uk/solutions/software/gold/) and Flare (Cresset, https://cresset-group.com/software/flare/) use evolutionary methods to search for the lowest-energy binding pose [166]. Systematic search methods, such as Glide (Schrödinger, https://www.schrodinger.com/platform/products/glide/), exhaustively search the conformational space of the ligand to identify all possible binding poses [166].

Monte Carlo simulations, such as LigandFit [167], randomly sample the conformational space of the ligand to determine likely binding poses [166]. Most flexible ligand docking programs assume the receptor to be rigid. However, several other tools also support receptor flexibility, such as AutoDock Vina, which allows side-chain flexibility (https://autodock-vina.readthedocs.io), RosettaLigand, which models both ligand and receptor flexibility [168], GOLD allows for flexibility in the receptor [166] and Schrödinger's Induced Fit Docking protocol that integrates Glide docking with Prime-based receptor refinement [169].

Molecular docking techniques are riddled with limitations and difficulties. The results they predict may often diverge from experimental findings due to complex factors [170,171].

- Understanding the myriad molecular features influencing molecule interactions can be intricate and computationally demanding. Each step of the docking process adds further complexity.
- The numerous conformational degrees of freedom in molecules make positioning a ligand in a macromolecule's binding pocket both time-consuming and challenging. This necessitates the use of sophisticated scoring functions.
- The availability of low-resolution crystallographic structural data for molecules or entirely unknown structural information compounds the difficulty of locating accurate receptor binding pockets.
- The molecules' inherent flexibility and geometry alterations during binding further complicate the docking process.
- Notably, many docking tools need help with hydrated docking, which involves considering water molecules in macromolecule-ligand interactions and removing water molecules from the receptor's binding pocket before docking, while a common practice, may only sometimes be appropriate, especially when water molecules are tightly bound or functionally active in the binding site.
- Existing force fields may not support all types of atoms, adding complexity in selecting appropriate force fields and parameters for docking.
- Many docking software tools only recognise biomolecules like DNA, RNA, proteins, and enzymes as receptors, while synthetic-organic, synthetic organic-inorganic hybrid and inorganic molecules are often overlooked.
- Some docking tools may accept ligands with specific metal atoms, but recognising all metal types could be more consistent. Similarly, while a few docking tools accommodate nanoparticles as ligands, not all kinds, such as gold (Au), silver (Ag), and iron (Fe), are universally recognised.
- Handling ions as ligands can be problematic, as automatic charge neutralisation by software may require manual verification.
- Docking tools may limit the number of atoms and torsions in a molecule, further constraining their applicability.

In addition to conventional docking, data-driven docking methods have emerged as a powerful alternative that integrates experimental constraints from techniques such as NMR, EPR, FRET, or cross-linking mass spectrometry. These constraints can be used during docking to guide the search toward biologically plausible conformations. Tools like HADDOCK (High Ambiguity Driven DOCKing) enable such data to restrain docking models and improve reliability, especially for flexible or partially disordered systems [172]. These integrative approaches significantly enhance the interpretability and accuracy of docking predictions in complex biological environments.

12. Applications in olfaction

Various computational methods have been employed to investigate olfactory receptors. The recently released cryo-EM structure of OR51E2 revealed a unique binding pocket (Fig. 6). The binding pocket of olfactory receptors (ORs) is located in a similar transmembrane region as that of class A GPCRs such as the β2-adrenergic receptor and rhodopsin; however, it differs structurally in that it is typically more occluded and compact, with unique residue arrangements that confer specificity to small, volatile odorants [55]. The binding pocket of olfactory receptors (ORs) is hydrophobic, which suggests that odorants bind to ORs through hydrophobic interactions [30].

In contrast, the binding site of the β-adrenergic receptor forms ionic bonds, hydrogen bonds, and electrostatic interactions with its ligands [30]. The odorant-binding pocket identified in OR51E2 is smaller than the β_2-AR and rhodopsin binding pocket and does not engage TM2 and TM7 [55]. However, the binding specificity of ORs is not only determined by the hyper-variable region of the central transmembrane domains; the N-termini and C-termini also play a role [30,173,174]. Both terminals are short, containing approximately 20 amino acids each [30].

One study utilised molecular modelling, fingerprint interaction analysis, and molecular dynamics simulations to examine the olfactory receptor Olfr73 [92]. The aim was to gain insights into the fundamental principles governing odorant binding to this olfactory receptor and to discover novel receptor-activating compounds beyond the scope of

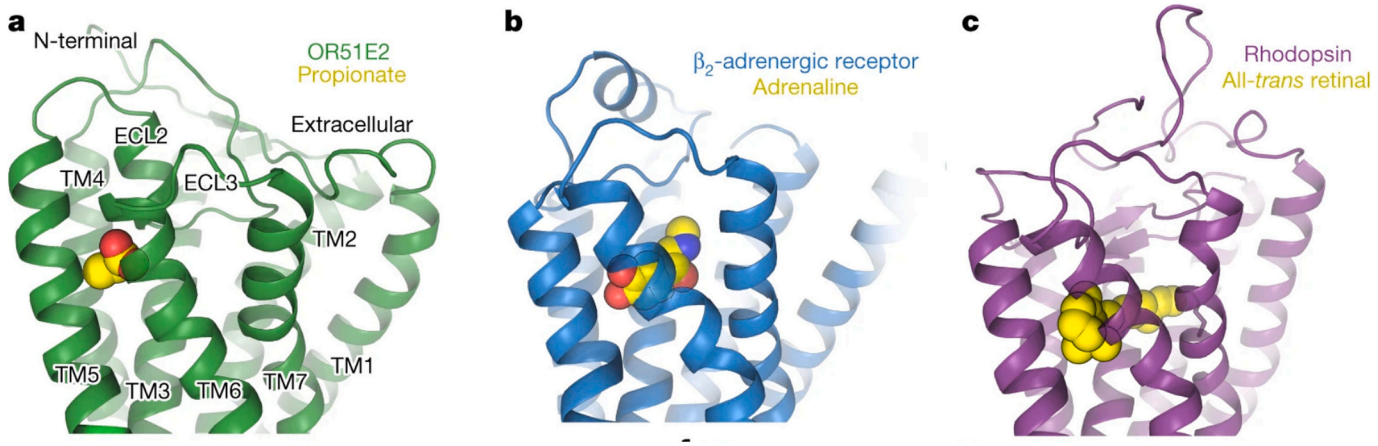


Fig. 6. Comparison of the odorant-binding pocket in OR51E2 with other class A GPCRs. (a) Cryo-EM structure of the human olfactory receptor OR51E2 bound to its ligand propionate (yellow spheres), showing key interactions with transmembrane helices TM3-TM6 and extracellular loop 2 (ECL2). (b) Structure of the β_2-adrenergic receptor (β_2AR) in complex with adrenaline (PDB ID: 4LDO), illustrating its ligand-binding site. (c) Structure of rhodopsin bound to all-trans-retinal (PDB ID: 6FUF), highlighting its orthosteric pocket. The figure illustrates similarities in the overall helical arrangement of the transmembrane domains while emphasising distinct binding pocket geometries and ligand environments [55]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

traditional chemical odorant libraries. The results revealed that the binding pocket of the homology-modelled Olfr73 is comparatively smaller but exhibits greater flexibility than the binding pockets found in typical non-olfactory G-protein-coupled receptors. The study's authors further screened a library of 1.6 million compounds against Olfr73. This screening process identified 25 potential agonists beyond conventional agonists. Among these, 17 compounds were verified through cell-based assays, with some exhibiting promising therapeutic potential. The study's overall conclusions propose a molecular explanation for the diminished interaction between an odorant and its olfactory receptor (OR), accounting for the typically low potency of OR-activating compounds. The results offer proof-of-principle for identifying novel therapeutic OR agonists.

In 2023, Alfonso-Prieto and Capelli introduced a protocol that involved conducting a series of molecular dynamics simulations using de novo structures predicted by state-of-the-art machine learning algorithms [175]. They applied this protocol to investigate the human OR51E2 receptor, which has been extensively studied [175]. The findings of this study highlighted the importance of MD simulations in refining and validating models generated de novo [175]. The study also revealed the significance of a sodium ion at a binding site close to $D^{2.50}$ and $E^{3.39}$ in stabilising the receptor's inactive state [175]. Remarkably, these two acidic residues exhibit vital conservation among human ORs, implying that this crucial requirement likely applies to the approximately 400 other members within this receptor family [175]. With the recent publication of a Cryo-EM structure of the same receptor in its active state, Alfonso-Prieto and Capelli presented this protocol as an *in-silico* method that complements ongoing efforts to unravel the structures of ORs [175].

In a recent follow-up study, Pirona et al. (2024) built upon the Alfonso-Prieto and Capelli OR51E2 modelling protocol [175] by performing extended microsecond-scale molecular dynamics simulations to investigate the effects of calcium binding on olfactory receptor inactivation. Using the experimentally determined active-state cryo-EM structure of OR51E2 as a starting point, they demonstrated that calcium (Ca^{2+}) binding to the conserved acidic residues D2.50 and E3.39 stabilises the receptor in an inactive conformation more effectively than sodium (Na^{+}) or protonation of these residues. This stabilisation was linked to reduced water permeability, suppressed lipid infiltration into the ion-binding pocket, and conformational shifts in TM5 and TM6—structural hallmarks of GPCR inactivation. Importantly, the findings suggest that calcium binding achieves better electrostatic complementarity at this site compared to monovalent ions, providing new mechanistic insights into the regulation of OR states. Given that over 90 % of human ORs possess this conserved D2.50/E3.39 motif, the study proposes calcium as a potential physiological modulator of OR activity, opening new avenues for functional studies and drug targeting in this receptor family [176].

Advanced machine learning applications have been revived due to advancements in artificial intelligence and growing research in decoding human olfactory perception from the chemical features of odorant molecules [177]. Achebouche et al. conducted a study wherein they devised Convolutional Neural Network (CNN) and Graphical Convolutional Network (GCN) models to analyse the relationships that map odorant molecules to odours and odorant molecules to olfactory receptors [177]. CNNs are deep learning architectures that capture spatial hierarchies in data, making them suitable for grid-like inputs such as molecular fingerprints, while GCNs are designed to learn directly from graph-structured data, such as molecular graphs, by propagating information across atom-level connectivity. The study employed an extensive dataset of 5955 molecules, 160 odours, and 106 olfactory receptors [177]. Their findings indicated encouraging performance results, with the GCN model achieving a precision-recall area under the curve (AUC-PR) of 0.66 for odorant-to-odour prediction, reflecting moderate predictive ability in a highly imbalanced, multi-label classification setting, and 0.91 for odorant-to-OR prediction, suggesting strong predictive power for ligand–receptor relationships [177]. In addition to their previous findings, the study reported an additional result by examining the correspondence of odours and ORs associated with 389 compounds [177]. They calculated a pairwise score for each odour-OR combination, providing further insights into the relationships between odours and olfactory receptors [177]. This analysis suggested a potential combinatorial relationship between olfactory receptors and odours [177]. The study highlights the potential of artificial intelligence in identifying the perception of smell and revealing the complete collection of receptors associated with a specific odorant molecule [177].

In 2013, Boyle, McInally, and Ray reported a cheminformatics pipeline that predicts receptor–odorant interactions from an extensive collection of chemical structures (>240,000) for receptors, which they tested on a smaller panel of odorants (~100) [178]. They use computational tools to identify shared structural features from known ligands of individual receptors [178]. Specifically, they calculated a panel of 32 molecular descriptors using Dragon software, which included physicochemical properties such as molecular weight, lipophilicity (logP), hydrogen bond donors and acceptors, and topological indices. A support vector machine (SVM) classifier was subsequently trained on these features to predict the activators and inhibitors of Drosophila olfactory receptors. These features were used to screen in silico new candidate ligands from >240,000 potential volatiles for several olfactory receptors in the Drosophila antenna [178]. Additional experiments involving nine olfactory receptors yielded a remarkable success rate of approximately 71 % for the screening process [178]. This led to the identification of numerous new activators and inhibitors [178]. In summary, their findings demonstrate that computational prediction of receptor-odour interactions holds the potential to facilitate systems-level analysis of olfactory receptors present in various organisms [178].

In another study, Gelis et al. utilised protein-ligand complex molecular dynamics (MD) simulations to better understand the dynamic interactions between proteins and odorants essential for receptor activation [179]. To achieve this, they developed a dynamic model of the human olfactory receptor hOR2AG1, a well-characterised functional receptor [179]. The homology model of hOR2AG1 was based on an X-ray structure of bovine rhodopsin with a resolution of 2.2 Å, chosen because both receptors belong to class A G protein-coupled receptors (GPCRs) and share hydrophobic ligands [179]. Through dynamic computational predictions and experimental analysis using site-directed mutagenesis, they refined the three-dimensional structure of the ligand-binding site within the hOR2AG1 homology model [179]. Their investigation identified a binding pocket between helices III, V, VI, and VII as the most promising site for ligand binding. This finding was experimentally validated by expressing wild-type (WT) and mutant hOR2AG1 receptors in Hana3a cells and conducting functional characterisations using single-cell Ca2+ imaging [179]. The study's results provide valuable insights into olfactory receptor activation and ligand binding mechanisms, contributing to understanding how olfactory receptors interact with odorants at the molecular level.

More recently, Nicoli et al. utilised experimental data to guide the investigation of OR5K1 ligand binding modes within the orthosteric binding site [180]. They incorporated structural information from AI-driven modelling, recently available in the AlphaFold Protein Structure Database, and homology modelling to gain insights into the binding process. Induced-fit docking simulations were employed to explore the conformational space of the binding site for the ensemble docking [180]. To refine their models, they used mutagenesis data to guide the sampling of side chain residues and model selection [180]. The resulting models exhibited improved rationalisation of the distinct activities of active (agonist) versus inactive molecules compared to the initial models, effectively capturing the subtle structural differences responsible for activity variations [180]. Notably, the study identified specific residues, $L104^{3.32}$ and $L255^{6.51}$, as crucial for the activity of OR5K1 agonists [180]. In conclusion, Nicoli et al. presented a robust model refinement protocol applicable for modelling the orthosteric binding site of

olfactory receptors and other G protein-coupled receptors (GPCRs) with low sequence identity to existing templates [180]. Their approach is valuable for understanding ligand binding mechanisms and receptor activity in challenging receptor structures.

In a separate investigation, Cong et al. aimed to elucidate how the amino acid sequences of ORs contribute to their diverse responses to various ligands [34]. The study involved constructing a proteochemometric (PCM) model using site-directed mutagenesis, in vitro functional assays, and molecular simulations. PCM is a computational modelling approach that integrates protein sequence information and ligand chemical descriptors into a unified framework to predict receptor–ligand interactions across multiple targets. This PCM model relied on OR sequence similarities and ligand physicochemical features to predict OR responses to odorants, employing supervised machine learning techniques [34]. Among these techniques, the Random Forest (RF) algorithm, an ensemble learning method based on decision trees, was chosen for its robustness in handling high-dimensional and nonlinear data. The research indicated that the ligand selectivity of the ORs is primarily encoded in the residues located within 8 Å of the orthosteric pocket [34]. Subsequent predictions utilising the Random Forest (RF) method achieved a hit rate of up to 58 %, as confirmed by in vitro functional assays involving 111 ORs and seven odorants with distinct scaffolds [34]. This led to the identification of sixty-four new OR–odorant pairs, with 25 ORs successfully being deorphanized [34]. The most effective model demonstrated a 56 % deorphanization rate [34]. The PCM-RF approach provides an accelerated method for mapping OR–odorant interactions and successfully deorphanizing ORs [34]. This advancement holds great promise for enhancing our understanding of olfactory receptor functionality and facilitating the identification of new receptor-ligand pairs for various therapeutic and sensory applications.

13. Conclusions and future directions

Computational modelling has become a powerful tool in studying olfactory receptors, enabling a deeper understanding of their structure, function, and dynamic interactions with odorants. Homology modelling has successfully predicted OR structures when experimental data is limited. Pharmacophore modelling aids in identifying critical residues involved in ligand binding and contributes to the identification of potential drug candidates. Molecular dynamics simulations and free energy calculations allow for investigating OR-odorant interactions, refining models, and predicting binding affinities. Molecular docking assists in predicting the binding modes and affinities between receptors and ligands, contributing to drug discovery efforts. Machine learning techniques have also successfully predicted receptor responses to various odorants, providing insights into OR selectivity.

The application of computational approaches in olfactory receptor modelling is poised to grow. With advancements in computational power and force field accuracy, molecular dynamics simulations will continue to play a vital role in understanding the dynamic behaviour of ORs. Furthermore, combining different computational methods, such as machine learning with docking or molecular dynamics simulations, holds promise for more accurate predictions of receptor-odorant interactions. Moreover, utilising experimental data in conjunction with computational models will further enhance the accuracy of predictions and facilitate the deorphanisation of ORs. As new experimental structures become available, improved homology models will contribute to a more comprehensive understanding of OR structure-function relationships. These integrated strategies collectively provide an increasingly powerful toolkit for exploring the structural basis of olfactory receptor function, ligand specificity, and receptor–ligand interactions across diverse biological systems.

Despite these advances, key challenges remain. Experimental validation of computationally predicted OR structures and ligand interactions is still limited, primarily due to the scarcity of high-resolution receptor structures and the difficulties in expressing and crystallising ORs. Additionally, the functional diversity and redundancy among ORs present further hurdles in reliably predicting receptor–ligand relationships, especially across species.

Future research will likely benefit from combining AI-based structure prediction tools (e.g., AlphaFold, RoseTTAFold) with dynamic simulation methods and structure-based screening to refine receptor models and capture biologically relevant conformations. Hybrid approaches integrating experimental constraints (e.g., cross-linking MS, FRET, NMR) with computational workflows could improve modelling accuracy and expand our understanding of OR activation and signalling.

Moreover, the emerging role of ectopically expressed ORs in non-olfactory tissues presents an exciting frontier for therapeutic discovery. Computational methods may help uncover novel ligands, elucidate tissue-specific receptor functions, and guide the design of selective modulators with clinical relevance.

Finally, with the increasing availability of OR–ligand datasets, the integration of machine learning techniques, including deep learning and graph neural networks, offers promise for large-scale deorphanization and ligand profiling, paving the way for a systems-level understanding of olfactory coding and its biomedical applications. These developments will advance basic olfactory biology and enable translational applications in drug discovery, diagnostics, and sensory modulation.

Acknowledgements/funding

This work was funded by joint UK Medical Research Council / US National Science Foundation grant "NeuroNex Next-Generation Networks: From Odor to Action" (MRC #MR/T046759/1, NSF #2014217).

CRediT authorship contribution statement

Chiemela S. Odoemelam: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Volker Steuber:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization. **Michael Schmuker:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: (Chiemela S Odoemelam, Volker Steuber, Michael Schmuker reports financial support was provided by UK Research and Innovation Medical Research Council. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.).

Data availability

No new data were generated or analysed in support of this research.

References

[1] L. Buck, R. Axel, A novel multigene family may encode odorant receptors: a molecular basis for odor recognition, Cell 65 (1) (1991) 175–187.
[2] S. Sachse, I. Manzini, Editorial for the special issue 'olfactory coding and circuitries', Cell Tissue Res. [Internet] 383 (1) (2021 Jan 1) 1–6, https://doi.org/10.1007/s00441-020-03389-1.
[3] B.W. Ache, J.M. Young, Olfaction: diverse species, conserved principles, Neuron 48 (3) (2005) 417–430.
[4] R. Sharma, Y. Ishimaru, I. Davison, K. Ikegami, M.S. Chien, H. You, et al., Olfactory receptor accessory proteins play crucial roles in receptor function and gene choice. Nathans J, editor, eLife [Internet] 6 (2017 Mar 6) e21895, https://doi.org/10.7554/eLife.21895.
[5] C.D. Derby, J. Caprio, What are olfaction and gustation, and do all animals have them? Chem. Senses [Internet] 49 (2024 Jan 1) https://doi.org/10.1093/chemse/bjae009 bjae009. [cited 2025 Apr 17].

[6] C.S. Porteus, C.C. Roggatz, Z. Velez, J.D. Hardege, P.C. Hubbard, Acidification can directly affect olfaction in marine organisms, J. Exp. Biol. [Internet] 224 (14) (2021 Jul 28), https://doi.org/10.1242/jeb.237941 jeb237941. [cited 2025 Apr 17].

[7] J.M. Sullivan, L. Bonnaud-Ponticelli, A. Di Cosmo, Sensory systems of aquatic animals, Front. Physiol. 15 (2024) 1432649.

[8] B.C. Jones, M.M. Rocker, R.S.J. Keast, D.L. Callahan, H.J. Redmond, R. P. Smullen, et al., Systematic review of the odorous volatile compounds that contribute to flavour profiles of aquatic animals, Rev. Aquac. [Internet] 14 (3) (2022 Jun 1) 1418–1477 [cited 2025 Apr 17]. Available from, https://doi.org/10.1111/raq.12657.

[9] U. Stockhorst, R. Pietrowsky, Olfactory perception, communication, and the nose-to-brain pathway, Physiol. Behav. [Internet] 83 (1) (2004 Oct 30) 3–11. Available from, https://www.sciencedirect.com/science/article/pii/S0031938404003439.

[10] C. Marneweck, A. Jürgens, A.M. Shrader, Dung odours signal sex, age, territorial and oestrous state in white rhinos, Proc. R. Soc. B Biol. Sci. 2017 (284) (1846), 20162376.

[11] S.P. Caro, J. Balthazart, Pheromones in birds: myth or reality? J. Comp. Physiol. A. 196 (2010) 751–766.

[12] S.K. Woodley, Pheromonal communication in amphibians, J. Comp. Physiol. A. 196 (2010) 713–727.

[13] R.C. Jimenez, N. Casajuana-Martin, A. García-Recio, L. Alcántara, L. Pardo, M. Campillo, et al., The mutational landscape of human olfactory G protein-coupled receptors, BMC Biol. 19 (1) (2021 Feb 5) 21.

[14] B. Malnic, P.A. Godfrey, L.B. Buck, The human olfactory receptor gene family, Proc. Natl. Acad. Sci. USA 101 (8) (2004 Feb 24) 2584–2589.

[15] Y. Niimura, A. Matsui, K. Touhara, Extreme expansion of the olfactory receptor gene repertoire in African elephants and evolutionary dynamics of orthologous gene groups in 13 placental mammals, Genome Res. 24 (9) (2014 Sep) 1485–1496.

[16] Y. Fukutani, M. Abe, H. Saito, R. Eguchi, T. Tazawa, C.A. de March, et al., Antagonistic interactions between odorants alter human odor perception, Curr. Biol. [Internet] 33 (11) (2023 Jun 5) 2235–2245 [cited 2023 Aug 15]. e4, https://doi.org/10.1016/j.cub.2023.04.072.

[17] C. Wu, M. Xu, J. Dong, W. Cui, S. Yuan, The structure and function of olfactory receptors, Trends Pharmacol. Sci. [Internet] (2025), https://doi.org/10.1016/j.tips.2024.01.004 [cited 2024 Feb 27].

[18] C. Trimmer, A. Keller, N.R. Murphy, L.L. Snyder, J.R. Willer, M.H. Nagai, et al., Genetic variation across the human olfactory receptor repertoire alters odor perception, Proc. Natl. Acad. Sci. USA 116 (19) (2019) 9475–9480. May 7.

[19] V. Slepak, A. Pronin, Ectopically expressed olfactory GPCRs OR51E1/2: ligand specificity, signal transduction, and cytostatic and cytotoxic activity, FASEB J. 35 (S1) (2021 May 1).

[20] Y. Niimura, M. Nei, Extensive gains and losses of olfactory receptor genes in mammalian evolution, PLoS ONE 2 (8) (2007 Aug 8) e708.

[21] R. Zhang, P. Wang, S. Yu, P. Hansbro, H. Wang, Computerized screening of G-protein coupled receptors to identify and characterize olfactory receptors, J. Toxicol. Environ. Health A 83 (1) (2020 Jan 2) 9–19.

[22] A. Pronin, V. Slepak, Ectopically expressed olfactory receptors OR51E1 and OR51E2 suppress proliferation and promote cell death in a prostate cancer cell line, J. Biol. Chem. 296 (2021).

[23] B.D. Shepard, H. Koepsell, J.L. Pluznick, Renal olfactory receptor 1393 contributes to the progression of type 2 diabetes in a diet-induced obesity model, Am. J. Physiol. Renal Physiol. 316 (2) (2019) F372–F381.

[24] S.J. Lee, I. Depoortere, H. Hatt, Therapeutic potential of ectopic olfactory and taste receptors, Nat. Rev. Drug Discov. [Internet] 18 (2) (2019 Feb 1) 116–138. Available from, https://doi.org/10.1038/s41573-018-0002-3.

[25] N.M. Dalesio, S.F. Barreto Ortiz, J.L. Pluznick, D.E. Berkowitz, Olfactory, taste, and photo sensory receptors in non-sensory organs: it just makes sense, Front. Physiol. 9 (2018) 1673.

[26] I. Ferrer, P. Garcia-Esparcia, M. Carmona, E. Carro, E. Aronica, G.G. Kovacs, et al., Olfactory receptors in non-chemosensory organs: the nervous system in health and disease, Front. Aging Neurosci. 8 (2016 Jul).

[27] J. Cheng, Z. Yang, X.Y. Ge, M.X. Gao, R. Meng, X. Xu, et al., Autonomous sensing of the insulin peptide by an olfactory G protein-coupled receptor modulates glucose metabolism, Cell Metab. [Internet] 34 (2) (2022 Feb 1) 240–255, e10. Available from, https://www.sciencedirect.com/science/article/pii/S1550413121006860.

[28] Z. Yang, J. Cheng, P. Shang, J.P. Sun, X. Yu, Emerging roles of olfactory receptors in glucose metabolism, Trends Cell Biol. [Internet] 33 (6) (2023 Jun 1) 463–476. Available from, https://www.sciencedirect.com/science/article/pii/S096289242200215X.

[29] D. Maßberg, H. Hatt, Human olfactory receptors: novel cellular functions outside of the nose, Physiol. Rev. [Internet] 98 (3) (2018 Jul 1) 1739–1763, https://doi.org/10.1152/physrev.00013.2017 [cited 2025 Apr 17].

[30] A. Sharma, R. Kumar, I. Aier, R. Semwal, P. Tyagi, P. Varadwaj, Sense of smell: structural, functional, mechanistic advancements and challenges in human olfactory research, Curr. Neuropharmacol. 17 (9) (2019) 891–911.

[31] B. Malnic, J. Hirono, T. Sato, L.B. Buck, Combinatorial receptor codes for odors, Cell 96 (5) (1999) 713–723.

[32] J. Porter, B. Craven, R.M. Khan, S.J. Chang, I. Kang, B. Judkewitz, et al., Mechanisms of scent-tracking in humans, Nat. Neurosci. [Internet] 10 (1) (2007 Jan 1) 27–29. Available from: https://doi.org/10.1038/nn1819.

[33] I. Glezer, B. Malnic, Chapter 5 - Olfactory receptor function, in: R.L. Doty (Ed.), Handbook of Clinical Neurology, Elsevier, 2019, pp. 67–78.

[34] X. Cong, W. Ren, J. Pacalon, R. Xu, L. Xu, X. Li, et al., Large-scale G protein-coupled olfactory receptor–ligand pairing, ACS Cent. Sci. 8 (3) (2022 Mar 23) 379–387.

[35] S.E. Ryu, T. Shim, J.Y. Yi, S.Y. Kim, S.H. Park, S.W. Kim, et al., Odorant receptors containing conserved amino acid sequences in transmembrane domain 7 display distinct expression patterns in mammalian tissues, Mol Cells [Internet] 40 (12) (2017 Dec 1) 954–965. Available from, https://www.sciencedirect.com/science/article/pii/S1016847823006878.

[36] S. Zozulya, F. Echeverri, T. Nguyen, The human olfactory receptor repertoire, Genome Biol. 2 (6) (2001). RESEARCH0018.

[37] G. Pándy-Szekeres, C. Munk, T.M. Tsonkov, S. Mordalski, K. Harpsøe, A. S. Hauser, et al., GPCRdb in 2018: adding GPCR structure models and ligands, Nucleic Acids Res. 46 (D1) (2018) D440–D446.

[38] G. Pándy-Szekeres, J. Caroli, A. Mamyrbekov, A.A. Kermani, G.M. Keserű, A. J. Kooistra, et al., GPCRdb in 2023: state-specific structure models using AlphaFold2 and new ligand resources, Nucl. Acids Res. [Internet] 51 (D1) (2023 Jan 6) D395–D402 [cited 2023 Aug 11]. Available from, https://doi.org/10.1093/nar/gkac1013.

[39] F. Horn, E. Bettler, L. Oliveira, F. Campagne, F.E. Cohen, G. Vriend, GPCRDB information system for G protein-coupled receptors, Nucleic Acids Res. 31 (1) (2003) 294–297.

[40] F. Horn, J. Weare, M.W. Beukers, S. Hörsch, A. Bairoch, W. Chen, et al., GPCRDB: an information system for G protein-coupled receptors, Nucleic Acids Res. 26 (1) (1998) 275–279.

[41] V. Isberg, S. Mordalski, C. Munk, K. Rataj, K. Harpsøe, A.S. Hauser, et al., GPCRdb: an information system for G protein-coupled receptors, Nucl. Acids Res. [Internet] 44 (D1) (2016 Jan 4) D356–D364 [cited 2023 Aug 11]. Available from, https://doi.org/10.1093/nar/gkv1178.

[42] V. Isberg, B. Vroling, R. van der Kant, K. Li, G. Vriend, D. Gloriam, GPCRDB: an information system for G protein-coupled receptors, Nucl. Acids Res. [Internet] 42 (D1) (2014 Jan 1) D422–D425 [cited 2023 Aug 11]. Available from, https://doi.org/10.1093/nar/gkt1255.

[43] C. Munk, V. Isberg, S. Mordalski, K. Harpsøe, K. Rataj, A.S. Hauser, et al., GPCRdb: the G protein-coupled receptor database – an introduction, Br. J. Pharmacol. [Internet] 173 (14) (2016 Jul 1) 2195–2207 [cited 2023 Nov 8]. Available from, https://doi.org/10.1111/bph.13509.

[44] A.J. Kooistra, S. Mordalski, G. Pándy-Szekeres, M. Esguerra, A. Mamyrbekov, C. Munk, et al., GPCRdb in 2021: integrating GPCR sequence, structure and function, Nucleic Acids Res. 49 (D1) (2021) D335–D343.

[45] H. Breer, Olfactory receptors: molecular basis for recognition and discrimination of odors, Anal. Bioanal. Chem. 377 (2003) 427–433.

[46] M. Spehr, S.D. Munger, Olfactory receptors: G protein-coupled receptors and beyond, J. Neurochem. 109 (6) (2009) 1570–1583.

[47] P. Mombaerts, Genes and ligands for odorant, vomeronasal and taste receptors, Nat. Rev. Neurosci. 5 (4) (2004) 263–278.

[48] J.A. Ballesteros, H. Weinstein, [19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors, in: S.C. Sealfon (Ed.), Methods in Neurosciences [Internet], Academic Press, 1995, pp. 366–428. Available from: https://www.sciencedirect.com/science/article/pii/S1043947105800497.

[49] L.R. Saraiva, X. Ibarra-Soria, M. Khan, M. Omura, A. Scialdone, P. Mombaerts, et al., Hierarchical deconstruction of mouse olfactory sensory neurons: from whole mucosa to single-cell RNA-seq, Sci. Rep. [Internet] 5 (1) (2015 Dec 16) 18178. Available from, https://doi.org/10.1038/srep18178.

[50] N. Kanageswaran, M. Demond, M. Nagel, B.S.P. Schreiner, S. Baumgart, P. Scholz, et al., Deep sequencing of the murine olfactory receptor neuron transcriptome, PLOS ONE [Internet] 10 (1) (2015 Jan 15) e0113170. Available from: https://doi.org/10.1371/journal.pone.0113170.

[51] David H. Gire, Kevin M. Franks, Joseph D. Zak, Kenji F. Tanaka, Jennifer D. Whitesell, Abigail A. Mulligan, et al., Mitral cells in the olfactory bulb are mainly excited through a multistep signaling path, J. Neurosci. [Internet] 32 (9) (2012 Feb 29) 2964. Available from: http://www.jneurosci.org/content/32/9/2964.abstract.

[52] R. Inoue, Y. Fukutani, T. Niwa, H. Matsunami, M. Yohda, Identification and characterization of proteins that are involved in RTP1S-dependent transport of olfactory receptors, Int. J. Mol. Sci. 24 (9) (2023).

[53] S. Serizawa, K. Miyamichi, H. Sakano, One neuron–one receptor rule in the mouse olfactory system, Trends Genet. [Internet] 20 (12) (2004 Dec 1) 648–653. Available from: https://www.sciencedirect.com/science/article/pii/S0168952504002720.

[54] H. Saito, Q. Chi, H. Zhuang, H. Matsunami, J.D. Mainland, Odor coding by a mammalian receptor repertoire, Sci. Signal. 2 (60) (2009 Mar 3) ra9.

[55] C.B. Billesbølle, C.A. de March, W.J.C. van der Velden, N. Ma, J. Tewari, C.L. del Torrent, et al., Structural basis of odorant recognition by a human odorant receptor, Nature [Internet] 615 (7953) (2023 Mar 1) 742–749. Available from, https://doi.org/10.1038/s41586-023-05798-y.

[56] P.D. Lyne, Structure-based virtual screening: an overview, Drug Discov. Today [Internet] 7 (20) (2002 Oct 15) 1047–1055. Available from, https://www.sciencedirect.com/science/article/pii/S1359644602024832.

[57] J. Lötsch, D. Kringel, T. Hummel, Machine learning in human olfactory research, Chemical Senses [Internet] 44 (1) (2019 Jan 1) 11–22 [cited 2025 Apr 17]. Available from: https://doi.org/10.1093/chemse/bjy067.

[58] A. E Lohning, S. M Levonis, B. Williams-Noonan, S. S Schweiker, A practical guide to molecular docking and homology modelling for medicinal chemists, Curr. Top. Med. Chem. 17 (18) (2017) 2023–2040.

[59] T. Schwede, Protein modeling: what happened to the "protein structure gap"? Structure 21 (9) (2013) 1531–1540.

[60] T. Schmidt, A. Bergner, T. Schwede, Modelling three-dimensional protein structures for applications in drug design, Drug Discov. Today [Internet] 19 (7) (2014 Jul 1) 890–897. Available from: https://www.sciencedirect.com/science/article/pii/S1359644613003942.

[61] G. Bitencourt-Ferreira, W.F. de Azevedo, Homology modeling of protein targets with MODELLER, in: W.F. de Azevedo Jr. (Ed.), Docking Screens for Drug Discovery [Internet], Springer New York, New York, NY, 2019, pp. 231–249, https://doi.org/10.1007/978-1-4939-9752-7_15.

[62] C. Chothia, A.M. Lesk, The relation between the divergence of sequence and structure in proteins, EMBO J. [Internet] 5 (4) (1986 Apr 1) 823–826 [cited 2025 Apr 17]. Available from, https://doi.org/10.1002/j.1460-2075.1986.tb04288.x.

[63] M.T. Muhammed, E. Aki-Yalcin, Homology modeling in drug discovery: overview, current applications, and future perspectives, Chem. Biol. Drug Des. [Internet] 93 (1) (2019 Jan 1) 12–20 [cited 2023 Jul 17]. Available from: https://doi.org/10.1111/cbdd.13388.

[64] T. Werner, M.B. Morris, S. Dastmalchi, W.B. Church, Structural modelling and dynamics of proteins for insights into drug interactions, Adv. Drug Deliv. Rev. 64 (4) (2012) 323–343.

[65] B. Webb, A. Sali, Comparative protein structure modeling using MODELLER, Curr. Protoc. Bioinform. [Internet] 54 (1) (2016 Jun 1), https://doi.org/10.1002/cpbi.3, 5.6.1-5.6.37. Available from:. [cited 2023 Jul 17].

[66] L.A. Kelley, S. Mezulis, C.M. Yates, M.N. Wass, M.J.E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis, Nat. Protoc. [Internet] 10 (6) (2015 Jun 1) 845–858. Available from: https://doi.org/10.1038/nprot.2015.053.

[67] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, et al., SWISS-MODEL: homology modelling of protein structures and complexes, Nucl. Acids Res. [Internet] 46 (W1) (2018 Jul 2) W296–W303 [cited 2023 Jul 17]. Available from, https://doi.org/10.1093/nar/gky427.

[68] S. Bienert, A. Waterhouse, T.A.P. de Beer, G. Tauriello, G. Studer, L. Bordoli, et al., The SWISS-MODEL repository—new features and functionality, Nucl. Acids Res. [Internet] 45 (D1) (2017 Jan 4) D313–D319 [cited 2023 Jul 17]. Available from: https://doi.org/10.1093/nar/gkw1132.

[69] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, et al., Highly accurate protein structure prediction with AlphaFold, Nature [Internet] 596 (7873) (2021 Aug 26) 583–589 [cited 2025 Apr 18]. Available from, https://www.nature.com/articles/s41586-021-03819-2.

[70] M. Baek, R. McHugh, I. Anishchenko, H. Jiang, D. Baker, F. DiMaio, Accurate prediction of protein–nucleic acid complexes using RoseTTAFoldNA, Nat. Methods [Internet] 21 (1) (2024) 117–121. Available from, https://www.nature.com/articles/s41592-023-02086-5.

[71] T. Liang, C. Jiang, J. Yuan, Y. Othman, X.Q. Xie, Z. Feng, Differential performance of RoseTTAFold in antibody modeling, Brief. Bioinform. [Internet] 23 (5) (2022 Sep 20), https://doi.org/10.1093/bib/bbac152/6589866 bbac152. Available from:. [cited 2025 Apr 18].

[72] S. Mansoor, M. Baek, D. Juergens, J.L. Watson, D. Baker, Zero-shot mutation effect prediction on protein stability and function using RoseTTAFold, Protein Sci. [Internet] 32 (11) (2023 Nov), https://doi.org/10.1002/pro.4780 e4780. Available from:. [cited 2025 Apr 18.

[73] M. Zalewski, B. Wallner, S. Kmiecik, Protein–peptide docking with ESMFold language model, J. Chem. Theor. Comput. [Internet] 21 (6) (2025 Mar 25) 2817–2821 [cited 2025 Apr 18]. Available from, https://doi.org/10.1021/acs.jctc.4c01585.

[74] Y. Song, Q. Yuan, S. Chen, Y. Zeng, H. Zhao, Y. Yang, Accurately predicting enzyme functions through geometric graph learning on ESMFold-predicted structures, Nat. Commun. [Internet] 15 (1) (2024 Sep 18) 8180 [cited 2025 Apr 18]. Available from: https://www.nature.com/articles/s41467-024-52533-w.

[75] Y. Fang, Y. Jiang, L. Wei, Q. Ma, Z. Ren, Q. Yuan, et al., DeepProSite: structure-aware protein binding site prediction using ESMFold and pretrained language model. Cowen L, editor, Bioinformatics [Internet] 39 (12) (2023 Dec 1) [cited 2025 Apr 18]. btad718. Available from: https://academic.oup.com/bioinformatics/article/doi/10.1093/bioinformatics/btad718/7453375.

[76] J.R. Jeliazkov, D. Del Alamo, J.D. Karpiak, ESMFold Hallucinates Native-Like Protein Sequences [Internet], 2023, https://doi.org/10.1101/2023.05.23.541774 [cited 2025 Apr 18]. Available from.

[77] G. Wang, X. Fang, Z. Wu, Y. Liu, Y. Xue, Y. Xiang, et al., HelixFold: An Efficient Implementation of AlphaFold2 using PaddlePaddle [Internet], arXiv. [cited 2025 Apr 18]. Available from, http://arxiv.org/abs/2207.05477, 2022.

[78] F. Zhu, A. Nowaczynski, R. Li, J. Xin, Y. Song, M. Marcinkiewicz, et al., ScaleFold: Reducing AlphaFold initial training time to 10 hours, in: Proceedings of the 61st ACM/IEEE Design Automation Conference [Internet], ACM, San Francisco CA USA, 2024, pp. 1–6 [cited 2025 Apr 18]. Available from: https://doi.org/10.1145/3649329.3657326.

[79] S. Wu, J. Skolnick, Y. Zhang, Ab initio modeling of small proteins by iterative TASSER simulations, BMC Biol. [Internet] 5 (1) (2007 Dec) 17 [cited 2025 Apr 19]. Available from: https://doi.org/10.1186/1741-7007-5-17.

[80] P.R. Suganya, K. Sudevan, S. Kalva, L.M. Saleena, Homology Modeling for HUMAN ADAM12 using Prime, I-TASSER and Easymodeller 6, 2025.

[81] J. Yang, Y. Zhang, I-TASSER server: new development for protein structure and function predictions, Nucl. Acids Res. [Internet] 43 (W1) (2015 Jul 1) W174–W181 [cited 2025 Apr 19]. Available from: https://doi.org/10.1093/nar/gkv342.

[82] J. Zhang, J. Yang, R. Jang, Y. Zhang, GPCR-I-TASSER: a hybrid approach to G protein-coupled receptor structure modeling and the application to the human genome, Structure [Internet] 23 (8) (2015 Aug) 1538–1549 [cited 2025 Apr 19]. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0969212615002324.

[83] W. Zheng, C. Zhang, Y. Li, R. Pearce, E.W. Bell, Y. Zhang, Folding non-homologous proteins by coupling deep-learning contact maps with I-TASSER assembly simulations, Cell Rep. Methods [Internet] 1 (3) (2021 Jul) 100014 [cited 2025 Apr 19]. Available from, https://linkinghub.elsevier.com/retrieve/pii/S266723752100014X.

[84] T. Hameduh, Y. Haddad, V. Adam, Z. Heger, Homology modeling in the time of collective and artificial intelligence, Comput. Struct. Biotechnol. J. [Internet] 18 (2020) 3494–3506 [cited 2025 Apr 18]. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2001037020304748.

[85] Y. Haddad, V. Adam, Z. Heger, Ten quick tips for homology modeling of high-resolution protein 3D structures, PLoS Comput. Biol. 16 (4) (2020 Apr) e1007449.

[86] G. Munsamy, M.E. Soliman, Homology modeling in drug discovery-an update on the last decade, Lett. Drug Des. Discov. 14 (9) (2017) 1099–1111.

[87] A. Kryshtafovych, B. Monastyrskyy, K. Fidelis, CASP prediction center infrastructure and evaluation measures in CASP10 and CASP ROLL, Proteins 82 (Suppl. 2) (2014 Feb) 7–13.

[88] P. Larsson, B. Wallner, E. Lindahl, A. Elofsson, Using multiple templates to improve quality of homology models in automated homology modeling, Protein Sci. 17 (6) (2008) 990–1002.

[89] C.N. Cavasotto, S.S. Phatak, Homology modeling in drug discovery: current trends and applications, Drug Discov. Today [Internet] 14 (13) (2009 Jul 1) 676–683. Available from: https://www.sciencedirect.com/science/article/pii/S1359644609001469.

[90] B. Berwal, P. Saha, R. Kumar, Improving olfactory Receptor Structure Modeling via Hybrid Methods. bioRxiv, 2024, 2024–05.

[91] A. Kato, K. Touhara, Mammalian olfactory receptors: pharmacology, G protein coupling and desensitization, Cell. Mol. Life Sci. 66 (23) (2009 Dec) 3743–3753.

[92] S. Yuan, T. Dahoun, M. Brugarolas, H. Pick, S. Filipek, H. Vogel, Computational modeling of the olfactory receptor Olfr73 suggests a molecular basis for low potency of olfactory receptor-activating compounds, Commun. Biol. 2 (1) (2019) 141.

[93] Y. Yu, Z. Ma, J. Pacalon, L. Xu, W. Li, C. Belloir, et al., Extracellular loop 2 of G protein–coupled olfactory receptors is critical for odorant recognition, J. Biol. Chem. 298 (9) (2022).

[94] R.O. Dror, D.H. Arlow, P. Maragakis, T.J. Mildorf, A.C. Pan, H. Xu, et al., Activation mechanism of the β2-adrenergic receptor, Proc. Natl. Acad. Sci. [Internet] 108 (46) (2011 Nov 15) 18684–18689 [cited 2025 Apr 17]. Available from, https://doi.org/10.1073/pnas.1110499108.

[95] V.V. Gurevich, E.V. Gurevich, Molecular mechanisms of GPCR signaling: a structural perspective, Int. J. Mol. Sci. 18 (12) (2017) 2519.

[96] Q. Zhou, D. Yang, M. Wu, Y. Guo, W. Guo, L. Zhong, et al., Common activation mechanism of class A GPCRs. Shan Y, Boudker O, editors. eLife [Internet] 8, 2019 Dec 19 e50279. Available from, https://doi.org/10.7554/eLife.50279.

[97] A.S. Hauser, A.J. Kooistra, C. Munk, F.M. Heydenreich, D.B. Veprintsev, M. Bouvier, et al., GPCR activation mechanisms across classes and macro/microscales, Nat. Struct. Mol. Biol. [Internet] 28 (11) (2021 Nov 1) 879–888. Available from, https://doi.org/10.1038/s41594-021-00674-7.

[98] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (3) (1993 Dec 5) 779–815.

[99] N. Eswar, B. Webb, M.A. Marti-Renom, M.S. Madhusudhan, D. Eramian, M. Shen, et al., Comparative protein structure modeling using modeller, CP Bioinformatics [Internet] 15 (1) (2006 Sep), https://doi.org/10.1002/0471250953.bi0506s15. Available from:. [cited 2025 Apr 19].

[100] K. Arnold, L. Bordoli, J. Kopp, T. Schwede, The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling, Bioinformatics 22 (2) (2006 Jan 15) 195–201.

[101] M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, et al., SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information, Nucleic Acids Res. 42 (Web Server issue) (2014 Jul). W252-258.

[102] N. Guex, M.C. Peitsch, SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling, Electrophoresis 18 (15) (1997 Dec) 2714–2723.

[103] L.A. Kelley, M.J.E. Sternberg, Protein structure prediction on the web: a case study using the Phyre server, Nat. Protoc. 4 (3) (2009) 363–371.

[104] Y. Zhang, M.E. DeVries, J. Skolnick, Structure modeling of all identified G protein–coupled receptors in the human genome. Murray D, editor, PLoS Comput. Biol. [Internet] 2 (2) (2006 Feb 17), https://doi.org/10.1371/journal.pcbi.0020013 e13. Available from. [cited 2025 Apr 19].

[105] D.M. Nikolaev, A.A. Shtyrov, M.S. Panov, A. Jamal, O.B. Chakchir, V. A. Kochemirovsky, et al., A comparative study of modern homology modeling algorithms for rhodopsin structure prediction, ACS Omega [Internet] 3 (7) (2018 Jul 31) 7555–7566 [cited 2025 Apr 19]. Available from: https://doi.org/10.1021/acsomega.8b00721.

[106] M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, et al., AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models, Nucl. Acids Res. [Internet] 50 (D1) (2022 Jan 7) D439–D444 [cited 2025 Apr 18]. Available from: https://academic.oup.com/nar/article/50/D1/D439/6430488.

[107] T.C. Terwilliger, D. Liebschner, T.I. Croll, C.J. Williams, A.J. McCoy, B.K. Poon, et al., AlphaFold predictions are valuable hypotheses and accelerate but do not replace experimental structure determination, Nat. Methods [Internet] 21 (1)

(2024 Jan) 110–116 [cited 2025 Apr 18]. Available from: https://www.nature.com/articles/s41592-023-02087-4.
[108] J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, et al., Accurate structure prediction of biomolecular interactions with AlphaFold 3, Nature [Internet] 630 (8016) (2024 Jun 13) 493–500 [cited 2025 Apr 18]. Available from: https://www.nature.com/articles/s41586-024-07487-w.
[109] M.L. Hekkelman, I. De Vries, R.P. Joosten, A. Perrakis, AlphaFill: enriching AlphaFold models with ligands and cofactors, Nat. Methods [Internet] 20 (2) (2023 Feb) 205–213 [cited 2025 Apr 18]. Available from: https://www.nature.com/articles/s41592-022-01685-y.
[110] Z. Lin, H. Akin, R. Rao, B. Hie, Z. Zhu, W. Lu, et al., Language models of protein sequences at the scale of evolution enable accurate structure prediction, 2025.
[111] M. Manfredi, C. Savojardo, G. Iardukhin, D. Salomoni, A. Costantini, P.L. Martelli, et al., Alpha&ESMhFolds: a web server for comparing AlphaFold2 and ESMFold models of the human reference proteome, J. Mol. Biol. [Internet] 436 (17) (2024 Sep) 168593 [cited 2025 Apr 18]. Available from, https://linkinghub.elsevier.com/retrieve/pii/S0022283624001888.
[112] Z. Li, X. Liu, W. Chen, F. Shen, H. Bi, G. Ke, et al., Uni-Fold: An Open-Source Platform for Developing Protein Folding Models beyond AlphaFold [Internet], 2022, https://doi.org/10.1101/2022.08.04.502811 [cited 2025 Apr 18]. Available from:.
[113] J. Yu, Z. Chen, Z. Li, M. Zeng, W. Lin, H. Huang, et al., OpenComplex I: Predicting Three-Dimensional Atomic-Level Structures of Biomolecular Complexes [Internet], 2025, https://doi.org/10.1101/2025.03.25.643589 [cited 2025 Apr 18]. Available from:.
[114] J. Lee, K. Han, J. Kim, H. Yu, Y. Lee, Solvent: A Framework for Protein Folding [Internet]. arXiv [cited 2025 Apr 18]. Available from: http://arxiv.org/abs/2307.04603, 2023.
[115] J.W. Lee, J.H. Won, S. Jeon, Y. Choo, Y. Yeon, J.S. Oh, et al., DeepFold: enhancing protein structure prediction through optimized loss functions, improved template features, and re-optimized energy function, Bioinformatics [Internet] 39 (12) (2023 Dec 1), https://doi.org/10.1093/bioinformatics/btad712 btad712. Available from:. [cited 2025 Apr 18].
[116] M. Hetmann, L. Parigger, H. Sirelkhatim, A. Stern, A. Krassnigg, K. Gruber, et al., Folding the human proteome using BioNeMo: a fused dataset of structural models for machine learning purposes, Sci. Data [Internet] 11 (1) (2024 Jun 6) 591 [cited 2025 Apr 18]. Available from, https://www.nature.com/articles/s41597-024-03403-z.
[117] J.E. King, D.R. Koes, Interpreting forces as deep learning gradients improves quality of predicted protein structures, Biophys. J. [Internet] 123 (17) (2024 Sep) 2730–2739 [cited 2025 Apr 18]. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0006349523041498.
[118] G.J. Huang, T.K. Parry, W.A. McLaughlin, Assessment of the performances of the protein modeling techniques participating in CASP15 using a structure-based functional site prediction approach: ResiRole, Bioengineering [Internet] 10 (12) (2023 Nov 30) 1377 [cited 2025 Apr 18]. Available from: https://www.mdpi.com/2306-5354/10/12/1377.
[119] S.Y. Yang, Pharmacophore modeling and applications in drug discovery: challenges and recent advances, Drug Discov. Today [Internet] 15 (11) (2010 Jun 1) 444–450. Available from: https://www.sciencedirect.com/science/article/pii/S135964461000111X.
[120] T. Kaserer, K.R. Beck, M. Akram, A. Odermatt, D. Schuster, Pharmacophore models and pharmacophore-based virtual screening: concepts and applications exemplified on hydroxysteroid dehydrogenases, Molecules 20 (12) (2015) 22799–22832.
[121] S. Jiang, M. Feher, C. Williams, B. Cole, D.E. Shaw, AutoPH4: an automated method for generating pharmacophore models from protein binding pockets, J. Chem. Inf. Model [Internet] 60 (9) (2020 Sep 28) 4326–4338. Available from, https://doi.org/10.1021/acs.jcim.0c00121.
[122] D. Schaller, D. Šribar, T. Noonan, L. Deng, T.N. Nguyen, S. Pach, et al., Next generation 3D pharmacophore modeling, WIREs Comput. Mol. Sci. [Internet] 10 (4) (2020 Jul 1), https://doi.org/10.1002/wcms.1468 e1468. Available from. [cited 2023 Jul 18].
[123] S. KB, A. Kumari, D. Shetty, E. Fernandes, C. DV, J. Jays, et al., Structure based pharmacophore modelling approach for the design of azaindole derivatives as DprE1 inhibitors for tuberculosis, J. Mol. Graph. Model. [Internet] 101 (2020 Dec 1) 107718. Available from: https://www.sciencedirect.com/science/article/pii/S1093326320305076.
[124] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, et al., The protein data bank, Nucl. Acids Res. [Internet] 28 (1) (2000 Jan 1) 235–242 [cited 2023 Jul 18]. Available from: https://doi.org/10.1093/nar/28.1.235.
[125] Y.C. Martin, M.G. Bures, E.A. Danaher, J. DeLazzer, I. Lico, P.A. Pavlik, A fast new approach to pharmacophore mapping and its application to dopaminergic and benzodiazepine agonists, J. Comput. Aided Mol. Des. 7 (1) (1993 Feb) 83–102.
[126] K. Poptodorov, T. Luu, R.D. Hoffmann, Pharmacophore model generation software tools, Pharmacophores and Pharmacophore Searches 32 (2006) 15–47.
[127] X. Qing, X. Yin Lee, J. De Raeymaeker, J. RH Tame, K. YJ Zhang, M. De Maeyer, et al., Pharmacophore modeling: advances, limitations, and current utility in drug discovery, J. Receptor Ligand Channel Res. [Internet] 7 (null) (2014 Nov 11) 81–92. Available from: https://doi.org/10.2147/JRLCR.S46843.
[128] R. Mannhold, H. Kubinyi, G. Folkers, Pharmacophores and pharmacophore Searches, John Wiley & Sons, 2006.
[129] J. Kirchmair, G. Wolber, C. Laggner, T. Langer, Comparative performance assessment of the conformational model generators omega and catalyst: a large-scale survey on the retrieval of protein-bound ligand conformations, J. Chem. Inf. Model. 46 (4) (2006) 1848–1861.
[130] J. Kirchmair, C. Laggner, G. Wolber, T. Langer, Comparative analysis of protein-bound ligand conformations with respect to catalyst's conformational space subsampling algorithms, J. Chem. Inf. Model. 45 (2) (2005) 422–430.
[131] A. Hospital, J.R. Goñi, M. Orozco, J.L. Gelpí, Molecular dynamics simulations: advances and applications, Adv. Appl. Bioinform. Chem. [Internet] (8) (2015 Dec 12) 37–47. Available from: https://doi.org/10.2147/AABC.S70333.
[132] R.C. Bernardi, M.C.R. Melo, K. Schulten, Enhanced sampling techniques in molecular dynamics simulations of biological systems, Biochim. Biophys. Acta (BBA) - Gen. Subj. [Internet] 1850 (5) (2015 May 1) 872–877. Available from: https://www.sciencedirect.com/science/article/pii/S0304416514003559.
[133] Y. Wang, J.M. Lamim Ribeiro, P. Tiwary, Machine learning approaches for analyzing and enhancing molecular dynamics simulations, Curr. Opin. Struct. Biol. [Internet] 61 (2020 Apr 1) 139–145. Available from: https://www.sciencedirect.com/science/article/pii/S0959440X19301551.
[134] X. Liu, D. Shi, S. Zhou, H. Liu, H. Liu, X. Yao, Molecular dynamics simulations and novel drug discovery, Expert Opin. Drug Discov. [Internet] 13 (1) (2018 Jan 2) 23–37. Available from: https://doi.org/10.1080/17460441.2018.1403419.
[135] Y.I. Yang, Q. Shao, J. Zhang, L. Yang, Y.Q. Gao, Enhanced sampling in molecular dynamics, J. Chem. Phys. [Internet] 151 (7) (2019 Aug 16), https://doi.org/10.1063/1.5109531, 070902. [cited 2023 Aug 23].
[136] P.C.T. Souza, S. Thallmair, P. Conflitti, C. Ramírez-Palacios, R. Alessandri, S. Raniolo, et al., Protein–ligand binding with the coarse-grained Martini model, Nat. Commun. [Internet] 11 (1) (2020 Jul 24) 3714 [cited 2025 Apr 18]. Available from: https://www.nature.com/articles/s41467-020-17437-5.
[137] L. Monticelli, S.K. Kandasamy, X. Periole, R.G. Larson, D.P. Tieleman, S. J. Marrink, The MARTINI coarse-grained force field: extension to proteins, J. Chem. Theor. Comput. [Internet] 4 (5) (2008 May 1) 819–834. Available from: https://doi.org/10.1021/ct700324x.
[138] J. Lee, X. Cheng, J.M. Swails, M.S. Yeom, P.K. Eastman, J.A. Lemkul, et al., CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field, J. Chem. Theory Comput. 12 (1) (2016 Jan 12) 405–413.
[139] S. Feng, S. Park, Y.K. Choi, W. Im, CHARMM-GUI Membrane Builder: past, current, and future developments and applications, J. Chem. Theor. Comput. [Internet] 19 (8) (2023 Apr 25) 2161–2185 [cited 2025 Apr 18];. Available from: https://doi.org/10.1021/acs.jctc.2c01246.
[140] N. Aho, P. Buslaev, A. Jansen, P. Bauer, G. Groenhof, B. Hess, Scalable constant pH molecular dynamics in GROMACS, J. Chem. Theor. Comput. [Internet] 18 (10) (2022 Oct 11) 6148–6160 [cited 2025 Apr 18]. Available from: https://doi.org/10.1021/acs.jctc.2c00516.
[141] A. Singharoy, I. Teo, R. McGreevy, J.E. Stone, J. Zhao, K. Schulten, Molecular dynamics-based refinement and validation for sub-5 Å cryo-electron microscopy maps, eLife [Internet] 5 (2016 Jul 7) [cited 2025 Apr 18]. e16105. Available from: https://elifesciences.org/articles/16105.
[142] F. Ballabio, C. Paissoni, M. Bollati, M. De Rosa, R. Capelli, C. Camilloni, Accurate and efficient SAXS/SANS implementation including solvation layer effects suitable for molecular simulations, J. Chem. Theor. Comput. [Internet] 19 (22) (2023 Nov 28) 8401–8413, https://doi.org/10.1021/acs.jctc.3c00864, cited 2025 Apr 18].
[143] C. Camilloni, A. Cavalli, M. Vendruscolo, Assessment of the use of NMR chemical shifts as replica-averaged structural restraints in molecular dynamics simulations to characterize the dynamics of proteins, J. Phys. Chem. B [Internet] 117 (6) (2013 Feb 14) 1838–1843, https://doi.org/10.1021/jp3106666 [cited 2025 Apr 18].
[144] P.K. Weiner, P.A. Kollman, AMBER: assisted model building with energy refinement. A general program for modeling molecules and their interactions, J. Comput. Chem. [Internet] 2 (3) (1981 Sep 1) 287–303, https://doi.org/10.1002/jcc.540020311 [cited 2025 Feb 17].
[145] T.D. Kühne, M. Iannuzzi, M. Del Ben, V.V. Rybkin, P. Seewald, F. Stein, et al., CP2K: an electronic structure and molecular dynamics software package - Quickstep: efficient and accurate electronic structure calculations, J. Chem. Phys. [Internet] 152 (19) (2020 May 19) 194103, https://doi.org/10.1063/5.0007045 [cited 2025 Feb 17].
[146] A. Smith, X. Dong, V. Raghavan, An overview of molecular dynamics simulation for food products and processes, Processes 10 (1) (2022).
[147] M.R. Challapa-Mamani, E. Tomás-Alvarado, A. Espinoza-Baigorria, D.A. León-Figueroa, R. Sah, A.J. Rodriguez-Morales, et al., Molecular docking and molecular dynamics simulations in related to Leishmania donovani: an update and literature review, Trop. Med. Infect. Dis. 8 (10) (2023).
[148] N. Hansen, W.F. van Gunsteren, Practical aspects of free-energy calculations: a review, J. Chem. Theor. Comput. [Internet] 10 (7) (2014 Jul 8) 2632–2647. Available from, https://doi.org/10.1021/ct500161f.
[149] A. Perez, J.A. Morrone, C. Simmerling, K.A. Dill, Advances in free-energy-based simulations of protein folding and ligand binding, Curr. Opin. Struct. Biol. [Internet] (36) (2016 Feb 1) 25–31. Available from: https://www.sciencedirect.com/science/article/pii/S0959440X15001888.
[150] D.L. Mobley, M.K. Gilson, Predicting binding free energies: frontiers and benchmarks, Annu. Rev. Biophys. [Internet] 46 (1) (2017 May 22) 531–558 [cited 2023 Jul 20]. Available from, https://doi.org/10.1146/annurev-biophys-070816-033654.
[151] R. Abel, L. Wang, E.D. Harder, B.J. Berne, R.A. Friesner, Advancing drug discovery through enhanced free energy calculations, Acc. Chem. Res. [Internet] 50 (7) (2017 Jul 18) 1625–1632. Available from: https://doi.org/10.1021/acs.accounts.7b00083.

[152] Z. Cournia, B. Allen, W. Sherman, Relative binding free energy calculations in drug discovery: recent advances and practical considerations, J. Chem. Inf. Model. 57 (12) (2017 Dec 26) 2911–2937.
[153] M.R. Shirts, D.L. Mobley, J.D. Chodera, Alchemical free energy calculations: ready for prime time? Annu. Rep. Comput. Chem. 3 (2007) 41.
[154] T.S. Lee, Z. Lin, B.K. Allen, C. Lin, B.K. Radak, Y. Tao, et al., Improved alchemical free energy calculations with optimized smoothstep softcore potentials, J. Chem. Theory Comput. 16 (2020) 5512.
[155] E. Wang, H. Sun, J. Wang, Z. Wang, H. Liu, J.Z.H. Zhang, et al., End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design, Chem. Rev. [Internet] 119 (16) (2019 Aug 28) 9478–9508, https://doi.org/10.1021/acs.chemrev.9b00055.
[156] R. Lazim, D. Suh, S. Choi, Advances in molecular dynamics simulations and enhanced sampling methods for the study of protein systems, Int. J. Mol. Sci. 21 (17) (2020).
[157] M. Papadourakis, H. Sinenka, P. Matricon, J. Hénin, G. Brannigan, L. Pérez-Benito, et al., Alchemical free energy calculations on membrane-associated proteins, J. Chem. Theor. Comput. [Internet] 19 (21) (2023 Nov 14) 7437–7458, https://doi.org/10.1021/acs.jctc.3c00365 [cited 2025 Apr 18].
[158] H.M. Baumann, V. Gapsys, B.L. de Groot, D.L. Mobley, Challenges encountered applying equilibrium and nonequilibrium binding free energy calculations, J. Phys. Chem. B [Internet] 125 (17) (2021 May 6) 4241–4261. Available from: https://doi.org/10.1021/acs.jpcb.0c10263.
[159] D.W. Borhani, D.E. Shaw, The future of molecular dynamics simulations in drug discovery, J. Comput. Aid. Mol. Des. [Internet] 26 (1) (2012 Jan 1) 15–26. Available from, https://doi.org/10.1007/s10822-011-9517-y.
[160] O. Khoruzhii, A.G. Donchev, N. Galkin, A. Illarionov, M. Olevanov, V. Ozrin, et al., Application of a polarizable force field to calculations of relative protein–ligand binding affinities, Proc. Natl. Acad. Sci. 105 (30) (2008) 10378–10383.
[161] J.W. Ponder, C. Wu, P. Ren, V.S. Pande, J.D. Chodera, M.J. Schnieders, et al., Current status of the AMOEBA polarizable force field, J. Phys. Chem. B 114 (8) (2010) 2549–2564.
[162] D. Cappel, M.L. Hall, E.B. Lenselink, T. Beuming, J. Qi, J. Bradner, et al., Relative binding free energy calculations applied to protein homology models, J. Chem. Inf. Model. 56 (12) (2016) 2388–2400.
[163] G.A. Ross, E. Russell, Y. Deng, C. Lu, E.D. Harder, R. Abel, et al., Enhancing water sampling in free energy calculations with grand canonical Monte Carlo, J. Chem. Theory Comput. 16 (2020) 6061.
[164] J. Wahl, M. Smieško, Assessing the predictive power of relative binding free energy calculations for test cases involving displacement of binding site water molecules, J. Chem. Inf. Model. 59 (2019) 754.
[165] J. Fan, A. Fu, L. Zhang, Progress in molecular docking, Quant. Biol. (2019) 1–7.
[166] N.S. Pagadala, K. Syed, J. Tuszynski, Software for molecular docking: a review, Biophys. Rev. 9 (2) (2017) 91–102.
[167] C.M. Venkatachalam, X. Jiang, T. Oldfield, M. Waldman, LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites, J. Mol. Graph. Model. [Internet] 21 (4) (2003 Jan 1) 289–307. Available from: https://www.sciencedirect.com/science/article/pii/S109332630200164X.
[168] I.W. Davis, D. Baker, RosettaLigand docking with full ligand and receptor flexibility, J. Mol. Biol. 385 (2) (2009) 381–392.
[169] E.B. Miller, R.B. Murphy, D. Sindhikara, K.W. Borrelli, M.J. Grisewood, F. Ranalli, et al., Reliable and accurate solution to the induced fit docking problem for protein–ligand binding, J. Chem. Theor. Comput. [Internet] 17 (4) (2021 Apr 13) 2630–2639 [cited 2025 Apr 18]. Available from: https://pubs.acs.org/doi/10.1021/acs.jctc.1c00136.
[170] M. Mohanty, P.S. Mohanty, Molecular docking in organic, inorganic, and hybrid systems: a tutorial review, Monatshefte für Chemie - Chemical Monthly [Internet] 154 (7) (2023 Jul 1) 683–707, https://doi.org/10.1007/s00706-023-03076-1.
[171] D.B. Kitchen, H. Decornez, J.R. Furr, J. Bajorath, Docking and scoring in virtual screening for drug discovery: methods and applications, Nat. Rev. Drug Discov. 3 (11) (2004) 935–949.
[172] Z. Orbán-Németh, R. Beveridge, D.M. Hollenstein, E. Rampler, T. Stranzl, O. Hudecz, et al., Structural prediction of protein models using distance restraints derived from cross-linking mass spectrometry data, Nat. Protoc. [Internet] 13 (3) (2018 Mar) 478–494 [cited 2025 Apr 18]. Available from: https://www.nature.com/articles/nprot.2017.146.
[173] S. Katada, T. Hirokawa, Y. Oka, M. Suwa, K. Touhara, Structural basis for a broad but selective ligand spectrum of a mouse olfactory receptor: mapping the odorant-binding site, J. Neurosci. 25 (7) (2005) 1806–1815.
[174] S. Katada, T. Hirokawa, K. Touhara, Exploring the odorant binding site of a G-protein-coupled olfactory receptor, Curr. Comput. Aided Drug Des. 4 (2) (2008) 123–131.
[175] M. Alfonso-Prieto, R. Capelli, Machine Learning-based modeling of olfactory receptors in their inactive state: human OR51E2 as a case study, J. Chem. Inf. Model [Internet] 63 (10) (2023 May 22) 2911–2917. Available from: https://doi.org/10.1021/acs.jcim.3c00380.
[176] L. Pirona, F. Ballabio, M. Alfonso-Prieto, R. Capelli, Calcium-driven in silico inactivation of a human olfactory receptor, J. Chem. Inf. Model [Internet] 64 (8) (2024 Apr 22) 2971–2978. Available from: https://doi.org/10.1021/acs.jcim.4c00249.
[177] R. Achebouche, A. Tromelin, K. Audouze, O. Taboureau, Application of artificial intelligence to decode the relationships between smell, olfactory receptors and small molecules, Sci. Rep. [Internet] 12 (1) (2022) 18817, https://doi.org/10.1038/s41598-022-23176-y.
[178] S.M. Boyle, S. McInally, A. Ray, Expanding the olfactory code by in silico decoding of odor-receptor chemical space. Luo L, editor, eLife [Internet] 2 (2013 Oct 1) e01120, https://doi.org/10.7554/eLife.01120.
[179] L. Gelis, S. Wolf, H. Hatt, E.M. Neuhaus, K. Gerwert, Prediction of a ligand-binding niche within a human olfactory receptor by combining site-directed mutagenesis with dynamic homology modeling, Angew. Chem. Int. Ed. 51 (5) (2012) 1274–1278.
[180] A. Nicoli, F. Haag, P. Marcinek, R. He, J. Kreißl, J. Stein, et al., Modeling the orthosteric binding site of the G protein-coupled odorant receptor OR5K1, J. Chem. Inf. Model. 63 (7) (2023) 2014–2029.