nature biotechnology

Review article

https://doi.org/10.1038/s41587-023-01987-2

Computational drug development for membrane protein targets

Received: 13 January 2023

Accepted: 13 September 2023

Published online: 15 February 2024

Check for updates

Haijian Li[1,5], Xiaolin Sun[1,5], Wenqiang Cui[1,2,5], Marc Xu[1,2,5], Junlin Dong[1,2,5], Babatunde Edukpe Ekundayo[3,5], Dongchun Ni[3,5], Zhili Rao[1,5], Liwei Guo[1,5], Henning Stahlberg[3] ✉, Shuguang Yuan[1] ✉ & Horst Vogel[1,4] ✉

The application of computational biology in drug development for membrane protein targets has experienced a boost from recent developments in deep learning-driven structure prediction, increased speed and resolution of structure elucidation, machine learning structure-based design and the evaluation of big data. Recent protein structure predictions based on machine learning tools have delivered surprisingly reliable results for water-soluble and membrane proteins but have limitations for development of drugs that target membrane proteins. Structural transitions of membrane proteins have a central role during transmembrane signaling and are often influenced by therapeutic compounds. Resolving the structural and functional basis of dynamic transmembrane signaling networks, especially within the native membrane or cellular environment, remains a central challenge for drug development. Tackling this challenge will require an interplay between experimental and computational tools, such as super-resolution optical microscopy for quantification of the molecular interactions of cellular signaling networks and their modulation by potential drugs, cryo-electron microscopy for determination of the structural transitions of proteins in native cell membranes and entire cells, and computational tools for data analysis and prediction of the structure and function of cellular signaling networks, as well as generation of promising drug candidates.

About one-third of human proteins are membrane proteins, with G-protein-coupled receptors (GPCRs) and channel proteins corresponding to the most abundant membrane proteins of the total proteins encoded by the human genome[1,2]. Membrane proteins have an important role in central cellular signal detection and transduction as well as in pathophysiological processes. It is therefore not surprising that the majority of the drugs used clinically to treat severe diseases target membrane proteins[2].

At present, the development of a first-in-class drug molecule typically takes 10–15 years, and the related costs are on the order of a few hundred million US dollars up to $4.5 billion, depending on which research and development (R&D) phases are included in the analysis[3]. Recent developments in the field of computational and structural biology paired with functional bioanalyses promise substantial savings of time and costs in the development of drugs[4].

[1]Center for Computer-Aided Drug Discovery, Faculty of Pharmaceutical Sciences, Shenzhen Institute of Advanced Technology/Chinese Academy of Sciences (SIAT/CAS), Shenzhen, China. [2]University of Chinese Academy of Sciences, Beijing, China. [3]Laboratory of Biological Electron Microscopy, IPHYS, SB, EPFL and Department of Fundamental Microbiology, Faculty of Biology and Medicine, University of Lausanne, Lausanne, Switzerland. [4]Institut des Sciences et Ingénierie Chimiques (ISIC), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. [5]These authors contributed equally: Haijian Li, Xiaolin Sun, Wenqiang Cui, Marc Xu, Junlin Dong, Babatunde Edukpe Ekundayo, Dongchun Ni, Zhili Rao, Liwei Guo. ✉e-mail: henning.stahlberg@epfl.ch; shuguang.yuan@siat.ac.cn; horst.vogel@epfl.ch

https://doi.org/10.1038/s41587-023-01987-2

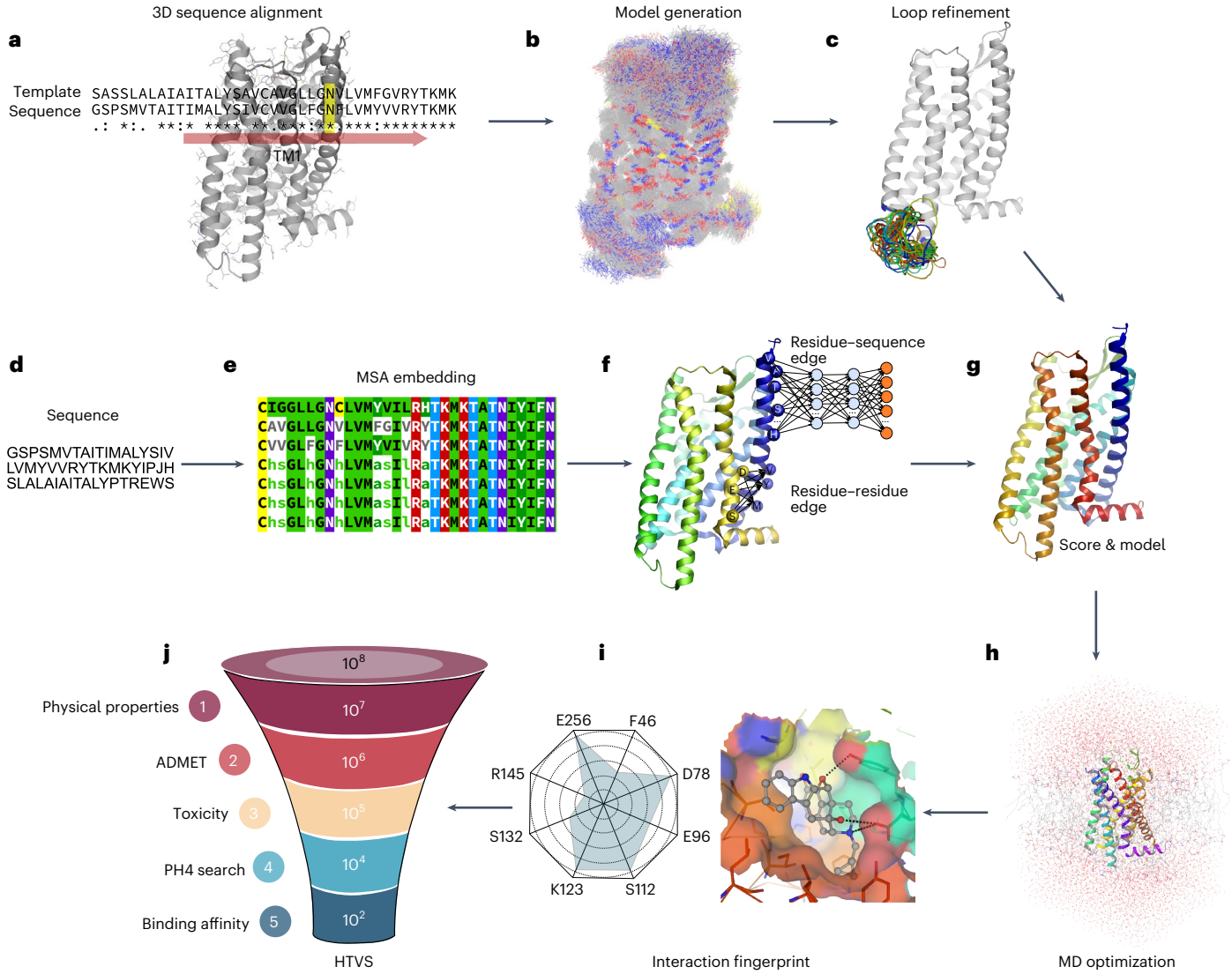

Fig. 1 | Workflow of virtual ligand screening for membrane proteins with a GPCR as a prototypical example. a, Homology modeling based on 3D structure–sequence alignment. **b**, Multiple-model generation. **c**, Refinement of flexible loop regions. **d**, Amino acid sequence of a given target protein. **e**, Homologous protein sequences from jointly embedded MSAs for subsequent structure modeling. **f**, Features for pairs of amino acids delivered from the machine learning tool. The residue–residue edges represent relative distances and relative angles between pairs of amino acids in 3D space. The sequence–residue edges might have sequence evolutionary information, which can be used to predict pair distances. This information is passed to the structure module to build the *xyz* coordinates of the structure. **g**,**h**, Selection of the protein structure model with the best score (**g**) for MD simulation optimization (**h**). **i**, Interaction fingerprint analysis between amino acids of the receptor's binding pocket and the inserted (docked) compound. **j**, Virtual screening for a particular GPCR in the selected compound library composed of up to 10^8 compound drug candidates. Applying stepwise selection filtering on the basis of different properties (1–5) decreases the number of predicted drug candidates successively to the order of about 10^2 compounds. Final experimental functional tests of the predicted compounds for the particular target GPCR reveal the success rate of real hits, which can be as high as 80%[137]. As predictions are not perfect, functional tests are repeated consecutively using computationally predicted, synthetically modified improved compounds until an optimal active compound is reached. HTVS, high-throughput virtual screening.

Computational methods have become increasingly important in modern drug discovery and are now being widely used in the pharmaceutical industry. This is mainly because computational tools can have essential roles in different stages of drug hunting (Fig. 1), including in building three-dimensional (3D) models for the target, virtual screening, prediction of drug-likeness profiles for the candidates and prediction of biological activities.

In this Review, we discuss the different steps in the pipeline during which computational tools accelerate the design of medicines, ranging from structure prediction and prediction of structural transitions to in silico screening of chemical compound libraries to predict highly active ligands and computer-aided planning of chemical syntheses to enlarge the chemical space of bioactive ligands for receptors and channels.

We then discuss how reliable computational biology approaches for drug design are at the moment. This central question can only be answered using experimental tools, which, besides structure elucidation, concerns functional analyses. Here we present a selection of recent developments of bioanalyses for probing the cellular signaling network mediated by membrane proteins such as GPCRs and ion channels, showing the importance of the intrinsic interplay between experimental and computational screening approaches for modern drug development.

Prediction of membrane protein structures

3D structures of biological macromolecules as stored in the Protein Data Bank (PDB)[5] are experimentally determined using nuclear magnetic resonance[6], X-ray crystallography[7] and cryo-electron microscopy

Table 1 | Tools for membrane protein structure prediction

Tool	Description and link	References
Primary sequence alignment tools		
tmAligner	Uses Wu–Manber and dynamic string-matching algorithms for MSA. It is an efficient and template-independent tool to generate MSAs and facilitates rapid classification of uncharacterized transmembrane sequences. http://skuastk.org/tmaligner/	203
Secondary structure prediction		
TMPSS	A deep learning-based tool to predict the secondary structure and folding topology of α-helical membrane proteins. It uses a convolutional neural network with a shared attention mechanism and captures global and local structural features. https://github.com/NENUBioCompute/TMP-SS	27
Tertiary structure prediction and structural databases		
PredMP	A membrane protein 3D structure prediction tool that uses a high-throughput deep learning method. A deep transfer learning model is exploited using non-membrane proteins as a pretrained dataset. The model is not efficient enough in predicting sequences for proteins with a large number of residues. http://www.predmp.com	28
ColabFold	User-friendly platform combining MMSeqs2 with AlphaFold2 or RoseTTAFold for fast prediction of protein folding. Lack of strong evidence in predicting transmembrane protein structures. https://github.com/sokrypton/ColabFold	204
AlphaFold	A deep learning-based method enabling prediction of the overall backbone structure of transmembrane proteins with high accuracy. Predictions are less accurate in structural details concerning the conformation of amino acid side chains, loop regions, the shape of ligand-binding pockets and the conformation of transducer-binding interfaces. https://github.com/deepmind/alphafold	9
RoseTTAFold	A three-track machine learning-based model enabling prediction of the overall backbone structure of transmembrane proteins with acceptable accuracy from amino acid sequence data. Predictions are less accurate in structural details concerning the conformation of amino acid side chains and extramembrane loop regions. https://github.com/RosettaCommons/RoseTTAFold	11,32
MemBrain	Predicts the structure of α-helical transmembrane proteins with high accuracy using multiscale deep learning. Predicting structural details of contact regions between transmembrane helices might be possible in the future. http://www.csbio.sjtu.edu.cn/bioinf/MemBrain/	205
RosettaGPCR	Multiple-structural template homology modeling of GPCRs using RoseTTA. Inactive receptor structures are accurately predicted, facilitating the docking of antagonist molecules. https://github.com/benderb1/rosettagpcr	206
ESMFold	Large language pretrained transformer model to predict the 3D structures of evolutionarily diverse proteins. Benchmarks validating its robustness in predicting transmembrane protein structures are lacking. https://github.com/facebookresearch/esm	24
OmegaFold	Protein sequence-based language model for predicting the structure of orphan proteins. Benchmarks validating its robustness in predicting transmembrane protein structures are lacking. https://hpc.nih.gov/apps/OmegaFold.html	26
MembraneFold	Prediction of 3D membrane protein structures based on topology and sequences of transmembrane regions. https://ku.biolib.com/MembraneFold/	207
Quaternary structure prediction		
AlphaFold-Multimer	AlphaFold extension to predict the structures of complexes made up of multiple proteins. Data for prediction of transmembrane protein assemblies are limited. https://github.com/deepmind/alphafold	10
DeepComplex	Web server for modeling the structures of multiprotein complexes based on deep learning optimization for inter-chain contact and distance data. Benchmarks in predicting transmembrane protein complexes are lacking. http://tulip.rnet.missouri.edu/deepcomplex/web_index.html	45
EquiDock	End-to-end rigid protein–protein docking. Point cloud-based model significantly shortens computational time. https://github.com/octavian-ganea/equidock_public	48
PatchDock	Web server for large-scale protein–protein docking based on protein shape and symmetry complementarity. Applicable to distinguish homogeneous and heterogeneous GPCR assemblies. http://bioinfo3d.cs.tau.ac.il/PatchDock/	43
GRAMM-X	Web server for grid matching protein–protein docking that integrates smoothed Lennard–Jones potential on global search fast Fourier transformation. Applicable for GPCRs. https://gramm.compbio.ku.edu/	208
RosettaDock	Web server for local docking using a Monte Carlo-based algorithm. Backbone flexibility is under consideration. https://rosie.graylab.jhu.edu/	209

(cryo-EM)[8]. Computational structural biology complements and facilitates experimental structure elucidation (Fig. 1). In particular, artificial intelligence (AI)-based methods have substantially improved protein folding prediction[9–11] and predict different structural levels of membrane proteins. AlphaFold2 and RoseTTAFold are two methods for structure prediction based on deep learning architectures, and they deliver surprisingly precise initial structural models for water-soluble proteins[9,11], as well as for membrane proteins[12–14] in spite of the relatively small database of membrane protein structures. Several recent computational methods are complementary to or can be used in combination with AlphaFold2 or RoseTTAFold (Table 1).

The relatively precise prediction of protein structures by AlphaFold2 can be attributed to the application of neural network architectures combined with machine learning based on the evolutionary and physical constraints of protein structures[9]. The core technique of AlphaFold2 is the multiple-sequence alignment (MSA) transformer. AlphaFold2 shows good performance for protein prediction owing to its self-attention capabilities, but it is not the only neural network architecture used for protein prediction. Procedures based on convolutional neural networks, recurrent neural networks and generative adversarial networks have also been used[15]. In general, MSA is the first step for many protein structure prediction tools[16,17].

Using an MSA as input allows for incorporation of information about evolutionary events such as mutations, insertions, deletions and rearrangements to address structural and functional relationships among protein families[9,13,17–21], but challenges remain for proteins for which alignments are not available, such as orphan, de novo-designed and rapidly evolving proteins. Tools such as AlphaFold2 also provide single-sequence prediction possibilities, but prediction accuracy is lower.

The rapid development of large language models (LLMs) fostered the development of single-sequence prediction tools. Recent LLM-based models have advanced protein prediction and provided competitive accuracy when compared to alignment-based approches[22–25]. OmegaFold is a representative tool to achieve end-to-end accurate structure prediction using single primary sequences[26]. An additional example concerns the pretrained model ESMFold developed by Meta, which resulted in comparably accurate tertiary structure prediction[24].

The continuing advancement of deep learning-based protein structure prediction tools such as AlphaFold2 and RoseTTAFold raises the question of whether older tools such as TMPSS[27] and PredMP[28] have become obsolete. AlphaFold's performance is reliable for proteins with known homologs, whereas it is less accurate in predicting the structure of proteins with no prior evolutionary information. Prediction of protein structures that do not rely on evolutionary information might still use tools like TMPSS and PredMP, and the robust prediction power of these tools on membrane proteins still attracts their use[29,30].

Although AlphaFold2 delivers precise overall structures for many proteins, the accuracy of the predicted structural details might not be sufficient for interpretation of protein function and for structure-based drug design. Although AlphaFold2 has released structural models for 98.5% of the human proteome, only 58% of the residues are modeled with high confidence[31]. This is partly due to intrinsically disordered protein regions. The human proteome is predicted to contain 30–50% disordered regions, which often can transform into ordered structures when interacting with other proteins in signaling events. The prediction of how the disordered protein sequences fold remains to be solved[31]. Additionally, the multiple domains that are present in many human proteins are not always located with sufficient accuracy by AlphaFold2 and RoseTTAFold[31].

The quality of the prediction of membrane proteins has been investigated in detail for several membrane proteins[12], including GPCRs[32–35], channel proteins[36,37] and transporters[35]. In general, AlphaFold2 performs well when predicting the overall backbone structure of membrane proteins. However, it is often not sufficiently accurate in details for structure-based drug discovery. For GPCRs, for example, the predicted structural models are usually less accurate when template structures are lacking. When good templates are available, the template-based software Modeller results in smaller modeling root mean square deviation, whereas AlphaFold and RoseTTAFold outperform Modeller when good templates are not available[32]. When AlphaFold2-predicted structures of GPCRs were compared to experimentally determined structures[33], AlphaFold2 could capture the overall backbone features of the receptors very well, but the predicted structures were sometimes less accurate in structural details concerning the shape of the ligand-binding pockets and the conformation of transducer-binding interfaces[33].

AlphaFold2 alone often fails to generate accurate functional alternative conformations when using default parameters. However, this situation could be improved by combining AlphaFold2 with other tools such as sequence clustering[38] and symmetrical docking[39] or biased interference by masking positions in MSAs[40] and introducing state-annotated template databases[41]. AlphaFold2 was shown to achieve accurate backbone conformations for functional alternative conformations of GPCRs and transporters when some procedures in the AlphaFold2 tool were modified on a case-by-case basis[35]. AlphaFold2 in combination with other machine learning tools has also been shown to improve structural modeling of ion channels[37] and an ABC transporter[12]. Although these procedures cannot be generalized yet, further improvements will continue to be made in this field.

The prediction of interfaces between subunits in multimolecular protein assemblies is of fundamental importance to model protein–protein interaction networks. Conventional approaches regard quaternary structure prediction as a protein–protein docking task accomplished by searching for global binding modes (MDockPP)[42] or shape complementarity (PatchDock)[43] or by using an experimental data-driven approach (HADDOCK)[44]. Recent years have seen the emergence of user-friendly tools based on modern machine learning architectures to predict protein–protein binding interfaces and inter-chain contacts[45]. AlphaFold2 and its variant, AlphaFold-Multimer, can provide large-scale predictions of the structures of multimeric protein complexes with acceptable quality when using optimized MSAs as input[46,47]. Furthermore, EquiDock, a deep learning model based on a graph-matching network, docks interacting proteins as rigid bodies into a final model of a multiprotein complex[48]. Although the prediction of multiprotein complexes has been focused on aqueous proteins, there is no limitation to also using these approaches for membrane proteins once sufficient experimental structural data are available.

AlphaFold2 has been used to predict the quaternary structure of membrane proteins for a monotopic membrane protein[49], and, despite the lack of close homologs, AlphaFold2 correctly predicted the protein folds of oligomeric assemblies at a granular level. To further improve the prediction of multimeric protein complexes, Gao et al. extended AlphaFold2 and developed AF2Complex[50]. The overall performance of this method was evaluated on the multimeric targets of CASP14 and on the membrane protein system involved in cytochrome *c* biogenesis. The results suggest that AF2Complex can obtain substantially better predictions than either some complex docking-centric approaches or the AlphaFold-Multimer tool. The authors demonstrated that network models developed for single proteins can be adapted to predict multimeric protein structures without retraining[50].

Most of the different computational tools discussed here have been used at different levels of the rational drug design pipeline. The future will show whether the recent progress of AlphaFold2 and RoseTTAFold in predicting protein structures will be a game-changer in structure-based drug development. All the computational tools rely on a large amount of high-quality input data, which are limited for membrane proteins. The following section discusses recent developments in structure determination for membrane proteins and how the mutual influences between experimental and computational tools have contributed to the recent progress in structure membrane biology.

AI-based structure determination from X-ray diffraction or cryo-EM data in combination with machine learning

As an alternative to computational analysis, X-ray crystallography and cryo-EM are the methods of choice for experiment-based structural analysis of membrane proteins[51,52]. Cryo-EM is the dominant method for membrane protein structure determination but is limited when studying proteins of small size or of lower stability. It has a lower size limit of approximately 60 kDa, which is often outside the size range of a single GPCR (for example, approximately 35 kDa for a class A GPCR). However, the structure of GPCRs can be determined by cryo-EM if the size of the particles is increased by binding to nanobodies, G proteins or arrestin[53–56]. Cryo-EM is very well suited to study ion channels, many of which form oligomeric assemblies of larger than 100 kDa[57]. In contrast, X-ray crystallography is not restricted to analysis of proteins above a lower size limit but has a major bottleneck from difficulties obtaining well-diffracting 3D crystals of the target protein. Therefore, elucidating the different structural states of a difficult target protein such as a GPCR can necessitate combination of these two experimental approaches.

https://doi.org/10.1038/s41587-023-01987-2

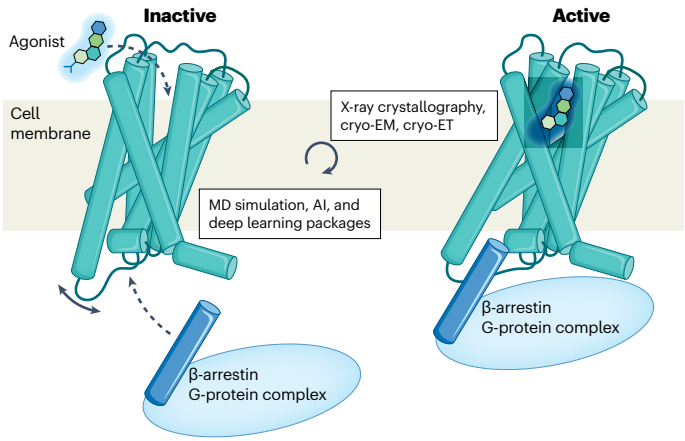

Fig. 2 | An integrated structural biology approach. Combining experimental structure determinations with computational and AI-assisted software will enable researchers to capture the entire signaling cascade of GPCRs from the inactive state to G-protein complex or β-arrestin-bound states and the intermediate structural transitions.

Table 2 | Number of experimentally determined GPCR structures

GPCR type	X-ray	Cryo-EM
Class A	50 (225)	169 (3)
Class B1	0 (14)	72 (2)
Class B2	0 (0)	9 (0)
Class C	0 (9)	15 (35)
Class F	0 (10)	6 (1)
Total	50 (258)	272 (41)

Data were sourced from the GPCRdb without accounting for unique GPCR structures[63]. The number of active-state structures is indicated for each GPCR class, with the number of inactive-state structures in parentheses.

The structural states of GPCRs in their signaling cascades involve structural transitions from an inactive apo state to a ligand-bound activated state, and the subsequent consecutive transmembrane conformational changes finally lead to the binding and activation of heterotrimeric G proteins or the binding of arrestin at the intracellular side of the receptor (Fig. 2)[58–62].

X-ray crystallography has been a robust tool for resolving the structures of GPCRs in their inactive state, accounting for 258 of the 299 deposited structures of inactive-state human GPCRs. In contrast, cryo-EM has been the primary tool for determination of active-state structures, accounting for 272 of the 322 deposited structures of active-state human GPCRs. Of the 41 inactive-state structures determined by cryo-EM, 35 are class C GPCRs that harbor an extracellular Venus fly trap domain and exist as dimers or oligomers, therefore having a much larger size than a typical class A GPCR (Table 2)[63]. In comparison, cryo-EM accounts for 117 of the 222 available structures of ligand-gated ion channels deposited since the release of the TRPV1 structure in 2013, the first-ever ion channel structure determined by cryo-EM to resolve side chains[64,65].

X-ray crystallography generally restricts the target protein to a single homogenous conformation. In only a few exceptional cases, ligand-gated activation of a receptor might be observed by using photoactivated ligands or by diffusing activating compounds into the existing crystal. Crystallization also often occurs under nonphysiological buffer conditions. Therefore, X-ray crystallography might not deliver the most physiologically relevant structure. In contrast, cryo-EM, which relies on a vitrified aqueous sample that contains target protein particles without any need of crystallization, can capture the landscape of different structural states from a single dataset. In addition, structures of a target protein under varying environmental conditions can be easily achieved from multiple cryo-EM datasets (for example, when the samples are quick-frozen from varying temperatures or under different salt concentrations or pH values). This is particularly interesting for ion channels, which undergo structural changes regulated by physiological agents and ligand binding, resulting in the channels opening and closing[66–71].

A remarkable example comes from studies of the TRPV1 ion channel, for which 25 cryo-EM structures were obtained in a series of snapshot images covering the dynamic structural transitions of the selectivity filter in response to changes in physiological agents, thereby revealing insights into our sensations in pain, heat and taste[72].

Another example is the structure of the mechanosensitive Piezo1 channel in lipid vesicles resolved by cryo-electron tomography (cryo-ET)[73]. Piezo channels are mechanically activated cation channels that act as versatile mechano-transducers in various cell types. Cryo-EM studies directly visualized the substantial deformability of the Piezo1–lipid bilayer system, showing an expansion of 300 nm^2 between a curved and flattened structure in the membrane plane. This study provided a general method to investigate single ion channels in closed 10- to 30-nm-sized lipid bilayer vesicles, opening the possibility to investigate channels under the influence of transmembrane potential. The list of experimentally determined ion channel structures in Table 3 shows the increasing importance of cryo-EM in this field, especially for large channel proteins.

Cryo-EM has benefitted from the development of algorithms such as 3D variability analysis (3DVA) implemented in the cryo-EM data processing software CryoSPARC and from deep learning tools such as CryoDRGN that have enabled the modeling of continuous heterogeneity from a single cryo-EM dataset of single-protein particles[74–76]. To sample the conformational space of particles obtained from a cryo-EM dataset, the 3DVA algorithm generates reaction coordinates (K) by computing the eigenvectors of the 3D covariance of a set of particle images and then positions each particle along each eigenvector[76]. 3DVA is already being used to study GPCRs. A notable example comes from the analysis of cryo-EM data collected on the cannabinoid–G_i and secretin–G_s receptor complexes, which uncovered consensus structural movements, despite the differences in biochemical composition[77]. CryoDRGN is a generative deep learning approach for analyzing structural heterogeneity. It encodes two-dimensional cryo-EM projection images into a latent space with the help of a variational autoencoder, from which slices or series of 3D reconstructions are decoded. CryoDRGN thereby evaluates the continuous 3D structural heterogeneity in the proteins imaged in the cryo-EM dataset. This rapidly advancing application of neural network computing shows superior performance over traditional algorithms, especially for small datasets of particles that are in heterogeneous conformations, as exemplified by cryoDRGN2 (ref. 78). It is foreseeable that further R&D in machine learning algorithms will make the analysis of 3D structures from cryo-EM data of heterogenous particle populations routine, to maximize the mechanistic information extracted from the image data.

The implementation of molecular dynamics (MD) simulation to study the structural transitions of GPCRs and ion channels has gained popularity over the last decade[79,80]. MD simulation uses the available limited structural data as starting input models for calculation of long- or short-lived structural transitions. MD simulation can provide femtosecond-resolution information on the structural movements of a membrane protein following environmental changes such as changes to temperature, pH and membrane lipid composition, complementing experimental data[79–81]. Such information is currently inaccessible to time-resolved structure determination techniques

Table 3 | Number of experimentally determined ion channel structures

Ion channel type	X-ray	Cryo-EM
Mechanosensitive	1	14
Potassium, sodium and proton selective	26	23
Calcium ion selective	0	4
Transient receptor potential (TRP)	20	229
Gap junctions and related channels	1	13
Fluc family (fluoride-selective channels)	15	0
Formate/nitrite transporter (FNT) family	8	0
Cys-loop receptor family	75	11
Glutamate receptor family	43	126

Data were sourced from the mpstruc database[210].

such as time-resolved cryo-EM, which is limited to the microsecond to millisecond range of resolution[82,83].

The combination of experimentally determined structures with MD simulation often results in captivating movies of the structural dynamics of GPCRs in action (Fig. 2). These data are now being deposited on the online platform GPCRmd, enabling user-friendly access to these movies[84]. However, how GPCRs organize in the cell during the different steps of signaling cascades remains to be discovered. GPCRs can form homo- and hetero-oligomers, which modulates their function in the cell[85,86]. Optical super-resolution imaging techniques have observed these oligomeric states at nanoscale resolution[87,88]. Recent improvements in cryo-ET for in situ structural biology now enable structure determination for membrane proteins on a resolution scale of sub-nanometers to a few Angstroms directly in the native membrane by sub-tomogram averaging and single-particle imaging[89–92]. This development, combined with computational tools including AI-assisted software packages such as AI-Tom for data processing, particularly in template matching to identify target molecules from the noisy tomograms, holds the promise of visualizing the oligomeric states of GPCRs in the cell at a resolution of single-digit Angstroms, opening new avenues for the structural biology of GPCRs and ion channels in the future[93]. In X-ray crystallography, near-accurate structure models generated from AI tools such as AlphaFold2 provide suitable templates for phasing diffraction data by molecular replacement, particularly in challenging cases[94]. Therefore, computational tools can substantially reduce the necessity for experimental phasing techniques and accelerate model building in the determination of protein structures, including for GPCRs and ion channels.

The integration of computational and machine learning tools in experimental structure determination enables researchers to generate more accurate models of membrane proteins in distinct functionally relevant states that will facilitate virtual screening approaches for drug discovery.

Virtual screening of large compound libraries for active ligands targeting membrane proteins

Based on rich compound libraries and the development of automated technologies, experimental high-throughput screening (HTS) has a critical role in drug development and already accounts for approximately 15% of drug development costs[95]. Virtual screening complements HTS by screening large compound databases or virtual compound libraries in silico and selecting a limited number of compounds for testing. Virtual screening is based on docking[96] or pharmacophore[97] modeling of the target protein structure with compounds to filter out expected candidates on the basis of their biophysical properties[98]. Since the introduction of computer-aided drug discovery (CADD) technology, more than 70 CADD-developed drugs have been approved for marketing, including several targeting GPCRs[99]. Virtual HTS improves the efficiency of subsequent functional screening and thus reduces the costs for R&D of a new drug.

Traditional drug development programs have often focused on single-target drugs and have obtained numerous high-selectivity drugs as a result. However, many diseases (neurodegenerative diseases, cardiovascular diseases and oncological diseases) involve multiple pathogenic factors and a single drug for a single target is often not sufficient for treatment. In such cases, multi-medicine therapy and multi-compound medication are used[100], but these treatments often require special attention to the drugs' side effects and drug interactions. In silico screening of ultra-large compound libraries has provided important contributions[101,102] to the development of multi-targeted and other challenging drugs.

For example, a multi-target drug was designed by virtual screening of large libraries for drugs for Parkinson's disease that could act with improved efficacy on both D2 dopamine receptors and A_{2A} adenosine receptors, two prototypical GPCRs[103]. Among the ten compounds that were identified by virtual screening, three showed affinity for both targets. Virtual screening provided a molecular scaffold, resulting in an optimized compound with nanomolar-range dual-target activity, as shown through analysis of pharmacodynamics in a rat model.

Other studies have focused on other GPCRs. One study looked to develop drugs with fewer side effects that more selectively target cellular pathways in the central nervous system[104] by performing an ultra-large library docking screen and identified 17 compounds targeting both α_{2A}- and α_{2B}-adrenergic receptors.

Another example comes from the development of drugs to alleviate sleep disorders and depression by targeting a certain melatonin receptor subtype known to regulate circadian rhythm[105]. In total, 150 million molecules with favorable physical properties were screened and the 300,000 top-scoring molecules were clustered according to topological similarity to retain diversity. Ultimately, 38 molecules were chemically synthesized, of which 15 showed activity in functional assays.

Virtual HTS has also stimulated the development of improved ion channel drugs, for example, ones targeting the sodium-activated potassium channel subunit $K_{Na}1.1$ to treat certain forms of epilepsy[106]. By comparing structural similarity to potent $K_{Na}1.1$–inhibitor complexes, a virtual screening approach delivered six compounds that inhibited $K_{Na}1.1$ by at least 40% in functional assays.

Tafluprost, used in the treatment of glaucoma, is a prostaglandin-like agonist compound that binds the prostaglandin $F_{2\alpha}$ receptor. A recent computational study for drug repurposing reported that tafluprost could act as a putative antiarrhythmic drug through both $Na_V1.5$ and TASK-1 (ref. 107). The study developed an algorithm for the deep exploration of similar 3D structure and sequence patterns between different proteins to find polypharmacological profiles of currently used drugs. The cardiac ion channels $Na_V1.5$ and TASK-1 were identified as prototypical examples to bind a common agonist, tafluprost.

In addition to screening existing chemical compound databanks, researchers have also attempted to expand the virtual compound database. For example, 170 million compounds involved in 130 well-characterized reactions for structural docking were selected to obtain a new library of scaffolds, which would otherwise be unavailable[108]. Each molecule of this library was docked computationally against the AmpC β-lactamase and the D4 dopamine receptor to discover the best noncovalent AmpC inhibitor available. Compounds with the highest score were then synthesized and tested experimentally for their interaction with AmpC and the dopamine receptor. In addition, novel chemotypes with sub-nanomolar affinity for dopamine receptor were identified.

It is also important to develop virtual screening tools with a faster screening speed. An open-source platform called VirtualFlow enables the docking of 1 billion compounds in just 2 weeks using 10,000

CPU cores[109]. In a demonstration, the VirtualFlow software for Virtual Screening (VFVS) model was used to screen 1.3 billion compounds for compounds targeting KEAP1. This was completed in 4 weeks using 8,000 CPU cores, resulting in compound iKeap1 with nanomolar activity. Another method called V-SYNTHES can increase the speed of docking by more than 5,000 times by using an iterative docking approach, and the hit rate of this method is twice the rate of traditional virtual screening of large compound libraries[110].

With the identification of an increasing number of high-resolution structures for GPCRs, virtual screening is becoming an important tool for the development of GPCR drugs (Table 4)[111].

Traditional docking programs leverage the geometric and electrostatic complementarity between protein–protein or protein–ligand pairs to quantify interactions[112]. With the help of deep learning, many powerful new tools have been developed for virtual screening and have been combined with traditional CADD methods. These include DiffDock[113], GNINA[114], PyRMD[115] and others[116,117]. The main challenge for traditional methods is a reasonable computing time when considering molecular flexibility and reflecting real binding. Machine learning can optimize scoring functions and accelerate identification of potential binding sites[118]. Despite the wide applications of traditional docking methods, new molecular docking approaches that integrate machine learning algorithms are rapidly evolving[118]. A recent comparative study evaluated the performance of these deep learning-based approaches and traditional docking in blind docking, molecular docking and ligand-binding pocket searching[119]. The study demonstrates that current deep learning-aided methods perform better in ligand-binding pocket searching and end-to-end blind docking. However, traditional approaches still outperform the deep learning models at docking on given ligand-binding pockets.

Computer-aided chemical synthesis to enlarge the chemical space of bioactive ligands for receptors and channels

As the throughput of virtual screening tools continues to grow, the demand for construction of libraries exploring a larger part of the chemical space is also increasing. Developing computer programs that design multistep chemical synthetic routes substantially different from prior approaches would not only boost general synthetic creativity but also be of practical value. This rationale has been widely accepted and integrated into commercialized software packages, including Chemaxon and Chematica, as a basis for modern drug design[120]. However, the demand to use the enormously expanding chemical space for synthesizing new drugs cannot be fulfilled by simply enumerating building blocks via acknowledged synthetic chemistry rules. With the accumulation of computational power and algorithms, machine learning techniques have emerged to assist chemical reaction route planning[121–125]. These new approaches are ideally suited for predicting high-throughput chemical reactions, including innovative synthesis routes[126].

Here retrosynthetic toolkits have an important role and are primarily used to discover the optimal routes for recursively disconnecting a target molecule until commercially available building blocks are reached[125,127]. The applied methodology has undergone substantial development from natural language processing to Monte Carlo tree searching to achieve predictions for the simplest reagents in each synthesis step[128,129]. For example, exploiting the transformer architecture, Lin et al. have reported 63% predictive accuracy for synthesis of the US Patent and Trademark Office chemical compound library and also used this approach for the synthesis of several complex bioactive compounds[128], which have been decomposed into valid and accessible fragments. Applications have been expanded toward prediction of reaction outcomes[129,130] and optimization of reaction conditions[131,132]. The optimal performance when predicting reaction products exhibited 71.8% rank 1 accuracy depending on the quality of the templates and the awareness of reaction types[133].

Table 4 | GPCR and ion channel drug candidates obtained by virtual screening

Target	Screened compounds (millions)	Tested compounds	Hits (≤10 μM)	Hit rate (%)
A_{2A} adenosine receptor	4.3	56	23	41[211]
A_{2A} adenosine receptor	0.01	10	3	30[103]
α_{2A}-adrenergic receptor	301	48	30	63[104]
β_2-adrenergic receptor	3.7	19	8	42[212]
CXCR4 receptor	2.4	6	3	50[213]
D3 dopamine receptor	4.1	25	14	56[214]
D4 dopamine receptor	138	124	32	26[108]
H1 histamine receptor	0.11	33	15	45[215]
$K_{Na}1.1$ potassium channel	0.1	17	6	35[106]
κ-opioid receptor	4.5	43	14	33[216]
Melatonin receptor	150	38	15	39[105]
$Na_V1.5$ cardiac ion channel	11.3	21	2	9.5[107]
Neurotensin receptor 1	0.5	25	8	32[217]
Olfactory receptor Olfr73	1.6	25	17	68[137]
Orexin receptor 2	11.3	43	11	26[218]
P2X7 receptor	0.1	73	3	4.1[219]
SMO receptor	3.2	21	4	19[220]
TRPA1 channel	0.01	6	2	33.3[221]
5-HT_{1B} receptor	1.3	22	11	50[222]

With the growing demand for molecular complexity and drug selectivity, computational strategies need to fulfill all requirements by generating chemically valid, purchasable and selective reactants. Predicting the synthesis of complex drug candidates while simultaneously considering chirality and stereoselectivity is still challenging[128,134]. Examples include rufinamide, which is supposed to modulate the activity of a voltage-sensitive sodium channel[128,135], and a novel derivative of benzopyran sulfonamide targeting a GPCR, the 5-HT_6 receptor[128,136]. The problems in this field of medicinal chemistry arise from the diversity of the potential bioactive compounds, which at present cannot be integrated into established types of reaction rules. Further, the lack of corresponding openly accessible datasets is a major hurdle when developing and evaluating target-specific reaction pathways in academic research[136]. Nonetheless, it can be expected that, with improvement in predictive accuracy, computational methods might emphasize and reweight specific types of reactions, for instance, Suzuki coupling in concatenating aromatic rings and protection/deprotection groups, thus promoting the systematic drug–target interaction synthetic relationship. With the advent of reaction route planning machine learning toolboxes, chemists can now concentrate more on what kind of molecules they want to synthesize rather than how to make them[121].

Probing cellular signaling mediated by GPCRs and channel proteins

The question of how reliable the different computational tools are for finding new medicines can only be answered by a functional analysis of computer-predicted drug candidates. In this context, modern bioanalytical techniques have a central role in the CADD chain[137–139]. Here we present a selection of recent experimental and computational developments from probing the cellular signaling network mediated by membrane proteins such as GPCRs and ion channels.

GPCR signaling

The human genome encodes more than 800 GPCRs, representing the largest family of human membrane proteins[140]. GPCRs have an important role in central cellular signal detection and transduction[141] as well as in pathophysiological processes, which explains their distinguished importance as drug targets[142–147].

The seven-transmembrane-domain helical receptors detect physical (photons, temperature, mechanical stimuli)[148–150] and chemical (odors, hormones, growth factors)[142] extracellular signals to initiate transmembrane signal transduction processes, which finally evoke intracellular responses (Fig. 2).

Bioassays are of central importance for probing in silico-designed compounds for their ability to influence specific reactions of the GPCR signaling pathways depicted in Fig. 2. In the following, recent trends in bioassay development will be discussed according to the biological function tested and the technology used to measure this function.

The considerable progress made in this field relies primarily on developments to selectively label proteins in living cells and on super-resolution microscopy techniques capable of detecting single molecules and molecular complexes with nanometer localization precision[151–153].

Most functional assays use fluorescent or bioluminescent probes, which are introduced genetically or post-translationally into the target protein (a GPCR, G protein, arrestin or downstream signaling proteins). Ligand-induced structural transitions in or interactions of the target proteins are detected as photophysical changes in single optical probes or between a specific couple of interacting probes within fluorescence or bioluminescence resonance energy transfer measurements[154,155].

These strategies have been applied to study GPCR oligomerization[85,86,156], GPCR activation[157,158], GPCR–G protein interactions[143,144,159] and activation of G proteins[160,161] and β-arrestin[162,163] in live cells, isolated plasma membranes and purified systems.

At present, most bioassays probing GPCR signaling are based on ensemble measurements, which have intrinsic limitations in resolving the conformational transitions of signaling proteins. Here single-molecule microscopy techniques offer complementary advancement, as they can resolve distinct processes of the GPCR signaling network that are hidden when measuring averaged properties in ensembles. Examples include the detection of a complex mobility distribution pattern for individual GPCRs in live cells[164–167] and resolving the distinct conformational functional states of GPCRs and G proteins[168–172].

Ion channel signaling

Unlike GPCRs, which are structurally well-defined seven-transmembrane-domain receptors, ion channels constitute a heterogeneous family of proteins and pharmacological targets. About 300 different ion channels have been identified in humans, which are classified according to gating mechanism, ion selectivity or cellular localization[173]. Currently, over 160 drugs are on the market targeting ion channels and about 50 further drugs are in preclinical development or clinical trials[174].

To probe the function of ion channels, the patch-clamp technique is the gold standard in electrophysiology research[175]. However, traditional patch-clamp methods suffer from low throughput. To overcome this limitation for drug screening, automatic patch-clamp (APC) techniques have been developed, replacing the traditional patch-clamp glass pipette with a planar glass slide comprising single- or multiple-micrometer-sized apertures. Heterologous or primary cells expressing the ion channel of interest are placed over the aperture to form high-resistance electrical seals. APC techniques allow the characterization of cell-based ion channels in high-throughput format[176,177]. In addition, purified ion channels have been reconstituted into freestanding planar lipid membranes for electrophysiological characterization under defined conditions outside of the complex cellular environment[178].

Ion channel function can also be measured indirectly using fluorescent probes incorporated at specific sites of the channel protein. For example, simultaneous observation of the electrical current during channel gating and the fluorescence of environment-sensitive probes can resolve major questions about the conformational transitions ion channels perform during activation, drug application and desensitization[179,180]. Using cell-derived plasma membrane vesicles, ligand binding-induced conformational changes have been observed through single-molecule imaging in a cyclic nucleotide-gated ion channel[181]. In addition, optical technologies based on measuring transmembrane potential with fast voltage-sensitive dyes have attracted interest for the study of drugs acting on ligand-gated ion channels[182]. Further, genetically encoded microbial rhodopsins can be used as fast fluorescent voltage indicators to measure neuronal activity when investigating fundamental principles of brain function and eventually the action of certain drugs for treatment of neurological disorders[183].

Computational tools for the analysis of experimental functional assays

As outlined elsewhere, computational methods have been used to analyze membrane dynamics data collected through single-molecule tracking[184]. Bayesian-based approaches along with machine and deep learning tools have been used to classify and characterize the modes of diffusion of different receptor subpopulations[185–189]. Application of these analyses revealed a complex heterogenous pattern of transient Brownian and restricted receptor mobilities, which were modulated differently by distinct ligands[164–166,190]. It is the combination of experimental and computational approaches that finally resolves how the dynamic state of a membrane protein is linked to its function. This might lead to the development of a new generation of drugs targeting specific receptor states.

In addition, numerous computational studies have been performed to simulate the experimentally determined activities of various ion channel proteins[37]. A combination of machine learning or deep learning approaches and molecular simulation strategies has been used to understand the molecular mechanism of channel activation and to predict potential ligands that could modulate channel activity allosterically and might act as potential drugs[37,191–193].

Super-resolution optical microscopy techniques have further resolved the nonuniform lateral distribution of cellular components and signaling reactions with the help of deep learning approaches[194–197]. In a complex interplay, lipids and proteins influence their heterogenous lateral distribution, leading to the formation of membrane nanodomains with important implications for many cellular processes such as membrane fusion, protein trafficking and signal transduction, which could be targeted by modern medicines[198]. Computational methods, in particular MD simulation, along with enhanced sampling techniques such as metadynamics provide important insight into the experimentally resolved organizational principles of cell membranes[199].

Conclusion

Computational and machine learning approaches promise to accelerate drug discovery tremendously. These promises are fostered by the ever-increasing number of success stories in various fields of CADD campaigns. The most prominent examples are the structure-based virtual screening of ultra-large chemical libraries and the recent prediction of protein structures from their amino acid sequences based on deep learning approaches such as AlphaFold. This is mirrored by the growing number of articles reporting on the application of AI methods for drug development[200].

It is worth noting that the reliability of an AI tool or algorithm depends on both a large amount and a high quality of data. If either is missing, the predictions from AI are less reliable. For example, AlphaFold2 can accurately predict the structure of water-soluble proteins owing to the large number of experimentally determined structures

available for the AI training phase. In contrast, for prediction of membrane protein structures, AlphaFold2 is often not sufficiently accurate in details for structure-based drug discovery owing to the limited number of experimentally determined structures, especially for structural transitions between different functional states of membrane proteins involved in transmembrane signaling. Here the accuracy of AI structure predictions would benefit from progress in experimental structural biology.

Further, an increasing number of new startup companies, as well as established digital giants (Google, Microsoft, Amazon, IBM and Meta), are entering the field of 'digital pharma' by using various applications of AI to accelerate the development of drugs[201]. In addition, traditional pharmaceutical companies are moving into the field of data-driven drug development by building their own in-house machine learning-driven R&D and collaborating with new biotechnology companies. The applications of these exciting developments for membrane proteins are still at a very early phase. There is room for substantial improvement to overcome current limitations.

A major limitation in experimental approaches is structure determination for small membrane proteins (<40 kDa) such as class A GPCRs in the inactive state without the bound G-protein complex. Although AI tools can be used to generate models of GPCRs in the inactive state, thorough benchmarking of the accuracy of these models and accurate models for machine learning would require more experimentally determined models. The limitations of obtaining suitable crystals for structure determination by X-ray crystallography and low signal-to-noise ratios in cryo-EM are important bottlenecks to be addressed in the future. Developing AI tools for denoising electron micrographs in cryo-EM could be one strategy to overcome this bottleneck[202]. However, further research in electron microscopy hardware development is also necessary. The combination of computational and hardware development, which has triggered the resolution revolution in cryo-EM, could trigger another resolution revolution in the determination of structures for small membrane proteins, which would greatly accelerate virtual screening drug discovery approaches in the future.

References

1. Fagerberg, L., Jonasson, K., von Heijne, G., Uhlen, M. & Berglund, L. Prediction of the human membrane proteome. *Proteomics* **10**, 1141–1149 (2010).
2. Gong, J. et al. Understanding membrane protein drug targets in computational perspective. *Curr. Drug Targets* **20**, 551–564 (2019).
3. Schlander, M., Hernandez-Villafuerte, K., Cheng, C. Y., Mestre-Ferrandiz, J. & Baumann, M. How much does it cost to research and develop a new drug? A systematic review and assessment. *Pharmacoeconomics* **39**, 1243–1269 (2021).
4. Vemula, D., Jayasurya, P., Sushmitha, V., Kumar, Y. N. & Bhandari, V. CADD, AI and ML in drug discovery: a comprehensive review. *Eur. J. Pharm. Sci.* **181**, 106324 (2022).
5. Burley, S. K. et al. RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. *Nucleic Acids Res.* **49**, D437–D451 (2021).
6. Purslow, J. A., Khatiwada, B., Bayro, M. J. & Venditti, V. NMR methods for structural characterization of protein–protein complexes. *Front. Mol. Biosci.* **7**, 9 (2020).
7. Li, F. et al. Highlighting membrane protein structure and function: a celebration of the Protein Data Bank. *J. Biol. Chem.* **296**, 100557 (2021).
8. Garcia-Nafria, J. & Tate, C. G. Cryo-electron microscopy: moving beyond X-ray crystal structures for drug receptors and drug development. *Annu. Rev. Pharmacol. Toxicol.* **60**, 51–71 (2020).
9. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
10. Evans, R. et al. Protein complex prediction with AlphaFold-Multimer. Preprint at *bioRxiv* https://doi.org/10.1101/2021.10.04.463034 (2022).
11. Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
12. Hegedus, T., Geisler, M., Lukacs, G. L. & Farkas, B. Ins and outs of AlphaFold2 transmembrane protein structure predictions. *Cell. Mol. Life Sci.* **79**, 73 (2022).
13. Tunyasuvunakool, K. et al. Highly accurate protein structure prediction for the human proteome. *Nature* **596**, 590–596 (2021).
14. Lupas, A. N. et al. The breakthrough in protein structure prediction. *Biochem. J.* **478**, 1885–1890 (2021).
15. Pakhrin, S. C., Shrestha, B., Adhikari, B. & Kc, D. B. Deep learning-based advances in protein structure prediction. *Int. J. Mol. Sci.* **22**, 5553 (2021).
16. Senior, A. W. et al. Improved protein structure prediction using potentials from deep learning. *Nature* **577**, 706–710 (2020).
17. Yang, J. et al. Improved protein structure prediction using predicted interresidue orientations. *Proc. Natl Acad. Sci. USA* **117**, 1496–1503 (2020).
18. Ju, F. et al. CopulaNet: learning residue co-evolution directly from multiple sequence alignment for protein structure prediction. *Nat. Commun.* **12**, 2535 (2021).
19. Zheng, W. et al. Deep-learning contact-map guided protein structure prediction in CASP13. *Proteins Struct. Funct. Bioinf.* **87**, 1149–1164 (2019).
20. Xu, J., Mcpartlon, M. & Li, J. Improved protein structure prediction by deep learning irrespective of co-evolution information. *Nat. Mach. Intell.* **3**, 601–609 (2021).
21. Lupo, U., Sgarbossa, D. & Bitbol, A.-F. Protein language models trained on multiple sequence alignments learn phylogenetic relationships. *Nat. Commun.* **13**, 6298 (2022).
22. Seo, S., Oh, M., Park, Y. & Kim, S. DeepFam: deep learning based alignment-free method for protein family modeling and prediction. *Bioinformatics* **34**, i254–i262 (2018).
23. Chowdhury, R. et al. Single-sequence protein structure prediction using a language model and deep learning. *Nat. Biotechnol.* **40**, 1617–1623 (2022).
24. Lin, Z. et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* **379**, 1123–1130 (2023).
25. Weissenow, K., Heinzinger, M. & Rost, B. Protein language-model embeddings for fast, accurate, and alignment-free protein structure prediction. *Structure* **30**, 1169–1177 (2022).
26. Wu, R. et al. High-resolution de novo structure prediction from primary sequence. Preprint at *bioRxiv* https://doi.org/10.1101/2022.07.21.500999 (2022).
27. Liu, Z. et al. TMPSS: a deep learning-based predictor for secondary structure and topology structure prediction of α-helical transmembrane proteins. *Front. Bioeng. Biotechnol.* **8**, 629937 (2020).
28. Wang, S. et al. PredMP: a web server for de novo prediction and visualization of membrane proteins. *Bioinformatics* **35**, 691–693 (2019).
29. Robson, B. Testing machine learning techniques for general application by using protein secondary structure prediction. A brief survey with studies of pitfalls and benefits using a simple progressive learning approach. *Comput. Biol. Med.* **138**, 104883 (2021).
30. Ismi, D. P., Pulungan, R. & Afiahayati Deep learning for protein secondary structure prediction: pre and post-AlphaFold. *Comput. Struct. Biotechnol. J.* **20**, 6271–6286 (2022).
31. David, A., Islam, S., Tankhilevich, E. & Sternberg, M. J. E. The AlphaFold database of protein structures: a biologist's guide. *J. Mol. Biol.* **434**, 167336 (2022).

32. Lee, C., Su, B. H. & Tseng, Y. J. Comparative studies of AlphaFold, RoseTTAFold and Modeller: a case study involving the use of G-protein-coupled receptors. *Brief. Bioinform.* **23**, bbac308 (2022).
33. He, X. H. et al. AlphaFold2 versus experimental structures: evaluation on G protein-coupled receptors. *Acta Pharmacol. Sin.* **44**, 1–7 (2023).
34. Sala, D., Hildebrand, P. W. & Meiler, J. Biasing AlphaFold2 to predict GPCRs and kinases with user-defined functional or structural properties. *Front. Mol. Biosci.* **10**, 1121962 (2023).
35. Del Alamo, D., Sala, D., McHaourab, H. S. & Meiler, J. Sampling alternative conformational states of transporters and receptors with AlphaFold2. *eLife* **11**, e75751 (2022).
36. Tikhonov, D. B. & Zhorov, B. S. P-loop channels: experimental structures, and physics-based and neural networks-based models. *Membranes* **12**, 229 (2022).
37. Zhu, Z. et al. Simulation and machine learning methods for ion-channel structure determination, mechanistic studies and drug design. *Front. Pharm.* **13**, 939555 (2022).
38. Wayment-Steele, H. K., Ovchinnikov, S., Colwell, L. & Kern, D. Prediction of multiple conformational states by combining sequence clustering with AlphaFold2. Preprint at *BioRxiv* https://doi.org/10.1101/2022.10.17.512570 (2022).
39. Jeppesen, M. & André, I. Accurate prediction of protein assembly structure by combining AlphaFold and symmetrical docking. Preprint at *bioRxiv* https://doi.org/10.1101/2023.06.22.546069 (2023).
40. Stein, R. A. & Mchaourab, H. S. SPEACH_AF: sampling protein ensembles and conformational heterogeneity with Alphafold2. *PLoS Comput. Biol.* **18**, e1010483 (2022).
41. Heo, L. & Feig, M. Multi-state modeling of G-protein coupled receptors at experimental accuracy. *Proteins* **90**, 1873–1885 (2022).
42. Huang, S. Y. & Zou, X. MDockPP: a hierarchical approach for protein–protein docking and its application to CAPRI rounds 15–19. *Proteins* **78**, 3096–3103 (2010).
43. Schneidman-Duhovny, D., Inbar, Y., Nussinov, R. & Wolfson, H. J. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* **33**, W363–W367 (2005).
44. Van Zundert, G. et al. The HADDOCK2. 2 web server: user-friendly integrative modeling of biomolecular complexes. *J. Mol. Biol.* **428**, 720–725 (2016).
45. Quadir, F., Roy, R. S., Soltanikazemi, E. & Cheng, J. DeepComplex: a web server of predicting protein complex structures by deep learning inter-chain contact prediction and distance-based modelling. *Front. Mol. Biosci.* **827**, 716973 (2021).
46. Bryant, P., Pozzati, G. & Elofsson, A. Improved prediction of protein–protein interactions using AlphaFold2. *Nat. Commun.* **13**, 1265 (2022).
47. Burke, D. F. et al. Towards a structurally resolved human protein interaction network. *Nat. Struct. Mol. Biol.* **30**, 216–225 (2023).
48. Ganea, O.-E. et al. Independent SE(3)-equivariant models for end-to-end rigid protein docking. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2111.07786 (2021).
49. Gulsevin, A. et al. Template-free prediction of a new monotopic membrane protein fold and assembly by AlphaFold2. *Biophys. J.* **122**, 2041–2052 (2023).
50. Gao, M., Nakajima An, D., Parks, J. M. & Skolnick, J. AF2Complex predicts direct physical interactions in multimeric proteins with deep learning. *Nat. Commun.* **13**, 1744 (2022).
51. Birch, J. et al. Changes in membrane protein structural biology. *Biology* **9**, 401 (2020).
52. Puthenveetil, R., Christenson, E. T. & Vinogradova, O. New horizons in structural biology of membrane proteins: experimental evaluation of the role of conformational dynamics and intrinsic flexibility. *Membranes* **12**, 227 (2022).
53. García-Nafría, J. & Tate, C. G. Structure determination of GPCRs: cryo-EM compared with X-ray crystallography. *Biochem. Soc. Trans.* **49**, 2345–2355 (2021).
54. Robertson, M. J. et al. Structure determination of inactive-state GPCRs with a universal nanobody. *Nat. Struct. Mol. Biol.* https://doi.org/10.1038/s41594-022-00859-8 (2022).
55. Danev, R. et al. Routine sub-2.5 Å cryo-EM structure determination of GPCRs. *Nat. Commun.* **12**, 4333 (2021).
56. Zhang, K., Wu, H., Hoppe, N., Manglik, A. & Cheng, Y. Fusion protein strategies for cryo-EM study of G protein-coupled receptors. *Nat. Commun.* **13**, 4366 (2022).
57. Cheng, Y. Membrane protein structural biology in the era of single particle cryo-EM. *Curr. Opin. Struct. Biol.* **52**, 58–63 (2018).
58. Latorraca, N. R., Venkatakrishnan, A. J. & Dror, R. O. GPCR dynamics: structures in motion. *Chem. Rev.* **117**, 139–155 (2017).
59. Weis, W. I. & Kobilka, B. K. The molecular basis of G protein-coupled receptor activation. *Annu. Rev. Biochem.* **87**, 897–919 (2018).
60. Knight, K. M. et al. A universal allosteric mechanism for G protein activation. *Mol. Cell* **81**, 1384–1396 (2021).
61. Chen, Q. & Tesmer, J. J. G. G protein-coupled receptor interactions with arrestins and GPCR kinases: the unresolved issue of signal bias. *J. Biol. Chem.* **298**, 102279 (2022).
62. Kankanamge, D., Tennakoon, M., Karunarathne, A. & Gautam, N. G protein γ subunit, a hidden master regulator of GPCR signaling. *J. Biol. Chem.* **298**, 102618 (2022).
63. Pándy-Szekeres, G. et al. GPCRdb in 2023: state-specific structure models using AlphaFold2 and new ligand resources. *Nucleic Acids Res.* **51**, D395–D402 (2023).
64. Liao, M., Cao, E., Julius, D. & Cheng, Y. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* **504**, 107–112 (2013).
65. Newport, T. D., Sansom, M. S. P. & Stansfeld, P. J. The MemProtMD database: a resource for membrane-embedded protein structures and their lipid interactions. *Nucleic Acids Res.* **47**, D390–D397 (2019).
66. Zhang, K., Julius, D. & Cheng, Y. A step-by-step protocol for capturing conformational snapshots of ligand gated ion channels by single-particle cryo-EM. *STAR Protoc.* **3**, 101732 (2022).
67. Singh, A. K. et al. Structural basis of temperature sensation by the TRP channel TRPV3. *Nat. Struct. Mol. Biol.* **26**, 994–998 (2019).
68. Matthies, D. et al. Cryo-EM structures of the magnesium channel CorA reveal symmetry break upon gating. *Cell* **164**, 747–756 (2016).
69. Zhang, Y. et al. Asymmetric opening of the homopentameric 5-HT_{3A} serotonin receptor in lipid bilayers. *Nat. Commun.* **12**, 1074 (2021).
70. Basak, S., Gicheru, Y., Rao, S., Sansom, M. S. P. & Chakrapani, S. Cryo-EM reveals two distinct serotonin-bound conformations of full-length 5-HT_{3A} receptor. *Nature* **563**, 270–274 (2018).
71. Polovinkin, L. et al. Conformational transitions of the serotonin 5-HT_3 receptor. *Nature* **563**, 275–279 (2018).
72. Zhang, K., Julius, D. & Cheng, Y. Structural snapshots of TRPV1 reveal mechanism of polymodal functionality. *Cell* **184**, 5138–5150 (2021).
73. Yang, X. et al. Structure deformation and curvature sensing of PIEZO1 in lipid membranes. *Nature* **604**, 377–383 (2022).
74. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
75. Zhong, E. D., Bepler, T., Berger, B. & Davis, J. H. CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. *Nat. Methods* **18**, 176–185 (2021).
76. Punjani, A. & Fleet, D. 3D flexible refinement: structure and motion of flexible proteins from cryo-EM. *Microsc. Microanal.* **28**, 1218 (2022).

77. Marino, J. & Schertler, G. F. X. A set of common movements within GPCR–G-protein complexes from variability analysis of cryo-EM datasets. *J. Struct. Biol.* **213**, 107699 (2021).
78. Zhong, E. D., Lerer, A., Davis, J. H. & Berger, B. in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)* 4046–4055 (IEEE, 2021).
79. MacKerell, A. D. Jr. Ions everywhere? Mg^{2+} in the μ-opioid GPCR and atomic details of their impact on function. *Biophys. J.* **118**, 783–784 (2020).
80. Ngo, V. et al. Polarization effects in water-mediated selective cation transport across a narrow transmembrane channel. *J. Chem. Theory Comput.* **17**, 1726–1741 (2021).
81. Hollingsworth, S. A. & Dror, R. O. Molecular dynamics simulation for all. *Neuron* **99**, 1129–1143 (2018).
82. Voss, J. M., Harder, O. F., Olshin, P. K., Drabbels, M. & Lorenz, U. J. Rapid melting and revitrification as an approach to microsecond time-resolved cryo-electron microscopy. *Chem. Phys. Lett.* **778**, 138812 (2021).
83. Frank, J. Time-resolved cryo-electron microscopy: recent progress. *J. Struct. Biol.* **200**, 303–306 (2017).
84. Rodríguez-Espigares, I. et al. GPCRmd uncovers the dynamics of the 3D-GPCRome. *Nat. Methods* **17**, 777–787 (2020).
85. Milligan, G., Ward, R. J. & Marsango, S. GPCR homo-oligomerization. *Curr. Opin. Cell Biol.* **57**, 40–47 (2019).
86. Gahbauer, S. & Böckmann, R. A. Membrane-mediated oligomerization of G protein coupled receptors and its implications for GPCR function. *Front. Physiol.* **7**, 494 (2016).
87. Joseph, M. D., Tomas Bort, E., Grose, R. P., McCormick, P. J. & Simoncelli, S. Quantitative super-resolution imaging for the analysis of GPCR oligomerization. *Biomolecules* **11**, 1503 (2021).
88. Møller, T. C. et al. Oligomerization of a G protein-coupled receptor in neurons controlled by its structural dynamics. *Sci. Rep.* **8**, 10414 (2018).
89. Sanchez, R. M., Zhang, Y., Chen, W., Dietrich, L. & Kudryashev, M. Subnanometer-resolution structure determination in situ by hybrid subtomogram averaging—single particle cryo-EM. *Nat. Commun.* **11**, 3709 (2020).
90. Floris, D. & Kühlbrandt, W. Molecular landscape of etioplast inner membranes in higher plants. *Nat. Plants* **7**, 514–523 (2021).
91. Zhu, Y. et al. Structure and activity of particulate methane monooxygenase arrays in methanotrophs. *Nat. Commun.* **13**, 5221 (2022).
92. Tao, X., Zhao, C. & MacKinnon, R. Membrane protein isolation and structure determination in cell-derived membrane vesicles. *Proc. Natl Acad. Sci. USA* **120**, e2302325120 (2023).
93. Zeng, X. & Xu, M. AITom: open-source AI platform for cryo-electron tomography data analysis. Preprint at *arXiv* https://doi.org/10.48550/arXiv.1911.03044 (2019).
94. McCoy, A. J., Sammito, M. D. & Read, R. J. Implications of AlphaFold2 for crystallographic phasing by molecular replacement. *Acta Crystallogr. D Struct. Biol.* **78**, 1–13 (2022).
95. Gulezian, E. et al. Membrane protein production and formulation for drug discovery. *Trends Pharmacol. Sci.* **42**, 657–674 (2021).
96. Hosseini, M., Chen, W., Xiao, D. & Wang, C. Computational molecular docking and virtual screening revealed promising SARS-CoV-2 drugs. *Precis. Clin. Med.* **4**, 1–16 (2021).
97. Zhao, S. et al. Ligand-based pharmacophore modeling, virtual screening and biological evaluation to identify novel TGR5 agonists. *RSC Adv.* **11**, 9403–9409 (2021).
98. Halperin, I., Ma, B., Wolfson, H. & Nussinov, R. Principles of docking: an overview of search algorithms and a guide to scoring functions. *Proteins* **47**, 409–443 (2002).
99. Sabe, V. T. et al. Current trends in computer aided drug design and a highlight of drugs discovered via computational techniques: a review. *Eur. J. Med. Chem.* **224**, 113705 (2021).
100. Cavalli, A. et al. Multi-target-directed ligands to combat neurodegenerative diseases. *J. Med. Chem.* **51**, 347–372 (2008).
101. Gentile, F. et al. Artificial intelligence-enabled virtual screening of ultra-large chemical libraries with deep docking. *Nat. Protoc.* **17**, 672–697 (2022).
102. Crunkhorn, S. Screening ultra-large virtual libraries. *Nat. Rev. Drug Discov.* **21**, 95 (2022).
103. Kampen, S. et al. Structure-guided design of G-protein-coupled receptor polypharmacology. *Angew. Chem. Int. Ed.* **60**, 18022–18030 (2021).
104. Fink, E. A. et al. Structure-based discovery of nonopioid analgesics acting through the α_{2A}-adrenergic receptor. *Science* **377**, eabn7065 (2022).
105. Stein, R. M. et al. Virtual discovery of melatonin receptor ligands to modulate circadian rhythms. *Nature* **579**, 609–614 (2020).
106. Cole, B. A. et al. Structure-based identification and characterization of inhibitors of the epilepsy-associated KNa1.1 (KCNT1) potassium channel. *iScience* **23**, 101100 (2020).
107. Valdes-Jimenez, A. et al. A new strategy for multitarget drug discovery/repositioning through the identification of similar 3D amino acid patterns among proteins structures: the case of tafluprost and its effects on cardiac ion channels. *Front. Pharm.* **13**, 855792 (2022).
108. Lyu, J. et al. Ultra-large library docking for discovering new chemotypes. *Nature* **566**, 224–229 (2019).
109. Gorgulla, C. et al. An open-source drug discovery platform enables ultra-large virtual screens. *Nature* **580**, 663–668 (2020).
110. Sadybekov, A. A. et al. Synthon-based ligand discovery in virtual libraries of over 11 billion compounds. *Nature* **601**, 452–459 (2022).
111. Ballante, F., Kooistra, A. J., Kampen, S., de Graaf, C. & Carlsson, J. Structure-based virtual screening for ligands of G protein–coupled receptors: what can molecular docking do for you? *Pharmacol. Rev.* **73**, 1698 (2021).
112. Pagadala, N. S., Syed, K. & Tuszynski, J. Software for molecular docking: a review. *Biophys. Rev.* **9**, 91–102 (2017).
113. Corso, G., Stärk, H., Jing, B., Barzilay, R. & Jaakkola, T. DiffDock: diffusion steps, twists, and turns for molecular docking. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2210.01776(2022).
114. McNutt, A. T. et al. GNINA 1.0: molecular docking with deep learning. *J. Cheminform.* **13**, 43 (2021).
115. Amendola, G. & Cosconati, S. PyRMD: a new fully automated AI-powered ligand-based virtual screening tool. *J. Chem. Inf. Model.* **61**, 3835–3845 (2021).
116. Alves, L. A. et al. Graph neural networks as a potential tool in improving virtual screening programs. *Front. Chem.* **9**, 787194 (2021).
117. Liu, Z. et al. DeepScreening: a deep learning-based screening web server for accelerating drug discovery. *Database* **2019**, baz104 (2019).
118. Crampon, K., Giorkallos, A., Deldossi, M., Baud, S. & Steffenel, L. A. Machine-learning methods for ligand–protein molecular docking. *Drug Discov. Today* **27**, 151–164 (2022).
119. Yu, Y., Lu, S., Gao, Z., Zheng, H. & Ke, G. Do deep learning models really outperform traditional approaches in molecular docking? Preprint at *arXiv* https://doi.org/10.48550/arXiv.2302.07134 (2023).
120. Molga, K., Dittwald, P. & Grzybowski, B. A. Navigating around patented routes by preserving specific motifs along computer-planned retrosynthetic pathways. *Chem* **5**, 460–473 (2019).
121. de Almeida, A. F., Moreira, R. & Rodrigues, T. Synthetic organic chemistry driven by artificial intelligence. *Nat. Rev. Chem.* **3**, 589–604 (2019).
122. Struble, T. J. et al. Current and future roles of artificial intelligence in medicinal chemistry synthesis. *J. Med. Chem.* **63**, 8667–8682 (2020).

123. Coley, C. W. et al. A graph-convolutional neural network model for the prediction of chemical reactivity. *Chem. Sci.* **10**, 370–377 (2019).
124. Chen, S. & Jung, Y. S. A generalized-template-based graph neural network for accurate organic reactivity prediction. *Nat. Mach. Intell.* **4**, 772–780 (2022).
125. Schreck, J. S., Coley, C. W. & Bishop, K. J. M. Learning retrosynthetic planning through simulated experience. *ACS Cent. Sci.* **5**, 970–981 (2019).
126. Mo, Y. M. et al. Evaluating and clustering retrosynthesis pathways with learned strategy. *Chem. Sci.* **12**, 1469–1478 (2021).
127. Liu, B. W. et al. Retrosynthetic reaction prediction using neural sequence-to-sequence models. *ACS Cent. Sci.* **3**, 1103–1113 (2017).
128. Lin, K. J., Xu, Y. J., Pei, J. F. & Lai, L. H. Automatic retrosynthetic route planning using template-free models. *Chem. Sci.* **11**, 3355–3364 (2020).
129. Segler, M. H. S. & Waller, M. P. Neural-symbolic machine learning for retrosynthesis and reaction prediction. *Chemistry* **23**, 5966–5971 (2017).
130. Jin, W. G., Coley, C. W., Barzilay, R. & Jaakkola, T. Predicting organic reaction outcomes with Weisfeiler–Lehman network. *Proc. 31st Int. Conf. Neural Information Processing Systems* 2604–2613 (2017); https://doi.org/10.5555/3294996.3295021
131. Gao, H. Y. et al. Using machine learning to predict suitable conditions for organic reactions. *ACS Cent. Sci.* **4**, 1465–1476 (2018).
132. Zhou, Z. P., Li, X. C. & Zare, R. N. Optimizing chemical reactions with deep reinforcement learning. *ACS Cent. Sci.* **3**, 1337–1344 (2017).
133. Coley, C. W., Barzilay, R., Jaakkola, T. S., Green, W. H. & Jensen, K. F. Prediction of organic reaction outcomes using machine learning. *ACA Cent. Sci.* **3**, 434–443 (2017).
134. Klucznik, T. et al. Efficient syntheses of diverse, medicinally relevant targets planned by computer and executed in the laboratory. *Chem* **4**, 522–532 (2018).
135. Padmaja, R. D. & Chanda, K. A short review on synthetic advances toward the synthesis of rufinamide, an antiepileptic drug. *Org. Process Res. Dev.* **22**, 457–466 (2018).
136. Segler, M. H. S., Preuss, M. & Waller, M. P. Planning chemical syntheses with deep neural networks and symbolic AI. *Nature* **555**, 604–610 (2018).
137. Yuan, S. et al. Computational modeling of the olfactory receptor Olfr73 suggests a molecular basis for low potency of olfactory receptor-activating compounds. *Commun. Biol.* **2**, 141 (2019).
138. Schardt, J. S. et al. Agonist antibody discovery: experimental, computational, and rational engineering approaches. *Drug Discov. Today* **27**, 31–48 (2022).
139. El Daibani, A. et al. Molecular mechanism of biased signaling at the κ opioid receptor. *Nat. Commun.* **14**, 1338 (2023).
140. Roth, B. L. & Kroeze, W. K. Integrated approaches for genome-wide interrogation of the druggable non-olfactory G protein-coupled receptor superfamily. *J. Biol. Chem.* **290**, 19471–19477 (2015).
141. Husted, A. S., Trauelsen, M., Rudenko, O., Hjorth, S. A. & Schwartz, T. W. GPCR-mediated signaling of metabolites. *Cell Metab.* **25**, 777–796 (2017).
142. Sriram, K. & Insel, P. A. G protein-coupled receptors as targets for approved drugs: how many targets and how many drugs? *Mol. Pharmacol.* **93**, 251–258 (2018).
143. Avet, C. et al. Effector membrane translocation biosensors reveal G protein and betaarrestin coupling profiles of 100 therapeutically relevant GPCRs. *eLife* **11**, e74101 (2022).
144. Hauser, A. S. et al. Common coupling map advances GPCR–G protein selectivity. *eLife* **11**, e74107 (2022).
145. Low, V., Li, Z. & Blenis, J. Metabolite activation of tumorigenic signaling pathways in the tumor microenvironment. *Sci. Signal.* **15**, eabj4220 (2022).
146. Drew, L. Olfactory receptors are not unique to the nose. *Nature* **606**, S14–S17 (2022).
147. Orecchioni, M. et al. Olfactory receptor 2 in vascular macrophages drives atherosclerosis by NLRP3-dependent IL-1 production. *Science* **375**, 214–221 (2022).
148. Hofmann, K. P. & Lamb, T. D. Rhodopsin, light-sensor of vision. *Prog. Retin. Eye Res.* **93**, 101116 (2023).
149. Marullo, S. et al. Mechanical GPCR activation by traction forces exerted on receptor N-glycans. *ACS Pharm. Transl. Sci.* **3**, 171–178 (2020).
150. Erdogmus, S. et al. Helix 8 is the essential structural motif of mechanosensitive GPCRs. *Nat. Commun.* **10**, 5784 (2019).
151. Zhang, R. & Xie, X. Tools for GPCR drug discovery. *Acta Pharmacol. Sin.* **33**, 372–384 (2012).
152. Nielsen, C. D., Dhasmana, D., Floresta, G., Wohland, T. & Cilibrizzi, A. Illuminating the path to target GPCR structures and functions. *Biochemistry* **59**, 3783–3795 (2020).
153. Zhou, Y., Meng, J., Xu, C. & Liu, J. Multiple GPCR functional assays based on resonance energy transfer sensors. *Front. Cell Dev. Biol.* **9**, 611443 (2021).
154. Tian, H., Furstenberg, A. & Huber, T. Labeling and single-molecule methods to monitor G protein-coupled receptor dynamics. *Chem. Rev.* **117**, 186–245 (2017).
155. Serfling, R. et al. Quantitative single-residue bioorthogonal labeling of G protein-coupled receptors in live cells. *ACS Chem. Biol.* **14**, 1141–1149 (2019).
156. Meral, D. et al. Molecular details of dimerization kinetics reveal negligible populations of transient μ-opioid receptor homodimers at physiological concentrations. *Sci. Rep.* **8**, 7705 (2018).
157. Schihada, H. et al. A universal bioluminescence resonance energy transfer sensor design enables high-sensitivity screening of GPCR activation dynamics. *Commun. Biol.* **1**, 105 (2018).
158. Schihada, H. et al. Development of a conformational histamine H_3 receptor biosensor for the synchronous screening of agonists and inverse agonists. *ACS Sens.* **5**, 1734–1742 (2020).
159. Maziarz, M. et al. Revealing the activity of trimeric G-proteins in live cells with a versatile biosensor design. *Cell* **182**, 770–785 (2020).
160. Olsen, R. H. J. & English, J. G. Advancements in G protein-coupled receptor biosensors to study GPCR–G protein coupling. *Br. J. Pharmacol.* **180**, 1433–1443 (2023).
161. Liu, J. et al. Biased signaling due to oligomerization of the G protein-coupled platelet-activating factor receptor. *Nat. Commun.* **13**, 6365 (2022).
162. Nuber, S. et al. β-Arrestin biosensors reveal a rapid, receptor-dependent activation/deactivation cycle. *Nature* **531**, 661–664 (2016).
163. Janetzko, J. et al. Membrane phosphoinositides regulate GPCR–β-arrestin complex assembly and dynamics. *Cell* **185**, 4560–4573 (2022).
164. Momboisse, F. et al. Tracking receptor motions at the plasma membrane reveals distinct effects of ligands on CCR5 dynamics depending on its dimerization status. *eLife* **11**, e76281 (2022).
165. Moller, J. et al. Single-molecule analysis reveals agonist-specific dimer formation of μ-opioid receptors. *Nat. Chem. Biol.* **16**, 946–954 (2020).
166. Sungkaworn, T. et al. Single-molecule imaging reveals receptor–G protein interactions at cell surface hot spots. *Nature* **550**, 543–547 (2017).
167. Kasai, R. S. & Kusumi, A. Single-molecule imaging revealed dynamic GPCR dimerization. *Curr. Opin. Cell Biol.* **27**, 78–86 (2014).

168. Lamichhane, R. et al. Single-molecule view of basal activity and activation mechanisms of the G protein-coupled receptor β2AR. *Proc. Natl Acad. Sci. USA* **112**, 14254–14259 (2015).
169. Gregorio, G. G. et al. Single-molecule analysis of ligand efficacy in β2AR–G-protein activation. *Nature* **547**, 68–73 (2017).
170. Walsh, S. M. et al. Single proteoliposome high-content analysis reveals differences in the homo-oligomerization of GPCRs. *Biophys. J.* **115**, 300–312 (2018).
171. Milstein, J. N., Nino, D. F., Zhou, X. & Gradinaru, C. C. Single-molecule counting applied to the study of GPCR oligomerization. *Biophys. J.* **121**, 3175–3187 (2022).
172. Asher, W. B. et al. Single-molecule FRET imaging of GPCR dimers in living cells. *Nat. Methods* **18**, 397–405 (2021).
173. Kondratskyi, A. Classification of ion channels. *Ion Channel Library* https://www.ionchannellibrary.com/classification-of-ion-channels/ (2019).
174. Kondratskyi, A. Drugs on the market. *Ion Channel Library* https://www.ionchannellibrary.com/ion-channel-drugs/ (2019).
175. Sakmann, B. & Neher, E. *Single-Channel Recording* (Plenum Press, 1983).
176. Seibertz, F. et al. A modern automated patch-clamp approach for high throughput electrophysiology recordings in native cardiomyocytes. *Commun. Biol.* **5**, 969 (2022).
177. Obergrussberger, A., Friis, S., Bruggemann, A. & Fertig, N. Automated patch clamp in drug discovery: major breakthroughs and innovation in the last decade. *Expert Opin. Drug Discov.* **16**, 1–5 (2021).
178. Gutsmann, T., Heimburg, T., Keyser, U., Mahendran, K. R. & Winterhalter, M. Protein reconstitution into freestanding planar lipid membranes for electrophysiological characterization. *Nat. Protoc.* **10**, 188–198 (2015).
179. Talwar, S. & Lynch, J. W. Investigating ion channel conformational changes using voltage clamp fluorometry. *Neuropharmacology* **98**, 3–12 (2015).
180. Andriani, R. T. & Kubo, Y. Voltage-clamp fluorometry analysis of structural rearrangements of ATP-gated channel P2X2 upon hyperpolarization. *eLife* **10**, e65822 (2021).
181. Patel, V. R. et al. Single-molecule imaging with cell-derived nanovesicles reveals early binding dynamics at a cyclic nucleotide-gated ion channel. *Nat. Commun.* **12**, 6459 (2021).
182. Menegon, A. et al. A new electro-optical approach for conductance measurement: an assay for the study of drugs acting on ligand-gated ion channels. *Sci. Rep.* **7**, 44843 (2017).
183. Zhang, X. M., Yokoyama, T. & Sakamoto, M. Imaging voltage with microbial rhodopsins. *Front. Mol. Biosci.* **8**, 738829 (2021).
184. Sankaran, J. & Wohland, T. Fluorescence strategies for mapping cell membrane dynamics and structures. *APL Bioeng.* **4**, 020901 (2020).
185. Monnier, N. et al. Bayesian approach to MSD-based analysis of particle motion in live cells. *Biophys. J.* **103**, 616–626 (2012).
186. Monnier, N. et al. Inferring transient particle transport dynamics in live cells. *Nat. Methods* **12**, 838–840 (2015).
187. Dosset, P. et al. Automatic detection of diffusion modes within biological membranes using back-propagation neural network. *BMC Bioinformatics* **17**, 197 (2016).
188. Granik, N. et al. Single-particle diffusion characterization by deep learning. *Biophys. J.* **117**, 185–192 (2019).
189. Kowalek, P., Loch-Olszewska, H. & Szwabinski, J. Classification of diffusion modes in single-particle tracking data: feature-based versus deep-learning approach. *Phys. Rev. E* **100**, 032410 (2019).
190. Veya, L., Piguet, J. & Vogel, H. Single molecule imaging deciphers the relation between mobility and signaling of a prototypical G protein-coupled receptor in living cells. *J. Biol. Chem.* **290**, 27723–27735 (2015).
191. Duncan, A. L., Song, W. & Sansom, M. S. P. Lipid-dependent regulation of ion channels and G protein-coupled receptors: insights from structures and simulations. *Annu. Rev. Pharmacol. Toxicol.* **60**, 31–50 (2020).
192. Kozlovskii, I. & Popov, P. Spatiotemporal identification of druggable binding sites using deep learning. *Commun. Biol.* **3**, 618 (2020).
193. Verkhivker, G., Alshahrani, M., Gupta, G., Xiao, S. & Tao, P. From deep mutational mapping of allosteric protein landscapes to deep learning of allostery and hidden allosteric sites: zooming in on "allosteric intersection" of biochemical and big data approaches. *Int. J. Mol. Sci.* **24**, 7747 (2023).
194. Ouyang, W., Aristov, A., Lelek, M., Hao, X. & Zimmer, C. Deep learning massively accelerates super-resolution localization microscopy. *Nat. Biotechnol.* **36**, 460–468 (2018).
195. Kim, T., Moon, S. & Xu, K. Information-rich localization microscopy through machine learning. *Nat. Commun.* **10**, 1996 (2019).
196. Qiao, C. et al. Rationalized deep learning super-resolution microscopy for sustained live imaging of rapid subcellular processes. *Nat. Biotechnol.* **41**, 367–377 (2023).
197. Mahecic, D. et al. Event-driven acquisition for content-enriched microscopy. *Nat. Methods* **19**, 1262–1267 (2022).
198. Scheefhals, N., Westra, M. & MacGillavry, H. D. mGluR5 is transiently confined in perisynaptic nanodomains to shape synaptic function. *Nat. Commun.* **14**, 244 (2023).
199. Marrink, S. J. et al. Computational modeling of realistic cell membranes. *Chem. Rev.* **119**, 6184–6226 (2019).
200. Callaway, E. What's next for AlphaFold and the AI protein-folding revolution. *Nature* **604**, 234–238 (2022).
201. Savage, N. Tapping into the drug discovery potential of AI. *Biopharma Dealmakers* https://doi.org/10.1038/d43747-021-00045-7(2021).
202. Bepler, T., Kelley, K., Noble, A. J. & Berger, B. Topaz-Denoise: general deep denoising models for cryoEM and cryoET. *Nat. Commun.* **11**, 5208 (2020).
203. Bhat, B., Ganai, N. A., Andrabi, S. M., Shah, R. A. & Singh, A. TM-Aligner: multiple sequence alignment tool for transmembrane proteins with reduced time and improved accuracy. *Sci. Rep.* **7**, 12543 (2017).
204. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
205. Feng, S. H., Zhang, W. X., Yang, J., Yang, Y. & Shen, H. B. Topology prediction improvement of α-helical transmembrane proteins through helix-tail modeling and multiscale deep learning fusion. *J. Mol. Biol.* **432**, 1279–1296 (2020).
206. Bender, B. J., Marlow, B. & Meiler, J. Improving homology modeling from low-sequence identity templates in Rosetta: a case study in GPCRs *PLOS Comput. Biol.* https://doi.org/10.1371/journal.pcbi.1007597 (2020).
207. Gutierrez, S., Tyczynski, W. G., Boomsma, W., Teufel, F. & Winther, O. MembraneFold: visualising transmembrane protein structure and topology. Preprint at *bioRxiv* https://doi.org/10.1101/2022.12.06.518085 (2022).
208. Tovchigrechko, A. & Vakser, I. A. GRAMM-X public web server for protein–protein docking. *Nucleic Acids Res.* **34**, W310–W314 (2006).
209. Lyskov, S. & Gray, J. J. The RosettaDock server for local protein–protein docking. *Nucleic Acids Res.* **36**, W233–W238 (2008).
210. White, S. Membrane proteins of known 3D structure. https://blanco.biomol.uci.edu/mpstruc/ (2023).
211. Katritch, V. et al. Structure-based discovery of novel chemotypes for adenosine A_{2A} receptor antagonists. *J. Med. Chem.* **53**, 1799–1809 (2010).
212. Scharf, M. M., Bünemann, M., Baker, J. G. & Kolb, P. Comparative docking to distinct G protein–coupled receptor conformations exclusively yields ligands with agonist efficacy. *Mol. Pharmacol.* **96**, 851 (2019).

213. Schmidt, D., Bernat, V., Brox, R., Tschammer, N. & Kolb, P. Identifying modulators of CXC receptors 3 and 4 with tailored selectivity using multi-target docking. *ACS Chem. Biol.* **10**, 715–724 (2015).
214. Lane, J. R. et al. Structure-based ligand discovery targeting orthosteric and allosteric pockets of dopamine receptors. *Mol. Pharmacol.* **84**, 794–807 (2013).
215. Kooistra, A. J. et al. Function-specific virtual screening for GPCR ligands using a combined scoring method. *Sci. Rep.* **6**, 28288 (2016).
216. Zheng, Z. et al. Structure-based discovery of new antagonist and biased agonist chemotypes for the κ opioid receptor. *J. Med. Chem.* **60**, 3070–3081 (2017).
217. Ranganathan, A. et al. Ligand discovery for a peptide-binding GPCR by structure-based screening of fragment- and lead-like chemical libraries. *ACS Chem. Biol.* **12**, 735–745 (2017).
218. Gunera, J., Baker, J. G., van Hilten, N., Rosenbaum, D. M. & Kolb, P. Structure-based discovery of novel ligands for the orexin 2 receptor. *J. Med. Chem.* **63**, 11045–11053 (2020).
219. Caseley, E. A., Muench, S. P., Fishwick, C. W. & Jiang, L. H. Structure-based identification and characterisation of structurally novel human P2X7 receptor antagonists. *Biochem. Pharmacol.* **116**, 130–139 (2016).
220. Lacroix, C. et al. Identification of novel Smoothened ligands using structure-based docking. *PLoS ONE* **11**, e0160365 (2016).
221. Iwata, H. et al. Discovery of natural TRPA1 activators through pharmacophore-based virtual screening and a biological assay. *Bioorg. Med. Chem. Lett.* **31**, 127639 (2021).
222. Rodriguez, D., Brea, J., Loza, M. I. & Carlsson, J. Structure-based discovery of selective serotonin 5-HT_{1B} receptor ligands. *Structure* **22**, 1140–1151 (2014).

Acknowledgements

Funding from the following sources is acknowledged: the National Natural Science Foundation of China and Swiss National Science Foundation (NSFC-SNF 32161133022 to H.V. and H.S.); the Shenzhen Key Laboratory of Computer-Aided Drug Discovery, Advanced Technology, Chinese Academy of Sciences, Shenzhen (funding no. ZDSYS20201230165400001 to S.Y. and H.V.); the Chinese Academy of Science President's International Fellowship Initiative (PIFI) (no. 2020FSB0003 to H.V.); Guangdong Retired Expert to H.V. (granted by Guangdong Province); Shenzhen Pengcheng Scientist to H.V; the AlphaMol and SIAT Joint Laboratory to S.Y. and H.V; and Shenzhen Government Top-Talent Working Funding and Guangdong Province Academician Work Funding to H.V.

Competing interests

S.Y. and H.V. are cofounders of AlphaMol Science Ltd. The remaining authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Henning Stahlberg, Shuguang Yuan or Horst Vogel.