SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Mark Rejna

Computational drug development for membrane protein targets

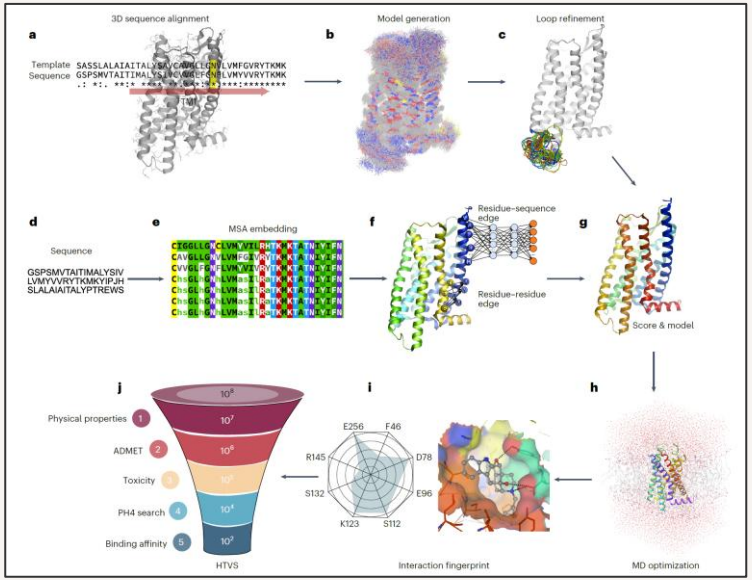
Table of Contents

Introduction

- Membrane proteins account for 1/3 of all human proteins, with G-Protein-coupled receptors (GPCRs) and ion channel proteins being the most common.
- The majority of FDA-approved drugs target membrane proteins to treat severe diseases.
- In 2023, the market for GPCR targeting was valued at $4.1 billion and it is estimated by BCC Research that the market will reach a value of $6.1 billion in 2029.
- Challenges in traditional drug development arise as membrane proteins are difficult to study and manipulate due to their dynamic behaviour, structural complexity, and dependence on lipid environments. Thus, experimental techniques such as X-ray crystallography and NMR are less effective.
- As a result, the development of a drug molecule takes 10–15 years with costs up to $4.5 billion.
- Developments in computational methods overcome these challenges by enabling 3D modelling, virtual screening, and prediction of drug-likeness and biological activity, ultimately reducing time and cost.

Abstract

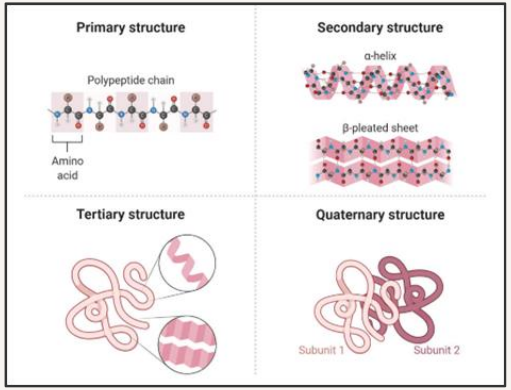
- Recent advances in deep learning and computational biology have accelerated drug development for membrane protein targets.
- Tools such as AlphaFold2 provide reliable predictions for soluble proteins, however, limitations remain in modeling dynamic structural transitions critical for membrane protein signaling.
- These limitations can be overcome by implementing a combination of computational tools and experimental techniques.
- This review discusses the steps in the pipeline during which computational tools accelerate the design of drugs and then moves onto the current reliability in these computational tools and how to improve it in the future.



Tools for membrane protein structure prediction

Primary sequence alignment tools
tmAligner

Secondary structure prediction
TMPSS



Tertiary structure prediction and structural databases
PredMP
ColabFold
AlphaFold
RoseTTAFold
MemBrain
RosettaGPCR
ESMFold
OmegaFold
Membrane Fold

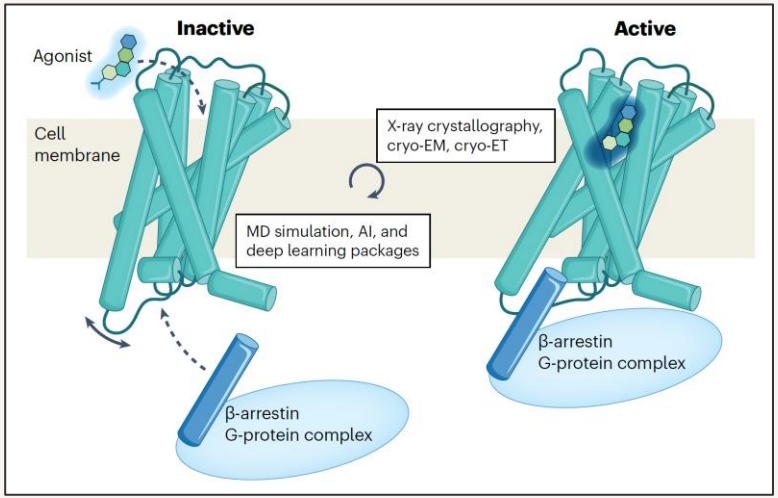
Quaternary structure prediction
AlphaFold-Multime
DeepComplex
EquiDock
PatchDock
GRAMM-X
RosettaDock

Predicting membrane protein structures

- AI-based computational tools enable faster structure predictions than experimental techniques.
- AlphaFold2 & RoseTTAFold use deep learning (neural networks, multiple sequence alignments (MSA)) to predict protein folds, even with limited structural data, good for backbone structure but struggle with proteins for which alignments are not available.
- AlphaFold2 has structural models for 98.5% of the human proteome, only 58% are modeled with high confidence due to intrinsically disordered protein regions.
- When MSA is not available Single-Sequence Predictors (LLMs) ESMFold and OmegaFold are useful.
- Older tools such as TMPSS, and PredMP remain useful for membrane proteins where less evolutionary data is available.
- Structure predictions can be improved by using a combination of tools, for GPCRs - better binding pocket modeling, for ion channels & transporters - better functional conformations.
- Current Challenges:
    - Accuracy drops for proteins with no evolutionary data or disordered regions.
    - Difficult to model alternative conformations, ligand-binding pockets, and multimeric assemblies.
    - Limited experimental data for membrane proteins reduces AI prediction reliability.

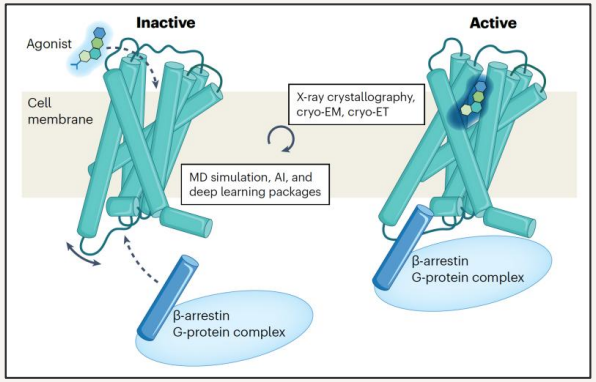
Experimental techniques enchaned with ML

- Cryo-EM captures membrane protein structure of flexible, large protein complexes with a lower limit of 60 kDa.
- The size of a class A single GPCR is approximately 35 kDa, outside of this range, although the size of the particle can be increased using G proteins or β-arrestin allowing the structures to be determined in their active state.
- By combining experimental and computational techniques, it is possible to model the structure of the entire signaling cascade of GPCRs from the inactive state to G-protein complex or β-arrestin-bound states as well as the intermediate structural transitions.



Experimental techniques enchaned with ML

- X-ray crystallography - tool for resolving the structures of GPCRs in their inactive state (258/299 human inactive GPCRs)
- X-ray crystallography provides high-resolution structures but requires stable crystals, often limiting it to rigid, inactive conformations.
- Cryo-EM - tool for determination of active-state structures (272/322 human active GPCRs), (117/222 ion channels)
- Advancements in Cryo-ET allow for the structure determination for membrane proteins on a single-particle imaging scale.
- Integration of computational and ML tools in the experimental structure determination allows for more accurate models of membrane proteins in distinct functionally relevant states aiding drug discovery.



Virtual screening in drug discovery

- Virtual screening complements experimental high-throughput screening (HTS) by screening large compound databases against target proteins and selecting a number of compounds for testing.
- Uses docking/pharmacophore modeling to filter compounds by biophysical properties.
- HTS accounts for approximately 15% of R&D costs - Virtual HTS improves efficiency thus reducing costs and time - 1 billion compounds in 2 weeks.
- Traditional drug development focuses on single-target, high-selective drugs which are insufficient for some treatment meaning multiple medicines are needed and side effects and drug interactions can be a problem. Virtual screening has allowed the development of multi-targeted drugs - Parkinson's disease.
- More than 70 computer aided drug designs have been approved for marketing.

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]

AI-Driven Chemical Synthesis

- As virtual screening identifies more potential ligands, there is an increasing demand for AI methods that can design new chemical structures beyond existing chemical libraries and plan how to synthesize them efficiently.
- Deep learning methods are being used to find the most optimal retrosynthesis routes to synthesise complex molecules from available compounds as well as predicting the simplest reagents and reaction conditions for each synthesis step - 71.8% accuracy.
- As molecular complexity and drug selectivity increases challenges arise as computational models need to fulfill requirements such as purchasable, selective, and valid reactants as well as dealing with chirality and stereoselectivity.
- Overall, as computational predictions improve in accuracy, chemists can shift their mindset from "how" to "what“ when discussing chemical synthesis.

GPCR and ion channel signaling

GPCR Signaling

- 800+ GPCRs regulate neurotransmission, cell metabolism and growth, immune responses.
- Fluorescent probes and super-resolution microscopy used to study GPCR oligomerisation, activation, protein interactions.

Ion Channel Signaling

- 300+ structurally diverse channels regulate electrical signaling, synaptic activity, and cell homeostasis.
- Automatic patch-clamp enable efficient measurements of ion channel activity.
- Optical fluorescent methods are used to identify ion channel dynamics, activation states, and drug interactions at single-molecule resolution.

Integrating computational analysis and experiments

- GPCRs: ML is used to classify receptor states from single-molecule tracking data as well as to link dynamic protein states to functions.
- Ion channels: ML and molecular simulations are used to predict activation mechanisms and identify allosteric modulators.
- Overall aiding drug development by targeting specific receptors or regulatory sites.

Author's Conclusion

- Computational and ML methods are rapidly accelerating drug discovery.
- Many start-up companies, as well as Google, Microsoft, Amazon, IBM, Meta, and traditional pharmaceutical companies are transitioning into data-driven drug development.
- Computational drug development industry is continuously growing as there is room for substantial improvement to overcome current limitations
    - Large amount of high-quality data for reliable predictions and for the AI training phase
    - More experimentally determined models to help predict structures of GPCRs in the inactive state - experimental methods are limited by size
    - X-ray crystallography and cryo-EM have low signal-to-noise ratios - need to develop AI denoising algorithms
    - Research in electron microscopy hardware development is also necessary to develop higher-resolution structures
- In the future, a resolution revolution in the determination of structures for small membrane proteins would greatly accelerate drug discovery processes.

Evaluation and conclusion

Feedback

+ Many clear real-world examples/case studies were given throughout the research paper.
+ Not only recent success but also limitations and future improvements were outlined.
^ More visual figures (flowcharts, diagrams) to highlight the relationships between computational tools and stages in drug discovery.

- Overall, this research paper effectively outlines how computational methods are improving drug discovery for membrane protein targets. It discusses how computational tools are implemented to aid experimental methods whilst highlighting the limitations and future of the drug development.

References

Li, H., Sun, X., Cui, W., Xu, M., Dong, J., Ekundayo, B.E., Ni, D., Rao, Z., Guo, L., Stahlberg, H., Yuan, S., & Vogel, H. (2024). *Computational drug development for membrane protein targets*. Published: February 15, 2024. Accessed: June 30, 2025.

BCC Publishing. (2025). *G-Protein Coupled Receptor (GPCR) Targeting: Technologies and Global Markets*. Published: April 2025. Accessed: June 30, 2025.

Aryal, S. (2022). *Protein Structure – Primary, Secondary, Tertiary, and Quaternary*. Microbe Notes. Retrieved from https://microbenotes.com/protein-structure-primary-secondary-tertiary-and-quaternary/. Accessed: July 2, 2025.

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Mark Rejna

Thank you for listening, questions are welcome

If you have any further questions, please contact me via

Email: mark_rejna@mymail.sutd.edu.sg
Telegram: @markrejna
Phone: +65 8580 2824