CellPress
REVIEWS

Special Issue: Rise of Machines in Medicine

Spotlight

Artificial Intelligence: A Novel Approach for Drug Discovery

Óscar Díaz, [1,2,3]
James A.R. Dalton, [1,2,3] and
Jesús Giraldo [1,2,3,*]

Molecular dynamics (MD) simulations can mechanistically explain receptor function. However, the enormous data sets that they may imply can be a hurdle. Plante and colleagues (*Molecules*, 2019) recently described a machine learning approach to the analysis of MD simulations. The approach successfully classified ligands and identified functional receptor motifs and thus it seems promising for mechanism-based drug discovery.

G protein-coupled receptors (GPCRs) are one of the most targeted protein families in current pharmaceutical research [1]. There can be various explanations: one is that these proteins are involved in many physiological functions and, therefore, their malfunctioning can be the cause of disease. Another is their structural and functional complexity, which is on the one hand a hurdle for drug discovery, but, on the other hand, it poses an extra level of richness to ligand space. The more subtle the desired functional effect, the more extensive and varied the molecular space for drug design will be. Different types of ligands can be found depending on the pharmacological effect sought: full and partial agonists, neutral antagonists, and full and partial inverse agonists. In addition, allosteric modulators and biased ligands are now attracting much attention because of their particular properties: the former, modulating the effects of endogenous ligands without replacing them from their binding sites [2] and the latter, differentially selecting one signaling pathway with respect to others [3].

Crystallography has contributed enormously to GPCR knowledge by determining the structures of receptors, either free or bound to agonists, antagonists, inverse agonists, and allosteric modulators, thereby helping to establish the general characteristics associated with GPCR active and inactive states (i.e., the structural motifs of receptor function) [4]. Nevertheless, crystal structures are snapshots, static views, of very flexible molecules. GPCRs, as proteins, can adopt multiple 3D forms (molecular conformations) and these forms will determine their ultimate physiological response.

Molecular dynamics (MD) simulations is a computational technique that can effectively explore the conformational space of GPCRs [5] (Figure 1). However, this technique inherently presents a 'Big Data' problem. To satisfactorily explore the GPCR conformational space, long-timescale MD simulations are needed and this implies the involvement of a large collection of molecular structures of the system. How can we manage such a large set? How can we extract relevant information from it? Although current state-of-the-art GPCR-ligand MD simulations can be analyzed effectively with conventional methods in the field of multivariate statistics, it can be assumed that if advances in computational power and force-field accuracy continue to be made as expected [6], MD simulations of GPCRs will become exclusively high throughput, as well as robustly predictive, in the near future. This means the functional effect of perhaps hundreds (or even thousands) of ligands could theoretically be characterized over microsecond (or millisecond) time-lengths in a single receptor and readily achieved in a manageable time-frame. The concomitant explosion of such simulation data will likely necessitate the implementation of automated analytical methods, such as that provided by machine learning (ML) or other artificial intelligence (AI) techniques [6,7].

In a recent article [8], Harel Weinstein's group describes a new ML approach to the analysis of MD simulations that is based on transforming MD simulation trajectories into a representation recognizable by deep learning object recognition technology. Two Class A GPCRs were chosen as targets: serotonin 5-HT2A and dopamine D2 receptors, which were studied bound to a collection of eight ligands: either full, partial, or inverse agonists. After MD simulations of each receptor–ligand complex, the atomic coordinates (XYZ) for each MD-generated conformation (or frame) were transformed into the colors of an image (RGB) that was easily recognizable by the ML algorithm. After training, the ML algorithm was able to classify the MD-generated GPCR conformations according to the pharmacological effect of the bound ligand (full, partial, inverse agonist) with high accuracy for the tested ligands (>98%). Furthermore, the contribution of each atom to the classification could be calculated, allowing the identification of the molecular determinants of receptor function (i.e., those atoms/residues of the receptor that are most important for classification according to ligand pharmacological effect). Remarkably, it was found that the most important regions responsible for receptor activation were common to both 5-HT2A and D2 receptors. In particular, the extracellular ends of transmembrane (TM) helices 4 and 5, extracellular loop 2, the intracellular ends of TM helices 5 and 6, and intracellular loops ICL2 and ICL3

CellPress
REVIEWS

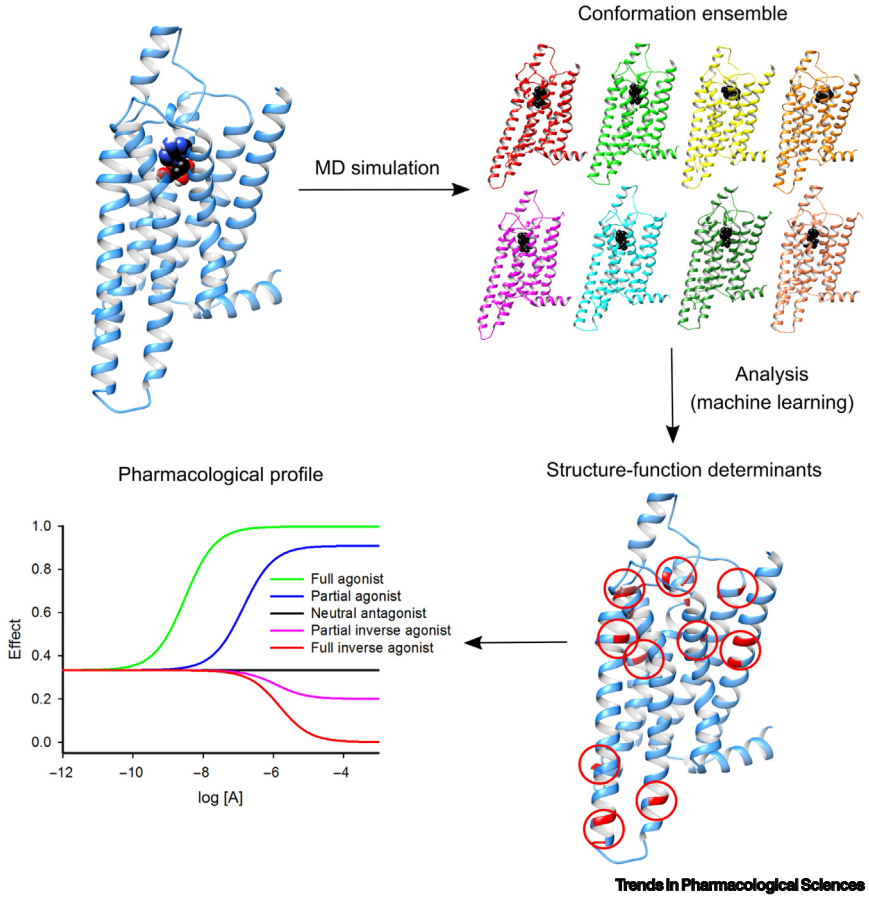

Figure 1. Workflow of Molecular Dynamics (MD) Simulations and Machine Learning-Based Approach for Ligand Classification. MD simulations of a ligand-bound G protein-coupled receptor generate multiple ligand–receptor conformations (conformational ensemble). Machine learning can be used as a powerful analysis tool to differentiate between ligand-induced receptor conformational changes, which allows the identification of residues involved in the pharmacological action of the ligand. When a large collection of ligands are studied, machine learning approaches could potentially classify the pharmacological profile of new drugs.

were identified as important for discrimination between different ligand types.

With this work, the authors added to ongoing work in the area of ligand bias, which is a pharmacological concept of fundamental importance, the appropriate quantification of which [9,10] in drug discovery programs can be key to finding drugs with the desired efficacy in the selected signaling pathway. It is expected that the coupling of MD simulations with AI approaches integrated in a general pharmacological model of drug action will be a promising tool in the design of more efficacious and safer medicines. Moreover, advances in computational capability can make this conceptual progress accessible to the simultaneous analysis of large data sets of ligands, thus making dynamic virtual screening a systematic tool for novel drug discovery.

Acknowledgments

This study was supported in part by Ministerio de Economía, Industria y Competitividad (SAF2017-87199-R).

[1]Laboratory of Molecular Neuropharmacology and Bioinformatics, Unitat de Bioestadística and Institut de Neurociències, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain
[2]Instituto de Salud Carlos III, Centro de Investigación Biomédica en Red de Salud Mental, CIBERSAM, 08193, Bellaterra, Spain
[3]Unitat de Neurociència Traslacional, Parc Taulí Hospital Universitari, Institut d'Investigació i Innovació Parc Taulí (I3PT), Institut de Neurociències, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

*Correspondence:
Jesus.Giraldo@uab.es (J. Giraldo).

https://doi.org/10.1016/j.tips.2019.06.005

References

1. Santos, R. *et al.* (2017) A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* 16, 19–34
2. Thal, D.M. *et al.* (2018) Structural insights into G-protein-coupled receptor allostery. *Nature* 559, 45–53
3. Kenakin, T. (2019) Biased receptor signaling in drug discovery. *Pharmacol. Rev.* 71, 267–315
4. Weis, W.I. and Kobilka, B.K. (2018) The molecular basis of G protein-coupled receptor activation. *Annu. Rev. Biochem.* 87, 897–919
5. Hollingsworth, S.A. and Dror, R.O. (2018) Molecular dynamics simulation for all. *Neuron* 99, 1129–1143
6. Perez, A. *et al.* (2018) Simulations meet machine learning in structural biology. *Curr. Opin. Struct. Biol.* 49, 139–144
7. Degiacomi, M.T. (2019) Coupling molecular dynamics and deep learning to mine protein conformational space. *Structure* 27, 1034–1040
8. Plante, A. *et al.* (2019) A machine learning approach for the discovery of ligand-specific functional mechanisms of GPCRs. *Molecules* 24, E2097
9. Kenakin, T. *et al.* (2012) A simple method for quantifying functional selectivity and agonist bias. *ACS Chem. Neurosci.* 3, 193–203
10. Zhou, B. *et al.* (2019) Can adding constitutive receptor activity redefine biased signaling quantification? *Trends Pharmacol. Sci.* 40, 156–160

Special Issue: Rise of Machines in Medicine

Forum

Deep Learning to Therapeutically Target Unreported Complexes

Ariel Fernández [1,2,3,4,*]

The disruption of large protein–protein (PP) interfaces remains a challenge in targeted therapy. Designing drugs that compete with binding partners is daunting, especially when the structure of the protein complex is unknown.