To explore drug space smarter: artificial intelligence in drug design for G protein-coupled receptors

Liu, X.

Citation

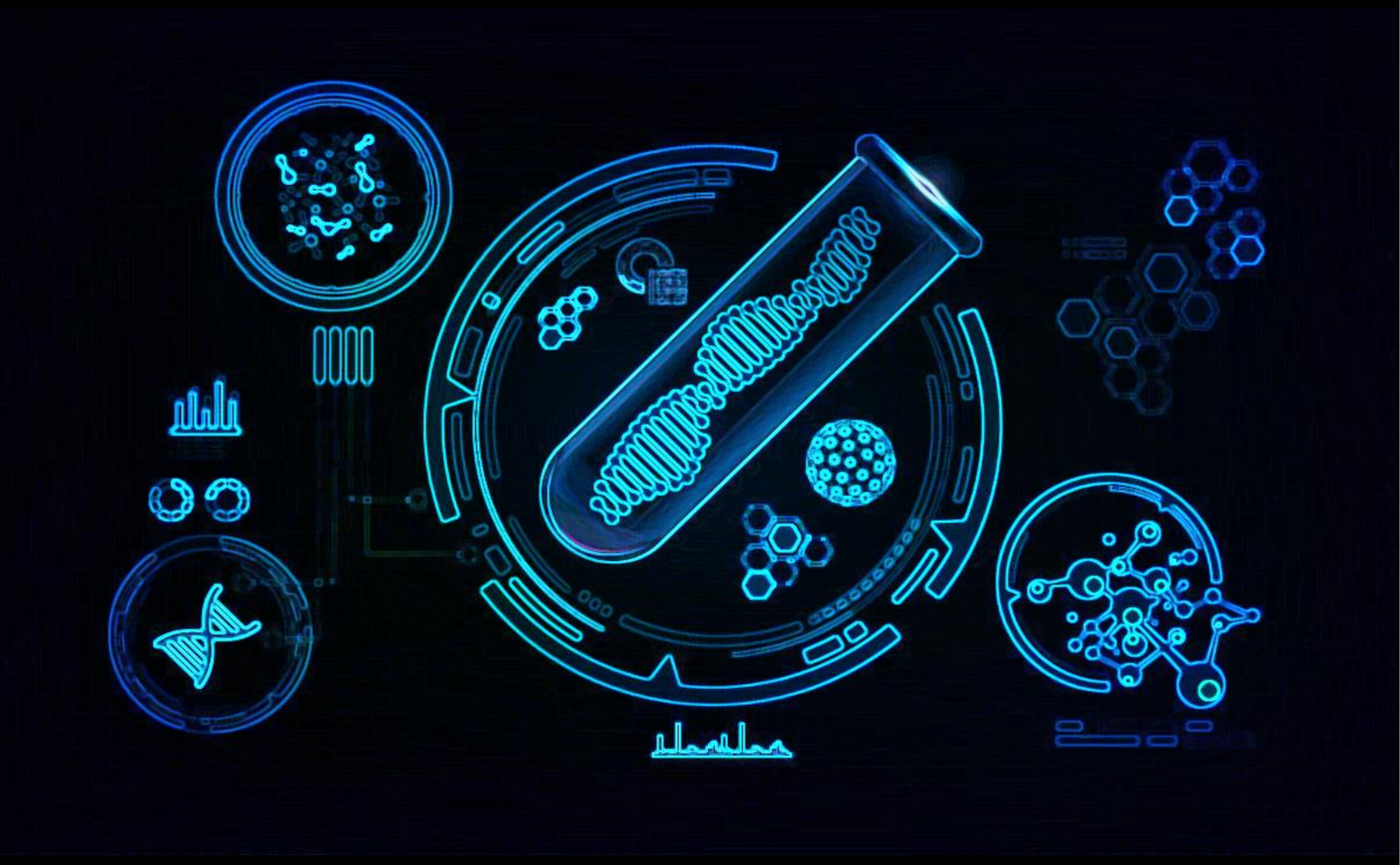
Liu, X. (2022, February 15). *To explore drug space smarter: artificial intelligence in drug design for G protein-coupled receptors*. Retrieved from https://hdl.handle.net/1887/3274010

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3274010

Note: To cite this publication please use the final published version (if applicable).

Chapter 1

General introduction



About this thesis

Drug discovery is a time- and resource-consuming process; from original idea to the regulatory approval of the finished product tends to take more than 10 years and bring costs in excess of $1 billion [1]. In order to decrease the cost and save time in this process, a plethora of computational methods have been developed [2,3]. Benefiting from the rapid growth of high throughput technologies, modern pharmacology has become a data-rich field with high-dimensional biological data accumulated in public databases, e.g. compound profiling data, gene expression data, transgenic phenotyping data, proteomics data, publications and patent information, *etc.* [4]. In order to analyze these 'big data' [5], artificial intelligence (AI) approaches are increasingly exploited in diverse scenarios to accelerate drug discovery [6].

AI is defined as the simulation of human intelligence to make machines think like humans and mimic their actions such as learning, reasoning, perception, and problem solving [7]. A typical subset of AI is machine learning (ML), which is trained on a large amount of (curated) data, and used to make predictions or decisions without being explicitly programmed [8]. One of the most popular ML methods are neural networks which consist of multiple layers of neurons (input, hidden, and output layers) that are mimics of neural activity in the human brain. Recently, neural networks have become extremely 'deep' because more and more hidden layers are added and organized with different architectures [9]. These so-called 'deep learning' methods have achieved major breakthroughs in e.g., image recognition, natural language processing, decision making, and other data-rich domains [10].

With the development of deep learning and rapid advances made in computational hardware, AI has further expanded its application scope in drug discovery [11]. In this thesis, I will introduce some of my projects about the application of AI in *de novo* drug design for G protein-coupled receptors and discuss its possible role in the future.

1.1. AI in drug discovery

The motivation behind launching a drug discovery project typically is that there is no or in some form limited suitable cure available to treat an existing disease or to meet clinical needs. The process of drug discovery can be seen as 'serendipity', *i.e.* it is about searching the optimal molecules from a huge chemical space (comprised of 10^{33}~10^{60} 'drug-like' molecules) [12]. In order to find drug candidates that are efficacious, safe, and meet clinical and commercial needs, the whole process of drug discovery includes target identification and validation, compound screening, lead optimization, preclinical development and finally the selection of drug candidates for clinical trials (Fig. 1.1) [13-15]. In all stages of drug discovery and development, AI algorithms are being developed and utilized in many aspects, such as understanding disease mechanisms to identify novel targets, providing target-disease associations, predicting properties of compounds, lead compound design and optimization, developing new biomarkers for prognosis, analyzing biometric and other data from wearable patient monitoring devices [14].

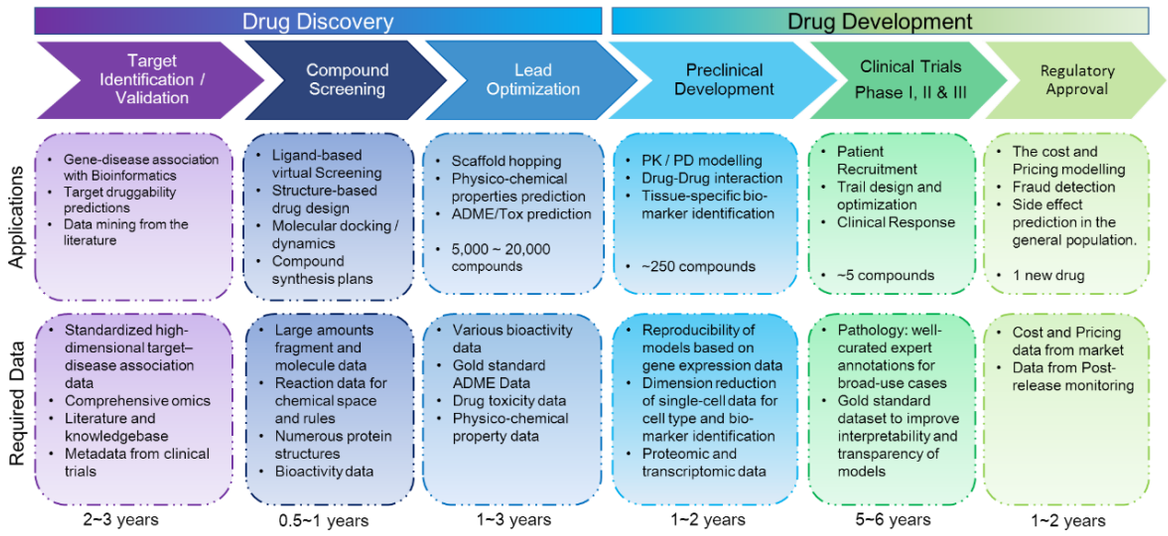


Fig. 1.1: Applications of AI in the pipeline of drug discovery and the required data. The information in this figure is collected from Ref. [14,15].

The initial phase of drug discovery starts from developing a hypothesis that the candidate drugs will lead to a therapeutic effect in a disease state when a protein or pathway is the inhibited or activated. Therefore the first and foremost step in developing a novel drug is

target identification and validation. AI approaches are extensively applied in gene-disease association modelling. For example, mRNA/protein levels could be examined to determine if they are expressed differently in disease and it thus can reflect whether they are correlated with the exacerbation or progression of diseases [16]. In addition, data mining and meta-analysis have resulted in a significant increase in target identification with literature and knowledge bases that contain successful and failed trails [17,18].

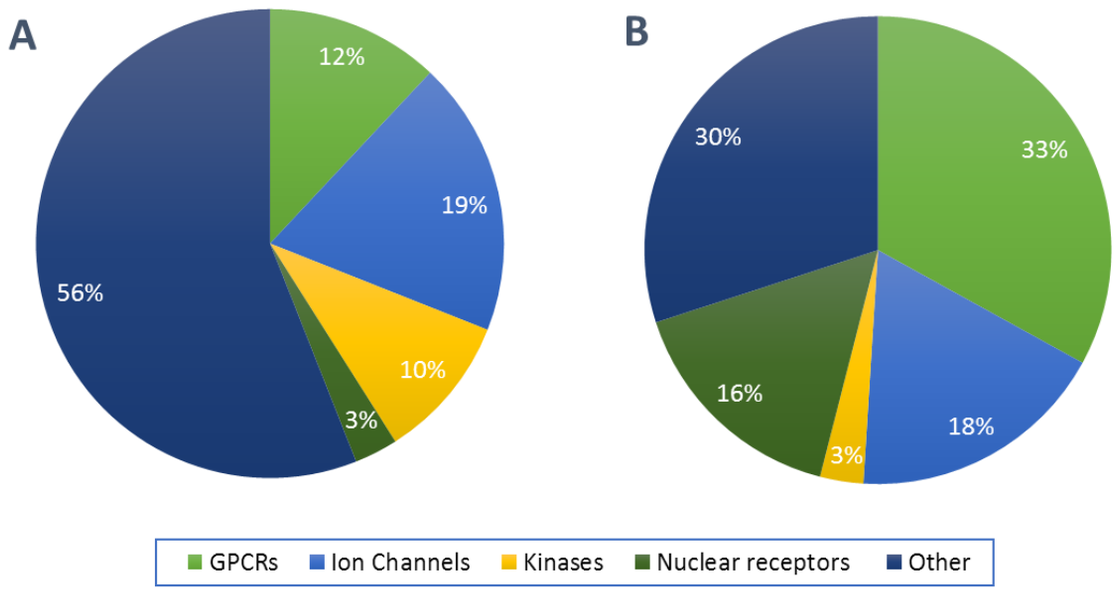


Fig. 1.2: The percentage of human drug targets in major protein families (A) and the proportion of small-molecule drugs targeting these families in human (B). The figure is adapted from Ref. [19].

After being identified, the role of these targets in a disease is validated using physiologically relevant in vitro/vivo models. An important question is how likely it is that a (small molecule) drug can be developed for the selected target, *i.e.* these proteins should bind small molecules as potential drugs. Through systems biology analysis druggable proteins have also been found in protein-protein interaction networks and these tend to be highly connected. For further investigations information on the 3D structure of the targets is very useful and can be obtained through X-ray crystallography [20], nuclear magnetic resonance spectroscopy [21] or cryogenic electron microscopy [22] technologies. It is worth mentioning that AI methods have recently been shown to be key in the prediction of 3D structure. For example, the DeepMind team exploited deep learning models to develop

the AlphaFold algorithm which has dramatically improved the accuracy of protein structure prediction [23]. It is now known that certain target classes are more amenable to small molecule drug discovery (Fig. 1.2), such as G protein-coupled receptors (GPCR), ion channels, kinases, *etc*. [19]. Despite the ultimate validation of the target in later clinical trials, target validation in the early phage is essential to make efforts on the completion of promising projects.

After the determination of the target, an appropriate organic small molecule library will be screened to search for feasible compounds as "hit compounds" that can be optimized into "drug leads" that can inhibit or activate the given target after binding. Besides random/trial-and-error experimental high throughput screening (HTS) technologies necessitating complex laboratory automation, virtual screening has also become a common paradigm to discover drug hits [24]. With structural and bioactivity data being accumulated, the quantitative relationship between structure and activity (QSAR) can be modelled with ML methods to predict the activity of the given molecules rapidly. In order to improve the application scope and the accuracy of ligand-based virtual screening, van Westen et al. proposed Proteochemometric modelling (PCM), updating classical QSAR models with target information, which was able to deal with different protein targets within one model [25,26]. In addition, deep learning can increase the performance substantially, when predicting the properties and activities of small organic molecules [27]. If the 3D structure of a protein target is available, molecular docking is also an effective method for structure-based virtual screening, which allows for the analysis of the ligand-target interaction by minimizing objective functions [28]. In this way it can provide a more mechanistic explanation of the interaction between ligand and target. Furthermore, molecular dynamics is applied as a computational method to study the dynamic mechanism of the drug-target interaction by calculating continuous motions using force fields [29].

For virtual screening, compound libraries need to be assembled containing small organic molecules that obey chemical guidelines such as the Lipinski Rule of Five [30]. The molecules in the library can be collected from public databases, such as ChEMBL [31],

ZINC [32], PubChem [33], *etc.*. On the other hand it is also common strategy to *de novo* design drug molecules with computational methods. This direction is my major study and I will give a detailed description of AI-based approaches applied in drug *de novo* design and make a comparison between different categories of methods in the next chapter. In order to run an experimental assay after virtual compound screening or *de novo* design, these drug leads are required to be synthesized in reality. AI algorithms can also be used to predict chemical synthesis routes which is comprised of "reversed" reactions decomposing the molecules (retrosynthesis). For example, Segler et al. scored the tree nodes in conjunction with deep learning and Monte Carlo tree search (MCTS) to search for the most promising synthetic pathways [34].

The selected drug leads need to be optimized to maintain favorable properties while eliminating deficiencies in the lead structure. In other words, drug discovery is a classical multi-objective optimization problem. The lead structure can be optimized either by changing or adding substituents through rational drug design or modifying the basic scaffold via scaffold hopping to avoid patent conflicts. In addition to the efficacy of drug leads, safety is a second important aspect to be persecuted. Typically, absorption, distribution, metabolism and excretion (ADME) are common metrics to evaluate drug safety. The principles of ADME are also important parameters in pharmacokinetics / pharmacodynamics (PK/PD) modelling, which is another approach to determine the toxicity and the drug dose required for the desired effect [35]. Side effects of drugs are often caused by binding to unwanted targets. For example, hERG (human Ether-à-go-go-Related Gene), a potassium ion channel, has an inclination to bind many drug molecules because of its large ligand binding pocket [36]. It may result in long QT syndrome (a change to the heart rhythm that can lead to fast chaotic heartbeats) when hERG is inhibited by potential drug candidates [37].

In the later stage of drug development, AI-based approaches for biomarker discovery have demonstrated to lead to a better understanding of the molecular mechanisms of drug effects and to identify the right drug for the right patients [38]. It will be most beneficial to build

and validate such AI models on datasets collected in the preclinical stage [39]. After being validated using independent datasets, corresponding biomarkers can be applied to stratify patients and identify potential indications [14]. In spite of the successful use of AI in early drug discovery and development there are a number of key issues that still need to be further addressed. Because they lack transparency, ML methods are often called 'black-box' approaches, which sparked criticism from end-users in the clinical adoption [40]. Interpretability is indeed a general weakness of ML-based predictive models [41]. One of the other key issues is the generalizability, namely that models need to be validated in the context of multi-site, multi-institutional datasets to prove the robustness of their predictions beyond the original training set. Faced with these challenges in drug discovery, the scientific community has nevertheless been making great contributions, including model training with parameter optimization approaches [42], interpretation of prediction results and deduction of their biological insights [43], and model reproducibility [44].

1.2. G protein-coupled receptors

G protein-coupled receptors (GPCRs), containing a characteristic structure of seven trans-membrane helices, are the largest family of cell-surface receptors. The superfamily consists of almost 900 members encoded by approximately 4% of human genes, which are classified into five families: rhodopsin, secretin, glutamate, adhesion and frizzled/Taste2 [45]. When binding signaling molecules from the extracellular environment, such as hormones or neurotransmitters, GPCRs will be activated. This activation will cause a conformational change and in turn activate the associated G protein by exchanging bound GDP for GTP [46]. Subsequently, α subunit of the G protein, bound with GTP, dissociates from the β and γ subunits to further regulate intracellular downstream biological pathways [46]. GPCRs are important switches to determine on-off states of numerous signaling pathways and they are involved in many biological processes, such as cell survival, proliferation, motility, *etc.* [47]. The aberrant activity or expression of certain GPCRs also contributes to some severe diseases, such as type II diabetes, Alzheimer's disease, hypertension, and heart failure [48]. In addition, the role of GPCRs in tumor growth, progression, and metastasis formation has become more apparent. Therefore,

approximately 34% of all FDA approved drugs have a GPCRs as drug target (Fig. 1.2) [48].

1.3. Adenosine receptors

Adenosine receptors (ARs) are a class of purinergic G protein-coupled receptors with adenosine as the endogenous ligand. Adenosine is an essential component for all of life because it is one of the four precursors to nucleic acids and one of its derivatives (*i.e.* ATP) is the "molecular unit of currency" for energy transfer in metabolic processes. In addition, adenosine can also form a signaling molecule (*i.e.* cAMP) to modulate a variety of biological pathways which are activated by the adenosine receptors when binding to adenosine. There are four subtypes of adenosine receptors namely A_1, A_{2A}, A_{2B} and A_3 which are widely distributed in human tissues and have been implicated in many physiological and pathological functions [49]. These dysfunctions include lipolysis, cardiac rhythm and circulation, immune function, renal blood flow, sleep regulation and angiogenesis, as well as inflammatory diseases, neurodegenerative disorders and ischemia–reperfusion damage [50].

As promising therapeutic targets ARs have been studied for a long time. First of all, adenosine itself can act as an agonist for the treatment of supraventricular tachycardia [51]. In addition, caffeine is the most commonly consumed “drug” in the world, but this AR antagonist is also used for treating apnoea in premature infants [52]. Interestingly, an inverse correlation between caffeine consumption and risk of Parkinson’s disease has been demonstrated [53]. However there are some serious challenges when taking ARs as drug targets due to their ubiquitous expression and the complexity of adenosine signaling throughout the human body. Tissue- and target-specific issues should be carefully taken into consideration [50]. Due to the complexity of adenosine signaling, it is crucial to understand the disease process when ARs are targeted in different cellular elements and disease courses

1.4. Research questions

The aim of this thesis is to apply AI approaches to increase the efficiency of the drug *de*

novo design process and thereby decrease cost and save time. Faced with the complexity of drug discovery and rapid development of AI technologies, I will investigate the following research questions which will be addressed in the following chapters:

1) Can AI support *de novo* drug design reliably and suggest active molecules for a single target as drug candidates?
2) Can AI be adjusted to polypharmacology to design desired molecules that bind multiple targets in order to balance efficacy and safety?
3) Can we improve the generality of AI models to design active molecules based on user-provided information such as fragments?
4) How to make these computational methods easily accessible by users who are not experts in computer coding skills?

1.5. Thesis outline

Up to now, there are numerous computational methods available for *de novo* drug design. In **Chapter 2**, I give a systematic overview of these methods, including optimization methods and deep learning methods. I conclude with describing the advantages and disadvantages of these methods and propose some possibilities of their combinations.

In **Chapter 3**, I introduce the first version of my proposed method, ***DrugEx***, which is a deep learning-based model for *de novo* drug design. In this method, a recurrent neural network was implemented to construct the generator. During the training process, the generator acted as the agent and the predictor acted as the environment interplay under the reinforcement learning framework. Here we add another pre-trained generator as the exploration strategy to improve the diversity of generated molecules. To evaluate the performance of my proposed method, the $A_{2A}AR$ is taken as an example, and most of the generated molecules are presumed to be active and located in the same region of chemical space occupied by ligands in the training set.

In **Chapter 4**, ***DrugEx*** was updated to the second version to include polypharmacology. In the first version it exclusively dealt with a single objective, while this second version has

the ability to deal with multiple objectives. The concept of evolutionary algorithms was merged into our method such that *crossover* and *mutation* operations were implemented by the same deep learning model of the *agent* to update the exploration strategy. In this chapter, three protein targets were chosen for the case study (two adenosine receptors, A_1AR and $A_{2A}AR$, and hERG in this study). Scores for all objectives provided by the *environment* constructed by three predictors were used for constructing Pareto ranks of the generated molecules with non-dominated sorting and Tanimoto-based crowding distance algorithms. The results demonstrate a generation of compounds with a diverse predicted selectivity profile toward multiple targets, offering the potential of high efficacy and lower toxicity.

In the first two versions our proposed method can handle multiple objectives, but these objectives have to be fixed during the training process. In other words, if the objectives are changed, the model has to be retrained. In **Chapter 5**, ***DrugEx*** was updated again for scaffold-constrained molecule generation. Here, we use end-to-end deep learning methods to implement the new ***DrugEx*** model, which has the ability to generate molecules containing pharmacological the scaffold containing multiple fragments given by users. We also trained the generator into reinforcement learning framework to ensure that the generated molecules are most likely active towards $A_{2A}AR$.

In order to access the abovementioned methods conveniently, we developed a web-based online toolkit. In **Chapter 6**, a detailed description of this toolkit is provided, which constitutes a graphic utility interface named ***GenUI***. It contains two main parts: client and backend components. In the backend components, we use Docker to make the installation automatically without complex configuration. In addition, it contains managers to distribute the computational resources and dispatch different tasks from different users in the task queue. In the client components, users can easily create their project and schedule a variety of tasks automatically, including data collection and preprocessing, QSAR modelling, Generator training, novel molecule design and chemical space visualization. The generated data for each user is stored in the server, and users can easily download it into a local machine.

Finally, I will draw general conclusions in **Chapter 7**. Moreover, I will also put forward some key points for future directions and expected trends regarding computational methods in *de novo* drug design.

References

1. Paul SM, Mytelka DS, Dunwiddie CT, Persinger CC, Munos BH, Lindborg SR, Schacht AL (2010) How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nat Rev Drug Discov 9 (3):203-214. doi:10.1038/nrd3078
2. Liu X, IJzerman AP, van Westen GJP (2021) Computational Approaches for De Novo Drug Design: Past, Present, and Future. Methods Mol Biol 2190:139-165. doi:10.1007/978-1-0716-0826-5_6
3. Schneider G, Fechner U (2005) Computer-based *de novo* design of drug-like molecules. Nat Rev Drug Discov 4 (8):649-663. doi:10.1038/nrd1799
4. Macarron R, Banks MN, Bojanic D, Burns DJ, Cirovic DA, Garyantes T, Green DV, Hertzberg RP, Janzen WP, Paslay JW, Schopfer U, Sittampalam GS (2011) Impact of high-throughput screening in biomedical research. Nat Rev Drug Discov 10 (3):188-195. doi:10.1038/nrd3368
5. Howe D, Costanzo M, Fey P, Gojobori T, Hannick L, Hide W, Hill DP, Kania R, Schaeffer M, St Pierre S, Twigger S, White O, Rhee SY (2008) Big data: The future of biocuration. Nature 455 (7209):47-50. doi:10.1038/455047a
6. Fleming N (2018) How artificial intelligence is changing drug discovery. Nature 557 (7707):S55-S57. doi:10.1038/d41586-018-05267-x
7. Rajaram NS (1990) Artificial intelligence: a technology review. ISA Trans 29 (1):1-3. doi:10.1016/0019-0578(90)90023-e
8. Jordan MI, Mitchell TM (2015) Machine learning: Trends, perspectives, and prospects. Science 349 (6245):255-260. doi:10.1126/science.aaa8415
9. Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85-117. doi:10.1016/j.neunet.2014.09.003
10. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521 (7553):436-444. doi:10.1038/nature14539
11. Yang X, Wang Y, Byrne R, Schneider G, Yang S (2019) Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery. Chem Rev 119 (18):10520-10594. doi:10.1021/acs.chemrev.8b00728
12. Polishchuk PG, Madzhidov TI, Varnek A (2013) Estimation of the size of drug-like chemical space based on GDB-17 data. J Comput Aided Mol Des 27 (8):675-679. doi:10.1007/s10822-013-9672-4
13. Hughes JP, Rees S, Kalindjian SB, Philpott KL (2011) Principles of early drug discovery. Br J Pharmacol 162 (6):1239-1249. doi:10.1111/j.1476-5381.2010.01127.x
14. Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, Li B, Madabhushi A, Shah P, Spitzer M, Zhao S (2019) Applications of machine learning in drug discovery and development. Nat Rev Drug Discov 18 (6):463-477. doi:10.1038/s41573-019-0024-5
15. Ashburn TT, Thor KB (2004) Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discov 3 (8):673-683. doi:10.1038/nrd1468
16. Katsila T, Spyroulias GA, Patrinos GP, Matsoukas MT (2016) Computational approaches in target identification and drug discovery. Comput Struct Biotechnol J 14:177-184. doi:10.1016/j.csbj.2016.04.004
17. Bravo A, Pinero J, Queralt-Rosinach N, Rautschka M, Furlong LI (2015) Extraction of relations between genes and diseases from text and large-scale data analysis: implications for translational research. BMC Bioinformatics 16:55. doi:10.1186/s12859-015-0472-9
18. Kim J, Kim JJ, Lee H (2017) An analysis of disease-gene relationship from Medline abstracts by DigSee. Sci Rep 7:40154. doi:10.1038/srep40154
19. Santos R, Ursu O, Gaulton A, Bento AP, Donadi RS, Bologa CG, Karlsson A, Al-Lazikani B, Hersey

A, Oprea TI, Overington JP (2017) A comprehensive map of molecular drug targets. Nat Rev Drug Discov 16 (1):19-34. doi:10.1038/nrd.2016.230

20. Thomas JM (2012) Centenary: The birth of X-ray crystallography. Nature 491 (7423):186-187. doi:10.1038/491186a
21. Pellecchia M, Sem DS, Wuthrich K (2002) NMR in drug discovery. Nat Rev Drug Discov 1 (3):211-219. doi:10.1038/nrd748
22. Renaud JP, Chari A, Ciferri C, Liu WT, Remigy HW, Stark H, Wiesmann C (2018) Cryo-EM in drug discovery: achievements, limitations and prospects. Nat Rev Drug Discov 17 (7):471-492. doi:10.1038/nrd.2018.77
23. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Zidek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D (2021) Highly accurate protein structure prediction with AlphaFold. Nature. doi:10.1038/s41586-021-03819-2
24. Fox S, Farr-Jones S, Sopchak L, Boggs A, Nicely HW, Khoury R, Biros M (2006) High-throughput screening: update on practices and success. J Biomol Screen 11 (7):864-869. doi:10.1177/1087057106292473
25. van Westen GJ, Hendriks A, Wegner JK, IJzerman AP, van Vlijmen HW, Bender A (2013) Significantly improved HIV inhibitor efficacy prediction employing proteochemometric models generated from antivirogram data. PLoS Comput Biol 9 (2):e1002899. doi:10.1371/journal.pcbi.1002899
26. van Westen GJ, Wegner JK, Geluykens P, Kwanten L, Vereycken I, Peeters A, IJzerman AP, van Vlijmen HW, Bender A (2011) Which compound to select in lead optimization? Prospectively validated proteochemometric models guide preclinical development. PLoS One 6 (11):e27518. doi:10.1371/journal.pone.0027518
27. Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T (2018) The rise of deep learning in drug discovery. Drug discovery today 23 (6):1241-1250. doi:10.1016/j.drudis.2018.01.039
28. Kitchen DB, Decornez H, Furr JR, Bajorath J (2004) Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov 3 (11):935-949. doi:10.1038/nrd1549
29. Hollingsworth SA, Dror RO (2018) Molecular Dynamics Simulation for All. Neuron 99 (6):1129-1143. doi:10.1016/j.neuron.2018.08.011
30. Mullard A (2018) Re-assessing the rule of 5, two decades on. Nat Rev Drug Discov 17 (11):777. doi:10.1038/nrd.2018.197
31. Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, Light Y, McGlinchey S, Michalovich D, Al-Lazikani B, Overington JP (2012) ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res 40 (Database issue):D1100-1107. doi:10.1093/nar/gkr777
32. Irwin JJ, Shoichet BK (2005) ZINC--a free database of commercially available compounds for virtual screening. Journal of chemical information and modeling 45 (1):177-182. doi:10.1021/ci049714+
33. Cincilla G, Thormann M, Pons M (2010) Structuring Chemical Space: Similarity-Based Characterization of the PubChem Database. Mol Inform 29 (1-2):37-49. doi:10.1002/minf.200900015
34. Segler MHS, Preuss M, Waller MP (2018) Planning chemical syntheses with deep neural networks and symbolic AI. Nature 555 (7698):604-610. doi:10.1038/nature25978

35. Ahmad I, Huang L, Hao H, Sanders P, Yuan Z (2016) Application of PK/PD Modeling in Veterinary Field: Dose Optimization and Drug Resistance Prediction. Biomed Res Int 2016:5465678. doi:10.1155/2016/5465678
36. Sanguinetti MC, Tristani-Firouzi M (2006) hERG potassium channels and cardiac arrhythmia. Nature 440 (7083):463-469. doi:10.1038/nature04710
37. Milnes JT, Crociani O, Arcangeli A, Hancox JC, Witchel HJ (2003) Blockade of HERG potassium currents by fluvoxamine: incomplete attenuation by S6 mutations at F656 or Y652. Br J Pharmacol 139 (5):887-898. doi:10.1038/sj.bjp.0705335
38. van Gool AJ, Bietrix F, Caldenhoven E, Zatloukal K, Scherer A, Litton JE, Meijer G, Blomberg N, Smith A, Mons B, Heringa J, Koot WJ, Smit MJ, Hajduch M, Rijnders T, Ussi A (2017) Bridging the translational innovation gap through good biomarker practice. Nat Rev Drug Discov 16 (9):587-588. doi:10.1038/nrd.2017.72
39. Kraus VB (2018) Biomarkers as drug development tools: discovery, validation, qualification and use. Nat Rev Rheumatol 14 (6):354-362. doi:10.1038/s41584-018-0005-9
40. Min S, Lee B, Yoon S (2017) Deep learning in bioinformatics. Brief Bioinform 18 (5):851-869. doi:10.1093/bib/bbw068
41. Murdoch WJ, Singh C, Kumbier K, Abbasi-Asl R, Yu B (2019) Definitions, methods, and applications in interpretable machine learning. Proc Natl Acad Sci U S A 116 (44):22071-22080. doi:10.1073/pnas.1900654116
42. Angermueller C, Parnamaa T, Parts L, Stegle O (2016) Deep learning for computational biology. Mol Syst Biol 12 (7):878. doi:10.15252/msb.20156651
43. Finnegan A, Song JS (2017) Maximum entropy methods for extracting the learned features of deep neural networks. PLoS Comput Biol 13 (10):e1005836. doi:10.1371/journal.pcbi.1005836
44. Hutson M (2018) Artificial intelligence faces reproducibility crisis. Science 359 (6377):725-726. doi:10.1126/science.359.6377.725
45. Lv X, Liu J, Shi Q, Tan Q, Wu D, Skinner JJ, Walker AL, Zhao L, Gu X, Chen N, Xue L, Si P, Zhang L, Wang Z, Katritch V, Liu ZJ, Stevens RC (2016) In vitro expression and analysis of the 826 human G protein-coupled receptors. Protein & cell 7 (5):325-337. doi:10.1007/s13238-016-0263-8
46. Trzaskowski B, Latek D, Yuan S, Ghoshdastider U, Debinski A, Filipek S (2012) Action of molecular switches in GPCRs--theoretical and experimental studies. Curr Med Chem 19 (8):1090-1109. doi:10.2174/092986712799320556
47. Dorsam RT, Gutkind JS (2007) G-protein-coupled receptors and cancer. Nature reviews Cancer 7 (2):79-94. doi:10.1038/nrc2069
48. Hauser AS, Attwood MM, Rask-Andersen M, Schioth HB, Gloriam DE (2017) Trends in GPCR drug discovery: new agents, targets and indications. Nature reviews Drug discovery. doi:10.1038/nrd.2017.178
49. Jacobson KA, Gao ZG (2006) Adenosine receptors as therapeutic targets. Nat Rev Drug Discov 5 (3):247-264. doi:10.1038/nrd1983
50. Chen JF, Eltzschig HK, Fredholm BB (2013) Adenosine receptors as drug targets--what are the challenges? Nat Rev Drug Discov 12 (4):265-286. doi:10.1038/nrd3955
51. Delacretaz E (2006) Clinical practice. Supraventricular tachycardia. N Engl J Med 354 (10):1039-1051. doi:10.1056/NEJMcp051145
52. Moro S, Gao ZG, Jacobson KA, Spalluto G (2006) Progress in the pursuit of therapeutic adenosine receptor antagonists. Med Res Rev 26 (2):131-159. doi:10.1002/med.20048

53. Saaksjarvi K, Knekt P, Rissanen H, Laaksonen MA, Reunanen A, Mannisto S (2008) Prospective study of coffee consumption and risk of Parkinson's disease. Eur J Clin Nutr 62 (7):908-915. doi:10.1038/sj.ejcn.1602788