6-23-2022

Molecular Modelling a Key Method for Potential Therapeutic Drug Discovery

Omar Karkoutly

Anupam Dhasmana

Vijian Dhevan

Subhash C. Chauhan

Manish Tripathi

Molecular Modelling a Key Method for Potential Therapeutic Drug Discovery

Omar Karkoutly[1,2,5], **Anupam Dhasmana**[2,5], **Vijian Dhevan**[3,4], **Subhash C. Chauhan**[2,5], and **Manish K. Tripathi**[2,5]*

[1]*Department of Biology, College of Sciences, The University of Texas Rio Grande Valley, McAllen, TX 78539, USA*

[2]*South Texas Center of Excellence in Cancer Research, School of Medicine, University of Texas Rio Grande Valley, McAllen TX 78504, USA*

[3]*Valley Baptist Hospital, Harlingen, USA*

[4]*Department of Surgery, School of Medicine, University of Texas Rio Grande Valley, Edinburg, TX 78501, USA*

[5]*Department of Immunology and Microbiology, School of Medicine, University of Texas Rio Grande Valley, USA*

*Corresponding author: Manish Tripathi, South Texas Center of Excellence in Cancer Research, Department of Immunology and Microbiology, School of Medicine, University of Texas Rio Grande Valley, McAllen TX 78504, USA

ARTICLE INFO

ABSTRACT

The well-defined and characterized 3D crystal structure of a protein is important to explore the topological and physiological features of the protein. The distinguished topography of a protein helps medical chemists design drugs on the basis of the pharmacophoric features of the protein. Structure-based drug discovery, specifically for pathological proteins that cause a higher risk of disease, takes advantage of this fact. Current tools for studying drug-protein interactions include physical, chromatographic, and electrophoretic methods. These techniques can be separated into either non-spectroscopic (equilibrium dialysis, ultrafiltration, ultracentrifugation, etc.) or spectroscopic (Fluorescence spectroscopy, NMR, X-ray diffraction, etc.) methods. These methods, however, can be time-consuming and expensive. On the other hand, *in silico* methods of analyzing protein-drug interactions, such as docking, molecular simulations, and High-Throughput Virtual Screenings (HTVS), are heavily underutilized by core drug discovery laboratories. These kinds of approaches have a great potential for the mass screening of potential small drugs molecules. Studying protein-drug interactions is of particular importance for understanding how the structural conformation of protein elements affect overall ligand binding affinity. By taking a bioinformatics approach to analyzing drug-protein interactions, the speed with which we identify potential drugs for genetic targets can be greatly increased.

Abbreviations: HTVS: High-Throughput Virtual Screenings; AYA: Adolescence And Young Adult; ER: Estrogen Receptor; PR: Progesterone Receptor; HER2: Human Epidermal Growth Factor Receptor; MS: mass-Spectroscopy; PDB: Protein Data Bank; NMR; Nuclear Magnetic Resonance; NPASS: Natural Product Activity & Species Source Database; ADMET: Absorption, Distribution, Metabolism, Excretion, and Toxicity

Introduction

Cancer is the second leading cause of death in the United States and serves as a great barrier to increasing life expectancy in many countries around the world [1]. In fact, below the age of 70, cancer is the first or second leading cause of death in 112 of 183 countries and is the third or fourth cause of death in another 23 countries [2]. Although the overall incidence and cancer mortality rate has been greatly reduced over the past couple of years due to advancements in early detection screenings and treatment options, cancer

remains prevalent [3]. Factors that increase risk of cancer, like obesity, diabetes, and aging, are on the rise, and as a result, people are at a higher risk of getting cancer, especially in the United States [4]. Furthermore, treatment options are becoming more limited due to the robust characteristics of aggressive cancer types that allows them to obtain drug resistance. This is especially true for the adolescence and young adult (AYA) population, as from 2006-2017, an increase in overall cancer incidence was seen, which is bad because treating younger cancer patients means that the cancer has more time to build up a resistance to the anticancer drug being used. Because of this, research into the development of new drugs for cancer treatment has been an ongoing effort.

Drug resistance for typical chemotherapeutic treatments is one of the main reasons these kinds of cancer therapies result in failure [5]. Despite the considerable progress being made in targeted cancer therapies, there is no treatment that is 100% effective in eliminating cancers, because of their innate resistance (to a broad range of anticancer drugs) or their acquired resistance (as existing therapies become more effective against them) [6]. One reason for this resistance may be due to cancer cell plasticity, which allows for cancer cells to switch between differentiated (limited tumorigenic potential) and undifferentiated (cancer stem cells) states [7]. Plasticity greatly contributes to tumor heterogeneity, which describes the differences between subpopulations of the same tumor type in different patients and is the reason why differential responses to therapies occur [8]. Furthermore, cancers are extremely complex, and their robustness [9,10] allows them to survive, adapt, and maintain their proliferative potential and functionality in the face of any internal or external perturbations (such as against a wide variety of anticancer therapies) [11]. One major solution to this is to develop novel drugs that are either better than their predecessors or that can result in deeper responses from being used sequentially or in combination with existing drugs [12].

Proteins as Therapeutic Targets

The diagnostic detection and measurement of cancer progression is essential for effective disease management, especially since the early stages of cancer have the highest therapeutic potential [13]. These early stages, however, are typically asymptomatic, and as a result, identifying novel biomarkers of various cancers is essential for early detection [14]. Cancer biomarkers can be any sort of tumor characteristic (like tumor tissue) or bodily response to cancer (like bodily fluids), that help indicate current or future cancer behavior, such as cancer risk, cancer type, and drug or treatment efficacy [15]. Not only can these biomarkers be used in diagnosis and early malignancy detection, but they may also be used as specific drug targets when designing novel anticancer drugs.

A prominent example of cancer biomarkers include the estrogen receptor (ER), the progesterone receptor (PR), and the human epidermal growth factor receptor (HER2), all of which are essential for the standard care of newly diagnosed, recurring, and malignant breast cancer patients [16]. Targeted HER2 drugs, such as Tratsuzumab, a humanized monoclonal antibody against HER2, have been developed and shown to increase time to progression and survival in both early stage and metastatic breast cancers [17]. As cancers continue to develop drug resistance, it becomes more critical to identify new biomarkers and develop more efficient drugs to help combat disease progression [18]. Current tools for doing this are time-consuming and inefficient, but the latest structural modeling tools can make this process easier and faster and can even take advantage of existing drug databases to potentially repurpose known drugs.

Methods for Drug Identification

The cancer proteome and metabolome (the entire set of proteins or small molecule metabolites, respectively, that are produced by a cancer), can contain important information relating to the discovery of novel biomarkers. Various methods such as electrophoresis, mass-spectroscopy techniques, and protein microarrays can be used to discover novel biomarkers. Additionally, many target-specific immunoassays and immunosensor techniques, including electrochemical, mass-sensitive, and optical have been used for tumor-related biomarker detection [19]. Traditional chemotherapies directly target the DNA of cells, but this can damage healthy cells, so modern approaches to anticancer drugs focus on molecular targeted therapy (i.e. monoclonal antibodies and small molecule inhibitors) to reverse abnormalities in the expression of kinases, tubulin proteins, extracellular matrix components, vascular targets, cancer stem cell pathways, or the tumor microenvironment (like acidity) as possible drug targets so that cancer cells can be selectively killed with a decreased toxicity towards normal cells [20].

Physical methods for studying drug-protein binding have been traditionally divided into either non-spectroscopic (like calorimetry, dialysis, filtration, electrophoresis and centrifugation) or spectroscopic (like UV and visible light absorption, NMR, X-rays, and fluorescence) [21]. These, however, have been replaced with more advanced and efficient methods such as a variety of mass-spectroscopy (MS) techniques including a direct approach, a structural approach, an enzymatic approach, an affinity-based approach, and a global proteomics approach [22]. These various MS approach make it possible to characterize drug target structures, screen large numbers of potential drug candidates (in metabolism and in pharmacokinetic studies), detect drug-target complexes, examine how protein structure is affected by the drug, and monitor the enzymatic activity of the target protein in relation to the drug [23]. Despite these major improvements in analyzing protein-drug interactions, these methods remain complex, time-

consuming, and costly [24]. As a result, more convenient tools, such as computational methods and structural modeling, should be used for estimating protein-drug binding affinities instead.

Structural Modeling and Drug Bank

The RCSB protein data bank (PDB) is an open access resource in biology and medicine for finding three-dimensional structural data on large biological molecules such as proteins and can be used to find the PDB ID for the crystal structure of a protein of interest (Ex: HER2) [25]. All 3D structures found on this resource are experimentally verified by either X-ray crystallography or nuclear magnetic resonance (NMR) and give an accurate depiction of the structure of the protein and/or its binding domain. This makes it perfectly valid for *in silico* use and for extrapolating that data towards *in vitro* and *in vivo* studies. Furthermore, a comprehensive list of potential inhibitors, agonists, and antagonists can be obtained from a variety of existing sources. For example, the Natural Product Activity & Species Source Database (NPASS) can be used to find potential nutraceuticals that are effective against the protein of interest or use that data to develop a novel drug that is analogous in structure [26]. Alternatively, the DrugBank library, a comprehensive open access database containing information on drugs, drug properties, and drug targets, may be used to screen approved and experimental drugs to find effective inhibitors of the protein of interest [27].

This can all be accomplished in a matter of days or weeks by performing a multi-layered High-Throughput Virtual Screening (HTVS) with the BIOVIA Discovery Studio Client 2020 software. In a multi-layered HTVS, several screening layers are performed in succession to identify the best molecule that will bind to the protein of interest [28]. This process takes place in three stages, each of which sequentially narrows down the list of potential inhibitors. First, a preliminary rigid docking analysis, using the DS LibDock extension, takes place by comparing the binding energies of the protein's crystalline structure to each ligand from the identified drug library in a rigid conformation to determine which ligands best fit at the binding site. Then, a flexible docking analysis, using the CDOCKER extension, takes place by mimicking the flexible nature of the binding site domain in nature and produces docked conformations with extreme precision. The number of potential drugs is then finalized after an Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) analysis is performed to determine the exact pharmacokinetic properties of the protein-drug interactions that were identified. Typically, after completion, anywhere from 10-20 drugs are identified and can then undergo further testing via *in vitro* and *in vivo* analyses to confirm their potential for inhibition. Figure 1 summarizes the steps in identification of drugs through structural modeling.
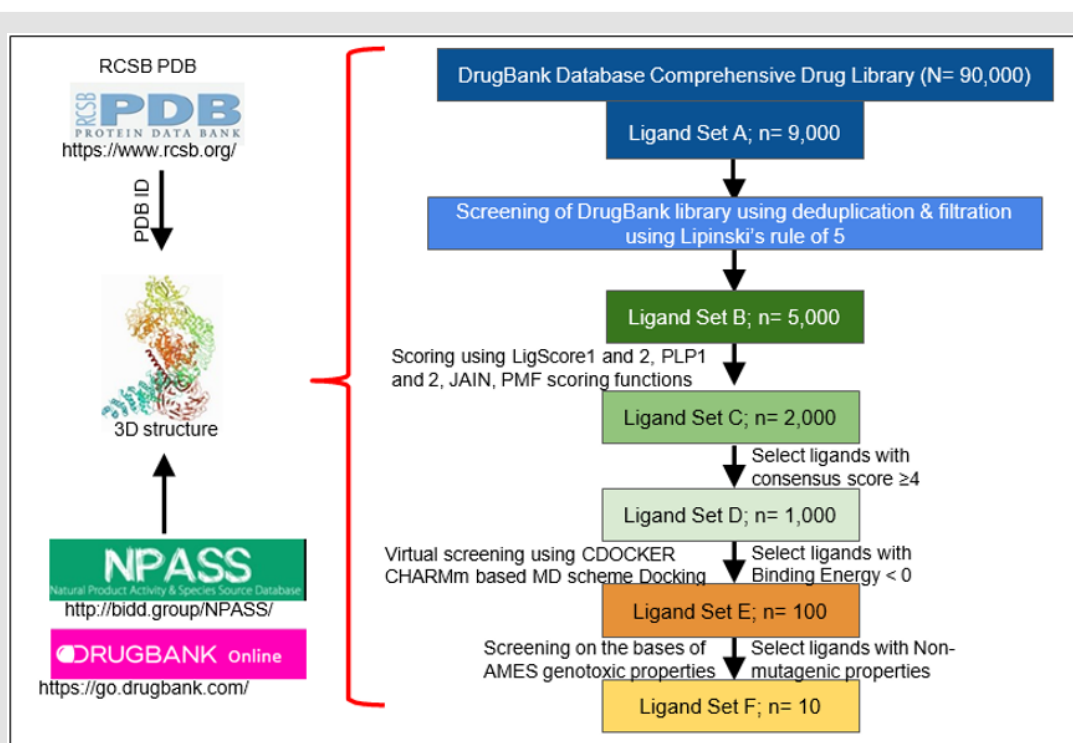

Figure 1: Steps in identification of drugs through structural modeling. RCSB PDB is analyzed for the specific protein ID. The 3D structure of the protein of interest, which is verified by X Ray crystallography or NMR, is used for further analysis. Identified binding domains are screened on different drug bank libraries on the basis of different parameters (n= is the depiction for the total target number).

Conclusion

Using these underutilized *in silico* tools will save a lot of time and money when considering the alternatives that are much more labor and resource intensive. For example, High-Throughput Screenings (HTS) are similar to HTVS except they are performed physically under wet-lab conditions. HTS is a drug discovery process that is popular amongst many pharmaceutical companies and takes advantage of robotics to autonomously screen a library of drugs and test their biological functions for pharmacological profiling [29]. The problem is that the equipment required (robots) and the bioactive drug screening libraries can cost tens of thousands of dollars, require technical training to use precisely, and can take months to finish a large screening analysis. Furthermore, from the ADMET analysis, HTVS can analyze pharmacokinetic parameters, like toxicity, of the identified drugs and their potential impact on certain tissues, like the liver, something which requires further testing after completion of HTS.

On top of all that, once this process is completed, a new drug is not discovered. Rather, these molecules are identified as "leads" for furthering and optimizing the drug discovery process, which takes too long to be feasible for immediate use. In fact, from lab experimentation to clinical testing and drug approval, novel drug development is a complex, time-consuming, and expensive process that can cost a manufacturer million, sometimes even billions [30] of dollars in resources and 12-15 years for completion [31]. By taking advantage of the latest bioinformatics techniques, such as HTVS, to analyze protein-drug interactions, small molecule inhibitors for cancer protein targets can be found with ease by repurposing existing drugs instead of waiting years for new drug approval. The potential for repurposing existing drugs as antagonists for novel cancer protein targets shows great promise and should be a more frequently explored option by pharmaceutical companies worldwide.

Acknowledgment

Contributors

OK and MKT wrote most of the manuscript. AD helped in OK training, MKT, AD, VD, and SCC contributed with editing and manuscript improvements. OK and MKT edited the figures. All authors read and approved the final manuscript.

Funding

Competing Interests

The authors declare that they have no competing interests.

References

1. Nagai H, Kim YH (2017) Cancer prevention from the perspective of global cancer burden patterns. Journal of thoracic disease 9(3): 448-451.

2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, et al. (2021) Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: a cancer journal for clinicians 71(3): 209-249.

3. Siegel RL, Miller KD, Jemal A (2020) Cancer statistics, 2020. CA: a cancer journal for clinicians 70(1): 7-30.

4. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet Tieulent J, et al. (2015) Global cancer statistics, 2012. CA: cancer journal for clinicians 65(2): 87-108.

5. Haider T, Pandey V, Banjare N, Gupta PN, Soni V (2020) Drug resistance in cancer: mechanisms and tackling strategies. Pharmacological reports : PR 72(5): 1125-1151.

6. Gottesman MM (2002) Mechanisms of cancer drug resistance. Annual review of medicine 53: 615-627.

7. da Silva Diz V, Lorenzo Sanz L, Bernat Peguera A, Lopez Cerda M, Muñoz P (2018) Cancer cell plasticity: Impact on tumor progression and therapy response. Seminars in cancer biology 53: 48-58.

8. Dagogo Jack I, Shaw AT (2018) Tumour heterogeneity and resistance to cancer therapies. Nature reviews Clinical oncology 15(2): 81-94.

9. Kitano H (2004) Cancer as a robust system: implications for anticancer therapy. Nature reviews Cancer 4(3): 227-235.

10. Kitano H (2003) Cancer robustness: tumour tactics. Nature 426(6963): 125.

11. Tian T, Olson S, Whitacre JM, Harding A (2011) The origins of cancer robustness and evolvability. Integrative biology : quantitative biosciences from nano to macro 3(1): 17-30.

12. Vasan N, Baselga J, Hyman DM (2019) A view on drug resistance in cancer. Nature 575(7782): 299-309.

13. Wulfkuhle JD, Liotta LA, Petricoin EF (2003) Proteomic applications for the early detection of cancer. Nature reviews Cancer 3(4): 267-275.

14. Etzioni R, Urban N, Ramsey S, McIntosh M, Schwartz S, et al. (2003) The case for early detection. Nature reviews Cancer 3(4): 243-352.

15. Ma Y, Gamagedara S (2015) Biomarker analysis for oncology. Biomarkers in medicine 9(9): 845-850.

16. Colomer R, Aranda López I, Albanell J, García Caballero T, Ciruelos E, et al. (2018) Biomarkers in breast cancer: A consensus statement by the Spanish Society of Medical Oncology and the Spanish Society of Pathology. Clinical & translational oncology : official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico 20(7): 815-826.

17. Serena Bertozzi APL, Luca Seriau, Roberta Di Vora, Carla Cedolini, Laura Mariuzzi (2018) Biomarkers in Breast Cancer. Biomarker - Indicator of Abnormal Physiological Process, Ghousia Begum. IntechOpen.

18. Ileana Dumbrava E, Meric Bernstam F, Yap TA (2018) Challenges with biomarkers in cancer drug discovery and development. Expert opinion on drug discovery 13(8): 685-690.

19. Wu J, Fu Z, Yan F, Ju H (2007) Biomedical and clinical applications of immunoassays and immunosensors for tumor markers. . TrAC Trends in Analytical Chemistry 26(7): 679-688.

20. Kumar B, Singh S, Skvortsova I, Kumar V (2017) Promising Targets in Anti-cancer Drug Development: Recent Updates. Current medicinal chemistry 24(42): 4729-4752.

21. (1971) C.F. C. Physical Methods for Studying Drug-Protein Binding. In: Brodie BB, Gillette JR, Ackerman HS (Edi.). Concepts in Biochemical Pharmacology. Handbook of Experimental Pharmacology, Springer, Berlin, Heidelberg.

22. Zinn N, Hopf C, Drewes G, Bantscheff M (2012) Mass spectrometry approaches to monitor protein-drug interactions. Methods (San Diego, Calif) 57(4): 430-440.

23. Campbell JL, Le Blanc JC (2011) Peptide and protein drug analysis by MS: challenges and opportunities for the discovery environment. Bioanalysis 3(6): 645-657.

24. Wanat K, Brzezińska E, Sobańska AW (2018) Aspects of Drug-Protein Binding and Methods of Analyzing the Phenomenon. Current pharmaceutical design 24(25): 2974-2985.

25. Burley SK, Berman HM, Kleywegt GJ, Markley JL, Nakamura H, et al. (2017) Protein Data Bank (PDB): The Single Global Macromolecular Structure Archive. Methods in molecular biology (Clifton, NJ) 1607: 627-641.

26. Zeng X, Zhang P, He W, Qin C, Chen S, et al. (2018) NPASS: natural product activity and species source database for natural product research, discovery and tool development. Nucleic acids research 46(D1): D1217-D1222.

27. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, et al. (2018) DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic acids research 46(D1): D1074-D1082.

28. Dhasmana A, Kashyap VK, Dhasmana S, Kotnala S, Haque S, et al. (2020) Neutralization of SARS-CoV-2 Spike Protein via Natural Compounds: A Multilayered High Throughput Virtual Screening Approach. Current pharmaceutical design 26(41): 5300-5309.

29. Pusterla T (2019) High-throughput screening (HTS).

30. DiMasi JA, Grabowski HG, Hansen RW (2016) Innovation in the pharmaceutical industry: New estimates of R&D costs. Journal of health economics 47: 20-33.

31. Van Norman GA (2016) Drugs, Devices, and the FDA: Part 1: An Overview of Approval Processes for Drugs. JACC Basic to translational science 1(3): 170-179.

Assets of Publishing with us

- Global archiving of articles
- I*mm*ediate, unrestricted online access
- Rigorous Peer Review Process
- Authors Retain Copyrights
- Unique DOI for all articles

https://biomedres.us/