PHARMA RESEARCH BULLETIN
Open Access Journal (ISSN: 2582-676X)

Review article
Volume 3, Issue 1, 2024, 11-29

Molecular Approaches and Significant Advances to Target Identification, Drug Discovery and Drug Development

Prabitha Prabhakaran[1], Sakshi Bhardwaj[2]*, Dhivya Shanmugarajan[2], Sharath Chandra[2].

[1] Department of Pharmaceutical Chemistry, JSS College of Pharmacy, JSS Academy of Higher Education and Research, Mysore, India.
[2] Department of Life Sciences, Altem Technologies, Bangalore, India

* Correspondence: sakshi.bhardwaj@altem.com

Received: 18 January 2024; Accepted: 28 February 2024; Published: 5 March 2024

Abstract: Molecular pharmacology is a multidisciplinary field that explores how drugs interact with molecular targets in biological systems to achieve therapeutic effects. It integrates molecular biology, biochemistry, pharmacology, cheminformatics, and bioinformatics. Advances in biotechnology and molecular biology have driven the discovery of new targets and the development of potent medications. The primary aim is to identify optimal therapies by understanding drug interactions with receptors, enzymes, and transport molecules. Drug design, based on knowledge of biological targets, involves rational approaches where molecules are tailored to bind effectively. Computer-aided drug design (in silico) accelerates this process using methods like molecular docking, dynamics, and pharmacophore modelling. The drug discovery and development process includes early discovery, preclinical studies, clinical trials, FDA approvals, and post-marketing surveillance. Computational techniques such as QSAR, ADMET profiling, and molecular modelling enhance target identification and ligand optimization. High-throughput screening, aided by robotics and automation, enables rapid testing of millions of molecules, leading to the discovery of therapies for cancer and neurological diseases. Biotechnology advancements have enabled the development of monoclonal antibodies and gene therapies, with technologies like CRISPR-Cas9 offering precise gene editing. Genomics and proteomics accelerate target discovery by identifying disease-related genes and proteins, fostering personalized medicine approaches. Molecular pharmacology continues to revolutionize drug design and discovery, providing new opportunities for treating complex diseases and advancing precision medicine.

Keywords: Molecular pharmacology, Drug designing, ADMET, QSAR, Target identification, CRISPR.

1. Introduction

Modern molecular biology and drug discovery, both aim to develop novel small molecules that potently and specifically alter the functions of target proteins in the biological system [1]. The challenging process of drug discovery can take decades and costs billions in resources for each clinical approval [2-3]. The testing of a variety of chemical compounds with the potential to be developed as medicinal treatments is constrained by the necessary infrastructure and financial expenditure [4]. The lack of systematic understanding is a significant factor for the failure of most experimental drugs that have entered clinical trials, as indistinct assessments of uncertain biological effects may have occurred before clinical stages. Moreover, post-approval problems might be just as catastrophic. Consequently, a crucial step in the approach to prevent both avoidable patient morbidity and pricey and time-consuming clinical failures is thoroughly comprehending the biological activity of possible therapies [2]. The key to effective drug designing and development is identifying and capturing the clinical spectrum of the disease as well as the precise function that a possible therapeutic target has in the disease. Target identification has improved in efficiency over the past few years, and it has contributed significantly to the discovery of numerous medications to treat many diseases [5]. One of the crucial milestones in the drug development process is the target identification phase [6]. Because of their interactions with small molecules, these target molecules are either enhanced or suppressed to produce a pharmacological effect.

Furthermore, insufficient early therapeutic target validation has been linked to both poor medication acceptance rates and expensive clinical failures [7-8]. Indeed, earlier proof-of-concept studies and the most efficient target validation were anticipated to cut phase II clinical trial attrition by 24%, resulting in a 30% reduction in the price of generating new molecular entities [9]. In 1905, John Langley proposed the hypothesis of receptive compounds, marking the beginning of the twentieth century's drug development process. The first rational drug discovery and development of chemical compounds were carried out by two scientists, Paul Ehrlich, the inventor of modern chemotherapy, and second, Sacachiro Hata created arsphenamine (Salvarsan) in 1910 by establishing a structure-activity relationship with the medication atoxyl, which was used to treat syphilis and the trypanosomiasis that causes sleeping sickness [10]. The term QSAR (quantitative structure-activity relationship) was first used in 1960 by Hansch & Fujita. Consequently, in the year 1970, the technology advanced because of the participation of molecular modelling and combinatorial chemistry. The finding of a specific target molecule protein, enzyme, receptor, or other molecules that are essential to the onset of a disease is done through the target identification process. Subsequently, drugs that can engage with these targets and alter their activity are developed [11]. However, the process of developing drugs that are specifically designed for any disease requires target identification. Researchers can create medications that interact with a target molecular structure in a highly precise way, reducing off-target effects and decreasing unusual side effects, by identifying the specific molecular targets [12].

The analysis of molecular-level biochemical and biophysical properties of therapeutics, as well as their molecular interactions and their impact on macromolecules and cellular structures and activities, is done scientifically. Many computational techniques, including cheminformatic, similarity searches and homology modelling techniques, are used throughout the target identification phases of the drug discovery process [13]. The two major approaches towards drug discovery and development include Structure-based drug designing/target-based drug designing and ligand-based drug designing. Target-based drug discovery (TDD) is enhanced by phenotypic drug discovery (PDD). PDD uses empirical, target-agnostic lead generation to find pharmacologically active compounds and new therapies that operate via novel drug mechanisms. The numerous methodologies for target identification and validation can be carried out by a variety of methods, including data mining Bioinformatics tools, which include discovering, choosing, and prioritizing possible disease targets, Genetic polymorphism and a link to the disease are examples of genetic association. Changes in mRNA and protein levels are included in the expression profile. Invitro cell-based mechanistic research and Functional screening may involve target-specific tools, knockdown, or knockout techniques [14-16].

2. Target Identification

The word "target," is referring to the proteins or including other biomolecules like DNA, RNA and peptides that the drug directly binds to, and which oversee the drug's therapeutic effectiveness. For the definition of "excellent" drug targets, druggability is not the only required quality. It's true that frequently, scientists are initially drawn to certain candidates based on their involvement in a biological process vital to the development of diseases [17-18]. The disease mechanism describes both genotype and phenotype change's impact in terms of ailment. Molecular targets are the naturally occurring, cellular or molecular structural elements engaged in the pathophysiology that causes disease. The researcher’s choice of targets is influenced by the ailment on which they focused. In the process of discovering and developing new drugs, the identification of the target molecule is a crucial necessity [19]. Target identification methodologies have greatly improved in recent years, and there are now more desirable pharmacological targets than before [20]. The two general methods for drug designing and development as mentioned earlier, both require target identification. Databases, biological assays, and machine learning/artificial intelligence algorithms play a vital role in target identification [21]. The tools and databases aid to identify the therapeutic target and comprehending the characteristics of small organic substances. The biological assays might gauge the impact of small compounds on the target. In comparison to conventional experimental tests, machine learning and artificial intelligence approaches may predict how medicines and targets interact [22]. These methods make the drug development process quicker and more understanding, especially while using computational approaches for the drug discovery and development process. The main target classes include transporter proteins, substrates, ribosomes, DNA, RNA, enzymes, ion channels, membrane receptors and monoclonal antibodies. The receptors are among them the most common pharmacological targets. Although most clinically effective medications are used to target proteins, there are now several new groups of pharmacological targets, including regulatory DNA elements, non-coding RNA, and nucleic acids (ncRNAs). In the areas of drug research and precision medicine, their significance is quickly expanding [23].

Phenotypic screening focuses on the objective of the discovery of bioactive small compounds in cellular systems that can reverse or treat aberrant phenotypes without knowing their exact mechanisms of action [24]. Numerous databases, including Swiss Target Prediction [25], PDB [26], UniProt [27], NCBI [28], Expasy [29], Rhea [30], IUBMB ExplorEnz Enzyme [31], BRENDA [32], KEGG [33], Metacyc [34], CheMBL [35], Pharos [36], BindingDB [37], can be used to choose a protein molecule or macromolecule. The three-dimensional structure of a protein molecule is shown in **Figure 1**.

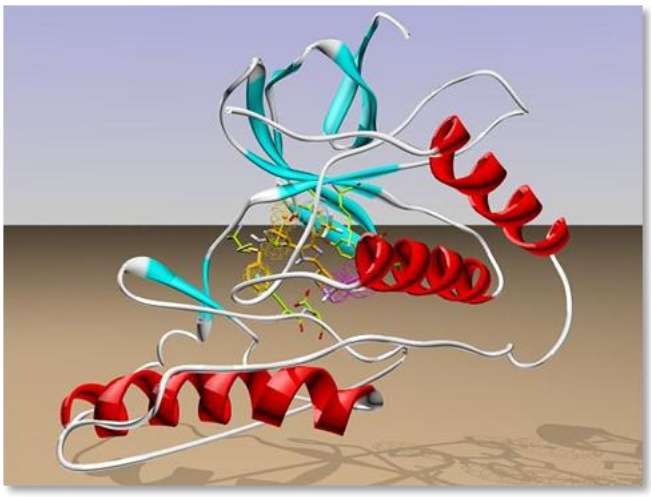

Figure 1: 3D structure of protein molecule.

Various data sources, like experimental, mechanistic, pharmacological, and, more recently, omics-based molecular profiles, are incorporated into modern bioinformatics approaches [38-39]. Omics technology has made it possible to examine biological samples at the level of genes, transcripts, proteins, metabolites, and their interactions in networks to find new targets. Three fields—genomics, proteomics, and metabolomics—are frequently used in the drug-development process. Whole genome sequencing, transcriptome analysis, and genome-wide association studies (GWASs) are crucial methods for identifying or validating new pharmacological targets since they allow for a methodical assessment of the therapeutic effectiveness and associated adverse effects. The study of human genetics allows for advances in drug discovery through a variety of avenues, including the understanding of disease onset and progression and providing a means for the identification of drug targets and even therapeutic agents by observing genes and their transcription products. Genes are at the centre of many biological processes [40]. The major ways that genomic methods are influencing drug research and development processes are through the application of genomics to fields like biomarkers, tissue expression profiling, side-effect profiling (SEP), pharmacogenomics (PGx), and genome-wide epigenetics. The application of proteomics to medication discovery and development similarly increases our understanding of how proteins interact, function, and are regulated. Proteomics is used across most of the drug development pipeline, including target and lead identification, chemical optimization, during clinical trials, and post-marketing analysis [41].

Although genomic techniques are widely established, proteomics has several significant benefits over genomics. Like proteins, RNA can fold in three dimensions, creating intricate structures that enable the binding of molecules with extreme specificity. A growing amount of attention is being paid to novel targets, which include ncRNAs (non-coding RNA). As targets for therapeutic development, ncRNAs can regulate gene expression [42]. Unfortunately, the process of developing new drugs is made much more difficult by the lack of knowledge regarding how ncRNAs act (or even if they do). When it comes to metabolomics, research has tended to concentrate on individual metabolic pathways and products due to the nearly overwhelming complexity of cells. This has improved our understanding of how cells work and revealed numerous important causes of disease. The action of a therapeutic molecule at a place other than the planned drug target is one of the main reasons why medications fail in clinical trials; this is a potentially serious issue that is typically not discovered until later in the development process [43]. Drug discovery and development can advance with the help of metabolomics, which provides a comprehensive "network" perspective of all biological pathways and an understanding of how these pathways interact with one another in a systems biology framework.

In the development of novel medications, the discovery of a pharmacological target and the subsequent validation procedure are the first steps. The data provided by the omics network shows how the target behaves in a normal setting, which aids in down streaming the target. Also, if the target has a vital role intervention's effects. Similarly, toxicology and pharmacological side effects need to be considered to prevent drug failure, as a cost-effective and time-saving process [44]. The main challenge is the integration of big data biomedical database's existing convergence with scientific literature. Several machine learning and artificial intelligence algorithms have been developed to lessen the process of gathering, processing, and analyzing data sources for drug discovery. These algorithms aid researchers working on the early stages of drug discovery by automatically identifying and extracting pertinent drug target-disease associations. The types of target molecules with their percentage usages are shown in **Figure 2**.

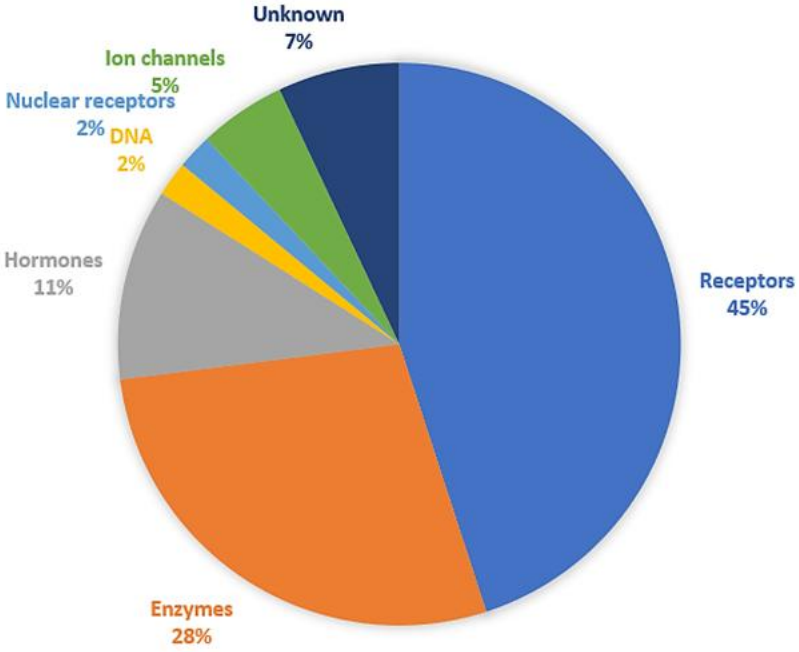

Figure 2: The types of target molecules with their percentage usages.

3. Genomics

There are several genomic approaches which allow DNA sequences for the drug designing and development processes. The most recent techniques that are being used in a pharmaceutical environment belong to the category of next-generation sequencing (NGS), and new approaches are constantly being developed. NGS is used to detect variants and mutations in DNA and RNA. The benefits of distinct sequencing chemistries, various sequencing matrices, and Bioinformatics technology are all combined with this technology [45]. A huge parallel sequencing of different length DNA or RNA sequences or even the entire genome is possible with such a combination in a reasonably short amount of time. It is a cutting-edge sequencing technique [46]. Deep sequencing is a cutting-edge method for monitoring, translation, and ribosome profiling offers a fresh way to track gene expression at the level of translation [47]. These days, ribosome profiling makes use of high-throughput sequencing to describe the complex pool of footprints that come from every translating ribosome in cells. Ribosome profiling offers a singular chance to map translation events objectively because each protected fragment represents the position of one ribosome on the mRNA [48]. Nowadays, Single Nucleotide Polymorphisms (SNPs) are the preferred marker due to their prevalence in almost all human groups. SNP markers have proven useful in human genomics, where complete genome sequencing resulted in the discovery of millions of SNP [49]. SNPs found among regulatory genes, transcripts, and expressed sequence tags have been utilized in functional genomic investigations [50]. The SNPs editing from gene sequences is shown in **Figure 3**. The Illumina (Solexa) sequencing, Roche sequencing, SOLiD sequencing, Ion semiconductor sequencing and SMRT sequencing, are the Massively parallel signature sequencing techniques being used for genomics in drug discovery and development processes. There are some other advanced techniques like Nanopore sequencing [51], Hybridization sequencing [52], and Sequencing with mass spectrometry [53] being popular for genomic sequences in drug discovery.

These various genomic approaches can be used as biomarkers, for tissue expression profiling, and toxicogenomic studies. Nevertheless, PGx is a genome-scale application of these methods that aims to clarify links between gene expression or single nucleotide polymorphism and therapeutic efficacy and toxicity [54].

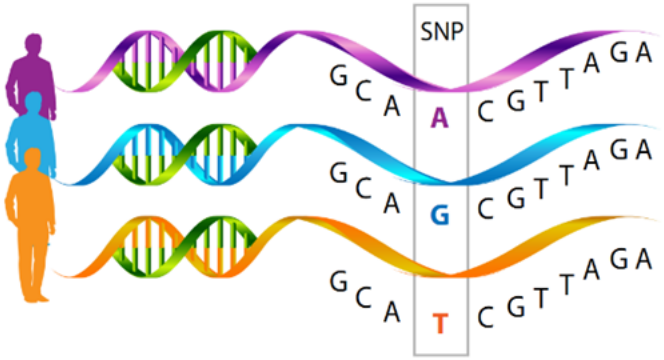

Figure 3: SNP editing from gene sequence.

4. Proteomics

The study of protein interaction, regulation, and function is known as proteomics in the field of drug discovery and development, and it has applications in a wide range of fields, such as target and lead identification, compound optimization, clinical trial design, and post-market analysis [55]. Proteomics has the following advantages over genomics: Firstly, gene expression does not always accurately represent protein regulation. Secondly, genomics ignores biological activities that follow gene transcription, like histone modification with acetyl, methyl, and phosphoryl groups at different amino acid residues [56-57]. The proteome will alter in response to physiological, environmental, and pharmacological interventions in addition to each other. Finding out information about proteins is difficult by nature. The main drawbacks are that protein concentration and activity may not always be correlated and that protein levels in cells change in response to a wide range of variables. Protein interactions with other proteins and molecules within the cell make analysis more difficult. Moreover, protein concentrations are frequently quite low, which makes it difficult to detect and accurately quantify data. There are many proteomic techniques in use, including Western blotting, ELISA, and chromatography-based methods. More advanced techniques include 2D gel electrophoresis, mass spectrometry, protein microarrays, yeast two-hybrid systems, Edman sequencing, quantitative techniques like ICAT, SILAC, and iTRAQ, isotopic labelling methods, as well as high throughput and bioinformatic methods like X-ray crystallography and NMR spectroscopy [58]. The development of methodologies is essential to the identification of proteomics, as protein microarrays, sometimes known as protein chips, enable high-throughput detection of small amounts of samples [59]. The most typical subset of analytic protein microarrays is the antibody microarray. Proteins are identified directly protein labelled after antibody capture. They are often employed to assess the degree of protein expression and binding affinities [60].

Some reverse-phase microarrays are being used to identify the changed or dysfunctional protein characteristic of a particular disease. The advancement in 2D gel electrophoresis is being used for the separation of proteins based on their mass and charge using the most effective and dependable two-dimensional polyacrylamide gel electrophoresis (2D-PAGE). The Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis of Proteins is a high-resolution method for dividing proteins based on their size, which makes it easier to approximate molecular weight [61]. For determining the three-dimensional structure of proteins, X-ray crystallography (Figure 4) is the method of choice. The size of the subunit that forms the crystal and the symmetry of the crystal packing is determined by processing the diffraction patterns that result from the exposure of the highly purified crystallized samples to X-rays. Many topics like the viral system, protein-nucleic acid complexes, and immunological complexes can be studied using X-ray crystallography. Also, precise information regarding the clarification of enzyme mechanism, drug design, site-directed mutagenesis, and protein-ligand interaction is provided by the three-dimensional protein structure [62]. Whereas the molecular weight of proteins can be ascertained by using mass spectroscopy to assess the mass-to-charge ratio (m/z) [63]. Nevertheless, the NMR is a prominent technique for examining molecular makeup, protein folding, and behaviour. NMR spectroscopy often comprises several steps, each utilizing a distinct set of extremely specialized procedures. In numerous fields of research, including structure-based drug design, homology modelling, and functional genomics, the protein structure is crucial [64].

Over the past few years, the use of bioinformatics for proteomics has become much more popular. The creation of a new algorithm that allows for the analysis of larger amounts of data with greater specificity and accuracy aids in the identification and quantification of proteins, making it possible to obtain detailed information about the expression of proteins. The key issue with these kinds of analyses is how to manage such a large amount of data. Finding connections between proteomic data and other omics technologies, such as genomics and metabolomics, is still challenging. Yet, database technology and new machine learning and artificial intelligence statistical algorithms are effective tools that could be helpful to get beyond these restrictions [65].

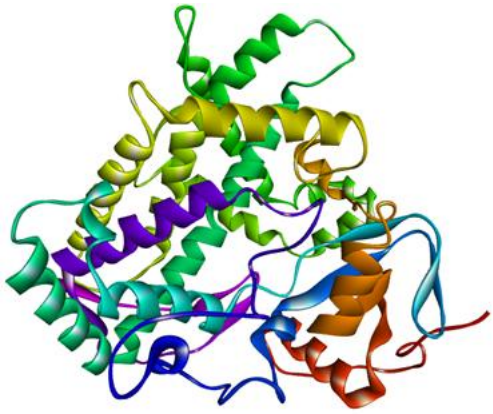

Figure 4: 3D crystal structure of the target molecule.

5. Metabolomics

The sheer variety, number, and concentration of metabolites and metabolite concentrations present in the cellular environment constitute the biggest obstacle to mapping the metabolome and its network of pathways. Yet, one significant advantage that metabolomics has over other 'omics is that because the pathways must adhere to precise stoichiometric limitations, metabolomics is more discriminating [66]. Moreover, metabolite concentration can be affected by both disease and pharmacological intervention, and small-scale changes in gene expression and protein concentration are amplified to cause significant large-scale changes in the metabolite concentration [67]. Chromatographic techniques, capillary electrophoresis, and NMR approaches are the three basic categories through which metabolite concentrations can be determined. To understand the metabolome, route databases are created using the information acquired from these methods. Through metabolomic profiling, we can better understand how metabolism and physiology work. Even though well-established laboratory biochemical procedures are fundamentally targeted quantitative metabolomics, some contend that metabolomics has long been an important component of medicine.

However, it is a novel approach to untargeted analysis using emerging quantitative methods that are unique. Moreover, metabolomics offers a paradigm shift in our perception of disease states as we move away from looking for single-molecule disease biomarkers and replace this with a search for more complex and dynamic patterns of metabolite concentrations. Metabolomics platforms and technology have a great deal of promise to be incorporated into standard clinical practice to support clinical judgements and to empower patients through a better understanding of underlying illness dynamics. In the end, metabolomics can provide a practical pathway for implementing the depth of innovation in treating human health and disease with the help of big data analytics and the improved translation of accurate mathematics to personalized medicine [68]. The processing of the enormous amounts of data that are generated and using this information to create a functional and coherent picture of the metabolome are two problems in metabolomic analysis. The large pathway databases are being created because of attempts to fuse newly generated information with existing data sets to get around the issue. Various databases are available for this like: The Human Metabolome Database, Lipid Maps, Mass Bank, National Institute of Standards and Technology, METLIN and the Kyoto Encyclopaedia of Genes and Genomes (KEGG) [69]. The most recent developments in systems biology can be used for medical applications by reconstructing and evaluating the fundamental mechanisms driving human biological activities, and this might all be accomplished by using computer-assisted methodologies in drug design and development. A list of various assessable databases of targets, ligands and protein-ligand complexes is given in **Table 1**.

Table 1: List of assessable databases for target, ligand and protein-ligand complex.

Type	Database Name	References
Target	SwissTarget Prediction	[25]
	PDB	[26]
	UniProt	[27]
	NCBI	[28]
	Expasy	[29]
	Rhea	[30]
	IUBMB ExplorEnz Enzyme	[31]
	BRENDA	[32]
	KEGG	[33]
	Metacyc	[34]
	CheMBL	[35]
	Pharos	[36]
	BindingDB	[37]
Drug	CheMBL	[35]
	Drug Bank	[70]
	PubChem	[71]
	Zinc	[72]
Protein-Ligand complex	AffinityDB	[73]
	iview	[74]
	FastDRH	[75]
	PDBligand	[76]
	PLD	[77]
	PDBSum	[78]
	Hic-Up	[79]
	Relibase	[80]
	MOAD	[81]
	Astex diverse set	[82]

Among all previously listed technologies, The Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated endonuclease 9 (Cas9) protein technology is gaining more popularity because this technology is a powerful gene-editing technique that won the Nobel Prize. All organisms' genomes can be altered with high precision and stability using this method. The most recent developments in technology include a genome library screening strategy that can find medication resistance and survival-essential genes by detecting function gain or loss. The flexible equipment enables genome screening for gene activation or inhibition and specifically targets non-coding regions, including promoters, miRNAs, and lncRNAs.

It was initially used as a weapon of natural defence against Streptococcus sp viral infections. By combining manually created trans-activating crRNA (tracrRNA) with CRISPR RNAs (crRNAs) to create single-guide RNAs, the researchers created the target DNA cleavage mechanism (sgRNAs) [83]. A double-strand break (DSB), a single-strand nick, or mutagenesis are the results of the endonuclease Cas9 protein cleaving the DNA at the target spot [84]. The Zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) are examples of previous gene-editing technologies that call for customizable, DNA sequence-binding modules coupled to the non-specific DNA endonuclease domain. It was useful to design new target locations for targeting proteins to modify any DNA sequence. CRISPR, in contrast to its forebears, uses a universal Cas9 protein that requires only a guided RNA (gRNA) to match the target. With this practical, quick, and adaptable genome editing approach, a potential library screen application to find crucial genes for cell survival and drug resistance is made possible [85]. Many high-throughput screens (HTS) can be created using CRISPR Cas9 editing to imitate the most precise targets in disease-relevant cells.

CRISPR technology is used by Charles River to create both pooled and arrayed HTS [86]. Nowadays, targeted CRISPR-Cas9 libraries covering the entire genome are used for critical gene screening or a group of genes with phenotypes or cellular activities are used to find targets. Machine learning and artificial intelligence algorithms made these cutting-edge technologies robust. All phases of the drug discovery process are moving more quickly thanks to CRISPR, which is also advancing the pre-clinical drug development stages. In conjunction with other high-throughput technologies, CRISPR seems to create the opportunity for extensive drug target screening to advance pharmacological research. Gene editing has become simpler and more accurate thanks to CRISPR, which has sped up the identification of therapeutic targets and perhaps sped up the drug discovery process. Hence, CRISPR technology plays a significant role in drug discovery, which will eventually lead to better and safer treatments. A description of the several omics sciences, such as genomics, transcriptomics, and proteomics. The "omics" cascade's final step, metabolomics, combines the environment's upstream input with the genome's downstream output is given in **Figure 5**.

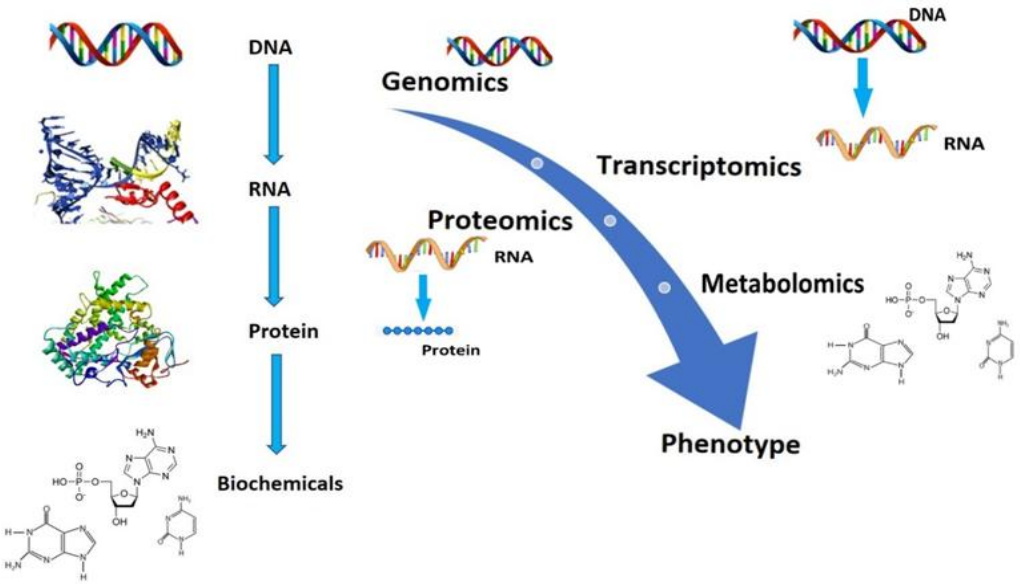

Figure 5: A summary of the several omics-sciences, including genomics, transcriptomics, and proteomics. Metabolomics, which is at the bottom of the "omics" cascade, represents both the downstream output of the genome and the upstream input from the environment.

6. Drug Discovery and Development

The healthcare systems constantly need new pharmaceuticals to treat unmet medical needs in a variety of therapeutic areas, and the drug industry primarily works to bring new drugs to market through the challenging process of drug research and development. The potential new medicinal entities are discovered through the process of drug discovery, which combines computational, experimental, translational, and clinical models. Drug discovery is still a drawn-out, expensive, challenging, and ineffective process with a high attrition rate of novel therapeutic discovery, despite advancements in biotechnology and understanding of biological systems.

The process of drug discovery includes several steps, including the selection of the target and its validation, the identification of hits, the generation and optimization of leads, and the identification of a candidate for further development. The search for potentially active compounds is usually the first step in the early drug discovery process. Once these substances have been discovered, testing for safety and efficacy must be done to see if they have a therapeutic effect on the intended ailment. Drug design, in its most basic form, is designing molecules that are complementary in charge and shape to the molecular target with which they interact and bind. In the big data era, drug design commonly relies on bioinformatics methods and computer modelling tools [87]. In conjunction with small molecules, biopharmaceuticals, particularly therapeutic antibodies, are a class of drugs that are becoming more and more significant, and computational techniques for enhancing the affinities, selectivity and stability of these protein-based therapeutics have made significant strides as well [88]. The preclinical studies using animal and cell-based models, human clinical trials, and regulatory approval are all steps in the development and discovery of new drugs. The identification of screening hits, medicinal chemistry, and optimization of those hits to improve their affinity, selectivity (to reduce the possibility of side effects), efficacy/potency, metabolic stability (to increase the half-life), and oral bioavailability are all essential components of modern drug discovery. Once a molecule satisfies all these criteria, medication development will start before clinical trials. There are numerous stages and steps involved in the drug discovery and development process (**Figure 6**). From discovery to approval, the expensive process of drug discovery typically takes more than ten years. Significantly, the related price and approval time vary depending on the drug being developed and the ailment it is intended to treat.

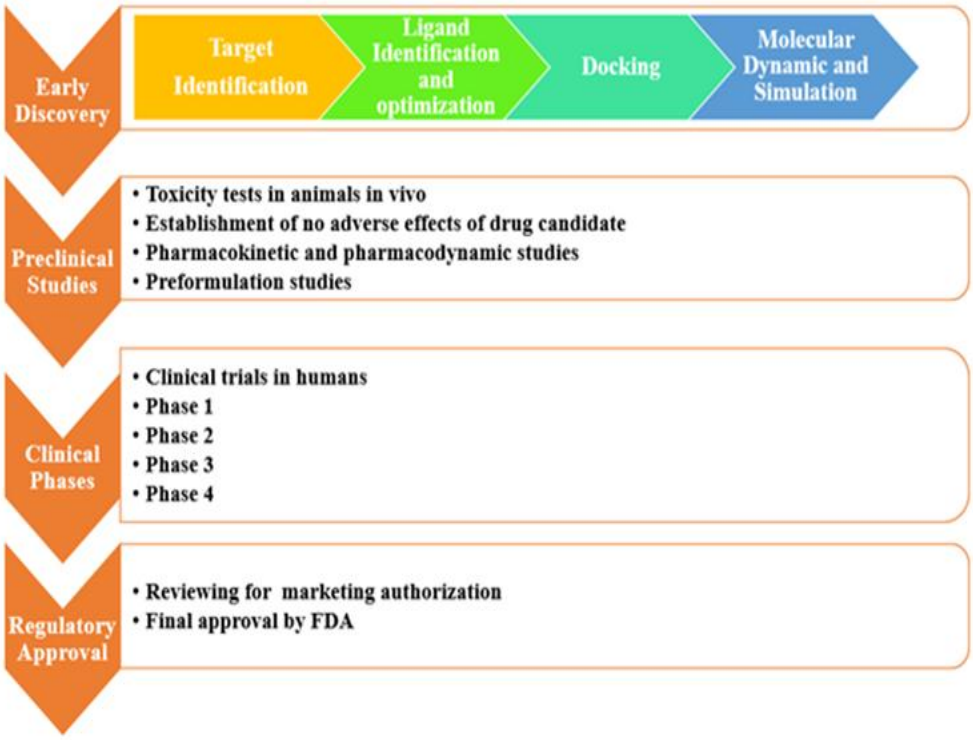

Figure 6: Drug Discovery and Development Process

On the other side, drug development includes improving chemical synthesis and formulation, doing animal toxicology studies, conducting clinical trials, and ultimately receiving regulatory approval. We will be discussing these by one as follows:

6.1 Early drug discovery

There are numerous steps (**Figure 7**) and tests involved in the early drug discovery process. To find and improve prospective leads to a particular target, researchers work together. Nowadays, in silico predictions or computational methods are being used for early drug discovery to process the following Target Identification and Validation, High Throughput Screening, Hit Identification, Assay Development and Screening, Hit-To-Lead (H2L), Lead Generation and Optimization, and In Vivo and In Vitro Assays as discussed below:

Figure 7: Steps involved in the early drug discovery process.

6.1.1 Target identification and validation

The significant step in the process of drug discovery is the identification of targets or macromolecules. Finding the biological cause of an illness and prospective targets for treatment is the crucial step in the discovery of a drug molecule. As it was previously mentioned, several target molecules in the biological system may have a role in a diseased condition and would be the target molecule in the drug discovery process of developing a therapeutic candidate against it. The molecular mechanisms that the target is intended to affect are then described after the target has been identified. An ideal target should be effective, secure, satisfy clinical and business criteria, and be "druggable." The methods used to identify targets may be based on concepts from various fields like molecular biology, biochemistry, genetics, or biophysics. Various approaches like data mining, genetic association, expression profiling, pathways and phenotypic analysis and functional screening are being used for the identification and selection of target molecules. Modern approaches and methods include disease association, bioactive compounds, cell-based models, protein interactions, signalling pathways analysis, functional analysis of genes, in vitro genetic manipulation, antibodies, and chemical genomics used by researchers to validate targets. The structure of a target molecule with its binding sites is shown in **Figure 8**.

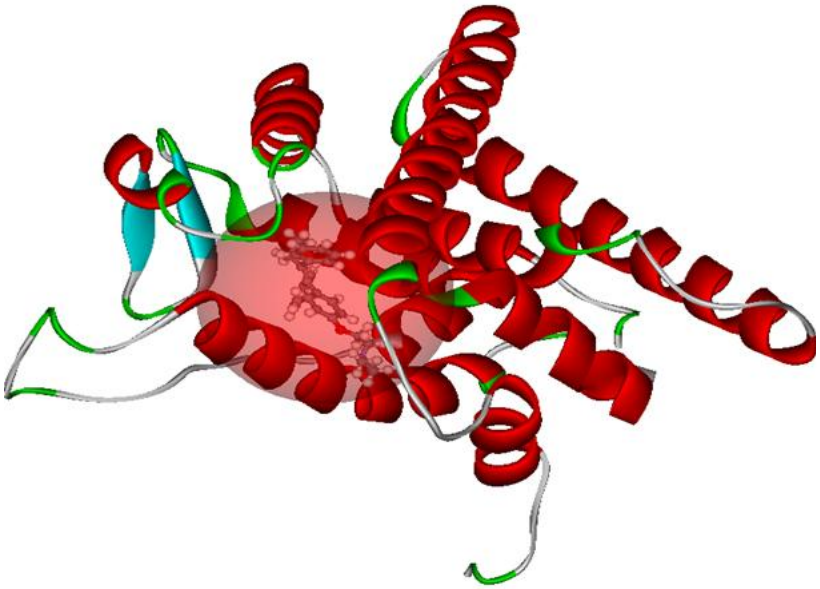

Figure 8: Protein structure with the identified binding site.

6.1.2 Lead identification and optimization

The process of finding or making a compound that can interact with the previously chosen target is known as lead compound identification. Researchers can do screening trials to find potential naturally occurring substances that can be converted into medicines. A synthetically stable, practical, drug-like molecule that is active in both primary and secondary assays and exhibits adequate specificity, affinities, and selectivities for the target receptor is referred to as a chemical lead. This step is to identify a molecule that will not interfere with other cellular processes while also targeting the predicted target. A lead compound's characteristics include its drug-likeness, synthetic viability, choice of mechanistic assays, evaluation of drug-resistant and efflux potential in vitro, proof of in vivo evaluation of the chemical class, and Pharmacokinetic and pharmacodynamic of the chemical class of drugs known from preliminary toxic effects or in silico investigations (**Figure 9**).

A drug ability evaluation is frequently performed to reduce the number of molecules that fail during the medication development process. A molecule must have the ability to bind to a particular target to be called druggable; nevertheless, the substance's pharmacokinetic profile addressing absorption, distribution, metabolism, and excretion is also crucial [89]. There are some assessable browsers available for ligand selection [70]. In addition, other tests, such as the Ames test and cytotoxicity assay, will assess the compound's potential toxicity. As the final phase in early-stage drug discovery, potential leads are assessed for a variety of qualities, such as selectivity and binding mechanisms, during lead optimization. For this downstream selectivity profiling and additional research, drug discovery researchers require quick ways to choose a smaller number of drug candidates. The comprehension and prediction of in vivo pharmacokinetics using in vitro tests have been made possible because of the development of high throughput DMPK (drug metabolism and pharmacokinetics) screens. The QSAR models and pharmacophoric models also play a major role in the down streaming of the ligand molecules.

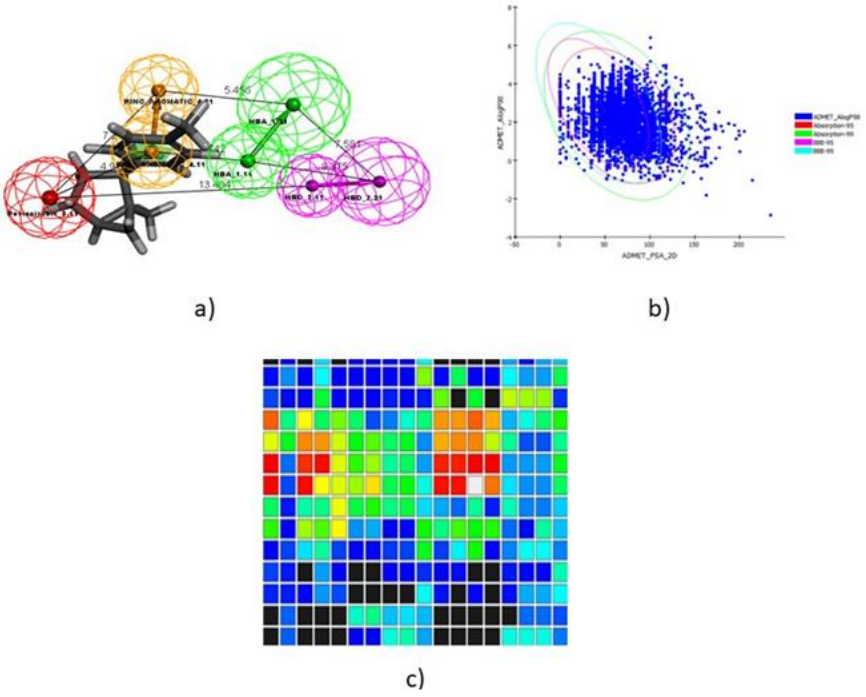

Figure 9: a) Pharmacophoric model generation for active molecules, b) Heat map generation between multiple ligand molecules and diverse target molecules, c) QSAR profiling of the ligand molecules.

6.1.3 Docking

Fundamentally, the goal of molecular docking is to use computational techniques to predict structural complements between ligand-receptor complexes. To perform docking, two related processes must be taken: first, the ligand conformations in the protein's active site must be sampled; next, these conformations must be ranked using a scoring function. It determines how a ligand and a target will ideally align when they are bound together to form a stable complex. It is thus possible to forecast the degree of connection or binding affinity between two molecules. The molecular docking method allows us to characterize how small molecules behave in the binding site of target proteins and to better understand basic biological processes by simulating the interaction between a small molecule and a protein at the atomic level. It’s a computational method for predicting the binding interactions between target and drug molecules to check their binding affinity. The flexibility of both the receptor and ligand, which permits one molecule to change its shape in response to the other, contributes to the size of this conformational space. The receptor is often kept stiff to restrict this broad conformational range. The scoring mechanism assesses the fitness of each sampled pose. The conformations which are kept after sampling are chosen based on fitness, which is then used to organize the retained postures according to how likely it is that they are correct. A list of the docked molecule's sorted poses is the result of docking. A small molecule's correct binding pose (**Figure 10**) must be identified to ascertain its binding affinity and to provide the possibility to use the stance for lead optimization [90].

There are several docking techniques, including flexible docking, fragment-based drug design, ligand-based drug design, and structure-based drug design.

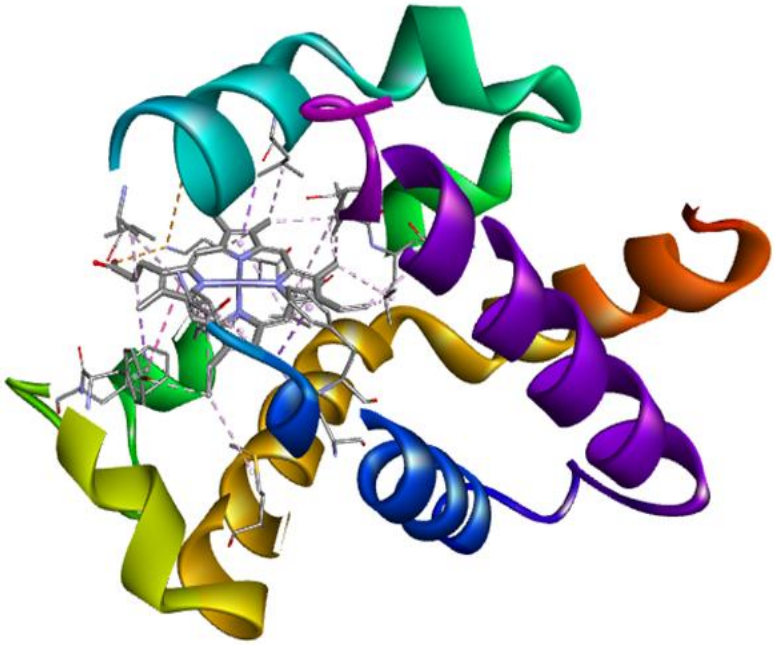

Figure 10: The docking pose of the Drug and target molecule.

6.1.4 Molecular dynamic and simulation

The molecular interactions that take place in biological settings give rise to biomolecular functions. It is difficult to comprehend how each component affects how cells work. Experimental methods have traditionally served as the primary entry point into the cellular environment. Researchers have created a comprehensive understanding of how a cell works by resolving biomolecular structures and investigating the kinetics of biomolecular processes, both in vivo and in vitro [91]. A biomolecule's atomic-level structure is incredibly useful and often yields significant insight into how the biomolecule behaves. Nevertheless, because the atoms in a biomolecule are constantly in motion, the dynamics of the individual molecules affect both their intramolecular connections and molecular function. Based on a broad model of the physics driving interatomic interactions, molecular dynamics (MD) simulations forecast how each atom in a protein or other molecular system will move over time. High-performance computing (HPC) and the ease of the fundamental Molecular dynamic simulation (**Figure 11**) method are two factors that contributed significantly to this astounding increase. A significant problem in system definition is solvent representation. Several strategies have been tested, however, the simplest strategy—the explicit and implicit depiction of solvent molecules—proves to be the most successful, even if it increases the size of the simulated systems [92].

Once the system has been created, the forces acting on each atom are determined by constructing force fields from equations, where potential energy is derived from the molecular structure. Despite being complicated equations, force fields are simple to calculate. Even for large systems, energy and force calculations are incredibly quick thanks to the force-field representation of molecular features' simplicity. Although not all force fields can represent all molecule types and parameters are not always interchangeable, simulations carried out utilizing modern force fields are typically equivalent. The effective length of the simulations is significantly increased since using a more basic description of the system allows for a much greater time step. Of course, doing so could compromise the ensemble's simulation's accuracy. MD simulation performance has been greatly enhanced by algorithmic developments, such as energy calculation, optimization, parallelization, or the use of graphical processing units (GPUs). Deep neural network (DNN) designs, which permit the example-driven formulation of arbitrarily complicated functions and their derivatives, have made machine learning (ML) potential particularly alluring. DNNs provide a very promising way to include quick yet accurate potential energy functions into MD simulations as a result [93]. On the other hand, structural models based on experimental structural biology data are frequently developed or improved using MD simulations.

The Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) approaches, which use MD simulation, but depend on continuum solvent models instead of an explicit depiction of water, are significantly quicker [94].

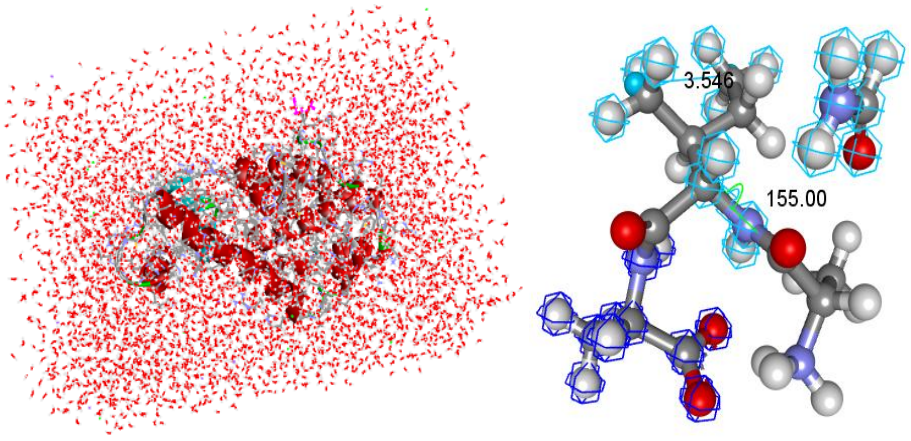

Figure 11: Molecular dynamic simulation for drug-target molecule.

The conformational changes or ligand binding is currently capable of being accurately modelled when routine simulations are reaching the microsecond scale. We can transition from the analysis of single structures, the foundation of molecular modelling as we know it, to the analysis of conformational ensembles thanks to advancements made in computational hardware, particularly the use of GPUs, and optimization of MD algorithms, including coarse-grained ones. Since they consider flexibility and dynamic features (including all thermodynamic information) and make it easier to match the results of experiments, conformational ensembles are a far better depiction of genuine macromolecules. One of the earliest issues in structural bioinformatics to be tackled is structure prediction. Although, MD, including the longest simulations run, has been widely employed for ab initio protein structure prediction, to simulate protein folding from scratch.

6.2 Preclinical Phase

As part of the drug development process, pre-clinical research examines a medicine's efficacy and safety in animal species with an observation towards potential human outcomes. General pharmacology and toxicology are the two methods which pre-clinical trials can be carried out. Pharmacology focuses on the pharmacodynamic and pharmacokinetic characteristics of drugs. It is crucial to investigate adverse pharmacological effects in appropriate animal models and to keep track of them in toxicological research. To determine the safety and effectiveness parameters in terms of absorption, distribution, metabolism, and excretion, pharmacokinetic studies are crucial. These studies provide data on absorption rates for various routes of administration, which aids in dosage form choice, distribution, rate of metabolism, and excretion, which determines the drug's half-life. The drug's half-life clarifies its safety profile, which is a requirement for a drug to be approved by regulatory bodies.

The drug's bioavailability and affinity are determined by the drug distribution mechanism, which explains the therapeutic effectiveness of the drug. Drug metabolism offers the possibility of passing through the stages of the biotransformation process and producing drug metabolites. It aids in understanding the processes and enzymes necessary for biotransformation. In-vitro and in-vivo tests can be used to conduct toxicological studies on the substance, which assess its toxicological effects. The direct impacts on cell proliferation and phenotype can be examined using in-vitro investigations. The determination of toxicological effects can be done in-vivo in both qualitative and quantitative ways. Computer-aided drug designing is a dynamically evolving area, and new approaches and methodologies are always being developed. Insilico technology has been widely used in recent years to assess the pertinent properties of drugs in the preclinical stage. This, combined with the quick advancement of computer science, has led to the creation of numerous software programs and in silico models, further encouraging the study of ADMET in vitro [95]. These methods can be applied in the preclinical stages to monitor ADMET, mechanism of action, optimal dosage and mode of administration, adverse effects, interaction with other medications, and effectiveness when compared to similar pharmaceuticals.

6.3 Clinical Phase

After preclinical research is finished, scientists go on to clinical drug development, which involves human volunteer studies and clinical trials to perfect the medication for use in patients. To ensure that the drug is as effective as feasible for its intended use, trials must be both safe and effective, conducted within the budget allotted for drug development, and follow a certain protocol. In this meticulous process to be successful, it needs to be well set up and have a large volunteer base. A clinical trial is frequently intended to determine whether a new treatment is more efficient or has less negative side effects than current treatments. Drugs go through several stages to evaluate their effectiveness, safety, and potential side effects. Four clinical stages—phase 1, phase 2, and phase 3, phase 4—are typically required for FDA drug approval. A novel drug or device is tested on a small number of individuals (between 20 and 80) in a Phase 1 study to determine its safety, including any side effects, and to determine the dosage (dosage). To determine whether a medicine is successful, more patients (between 100 and 300) are involved in a Phase 2 trial. This stage aims to gather preliminary information regarding the effectiveness of the drug or technology in patients with a particular ailment or condition. Safety, including short-term adverse effects, is still being investigated in these trials. A Phase 3 trial collects further data from a few hundred to a few thousand individuals about safety and efficacy, evaluating various populations and dosages, and contrasting the intervention with additional medications or therapy modalities. The FDA will approve the experimental medication or technology if it determines that the trial results support its use in treating a specific medical condition. The FDA must approve the medication or device before a Phase 4 trial may begin. With vast, diverse populations, the treatment's efficiency and safety are tracked. Sometimes, it takes using a drug or technology for a longer period before adverse effects become obvious to more people [96]. The process moves onto what is known as Phase 4 Clinical Trial/Post-Market Surveillance/Report Adverse Events after receiving FDA approval and being manufactured extensively by the sponsor. The FDA keeps watch over public safety and potentially dangerous side effects for at least the entire time treatment is on the market.

6.4 Regulatory Approval

When the FDA finds that a drug has been demonstrated to be both safe and effective for the purpose for which it was developed, it is then required to collaborate with the applicant to produce and improve prescribing instructions. "Labeling" is the term used for this. The optimal way to utilize the drug is explained on the label objectively and factually. But frequently, problems need to be fixed before the medicine is authorized for commercialization. FDA will occasionally ask the developer to respond to inquiries based on available information. In other situations, the FDA demands more research. This procedure keeps going until the developer chooses to stop clinical studies or submit a marketing application. A developer must have enough data from two significant, controlled clinical trials before submitting a marketing application. As a result, the actual assessment of a drug's safety must virtually span the months and even years that make up its shelf life. FDA evaluates reports of problems with prescription and over-the-counter medications, and it may decide to add warnings to dosing instructions or other practical information, as well as other events for more severe adverse drug reactions [97].

6.4.1 IV Case Study: Marine secondary metabolite activity against Methicillin-resistant Staphylococcus aureus (MRSA)

A past decade has reported various antibiotic-resistant bacteria, Staphylococcus aureus is a common bacteria that causes a wide range of clinical infections also known as a major human pathogen. It was reported that S.aureus was the leading bacterial cause of death in 135 countries associated with the highest number of deaths among persons over the age of 15 in the world. There are many drug targets associated with treating MRSA [98]. A pyruvate kinase enzyme is a hub protein in MRSA, conserved evolutionarily, and required for the growth of organisms. Significantly pyruvate kinase in bacteria regulates carbohydrate metabolism, inhibition of this enzyme impedes the catalyzing process and biosynthesis of pyruvate and Adenosine triphosphate (ATP) from phosphoenolpyruvate [99]. Thus, the reduction of ATP levels in the bacteria affects the metabolic growth of the microbe and leads to death so, it was chosen as a crucial drug target. On the other hand, marine secondary metabolites are known for treating various diseases. One of the metabolites Fijimycins A is a class of antibacterial cyclic depsipeptides derived from Streptomyces sp. The concept of structure-based drug design was implemented and docked using the molecular dynamics docking algorithm CDOCKER via the BIOVIA Discovery studio (**Figure 12**).

The protocol was validated, parameter optimised and docked with simulated annealing. The outcome of the results shows both hydrogen and hydrophobic interaction(Fig) with CDOCKER energy and interaction energy of - 29.4392 kcal/mol and -34.7287 kcal/mol respectively.

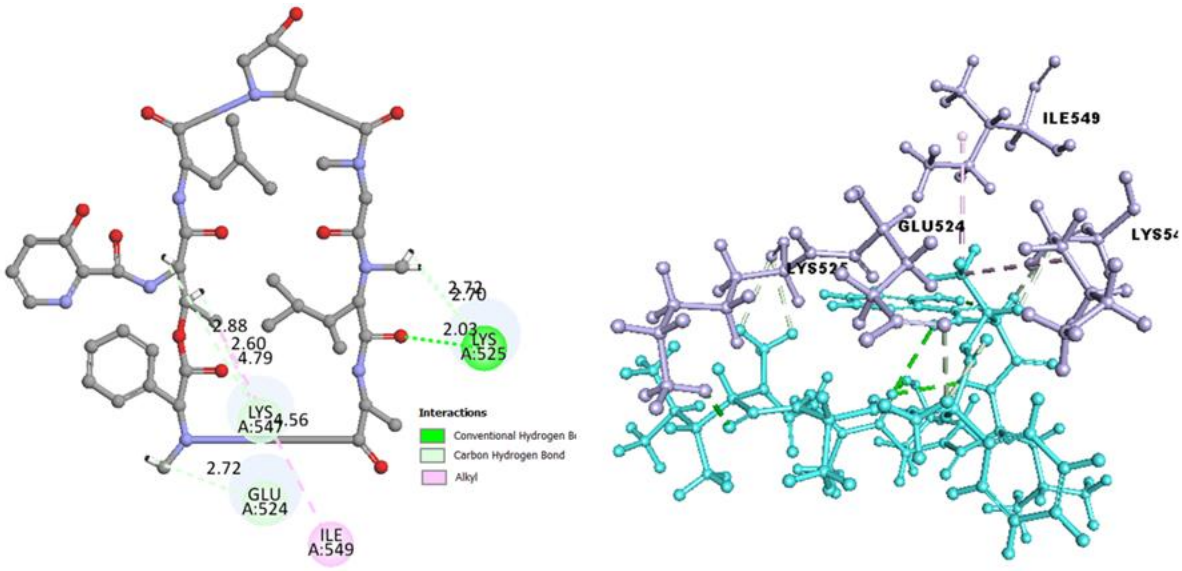

Figure 12: 2D and 3D interaction of Fijimycins A interaction with pyruvate kinase of MRSA.

7. Conclusion

We can conclude that in molecular pharmacology, molecular techniques have produced notable improvements in target identification, drug discovery, and drug development processes. Drug discovery has already been significantly impacted by the molecular revolution in biology and medicine, and development is likely to see a similar impact as well. The advent of methodologies like genomics, proteomics, and metabolomics has allowed researchers to better comprehend the intricate molecular processes underlying disease and medication responses. The capability to pinpoint precise targets for drug development is one of the key benefits of molecular methods. As a result, it is possible to develop medications that specifically target disease-causing compounds while bypassing healthy cells, minimizing adverse effects, and enhancing efficacy. Moreover, new medicine classes, including biologics and gene treatments have been developed because of molecular techniques. The drug development process has also been transformed by molecular techniques, becoming quicker and more effective. While computer-aided drug design and virtual screening approaches enable the identification of interesting therapeutic candidates without the need for costly and time-consuming laboratory studies, high-throughput screening technologies allow researchers to test a huge number of molecules fast. The link between protein sequence, structure, and function is widely known and shows that atomic-level structural information aids in understanding the molecular function of proteins.

The structural characterization of biomacromolecules, computer sciences, and molecular biology, among other fields, have made significant technological advancements possible, enabling rational drug design and presenting a comprehensive strategy. With the aid of computational methods and algorithms, it is now relatively simple to investigate the characteristics and determine the crucial component of drug molecules, i.e., their pharmacokinetic and pharmacodynamic profile. In silico techniques can be used to analyse the affinity between the drug and target molecules to predict the chemical interactions between the target and drug molecules. Not just the interactions, but also the atomic level simulations of the drug and receptor combination can be easily predicted via using these computational approaches. Furthermore, improvements in drug delivery technology have enabled the development of novel drug formulations that can boost drug effectiveness, lessen adverse effects, and boost patient compliance. For instance, sustained-release formulations can prolong medication activity and minimize dose, and frequency, while nanoparticles and liposomes can be employed to target certain cells or tissues. Moreover, improvements in computational modelling and high-throughput screening have sped up the drug development process and made it possible to identify new drug candidates more quickly. By evaluating sizable data sets and forecasting drug-target interactions, machine learning and artificial intelligence have also become more and more crucial in the discovery and development of new

medications. Overall, molecular methods have been crucial in expanding the area of molecular pharmacology and have resulted in the creation of numerous innovative and potent medications. The development of new drugs is anticipated to continue to be significantly impacted by these strategies as they progress.

Acknowledgement

The authors are thankful to the management of Altem Technologies for their valuable support.

Funding

It was declared to be none.

Conflict of interest

It was declared to be none.

References

1. Schenone M, Dančík V, Wagner BK, Clemons PA. Target identification and mechanism of action in chemical biology and drug discovery. Nat Chem Biol, 2013;9:232–4.
2. Weigelt J. The case for open-access chemical biology. EMBO Rep, 2009;10:941–5.
3. Lesko LJ, Woodcock J. Translation of pharmacogenomics and pharmacogenetics: a regulatory perspective. Nat Rev Drug Discov, 2004;3:763–9.
4. Messeguer A, Cortes N. Combinatorial chemistry in cancer research. Clin Transl Oncol, 2007;9:83–92.
5. Overington JP, Al-Lazikani B, Hopkins AL. How many drug targets are there? Nat Rev Drug Discov, 2006;5(12):993–6.
6. Horien C, Yuan P. Drug development. Yale J Biol Med, 2017;90(1):1–3.
7. Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J. Clinical development success rates for investigational drugs. Nat Biotechnol, 2014;32:40–51.
8. Dowden H, Munro J. Trends in clinical success rates and therapeutic focus. Nat Rev Drug Discov, 2019;18:495–6.
9. Paul SM, Mytelka DS, Dunwiddie CT, Persinger CC, Munos BH, Lindborg SR, et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nat Rev Drug Discov, 2010;9:203–14.
10. Athar M, Das A. Current trends in drug discovery: Target identification to clinical development of the drug. Int Res J Pharm, 2012;3:23–7.
11. Emmerich CH, Gamboa LM, Hofmann MCJ, De Souza DP, Frejno M, Vizcaíno JA, et al. Improving target assessment in biomedical research: the GOT-IT recommendations. Nat Rev Drug Discov, 2021;20:64–81.
12. Mohs RC, Greig NH. Drug discovery and development: Role of basic biological research. Alzheimers Dement (NY), 2017;3(4):651–7.
13. Ashenden SK, Kurbatova N, Bartosik A. Target identification and validation. In: Ashenden SK, editor. The Era of Artificial Intelligence, Machine Learning, and Data Science in the Pharmaceutical Industry. Academic Press; 2021. p. 61–80.
14. Kuhn M, Ian P, Stan A. Nonclinical statistics for pharmaceutical and biotechnology industries. Springer Nature; 2016.
15. Hughes J, Rees S, Kalindjian S, Philpott K. Principles of early drug discovery. Br J Pharmacol, 2011;162(6):1239–49.
16. Gashaw I, Ellinghaus P, Sommer A, Asadullah K. What makes a good drug target? Drug Discov Today, 2012;17(Suppl 24):S24–30.
17. Bakheet TM, Doig AJ. Properties and identification of human protein drug targets. Bioinformatics, 2009;25(4):451–7.
18. Kim B, Jo J, Han J, Lee H, Lee S. In silico re-identification of properties of drug target proteins. BMC Bioinformatics, 2017;18(Suppl 7):248.
19. Horien C, Yuan P. Drug development. Yale J Biol Med, 2017;90(1):1–3.
20. Santos R, Ursu O, Gaulton A, Bento AP, Donadi RS, Bologa CG, et al. A comprehensive map of molecular drug targets. Nat Rev Drug Discov, 2017;16(1):19–34.
21. Shangguan Z. J. Phys.: Conf Ser, 2021;1893:012013.

22. Kim E, Choi AS, Nam H. Drug repositioning of herbal compounds via a machine-learning approach. BMC Bioinformatics, 2019;20(10):247.
23. Diamantopoulos MA, Tsiakanikas P, Scorilas A. Non-coding RNAs: the riddle of the transcriptome and their perspectives in cancer. Ann Transl Med, 2018;6(12):241.
24. Wagner BK. The resurgence of phenotypic screening in drug discovery and development. Expert Opin Drug Discov, 2016;11:121–5.
25. Gfeller D, Grosdidier A, Wirth M, Daina A, Michielin O, Zoete V. SwissTargetPrediction: a web server for target prediction of bioactive small molecules. Nucleic Acids Res, 2014;42(W1):W32–8.
26. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Res, 2000;28(1):235–42.
27. Coudert E, Gehant S, de Castro E, Pozzato M, Baratin D, Neto T, et al. Annotation of biologically relevant ligands in UniProtKB using ChEBI. Bioinformatics, 2023;39:btac793.
28. Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, et al. RefSeq: expanding the Prokaryotic Genome Annotation Pipeline reach with protein family model curation. Nucleic Acids Res, 2021;49(D1):D1020–8.
29. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res, 2003;31:3784–8.
30. Bansal P, Morgat A, Axelsen KB, Muthukrishnan V, Coudert E, Aimo L, et al. Rhea, the reaction knowledgebase in 2022. Nucleic Acids Res, 2022;50(D1):D693–700.
31. McDonald AG, Boyce S, Tipton KF. ExplorEnz: the primary source of the IUBMB enzyme list. Nucleic Acids Res, 2009;37:D593–7.
32. Chang A, Jeske L, Ulbrich S, Hofmann J, Koblitz J, Schomburg I, et al. BRENDA, the ELIXIR core data resource in 2021: new developments and updates. Nucleic Acids Res, 2021;49(D1):D498–508.
33. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Res, 2023;51:D587–92.
34. Caspi R, Altman T, Billington R, Dreher K, Foerster H, Fulcher CA, et al. The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. Nucleic Acids Res, 2014;42(1):D459–71.
35. Mendez D, Gaulton A, Bento AP, Chambers J, De Veij M, Félix E, et al. ChEMBL: towards direct deposition of bioassay data. Nucleic Acids Res, 2019;47(D1):D930–40.
36. Kelleher KJ, Sheils TK, Mathias SL, Yang JJ, Metzger VT, Siramshetty VB, et al. Pharos 2023: an integrated resource for the understudied human proteome. Nucleic Acids Res, 2023;51(D1):D1405–16.
37. Liu T, Lin Y, Wen X, Jorissen RN, Gilson MK. BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. Nucleic Acids Res, 2007;35(Database issue):D198–201.
38. Matthews H, Hanison J, Nirmalan N. "Omics"-Informed Drug and Biomarker Discovery: Opportunities, Challenges and Future Perspectives. Proteomes, 2016;4(3):28.
39. Russell C, Rahman A, Mohammed AR. Application of genomics, proteomics and metabolomics in drug discovery, development and clinic. Ther Deliv, 2013;4(3):395–413.
40. Burbaum J, Tobal GM. Proteomics in drug discovery. Curr Opin Chem Biol, 2002;6:427–33.
41. Sun D, Gao W, Hu H, Zhou S. Why 90% of clinical drug development fails and how to improve it? Acta Pharm Sin B, 2022;12(7):3049-62.
42. Dersch P, Khan MA, Mühlen S, Görke B. Roles of regulatory RNAs for antibiotic resistance in bacteria and their potential value as novel drug targets. Front Microbiol, 2017;8:803.
43. Fogel DB. Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: A review. Contemp Clin Trials Commun, 2018;11:156-64.
44. Apic G, Ignjatovic T, Boyer S, Russell RB. Illuminating drug discovery with biological pathways. FEBS Lett, 2005;579:1872-7.
45. Qin D. Next generation sequencing and its clinical application. Cancer Biol Med, 2019;16(1):4-10.
46. Buyanbileg SI, Young EM, Hoon M, Ho LJ, Min KS. Next generation sequencing-based salivary biomarkers in oral squamous cell carcinoma. JKAOM, 2022;48(1):3-12.
47. Ingolia NT, Lareau LF, Weissman JS. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell, 2011;147:789-80.
48. Stern-Ginossar N. Decoding viral infection by ribosome profiling. J Virol, 2015;89:6164-6.
49. Kumar S, Banks TW, Cloutier S. SNP discovery through next-generation sequencing and its applications. Int J Plant Genomics, 2012;2012:831460.

50. Chorley BN, Wang X, Campbell MR, Pittman GS, Noureddine MA, Bell DA. Discovery and verification of functional single nucleotide polymorphisms in regulatory genomic regions: current and developing technologies. Mutat Res, 2008;659(1-2):147-57.
51. Wang Y, Zhao Y, Bollas A, et al. Nanopore sequencing technology, bioinformatics, and applications. Nat Biotechnol, 2021;39:1348-65.
52. Radom M, Formanowicz P. An algorithm for sequencing by hybridization based on an alternating DNA chip. Interdiscip Sci, 2018;10(3):605-15.
53. Peng W, Pronker MF, Snijder J. Mass spectrometry-based de novo sequencing of monoclonal antibodies using multiple proteases and a dual fragmentation scheme. J Proteome Res, 2021;20(7):3559-66.
54. Hardiman G. Application of ultra-high throughput sequencing and microarray technologies in pharmacogenomics testing. In: Dasgupta A, editor. Therapeutic Drug Monitoring: Newer Drugs and Biomarkers. London: Elsevier Science Publishing Co Inc.; 2012. p. 144-59.
55. Al-Amrani S, Al-Jabri Z, Al-Zaabi A, Alshekaili J, Al-Khabori M. Proteomics: concepts and applications in human medicine. World J Biol Chem, 2021;12(5):57-69.
56. Sidoli S, Cheng L, Jensen ON. Proteomics in chromatin biology and epigenetics: elucidation of post-translational modifications of histone proteins by mass spectrometry. J Proteomics, 2012;75:3419-33.
57. Ohtsuki S, Schaefer O, Kawakami H, et al. Simultaneous absolute protein quantification of transporters, cytochromes P450, and UDP-glucuronosyltransferases as a novel approach for the characterization of individual human liver: comparison with mRNA levels and activities. Drug Metab Dispos, 2012;40(1):83-92.
58. Aslam B, Basit M, Nisar MA, Khurshid M, Rasool MH. Proteomics: technologies and their applications. J Chromatogr Sci, 2017;55(2):182-96.
59. Sutandy FXR, Qian J, Chen S, Zhu H. Overview of protein microarrays. Curr Protoc Protein Sci, 2013;72(1):27.1.1-27.1.16.
60. Rosenberg JM, Utz PJ. Protein microarrays: a new tool for the study of autoantibodies in immunodeficiency. Front Immunol, 2015;6:138.
61. Kielkopf CL, Bauer W, Urbatsch IL. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis of proteins. Cold Spring Harb Protoc, 2021;2021(12).
62. Smyth MS, Martin JH. X-ray crystallography. Mol Pathol, 2000;53(1):8-14.
63. Yates J 3rd. A century of mass spectrometry: from atoms to proteomes. Nat Methods, 2011;8:633-7.
64. Hu Y, Cheng K, He L, Zhang X, Jiang B, Jiang L, et al. NMR-based methods for protein analysis. Anal Chem, 2021;93(4):1866-79.
65. Shangguan Z. A review of target identification strategies for drug discovery: from database to machine-based methods. J Phys Conf Ser, 2021;1893:012013.
66. Ashrafian H, Sounderajah V, Glen R, Ebbels T, Blaise BJ, Kalra D, et al. Metabolomics: the stethoscope for the twenty-first century. Med Princ Pract, 2021;30:301-10.
67. D'Alessandro A, Zolla L. Metabolomics and cancer drug discovery: let the cells do the talking. Drug Discov Today, 2012;17(1-2):3-9.
68. Bonomo R, Cavaletti G, Skene DJ. Metabolomics markers in neurology: current knowledge and future perspectives for therapeutic targeting. Expert Rev Neurother, 2020;20(7):725-38.
69. Russell C, Rahman A, Mohammed AR. Application of genomics, proteomics, and metabolomics in drug discovery, development, and clinic. Ther Deliv, 2013;4(3):395-413.
70. DrugBank [Internet]. Available from: http://www.drugbank.ca
71. PubChem [Internet]. Available from: http://pubchem.ncbi.nlm.nih.gov
72. ZINC [Internet]. Available from: http://zinc.docking.org
73. Block P, Sotriffer CA, Dramburg I, Klebe G. AffinDB: a freely accessible database of affinities for protein-ligand complexes from the PDB. Nucleic Acids Res, 2006;34(Database issue):D522-6.
74. Li H, Leung KS, Nakane T, Wong MH. iview: an interactive WebGL visualizer for protein-ligand complex. BMC Bioinformatics, 2014;15:56.
75. Wang Z, Pan H, Sun H, Kang Y, Liu H, Cao D, et al. fastDRH: a webserver to predict and analyze protein-ligand complexes based on molecular docking and MM/PB(GB)SA computation. Brief Bioinform, 2022;23(5):bbac201.
76. Shin JM, Cho DH. PDB-Ligand: a ligand database based on PDB for the automated and customized classification of ligand-binding structures. Nucleic Acids Res, 2005;33:D238-41.
77. Puvanendrampillai D, Mitchell JB. L/D Protein Ligand Database (PLD): additional understanding of the nature and specificity of protein-ligand complexes. Bioinformatics, 2003;19:1856-7.

78. Laskowski RA. PDBsum: summaries and analyses of PDB structures. Nucleic Acids Res, 2001;29:221-2.
79. Kleywegt G, Jones T. Databases in protein crystallography. Acta Crystallogr D Biol Crystallogr, 1998;54:1119-31.
80. Hendlich M, Bergner A, Günther J, Klebe G. Relibase: design and development of a database for comprehensive analysis of protein-ligand interactions. J Mol Biol, 2003;326:607-20.
81. Smith RD, Hu L, Falkner JA, Benson ML, Nerothin JP, Carlson HA. Exploring protein-ligand recognition with Binding MOAD. J Mol Graph Model, 2006;24:414-25.
82. Astex diverse set [Internet]. Available from: http://www.ccdc.cam.ac.uk
83. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science, 2012;337:816-21.
84. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. Science, 2013;339:823-6.
85. Dominguez AA, Lim WA, Qi LS. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. Nat Rev Mol Cell Biol, 2016;17:11.
86. Han K, Jeng EE, Hess GT, Morgens DW, Li A, Bassik MC. Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. Nat Biotechnol, 2017;35:463-74.
87. Zhou SF, Zhong WZ. Drug design and discovery: principles and applications. Molecules, 2017;22(2):279.
88. Luu KT, Kraynov E, Kuang B, Vicini P, Zhong WZ. Modeling, simulation, and translation framework for the preclinical development of monoclonal antibodies. AAPS J, 2013;15:551-8.
89. Deore A, Dhumane J, Wagh R, Sonawane R. The stages of drug discovery and development process. Asian J Pharm Res Dev, 2019;7:62-7.
90. McNutt AT, Francoeur P, Aggarwal R, et al. GNINA 1.0: molecular docking with deep learning. J Cheminform, 2021;13:43.
91. Stevens JA, Grünewald F, van Tilburg PAM, König M, Gilbert BR, Brier TA, et al. Molecular dynamics simulation of an entire cell. Front Chem, 2023;11.
92. Zhang J, Zhang H, Wu T, Wang Q, van der Spoel D. Comparison of implicit and explicit solvent models for the calculation of solvation free energy in organic solvents. J Chem Theory Comput, 2017;13(3):1034-43.
93. Doerr S, Majewski M, Pérez A, Krämer A, Clementi C, Noe F, et al. TorchMD: a deep learning framework for molecular simulations. J Chem Theory Comput, 2021;17(4):2355-63.
94. Poli G, Granchi C, Rizzolio F, Tuccinardi T. Application of MM-PBSA methods in virtual screening. Molecules, 2020;25:1971.
95. Wu F, Zhou Y, Li L, Shen X, Chen G, Wang X, et al. Computational approaches in preclinical studies on drug discovery and development. Front Chem, 2020;8.
96. What are clinical trials and studies? [Internet]. National Institute on Aging. Available from: https://www.nia.nih.gov/health/what-are-clinical-trials-and-studies
97. Adams CP, Brantner VV. New drug development: estimating entry from human clinical trials. Bureau of Economics, Federal Trade Commission. 2003.
98. The Lancet article on clinical trials [Internet]. Available from: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(22)02185-7/fulltext
99. Article from Chemistry Select [Internet]. Available from: https://doi.org/10.1002/slct.202201403.

How to cite this article: Prabhakaran P, Bhardwaj S, Shanmugarajan D, Chandra S. Molecular Approaches and Significant Advances to Target Identification, Drug Discovery and Drug Development. Pharma Research Bulletin, 2024;3(1):11-29.