Drug Discovery and Development, a New Hope

Azra Khan
Faculty of Health Science, SHIATS, Allahabad, UP., India

Neda Firoz
Deptt of Computational Biology & Bio Informatics, SHIATS, Allahabad, UP., India

Abstract—This Paper gives the review of Drug Discovery and Development in the Computational Biological era of medical science, advancements, and gives a new hope of better public health in the near future. It emphasizes on how we can develop better and competitive drugs with the use of software and wet lab synchronization and the hope of developing better tools to facilitate human life the comfort and disease competitive.

Keywords—Drug, discovery and history, sources, lead, ADME

I. INTRODUCTION

WHAT IS A DRUG?

Any chemical compound - sugar ???

Anything which produces a change in the body - an axe

So to understand Drug, we can define by characteristics:
1. use or potential use in diagnosis or treatment of disease
2. selective in their actions

A drug is a substance which has a physiological effect when ingested or otherwise introduced into the body or a Substance that is used, "primarily to bring about a change in some existing process or state, be it psychological, physiological or biochemical.

CHANGED CONTEXT OF DRUG DISCOVERY AND DEVELOPMENT

Drug discovery is the process by which new candidate medications are discovered.

Historically, drugs were discovered through identifying the active ingredient from traditional remedies or by serendipitous discovery. Later chemical libraries of synthetic small molecules, natural products or extracts were screened in intact cells or whole organisms to identify substances that have a desirable therapeutic effect in a process known as classical pharmacology. Since sequencing of the human genome which allowed rapid cloning and synthesis of large quantities of purified proteins, it has become common practice to use high throughput screening of large compounds libraries against isolated biological targets which are hypothesized to be disease modifying in a process known as reverse pharmacology. Hits from these screens are then tested in cells and then in animals for efficacy.

Modern drug discovery involves the identification of screening hits, medicinal chemistry and optimization of those hits to increase the affinity, selectivity (to reduce the potential of side effects), efficacy/potency, metabolic stability (to increase the half-life), and oral bioavailability. Once a compound that fulfills all of these requirements has been identified, it will begin the process of drug development prior to clinical trials. One or more of these steps may, but not necessarily, involve computer-aided drug design. Modern drug discovery is thus usually a capital-intensive process that involves large investments by pharmaceutical industry corporations as well as national governments (who provide grants and loan guarantees). Despite advances in technology and understanding of biological systems, drug discovery is still a lengthy, "expensive, difficult, and inefficient process" with low rate of new therapeutic discovery. In 2010, the research and development cost of each new molecular entity (NME) was approximately US$1.8 billion. Drug discovery is done by pharmaceutical companies, with research assistance from universities. The "final product" of drug discovery is a patent on the potential drug. The drug requires very expensive Phase I, II and III clinical trials, and most of them fail. Small companies have a critical role, often then selling the rights to larger companies that have the resources to run the clinical trials.

Discovering drugs that may be a commercial success, or a public health success, involves a complex interaction between investors, industry, academia, patent laws, regulatory exclusivity, marketing and the need to balance secrecy with communication. Meanwhile, for disorders whose rarity means that no large commercial success or public health effect can be expected, the orphan drug funding process ensures that people who experience those disorders can have some hope of pharmacotherapeutic advances.

The idea that the effect of a drug in the human body is mediated by specific interactions of the drug molecule with biological macromolecules, (proteins or nucleic acids in most cases) led scientists to the conclusion that individual chemicals are required for the biological activity of the drug. This made for the beginning of the modern era in
pharmacology, as pure chemicals, instead of crude extracts, became the standard drugs. Examples of drug compounds isolated from crude preparations are morphine, the active agent in opium, and digoxin, a heart stimulant originating from Digitalis lanata. Organic chemistry also led to the synthesis of many of the natural products isolated from biological sources.

II. HISTORY OF DRUG DEVELOPMENT

Historically substances, whether crude extracts or purified chemicals were screened for biological activity without knowledge of the biological target. Only after an active substance was identified was an effort made to identify the target. This approach is known as classical pharmacology, forward pharmacology, or phenotypic drug discovery. Later, small molecules were synthesized to specifically target a known physiological/pathological pathway, rather than adopt the mass screening of banks of stored compounds. This led to great success, such as the work of Gertrude Elion and George H. Hitchings on purine metabolism, the work of James Black on beta blockers and cimetidine, and the discovery of statins by Akira Endo. Another champion of the approach of developing chemical analogues of known active substances was Sir David Jack at Allen and Hanbury's, later Glaxo, who pioneered the first inhaled selective beta2-adrenergic agonist for asthma, the first inhaled steroid for asthma, ranitidine as a successor to cimetidine, and supported the development of the triptans.

Gertrude Elion, working mostly with a group of fewer than 50 people on purine analogues, contributed to the discovery of the first anti-viral; the first immunosuppressant (azathioprine) that allowed human organ transplantation; the first drug to induce remission of childhood leukaemia; pivotal anti-cancer treatments; an anti-malarial; an anti-bacterial; and a treatment for gout.

Cloning of human proteins made possible the screening of large libraries of compounds against specific targets thought to be linked to specific diseases. This approach is known as reverse pharmacology and is the most frequently used approach today.

Fig 1. History of Drug Development

SOURCES OF DRUGS

The records of using drug is found dating back to 2,700 B.C. in the Middle East and China! The drugs most commonly used were laxatives and anti-emetics. To relieve pain, Opium extract was used. Ephedrine was used for the treatment of respiratory tract disorders.

Until the beginning of twentieth century, the substances used for the treatment of diseases were obtained from natural sources. Natural sources include plants, animals, and minerals. Among the natural sources, plants were mainly used. Sometimes minerals and occasionally animals were used for the same purpose. Nowadays most of the drugs are manufactured in the laboratory, i.e. synthetic drugs. Microorganisms also serve as a source of a large number of drugs.

So, what are the sources of drugs?

The sources of drugs can be grouped according to the following:

1. Plant source

Plants As Sources of Drugs

Believe it or not, there used to be a time when the leaves that had the shape of the liver, were used for the treatment of liver diseases! Subsequently various parts of the plant such as root, bark, stem, leaf, seed and flower were used.
All parts of a specific plant do not equally contain a specific drug. For example, atropine, caffeine, cocaine, digoxin, and pilocarpine are obtained from the leaves of specific plants. Seeds of some plants are used to extract castor oil, colchicine, morphine, strychnine and theobromine. Barks of some plants are used for the extraction of drugs like cinnamon, quinidine, and quinine. Roots of some plants are used to extract reserpine and atropine. Nowadays plants that were used as drugs, with some expectation (digitalis, belladonna), are no longer considered for rational treatment. Rather the pharmacologically active constituents (e.g. atropine from the roots) are extracted and used.

The pharmacologically active constituents of different plants are grouped according to their physio-chemical properties and include:

A. Alkaloid
B. Glycoside
C. Oil
D. Gum
E. Mucilage
F. Carbohydrate and related compounds

Some of these active constituents can be extracted by soaking the plant in alcohol.

1) The purpose of extracting the active constituents are:

1. Identification of the active constituents.
2. Analysis of the pharmacodynamic and pharmacokinetic properties of the active constituents
3. Ensuring a precise and constant dosage in the therapeutic use of chemically pure constituents
4. The possibility of chemical synthesis

2. Animal source

Animals As Sources of Drugs

There was a time when the Chinese people used the dried skin of toad to treat toothache and bleeding in gum. Later it was found that toad skin contains adrenaline.

The liver of cod fish (cod liver oil) contains high levels of omega 3 fatty acids, vitamin A and vitamin D.

Insulin is extracted from the pancreas of bovine or porcine.

Immunoglobulin G is prepared by the injecting antigen into an animal and collecting the antibody formed as a reaction to the antigen. Immunoglobulin of animal origin (antisera) is frequently associated with hypersensitivity reactions which has led to its virtual abandonment. For example, horse globulin containing anti-tetanus and anti-diphtheria toxin has been extensively used at one time, but nowadays their use is more restricted as they give rise to complications like serum sickness.

So antisera is replaced by human immunoglobulin. Human immunoglobulin is prepared from pools of at least 1000 donations of human plasma containing the antibody to measles, mumps, hepatitis A and other viruses. Injection of human immunoglobulin produces immediate passive immunity lasting for about 4 to 6 weeks. Specific immunoglobulin (hepatitis B immunoglobulin, rabies immunoglobulin, tetanus immunoglobulin) are prepared by pooling the plasma of selected donors with high levels of the specific antibody required.

Human menopausal gonadotropins (hMG) is isolated from the urine of postmenopausal women and contains a mixture of follicle stimulating hormone (FSH) and luteinizing hormone (LH).

Human chorionic gonadotropin (hCG) is produced by the placenta and can be isolated and purified from the urine of pregnant woman. The hCG is nearly identical in activity to LH but it differs in sequence and carbohydrate content.

Heparin is commonly extracted from porcine intestinal mucosa or bovine lung.

3. Mineral source

Minerals As Sources of Drugs

The sword symbolizes strength and power, the early Greek physicians attempt to use iron therapy against weakness and anemia.

Various clay have been used for the treatment of diarrhea. One remedy, called for the powdering of the bowls with old clay pipes. The principal ingredients of such pipes would be kaolin and activated charcoal, both of which are used today for the treatment of diarrhea.

Calomel was used for the treatment of constipation. It contains mercury and subsequently found to have a diuretic effect and was used with digitals for the treatment of congestive cardiac failure. The diuretic effect of mercury was also observed following the used of that compound in the treatment of syphilis.

Iodine is used for the treatment of goiter.

Gold is used for the treatment for the arthritis.

Sulfur is used externally in skin diseases.

Aluminum hydroxide and magnesium trisilicate are widely used as antacids.

Magnesium sulfide is used to relieve constipation and to control eclamptic seizure.

4. Laboratory source

Laboratory As Sources of Drugs
Nowadays most drugs are produced artificially by combining two or more compounds or elements. It may be partially or totally synthesized. The structural alteration of the natural substance by the addition of a pure chemical substance leads to the production of a partially synthetic substance.

With the improvement of organic chemical industry, the synthesis of chemical substances in the laboratory has become extremely advanced. In most cases, drugs produced in laboratories are of high quality, less expensive, produced in large scale within short time, safer, and more effective than drugs extracted from plants or animals.

For example, 1 mg of digoxin produced in the laboratory has the same pharmacological effect as produced from 1000 mg of crude leaves of purple foxgloves. That is 1 mg synthetic digoxin is equivalent to 1000 mg of crude leaves of purple foxgloves.

Salicylates originally extracted from the plant source are nowadays produced in the laboratory.

The synthesis of sulfonamide began from protonsil dye. One of the adverse effects of sulfonamides was hypoglycemia, which led to the development of sulfonylurea drugs. Acetazolamide (carbonic anhydrase inhibitor), hydrochlorothiazide, and frusemide are also developed from sulfanilamide. Nowadays sulfonylureas are used to lower blood sugar level in non-insulin dependent diabetes mellitus.

Human insulin is produced by modification of porcine insulin or by bacteria using recombinant DNA technology. It is known to us that insulin contains 51 amino acids in two chains, A and B. A chain contains 21 amino acids and B chain contains 30 amino acids. Bovine insulin differs from human insulin at 3 amino acid sites whereas porcine insulin at 1 amino acid site. By changing the amino acids alanine or porcine insulin at position 30 of B chain with threonine, we can convert it to human insulin. Human insulin is absorbed more rapidly from the site of administration re the bovine or porcine insulin. But the duration of effect of human insulin is shorter and doses must be adjusted.

The actual production of insulin (see the diagram below) involves the introduction of human insulin gene into a non-pathogenic strain of the bacteria Escherichia coli K12. Insulin gene is separated from the chromosome using restriction enzymes. Then bacteria containing human gene are cultured in huge vats of nutrients until they are ready to have the insulin extracted from them.

In 1948, the antibiotic 7-chlortetracycline was isolated from the Streptomyces aurefaciens. The catalytic removal of chlorine from 7-chlortetracycline gave tetracycline. Tetracycline is superior than 7-chlortetracycline and has replaced it. Studies on the structure and synthesis of penicillin led to the development of the naturally synthetic penicillin and later to cephalosporin.

Most of the currently used analgesics, chemotherapeutic drugs, hypnotics and local anesthetics are produced in the laboratory.

The natural source of caffeine is the tea or coffee. Large amount of caffeine is nowadays obtained as the byproduct of manufacturing decaffeinated coffee.

Theophylline can be produced by methylation of theobromine (partial synthesis) or from urea (total synthesis).

5. Microorganisms

Microorganisms As Sources of Drugs

Well-known antibiotics produced by the actinomycetes are actinomycin, amphotericin, chloramphenicol, erythromycin, kanamycin, neomycin, gentamicin, streptomycin and tetracycline.

Aspergillate group of fungi produce antibiotics such as penicillin, griseofulvin and cephalosporin. Among the bacteria, genus Bacillus produces antibiotics such as polymyxin B and bacitracin.

Fig 2. Timeline of Drug Discovery & Development
III. DRUG DISCOVERY PROCESS

Research and Development of pharmaceutical products take a lengthy period that spans over ten years and requires large investment. Prior to market launch, new drugs undergo a long and complicated series of steps, including an evaluation of efficacy and safety, application for approval, and investigation and approval of drug applications by regulatory authorities. Pharmaceutical companies are in the business of identifying compounds that may be useful in new drugs. It takes tens or hundreds of thousands of compounds are made and tested every year (“screening”)

- tests are usually simple binding assays (does the molecule bind to a target protein?)

The testing is done in two stages

- Lead Generation (find a compound that binds)
- Lead Optimization (find a compound that binds better)

The chemical similarity is important at both these stages.

LEAD GENERATION

In the early stage of a drug discovery process, researchers may be faced with little or no structure activity relationship (SAR) information. At this point, assay development and screening should be undertaken immediately by the high-throughput screening (HTS) group. The aim of these analyses is to select and test fewer compounds, whilst gaining as
much information as possible about the dataset. However if a lead is known, then more focused approach can be adopted by searching for compounds with similar (two or three-dimensional) structures to the lead candidate or by substructure searching. In substructure searching the query will retrieve those structures from the database that contain groups present in the primary lead. These molecules can then be screened in a biological assay.

**LEAD OPTIMIZATION**

In medicinal chemistry the lead optimization process concerns many aspects such as the optimization of the affinity for the biological target, the toxicity, the oral bioavailability, the cell permeability, the plasma binding, the ease of metabolism. The principle employed is that any incremental change in the chemical structure produces incremental (positive or negative) changes in bio-activity and a systematic study of such cause and effect relationship is called structure activity relationship (SAR) study. The process is highly iterative and traditionally based on trial-and-error. When no structural data about the target is available, the lead optimization process can be made more methodological by using quantitative structure activity relationship (QSAR) studies. QSAR methods are used to attempt to correlate two different approaches can be used in QSAR depending on the available compounds:

1. Two-dimensional QSAR (2D-QSAR)
2. Three-dimensional QSAR (3D-QSAR)

This problem limits the applicability of CoMFA. In order to overcome this problem, some new approaches, which do not depend on a common alignment of the molecules, have been recently developed. Comparative molecular moment analysis CoMMA9, EVA or the WHIM are used because they provide three dimensional descriptors that are independent of the orientation of the molecules in space; they do not have to be aligned.

**IV. DRUG TARGETS**

What is a drug target? And how many such targets are there? Here, we consider the nature of drug targets, and by classifying known drug substances on the basis of the discussed principles we provide an estimation of the total number of current drug targets.

Estimations of the total number of drug targets are presently dominated by analyses of the human genome, which are limited for various reasons, including the inability to infer the existence of splice variants or interactions between the encoded proteins from gene sequences alone, and the fact that the function of most of the DNA in the genome remains unclear. In 1997, when 100,000 protein-coding sequences were hypothesized to exist in the human genome, Drews and Ryser estimated the number of molecular targets ‘hit’ by all marketed drug substances to be only 482. In 2002, after the sequencing of the human genome, others arrived at ~8,000 targets of pharmacological interest, of which nearly 5,000 could be potentially hit by traditional drug substances, nearly 2,400 by antibodies and ~800 by protein pharmaceuticals. And on the basis of ligand-binding studies, 399 molecular targets were identified belonging to 130 protein families, and ~3,000 targets for small-molecule drugs were predicted to exist by extrapolations from the number of currently identified such targets in the human genome.

In summary, current target counts are of the order of 10^2, whereas estimations of the number of potential drug targets are an order of magnitude higher. In this paper, we consider the nature of drug targets, and use a classification based on this consideration, and a list of approved drug substance to estimate the number of known drug targets, in the following categories:

**Drug discovery and computer aided drug design**

It is estimated that a typical drug discovery cycle, from lead identification through to clinical trials, can take 14 years with cost of 800 million US dollars. In the early 1990s, rapid developments in the fields of combinatorial chemistry and high-throughput screening technologies have created an environment for expediting the discovery process by enabling huge libraries of compounds to be synthesized and screened in short periods of time. However, these concerted efforts not only failed to increase the number of successfully launched new molecular entities, but seemingly aggravated the situation. Hit rates are often low and many of these identified hits fail to be further optimized into actual leads and preclinical. Among the late stage failures, 40–60% was reportedly due to absorption, distribution, metabolism, excretion and toxicity (ADME/Tox) deficiencies. Collectively, these issues underscore the need to develop alternative strategies that can help remove unsuitable compounds before the exhaustion of significant amount of resources. The more recent foundations of CADD were established in the early 1970s with the use of structural biology to modify the biological activity of insulin and to guide the synthesis of human haemoglobin ligands.

At that time, X-ray crystallography was expensive and time-consuming, rendering it infeasible for large-scale screening in industrial laboratories. Over the years, new technologies such as comparative modeling based on natural structural homologues have emerged and began to be exploited in lead design. These, together with advances in combinatorial chemistry, high-throughput screening technologies and computational infrastructures, have rapidly bridged the gap between theoretical modeling and medicinal chemistry. Numerous successes of designed drugs were reported, including Dorzolamide for the treatment of cystoid macular edema, Zanamivir for therapeutic or prophylactic treatment of influenza infection, Sildenafil for the treatment of male erectile dysfunction, and Amprenavir for the treatment of HIV. There are two major types of drug design. The first is referred to as *structure-based drug design* and the second, *ligand-based drug design*. 
1) **Structure-based**

Structure-based drug design (or **direct drug design**) relies on knowledge of the three-dimensional structure of the biological target obtained through methods such as x-ray crystallography or NMR spectroscopy.\[^5\] If an experimental structure of a target is not available, it may be possible to create a homology model of the target based on the experimental structure of a related protein. Using the structure of the biological target, candidate drugs that are predicted to bind with high affinity and selectivity to the target may be designed using interactive graphics and the intuition of a medicinal chemist. Alternatively, various automated computational procedures may be used to suggest new drug candidates.

The 3D structures of biomolecular targets are obtained from X-ray crystallography and NMR. In parallel, information about the structural dynamics and electronic properties about ligands are obtained from calculations. This has encouraged the rapid development of the structure-based drug design. Current methods for structure-based drug design can be divided roughly into two categories. The first category is about “finding” ligands for a given receptor, which is usually referred to as database searching. In this case, a large number of potential ligand molecules are screened to find those fitting the binding pocket of the receptor. This method is usually referred to as ligand-based drug design. The key advantage of database searching is that it saves synthetic effort to obtain new lead compounds. Another category of structure-based drug design methods is about “building” ligands, which is usually referred to as receptor-based drug design. In this case, ligand molecules are built up within the constraints of the binding pocket by assembling small pieces in a stepwise manner. These pieces can be either individual atoms or molecular fragments. The key advantage of such a method is that novel structures, not contained in any database, can be suggested.

**Three methods are used in this strategy:**
- De novo Drug Design
- Virtual Screening
- Docking

2) **Ligand-based**

Ligand-based drug design (or **indirect drug design**) relies on knowledge of other molecules that bind to the biological target of interest. These other molecules may be used to derive a pharmacophore model that defines the minimum necessary structural characteristics a molecule must possess in order to bind to the target. In other words, a model of the biological target may be built based on the knowledge of what binds to it, and this model in turn may be used to design new molecular entities that interact with the target. Alternatively, a quantitative structure-activity relationship (QSAR), in which a correlation between calculated properties of molecules and their experimentally determined biological activity, may be derived. These QSAR relationships in turn may be used to predict the activity of new analogs. Two methods are used in this strategy:
- QSAR Analysis
- Pharmacophore modeling

---

V. **ORPHAN DRUGS**

An orphan drug is a pharmaceutical agent that has been developed specifically to treat a rare medical condition, the condition itself being referred to as an orphan disease. In the US and EU it is easier to gain marketing approval for an orphan drug, and there may be other financial incentives, such as extended exclusivity periods, all intended to encourage the development of drugs which might otherwise lack a sufficient profit motive. The assignment of orphan status to a disease and to any drugs developed to treat it is a matter of public policy in many countries, and has resulted in medical breakthroughs that may not have otherwise been achieved due to the economics of drug research and development. Orphan drugs generally follow the same regulatory development path as any other pharmaceutical product, in which testing focuses on pharmacokinetics and pharmacodynamics, dosing, stability, safety and efficacy. However, some statistical burdens are lessened in an effort to maintain development momentum. For example, orphan drug regulations generally acknowledge the fact that it may not be possible to test 1,000 patients in a phase III clinical trial, as fewer than that number may be afflicted with the disease in question.

Since the market for any drug with such a limited application scope would, by definition, be small and thus largely unprofitable, government intervention is often required to motivate a manufacturer to address the need for an orphan drug. Critics of free market enterprise often cite this as a failure of free market economic systems. The intervention by government on behalf of orphan drug development can take a variety of forms:
- Tax incentives.
- Enhanced patent protection and marketing rights.
- Clinical research financial subsidization.
- Creating a government-run enterprise to engage in research and development (see Crown corporation).

Currently there are more than 400 orphan designated drugs in clinical trial process. Majority of these drugs are being developed in US followed by Europe. US dominates the development of orphan drugs with more than 300 orphan designated drugs being under clinical trial process.
VI. CONCLUSION

Computer Aided Drug Design (CADD) and Delivery Systems offers an in-depth discussion of the computer-assisted techniques used to discover, design, and optimize new, effective, and safe drugs. Recent technological developments in biochemistry, biomedical science, and nanotechnology have made computer-aided drug design and delivery systems possible on a molecular basis. This in-depth treatise covers this pioneering advances.

The objective of drug design is to find a chemical compound that can fit to a specific cavity on a protein target both geometrically and chemically. It is generally recognized that drug discovery and development are very time and resources consuming processes. There is an ever growing effort to apply computational power to the combined chemical and biological space in order to streamline drug discovery, design, development and optimization. In biomedical arena, computer-aided or in silico design is being utilized to expedite and facilitate hit identification, hit-to-lead selection, optimize the absorption, distribution, metabolism, excretion and toxicity profile and avoid safety issues. The development of any potential drug begins with years of scientific study to determine the biochemistry behind a disease, for which pharmaceutical intervention is possible. The result is the determination of specific receptors (targets). In the post genomic era, computer-aided drug design (CADD) has considerably extended its range of applications, spanning almost all stages in the drug discovery pipeline, from target identification to lead discovery, from lead optimization to preclinical or clinical trials.

One method that was quickly adopted by industry was the use of combinatorial chemistry and HTS. In HTS, large libraries of compounds are screened against drug targets to identify lead compounds that can modulate a particular outcome. However, setting up a combinatorial chemistry program and HTS is costly and not able to address the specific needs of many biological (drug target) systems. Compounds identified in such screenings are not always amenable to further medicinal chemistry development, with poor ADME (absorption, distribution metabolism and elimination) properties. Although these methods have increased the rate at which lead compounds can be identified, there has not been a commensurate increase in the rate of introduction of new chemical entities (NCE) into the world drug market. As an attractive alternative, in silico methods show promise in identifying new lead compounds faster and at a fraction of the cost of combinatorial approaches and HTS. The addition of computer aided drug design technologies to the R&D approaches of a company, could lead to a reduction in the cost of drug design and development by up to 50%. THERE IS a lot more to discover, with the new integration of information science and drug development. This gives a hope of better and quicker health care facilities.

ACKNOWLEDGEMENT

We would like to thank our parents, Department faculty, for the source of inspiration in the review of the paper done. We hope to have a near advancements in this field of study to make our community better in health care and disease competitive. We are also thankful to the department library of SHIATS to provide us the full help in studying different aspects of Drug Discovery.
REFERENCES


