CHAPTER 1

DRUG DISCOVERY: A BRIEF OVERVIEW

CHAPTER 1

DRUG DISCOVERY: A BRIEF OVERVIEW

CHAPTER 1

DRUG DISCOVERY: A BRIEF OVERVIEW

1 General Introduction

The development of new drugs with potential therapeutic applications is one of the most complex and difficult processes in the pharmaceutical industry. Millions of dollars and man-hours are devoted to the discovery of new therapeutic agents. As, the activity of a drug is the result of a multitude of factors such as bioavailability, toxicity and metabolism, rational drug design has been utopias for centuries. Very recently, impressive technological advances in areas such as structural characterization of biomacromolecules, computer sciences and molecular biology have made rational drug design feasible[1,2]

From vaccines that have reduced many fatal diseases to the realm of history, to psychoactive agents that have permitted many people to live normal lives, to antihypertensive drugs that prolong and improve the quality of life, to contraceptives that have permanently changed sexual mores, the products of the pharmaceutical industry have had a major and irrevocable impact on the society. Yet many disease states still remain untreated or are currently treated by agents whose side-effect profile leaves significant cause to worry. The aging of the population, partly a result of the availability of more effective medications, has presented additional challenges in therapeutic agent development, with an urgent need for drugs to treat arthritis, **cancer,** neurodegeneration, and cognition impairment.[3,4]

However, the process of drug discovery is complex and costly, and innovative science is especially so. But drug discovery is more than the application of basic

research knowledge and technologies[3];it involves many facets of project management and focus. To precisely define the drug discovery process and the parameters crucial to success is a difficult task. For each successful drug there are many accounts, both corporate and scientific, as to the process and the visionaries and facilitators involved in bringing the compound to the marketplace [5-7].

1.1Growth of Medicinal Chemistry as Drug Discovery Process[8]

The roots of medicinal chemistry lie in many branches of chemistry and biology. In the so called prescientific era, natural products having a history as folk remedies were in use but little of drug therapy of today is based on these remedies. Some of the natural products that are used today were used originally for other purposes such as arrow poisions, part of religious or even cosmetics for eg.opium, ergot etc.Before the development of chemistry as science, the drugs used were either natural organic product or inorganic material. Herbals that had been in use for thousands of years and resulted from various societies that had been made in the search to find cures for various ills and disease encountered and recognized. Among the earliest recognized uses of medicinal plant were herb called ***ma huang, a*** species of Ephedra, used medicinally in China for over 5000 years.

Paracelsus, the Swiss physician, made 1st recorded challenge to the use of herb and urged the alchemist of his era to use their knowledge for developing chemical medicines, primarily from mineral. With the gradual decline of magic and superstition that had accompanied the use of substances believed to have medicinal properties, physician began to look for evidence of effectiveness in their preparation. English physician had also experimented with the use of **Priestley's** "fixed air" or carbon dioxide and **Lavoisiers** "dephleogisticatredair" or oxygen. Carbonated drinks

were administered with a hope to dissolve kidney stone and oxygen for resuscitation .In 1840 **Long and wells** showed that ether and nitrous oxide could be used as anaesthetics for surgery. In the late 18^{th} and early 19^{th} century chemical experiments led to its use in the discovery of new drugs.Drug design has been aided by increasing understanding of biochemical metabolism and biosynthesis, and by statistical analysis of some relationship of physical properties of chemicals and their biological performance. The modern drug design is now being done a more rational basis for which more and more information is being obtained in cell biochemistry and cell biology at the molecular level. The rational approach envisages a physiological basis of a disease. Over the past decades advances in biotechnology have guided to a new approach to drug discovery called ***Structure based drug design.*** The basic consideration for structure based drug design is the generation of reliable three dimensional atomic models. For the medicinal chemists, the concept of interference of drugs with the bioconversions of substrate in enzymic reactions has had a profound didactic effect. It has enabled chemical drug design to be based on the structure of known substrates i.e amino acids, carbohydrates, hormones etc.The study of structure activity relationship of new chemical entities provide the basis for the development of better medicinal agents from the lead compound. Now computers have pressed into the service of medicinal chemists. Computers do not interpret data but they assist the scientist in collecting, storing, manipulating, analyzing and viewing the data, and increase the efficiency of the drug discovery process. The role of the medicinal chemist is that of increasing the duration of action and potency, and decreasing the adverse effect of newly discovered compounds.

1.1.1. Relationship between Functional group and Pharmacological activity

To find the influence of chemical structure on biological activity, it is necessary for medicinal chemist to understand the physicochemical properties of molecule that is the influence of organic functional group present within the molecule on its acid /base properties, water solubility, partition coefficient, crystal structure, stereochemistry etc. All these properties influence the absorption, distribution, metabolism and excretion (ADME) of the molecule. Therefore in order to design better chemical entity the medicinal chemist should know the relative contribution that each functional group makes to the overall physical chemical properties of molecule. Such studies are known as **Structure Activity Relationship Studies..**Functional group plays an important role in binding drugs to their targets and the SAR studies can identify whether these groups are important or not.

1.1.2. Stereochemistry and Drug Action

The physico chemical properties of drug molecules are not only dependent upon nature and type of functional group present in molecule but also on the spatial arrangement of these groups. It becomes an important factor when a drug molecule is introduced to an asymmetric environment as human body. Since the biological macromolecules are asymmetric in nature, how a drug molecule interacts with these molecules is determined by three dimensional orientations of groups present in drug molecule. If functional groups do not occupy the proper spatial arrangement, then productive bonding interactions with the biological macromolecules will not be possible, thereby the required physiological effect can’t be obtained. However the

drug can produce strong interaction with the receptor molecule if the functional groups are properly oriented. Thus for developing a new molecule it is very important for a medicinal chemist to understand what three dimensional orientation of groups is needed for drug action.

1.1.3. Role of membrane in drug action[9-10]

The role of membranes in drug action was recognized by **Seydel**.The membrane acts on the drug molecules, in the drug membrane interaction by the following ways:

1. The membrane may prevent diffusion to the active site.
2. The membrane may bind or accumulate drugs.
3. The solvation of the drug in the membrane may lead to a conformational change in its structure.
4. The diffusion through the membrane may become the rate limiting step.

Vice versa the drug acts on membrane properties*:*

1. The drug may change the thickness of the membrane.
2. The drug may increase the membrane surface.
3. The drug may change the fluidity of the membrane.
4. The drug may change the membrane potential and the hydration of the polar head groups.

1.2 Drug discovery process

The discovery of a new drug involves several steps which are given as:

1.2. 1. Strategy

A uniform strategy for the drug discovery process has remained elusive despite considerable analysis. Like the science on which the drug discovery process is based, the approach to a problem involves many paths, with considerable trial and error. Strategies for drug discovery are dependent on interrelated factors that include:

- Corporate, research, and marketing department cultures
- Individual scientists working within a company
- Synergies among various research disciplines
- Morale and quality of research management
- Research and Development leadership
- Individual and corporate experience
- Extent to which the management of a company is accustomed to risk
- Serendipity
- A particular company's market franchise (the therapeutic areas in which it markets drugs)

Inevitably, many drugs are described in terms of the individuals who were associated with improving the research effort, often against considerable scientific and corporate differences. Albeit making for interesting (and often exciting) reading , drug discovery is inevitably a team effort requiring constant iteration based on

experimental findings, planning, and synergies across many different scientific, development, and marketing disciplines.

The drug discovery and development process can be divided into distinct phases that interface with each other. ***The first step*** involves deciding on a given therapeutic target, which entails an iterative process involving the current state of the understanding of disease etiology, scientific knowledge and available technology, unmet medical need, and commercial opportunity. In diseases such as hypertension, for which there are many effective medications already in the marketplace, new programs directed at this target must focus on significant additional benefits. For agents being developed for diseases such as Alzheimer's disease (AD) or cancer, for which there are currently no effective or safe treatments, the decision to target the disease state is an easier one to justify even though the science involved represents a significantly greater risk.

The second phase of drug discovery emerged with advances in enzymology and protein biochemistry. Many of the biological pathways and processes were identified as enzymologists and pharmacologists defined new enzymes, receptor ligands, and their functions. As a result, although serendipity was still a major factor, drugs were now directed toward distinct molecular targets, the evolution of this phase was rapid, with increased substantial support of the biomedical sciences, growing sophistication in computational processing, and the availability of the personal computer.

The third phase of drug discovery is one represented by the popular vision of computer-driven discovery and development of compounds that not only treat the symptoms of the disease but also lead to its cure. The potential for the intellectually rigorous targeting of therapeutic agents to molecular targets whose genetics, structure, function, and pathophysiology are well understood is a noble goal. Unfortunately, this somewhat naive view of several promising yet still emerging technologies presumes more than biomedical science has learned to date. Although there are superior medications to treat hypertension, asthma, schizophrenia, and anxiety, for example, their etiology remains unknown, leading to a major dependence on hypothesis testing. Interestingly, the drug discovery wheel has come full circle as fermentation, plant, marine, and invertebrate sources have again emerged as important sources of novel therapeutic entities. Among the compounds identified in the past decade from such sources are the cholecystokinin antagonist asperlicin and the immunosuppressant FK 506

Once a disease has been targeted and the appropriate test protocols have been assembled, it is still uncertain whether a particular approach will result in a drug. Due to this drug discovery that is frequently unappreciated and unreported. If drug targets and the design of their ligands were as simple as the application of new technologies, more than 20 percent of the effort in drug discovery would reach fruition in the identification of clinical moieties. As it stands, 80 percent of the effort is valuable, if ultimately nonproductive, hypothesis testing

1.2.2. Needs of Medicinal Chemist

The goals of medicinal chemist are, to design and synthesize novel bioactive compounds. Trial and error screenings is very costly affair as well as less efficient and time consuming. At present, to discover a single marketable compound requires approximately >10,000 molecules to be tested {ratio is 1 in 10,000} and the expenses have increased exponentially

.

Hence the days of trial and error screening have gone, as chemists can no longer rely on their own experience and knowledge. Therefore, only molecules with a good chance of activity should be prepared and tested for which necessary is the knowledge of physico chemical properties of the target molecules and the interactions between a drug and a receptor

The ideal system to enable medicinal chemists to achieve their targets should be able to:

- obtain the stable conformation for a given molecule with its energy;
- Calculate atomic charges and interaction, electrostatic potentials, frontier orbitals, and indexes related to reactivity,
- describe drug – receptor interactions and compute their energies
- Superimpose and compare geometric and electronic molecular models
- calculate physicochemical properties of a molecule such as partition coefficient

- Dipole moment pKa's for different proton, spectroscopic coupling constant etc

- display molecular models in several different ways, special features for complex molecules like proteins

- Find and display quantitative and qualitative relationships between molecules and biological activities.

- Stimulate the course of a chemical reaction and to consider the possible byproducts;

- Make as easy as possible exchange of data between different systems

1.2.3. Role of Medicinal Chemist in Drug Discovery

The modern medicinal chemist, although part of a team, has particularly a crucial role in the early phases of drug discovery. The chemist trained to prepare new chemicals, with an acquired knowledge of the target disease and of competitive drug therapies, has an important part in framing the hypothesis for the new drug project, which then sets the objectives for the project. The chemist also helps to decide which existing chemicals to screen for a lead compound and which screening hits need to be re-synthesized for biological evaluation. Purification and proper characterization of the new chemicals is also the responsibility of the chemist. When an *in vitro* 'HIT' is identified, the chemist decides on what analogous compounds should be obtained or synthesized to explore the SARs for the structural family of the compound with an effort to maximize the desired activity. Developing *in vivo* activity for the hit compound in an appropriate animal model is also mainly the responsibility of the chemist. This can often be one of the most difficult steps to accomplish because

several factors, such as absorbability, distribution *in viv*o, rate of metabolism and rate of excretion (ADME), all present hurdles for the chemist to solve in the design and preparation of new, analogous chemicals for testing. The goal at this stage is to maximize efficacy while minimizing side effects in an animal model.

For the medicinal chemist to overcome all the challenges outlined above, several skills are required. These include a thorough knowledge of modern organic chemistry and medicinal chemistry, an understanding of the biology that relates to the target disease, an understanding of the pharmacological tests used in the project and sufficient knowledge of the factors that influence ADME characteristics of chemicals *in viv*o. Furthermore, they should also have an understanding of clinical medicine that pertains to the target disease; knowledge of the regulatory requirements for related drugs; a current knowledge of competitive therapies, both in the market and under development by competitors; a thorough knowledge of the literature that is relevant to the target disease; familiarity with the many newer technologies available to facilitate drug discovery; and an entrepreneurial attitude in behaving as an innovator and inventor. Finally —and of crucial importance to the timely success of the project — the chemist must show superior interpersonal skills throughout the life of the project to interact effectively with colleagues from other disciplines to achieve project goals.

1.2. 4 Influence of Drug Metabolism on Drug Development

Preclinical drug metabolism and pharmacokinetic studies play an important role in lead identification and optimization.[11, 12] After administration the drug reaches to blood stream **(fig1.1).** This process is called ***absorption***.

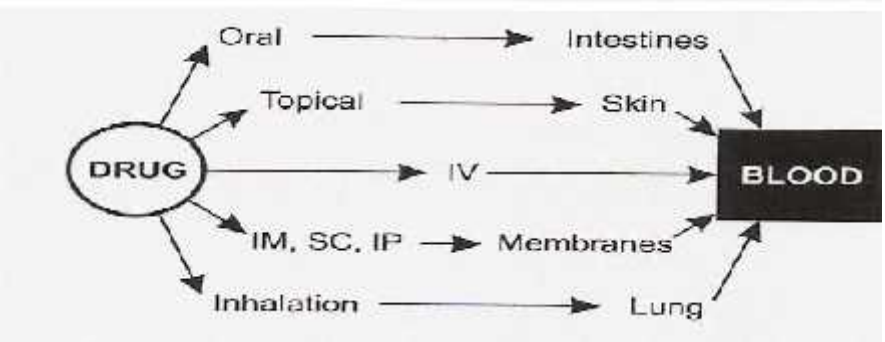

Fig: 1.1 Absorption of drug after administration

The drug then distributes rapidly between plasma and blood cells and also between plasma proteins. Lipid soluble drug cross the cell membrane and distributes into intracellular fluid of various tissues. This process is called ***distribution of drug***. Then comes the process of ***elimination***. The drug can be eliminated either by excretory route such as urine, bile etc or indirectly through enzymatic or biochemical transformation by liver. The later part is called ***metabolism***. The study of whole process is called ***ADMEstudies*.(Fig1.2)**

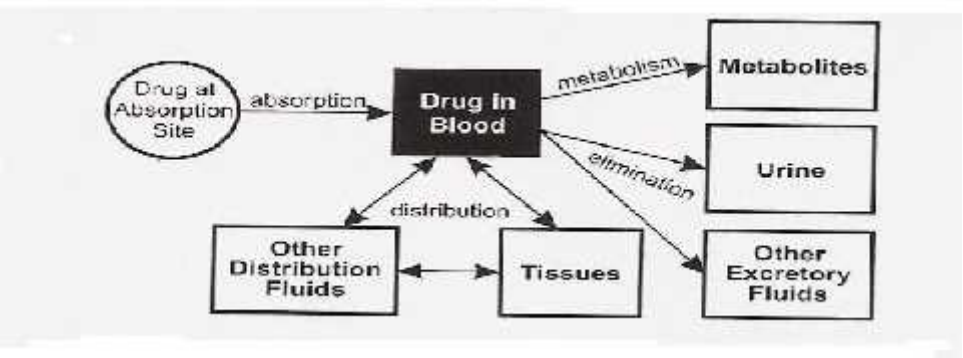

Fig1.2 Schematic representation of drug's path from blood

Information on drug metabolism plays a significant role in selection and further characterization of drug. In drug development it is important to have information about the enzymes responsible for metabolism, toxicity, metabolic stability, and pathway. Metabolite profiles are important for designing prodrugs and pharmacologically active metabolite. The drug metabolism can be altered by structural modification. Highly hydrophilic or highly lipophilic groups are not good as they result in very slow or very fast excretion rates.

The information obtained from preclinical drug metabolism studies are used to introduce functional groups which alter the physical properties to make the compound more metabolically stable. Thus it can be concluded that final selection of a successful drug lead depends on the drug metabolism studies

1.2.5 Development and introduction of drugs

Technological innovation and pressure of competition have caused enormous changes in drug discovery process. The drug discovery process is rapidly evolving due to technological development in target identification.**(Fig 1.3)** New chemical entities enter the drug discovery pipeline through combinatorial synthesis and rational drug design where information about the target of action is used to design the lead compound High Throughput Screening(HTS) helps in the identification of lead. In the second stage physiological properties as lipophilicity, solubility and stability are determined. These properties are helpful in predicting the protein binding and absorption in gastrointestinal tract. During lead optimization the selected leads are further screened. Most candidate fail at this stage and very few proceed for further

development Both ***invitro*** and ***invivo*** studies are carried out with a view to find most active lead with most appropriate safety profile. Knowledge of pharmacokinetic and metabolic characteristics is required in designing appropriate clinical trials.[13,14]

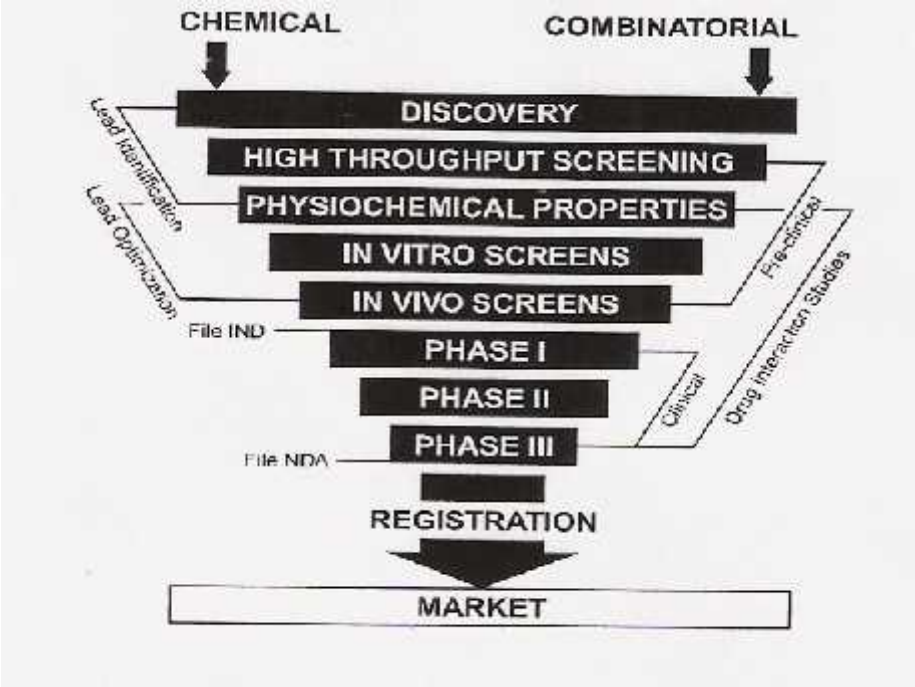

Fig: 1.3: Drug development process

In order to determine the therapeutic effects, side effects and toxicity, the drug discovery and modification process is accompanied by several tests in laboratories. Simultaneously the organic chemical reactions are carried out to study the stability of drugs in various solvents and under different conditions and dosage forms. When clinical trials are initiated, the manufacturer usually enlisted several groups of clinicians and clinical pharmacologists to test the drug. These tests take one to two years. As soon as the medicinal values in the clinic are established with reasonable certainty, the drugs are distributed to pharmacists and physicians for use.

1.3 Historical Background

At the time when Mendeleev was constructing the famous periodic table, Crum-Brown and Fraser[15] (1868) put forward the suggestion that physiological activity "Φ" of a series of quantemized strychnine derivatives depend on constitution (c), and suggested the following mathematical equation- (1)

$$\Phi = f(c) \text{----------} \quad \textbf{(1)}$$

Later Meyer[16], Overton[17] and Ferguson[18] applied the above principle and attempted the correlation of narcotic potencies with partition coefficient and thermodynamic parameters.

Later, in the era of the advancement of physical organic chemistry, Hammett [19] in 1937 expressed the chemical reactivity of meta- and para-substituted benzene derivatives by eq. (2), where KH is the rate constant for the parent (unsubstituted) molecule and Kx is the rate constant for the derivative.

$$\log (Kx/KH) = \rho\sigma x \text{----------} \quad \textbf{(2)}$$

The substitution constant σx refers to the electronic effect (potential) of the substituent relative to hydrogen. It is a parameter applicable to many different types of reactions -characterized by different ρ values –whose relative rates depend on the electron donating or withdrawing power of the substituent.

Since σ is defined in terms of ionization constants, which can be related to, free energies through the familiar equation (3), the eq. (2) or any of its transformation can be considered as a linear free energy relationship (LFER).

$\Delta G = -RT\ln K$ ----------**(3)**

The task of setting up a 'biological Hammett equation' was approached by three groups: Hansen[20], Zahradnik[21] and Hansch[22]simultaneously in 1962. The results of the first two groups were disappointing but Hansch and colleagues, by their many brilliant contributions showed that lipophilicity of the substituents in indole acetic acid and phenoxyacetic acid was an important determinant of plant growth regulatory activity. Octanol-water system was chosen as the standard system for determining lipophilicity (Log P) or hydrophobic substituent constant (π_x), the latter is defined by equation 4, analogous to Hammett[23] equation, for practical reasons. The choice of choosing octanol as substitute for lipid has been justified on many grounds[24, 25]

$\log (P_X/P_H) = \pi_{x--}$ ----------**(4)**

1.4 Emerging technologies in drug discovery process

As noted, a major target in the drug discovery process is to develop a high degree of “rationality.” This would represent an intellectually rigorous approach incorporating computer-assisted molecular design (CAMD), limited but sophisticated chemical synthetic effort, and highly focused biological assays. To many, rational drug design suggests that it is now possible, with the knowledge of the three-

dimensional (3-D) structure and sequence of various drug targets, to design new compounds, by iteration, on a computer. Thus, the drug design process may become significantly less of a risk and more resource efficient [26-27]

However, many of the enabling technologies that support the rational design approach are still in the emergent stage; that is, they have theoretical promise usually based on their use in the retrospective analysis of known compounds. With the exception of the enzyme thymidylate synthetase, molecular modeling techniques have yet to be generally used in a predictive manner.[28,29]

Receptor binding—the use of radiolabeled ligands to "tag" drug targets, receptors, and enzymes-have revolutionized the ability to determine compound structure-activity relationship (SAR) in a rapid, cost-effective manner. This technology resulted from early work on characterizing the insulin receptor by **Roth and Cuatrecasas**. In its present form (and diversity), receptor binding was driven by the work of **Snyder and coworkers** such that binding assays for nearly 100 receptors or enzymes have been developed. The technology has also been used in the identification of new receptors and receptor subtypes.[30-32]

Targeted screening involves the use of the receptor-binding or enzyme inhibition technique to evaluate large numbers of compounds (20,000 to 50,000 per year) in multiple assays to identify new pharmacophores. Compound sources include herbal, marine, and bacterial fermentations as well as chemical compound libraries; the latter include dissimilar pharmacophores from chemical companies, compounds synthesized as part of a directed chemical effort, and novel structures with no known biological activity. Targeted screening is an iterative process dependent on a finite

availability of compounds and binding assays. Ideally, as newer targets are identified and assay systems for them are developed, compounds should be reevaluated in a continuous manner [33,34]

Molecular modeling, or CAMD, is an emerging technology that makes use of knowledge of the steric and electronic aspects of the receptor/ligand, enzyme/substrate interaction to identify pharmacophores or aid in their design or both[35,36]

The target/ligand interaction can be studied from three advantage points[37]

- Knowledge of the SAR within a series and among series of pharmacophores, in effect approaching the receptor or enzyme from the drug perspective.
- Knowledge of the structure of the receptor or enzyme, approaching the problem from the receptor viewpoint
- Information regarding the receptor/ligand, enzyme/substrate interaction derived by 2- or 3-D nuclear magnetic resonance (NMR), x-ray crystallographic, or other structural protein analysis methods.[38]
- Each of these approaches has inherent limitations. The approach based on structure activity relationship (SAR) studies is limited in that the protein target (receptor or enzyme) is normally configured on a computer database in a minimal energy configuration with an approximation of water content. This approach has traditionally assumed that the protein and ligand have limited degrees of flexibility, a constraint that reflected the computational power available. It was noted that the receptor/ligand interaction could involve more than a single step and that receptors can induce changes in ligand

conformation features that had been known for enzymes for a number of years. These facets of the receptor/ligand interaction present additional dimensions to the CAMD process[39].

- Knowledge of the structure and 3-D conformation of the protein target provides an opportunity to identify the amino acid sequences and conformations that are responsible for ligand recognition and efficacy. These can be derived by knowledge of the primary sequence and, for a receptor, knowledge of which transmembrane helices are involved in ligand recognition. The interaction of various pharmacophores and compounds within a pharmacophore series can be used to identify the critical amino acids, when molecular biology is used to change these critical amino acids as point mutations, their importance in defining the ligand recognition parameters can be assessed[40].
- Additional information regarding the biophysical aspects of the protein/ligand interaction in real time using NMR can then be used to hierarchically integrate information from the previous two approaches to gain a more concise understanding of those properties of a molecule that impart selectivity, activity, and efficacy at a given protein[41]

1.5 Drug Designing

In the present era, there has been momentum in the process of drug research and development. At the center of the process of drug discovery are the element of drug design with which the process starts and will continue to guide the course of drug discovery and development, to the stage when the product is introduced into the clinical trials.

Activity of any drug depends upon the physicochemical properties. To develop a mathematical model, different descriptors of physical properties are taken and the activity of the drug is predicted accordingly. QSAR is most applicable when 3D structure of the target is not known; SAR information is obtained based on the compounds synthesized and their biological activities.

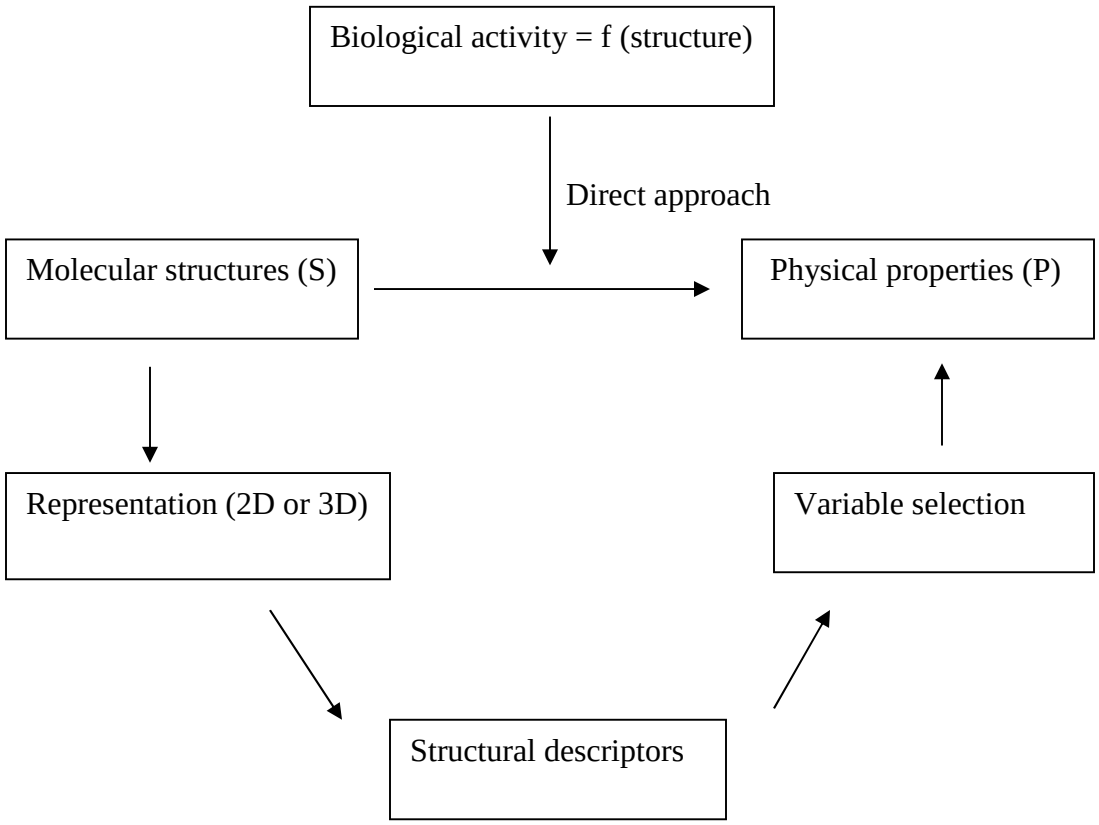

Figure 1.4: Flow of Drug Designing.

Drug design basically encompasses ***lead identification***, ***design of prototype***, based on this *lead* ***and lead optimization***. In a broader sense drug design implies **random evaluation of synthesis as well as natural products in bioassay systems, creation of newer drug molecules based on biologically active prototype** derived from either plant or animal kingdom, **synthesis of congeners** displaying interesting biological actions, the basic **concept of isosterism and bioisosterism** and finally precise design of a drug to enable it to **interact with a receptor site efficaciously**.

Earlier drug discovered by **classical methods** were designed with an almost **trial and error methods.** This made the process of drug discovery tedious, expensive and lengthy; the availability of computerized molecular modeling system has simplified the conditions for design of a new drug molecule**.**

1.5.1Challenges in drug design

Medicinal chemist faces certain difficulties and challenges in designing drugs towards specific target receptors. These are:

1. **Characterize medical condition and determine receptor targets**

   Nearly all biological processes in human body are tightly interconnected. There are side effects of nearly all drugs. The human body is a homeostatic machine, and tries to attain equilibrium. As a result, the body will attempt to counteract any pharmacotherapeutic intervention. The scientific study must first be done in order to determine the biological and biochemical problems that underlie the disease process This often takes years of study in order to characterize the targets for a potential drug

2. **Achieve active site complementarity: steric, electrostatic, and hydrophobic.**

   The process of ligand design starts just after discovering the receptor target. The designed ligand must be complimentary with the active site of receptor target. The designed ligand will have no chance of interacting with the receptor, if the pharmacophore is absent in receptor in order for recognition and binding to occur.

3. Consider biochemical mechanism of receptor.

The next challenge is to find the biochemical mechanism of the receptor target

4. Adhere to laws of chemistry

The challenge now comes is designing a suitable ligand.It is most important task of entire drug discovery process

5. Synthetic feasibility

The challenge then comes is the synthesis of novel ligand.It is of no use to design the ultimate drug, if it can not be manufactured The laws of chemistry state that each atom type has a specific size, charge, and geometry with respect to the number and types of neighboring atoms that it can be joined to. Within these rules, the drug developer must creatively propose suitable chemical structures that satisfy the requirements.

6. Biological considerations

Last point is the biological considerations to the development of new drugs. Since liver is the major organ of detoxification in the human body. Any drug that is taken undergoes a number of chemical reactions in the liver as the body attempts to neutralize foreign substance. These reactions give us knowledge as how drugs are modified as the body eliminates them. These constraints must also be taken under consideration as novel drugs are developed.

1.5.2 COMPUTER ASSISTED DRUG DESIGN [42-46]

Computer-assisted drug design (CADD), also called computer-assisted molecular design (CAMD), and represents more recent applications of computers as tools in the drug design process. In CADD, attempts are made to find a ligand (the putative drug) that will interact favourably with a receptor that represents the target

site. Binding of ligand to the receptor may include hydrophobic, electrostatic, and hydrogen-bonding interactions. In addition, solvation energies of the ligand and receptor site are also important because partial to complete desolvation must occur prior to binding. Computers are an integral part of the drug design process and have a large number of applications, which include structure analysis, structure comparisons, lead compound design, identification of active conformations and pharmacophores,combinatorial library design, protein and binding site structure,ligand binding,QSAR and 3-D QSAR.,

Drug design is an iterative process, which begins with a compound that displays an interesting biological profile and ends with optimizing both the activity profile for the molecule and its chemical synthesis. The process is initiated when the chemist conceives a hypothesis, which relates the chemical features of the molecule (or series of molecules) to the biological activity. Over the past two decades, the center of gravity (the intellectual focus) of medicinal chemistry has shifted dramatically from, how to make a molecule, to what molecule to make. The information feeding the drug design effort is increasingly quantitative, building upon recent developments in molecular structure description, combinatorial mathematics, statistics, and computer simulations. Collectively these areas have led to a new paradigm in drug design, which has been referred to as quantitative information analysis.

This approach to CADD optimizes the fit of a ligand in a receptor site. However, an optimum fit in a target site does not guarantee that the desired activity of the drug will be enhanced or that undesired side effects will be diminished. Moreover, this approach does not consider the pharmacokinetics of the drug. The approach used in

CADD is dependent upon the amount of information that is available about the ligand and receptor. Ideally, one would have 3-dimensional structural information for the receptor and the ligand-receptor complex from X-ray diffraction or NMR.

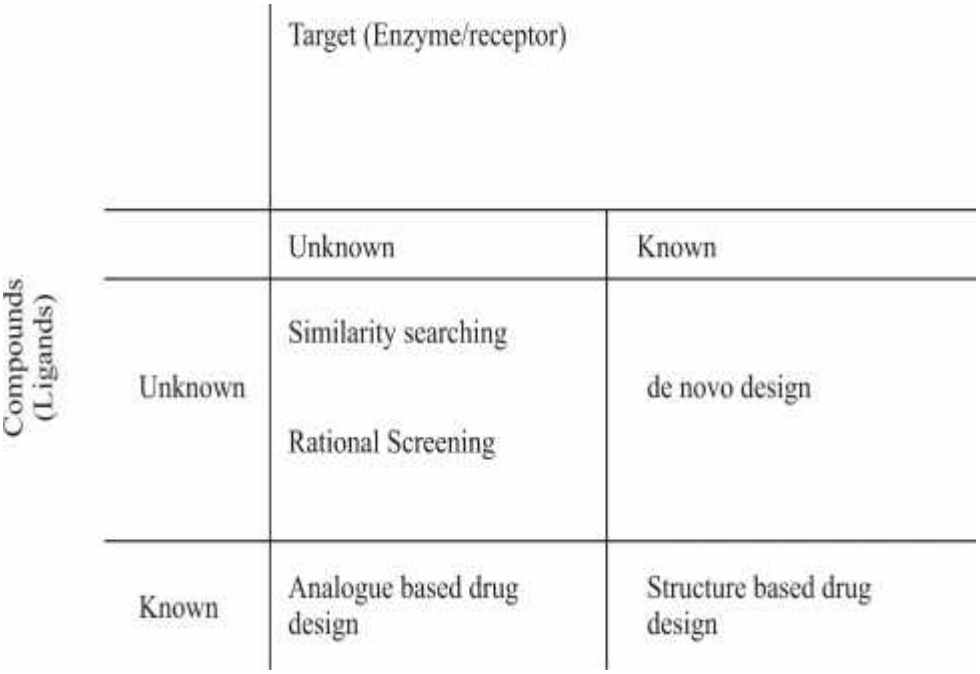

Figure 1.5: Four major cases in CADD also known as "direct" and "indirect design when the structure of the target is respectively known or unknown

Based on the information that is available, one can apply ligand-based or receptor based molecular design methods.

The **ligand-based** approach is applicable when the structure of the receptor site is unknown, but when a series of compounds have been identified that exert the activity of interest. To be used most effectively, one should have structurally similar compounds with high activity, with no activity, and with a range of intermediate activities. For this approach one finds the lowest energy conformers of the most rigid compounds and superimposes them. Conformational searching on the more flexible

compounds is then done while applying distance constraints derived from the structures of the more rigid compounds. Ultimately, all of the structures are superimposed to generate the pharmacophore. This template may then be used to develop new compounds with functional groups in the desired positions. Here, an assumption is made that the minimum energy conformers will bind most favourably in the receptor site.

The **receptor-based** approach to CADD applies when a reliable model of the receptor site is available, as from X-ray diffraction, NMR, or homology modeling. With the availability of the receptor site, the problem is to design ligands that will interact favourably at the site, which is a docking problem. Docking is the superimposition of the designed ligands with the present receptor model. This gives almost an exact picture of the receptor –ligand model which helps to understand the interactions well.

The **QSAR approach** is a rational approach to lead optimization when the structure of the target is not known. The underlying premise of QSAR is that there is a relationship between the biological and pharmacological activity of a compound, and its structural, physical and chemical properties. The steps followed are:

1) Calculation of various physicochemical and structural parameters.
2) Regression analysis and generation of QSAR equations.
3) Activity prediction of new designed compounds.

The computer aided modeling (CAM), which includes both molecular modeling and quantitative structure activity relationships, accelerated the development at

the drug discovery phase in terms of reducing time and money. In the present scenario of fast advancement in molecular structural biology and computer technology, the structure activity relationship studies based techniques can be used for both direct (where 3dimensional molecular structure of biological target is known) and for indirect (where the molecular target structure is unknown) design.

1.6 Quantitative Structure-Activity Analysis: The Center of Gravity in Modern Drug Design

Over the past two decades, the center of gravity (the intellectual focus) of medicinal chemistry has shifted dramatically from, how to make a molecule, to what molecule to make[47]. The challenge now is the gathering of information to make decisions regarding the use of resources in drug design. The information feeding the drug design effort is increasingly quantitative, building upon recent developments in molecular structure description, combinatorial mathematics, statistics, and computer simulations. Collectively these areas have led to a new paradigm in drug design, which has been referred to as quantitative information analysis[47].Within this paradigm there have evolved two broad approaches to drug design. ***The first is a mechanism-based approach*** where the location of pathology is found, followed by the identification of a macromolecule possessing effectors responsible for a critical biochemical event. Isolation, X-Ray analysis and amino acid identification of the effector region is the next sequence of studies. Drug design is then built around this information using computer simulations.

The second broad approach to drug design is employed when there is a paucity of information about an effector and its structure. This is often the case especially when the design goal is a physical or pharmacodynamic properties where there is no effector. This approach depends upon the illumination of information by probing a biological system with series of molecules with known structures. The relationship of these structures to some measured properties of the biological system is crafted into a mathematical model. This approach had its origins in the 60's and has become very important in industrial and academic drug design and basic research. The term used for this approach is quantitative structure-activity relationship (**QSAR**). The importance of QSAR has accelerated over the past decade with the explosive growth of combinatorial chemistry. With this technology, it is possible to synthesize and test thousands of compounds in a short time. The QSAR methodology is essential in the processing of this enormous amount of information into predictive models.

QSAR is a process whereby the structures of a set of compounds are quantified and then compared to the numerical values of a biological activity or a physical property. The challenge here has been to find some numerical code for a molecule or a fragment that is information-rich. This structure information and the measured property or activities are then processed into a mathematical model of relationship. From a quality model it is possible to predict and to design compounds for synthesis and testing that have a good possibility for activity.

Some pioneer work in molecular structure description includes the use of molecular orbital theory[48] molecular topology[49,50] and electronegativity coupled with topology.[51]A medicinal chemist must have a rich background in all of these methods for their role in drug design is central. The medicinal chemist is a decision-maker, a

leader of a team directing the choice of direction of a project, the synthesis and the testing of compounds. The department has a rich background in these methods and continues to explore new methods of information generation for drug design.

It is assumed that the potency of certain biological activity exerted by a series of congener compounds is expressible in terms of a function of various physicochemical parameters of the compounds if the function could be formulated showing that certain physiochemical properties are favorable to the activity. The structural modification, which enhanced such properties, would be expected to generate compounds of potent activity. From the structure activity relationship, at present, it is widely accepted that the chemical formula of a substance contains important information on its physical and biological properties. This information can be made available by derivation of QSAR equations. The knowledge of relation between structure and property make possible the design of substance with predetermined physico chemical parameters, which is an important task in discipline of chemistry, biology and medicine. For each of the analysis correlation with such physicochemical parameters such as π,σ, F, MR, Es values are examined and statistically best equation is selected.[52]

1.6.1 Principle of Quantitative Structure Activity Analysis

The resulting biological activity parameters “A” and molecular parameters “XI” are related. Since biological activity is dependent on molecular structure and the resulting properties, mathematical analysis reveals such connection in the form of so called

Quantitative Structure Activity Relationships

QSARs can be constructed for different purposes and according to different methods. Structure –response relationships describe the connections between the magnitude of a given biological affect and the drug structure in a set congeners, they can therefore be employed to optimize the effect on the basis of structure variations.

In structure–selectivity relationship, a new dimension becomes apparent; the simultaneous consideration of a variety of biological effects may allow an integral optimization. Structure activity relationships describe connection between the structure and the nature of the biological effect to be accepted. They may .in principle, lead to the prediction of new lead compounds.

The ultimate objective of QSARs is the prediction of either hypothesis on the mechanism of action or new analogues with high potency (fig: 1.6). After a lead has been discovered by any of the above methods. QSARs can help in recalling similar structure of biological activity profiles by computer analysis. QSAR may calculate the significance of the analogs at hand and there by suggest additional analogs to be synthesized. The contribution of QSAR after1964 was to quantify and to evaluate relationships between such physical properties and biological activities and to improve the methodology of measurement

It must always be remembered that the prediction depends on QSARs which have a statistical characteristic and thus there is always a certain probability of being in error. A basic limitation of structure activity analysis lies in fact that only such information can be extracted as is present in the biological data. Thus structure activity analysis by itself cannot lead to new concept in therapy. Though the manipulation of the information contained in biological data and chemical structure through the use of large computers and appropriate programs may also aid in developing new concepts

In recent years QSAR methods have been developed more to obtain much improved results .In this connection three dimensional Quantitative structure activity relationship (3D-QSAR) has been introduced.

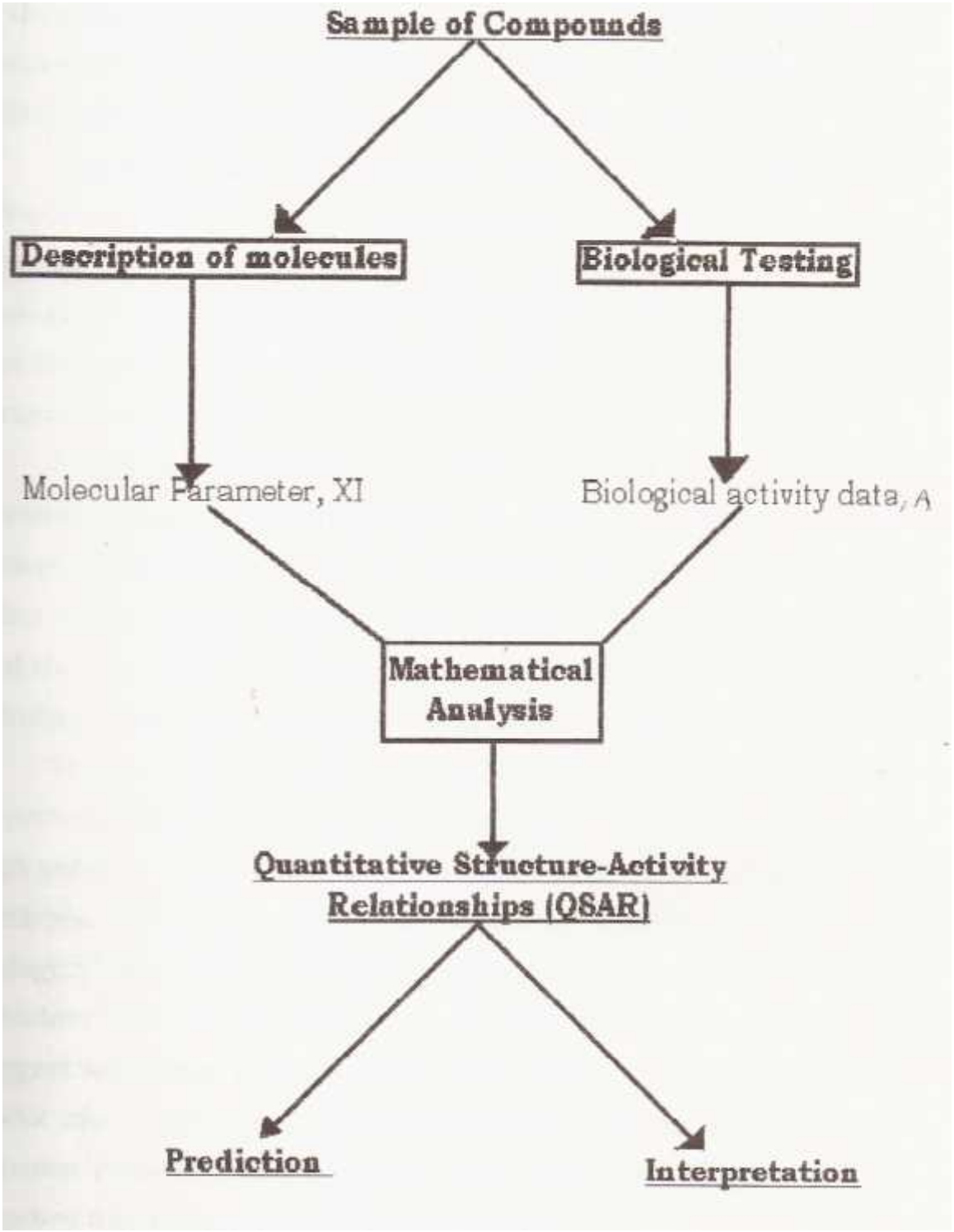

Fig: 1.6 Principle of Quantitative structure activity analysis

1.6.2 Importance of QSAR

The goal of a medicinal chemist is to design and synthesize novel bioactive compounds by employing rational methods. With the advancement in the field of molecular biology in terms of new biological specific targets and computer technology, the demands of selectivity and specificity have increased. The expenses in drug development process have increased exponentially. The two most important steps in drug design are ***lead identification and lead optimization.*** Generally lead molecule is obtained from prior old medicinal literatures of ancient civilizations i.e. folk medicine, screening of natural products, biochemical knowledge of disease processes i.e. pathogenesis or by high throughput random screening (HTS) programmes. But the next and crucial step: lead optimization is the rate limiting step in the process of selecting the candidate for drug development because it consumes a tremendous amount of wealth in terms of time, effort and money. **The QSAR studies help to** accelerate the drug development at this stage by reducing the above inputs and particularly the time. In this approach the biological activity is correlated with structural and physicochemical parameters in a series of compounds by mathematical models. These studies are useful in predicting the activity of unknown compounds and also in understanding the drug receptor interaction and mechanism of action particularly in cases where the structure of biological target is unknown.

1.6.3 Basic Requirements for the QSAR Analysis

- All analogues belong to a congeneric series exerting the same mechanism of action: This is a series of compounds with a similar basic structure but with varying substituents. These are known to act on the same target in the same manner.
- All analogues bind in a comparable manner.
- Same quantitative activity data for a series of compounds. This data is generated experimentally, assuming that all compounds have the same mode of action.
- The effects of isosteric replacement can be predicted.
- Binding affinity is correlated to interaction energies.
- Biological activity is correlated to binding affinity.

1.6.4 Drug Development Process (Need of QSAR)

The two important steps in drug design are lead generation and lead optimization. The steps involved in the process of drug discovery are

1. **New lead discovery**: Isolation of active substances from natural products, derivation and application of structure activity data, structure directed molecular design, modification of natural products, broad screening of synthetic compounds.
2. **Lead optimization**: Synthesis and testing of congeneric structures to develop structure activity relationships and mechanism of action based models,

calculate physical properties and correlate them with activity and comparison with standard drugs present.

3. **Pre-clinical lead development:** Drug formulation experiments, in vivo studies in animals, animal safety studies, drug metabolism studies, large scale synthesis.
4. **Clinical lead development**: Small scale safety and dose ranging test in healthy humans, develop clinical study protocols, recruit clinical investigators and patients for study, and carry out the study.

Considering both the potential benefits to human health and the enormous costs in terms of time and money of drug discovery, any tool or technique that increases the efficiency of any stage of the drug discovery enterprise will be highly prized. Computer-aided drug design (CADD) is one of these tools, which can be used to increase the efficiency of the drug discovery process. CADD is, however, not a direct route to new drugs, but it provides a somewhat more detailed map to the goal. The hope is that by providing bit and pieces of information and by helping to coordinate the information, CADD will help to save days and money for drug discovery projects**.**[53-55]

1.6.5 Applications of QSAR

The essential applications of QSAR are:

1.As an instrument for prediction

a) Estimation of physicochemical properties using substituent constant

b) Reduction of the number of compounds to be synthesized

c) Faster detection of the most active compound

d) Avoidance of synthesis of compound with same activity

2. As a diagnostic instrument

a) Information on possible types of interaction forces

b) Information on nature of the receptor

c) Information on the mechanism of action

d) Detection of inactive compounds

1.6.6 Significance of QSAR

A) Significance of results

The cross-validated r^2 or q^2 is obtained from regression analysis. It may be useful to remember that a value of 1.0 corresponds to perfect predictions, while a value of 0.0 implies there is no model at all. Again a q^2 of 0.5, halfway between no model and a perfect model, is likely to be helpful in decision-making. With low or negative values of q^2 one has to examine the residual plot for points lying especially far from the diagonal running from its lower left to upper right. Such “outlier” points indicate compounds whose target values were badly predicted. It is important to mention that cross validation calls attention to truly unique compounds (very different from all others in both shape and value of the target property). Thus with a negative

value of q^2 the safest thing to do is to try CoMFA after omitting the unique compound..

The q^2 values can be separated into three categories:

$q^2 > 0.6$ then the model is fairly good.

$q^2 = 0.4 – 0.6$ then the model is questionable.

$q^2 < 0.4$ then the model is poor.

There is no need to use the optimal number of components suggested by the program. One has to examine the r^2 values for each component listed from cross validation run. Each additional component should increase r^2 by 5-10%. If the increase in r^2 with an additional component is less than 5%, use the lower number of components instead. This is to be followed until the r^2difference between two components is larger.

B) Validity of QSAR Regression Equations

The following points are taken into consideration for

- ***Accepting regression equations.***
    - If the correlation coefficient ranges from 0.7-0.9.
    - If the F value indicates the overall significance is better than 95%.
    - If the standard deviation *s* is not much larger than the standard deviation of the biological data.

❖ ***The equation is rejected if:***

- If the number of variables included in the regression equation is unreasonably large.
- If the standard deviation ***s*** is smaller than the error in the biological data (over prediction).

1.6.7 Advantages of QSAR:

- Quantifying the relationship between structure and activity provides an understanding of the effect of structure on activity, which may not be straight forward when large amounts of data are generated.
- There is also the potential to make predictions leading to the synthesis of novel analogues. Interpolation is readily justified, but great care must be taken not to use extrapolation outside the range of the data set.
- The results can be used to understand interactions between functional groups in the molecules of greatest activity, with those of their target. To do this it is important to interpret any derived QSAR in terms of the fundamental chemistry of the set of analogues, including any outliers.

1.6.8 Disadvantages of QSAR:

- **False correlations** may arise through too heavy a reliance being placed on biological data, which, by its nature, is subject to considerable experimental error.
- Frequently, experiments upon which QSAR analyses depend**, lack design in the strict sense of experimental design**. Therefore the data collected may not reflect the complete property space. Consequently, many QSAR results cannot be used to confidently predict the most likely compounds of best activity.
- Various **physicochemical parameters are known to be cross-correlated**. Therefore only variables or their combinations that have little covariance should be used in a QSAR analysis; similar considerations apply when correlations are sought for different sets of biological data
- One of the limitation of QSAR is it is difficult to account for differences in the activities observed in case of the enantiomers and different conformers of the same molecules

The great advantage of the QSAR paradigm lies not in the extrapolation that can be made from known QSAR to fantastically potent new drug, but in the less spectacular slow development of science in medicinal chemistry”

Corwin Hansch C

Department of chemistry

Pomona college Claremont, California