CHAPTER I

General Introduction
1. Introduction

1.1. Computer Aided Drug Design (CADD)

Introduction of new therapeutic solutions is an expensive and time-consuming process. It is estimated that a typical drug discovery cycle, from lead identification through to clinical trials, can take 14 years [1] with cost of 800 million US dollars [2]. 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 [3, 4]. Hit rates are often low and many of these identified hits fail to be further optimized into actual leads and preclinical tests [5–7]. Among the late-stage failures, 40–60% was reportedly due to absorption, distribution, metabolism, excretion and toxicity (ADME/Tox) deficiencies [8–10]. Collectively, these issues underscore the need to develop alternative strategies that can help remove unsuitable compounds before the exhaustion of significant amount of resources [7].

In time, a new paradigm in drug discovery came underway, calling for early assessment of potency (activity) and selectivity of lead candidates, as well as their potential ADME/Tox liabilities. This helps reduce costly late-stage failures and accelerates successful development of new molecular entities. At the core of this, paradigm shift is the application of computational techniques to facilitate the discovery of new molecular entities. Computer-aided drug design (CADD) is a widely-used term that represents computational tools and resources for the storage, management, analysis and
modeling of compounds. It includes development of digital repositories for the study of chemical interaction relationships, computer programs for designing compounds with interesting physicochemical characteristics, as well as tools for systematic assessment of potential lead candidates before they are synthesized and tested. 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 [11] and to guide the synthesis of human haemoglobin ligands [12]. At that time, X-ray crystallography was expensive and time-consuming, rendering it infeasible for large-scale screening in industrial laboratories [13]. Over the years, new technologies such as comparative modeling based on natural structural homologues have emerged and began to be exploited in lead design [14]. 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 [15], Zanamivir for therapeutic or prophylactic treatment of influenza infection [16], Sildenafil for the treatment of male erectile dysfunction [17], and Amprenavir for the treatment of HIV infection [18]. CADD now plays a critical role in the search for new molecular entities [7, 13, 19]. Current focus includes improved design and management of data sources, creation of computer programs to generate huge libraries of pharmacologically interesting compounds, development of new algorithms to assess the potency and selectivity of lead candidates, and design of predictive tools to identify potential ADME/Tox liabilities. Here, we review major tools and resources that have been developed for expediting the search for novel drug candidates. The pipeline anatomy of a
typical virtual screening campaign from data preparation to post-screening analysis is discussed and related information have been accumulated in scientific literature and case reports. These data are collected and stored in a structured way in a number of databases. Every year, hundreds of biological data-bases are described [20]. At the same time, computational algorithms are actively developed to facilitate the design of combinatorial libraries. The most important data sources are reviewed in this section.

1.2. Quantitative Structure Activity Relationships (QSAR)

QSARs (Quantitative Structure–Activity Relationships) are based on the assumption that the structure of a molecule (i.e. its geometric, steric and electronic properties) must contain the features responsible for its physical, chemical, and biological properties, and on the ability to represent the chemical by one, or more, numerical descriptor(s). By QSAR models, the biological activity (or property, reactivity, etc.) of a new or untested chemical can be inferred from the molecular structure of similar compounds whose activities (properties, reactivities, etc.) have already been assessed.

The QSPR (Quantitative Structure–Property Relationship) acronym is used when a property is modeled. It has been nearly 40 years since the QSAR modeling firstly was used into the practice of agrochemistry, drug design, and toxicology, industrial and environmental chemistry. Its growing power in the following years may be attributed also to the rapid and extensive development in methodologies and computational techniques that have allowed to delineate and refine many variables and approaches used in this modelling approach. QSAR modelling is born in toxicology field. In fact, attempts to quantify relationships between chemical structure and acute toxic potency have been part
of the toxicological literature for more than 100 years. This relationship demonstrated the central axiom of structure–toxicity modeling— the toxicity of substances is governed by their properties, which in turn are determined by their chemical structure. Therefore, there are interrelationships between structure, properties, and toxicity. More than a century ago, Crum-Brown and Fraser [21] expressed the idea that the physiological action of a substance in a certain biological system (F) was a function (f) of its chemical constitution C:

\[ F = fC \]  \hspace{1cm} \text{[I]} \]

Thus, an alteration in chemical constitution, \( \Delta C \), would be reflected by an alteration in biological activity \( \Delta F \). At the turn of the 20th century, Meyer and Overton [22, 23] independently suggested that the narcotic (depressant) action of a group of organic compounds paralleled their olive oil/water partition coefficients.

In following years on the physical organic front, the seminal work of Hammett gave rise to the “s-r” culture [24, 25] in the delineation of substituent effects on organic reactions, while Taft devised a way for separating polar, steric, and resonance effects and introducing the first steric parameter, ES [26].

In 1962 Hansch et al. published their study on the structure-activity relationships of plant growth regulators and their dependency on Hammett constants and hydrophobicity [27]. Using the octanol/water system, a whole series of partition coefficients were measured, and thus a new hydrophobic scale was introduced. The parameter p, which is the relative hydrophobicity of a substituent, was defined in a manner analogous to the definition of sigma (18).
\( p = \log PX - \log PH \) ................................................................. [II]

PX and PH represent the partition coefficients of a derivative and the parent molecule, respectively. The contributions of Hammett and Taft together laid the basis for the development of the QSAR paradigm by Hansch and Fujita, which combined the hydrophobic constants with Hammett’s electronic constants to yield the linear Hansch equation and its many extended forms. There is a consensus among current predictive toxicologists that Corwin Hansch is the founder of modern QSAR. In the classic article [28] it was illustrated that, in general, biological activity for a group of ‘congeneric’ chemicals can be described by a comprehensive model:

\[
\log 1/C50 = a p + b e + cS + d \ ................................................................. [III]
\]

in which C, the toxicant concentration at which an endpoint is manifested (e.g. 50% mortality or effect), is related to a hydrophobicity term, p, (this is a substituent constant denoting the difference in hydrophobicity between a parent compound and a substituted analog, it has been replaced with the more general molecular term the log of the 1-octanol/water partition coefficient, log Kow), an electronic term, 1, (originally the Hammett substituent constant, s) and a steric term, S, (typically Taft’s substituent constant, ES). Due to the curvilinear, or bilinear, relationship between log1/C50 and hydrophobicity normally found in single dose tests, the quadratic p^2 term was later introduced to the model.

The rationale for Eq. (III) was given by McFarland [29]. He hypothesized that the relative activity of a biological active molecule, such as a toxicant, is dependent on: (I) the probability (Pr1) that the toxicant reaches its site of action, (II) the probability (Pr2)
that the toxicant will interact with the target at this site, and (III) the external concentration or dose. The delineation of these models led to explosive development in QSAR analysis and related approaches [30].

Besides the Hansch approach, other methodologies were also developed to tackle structure-activity questions. The Free-Wilson approach addresses structure-activity studies in a congeneric series as described in Equation (IV):

\[ BA = \sum ai xi + u \]  

Where \( BA \) is the biological activity, \( u \) is the average contribution of the parent molecule, and \( ai \) is the contribution of each structural feature; \( xi \) denotes the presence (\( xi = 1 \)) or absence (\( xi = 0 \)) of a particular structural fragment structure. In the years after the 1960s, the need to solve new problems, together with the contributions of many other investigators, generated thousands of variations of the Hansch approach to QSAR modelling, as well as approaches that are formally completely new. Hans Konemann [31] and Gilman Veith [32], who in the early 1980s developed multi-class-based, hydrophobic- dependent models for industrial organic chemicals, must share credit for the revival of QSAR. At present, the QSAR science, founded on the systematic use of mathematical models and on the multivariate point of view, is one of the basic tools of modern drug and pesticide design and has an increasing role in environmental sciences. QSAR models exist at the intersection of chemistry, statistics and biology, in toxicological studies. The development of a QSAR model requires these three components:
1) Data set that provides experimental measures of a biological activity for a group of chemicals.

2) Molecular structure and/or property data (i.e. the descriptors, variables, or predictors) for this group of chemicals; and

3) Statistical methods, to find the relationship between these two data sets.

The limiting factor in the development of QSARs is the availability of high quality experimental data. In QSAR analysis, it is imperative that the input data be both accurate and precise to develop a meaningful model. In fact, it must be realized that any resulting QSAR model that is developed is only as valid statistically as the data that led to its development. Data used in QSAR evaluations are obtained either from the literature or generated specifically for QSAR-type analyses. These data can consist of congeneric series of chemicals or assure structural diversity even within a chemical class. This diversity has allowed the generalization of more robust QSARs, applicable in an extended way. A structure–activity model is defined and limited by the nature and quality of the data used in model development and should be applied only within the model’s applicability domain. The ideal QSAR should:

(1) Consider an adequate number of molecules for sufficient statistical representation.

(2) Have a wide range of quantified end-point potency (i.e. several orders of magnitude) for regression models or adequate distribution of molecules in each class (i.e. active and inactive) for classification models.
(3) Be applicable for reliable predictions of new chemicals (validation and applicability domain) and

(4) Allow to obtain mechanistic information on the modelled end-point. Chemical descriptor(s) include empirical, quantum chemical or non-empirical parameters.

Empirical descriptors may be measured or estimated and include physico-chemical properties (such as for instance log P). Non-empirical descriptors can be based on individual atoms, substituents, or the whole molecule, they are typically structural features. They can be based on topology or graph theory and, as such, they are developed from the knowledge of 2D structure, or they can be calculated from the 3D structural conformations of a molecule.

A variety of properties have been also used in QSAR modeling, these include physico-chemical, quantum chemical, and binding properties. Examples of molecular properties are electron distribution, spatial disposition (conformation, geometry, and shape), and molecular volume. Physicochemical properties include descriptors for the hydrophobic, electronic, and steric properties of a molecule as well as other properties including solubility and ionization constants. Quantum chemical properties include charge and energy values. Binding properties involve biological macromolecules and are important in receptor-mediated responses. A big problem related to molecular descriptors is their reproducibility: experimental values can differ greatly even when referred to the same compound [33]. Several approaches have been developed for the theoretical calculation of logP [34-38], but also in these calculations it is not uncommon to have differences of several orders of magnitude [39]. In modern QSAR approaches, it is
becoming quite common to use a wide set of theoretical molecular descriptors of different kinds, able to capture all the structural aspects of a chemical to translate the molecular structure into numbers. Different descriptors are different ways or perspectives to view a molecule, taking into account the various features of its chemical structure, not only monodimensional as the simple counts of atoms and groups, but also bi-dimensional from the topological graph or three-dimensional from a minimum energy conformation.

A lot of software calculates wide sets of different theoretical descriptors, from SMILES, 2D-graphs to 3D-x,y,z-coordinates. Some of the most used are mentioned here: ADAPT [40, 41], OASIS [42], CODESSA [43], MolConnZ [44], and DRAGON [45]. It has been estimated that more than 3000 molecular descriptors are now available, and most of them have been summarized and explained [46-48]. The great advantage of theoretical descriptors is that they can be calculated homogeneously by a defined software for all chemicals, even those not yet synthesized, the only need being a hypothesized chemical structure, thus they are reproducible. Modeling methods used in the development of QSARs are of two types in relation to the modelled response: a potency of an end-point (a defined value of EC50) or a category/class (for instance Mutagen/Not mutagen). For the potency modelling, the most widely used mathematical technique is multiple regression analysis (MRA). Regression analysis is a simple approach that leads to a result that is easy to understand and, for this reason; most QSARs are derived using regression analysis. Regression analysis is a powerful means for establishing a correlation between independent variables (molecular descriptors X) and a dependent variable Y, such as biological activity Eq (V):

\[ Y = b + aX_1 + cX_2 + \ldots \]
For the modelling of categories, different quantitative models of classification can be applied. A wide range of classification methods exists, including: discriminant analysis (DA; linear, quadratic, and regularized DA), SIMCA (Soft Independent Modeling of Class Analogy), k-NN (k-Nearest Neighbours), CART (Classification and Regression Tree), Artificial Neural Network, Support Vector Machine, etc. In these techniques, the term “quantitative” is referred to the numerical value of the variables (the molecular descriptors) necessary to classify the chemicals in the qualitative classes. It is evident from the literature analysis that the QSAR world has undergone profound changes since the pioneering work of Corvin Hansch, considered the founder of modern QSAR modeling [27, 28, 30]. The main change is reflected in the growth of a parallel and quite different conceptual approach to the modeling of the relationships among a chemical’s structure and its activity/properties. In the Hansch approach, still applied widely and followed by many QSAR modelers, for instance Schultz et al. [49], Veith and Mekenyan [50], Benigni [51], molecular structure is represented by only a few molecular descriptors (typically log Kow, Hammett constants, HOMO/LUMO, some steric parameters) selected personally by the modeler and inserted in the QSAR equation to model a studied end-point. Alternatively, in a different approach chemical structure is represented, in the first preliminary step, by a large number of theoretical molecular descriptors which are then, in a second step, selected by different chemometric methods as the best correlated with response and included in the QSAR model (the algorithm).

The fundamental aim is the optimization of model performance for prediction. According to the Hansch approach, descriptor selection is guided by the modeler’s conviction to have a priori knowledge of the mechanism of the studied activity/property,
and the presumption to assign mechanistic meaning to any used molecular descriptor selected by the modeler from among a limited pool of potential modeling variables, normally well known and repeatedly used (for instance: Log Kow is a universal parameter miming cell membrane permeation, thus it is used in a lot of toxicity models, but it is also related to various partition coefficients such as bioconcentration/bioaccumulation, soil sorption coefficient, etc.; HOMO/LUMO are always selected for modeling chemical reactivity, etc.). On the other hand, the ‘statistical’ or chemometric approach, an approach parallel to the previous so-called ‘mechanistic’ one, is based on the fundamental conviction that the QSAR modeler should not influence, a priori and personally, the descriptor selection through mechanistic assumptions, but should apply unbiased mathematical tools to select from a wide pool of input descriptors, those descriptors most correlated to the studied response. The number and typology of the available input descriptors must be as wide and different as possible in order to guarantee the possibility of representing any aspect of the molecular structure. Different descriptors are different ways or perspectives to view a molecule, however the models must be developed taking into account the principle of parsimony, named the Ockham’s Razor: "entities should not be multiplied beyond necessity" or “avoid complexity if not necessary”. This principle is often paraphrased as “the simplest solution is the best.” Thus, descriptor selection must be performed by applying mathematical approaches (such as for instance evolutionary techniques, Genetic Algorithms, etc) with the final and crucial aim to maximize, as an optimization parameter, the predictive power of the QSAR model, as the real utility of any model is considered its predictivity. Regarding the interpretability of the descriptors, it is important to take into account that modeled response is frequently the result of a
series of complex biological or physico-chemical mechanisms, thus it is very difficult and reductionist to ascribe too much importance to the mechanistic meaning of the molecular descriptors used in a QSAR model. Moreover, it must also be highlighted that in multivariate models such as MLR models, even though the interpretation of the singular molecular descriptor can be certainly useful, it is only the combination of the selected set of descriptors that is able to model the studied end-point. If the main aim of QSAR modeling is to fill the gaps in available data, the modeler attention should be focused on model quality. In relation to this point, Livingstone states [52]: “The need for interpretability depends on the application, since a validated mathematical model relating a target property to chemical features may, in some cases, be all that is necessary, though it is obviously desirable to attempt some explanation of the “mechanism” in chemical terms, but it is often not necessary, per se”. Zefirov and Palyulin [53] took the same position, differentiating predictive QSAR, where attention essentially concerns the best prediction quality, from descriptive QSARs where major attention is paid to descriptor interpretability.

In fact, the first aim of any modeler should be validation for the predictive application of the QSAR model, for both the mechanistic approach and the statistical one. The famous “Kubinyi Paradox” [54, 55], emphasized also by Tropsha et al. in their famous papers: Beware of Q2 [56] and The Importance of being Earnest [57] is that: The “best fit” models are not the best ones for prediction! In fact, a QSAR model must, first of all, be a real model, robust and predictive, to be considered a reliable model[58]; only a stable and predictive model can be usefully interpreted for its mechanistic meaning, even so this is not always easy or feasible. QSAR model validation has been recognized
by specific OECD expert groups as a crucial and urgent point in recent years, and this has led to the development, for regulatory purposes, of the “OECD principles for the validation of QSAR models”[59]. The need for this important action was mainly due to the recent new chemicals policy of the European Commission (REACH: Registration, Evaluation and Authorization of Chemicals) [60], that explicitly states the need to use QSAR models to reduce experimental testing (including animal testing). Obviously, to meet the requirements of the REACH legislation it is essential to use QSAR models that produce reliable estimates, i.e., validated QSAR models. Thus, reliable QSAR model must be associated with the following information:

1. A defined endpoint.

2. An unambiguous algorithm.

3. A defined domain of applicability.

4. Appropriate measures of goodness-of-fit, robustness and predictivity.

5. A mechanistic interpretation, if possible.

The need for interpretability depends on the application, as a validated mathematical model relating a target property to chemical features may be all that is necessary, particularly when predicted data are needed for screening of large libraries of chemicals, though it is obviously desirable to attempt some explanation of the ‘mechanism’ in chemical terms[52,53].

To enhance their effectiveness and reduce the cost of development of the final medicinal products; that is, the QSAR technique can be considered as a low-cost/high-
return technique [61]. QSAR models thus can be considered to play a critical role in drug discovery and development by being involved in both the opening and the endgame phases of lead optimization [62].

The selection/utilization of an appropriate statistical methodology and structural descriptors is always vital in the development of a predictive QSAR/QSPR model that encodes the relationship between the structure of a molecule and its biological activity, chemical reactivity, or physical characteristics. It should be borne in mind that the predictive QSAR/QSPR models based only on the descriptors calculated from molecular structure are preferable for an early stage of screening in the drug discovery/development processes.

1.2.1. Purpose of QSAR

QSAR should not be seen as an academic tool to allow for the post-rationalization of data. We wish to derive the relationships between molecular structure, chemistry and biology for good reason. From these relationships we can develop models, and with luck, good judgment and expertise these will be predictive. There are many practical purposes of a QSAR and these techniques are utilized widely in many situations. The purpose of in silico studies, therefore, includes the following:

• To predict biological activity and physico-chemical properties by rational means.

• To comprehend and rationalize the mechanisms of action within a series of chemicals.

Underlying these aims, the reasons for wishing to develop these models include:
• Savings in the cost of product development (e.g. in the pharmaceutical, pesticide, personal products, etc.).

• Predictions could reduce the requirement for lengthy and expensive animal tests.

• Reduction (and even, in some cases, replacement) of animal tests, thus reducing animal use and obviously pain and discomfort to animals.

• Other areas of promoting green and greener chemistry to increase efficiency and eliminate waste by not following leads unlikely to be successful.

1.2.2. Applications of QSAR

The ability to predict a biological activity is valuable in a number of industries. Whilst some QSARs appear to be little more than academic studies, there are a large number of applications of these models within industry, academia and governmental (regulatory) agencies. A small number of potential uses are listed below:

• The rational identification of new leads with pharmacological, biocidal or pesticidal activity.

• The optimization of pharmacological, biocidal or pesticidal activity.

• The rational design of numerous other products such as surface-active agents, perfumes, dyes, and fine chemicals.

• The identification of hazardous compounds at early stages of product development or the screening of inventories of existing compounds.

• The designing out of toxicity and side-effects in new compounds.
• The prediction of toxicity to humans through deliberate, occasional and occupational exposure.

• The prediction of toxicity to environmental species.

• The selection of compounds with optimal pharmacokinetic properties, whether it be stability or availability in biological systems.

• The prediction of a variety of physico-chemical properties of molecules (whether they be pharmaceuticals, pesticides, personal products, fine chemicals, etc.).

• The prediction of the fate of molecules which are released into the environment.

• The rationalization and prediction of the combined effects of molecules, whether it be in mixtures or formulations. The key feature of the role of in silico technologies in all of these areas is that predictions can be made from molecular structure alone.

1.3. Molecular Docking

Molecular docking is a frequently used tool in computer-aided structure-based rational drug design. It evaluates how small molecule (substrate, inhibitor, drug or drug candidate) and the target macromolecule (receptor, enzyme or nucleic acid) fit together. This can be useful for developing better drug candidates and also for the understanding the nature of the binding. Herein in-silico COX2 docking study aims to rationalize the obtained biological data and to explain the possible interactions of the tested derivatives into the crystal structure of COX2 enzyme. Computer simulated automated docking study was performed using the widely distributed molecular grid-based docking program.
AutoDock 4.2, for docking of flexible ligand within flexible protein, where flexibility of the target protein is taken into account in the later type and ignored in the former one.

AutoDock scans the active site for low energy binding models and for orientations of the probe molecule, using a modified genetic algorithm that employs a local search (GALS) and pre-computed grids for the evaluation of the interaction energy. Both of the target Cox2 (PDB code: 1CX2) proteins were handled by using Accelrys Discovery Studio visualize v2.5 software [Accelrys Inc., San Diego, CA (2005)] and the representative amino acids of the ligand-binding site were selected within 5 Å neighborhood surrounding the embedded ligand; ibuprofen. The results of 10 randomly seeded runs were analyzed for each of the docked inhibitors. The docked inhibitors were assigned to a cluster if the atomic coordinates of the docked inhibitors exhibited a root mean square deviation (RMSD) of less than 0.5 Å difference from each other (RMSD-tolerance of 0.5 Å). The analysis was carried out for the top 10 docking clusters. The clusters were ranked from the averaged lowest energy obtained for members of the cluster to the highest. Each of the clusters that exhibited significant negative interaction energies were examined by Accelrys Discovery Studio visualize modeling program to determine their binding orientations.
1.3.1. Validation of the accuracy and performance of AutoDock

As cited in literature, if the RMSD (root mean square deviation) of the best docked conformation of the native ligand is $\leq 2.0$ Å from the experimental one, the used scoring function is successful. Therefore the validation of the docking accuracy was investigated by docking of the native co-crystallized ligands as to how closely the best docked conformation resembles the bound ligand in the biological method.[63,64]

The obtained success rates of AutoDock were highly excellent. Where the co-crystallized ligands of COX-2, namely ibuprofen and S58, respectively seem exactly superimposed on the native bound ones as shown in Figure 3. The RMSD of the docked ibuprofen into Cox1 were 0.91 and 2.51 Å, and that of the docked S58 into COX-2 were 0.40 and 0.34 Å by rigid and flexible docking for each receptor, respectively. Moreover, flexible docking involving AutoDock 4.2 seems to be more accurate, being of smaller RMSD values, to be of more resemblance to the biological co-crystallization.

1.4. The Non-steroidal Anti-inflammatory Drugs (NSAIDs)

Non-steroidal anti-inflammatory drugs (NSAIDs) are a group of drugs of diverse chemical composition and different therapeutic potentials having a minimum of three common features: identical basic pharmacological properties, similar basic mechanism of action as well as similar adverse effects. Moreover, all drugs in this group exhibit acidic character. Most NSAIDs are weak acids, with a pKa values in the range of 3.0–5.0 (acids of medium strength). NSAID molecules contain hydrophilic groups (carboxylic or enolic group) and lipophilic ones (aromatic ring, halogen atoms). In accordance with their acidic character, NSAIDs occur in the gastric juice in the protonated (lipophilic) form. Also in
the small intestine, there are conditions favorable for absorption of weak acids. NSAIDs exist in highly ionized forms in plasma. Low values defining NSAIDs distribution volume (from 0.1 to 1.0) in tissues may be a proof of poor distribution of these drugs in extra vascular systems. A very high degree of binding with plasma proteins (>97%) is the result of favorable amilphilic properties and accounts for the fact of displacing other drugs from protein binding of NSAIDs. Most NSAIDs are metabolized in the liver by oxidation and conjugation to inactive metabolites which are typically excreted in the urine, although some drugs are partially excreted in bile. Metabolism may be abnormal in certain disease states, and accumulation may occur even with normal dosage [65].

Non-steroidal anti-inflammatory drugs (NSAIDs) are better in use than steroidal drugs due to following properties:

1.4.1. Properties of NSAIDs

- Mildly analgesic
- Antipyretic
- Anti-inflammatory
- Act on sub-cortical sites such as thalamus and hypothalamus
- No affinity for morphine receptors
- In addition, tolerance and drug dependence do not develop to these drugs in patients.

1.4.2. Classification of NSAIDs

NSAIDs are classified according to their chemical structure into the following groups:

Figure (1-1)

1- Salicylic acid derivatives:

e.g., Acetylsalicylic acid (aspirin), sodium salicylate and methyl salicylate.
2. **Para-aminophenols:**

e.g., Paracetamol and phenacytne

3. **Pyrazolone** derivatives:

e.g., Phenylbutazone, oxyphenbutazone and propyphenazone.

4. **Indole** derivatives:

e.g., Indomethacin, sulindac.

5. **2-arylprippionic acid** derivatives (Profens):

e.g., Ibuprofen, fenoprofen, ketoprofen, naproxen, flurbiprofen and tiaprofenic acid

6. **N-arylanthranilic acids** (fenamic acids derivatives):

e.g., Mefanamic acid, tolfenamic acid, flufenamic acid, meclofenamic acid

7. **Arylacetic acid** derivatives:

e.g., Diclofenac, fenclofenac, aceclofenac, etodolac, indometacin, nabumetone, sulindac, tolmefin

8. **Oxicams:**

e.g., Piroxicam, Meloxicam, tenoxicam, lornoxicam and droxicam
Most NSAIDs act as non-selective inhibitors of the enzyme cyclooxygenase, inhibiting both the cyclooxygenase-1 (COX-1) and cyclooxygenase-2 (COX-2) isoenzymes. Cyclooxygenase catalyzes the formation of prostaglandins and thromboxane from arachidonic acid (itself derived from the cellular phospholipid bilayer by phospholipase A). Prostaglandins act (among other things) as messenger molecules in the process of inflammation [66, 68]
1.4.3. NSAID Mechanism of Action

The major mechanism by which the NSAIDs elicit their therapeutic effects (antipyretic, analgesic, and anti-inflammatory activities) is inhibition of prostaglandin (PG) synthesis.

Specifically NSAIDs competitively (for the most part) inhibit cyclooxygenases (COXs), the enzymes that catalyze the synthesis of cyclic endoperoxides from arachidonic acid to form prostaglandins.

Two COX isoenzymes have been identified: COX-1 and COX-2. COX-1, expressed constitutively, is synthesized continuously and is present in all tissues and cell types, most notably in platelets, endothelial cells, the GI tract, renal microvasculature, glomerulus, and collecting ducts. Thus COX-1 is important for the production of prostaglandins of homeostatic maintenance, such as platelet aggregation, the regulation of blood flow in the kidney and stomach, and the regulation of gastric acid secretion. Inhibition of COX-1 activity is considered a major contributor to NSAID GI toxicity. COX-2 is considered an inducible isoenzyme, although there is some constitutive expression in the kidney, brain, bone, female reproductive system, neoplasias, and GI tract. The COX-2 isoenzyme plays an important role in pain and inflammatory processes.
Figure (1-2): Inhibition of COX by NSAIDs

Generally, the NSAIDs inhibit both COX-1 and COX-2. Most NSAIDs are mainly COX-1 selective (e.g., aspirin, ketoprofen, indomethacin, piroxicam, sulindac). Some are considered slightly selective for COX-1 (e.g., ibuprofen, naproxen, diclofenac) and some others may be considered slightly selective for COX-2 (e.g., etodolac, nabumetone, and meloxicam). The mechanism of action of celecoxib and rofecoxib is primarily selective inhibition of COX-2; at therapeutic concentrations, the COX-1 isoenzyme is not inhibited thus GI toxicity may be decreased.

Other mechanisms that may contribute to NSAID anti-inflammatory activity include the reduction of superoxide radicals, induction of apoptosis, inhibition of adhesion molecule expression, decrease of nitric oxide synthase, decrease of pro-inflammatory cytokine levels (tumor necrosis factor-a, interleukin-1), modification of lymphocyte activity, and alteration of cellular membrane functions[69].

Central analgesic activity has been demonstrated in animal pain models by some NSAIDs such as diclofenac, ibuprofen, indomethacin, and ketoprofen. This may be because of the interference of prostaglandin (PGE1, F2 and F2a) mediated pain formation or with
transmitters or modulators in the nociceptive system. Other proposals include the central action mediated by opioid peptides, inhibition of serotonin release, or inhibition of excitatory amino acids or N-methyl-D-aspartate receptors. NSAIDs are mainly effective against the type of pain in which PGs sensitize pain receptors (inflammation and tissues) including the pain of arthritis, bursitis, pain of muscular and vascula origin and dysmenorrhea. The effectiveness of these agents against headache may result from their ability to inhibit PG-mediated cerebral vascular vasodilation. Antipyretic activity of NSAIDs results from inhibition of prostaglandin E2 (PGE2) synthesis in circumventricular organs in and near the preoptic hypothalamic area. Infections, tissue damage, inflammation, graft rejection, malignancies, and other disease states enhance the formation of cytokines that increase PGE2 production. PGE2 triggers the hypothalamus to promote increases in heat generation and decreases in heat loss.

1.4.4. Cyclooxygenase (COX)

The 70 kDa cyclooxygenase (COX) enzyme provides a fascinating and industrially relevant point of focus for the computational studies described in this dissertation. Described most simply, COX functions as a membrane-associated homodimer, catalyzing the committed step in the conversion of arachidonic acid (AA) to prostaglandin H2 (PGH2), following AA's release from membrane phospholipids. An overview of this reaction is depicted in Scheme I-1 below:
Figure (1-3): COX synthesizes prostaglandin G2 (PGG2) from arachidonic acid at the cyclooxygenase active site. PGH2 is reduced to prostaglandin H2 (PGH2) at the enzyme's peroxidase site.

COX is a binfunctional enzyme with two active sites. At the cyclooxygenase active site, COX exquisitely controls the region- and stereo-selective bis-dioxygenation and cyclization of AA to form PGG2. The reaction concludes at the entirely distinct peroxidase site, where intermediate PGG2’s C15 hydroperoxide is reduced to an alcohol to form PGH2. The genome codes two 60% sequence-identical isoforms of COX. Both COX-1 and COX-2 catalyze the same reaction, and the discussion which follows applies to both. COX-2 will be introduced in much greater detail later. Figure 1-3 below introduces the three-dimensional structure of COX.
Figure (1-4): The protein is shown on the left in white bonds, and the grid box is shown on the right side. The blue contours surround area in the box that are most favorable for binding of carbon atoms, and the red contours show area that favor oxygen atoms. A ligand is shown inside the box at upper right.

The COX reaction product PGH2 is unstable [70, 71] and its bioactivity is imparted downstream by tissue specific synthases which swiftly convert PGH2 to other bioactive prostaglandins, as shown in Figure 1-4 below. These prostaglandins, in turn, bind to G-protein-coupled receptors and effect diverse biological responses [72].
**Figure (1-5):** Major prostaglandins produced from PGH2

COX inhibitors (Aspirin being most famous) are therapeutically useful as anti-inflammatory, analgesic, anti-pyretic, and anti-coagulant agents. As a group, these inhibitors are commonly known as non-steroidal anti-inflammatory drugs (NSAIDs).

The physiological impacts of NSAIDs result from the drop in downstream prostaglandin concentrations which follow COX inhibition. To understand how COX inhibitors bind to the enzyme active site, it is helpful to review both the reactions catalyzed by cyclooxygenase, and the features that nature has evolved in the cyclooxygenase active site to bind arachidonic acid and tightly control its conversion to PGG2.
1.4.5. COX-2

In the 1980s, Needleman showed that cytokines and growth factors resulted in heightened COX expression (and activity) in fibroblasts [73, 74]. In 1991, a Harvard team cloned a cDNA of this intriguing up-regulated COX enzyme, and revealed it to be 60% sequence identical to mammalian COX [75]. The following year, a UCLA team expressed this inducible enzyme in COS cells and showed it to have cyclooxygenase activity [76]. Two COXs (now named COX-1 and COX-2) with differing expression profiles were firmly identified [77] and it was subsequently reported that a COX-2 selective inhibitor reduced inflammation without reducing prostaglandin production in the stomach [78]. These data, in conjunction with a large number of other supporting studies, established and validated the COX-2 inhibitory hypothesis. Namely, whereas COX-1 is constitutively expressed ubiquitously and most strongly associated with prostaglandin production in gastric mucosa and thromboxane production in platelets, COX-2 is expressed primarily in the CNS and at sites of inflammation. Thus, the development of COX-2 selective inhibitors should afford a therapeutic strategy for chronic inflammation treatment, without the gastric toxicity observed by non-selective inhibitors [79, 80].

1.4.6. Methods for Evaluation of anti – inflammatory activity [81, 82]

Both in-vivo and in-vitro methods are available for the evaluation of anti-inflammatory agents but among them in-vivo method the carrageenan induced rat paw oedema assay is believed to be one of the most reliable and also the most widely used. Carrageenan is a mixture of polysaccharides composed of sulfated galactose units and is derived from Irish
Sea moss, *Chondrus crispus*. The oedema, which develops in rat paw after carrageenan injection, is a biphasic event. The initial phase is attributed to the release of histamine and serotonin, the oedema maintained between the 1\textsuperscript{st} and 2\textsuperscript{nd} phase to kinin like substances and the 2\textsuperscript{nd} phase to prostaglandins like compound.

The carrageenan assay method is advantageous because,

1. The oedema is specifically inhibited by anti-inflammatory compounds.
2. Single oral dose of drugs at non-toxic levels are effective.
3. Low variability.
5. Carrageenan itself is neither antigenic nor causes any systemic effects.

\textbf{1.4.6.1. In vivo animal models: [83]}

The inflammatory response is accompanied by clinical signs of erythema, edema, hyperalgesia and pain. Inflammatory responses occur in three different phases, each apparently mediated by different mechanisms:

1. An acute transient phase characterized by local vasodilation and increased capillary permeability.
2. A sub acute phase, characterized by infiltration of leukocytes and phagocytic cells.
3. A chronic proliferative phase, in which tissue degeneration and fibrosis occurs.
1.4.6.2. Methods for testing acute and sub acute inflammation are:

• UV-erythema in guinea pigs

• Vascular permeability

• Oxazolone-induced ear edema in mice

• Croton-oil ear edema in rats and mice

• Paw edema in rats (various modifications and various irritants)

• Pleurisy tests

• Granuloma pouch technique (various modifications and various irritants)

The proliferative phase is measured by methods for testing granuloma formation, such as:

• Cotton wool granuloma

• Glass rod granuloma

1. UV-erythema in guinea pigs:

Prostaglandin E (PGE) levels in the skin have been shown to be elevated during the 24 h period following exposure of guinea pig skin to ultraviolet radiation from 280-320 nm. The development of increased PGE levels paralleled the development of the delayed phase of erythema. Delay the development of ultraviolet erythema on albino guinea pig skin by systemic pretreatment with clinically equivalent doses of phenylbutazone and other nonsteroidal anti-inflammatory agents.
Erythema (redness) is the earliest sign of inflammation, not yet accompanied by plasma exudation and edema. This model depicts the delay in development of UV erythema on albino guinea pig skin by systemic pretreatment with clinically equivalent doses of phenylbutazone and other NSAIDs.

**Procedure:** Albino guinea pigs of both sexes with an average weight of 350g are used. Four animals are used each for treatment and control group. 18 hr prior testing, the animals are shaved on both the flanks and on the back. Then they are chemically depilated by a commercial depilation product or by a suspension of barium sulphide. 20 min later, the depilation paste and the fur are rinsed off in running warm water. On the next day the test compound is dissolved in the vehicle and half of the test compound is administered by gavage (at 10 ml/kg) 30 min before UV exposure. Control animals are treated with the vehicle alone. The guinea pigs are placed in a leather cuff with a hole of 1.5×2.5cm size punched in it, allowing the UV radiation to reach only this area. Then animals are exposed to UV radiation. After 2 min of exposure the remaining half of the test compound is administered. The erythema is scored 2 and 4 hr after exposure.

**Evaluation:** The degree of erythema is evaluated visually by 2 different investigators in a double blinded manner. The followings scores are given:

- 0 = no erythema
- 1 = weak erythema
- 2 = strong erythema
- 4 = very strong erythema

Animals with a score of 0 or 1 are considered to be protected. The scoring after 2 and after 4 h gives some indication of the duration of the effect. \( ED_{50} \) values can be calculated. [83][84]
CHAPTER I

2. Vascular permeability: During inflammation, vascular permeability increases to allow plasma constituents such as antibodies and complement to access injured or infected tissues. The test is used to evaluate the inhibitory activity of drugs against increased vascular permeability which is induced by phlogistic substances. Mediators of inflammation, such as histamine, prostaglandins and leukotrienes are released following stimulation e.g. of mast cells. This leads to a dilation of arterioles and venules and to an increased vascular permeability. As a consequence, fluid and plasma proteins are extravasated and edemas are formed. The increase of permeability can be recognized by the infiltration of the injected sites of the skin with the vital dye Evan’s blue.

Procedure: Albino Wistar are used each group containing 4 rats. Control group will receive distilled water 1%w/v 1ml/100g by oral route and other group will receive test compound by oral route and standard group will receive Diclofenac 10ml/kg by intraperitoneal route. After 1h of these administration rats are injected with 0.25ml of 0.6% v/v solution of acetic acid intraperitoneally. Immediately, 10 ml/kg of 10%w/v Evans blue is injected intravenously via tail vain. After 30 min, the animals are anesthetized with ether anaesthesia and sacrificed. The abdomen is cut open and exposed viscera. The animals are held by a flap of abdominal wall over a Petri dish. The peritoneal fluid (exudates) collected, filtered and made up the volume to 10 ml using normal saline solution and centrifuged at 3000 rpm for 15 min. The absorbance (A) of the supernatant is measured at 590 nm using spectrophotometer.

Evaluation: Decreased concentration of dye with respected to absorbance indicates reduction in permeability. The result of test is compared with that of standard. ED50 values can also be calculated. [83][85]
3. Oxazolone-induced ear edema in mice: The oxazolone-induced ear edema model in mice is a model of delayed contact hypersensitivity that permits the quantitative evaluation of the topical and systemic anti-inflammatory activity of a compound following topical administration. The oxazolone-repeated challenge increased the level of Th2 cytokines and decreased that of a Th1 cytokine in the lesioned skin. The Th2 cytokines, especially IL-4, play major roles in the development of dermatitis in the present mouse model.

**Procedure:** Using 12 mice in each group, the same skin site of the right ear was sensitized by a single application of 10 μl (each 5 μl for inner and outer of ear) of 0.5% oxazolone in acetone 7 days before the first challenge (day 0), and 10 μl of 0.5% oxazolone in acetone was repeatedly applied to the sensitized right ear 3 times per week. In the nonsensitized animals, acetone alone was applied to the right ear. The mice are challenged 8 days later again under anesthesia by applying 0.01 ml 2% oxazolone solution to the inside of the right ear (control) or 0.01 ml of oxazolone solution, in which the test compound or the standard is solved. Groups of 10 to 15 animals are treated with the irritant alone or with the solution of the test compound. The left ear remains untreated. The maximum of inflammation occurs 24 h later. At this time the animals are sacrificed under anesthesia and a disc of 8 mm diameter is punched from both sides. The discs are immediately weighed on a balance. The weight difference is an indicator of the inflammatory edema.

**Evaluation:** Average values of the increase of weight are calculated for each treated group and compared statistically with the control group. [83][86]
4. Croton-oil ear edema in rats and mice: Croton oil contains 12-o-tetracanoilphorbol-13-acetate (TPA) and other phorbol esters as main irritant agents. TPA is able to activate protein kinase C (PKC), which activates other enzymatic cascades in turn, such as mitogen activated protein kinases (MAPK), and phospholipase A2 (PLA2), leading to release of platelet activation factor (PAF) and AA.

This cascade of events stimulates vascular permeability, vasodilation, polymorphonuclear leukocytes migration, release of histamine and serotonin and moderate synthesis of inflammatory eicosanoids by cyclooxygenase (COX) and 5-lipoxygenase (5-LOX) enzymes. COX and 5-LOX inhibitors, leukotriene B4 (LTB4) antagonists and corticosteroids show topical anti-inflammatory action in animal models of Croton oil or TPA induced skin inflammation.

**Procedure:** A total of 15μl of an acetonic solution containing 75μg of croton oil is applied to the inner surface of right ear of each mouse. Left ear remains untreated. Control animals receive only the irritant while indomethacin (100μg/ear) serves as reference. Varying dose levels of test drug are applied to the inner surface of right ear of each mouse by dissolving them in inflammation inducing solution. Animals are sacrificed by cervical dislocation 6 hr later and a plug (6 mm in diameter) is removed from both the treated and untreated ear. The difference in weight between the two plugs is taken as measure of edematous response. Since tetradecanoyl porbol acetate (TPA) is the chief ingredient of croton oil, purified TPA has also been used to induce ear edema in mice.

**Evaluation:** The antiphlogistic effect can be determined by expressing the increase in weight of the treated ear as percentage of the weight of the contralateral control ear. The
difference between both ears or excised discs is calculated as the average values for treated and control groups and the effect is evaluated by statistical methods.[83][87][88]

5. **Paw edema in rats:** The ability of anti-inflammatory drugs to inhibit the edema produced in the hind paw of the rat after injection of a phlogistic agent. Many phlogistic agents (irritants) have been used, such as brewer’s yeast, formaldehyde, dextran, egg albumin, kaolin, Aerosil, sulfated polysaccharides like carrageenan or naphthoylheparamine. The volume of the injected paw is measured before and after application of the irritant and the paw volume of the treated animals is compared to the controls. [Plethysmograph]. Carrageenan-induced rat paw edema is associated with three distinct phases. The first phase is early mediated by mast cell degranulation and histamine and serotonin release (1 h), the second phase (60 to 150 min) is characterized by bradykinin release and pain, and further eicosanoid production in the late phase (3-4h). So here the anti-inflammatory effect of the test compound is due to inhibition of which mediator can also be known.

**Procedure:** Male or female Sprague-Dawley rats with a body weight between 100 and 150 g are used. The animals are starved overnight. To insure uniform hydration, the rats receive 5 ml of water by stomach tube (controls) or the test drug dissolved or suspended in the same volume. Thirty minutes later, the rats are challenged by a subcutaneous injection of 0.05 ml of 1% solution of carrageenan into the plantar side of the left hindpaw. The paw is marked with ink at the level of the lateral malleolus and immersed in mercury up to this mark. The paw volume is measured plethysmographically immediately after injection, again 3 and 6 h, and eventually 24 h after challenge. Various devices have been developed for plethysmography of the paw, like mercury for
immersion of the paw, more sophisticated apparatus based on the principle of transforming the volume being increased by immersion of the paw into a proportional voltage using a pressure transducer, sensitive method of measuring mouse paw volume by interfacing a Mettler Delta Range top-loading balance with a microcomputer, commercially available plethysmometer.

**Evaluation:** The increase of paw volume after 3 or 6 h is calculated as percentage compared with the volume measured immediately after injection of the irritant for each animal. Effectively treated animals show much less edema. The difference of average values between treated animals and control groups is calculated for each time interval and statistically evaluated. The difference at the various time intervals gives some hints for the duration of the anti-inflammatory effect. A dose- response curve is run for active drugs and $ED_{50}$ values can be determined. [83][88][89]

6. **Pleurisy tests:** In experimental animals pleurisy can be induced by several irritants, such as histamine, bradykinin, prostaglandins, mast cell degranulators, dextran, enzymes, antigens, microbes, and nonspecific irritants, like turpentine and carrageenan. Carrageenan-induced pleurisy in rats is considered to be an excellent acute inflammatory model in which fluid extravasation, leukocyte migration and the various biochemical parameters involved in the inflammatory response can be measured easily in the exudate.

**Procedure:** The mouse pleurisy was induced by a single intrapleural injection of 0.1 ml of carrageenan (1%). After 4 h the animals were killed with an overdose of ether, the thorax was opened and the pleural cavity was washed with 1.0 ml of sterile PBS, containing heparin (20 IU per ml). Samples of the pleural lavage were collected for
determination of exudation, myeloperoxidase, adenosine-deaminase activities, and nitric oxide levels, as well as for determination of total and differential leukocyte counts. Total leukocyte counts were performed in a Neubauer chamber.

The cytospin preparations of pleural wash were stained with May–Grunwald Giemsa for the differential count which was performed under an oil immersion objective. The serum level of the creative protein was also analyzed. In another set of experiment animals were treated 30 min before carrageenan with a solution of Evans blue dye (25 mg/kg, i.v.) in order to evaluate the degree of exudation in the pleural space. A sample (500 μl) of the fluid leakage collected from the pleural cavity was stored in a freezer (−20 °C) to further determine the concentration of Evans blue dye. To this end, on the day of the experiments, a batch of samples was thawed at room temperature and the amount of dye was estimated by colorimetry using an Elisa plate reader at 600 nm, by interpolation from a standard curve of Evans blue dye in the range of 0.01 to 50 μg/ml.

**Evaluation:** One ml (the added Hank’s solution) is subtracted from the measured volume. The values of each experimental group are averaged and compared with the control group. ED50 values can be calculated using various doses. [83][90]

**7. Granuloma pouch technique:** With the introduction of an irritant substance into an s.c. air pocket, granulation tissue begins to proliferate and soon covers the whole inside of the pouch. This tissue consists of fibroblasts, endothelial cells and an infiltrate of macrophages and polymorphonuclear leukocytes. In the GPA this rapidly growing tissue can be exposed to carcinogenic and mutagenic sub-stances. One of the major advantages of the system is the possibility of bringing the test compounds into direct contact with the
target cells, by injecting them into the air pocket. It is also possible to administer the material by the oral and parenteral routes. It does not provide quantitative information on cytotoxicity of the test compounds in vivo.

**Procedure:** Male or female Sprague-Dawley rats with a body weight between 150 and 200 g are used. Ten animals are taken for controls and for test groups. The back of the animals is shaved and disinfected. With a very thin needle a pneumoderma is made in the middle of the dorsal skin by injection of 20 ml of air under ether anesthesia. Into the resulting oval airpouch 0.5 ml of a 1% solution of Croton oil in sesame oil is injected avoiding any leakage of air. Forty-eight hours later the air is withdrawn from the pouch and 72 h later any resulting adhesions are broken.

Instead of croton oil 1 ml of a 20% suspension of carrageenan in sesame oil can be used as irritant. Starting with the formation of the pouch, the animals are treated every day either orally or subcutaneously with the test compound or the standard. For testing local activity, the test compound is injected directly into the air sac at the same time as the irritant. On the 4th or the 5th day the animals are sacrificed under anesthesia. The pouch is opened and the exudate is collected in glass cylinders. The pouches are washed with 1 ml of saline, exudates are immediately cooled on ice and the volume is recorded. Total no. of leukocytes migrated into the pouch are evaluated after staining with Erythrosine B and remaining exudates is centrifuged at 3000 rpm for 10 min at 4 degrees and supernatant stored at -20 degrees until use.

**Evaluation:** The average value of the exudate of the controls and the test groups is calculated. Comparison is made by statistical means. [83][91][92]
1.5. References


[37] http://www.vcclab.org;


[41] http://research.chem.psu.edu/pcjgroup/ADAPT.html


[44] MolConnZ, Ver. 4.05, (2003), Hall Ass. Consult, Quincy, MA


[59] http://www.oecd.org/document/23/0,2340, _2649_201185_33957015__1_1,00.html


CHAPTER I


[65] Starek, M., Krzek, J., Talanta (2009); 77, 925.


[82] Kulkarni SK. In; Handbookss of Experimental Pharmacology, 2nd ed.Vallabh Prakashan, New Delhi; (2005); 127.


