SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF BIO AND CHEMICAL ENGINEERING

DEPARTMENT OF BIOINFORMATICS

UNIT – 1- SBIA5302 – Computer aided drug design

UNIT I

Drug Discovery

Drug discovery is a multifaceted process, which involves identification of a drug chemical therapeutically useful in treating and management of a disease condition. Typically, researchers find out new drugs through new visions into a disease process that permit investigator to design a medicine to stopover or contrary the effects of the disease.[1] The process of drug discovery includes the identification of drug candidates, synthesis, characterization, screening, and assays for therapeutic efficacy. When a molecule avails its satisfactory results in these investigations, it will commence the process of drug development subsequent to clinical trials. Drug discovery and development is an expensive process due to the high budgets of R&D and clinical trials. It takes almost 12-15 years to develop a single new drug molecule from the time it is discovered when it is available in market for treating patients. The average cost for research and development for each efficacious drug is likely to be $900 million to $2 billion. This figure includes the cost of the thousands of failures: For every 5,000-10,000 compounds that enter the investigation and development pipeline, ultimately only one attains approval. These statistics challenge imagination, but a brief understanding of the R&D process can explain why so many compounds don‘t make it and why it takes such a large, lengthy effort to get one medicine to patients. The Success requires immense resources the best scientific and logical minds, highly sophisticated laboratory and technology; and multifaceted project management. It also takes persistence and good fortune. Eventually, the process of drug discovery brings hope, faith and relief to billions of patients.

Stages of drug discovery and development include:

- Target identification
- Target validation
- lead identification
- lead optimization
- Product characterization
- Formulation and development
- Preclinical research
- Investigational New Drug
- Clinical trials
- New Drug Application
- Approval

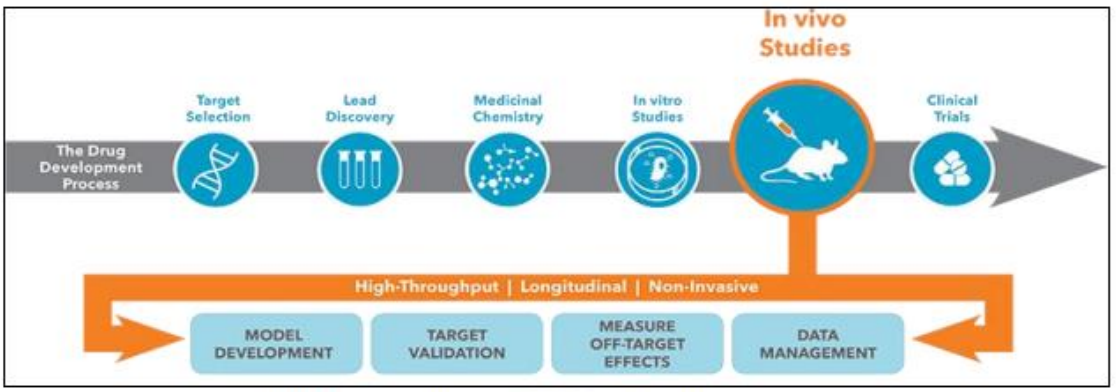


Figure 1: Stages of drug discovery and development process

Target Identification

The first step in the discovery of a drug is identification of the biological origin of a disease, and the potential targets for intervention. Target identification starts with isolating the function of a possible therapeutic target (gene/nucleic acid/protein) and its role in the disease. Identification of the target is followed by characterization of the molecular mechanisms addressed by the target. An ideal target should be efficacious, safe, meet clinical and commercial requirements and be _druggable'. The techniques used for target identification may be based on principles of molecular biology, biochemistry, genetics, biophysics, or other disciplines.

Approaches:

- Data mining using bioinformatics — identifying, selecting and prioritizing potential disease targets
- Genetic association — genetic polymorphism and connection with the disease
- Expression profile — changes in mRNA/protein levels
- Pathway and phenotypic analysis — In vitro cell-based mechanistic studies
- Functional screening — knockdown, knockout or using target specific tools

Target Validation

Target validation is the process by which the expected molecular target – for example gene, protein or nucleic acid of a small molecule is certified. Target validation includes: determining the structure activity relationship (SAR) of analogs of the small molecule; generating a drug-resistant mutant of the presumed target; knockdown or over expression of the presumed target; and monitoring the known signaling systems downstream of the presumed target. Target validation is the process of demonstrating the functional role of the identified target in the disease phenotype. Whilst the validation of a drug's efficacy and toxicity in numerous disease-relevant cell models and animal models is extremely valuable – the ultimate test is whether the drug works in a clinical setting. Target validation can be broken down in to two key steps.

Reproducibility: Once a drug target is identified, whether it be via a specific technique or from review of literature, the first step is to repeat the experiment to confirm that it can be successfully reproduced. The target validation technique includes affinity chromatography, expression-cloning, protein microarray, reverse transfected cell microarray, biochemical suppression, siRNA, DNA microarray, system biology and study of existing drugs.

Introduce variation to the ligand (drug)-target environment

- Genetic manipulation of target genes (in vitro) knocking down the gene (shRNA, siRNA, miRNA), knocking out the gene (CRISPR), knocking in the genes (viral transfection of mutant genes)
- Antibodies interacting to the target with high affinity and blocking further interactions
- Chemical genomics chemical approaches against genome encoding protein[13]

Identification of Lead

A chemical lead is defined as a synthetically stable, feasible, and drug like molecule active in primary and secondary assays with acceptable specificity, affinity and selectivity for

the target receptor. This requires definition of the structure activity relationship as well as determination of synthetic feasibility and preliminary evidence of in vivo efficacy and target engagement. Characteristics of a chemical lead are:

- SAR defined
- Drug ability (preliminary toxicity, hERG)
- Synthetic feasibility
- Select mechanistic assays
- In vitro assessment of drug resistance and efflux potential
- Evidence of in vivo efficacy of chemical class
- PK/Toxicity of chemical class known based on preliminary toxicity or in silico studies

In order to decrease the number of compounds that fail in the drug development process, a drug ability assessment is often conducted. This assessment is important in transforming a compound from a lead molecule into a drug. For a compound to be considered druggable it should have the potential to bind to a specific target; however, also important is the compound's pharmacokinetic profile regarding absorption, distribution, metabolism, and excretion. Other assays will evaluate the potential toxicity of the compound in screens such as the Ames test and cytotoxicity assay. [14]

Lead Optimization

Lead optimization is the process by which a drug candidate is designed after an initial lead compound is identified. The process involves iterative series of synthesis and characterization of a potential drug to build up a representation of in what way chemical structure and activity are related in terms of interactions with its targets and its metabolism.

In initial drug discovery, the resulting leads from hit-to-lead high throughput screening tests undergo lead optimization, to identify promising compounds. Potential leads are evaluated for a range of properties, including selectivity and binding mechanisms during lead optimization, as the final step in early stage drug discovery. The purpose of lead optimization is to maintain favorable properties in lead compounds, while improving on deficiencies in lead structure. In order to produce a pre-clinical drug candidate, the chemical structures of lead compounds (small molecules or biologics) need to be altered to improve target specificity and selectivity. Pharmacodynamic and pharmacokinetic parameters and toxicological properties are also evaluated. Labs must acquire data on the toxicity, efficacy, stability and bioavailability of leads, in order to accurately characterize the compound and establish the route of optimization.[15]

Researchers in drug discovery need rapid methods to narrow down the selection of drug candidates for this downstream selectivity profiling and further investigation. High throughput DMPK (drug metabolism and pharmacokinetics) screens have become an essential part of lead optimization, facilitating the understanding and prediction of in vivo pharmacokinetics using in vitro tests. In order to make new drugs with higher potency and safety profiles, chemical modifications to the structure of candidate drugs are made through optimization.

Automated screening systems are becoming an important part of pharmaceutical and biopharmaceutical drug discovery labs. Mass spectrometry is used for the detection and quantitation of metabolites. MALDI imaging is a key technique for evaluating drug candidates and their metabolites in tissue structure rapidly and accurately. Additionally, NMR Fragment-based Screening (FBS) in the pharmaceutical industry has become a widely applied method for the discovery and optimization of lead molecules in targeted screening campaigns.[16]

Product Characterization

When any new drug molecule shows a promising therapeutic activity, then the molecule is characterized by its size, shape, strength, weakness, use, toxicity, and biological activity. Early stages of pharmacological studies are helpful to characterize the mechanism of action of the compound.

Formulation and Development

Pharmaceutical formulation is a stage of drug development during which the physicochemical properties of active pharmaceutical ingredients (APIs) are characterized to produce a bioavailable, stable and optimal dosage form for a specific administration route.

During preformulation studies the following parameters are evaluated:

- Solubility in different media and solvents
- Dissolution of the active pharmaceutical ingredient (API)
- Accelerated Stability Services under various conditions
- Solid state properties (polymorphs, particle size, particle shape etc.)
- Formulation services and capabilities
- Formulation development of new chemical entities (NCE)
- Optimization of existing formulations
- Process development for selected dosage forms
- Novel formulations for improved delivery of existing dosage forms
- Controlled release and sustained release formulations
- Self-emulsifying drug delivery systems
- Colloidal drug delivery systems
- Sub-micron and nano-emulsions

Preclinical Testing

Pre-clinical research in drug development process involves evaluation of drug's safety and efficacy in animal species that conclude to prospective human outcome. The pre-clinical trials also have to acquire approval by corresponding regulatory authorities. The regulatory authorities must ensure that trials are conducted in safe and ethical way and would give approval for only those drugs which are confirm to be safe and effective. ICH has established a basic guideline for technical necessities of acceptable preclinical drug development.[17]

The pre-clinical trials can be conducted in two ways: General pharmacology and Toxicology. Pharmacology deals with the pharmacokinetic and pharmacodynamic parameters of drug. It is essential to explore unwanted pharmacological effects in suitable animal models and monitoring them in toxicological studies. Pharmacokinetic studies are very important to make known the safety and efficacy parameters in terms of

absorption, distribution, metabolism and excretion. These studies give information on absorption rate for diverse routes of administration, which helps in selection of dosage form, distribution, rate of metabolism and elimination; which governs the half-life of the drug. Half-life of the drug clarifies the safety outline of the drug which is the obligatory for a drug to get approved by regulatory agencies. The drug distribution mechanism elucidates the therapeutic effectiveness of the drug as it depends on the drugs bioavailability and its affinity. Drug metabolism provides the probability of through phases of biotransformation process and formation of drug metabolites. It also helps in understanding the reactions as well as enzymes involved in biotransformation. [18]

Toxicological studies of the drug can be performed by in-vitro and in-vivo test which evaluate the toxicological effects of the drug. In-vitro studies can be performed to inspect the direct effects on cell proliferation and phenotype. In-vivo studies can be performed for qualitative and quantitative determination of toxicological effects. As many drugs are species specific, it is essential to select appropriate animal species for toxicity study. In-vivo studies to evaluate pharmacological and toxicological actions, including mode of action, are often used to support the basis of the proposed use of the product in clinical studies. [19]

The Investigational New Drug Process (IND)

Drug developers must file an Investigational New Drug application to FDA before commencement clinical research.[20] In the IND application, developers must include:

- Preclinical and toxicity study data
- Drug manufacturing information
- Clinical research protocols for studies to be conducted
- Previous clinical research data (if any)
- Information about the investigator/ developer[21]

Clinical Research

Clinical trials are conducted in people (volunteer)and intended to answer specific questions about the safety and efficacy of drugs, vaccines, other therapies, or new methods of using current treatments. Clinical trials follow a specific study protocol that is designed by the researcher or investigator or manufacturer. As the developers design the clinical study, they will consider what they want to complete for each of the different Clinical Research Phases and starts the Investigational New Drug Process (IND), a process they must go through before clinical research begins. Before a clinical trial begins, researchers review prior information about the drug to develop research questions and objectives.[22] Then, they decide:

- Selection criteria for participants
- Number of people take part of the study
- Duration of study
- Dose and route of administration of dosage form
- Assessment of parameters
- Data collection and analysis

Phase 0 clinical trial

Phase 0 implicates investigative, first-in-human (FIH) trials that are conducted according to FDA guidelines. Phase 0 trials besides termed as human micro dose studies, they have single sub-therapeutic doses given to 10 to 15 volunteers and give pharmacokinetic data or help with imaging specific targets without exerting pharmacological actions. Pharmaceutical industries perform Phase 0 studies to pick which of their drug applicants has the preeminent pharmacokinetic parameters in humans.[24]

Phase 1: Safety and dosage

Phase I trials are the first tests of a drug with a lesser number of healthy human volunteers. In most cases, 20 to 80 healthy volunteers with the disease/condition participate in Phase 1. Patients are generally only used if the mechanism of action of a drug indicates that it will not be tolerated in healthy people. However, if a new drug is proposed for use in diabetes patients, researchers conduct Phase 1 trials in patients with that type of diabetes. Phase 1 studies are closely monitored and collect information about Pharmacodynemics in the human body. Researchers adjust dosage regimen based on animal study data to find out what dose of a drug can tolerate the body and what are its acute side effects. As a Phase 1 trial continues, researchers find out research mechanism of action, the side effects accompanying with increase in dosage, and information about effectiveness. This is imperative to the design of Phase 2 studies. Almost 70% of drugs travel to the next phase.

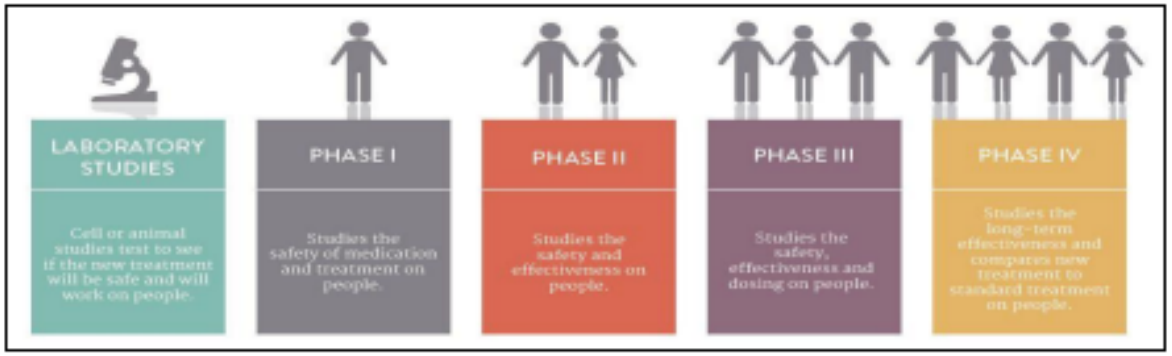


Figure 2: Phases of clinical trials

Phase 2: Efficacy and side effects

Phase II trials are conducted on larger groups of patients (few hundreds) and are aimed to evaluate the efficacy of the drug and to endure the Phase I safety assessments. These trials aren'tsufficient to confirm whether the drug will be therapeutic. Phase 2 studies provide with additional safety data to the researchers. Researchers use these data to refine research questions, develop research methods, and design new Phase 3 research protocols. Around 33% of drugs travel to the next phase.

Most prominently, Phase II clinical studies aid to found therapeutic doses for the large-scale Phase III studies.

Phase 3: Efficacy and adverse drug reactions monitoring

Researchers plan Phase 3 studies to prove whether a product deals anaction benefit to a specific peopleor not. Sometimes known as pivotal studies, these studies comprise 300 to 3,000 volunteers. Phase 3 studies deliver most of the safety data. Thepreviously study might not able to detect less common side effects.Butphase 3 studies are conducted on large no. of volunteers and longer in duration, the results are more probable to detect long-term or uncommon side effects. Around 25-30% of drugs travel to the next phase of clinical research.

If a drug developer has data from its previous tests, preclinical and clinical trials that a drug is safe and effective for its intended use, then the industry can file an application to market the medicine. The FDA review team comprehensivelyinspects all submitted data on the drug and makes a conclusion to approve or not to approve it.[25]

New Drug Application

A New Drug Application (NDA) expresses the full story of a drug molecule. Its purpose is to verify that a drug is safe and effective for its proposed use in the people studied. A drug developer must include all about a drug starting from preclinical data to Phase 3 trial datain the NDA. Developers must include reports on all studies, data, and analysis.[26]Beside with clinical trial outcomes, developers must include:

- Proposed labeling
- Safety updates
- Drug abuse information
- Patent information
- Institutional review board compliance information
- Directions for use

FDA Review

Once FDA obtains a complete NDA then FDA team of review may require about 6 to 10 months to take a pronouncement on whether to approve the NDA. If Once FDA obtains a incomplete NDA then FDA team of review refuse the NDA.

If FDAgoverns that a drug has been revealed to be safe and effective for its proposed use, it is then essential to work with the developerforupgrade prescribing information. This is denoted as "labeling." Labeling preciselydefines the basis for approval and directionhow to use the drug. Although, remaining issues required to be fixed before the drug to be approved for marketing. In other cases, FDA have need of additional studies. At this situation, the developer can choose whether to continue further developmentor not. If a developer distresses with an FDA decision, there are tools for official appeal.[27]

Phase 4: Post-Market Drug Safety Monitoring

Phase 4 trials are conductedwhen the drug or devicehas been approved by FDA.These trials are also recognized as post-marketing surveillance involving pharmacovigilance and continuing technical support after approval. There are numerous observational strategies and assessmentpatterns used in Phase 4trials to evaluate the efficacy, cost-effectiveness, and safety of an involvement in real-world settings. Phase IV studies may be required by regulatory authorities (e.g. change in labelling, risk management/minimization action plan) or may be undertaken by the sponsoring company for competitive purposes or other reasons. Therefore, the true illustration of a drug's safety essentiallyrequires over the months and even years that mark up a drug'slifespan in the market. FDA reviews reports of complications with prescription and OTC drugs, and can decide to add precautions to the dosage or practice information, as well as other events for more serious adverse

UNIT – 2- SBIA5302 – Computer aided drug design

Drug targets

A **biological target** is anything within a living organism to which some other entity (like an endogenous ligand or a drug) is directed and/or binds, resulting in a change in its behavior or function. Examples of common classes of biological targets are proteins and nucleic acids. The definition is context-dependent, and can refer to the biological target of a pharmacologically active drug compound, the receptor target of a hormone (like insulin), or some other target of an external stimulus. Biological targets are most commonly proteins such as enzymes, ion channels, and receptors.

Mechanism

The external stimulus (*i.e.*, the drug or ligand) physically binds to ("hits") the biological target.[1][2] The interaction between the substance and the target may be:

- noncovalent – A relatively weak interaction between the stimulus and the target where no chemical bond is formed between the two interacting partners and hence the interaction is completely reversible.
- reversible covalent – A chemical reaction occurs between the stimulus and target in which the stimulus becomes chemically bonded to the target, but the reverse reaction also readily occurs in which the bond can be broken.
- irreversible covalent – The stimulus is permanently bound to the target through irreversible chemical bond formation.

Depending on the nature of the stimulus, the following can occur:[3]

- There is no direct change in the biological target, but the binding of the substance prevents other endogenous substances (such as activating hormones) from binding to the target. Depending on the nature of the target, this effect is referred as receptor antagonism, enzyme inhibition, or ion channel blockade.
- A conformational change in the target is induced by the stimulus which results in a change in target function. This change in function can mimic the effect of the endogenous substance in which case the effect is referred to as receptor agonism (or channel or enzyme activation) or be the opposite of the endogenous substance which in the case of receptors is referred to as inverse agonism.

The term "biological target" is frequently used in pharmaceutical research to describe the native protein in the body whose activity is modified by a drug resulting in a specific effect, which may be a desirable therapeutic effect or an unwanted adverse effect. In this context, the biological target is often referred to as a **drug target**. The most common drug targets of currently marketed drugs include:

Drug targets

- proteins
  - G protein-coupled receptors (target of 50% of drugs)[7]
  - enzymes (especially protein kinases, proteases, esterases, and phosphatases)
  - ion channels
    - ligand-gated ion channels
    - voltage-gated ion channels
  - nuclear hormone receptors
  - structural proteins such as tubulin
  - membrane transport proteins
- nucleic acids

Databases containing biological targets information:

- Therapeutic Targets Database (TTD)
- DrugBank
- Binding DB

Membrane proteins as drug targets

Membrane proteins are common proteins that are part of, or interact with, biological membranes. Membrane proteins fall into several broad categories depending on their location. Integral membrane proteins are a permanent part of a cell membrane and can either penetrate the membrane (transmembrane) or associate with one or the other side of a membrane (integral monotopic). Peripheral membrane proteins are transiently associated with the cell membrane.

Membrane proteins are common, and medically important—about a third of all human proteins are membrane proteins, and these are targets for more than half of all drugs.[1] Nonetheless, compared to other classes of proteins, determining membrane protein structures remains a challenge in large part due to the difficulty in establishing experimental conditions that can preserve the correct conformation of the protein in isolation from its native environment.

Function

Membrane proteins perform a variety of functions vital to the survival of organisms:[2]

- Membrane receptor proteins relay signals between the cell's internal and external environments.
- Transport proteins move molecules and ions across the membrane. They can be categorized according to the Transporter Classification database.
- Membrane enzymes may have many activities, such as oxidoreductase, transferase or hydrolase.[3]
- Cell adhesion molecules allow cells to identify each other and interact. For example, proteins involved in immune response

The localization of proteins in membranes can be predicted reliably using hydrophobicity analyses of protein sequences, i.e. the localization of hydrophobic amino acid sequences.

Intergral membrane proteins

Integral membrane proteins are permanently attached to the membrane. Such proteins can be separated from the biological membranes only using detergents, nonpolar solvents, or sometimes denaturing agents. One such example of this type of protein which has not been functionally characterized yet is SMIM23. They can be classified according to their relationship with the bilayer:

- Integral polytopic proteins are transmembrane proteins that span across the membrane more than once. These proteins may have different transmembrane topology.[4][5] These proteins have one of two structural architectures:
  - Helix bundle proteins, which are present in all types of biological membranes;
  - Beta barrel proteins, which are found only in outer membranes of Gram-negative bacteria, and outer membranes of mitochondria and chloroplasts.[6]
- Bitopic proteins are transmembrane proteins that span across the membrane only once. Transmembrane helices from these proteins have significantly different amino acid distributions to transmembrane helices from polytopic proteins.[7]
- Integral monotopic proteins are integral membrane proteins that are attached to only one side of the membrane and do not span the whole way across

Peripheral membrane proteins

Peripheral membrane proteins are temporarily attached either to the lipid bilayer or to integral proteins by a combination of hydrophobic, electrostatic, and other non-covalent interactions. Peripheral proteins dissociate following treatment with a polar reagent, such as a solution with an elevated pH or high salt concentrations.

Integral and peripheral proteins may be post-translationally modified, with added fatty acid, diacylglycerol[8] or prenyl chains, or GPI (glycosylphosphatidylinositol), which may be anchored in the lipid bilayer.

THE ROLE MEMBRANE PROTEINS PLAY AS DRUG TARGETS

The ligands can target membrane proteins to regulate their biological activity, as drugs. The percentage of the es-tablished drugs acting on proteins to achieve therapeutic ef-fects has reached 88%. For rational drug target discov-ery, we should understand the role of MP target in biological pathways. In the review, we surveyed the drug binding tar-gets in the DrugBank [21], if it could be traced from UniProt [28], reviewed protein and annotated 'membrane' (kw-0472), the drug binding protein could be seen as membrane protein drug-target (statistics from April 20, 2018). Conse-quently, for all drugs comprising approved and experi-mented, the proportion of MPs is 42% (Fig. 1A). Whereas for approved drugs binding targets, 54% of them are MPs (Fig. 1B). Protein targets are separated into different classes in dif-ferent kinds of literature or databases according to their phylogenetics, functions and tissue expression etc. Latest researches have shed light on the viewpoint that conven-tional therapeutic drug targets are classified into about 130 protein families [29, 30]. Global academics and experts gen-erally divide those protein targets mainly in four classes: enzymes, transporters, ion channels, and receptors. Some articles hold one standpoint that most drug targets are consti-tuted by GPCRs, ion channels, proteases, kinases or nuclear hormone receptors [31].

Another classification method is DTO, the four sorts were GPCRs, kinases, ion channels and nuclear receptors. In DrugBank, each drug can have one or more targets, enzymes, transporters, and carriers asso-ciated with it. It divided a drug binding target as one or more kind of the four types, target, enzyme, transporter, carrier, respectively. After statistics, demonstrates the num-ber of nonredundant approved drug binding proteins (DBPs) and MPs (only human proteins) in four types respectively. In (Fig. 3), approved drug binding targets have been classified according to the DrugBank that are enzyme, transporter, tar-get, carrier and their Venn diagram. (Fig. 3) clearly shows that a protein belongs to one or more categories, which means that the data cannot be used directly because of the redundancy. Overwhelming evidence now implicates that GPCRs, ion channels and transporters are common MP tar-gets, all of which play crucial roles in biochemical pathways and have an extraordinary tendency on diseases treatment and to be drug targets [1]. The introduction of GPCRs, ion channels and transporters

are as follows. 2.1. G Protein-coupled Receptors (GPCRs) GPCRs also called 7TMs has the similar architectural characteristic referring to a domain in which the protein passes through the cell membrane seven times alternatively from the extracellular side to intracellular and back [33]. GPCRs have been the mainstay biological target for pharmacological therapy because they are the primary surface receptors to delivery signal and little substance. Therefore, GPCRs are associated with many diseases, and there has been a large number of researches focusing on the GPCR mechanisms and their biological activities, such as interacting with drugs. Hauser et al. make statistics and manually curate drugs in Clinical Trials database and other sources, resulting in that 475 drugs about 34% of the drugs that the US FDA approved act at GPCRs [34]. In 2008, twelve of the top-15 prescription drugs and top-15 generic drugs targeted GPCRs [35, 36]. The above evidence indicates the importance of GPCRs.

Structure and function of receptors

• Globular proteins acting as a cell Globular proteins acting as a cell's 'letter boxes etter boxes'

• Located mostly in the cell membrane Located mostly in the cell membrane

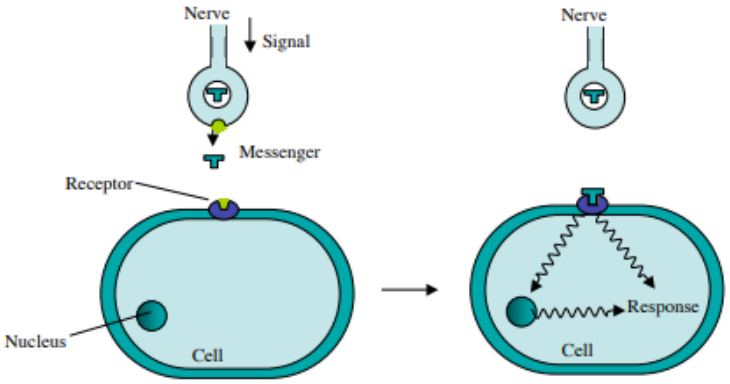
• Receive messages from chemical messengers coming from other Receive messages from chemical messengers coming from other cells

• Transmit a message into the cell leading to a cellular effect Transmit a message into the cell leading to a cellular effect

• Different receptors specific for different chemical messengers Different receptors specific for different chemical messengers

• Each cell has a range of receptors in the cell membrane making it Each cell has a range of receptors in the cell membrane making it responsive to different chemical messengers responsive to different chemical messengers

1. Structure and function of receptors



1. Structure and function of receptors

Chemical Messengers

Neurotransmitters: Chemicals released from nerve endings which travel across a nerve synapse to bind with receptors on target cells, such as muscle cells or another nerve. Usually short lived and responsible for messages between individual cells

Hormones: Chemicals released from cells or glands and which travel some distance to bind with receptors on target cells throughout the body

- Chemical messengers 'switch on' receptors without undergoing a reaction

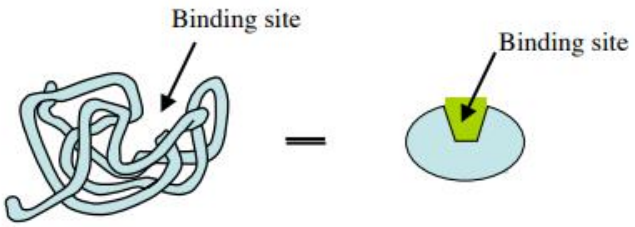
1. Structure and function of receptors

Mechanism

- Receptors contain a binding site (hollow or cleft in the receptor surface) that is recognised by the chemical messenger
- Binding of the messenger involves intermolecular bonds
- Binding results in an induced fit of the receptor protein
- Change in receptor shape results in a 'domino' effect
- Domino effect is known as Signal Transduction, leading to a chemical signal being received inside the cell
- Chemical messenger does not enter the cell. It departs the receptor unchanged and is not permanently bound

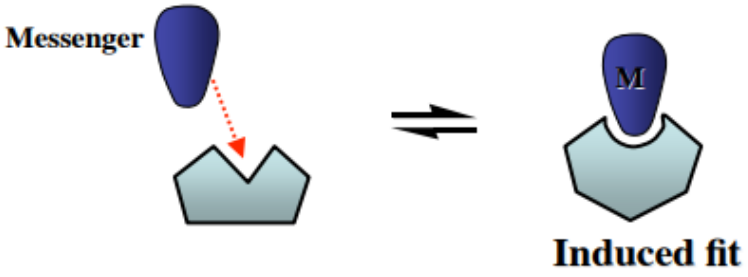
2. The binding site

- A hydrophobic hollow or cleft on the receptor surface - equivalent to the active site of an enzyme
- Accepts and binds a chemical messenger
- Contains amino acids which bind the messenger
- No reaction or catalysis takes place



3. Messenger binding

3.1 Introduction



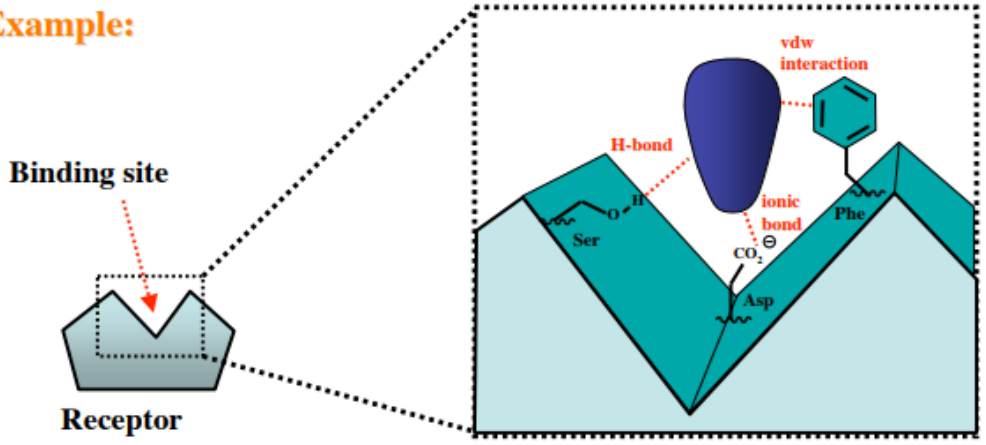
- Binding site is nearly the correct shape for the messenger
- Binding alters the shape of the receptor (induced fit)
- Altered receptor shape leads to further effects - signal transduction

3. Messenger binding

3.2 Bonding forces

- Ionic
- H-bonding
- van der Waals

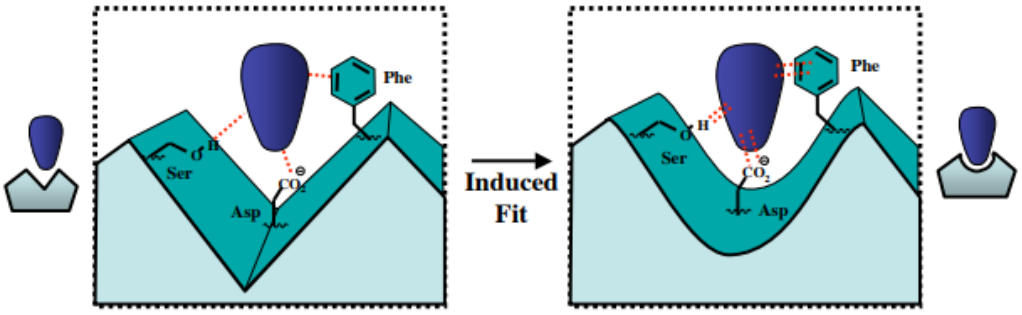
Example:



3. Substrate binding

3.2 Bonding forces

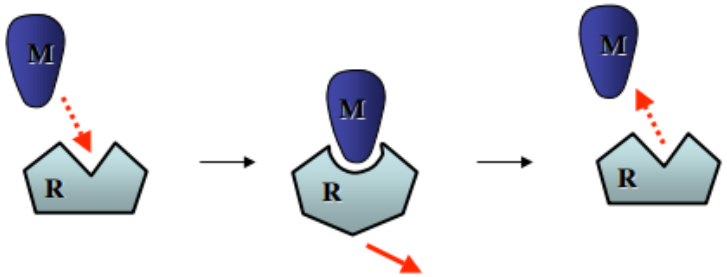
- Induced fit - Binding site alters shape to maximise intermolecular bonding



Intermolecular bonds not optimum length for maximum binding strength

Intermolecular bond lengths optimised

4. Overall process of receptor/messenger interaction



- Binding interactions must be:
  - strong enough to hold the messenger sufficiently long for signal transduction to take place
  - weak enough to allow the messenger to depart
- Implies a fine balance
- Drug design - designing molecules with stronger binding interactions results in drugs that block the binding site - antagonists

1. Receptor superfamilies

		RESPONSE TIME
• ION CHANNEL RECEPTORS	MEMBRANE BOUND	msecs
• G-PROTEIN COUPLED RECEPTORS		seconds
• KINASE LINKED RECEPTORS		minutes
• INTRACELLULAR RECEPTORS		

Receptors, which locate on both the cell surface and within the cell, are drug targets where medicine produce their beneficial effects in various disease states.

Receptors are typically envisaged as cell surface recognition sites for endogenous hormones, neurotransmitters, and neuromodulators. They are coupled to various signal transduction systems located both within the membrane and intracellularly, and can therefore regulate responses to the cellular/tissue microenvironment.

Receptors can be defined in terms of their selectivity, the saturability and reversibility of ligand binding, and functionality. The definition of a receptor in both pharmacological and physiological terms requires that it has specific interactions with ligands that belong to a given pharmacological class.

Receptors are complex proteins with multiple potential ligand recognition sites, including sites that may be distinct from the endogenous agonist recognition site and may actually reside on distinct proteins that are part of the receptor complex.

Such receptor modulatory sites may represent novel drug targets, e.g., allosteric or modulatory sites. The effect of benzodiazepines (BZs) on GABAA receptor function illustrates the conceptualization of ancillary drug targets and the elusive nature of the proposed endogenous modulator, presumed to be a "BZ-like" substance.

In septic shock, the induction of a toxic cytokine receptor-mediated cascade has significantly complicated the search for new drugs to treat this condition. This emphasizes the need to define key targets in critical pathways rather than attempt to treat their sequelae.

It is possible that many diseases are the result of multifactorial events that vary during the pathophysiological course of the illness. For instance, >32 discrete gene loci have been associated with schizophrenia. Therefore, drug targets that are downstream from key points in the disease transduction pathway may not be the optimal targets for treating the disorder.

G protein-coupled receptors

G protein-coupled receptors (GPCRs), which include ~900 members, represent the most leading family of validated **drug targets** in biomedicine. G protein-coupled receptors (GPCRs) have an important role in multiple diseases, including the development of cancer and cancer metastasis, and that's what makes GPCRs perfect drug targets for modern medicinal drugs.

When some GPCRs ligands such as chemokine, thrombin, lysophosphatidic acid (LPA), gastrin-releasing peptide and endothelin bind to their receptors, it causes a comformational change in G protein-coupled receptors (GPCRs) , which are involved in two signal transduction pathways: cAMP signaling pathway and phosphatidylinositol signaling pathway.

Recently, G protein-coupled receptors (GPCRs) have been an uprising star in drug therapy. One example of such GPCRs is chemokine receptor type 4 (CXCR4) which has great potential in drug target research. CXCR4 is thought to be involved in many disease states including more than 23 types of cancer and several immunodeficiency disorders, including head and neck cancer, breast cancer, small-cell lung cancer, non-small-cell lung cancer and Acquired Immune Deficiency Syndrome (AIDS).

G-protein-coupled receptor/GPCR ligands are compounds that bind to a GPCRs. These may be conceptualized as "keys" that fit into a specific "lock" on the cell surface, the latter being the receptor, to modify cellular activity at the biophysical, biochemical, and/or genomic level.

The complex signaling pathways modulated by G-protein-coupled receptors/GPCRs offer a variety of potential drug targets. GPCRs are coupled to various members of the G protein superfamily, so named because of their functional dependence on the hydrolysis of the purine nucleotide, GTP, for activity.

GPCRs in mammals are classified into five main families, named Glutamate, Rhodopsin, Adhesion, Frizzled, and Secretin according to the GRAFS classification.

GPCRs regulate physiological responses to a variety of stimuli that include endogenous ligands such as biogenic amines, peptides, glycoproteins, lipids, nucleotides, Ca2+ ions, and various exogenous ligands for sensory perception such as odorants, pheromones, and even photons.

As a consequence, these receptors mediate multiple physiological processes such as neurotransmission, cellular metabolism, secretion, cellular differentiation, growth, inflammatory, and immune responses.

It is estimated that ~50% of clinically prescribed drugs and 25 of the 100 top selling drugs target GPCRs. Yet, only a small fraction of all GPCRs are presently targeted by drugs.

Receptor Serine/Threonine Kinases (RSTKs

Receptor Serine/Threonine Kinases (RSTKs), respond to specific cytokines, including the transforming growth factor β (TGFβ) and bone morphogenetic protein (BMP) families. With the recent clinical success of drugs targeting protein kinase activity, drug discovery efforts are focusing on the role of reversible protein phosphorylation in disease states.

Activins signal through a combination of type I and II transmembrane serine/threonine kinase receptors. Activin receptors are shared by multiple transforming growth factor-β (TGF-β) ligands such as myostatin, growth and differentiation factor-11 and nodal.

3. DRUGS ACTING ON DNA

3.1 Intercalating agents

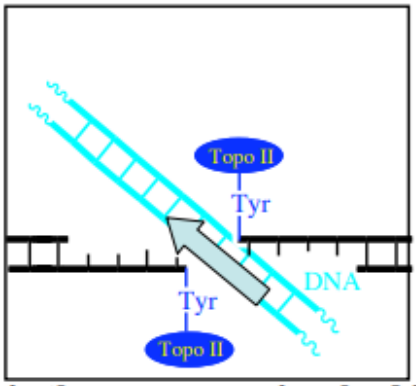
Mechanism of action

- **Contain planar aromatic or heteroaromatic ring systems**
- **Planar systems slip between the layers of nucleic acid pairs and disrupt the shape of the helix**
- **Preference is often shown for the minor or major groove**
- **Intercalation prevents replication and transcription**
- **Intercalation inhibits topoisomerase II- see Doxorubicin, p.198 *Corey*.**

3. DRUGS ACTING ON DNA

Topoisomerase II

- **Relieves the strain in the DNA helix by temporarily cleaving the DNA chain and crossing an intact strand through the broken strand**

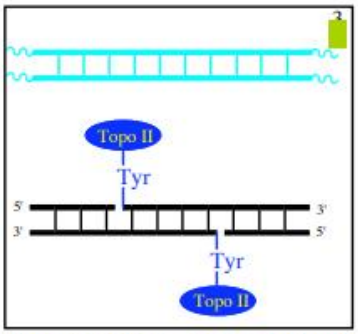


- **Tyrosine residues in the enzyme are involved in the chain breaking process**
- **The residues form covalent bonds to DNA**
- **The enzyme pulls the chains apart to create a gap**
- **The intact strand of DNA is passed through the gap**

3. DRUGS ACTING ON DNA

Topoisomerase II

- **Relieves the strain in the DNA helix by temporarily cleaving the DNA chain and crossing an intact strand through the broken strand**

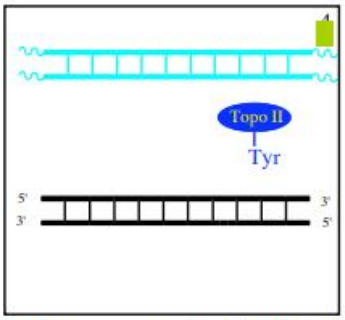


- **Tyrosine residues in the enzyme are involved in the chain breaking process**
- **The residues form covalent bonds to DNA**
- **The enzyme pulls the chains apart to create a gap**
- **The intact strand of DNA is passed through the gap**
- **The break is resealed**

3. DRUGS ACTING ON DNA

Topoisomerase II

- **Relieves the strain in the DNA helix by temporarily cleaving the DNA chain and crossing an intact strand through the broken strand**

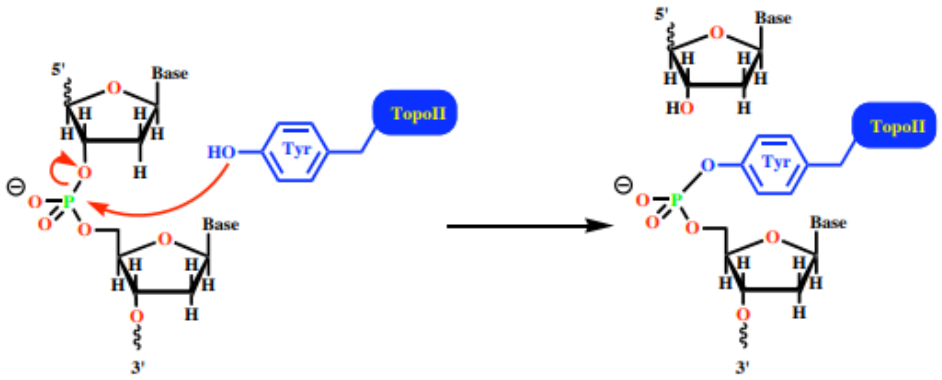


- **Tyrosine residues in the enzyme are involved in the chain breaking process**
- **The residues form covalent bonds to DNA**
- **The enzyme pulls the chains apart to create a gap**
- **The intact strand of DNA is passed through the gap**
- **The break is resealed**

3. DRUGS ACTING ON DNA

Topoisomerase II

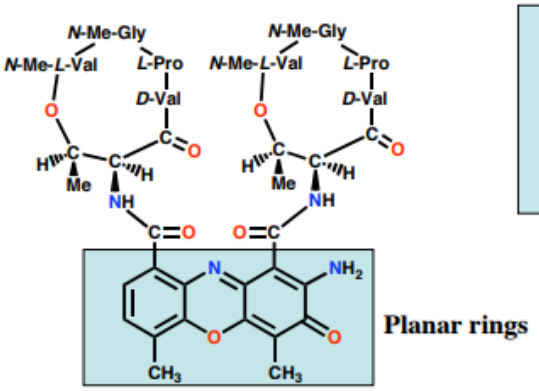
Mechanism of chain cutting



3. DRUGS ACTING ON DNA

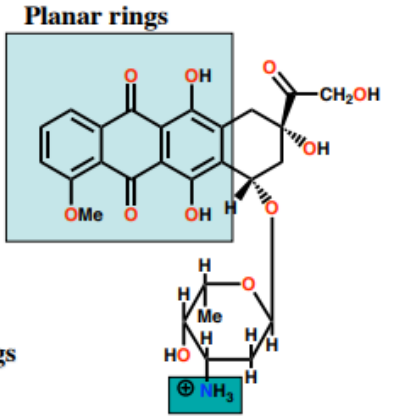
3.1 Intercalating agents

Examples



Dactinomycin

Extra binding to sugar phosphate backbone by cyclic peptide



Doxorubicin (Adriamycin)

Extra binding to sugar phosphate backbone by NH_3

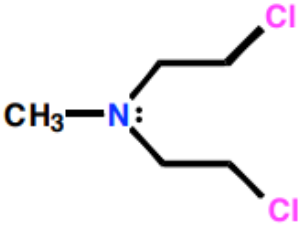
3. DRUGS ACTING ON DNA

3.2 Alkylating agents

- **Contain highly electrophilic groups**
- **Form covalent bonds to nucleophilic groups in DNA (e.g. 7-N of guanine)**
- **Prevent replication and transcription**
- **Useful anti-tumour agents**
- **Toxic side effects (e.g. alkylation of proteins)**

Example

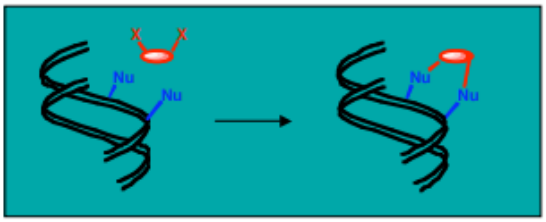
Mechlorethamine (nitrogen mustard)



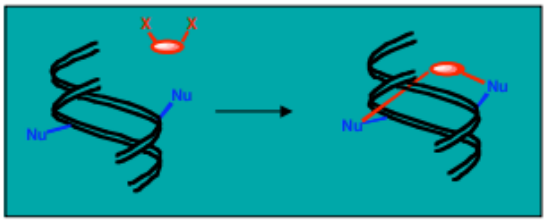
3. DRUGS ACTING ON DNA

3.2 Alkylating agents

Cross linking



Intrastrand cross linking

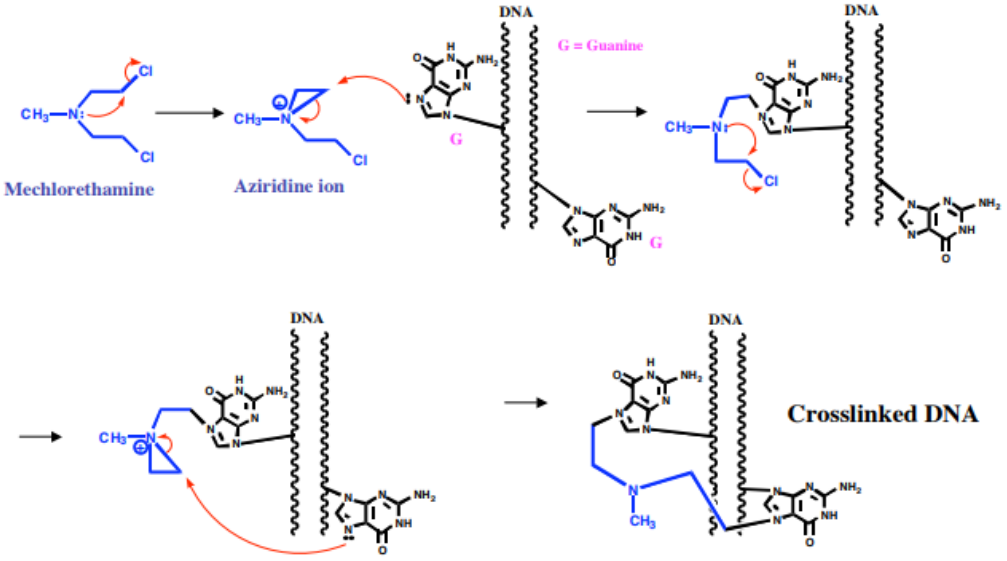


Interstrand cross linking

3. DRUGS ACTING ON DNA

3.2 Alkylating agents

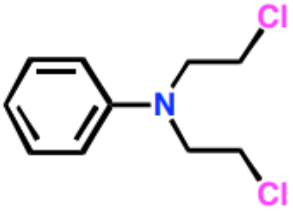
Mechanism of action

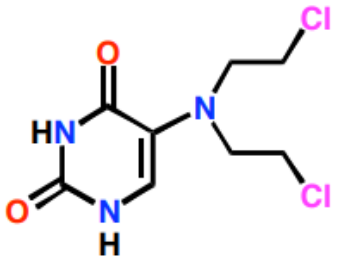


3. DRUGS ACTING ON DNA

3.2 Alkylating agents

Mechlorethamine analogues



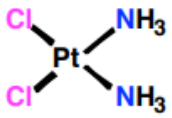


Aromatic ring - e withdrawing effect N is less nucleophilic Less reactive alkylating agent Selective for stronger nucleophiles (e.g. guanine)	Uracil mustard Used vs leukemia Attached to a nucleic acid building block Concentrated in fast growing cells (tumours) Some selectivity

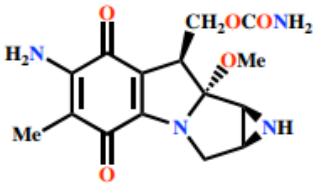
3. DRUGS ACTING ON DNA

3.2 Alkylating agents

Cisplatin



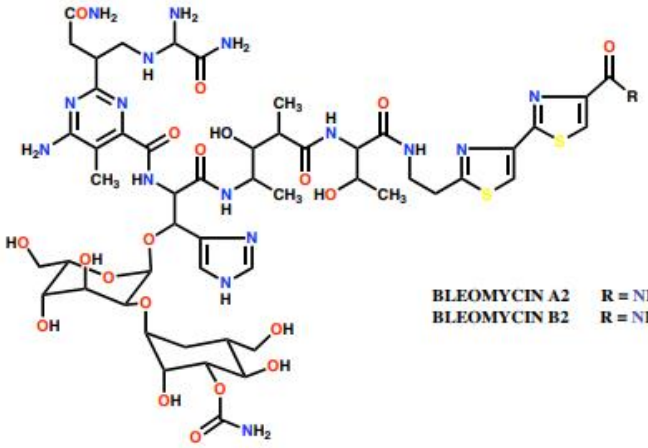
Mitomycin C



Binds to DNA in regions rich in guanine units Intrastrand links rather than interstrand Inhibits transcription	Converted to alkylating agent in the body

3. DRUGS ACTING ON DNA

3.3 Chain cutters



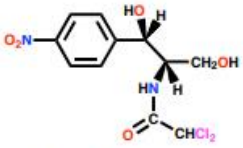
Bleomycin
Used vs skin cancer

BLEOMYCIN A2 R = $NHCH_2CH_2CH_2{}^+SMe_2$
BLEOMYCIN B2 R = $NHCH_2CH_2CH_2CH_2NHC(NH_2)=NH$

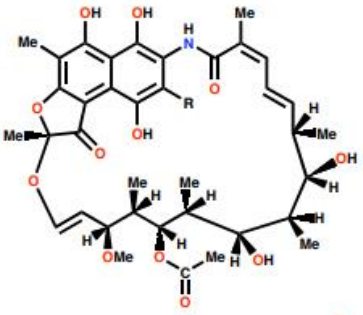
- **Abstracts H from DNA to generate radicals**
- **Radicals react with oxygen resulting in chain cutting**
- **Bleomycin also inhibits repair enzymes**

4. DRUGS ACTING ON rRNA

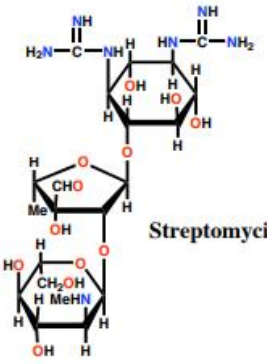
Antibiotics



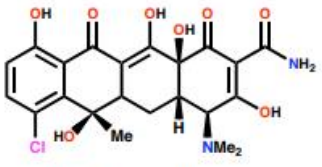
Chloramphenicol
(vs typhoid)



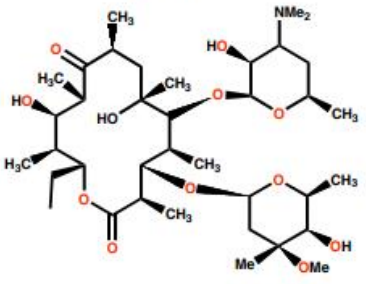
Rifamycins



Streptomycin



Chlortetracycline
(Aureomycin)



Erythromycin

5. DRUGS ACTING ON mRNA

Antisense Therapy

Advantages

- **Same effect as an enzyme inhibitor or receptor antagonist**
- **Highly specific where the oligonucleotide is 17 nucleotides or more**
- **Smaller dose levels required compared to inhibitors or antagonists**
- **Potentially less side effects**

Disadvantages

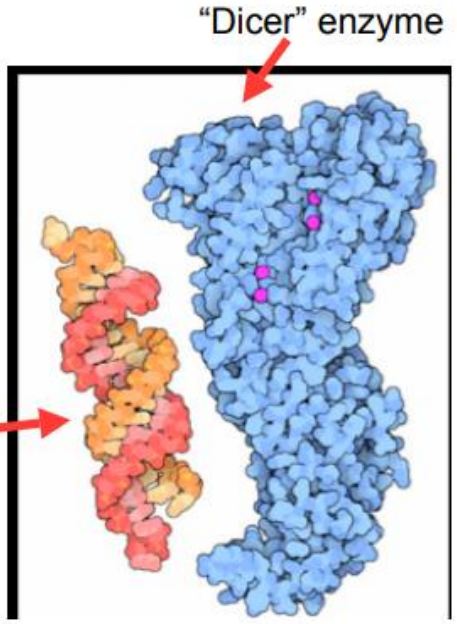
- **'Exposed' sections of mRNA must be targeted**
- **Instability and polarity of oligonucleotides (pharmacokinetics)**
- **Short lifetime of oligonucleotides and poor absorption across cell membranes**

5. DRUGS ACTING ON *or through* RNA

Small Interfering RNA-siRNA

Targets specific RNA sequences for destruction

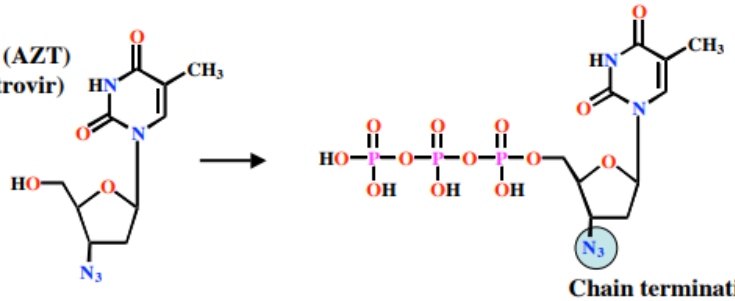
si RNA-
ds, ~21BP, 2 base overhang + phosphate



6. Drugs related to nucleic acid building blocks

Examples: Antiviral agents

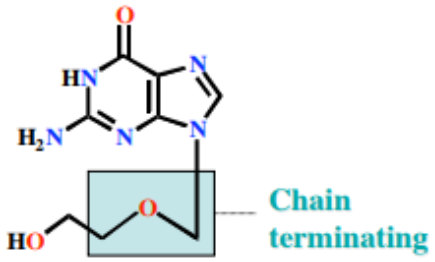
Azidothymidine (AZT)
(Zidovudine;Retrovir)



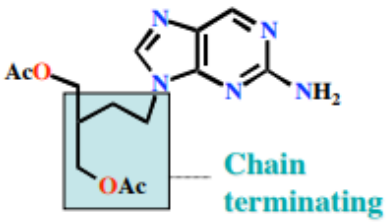
- Enzyme inhibitor
- AZT is phosphorylated to a triphosphate in the body
- Triphosphate has two mechanisms of action
  - inhibits a viral enzyme (reverse transcriptase)
  - added to growing DNA chain and acts as chain terminator

6. Drugs related to nucleic acid building blocks

Examples: Antiviral agents



Acyclovir
(Zovirax)



Famciclovir
(Famvir)

Notes:

Same mechanisms of action as AZT

Used vs herpes simplex and shingles

Medicine in twenty first century has become a science in which drug molecules are directed against the macromolecules. Enzymes hold a prominent position among the biological macromolecules that can be used as drug targets.

Some features of enzyme structures and reaction pathways which make them attractive and primary target for drug action are as follows

- They play an essential role in biological life processes and pathophysiology
- The structures of active sites of enzymes and the ligand binding pockets are highly amenable for high affinity interactions with small drug like molecules
- Presence of potential allosteric sites
- Conformational variations in the binding sites
- Can be directly used as a therapeutic agent

Enzymes are biological catalysts as well as can act as receptors by acting with the substrates. These are the most efficient catalysts known in nature. They have the ability to enhance reaction rates by lowering the activation energy of reactions and by stabilizing the reacting molecules at their activated complex states. Stabilization theory of the activated complex by enzymes was first proposed by Pauling. He concluded that the active site of enzyme is complementary to the structure of the activated complex so that the binding of the enzyme to the activated complex is extremely tight, which reduces the activation energy and enhance the reaction rate.

Structure of Enzymes

• Enzymes are proteins which catalyze the biochemical reactions.

• Some enzymes are inactive alone (apoenzymes) and they show catalytic activity in the presence of another chemical compound (cofactors).

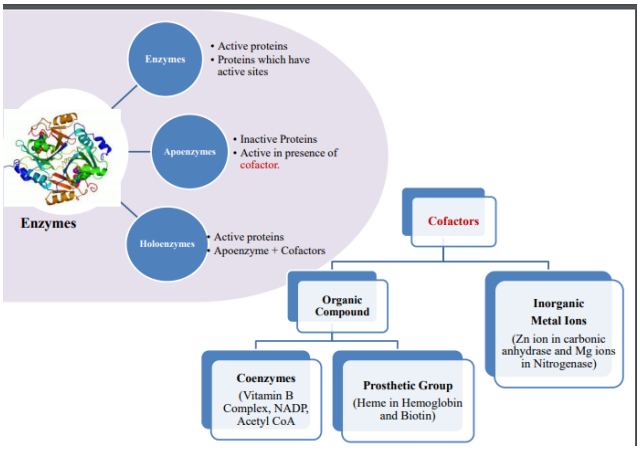
• Many enzymes require the presence of other compounds - cofactors - before their catalytic activity can be exerted. This entire active complex is referred to as the holoenzyme; *i.e.*, apoenzyme (protein portion) plus the cofactor (coenzyme, prosthetic group or metal-ion-activator) is called the holoenzyme.

Apoenzyme + Cofactor = Holoenzyme

According to Holum, the cofactor may be:

1. **Coenzyme** - a non-protein organic substance which is thermostable and loosely attached to the protein part. Examples: NAD^+ or NADP, many Vitamins, Acetyl CoA.
2. **Prosthetic group** - an organic substance which is thermostable and tightly bound to the protein or apoenzyme portion. Examples: heme (Porphyrin ring) in Haemoglobin.
3. **Metal-ion-activator** - these include K^+, Fe^{++}, Fe^{+++}, Cu^{++}, Co^{++}, Zn^{++}, Mn^{++}, Mg^{++}, Ca^{++}, and Mo^{+++}. Examples: Metalloenzymes such as carbonic anhydrase active only in the presence of a Zinc ion.

Apoenzyme (Inactive enzyme) + Cofactor (Organic or Inorganic molecule) = Holoenzyme (Active enzyme)



Classification and Nomenclature of Enzymes

➢ Trivial names of enzymes like Trypsin, Chemotrypsin and Pepsin gives no information about the source, substrate and reaction catalyzed.

➢ For systematic name: According to the International union of Biochemistry an enzyme name has two parts:

-First part is the name of the substrates for the enzyme.

-Second part is the type of reaction catalyzed by the enzyme.This part ends with the suffix "ase".

Example: Lactate dehydrogenase

Classification: Enzymes are classified into six different groups according to the reaction being catalyzed. The nomenclature was determined by the Enzyme Commission in 1961 (with the latest update having occurred in 1992), hence all enzymes are assigned an "EC" number. The classification does not take into account amino acid sequence, protein structure, or chemical mechanism.

EC Number: EC numbers are four digits, for example a.b.c.d, where "a" is the class, "b" is the subclass, "c" is the sub-subclass, and "d" is the sub-sub-subclass. The "b" and "c" digits describe the reaction, while the "d" digit is used to distinguish between different enzymes of the same function based on the actual substrate in the reaction.

Example: for Alcohol: NAD^+ oxidoreductase EC number is 1.1.1.1

Enzymes are divided into six major classes with several sub-classes

•**EC 1. Oxidoreductases:** Enzymes which catalyze oxidation-reduction reactions. *E.g.*, Oxidases, Oxigenases, dehydrogenases and peroxidases.

•**EC 2. Transferases:** Transfer of functional groups (amino, phosphate *etc.*). *E.g.*, methyltransferase, aminotransferase, kinase and phosphorylase.

•**EC 3. Hydrolases:** Hydrolysis of substrate. *E.g.*, Protease, Phosphatase, Phosphodiesterase, Lipase, Maltase and Sucrase.

•**EC 4. Lyases:** Add or remove molecules (ammonia, water, Carbon dioxide) to or from double bond, *i.e.*, Addition-Elimination reaction. *E.g.*, Decarboxylase, aldolase and synthase.

•**EC 5. Isomerases:** Catalyze rearrangement of atoms within a molecule. Isomerization reaction. *E.g.*, Racemase and Mutase.

•**EC 6. Ligases:** Joining of two molecules. Coupling reactions. *E.g.*, Synthetase, carboxylase.

Mechanism of Action

➢ Enzyme speed up chemical reactions in living organisms by decreasing the energy needed to start the reaction (activation energy).

➢ Important Terms:

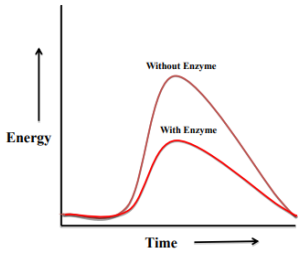
(i) Substrate: Chemical compound that bind to the active site of an enzyme.

(ii) Active Site: area on enzyme where substrate binds.

(iii) Product: what the enzyme produces. Products of enzyme catalyzed reaction.

An enzyme (E) functions by combining with the reactant (termed as *substrate*) to form an activated complex known as *enzyme-substrate* complex. The complex dissociates further to form product (P) and enzyme (E). The enzyme once released from the complex is free to combine with other substrate molecules.

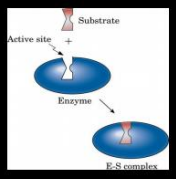
$E + S \longrightarrow E\text{-}S \longrightarrow P + E$



Two types of models for the formation of *Enzyme-Substrate* complex

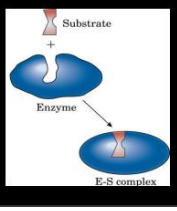
1. Lock and Key Model of enzyme mechanism

   - The enzyme is a rigid 3-D molecule.
   - The enzyme surface contains the active site.
   - Believes that each enzyme is specific to only one molecule (or somewhat identical molecule).



2. Induced-Fit Model of enzyme mechanism

   - The enzyme is flexible 3-D molecule.
   - The active site becomes modified to accommodate the substrate.
   - Enzyme is not rigid and conforms to the substance much like a hand grasping an object



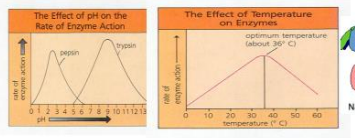
Factors affecting the enzyme activity

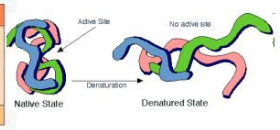
1. **Temperature**- For each enzyme, there is a specific temperature at which the activity of enzyme is a maximum. This is termed as the optimum temperature of enzyme. In general, the enzymes in human system have an optimum temperature in the range 35-40 °C.

2. **pH**- The pH at which an enzyme has maximum activity is termed as its optimum pH. For most of the enzymes, the optimum pH ranges 4-9.

3. **Enzyme concentration**- The rate of enzymatic reactions increases with an increase in the concentration of enzyme.

4. **Concentration of product**- An increased product concentration lowers down the enzymatic reaction. The products formed may inhibit the active site of enzyme or can make the reaction proceed in backward direction.





Specificity of Enzyme Action

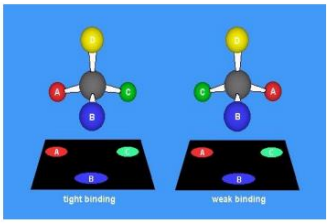
Biocatalysts are different from inorganic catalysts in that these are extremely specific in their catalytic actions. Mostly, each enzyme catalyzes a single reaction, for example, urease attack urea only and carboxypeptidase attack C-terminal peptide bond. Most of the enzymes can catalyze the same reaction with structurally related substrates.

Stereospecificity

Proteins, hormones and receptors are chiral

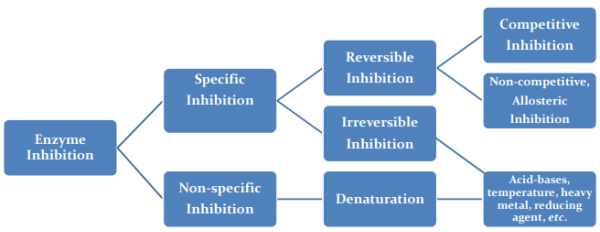
MOLECULES INTERACT WITH BIOMOLECULES STEREO SPECIFICALLY

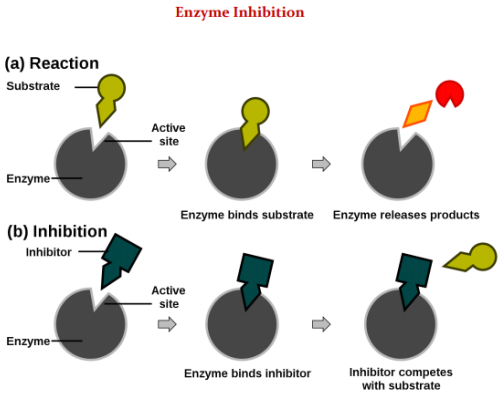
"Nature has a way of knowing how to make things work. Reactions often run in a catalytic mode, and material use, energy, and waste are minimized. Many molecules are chiral, and their unique handedness has both intricate and dramatic influences on how they interact with biological systems."



Enzyme Inhibition

Enzyme Inhibition refers to the halting or reduction of enzyme activity. This is the opposite of enzyme induction, which triggers or increases production. Molecules or factors which reduce the enzyme activity are called **enzyme inhibitors**.

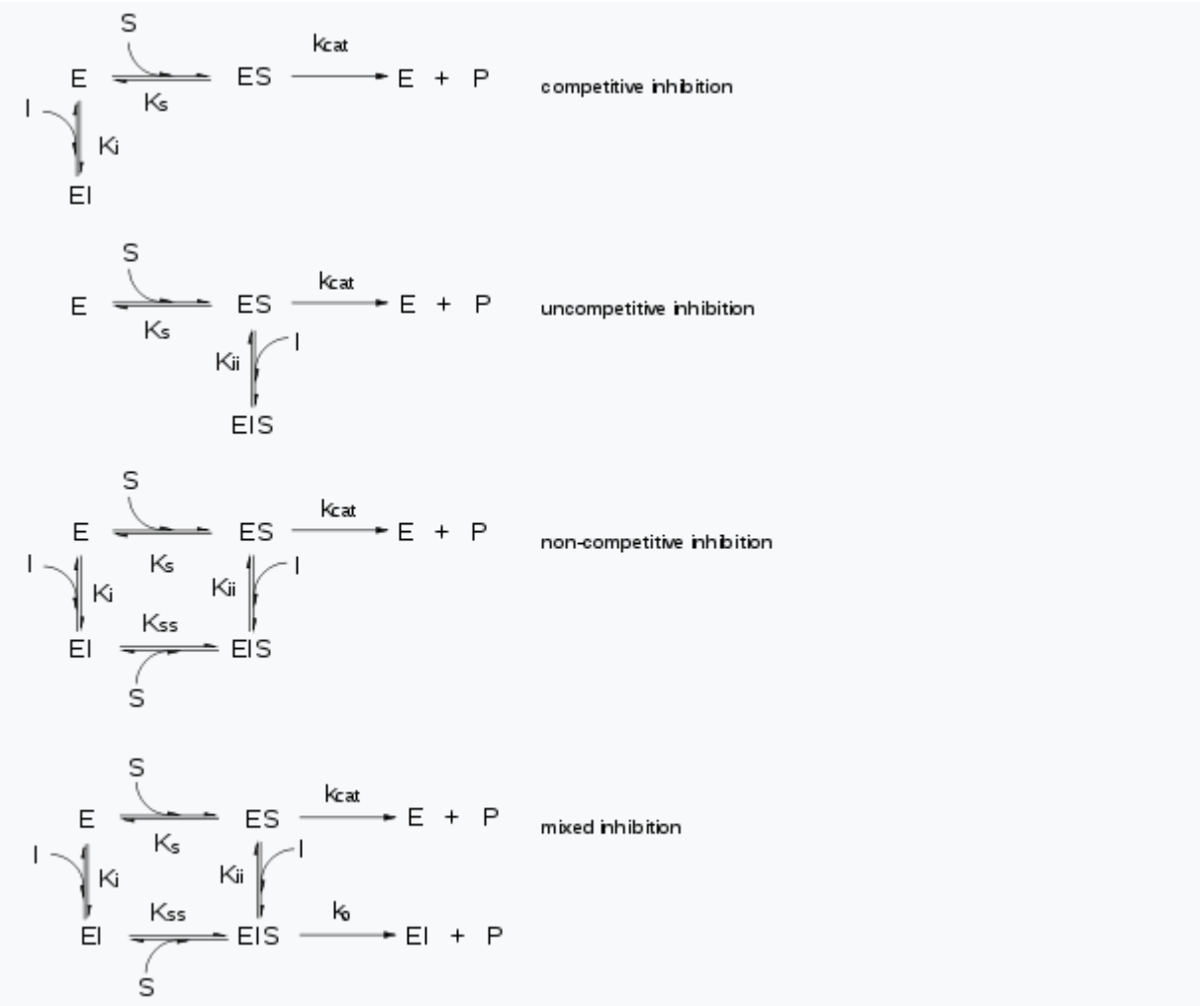




Types of reversible inhibitors

Reversible inhibitors attach to enzymes with non-covalent interactions such as hydrogen bonds, hydrophobic interactions and ionic bonds. Multiple weak bonds between the inhibitor and the active site combine to produce strong and specific binding. In contrast to substrates and irreversible inhibitors, reversible inhibitors generally do not undergo chemical reactions when bound to the enzyme and can be easily removed by dilution or dialysis.

There are four kinds of reversible enzyme inhibitors. They are classified according to the effect of varying the concentration of the enzyme's substrate on the inhibitor.



Types of inhibition. This classification was introduced by W.W. Cleland.[5]

- In **competitive inhibition**, the substrate and inhibitor cannot bind to the enzyme at the same time, as shown in the figure on the right. This usually results from the inhibitor having an affinity for the active site of an enzyme where the substrate also binds; the substrate and inhibitor *compete* for access to the enzyme's active site. This type of inhibition can be overcome by sufficiently high concentrations of substrate (V_{max} remains constant), i.e., by out-competing the inhibitor. However, the apparent K_m will increase as it takes a higher concentration of the substrate to reach the K_m point, or half the V_{max}. Competitive inhibitors are often similar in structure to the real substrate (see examples below).
- In **uncompetitive inhibition**, the inhibitor binds only to the substrate-enzyme complex. This type of inhibition causes V_{max} to decrease (maximum velocity decreases as a result of removing activated complex) and K_m to decrease (due to better binding efficiency as a

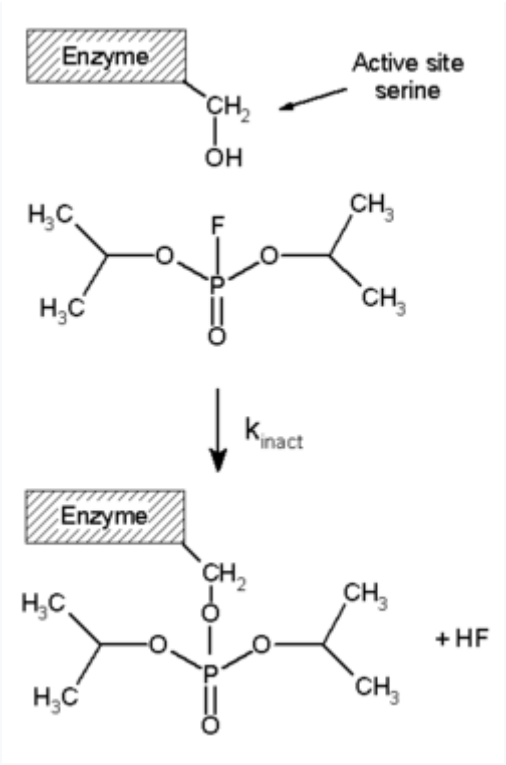
result of Le Chatelier's principle and the effective elimination of the ES complex thus decreasing the K_m which indicates a higher binding affinity).

- In **non-competitive inhibition**, the binding of the inhibitor to the enzyme reduces its activity but does not affect the binding of substrate. As a result, the extent of inhibition depends only on the concentration of the inhibitor. V_{max} will decrease due to the inability for the reaction to proceed as efficiently, but K_m will remain the same as the actual binding of the substrate, by definition, will still function properly.
- In **mixed inhibition**, the inhibitor can bind to the enzyme at the same time as the enzyme's substrate. However, the binding of the inhibitor affects the binding of the substrate, and vice versa. This type of inhibition can be reduced, but not overcome by increasing concentrations of substrate. Although it is possible for mixed-type inhibitors to bind in the active site, this type of inhibition generally results from an allosteric effect where the inhibitor binds to a different site on an enzyme. Inhibitor binding to this allosteric site changes the conformation (i.e., tertiary structure or three-dimensional shape) of the enzyme so that the affinity of the substrate for the active site is reduced.

These types can also be distinguished by the effect of increasing the substrate concentration [S] on the degree of inhibition caused by a given amount of inhibitor. For competitive inhibition the degree of inhibition is reduced by increasing [S], for noncompetitive inhibition the degree of inhibition is unchanged, and for uncompetitive (also called anticompetitive) inhibition the degree of inhibition increases with [S]

Irreversible inhibitors

Types of irreversible inhibition (covalent inactivation)



Reaction of the irreversible inhibitor diisopropylfluorophosphate (DFP) with a serine protease

Irreversible inhibitors usually covalently modify an enzyme, and inhibition can therefore not be reversed. Irreversible inhibitors often contain reactive functional groups such as nitrogen mustards, aldehydes, haloalkanes, alkenes, Michael acceptors, phenyl sulfonates, or fluorophosphonates. These nucleophilic groups react with amino acid side chains to form covalent adducts. The residues modified are those with side chains containing nucleophiles such as hydroxyl or sulfhydryl groups; these include the amino acids serine (as in DFP, right), cysteine, threonine, or tyrosine.

Irreversible inhibition is different from irreversible enzyme inactivation. Irreversible inhibitors are generally specific for one class of enzyme and do not inactivate all proteins; they do not function by destroying protein structure but by specifically altering the active site of their target. For example, extremes of pH or temperature usually cause denaturation of all protein structure, but this is a non-specific effect. Similarly, some non-specific chemical treatments destroy protein structure: for example, heating in concentrated hydrochloric acid will hydrolyse the peptide bonds holding proteins together, releasing free amino acids.[26]

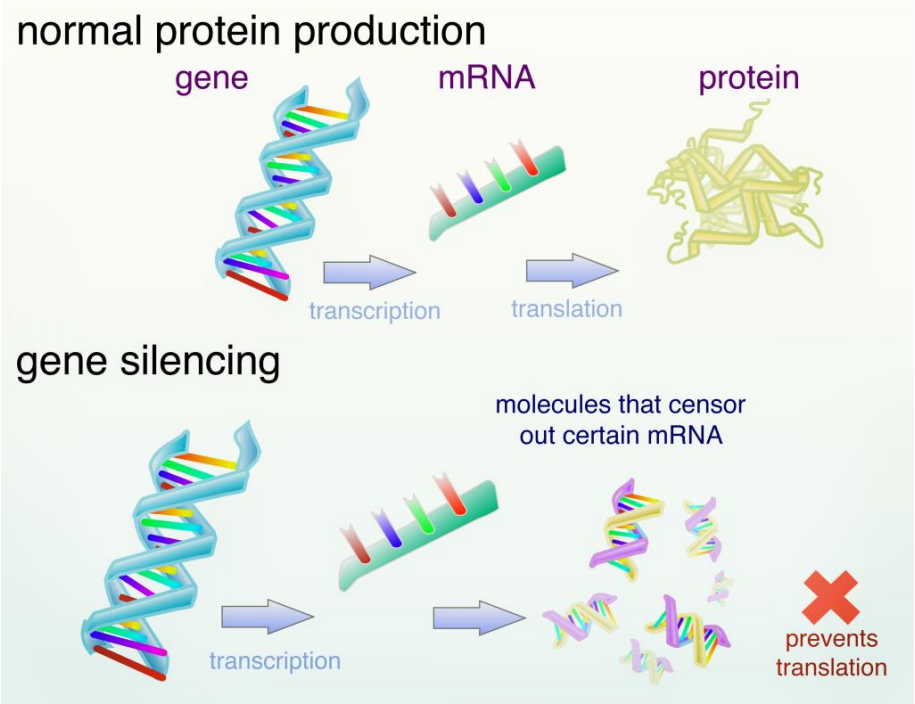
Irreversible inhibitors display time-dependent inhibition and their potency therefore cannot be characterised by an IC_{50} value.[1][27] This is because the amount of active enzyme at a given concentration of irreversible inhibitor will be different depending on how long the inhibitor is pre-incubated with the enzyme. Instead, $k_{obs}/[I]$ values are used,[28] where k_{obs} is the observed

pseudo-first order rate of inactivation (obtained by plotting the log of % activity vs. time) and [*I*] is the concentration of inhibitor. The $k_{obs}/[I]$ parameter is valid as long as the inhibitor does not saturate binding with the enzyme (in which case $k_{obs} = k_{inact}$).

Gene silencing

As the name implies, gene silencing is a technique that aims to reduce or eliminate the production of a protein from its corresponding gene. Genes are sections of DNA that contain the instructions for making proteins. Proteins are essential molecules that perform an array of functions including signaling between cells, speeding up biochemical reactions, and providing structural support for the cell. Each gene is responsible for producing a corresponding protein in a two-step process. First, a copy of the information encoded in a gene is made in the form of messenger RNA (mRNA), a process known as transcription. This occurs in the nucleus of the cell, the cellular structure where all of the cell's genetic material is contained. The mRNA subsequently travels out of the nucleus, and the genetic information it carries is used to produce a specific protein, a process known as translation. (For more information about proteins and how they are made, click here.)

Instead of directly editing DNA or inhibiting the transcription process, the key idea behind gene silencing is intervening in gene expression prior to translation. By designing a molecule that can specifically identify and breakdown the mRNA carrying instructions for making a certain protein, scientists have been able to effectively decrease levels of that protein. Imagine the gene silencing molecule as a censor and mRNA as messages from genes that are broadcast into proteins: the molecule will censor out a specified mRNA message, preventing the corresponding protein from being broadcast into the cell, and thus silencing the gene that is providing these instructions. The ability to significantly lower the levels of a specific protein opens up many possibilities in scientific research and drug development, since proteins are critically involved in the proper function and structure of cells.



Types of Gene Silencing Techniques^

There are various gene silencing methods currently employed in research and being developed as potential disease therapeutics. Nearly all of them involve disabling the function of mRNA by preventing it from being translated into a protein. However, they differ in the design of the molecule used to disrupt mRNA and the manner of mRNA breakdown. As a result, different silencing methods have specific advantages and drawbacks. Two of the leading and most understood methods of gene silencing are RNA interference (RNAi) and antisense oligonucleotides (ASOs).

RNA Interference

In RNAi, the molecules that identify the target mRNA are called small-interfering RNAs (siRNAs). Unlike normal single-stranded RNA found in cells – such as mRNA – siRNAs are short, synthetically made double-stranded RNA molecules designed to pair with a specific mRNA strand. This association of the siRNAs with a particular target mRNA causes the breakdown of the target mRNA by recruiting other proteins that degrade the mRNA target. Because siRNAs are double-stranded, they are more stable and less susceptible to degradation than ASOs, allowing them to continue to perform their silencing function for a longer period of time in the cell. For a more detailed description of how RNAi works, click here.

Antisense Oligonucleotides

Similar to siRNAs, ASOs are engineered by scientists to associate with a target mRNA strand. The binding of the ASO to mRNA directs a protein to breakdown the mRNA. However, unlike siRNAs, ASOs are smaller, single-stranded RNA molecules. As mentioned above, single-stranded RNAs are not as stable as double-stranded ones; thus, ASOs are often chemically modified to increase their durability in a biological environment. However, their smaller size and chemical structure allow ASOs to be transported in cells and living tissues much more effectively than siRNAs. For a more detailed description of how ASOs work, click here.

Is one gene silencing method better than the other?

In terms of developing a drug therapy based on gene silencing, how do RNAi and ASOs compare to each other in effectiveness? In cell culture experiments, gene silencing is often used to intentionally decrease levels of a certain protein for research purposes. In such applications, siRNAs have sometimes been shown to produce stronger and longer lasting gene silencing than ASOs. However, when developing silencing therapeutics, the strength and duration of gene silencing needed for treatment may vary; sometimes a shorter-acting or less complete gene silencing may be required. Furthermore, when considering the efficacy of each method in live animal models, the results are not as clear-cut. For example, as mentioned earlier, ASOs can often be distributed more easily than siRNAs throughout the target tissue because of their size and structure. This observation would be expected to simplify delivery and lower costs of a therapeutic application. The fact that there is no definitive answer to which gene silencing method is more effective has resulted in continued active research and development of both areas.

Antisense technology:

- Antisense technology is a recent approach to specific modification or inhibition of gene expression in vitro or in vivo.
- It is a tool to study gene function and utilize it to manipulate the gene expression within cells to treat an endless number of diseases.
- The antisense approach utilizes **antisense agents** to alter the expression of viral genome inside the host cell or regulate the expression of specific genes that causes that particular disease.

- **Sense strand or sequence**: it is the coding strand within double-stranded DNA that carries the translatable code in the 5′ to 3′ direction. It is complementary to the template strand. The sense strand have sequences similar to that of mRNA.
- **Antisense strand**: it is the template strand of ds DNA, from which mRNA is transcribed. Thus, the antisense strand is complementary to mRNA.
- In simple term 'sense' refers to original sequence of DNA or RNA molecule and 'antisense' refers to Complementary copy.
- The antisense strand base pairs with its complementary mRNA strands (sense RNA) and thus prevents it from being translated into a protein.
- Therefore in antisense technology, the complementary nucleic acid sequence (antisense agents) is utilized to silence gene expression. The binding, or hybridization, of antisense nucleic acid sequences to a specific mRNA target will inhibit normal gene expression (ie. either interrupt transcription or translation) resulting in flow of message from DNA to protein.
- There are several mechanisms to interrupt the gene expression by antisense technology. Sometime the gene expression is completely inhibited known as knock-out and other time it is partially interrupted known as knock-down.
- **Some of the antisense agents are:**
    - **Antisense oligonucleotide**: like oligodeoxyribonucleotides (ODN) having less than 30 mucleotides or longer antisense RNA (a RNA) sequences
        - First generation
        - Second generation
        - **Third generation**
    - **Ribozymes**
    - **RNA interference (RNAi)**

1. Antisense oligonucleotide:

- Zamecnik and Stephenson first demonstrated the antisense effect of synthetic oligonucleotide
- Zamecnik and Stephenson identified a repeated sequence of 21 nucleotides that was crucial to viral integration with the help of nucleotide sequences from the 5' and 3' ends of the 35S RNA of Rous sarcoma virus (RSV).
- They synthesized a 13-mer oligonucleotide, d(AATGGTAAAATGG), complement to the portion of this viral sequence.

- Viral production got inhibited when synthetic oligonucleotide was introduced into cultured fibroblast cells. Thus, they concluded that oligonucleotide was inhibiting viral integration by hybridizing to the crucial sequences and blocking them. They introduced the term 'hybridon' to describe such oligonucleotides.
- At the same time, Tennant et al and Miller et al reported similar effects for synthetic oligonucleotides in other systems.
- **Criteria for successful oligonucleotide**
  - Specific target recognition by Watson-Crick pairing
  - Good structural Mimicry
  - Activation of RNaseH
  - Enhanced cellular uptake
  - Enhanced resistance to various nucleases: Synthetic oligonucleotides are foreign to the cells into which they are introduced and thus becomes prey for endogenous nucleases.
- Synthetic oligonucleotides were protected from endogenous nuclease when they attained the persistence level in cell.
- There are three possible sites on a nucleotide where protective modifications could be introduced.

Antisense Oligonucleotide modification: first, second and third generation

- The three possible sites for oligonucleotide modification are- at the position of Nitrogenous Base, Ribose sugar (2' OH group) and the Phosphate backbone.
- The main purpose of modification is to protect the antisense nucleotide from nuclease degradation when introduced inside cells. And at the same time it should be considered that the modification do not alter the inhibit hybridization ability of the antisense nucleotide.

First generation modification:

- The first generation antisense-motivated nucleotide modification is made by Eckstein and colleagues in the late 1960s by replacing one of the oxygen atom (non-bridging oxygen) of the phosphate backbone with a sulfur atom.
- This modified antisense agent is known as phosphorothioate.
- Phosphorothioate oligonucleotide is more nuclease resistant than the original oligonucleotide. The nuclease resistance was measured by an increased half-life for a phosphorothioated oligonucleotide upto ten hours in human serum as compared to that of one hour of an unmodified oligonucleotide having the same sequence.

- However, the phosphorothionated nucleotide displayed slight reduced hybridization ability and also a tendency to bind un-specifically to certain proteins in cell. High concentration of phosphrothionated nucleotide thus result in cytotoxicity.
- Matsukura and colleagues demonstrated that phosphorothioated oligonucleotides were effective hybridons against the HIV replication in the cultured cells.
- The first FDA-approved antisense drug is **Vitravene** from ISIS (Carlsbad, CA, USA.)

Second generation modification:

- The second generation modification focused on non-specific bind with certain proteins and cytotoxic effect of phosphorothioated nucleotides.
- In this modification, the antisense oligonucleotide undergoes alkyl modifications at the carbon no 2 (C2 position) of the ribose sugar.
- The two most important of these modifications are 2'-O-methyl and 2'-O-methoxy-ethyl at the C2 position.
- After alkylation at the C2 position of ribose sugar, the antisense oligonucleotides become resistant to nuclease degradation and shows low cytotoxicity effect.
- However, the antisense oligonucleotide with 2'-O-alkyl modification are unavailable for RNase H cleavage after hybridization with sense RNA.
- Since RNase H cleavage is the most desirable mechanism for antisense effect, A hybrid oligonucleotide is constructed containing both the desirable characteristics of nuclease resistance and RNase cleavage and it is known as gapmer antisense oligonucleotide.
- **Gapmer antisense oligonucleotide:**
    - This hybrid antisense oligonucleotide contains a central block of deoxynucleotides sufficient to induce RNase H cleavage flanked by blocks of 2'-O-methyl modified nuclease resistance ribonucleotide.

Third generation modification:

- In this modification a.
- Antisense oligonucleotide forms either DNA: DNA homo-duplex or DNA: RNA hetero-duplex depending upon the nature of oligonucleotide.
- The unmodified oligo-deoxynucleotides form such desired DNA: DNA or DNA:RNA duplexes.
- However variety of nucleic acid analogs have been developed having high affinity with target DNA or RNA and as a modification these analogs are utilized for antisense oligonucleotide construction.

- Some of these nucleic acid analogs are peptide nucleic acids (PNAs), 2'-fluoro N3-P5'-phosphoramidites, 1', 5'- anhydrohexitol nucleic acids (HNAs) and locked nucleic acids.
- Among these analogs locked nucleic acid (LNA) is very desirable as shows promising effect. The LNA is composed of nucleotides that is locked into a single conformation through a 2'-0', 4'-C methylene linkage in 1,2:5,6-di-O-isopropylene-α-allofuranose. LNA has increased the thermodynamic stability and enhanced nucleic acid recognition.

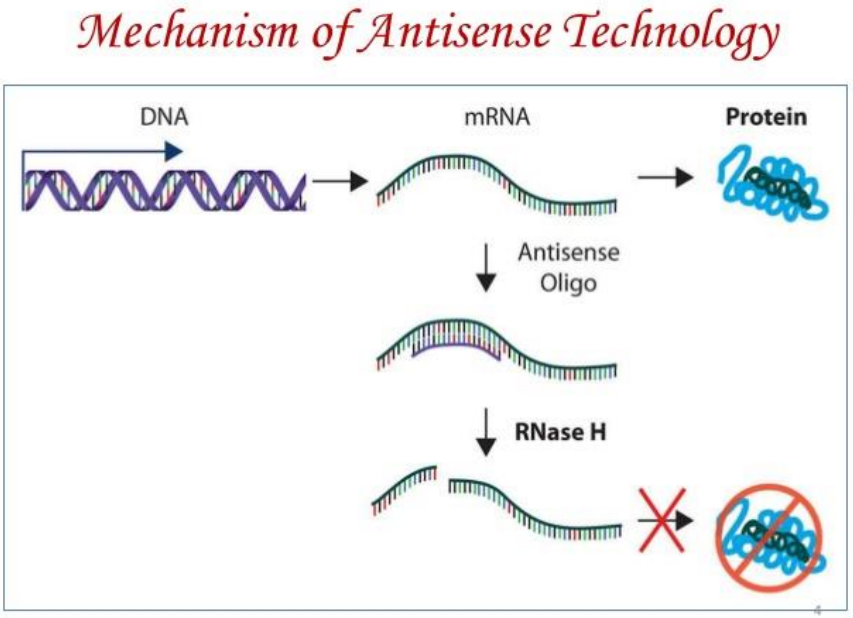
2. Ribozymes:

- Ribozymes are known as catalytic RNA, first described by Tom Cech (1982) from ribosomal RNA precursor from *Tetrahymena thermophilia*.
- Ribozymes acts as an enzymes to processes RNA precursors. And catalyze the modification or alteration of RNA or even DNA.

3. RNA Interference (RNAi):

- RNA interference (RNAi) was first described by Fire and colleagues in *Caenorhabditis elegans.*
- Several types of very short RNAs repress or silence the expression of genes and such silencing is known as RNA interference.
- RNA interference manifest in different ways; some time by inhibiting translation of mRNA and in other case by destruction of mRNA or silencing of promoter.

Antisense mediated gene silencing



- The basic concept in gene silencing utilizing antisense technology is that an antisense oligonucleotide (a short DNA or RNA) is synthesized and introduced into a cell. Since the antisense oligonucleotide is complementary to the targeted mRNA, it will bind forming RNA dimers in the cytoplasm and halts protein synthesis. This occurs because the mRNA no longer has access to the ribosome and cytoplasm. In the same time, it is degraded by nucleases.
- Therefore, the introduction of antisense oligonucleotide results in silencing the expression of gene.
- The events of gene expression is the flow of information from DNA into proteins through transcription (post transcriptional modification) and translation (post-translational modification).
- In the first step DNA is transcribed into pre-mRNA
- Pre-mRNA undergoes post transcriptional modification including- 5' capping, removal of intron (intron excision) and poly-adenylation to form mature functional mRNA.
- Then mRNA is transported to ribosome for translation.
- For silencing gene expression, the antisense oligonucleotide acts on each step of gene expression to achieve antisense knock-down or knock-out of the targeted gene.
- Gene silencing by antisense mechanism includes-Blocking RNA splicing, accelerating degradation of the RNAmolecule, and preventing introns from being spliced out of the mRNA, impeding the exportation of mRNA into the cytoplasm, hindering translation, and the triplex formation in DNA.

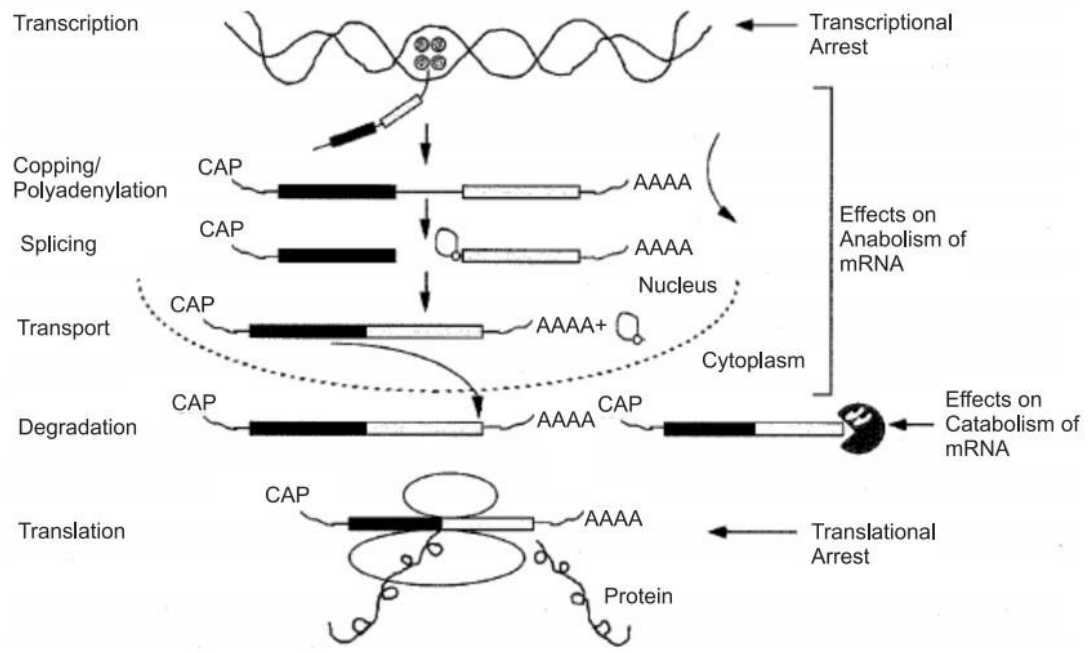


Figure 13.5 Pre-mRNA is transcribed from a gene. It is then processed through a carefully choreographed set of steps to the mature mRNA which is exported to the cytoplasm. These steps include 5' and 3' modifications, (5' cap: 3' poly A) splicing and specific transport activities. Numerous mechanisms of action for antisense drugs have been identified and they affect many steps post transcription and/or enhance cellular degradation of the RNA (Source: Kurreck, 2003).

Silencing transcription:

- In this step, antisense oligonucleotide turn off the gene expression by binding either of the three region- minor groove binding, Strand displaying by nucleic acid analogs, and major groove binding , triplex forming oligonucleotides.
- Pyrrole-imidazole are minor groove binding antisense polypeptides that binds with specific sequence in minor groove.
- Nucleic acid analogs containing antisense oligonucleotide binds with complementary strand of DNA helix, displacing the other strand. This is because the affinity and stability of nucleic acid analog with DNA is more than DNA: DNA duplex.
- Some antisense oligonucleotide are triplex forming and they create stable triplex DNA helix. Triplex forming oligonucleotides binds to duplex through Hoogsteen hydrogen binding: T-A:T and C-G:C triplets.
- **Silencing post-transcription:**
- In this step, antisense oligonucleotides inhibit post-transcriptional modification or RNA splicing.
- Once the mRNA is transcribed from DNA, it undergoes several modification including 5’ capping, polyA tail and intron removal. After post transcriptional modification,

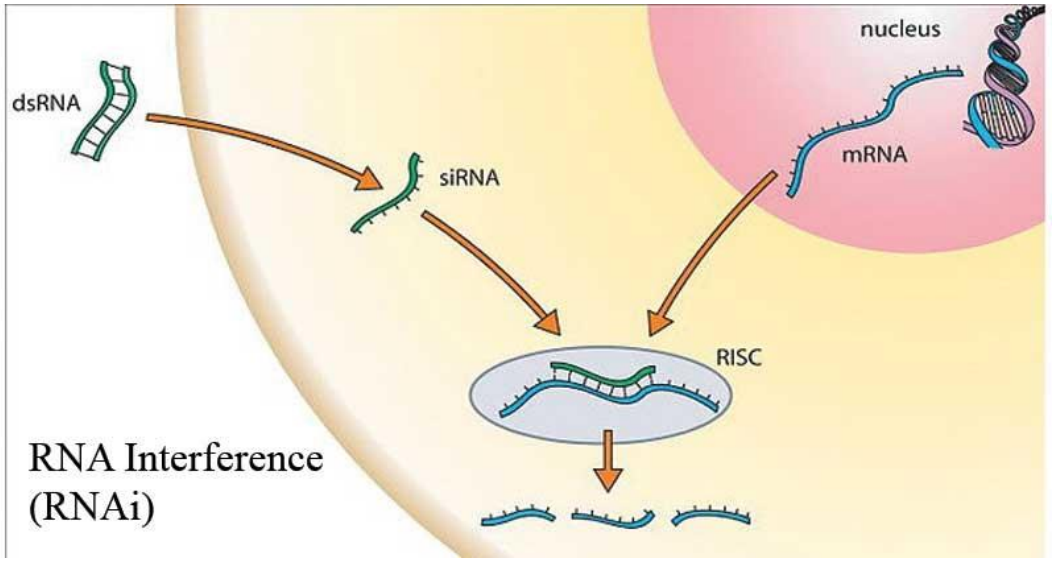
mRNA is transported out from nucleus to cytoplasm. Once the mRNA is hybridized with antisense oligonucleotide, the double stranded RNA duplex cannot be transported out to cytoplasm.

- Removal of intron is essential for RNA splicing because intron are non-coding sequences. In this process, the oligonucleotide-based antisense agent is used which binds with specific sequence of pre-mRNA preventing intron-excision.
- **Silencing translation:**
- In this step, antisense oligonucleotide inhibit translation by complement pairing with targeted mRNA.
- Binding of antisense oligonucleotide with mRNA results in duplex formation which interferes with the transcription apparatus and prevents formation of the ribosomal complex. Thus translation is halted.
- When an antisense strand binds to a mRNA strand, it form a double helix structure which is recognized as faulty and degraded by RNase H preventing the production of undesired protein.
- RNase H degradation mechanism is most used mechanism in gene silencing. RNase H is an endogenous enzyme which cleaves the RNA moiety of an RNA: DNA duplex.
- RNase is present in both nucleus and cytoplasm. It cannot degrade single stranded mRNA but can degrade dsRNA duplex.

RNA interference

- RNA interference is a natural phenomenon by which an mRNA is silenced thereby inhibiting the protein coded from that particular mRNA.
- Scientists have been working on this field from 1990 and finally, two scientists named Andrew Fire and Creg C. Mello shared the Nobel Prize in Physiology or Medicine for their work on RNAi on the nematode worm, ***Caenorhabditis elegans***, which they had published on 1998.
- It has been serving as a very effective tool for the suppression of the desired protein by silencing the required mRNA.
- The basic macro-molecules involved in the RNAi mechanism are two types of RNA molecules: **Micro RNA (miRNA)** and **small interfering RNA (siRNA)**.
- These two types of RNA can bind to specific mRNA through complementary bonding and inhibit the translation by ribosomes.

- The miRNA are short non-coding RNA nucleotides produced from the chromosome inside the nucleus and have hairpin sequences thereby folding it into double strands.
- The pre-miRNA is bound by a complex of protein called microprocessing unit which contains an RNase III enzyme and a dsRNA binding subunit; this **microprocessing unit** takes the pre-miRNA from the nucleus to the cytoplasm through the Nuclear pore complex (NPC) where it is subjected to a protein called **Dicer**, which contains a dsRNA-specific endonuclease subunit, which cleaves some portions of the pre-miRNA and a mature miRNA is produced.
- On the other hand, siRNAs molecules which are exogenous dsRNA nucleotides are similar in sequences to that of miRNA but are not produced inside the nucleus of the cell.
- It comes from external sources. After the processing of miRNA, both miRNA and siRNA contains a passenger strand and a guide strand.
- A complex of protein and RNA called **RNA induced silencing complex (RISC)** is bound to the guide RNA strand and the passenger RNA strand is degraded.
- This guide strand contains sequences that are complementary to a specific mRNA. It binds to the mRNA and the RISC makes cleavages on the mRNA separating the guide-mRNA double strand.
- The remaining mRNA is degraded by cytosolic exonucleases.
- Hence, the protein coded by the mRNA does not come into existence as the mRNA is silenced.



miRNA

A **microRNA** (abbreviated **miRNA**) is a small single-stranded non-coding RNA molecule (containing about 22 nucleotides) found in plants, animals and some viruses, that functions in RNA silencing and post-transcriptional regulation of gene expression.[1] miRNAs function via base-pairing with complementary sequences within mRNA molecules.[2] As a result, these mRNA molecules are silenced, by one or more of the following processes: (1) Cleavage of the mRNA strand into two pieces, (2) Destabilization of the mRNA through shortening of its poly(A) tail, and (3) Less efficient translation of the mRNA into proteins by ribosomes.[2][3]

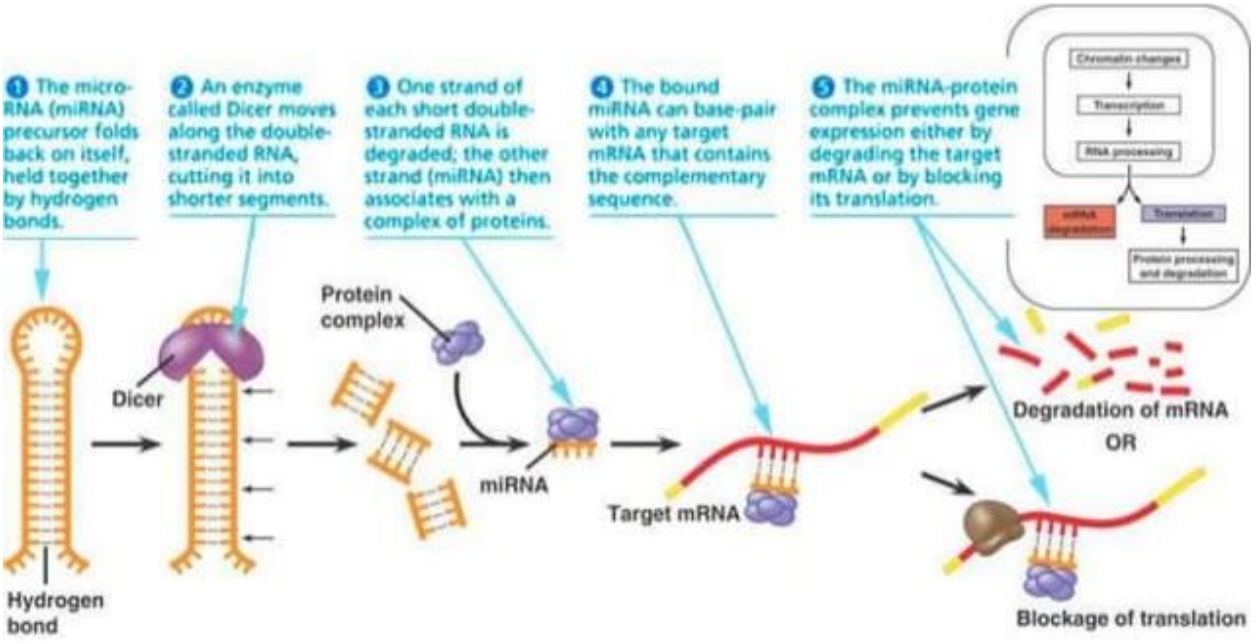
miRNAs resemble the small interfering RNAs (siRNAs) of the RNA interference (RNAi) pathway, except miRNAs derive from regions of RNA transcripts that fold back on themselves to form short hairpins, whereas siRNAs derive from longer regions of double-stranded RNA.[4] The human genome may encode over 1900 miRNAs,[5] although more recent analysis indicates that the number is closer to 600.[6]

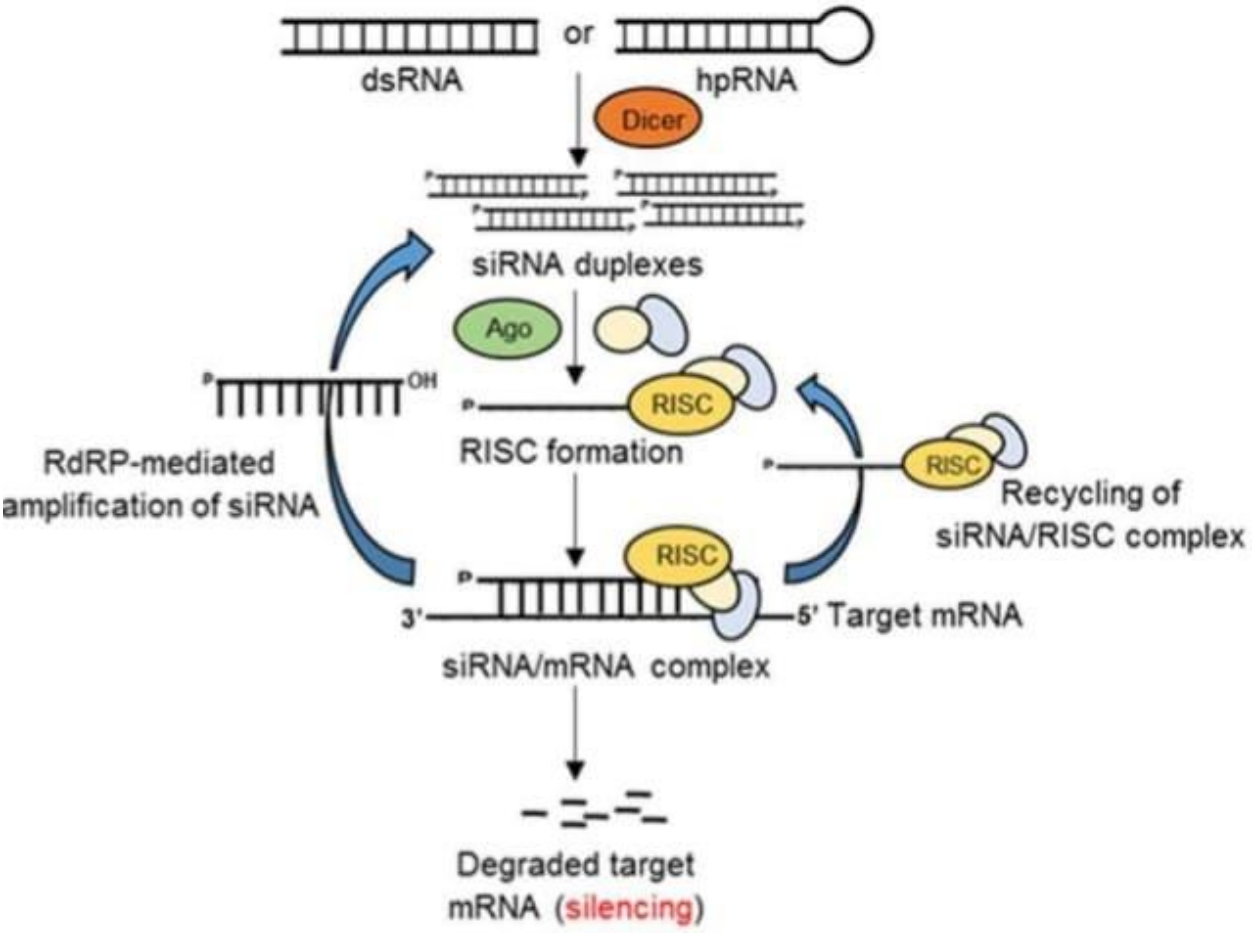
miRNAs are abundant in many mammalian cell types[7][8] and as extracellular **circulating miRNAs**.[9] Circulating miRNAs are released into body fluids including blood and cerebrospinal fluid and have the potential to be available as biomarkers in a number of diseases.[9][10] MiRNAs appear to target about 60% of the genes of humans and other mammals.[11][12] Many miRNAs are evolutionarily conserved, which implies that they have important biological functions.[6][1] For example, 90 families of miRNAs have been conserved since at least the common ancestor of mammals and fish, and most of these conserved miRNAs have important functions, as shown by studies in which genes for one or more members of a family have been knocked out in mice

In interference:-

1. RNAi :-
    1. siRNA :- dsRNA 21-22nt.
    2. miRNA:- ssRNA 19-25nt. Encoded by non- protein coding genome.
2. RISC :-
    1. RNA induced silencing complex, that cleaves mRNA.

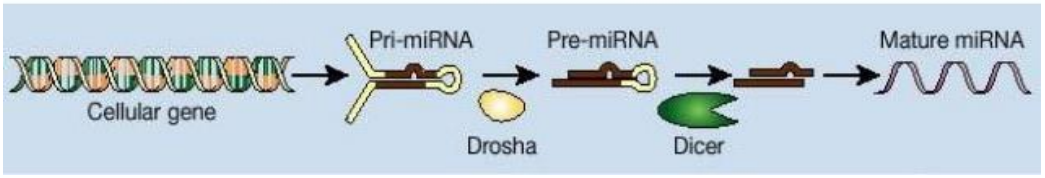
3. ENZYMES:-
    1. DICER:- Produces 20-21nt cleavages that initiate RNAi.
    2. DROSHA:- Cleaves base hairpin in to form pre miRNA; Which is later processed by Dicer.

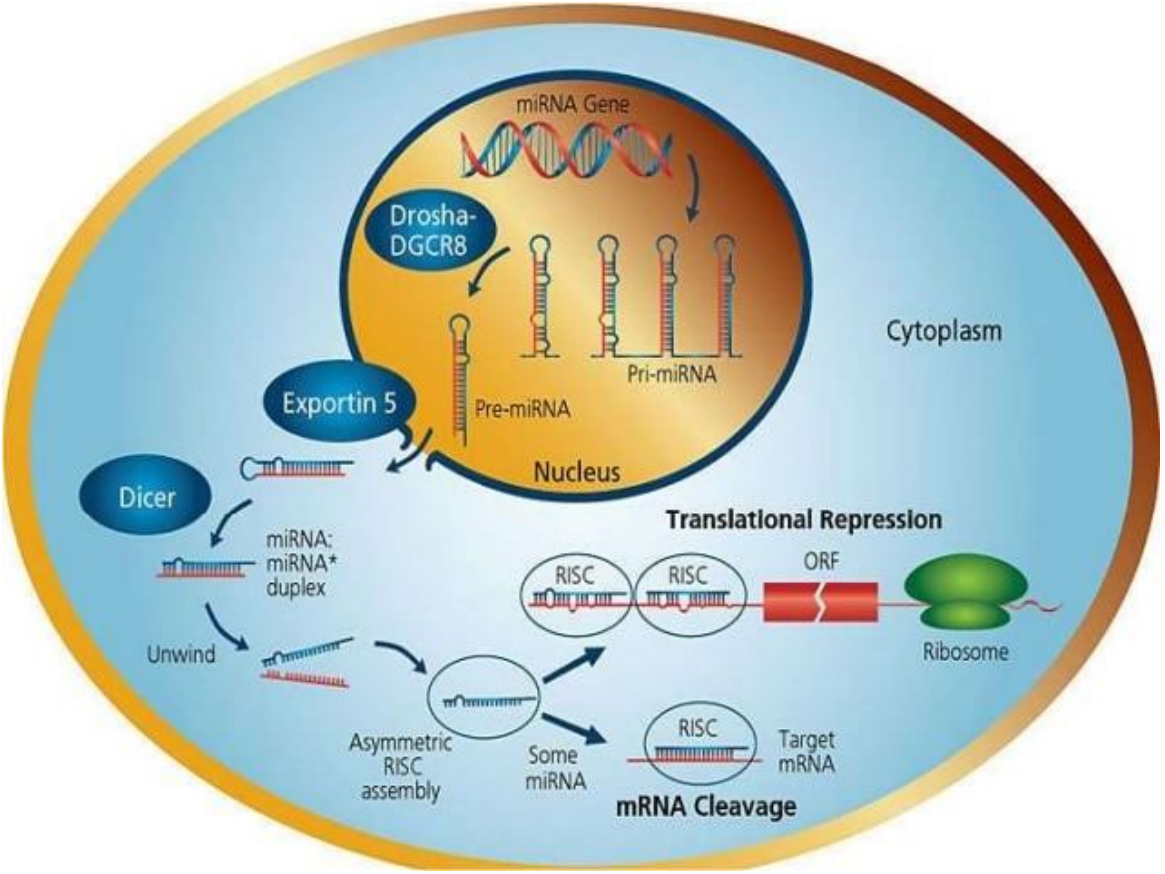

1 The micro-RNA (miRNA) precursor folds back on itself, held together by hydrogen bonds.
2 An enzyme called Dicer moves along the double-stranded RNA, cutting it into shorter segments.
3 One strand of each short double-stranded RNA is degraded; the other strand (miRNA) then associates with a complex of proteins.
4 The bound miRNA can base-pair with any target mRNA that contains the complementary sequence.
5 The miRNA-protein complex prevents gene expression either by degrading the target mRNA or by blocking its translation.
Chromatin changes
Transcription
RNA processing
Translation
Protein processing and degradation
Protein complex
Dicer
miRNA
Target mRNA
Degradation of mRNA
OR
Blockage of translation
Hydrogen bond


dsRNA
or
hpRNA
Dicer
siRNA duplexes
Ago
RISC
RISC formation
OH
RdRP-mediated amplification of siRNA
RISC
Recycling of siRNA/RISC complex
RISC
3'
5' Target mRNA
siRNA/mRNA complex
Degraded target mRNA (silencing)

miRNA

- It is a non coding RNA molecule of approx 21-23 nt. that inhibits the mRNA expression is known as miRNA.
- The formation of micro RNA (miRNA) consists of three important steps:-
  - Formation of primary miRNA.
  - Formation of precursor miRNA.
  - Formation of mature functional miRNA.





siRNA

- also as short interfering RNA or silencing RNA.
- It is a class of double - stranded RNA molecules.
- It is 20-25 bp in length.
- It is similar to miRNA.
- Operating withing the RNA interference RNAi pathway by the enzyme Dicer.
- It interferes with the expression of specific genes with complementary nucleotide sequences by degrading mRNA after transcription, resulting no translation.

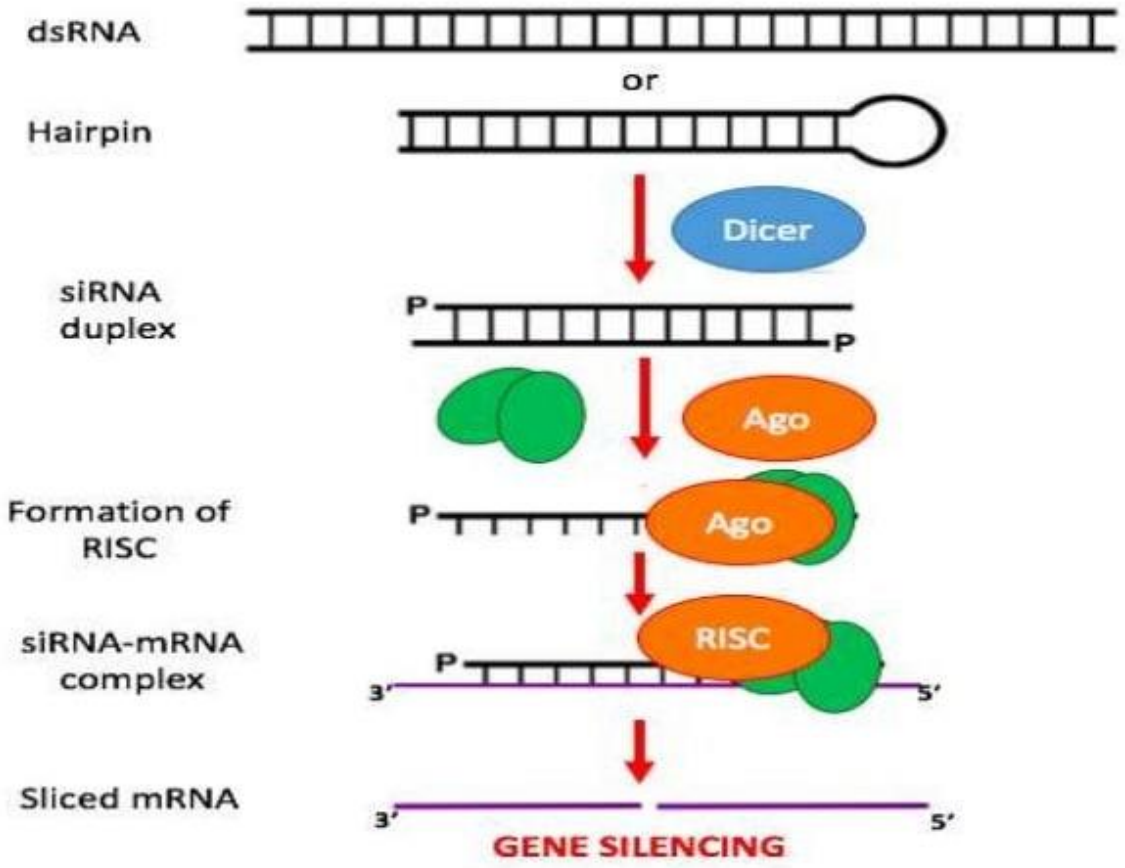


TABLE 20.1. Comparison of small interference RNA (siRNA) and microRNA (miRNA)

S.No.	Particulars	siRNA	miRNA
1.	Function	Gene regulation	Gene regulation
2.	Size in nucleotides	20-25	21-23
3.	Strands	Two complementary	Single
4.	Attachment to mRNA	Coding region	Non-Coding region
5.	Effect on translation	Translation is blocked	Translation is blocked

Property	miRNAs	siRNAs	Ref.
Definition	Regulators of endogenous genes	Defenders of genome integrity in response to foreign or invasive nucleic acids	*29*
Discovery time	1993	1999	*28*
Length	20-22 nt	21-24 nt	*18*
Precursors	Hairpin shaped ssRNAs	Long dsRNAs	*12*
Nature of precursors	Endogenous precursor gene of host's genome	Transposons, transgenes, repeat elements or viruses, *i.e.*, exogenous precursor	*29*
Mode of action	mRNA degradation, translational repression	DNA methylation, histone modification and mRNA degradation	*18*
Argonaute required	AGO1, AGO10	AGO1, AGO4, AGO6, AGO7	*14, 69*
Mechanism of gene regulation	Post-transcriptional only	Transcriptional as well as post-transcriptional	*18*
Complementarity with target sequences	Partially or fully complementary	Fully complementary	*29*
Functions	Cell development and cell differentiation, regulation of development processes, biotic and abiotic stress responses	Defense against transposons and viruses, stress adaptation	*17, 18, 92*

Applications

- Testing hypothesis of gene function.
- Target validation.
- Pathway analysis.
- Gene Redundancy.
- Functional screening.
- siRNA as therapeutics.

Significance of RNAi

- RNAi protects against viral infection.
- RNAi secures genome stability by keeping mobile elements silent.
- RNAi mechanisms repress protein synthesis and regulate the development of organism.
- RNAi offers a new experimental tool to repress gene specifically.
- RNAi might be a useful approach in future gene therapy.

Drug Designing is

1)challenging

2)Expensive

3)Time consuming

So, Multidisciplinary approach:

Computational tools, methodologies for structure guided approach + Global gene expression data analysis by softwares.

Hence,

1) Efficiency increased
2) Cost effectiveness
3) Time saved
4) Strategies to overcome toxic side effects

Mechanism Based Drug Design

- When the disease process is understood at the molecular level and the target molecule(s) are defined
- Drugs can be designed specifically to interact with the target molecule in such a way as to disrupt the disease.

Structure-Based Drug Design

- First techniques to be used in drug design.
- Helped in the discovery process of new drugs.
- Information about the structural dynamics and electronic properties about ligands are obtained from calculations.
- Structure-based drug design can be divided roughly into two categories:

I. Ligand based

II. Receptor Based

Ligand-Based Drug Design

- The first category is about "finding" ligands for a given receptor.
- A large number of potential ligand molecules are screened
- This method is usually referred as ligand-based drug design.
- It saves synthetic effort to obtain new lead compounds.

- **Rational drug design** or simply rational design, is the inventive process of finding new medications based on the knowledge of a biological target
- The drug is most commonly an organic small molecule that activates or inhibits the function of a biomolecule such as a protein, which in turn results in a therapeutic benefit to the patient.
- Drug design involves the design of small molecules that are complementary in shape and charge to the biomolecular target with which they interact and therefore will bind to it.
- Drug design frequently but not necessarily relies on computer modeling techniques.
- This type of modeling is often referred to as **computer-aided drug design**.
- Finally, drug design that relies on the knowledge of the three-dimensional structure of the biomolecular target is known as **structure-based drug design**.
- A more accurate term is ligand design (i.e., design of a small molecule that will bind tightly to its target)

- Although modeling techniques for prediction of binding affinity are reasonably successful, there are many other properties, such as bioavailability, metabolic half-life, side effects, etc., that first must be optimized before a ligand can become a safe and efficacious drug.
- Typically a drug target is a key molecule involved in a particular metabolic or signaling pathway that is specific to a disease condition or pathology or to the infectivity or survival of a microbial pathogen.
- Some approaches attempt to inhibit the functioning of the pathway in the diseased state by causing target molecule to stop functioning.
- Drugs may be designed that bind to the active region and inhibit this target molecule.
- Another approach may be to enhance the normal pathway by promoting specific molecules in the normal pathways that may have been affected in the diseased state.
- All drugs should also be designed so as not to affect any other important "off-target" molecules or antitargets, since drug interactions with off-target molecules may lead to undesirable side effects. Sequence homology is often used to identify such risks.
- Most commonly, drugs are organic small molecules produced through chemical synthesis, but biopolymer-based drugs (also known as biologics) produced through biological processes are becoming increasingly more common. In addition, mRNA-based gene silencing technologies may have therapeutic applications

Types

There are two major types of drug design. The first is referred to as **ligand-based drug design** and the second, **structure-based drug design**

Ligand-based

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

determined biological activity, may be derived. These QSAR relationships in turn may be used to predict the activity of new analogs.

Structure-based

- Structure-based drug design (or **direct drug design**) relies on knowledge of the three dimensional structure of the biological target obtained through methods such as x-ray crystallography or NMR spectroscopy.
- If an experimental structure of a target is not available, it may be possible to create a homology model of the target based on the experimental structure of a related protein.
- Using the structure of the biological target, candidate drugs that are predicted to bind with high affinity and selectivity to the target may be designed using interactive graphics and the intuition of a medicinal chemist. Alternatively various automated computational procedures may be used to suggest new drug candidates.
- Current methods for structure-based drug design can be divided roughly into two categories.
- The first category is about "finding" ligands for a given receptor, which is usually referred as database searching. In this case, a large number of potential ligand molecules are screened to find those fitting the binding pocket of the receptor. This method is usually referred as ligand-based drug design. The key advantage of database searching is that it saves synthetic effort to obtain new lead compounds.
- Another category of structure-based drug design methods is about "building" ligands, which is usually referred as receptor-based drug design. In this case, ligand molecules are built up within the constraints of the binding pocket by assembling small pieces in a stepwise manner. These pieces can be either individual atoms or molecular fragments. The key advantage of such a method is that novel structures, not contained in any database, can be suggested

Active site identification

- Active site identification is the first step in this program.
- It analyzes the protein to find the binding pocket, derives key interaction sites within the binding pocket, and then prepares the necessary data for Ligand fragment link.
- The basic inputs for this step are the 3D structure of the protein and a pre-docked ligand in PDB format, as well as their atomic properties.
- Both ligand and protein atoms need to be classified and their atomic properties should be defined, basically, into four atomic types:
- **hydrophobic atom**: All carbons in hydrocarbon chains or in aromatic groups.

- **H-bond donor**: Oxygen and nitrogen atoms bonded to hydrogen atom(s).
- **H-bond acceptor**: Oxygen and sp^2 or sp hybridized nitrogen atoms with lone electron pair(s).
- **Polar atom**: Oxygen and nitrogen atoms that are neither H-bond donor nor H-bond acceptor, sulfur, phosphorus, halogen, metal, and carbon atoms bonded to hetero-atom(s).
- The space inside the ligand binding region would be studied with virtual probe atoms of the four types above so the chemical environment of all spots in the ligand binding region can be known

Ligand fragment link

- A fragments database can enable drug design. The term "fragment" refers to functional groups or portions of molecules which might have bioactivity.
- Organic molecules can be decomposed into basic chemical fragments. The number of kinds of fragment structures is limited.
- There are a large number of possible fragment combinations. A small perturbation of the previous fragment conformation would cause great difference in activity. In order to find the lowest binding energy on the Potential energy surface (PES) between fragments and a receptor pocket, the scoring function calculation would be performed for every step of conformation change of the fragments derived from every type of possible fragments combination. Since this requires a large amount of computation, using different tricks may use less computing power and let the program work more efficiently.
- When a ligand is inserted into the pocket site of a receptor, groups on the ligand that bind tightly with the receptor should have the highest priority in finding their lowest-energy conformation. This allows us to put several seeds into the program at the same time and optimize the conformation of those seeds that form significant interactions with the receptor, and then connect those seeds into a continuous ligand in a manner that make the rest of the ligand have the lowest energy.
- The pre-placed seeds ensure high binding affinity and their optimal conformation determines the manner in which the ligand will be built, thus determining the overall structure of the final ligand. This strategy efficiently reduces the calculation burden for fragment construction.

- The two strategies above are widely used in most structure-based drug design programs. They are described as "**Grow**" and "**Link**". The two strategies are always combined in order to make the construction result more reliable

Scoring method

- Structure-based drug design attempts to use the structure of proteins as a basis for designing new ligands by applying accepted principles of molecular recognition.
- The basic assumption underlying structure-based drug design is that a good ligand molecule should bind tightly to its target. Thus, one of the most important principles for designing or obtaining potential new ligands is to predict the binding affinity of a certain ligand to its target and use it as a criterion for selection.
- One early method was developed by Böhm to develop a general-purposed empirical scoring function in order to describe the binding energy. The following "Master Equation" was derived:
- where:
- desolvation – enthalpic penalty for removing the ligand from solvent
- motion – entropic penalty for reducing the degrees of freedom when a ligand binds to its receptor
- configuration – conformational strain energy required to put the ligand in its "active" conformation
- interaction – enthalpic gain for "resolvating" the ligand with its receptor
- The basic idea is that the overall binding free energy can be decomposed into independent components that are known to be important for the binding process. Each component reflects a certain kind of free energy alteration during the binding process between a ligand and its target receptor. The Master Equation is the linear combination of these components. According to Gibbs free energy equation, the relation between dissociation equilibrium constant, K_d, and the components of free energy was built.

Lead Molecule

- Target validation (TV) → assay development → high-throughput screening → hit to lead (H2L) → lead optimization (LO) → preclinical drug development → clinical drug development
- **Hit to lead** (H2L) also known as **lead generation** is a stage in early drug discovery where small molecule hits from a high throughput screen (HTS) are

evaluated and undergo limited optimization to identify promising lead compounds. These lead compounds undergo more extensive optimization in a subsequent step of drug discovery called lead optimization(LO)

- The hit to lead stage starts with confirmation and evaluation of the initial screening hits and is followed by synthesis of analogs (hit expansion).
- Typically the initial screening hits display binding affinities for their biological target in the micromolar (10^{-6} molar concentration) range.
- Through limited H2L optimization, the affinities of the hits are often improved by several orders of magnitude to the nanomolar (10^{-9} M) range.
- The hits also undergo limited optimization to improve metabolic half life so that the compounds can be tested in animal models of disease and also to improve selectivity against other biological targets binding that may result in undesirable side effects.

Hit confirmation

After hits are identified from a high throughput screen, the hits are confirmed and evaluated using the following methods:

- Re-testing: compounds that were found active against the selected target are re-tested using the same assay conditions used during the HTS.
- Dose response curve generation: several compound concentrations are tested using the same assay, an IC_{50} or EC_{50}value is then generated. Methods are being developed that may allow the reuse of the compound that generated the hit in the initial HTS step. These molecules are removed from beads and transferred to a microarray for quantitative assessment of binding affinities in a "seamless" approach that could allow for the investigation of more hits and larger libraries.[4]
- Orthogonal testing: Confirmed hits are assayed using a different assay which is usually closer to the target physiological condition or using a different technology.
- Secondary screening: Confirmed hits are tested in a functional assay or in a cellular environment. Membrane permeability is usually a critical parameter.
- Chemical amenability: Medicinal chemists evaluate compounds according to their synthesis feasibility and other parameters such as up-scaling or costs
- Biophysical testing: Nuclear magnetic resonance (NMR), Isothermal Titration Calorimetry, dynamic light scattering,surface plasmon resonance, dual polarisation interferometry, microscale thermophoresis (MST) are commonly used to assess

whether the compound binds effectively to the target, the stoïchiometry of binding, any associated conformational change and to identify promiscuous inhibitors.

- Hit ranking and clustering: Confirmed hit compounds are then ranked according to the various hit confirmation experiments.
- Freedom to operate evaluation: Hit compound structures are quickly checked in specialized databases to determine if they are patentable

Hit expansion

Following hit confirmation, several compound clusters will be chosen according to their characteristics in the previously defined tests. An Ideal compound cluster will:

- have compound members that exhibit a high affinity towards the target (less than 1 µM)
- Moderate molecular weight and lipophilicity (usually measured as cLogP). Affinity, molecular weight and lipophilicity can be linked in single parameter such as ligand efficiency and lipophilic efficiency to assess druglikeness
- show chemical tractability
- be free of Intellectual property
- not interfere with the P450 enzymes nor with the P-glycoproteins
- not bind to human serum albumin
- be soluble in water (above 100 µM)
- be stable
- have a good druglikeness
- exhibit cell membrane permeability
- show significant biological activity in a cellular assay
- not exhibit cytotoxicity
- not be metabolized rapidly
- show selectivity versus other related targets

LEAD OPTMIZATION

- A lead compound (i.e. a "leading" compound, not lead metal) in drug discovery is a chemical compound that has pharmacological or biological activity likely to be therapeutically useful, but may still have suboptimal structure that requires modification to fit better to the target.

- Its chemical structure is used as a starting point for chemical modifications in order to improve potency, selectivity, or pharmacokinetic parameters.
- Furthermore, newly invented pharmacologically active moieties may have poor druglikeness and may require chemical modification to become drug-like enough to be tested biologically or clinically.

Discovering lead compounds

- A lead compound may arise from a variety of different sources.
- Lead compounds are found by characterizing natural products, employing combinatorial chemistry, or by molecular modeling as in rational drug design.
- Lead compounds are often tested by high-throughput screenings ("hits") which can screen compounds for their ability to inhibit (antagonist) or stimulate (agonist) a receptor of interest as well as determine their selectivity for them
- Traditional library screening, in house library, out sourced library, fragment based screening
- Virtual screening or *in silico* screening –computerised models of both the target and the molecule – no library required; followed by docking – estimate the binding energy using a scoring function

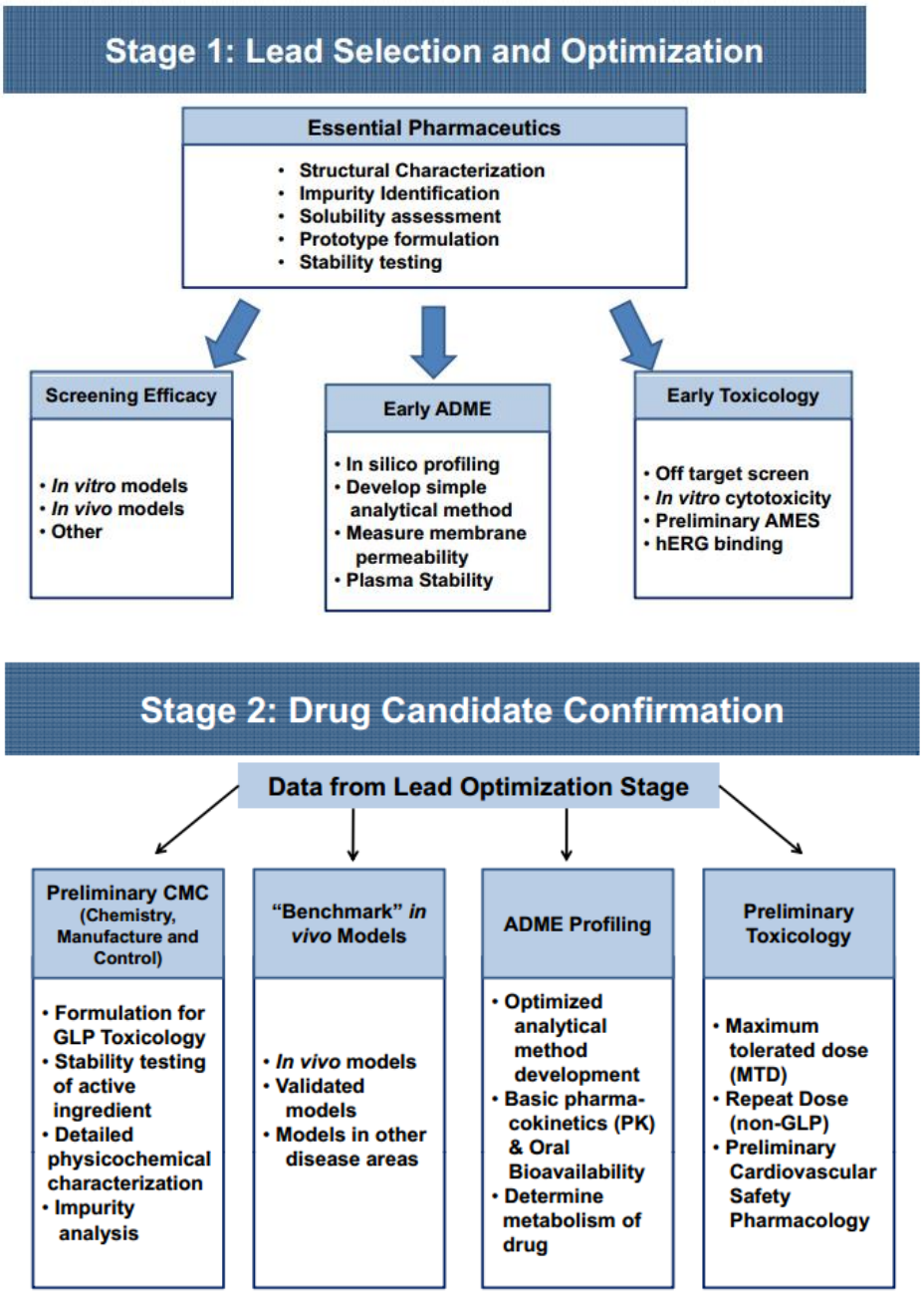
What makes a good Lead

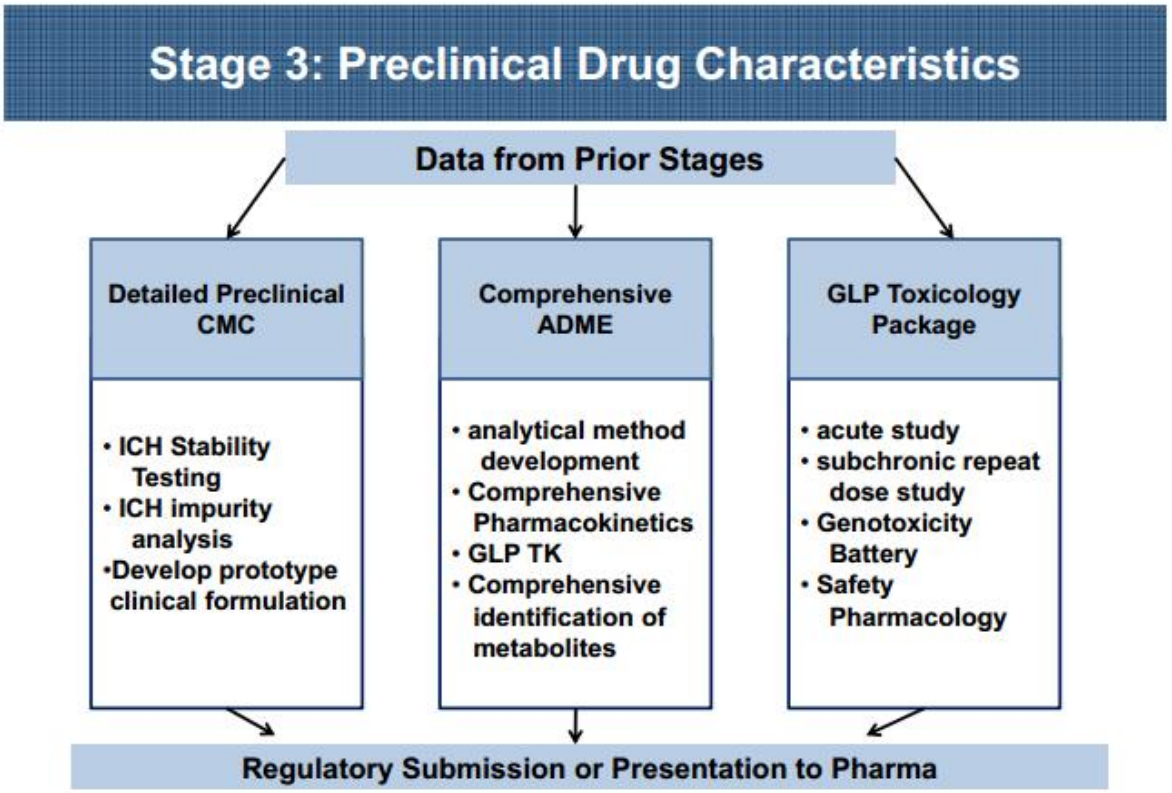
- Ideally multiple discrete series from High Throughput Screening
- Confirmed activity and structure using fresh, pure sample
- < 3uM *in vitro*, appropriate functional activity
- Reversibility
- Key assays available for selectivity
- Singletons need to be confirmed by small library synthesis
- Need to be 5-10 compounds to confirm options for diveristy
- No undesirable functional groups or chemical reactivity
- Toxicologically suspect groups
- Scope for expansion (IP position and suitable chemistry)
- Tractable physical properties
- Molecular Weight, Solubility, Polarity
- ADME profiles available for representative analogues
- Particularly oral absorption, solubility for i.v. and CNS penetration as appropriate.

- Evidence that any ADME issues are tractable

Lead optimization

- Lead identification/optimization is the one of the most important steps in drug development.
- The chemical structure of the lead compound is used as a starting point for chemical modifications in order to improve potency, selectivity, or pharmacokinetic parameters.
- Once a molecule is identified, the next step is to check its ADMET (Adsorption, Distribution, Metabolism, Excretion and Toxicity) properties.
- If the molecule has no toxicity and no mutagenicity either, it has potential for use as lead molecule. Further optimization gives better quality of lead molecules. These may subsequently be developed as drug(s).
- The objective of this drug discovery phase is to synthesize lead compounds, new analogs with improved potency, reduced off-target activities, and physiochemical/metabolic properties suggestive of reasonable in vivo pharmacokinetics. This optimization is accomplished through chemical modification of the hit structure, with modifications chosen by employing knowledge of the structure–activity relationship (SAR) as well as structure-based design if structural information about the target is available.
- Lead optimization is concerned with experimental testing and confirmation of the compound based on animal efficacy models and ADMET (in vitro and in situ) tools that may be followed by target identification and target validation.
-


Stage 1: Lead Selection and Optimization
Essential Pharmaceutics
• Structural Characterization
• Impurity Identification
• Solubility assessment
• Prototype formulation
• Stability testing
Screening Efficacy
• In vitro models
• In vivo models
• Other
Early ADME
• In silico profiling
• Develop simple analytical method
• Measure membrane permeability
• Plasma Stability
Early Toxicology
• Off target screen
• In vitro cytotoxicity
• Preliminary AMES
• hERG binding
Stage 2: Drug Candidate Confirmation
Data from Lead Optimization Stage
Preliminary CMC (Chemistry, Manufacture and Control)
• Formulation for GLP Toxicology
• Stability testing of active ingredient
• Detailed physicochemical characterization
• Impurity analysis
"Benchmark" in vivo Models
• In vivo models
• Validated models
• Models in other disease areas
ADME Profiling
• Optimized analytical method development
• Basic pharma-cokinetics (PK) & Oral Bioavailability
• Determine metabolism of drug
Preliminary Toxicology
• Maximum tolerated dose (MTD)
• Repeat Dose (non-GLP)
• Preliminary Cardiovascular Safety Pharmacology



Pharmacophore

- A pharmacophore is an abstract description of molecular features which are necessary for molecular recognition of a ligand by a biological macromolecule.
- The IUPAC defines a pharmacophore to be "an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response"
- Group on a molecule that interacts with receptor and responsible for biological activity
- A pharmacophore model explains how structurally diverse ligands can bind to a common receptor site. Furthermore pharmacophore models can be used to identify through denovo design or virtual screening novel ligands that will bind to the same receptor.

FEATURES

- Typical pharmacophore features include hydrophobic centroids, aromatic rings, hydrogen bond acceptors or donor, cations, and anions.

- These pharmacophoric points may be located on the ligand itself or may be projected points presumed to be located in the receptor.
- The features need to match different chemical groups with similar properties, in order to identify novel ligands.
- Ligand-receptor interactions are typically “polar positive”, “polar negative” or “hydrophobic”. A well-defined pharmacophore model includes both hydrophobic volumes and hydrogen bond vectors.

Process of pharmacophore development

The process for developing a pharmacophore model generally involves the following steps:

- Select a set of molecules

Choose a set of molecules that will be used for developing the pharmacophore model. As a pharmacophore model should be able to discriminate between molecules with and without bioactivity, the set of molecules should include both active and inactive compounds.

- Conformational analysis

Generate a set of low energy conformations that is likely to contain the bioactive conformation for each of the selected molecules.

- Molecular superimposition

Superimpose ("fit") all combinations of the low-energy conformations of the molecules. Similar functional groups common to all molecules in the set might be fitted (*e.g.*, phenyl rings or carboxylic acid groups). The conformation that results in the best fit is presumed to be the active conformation.

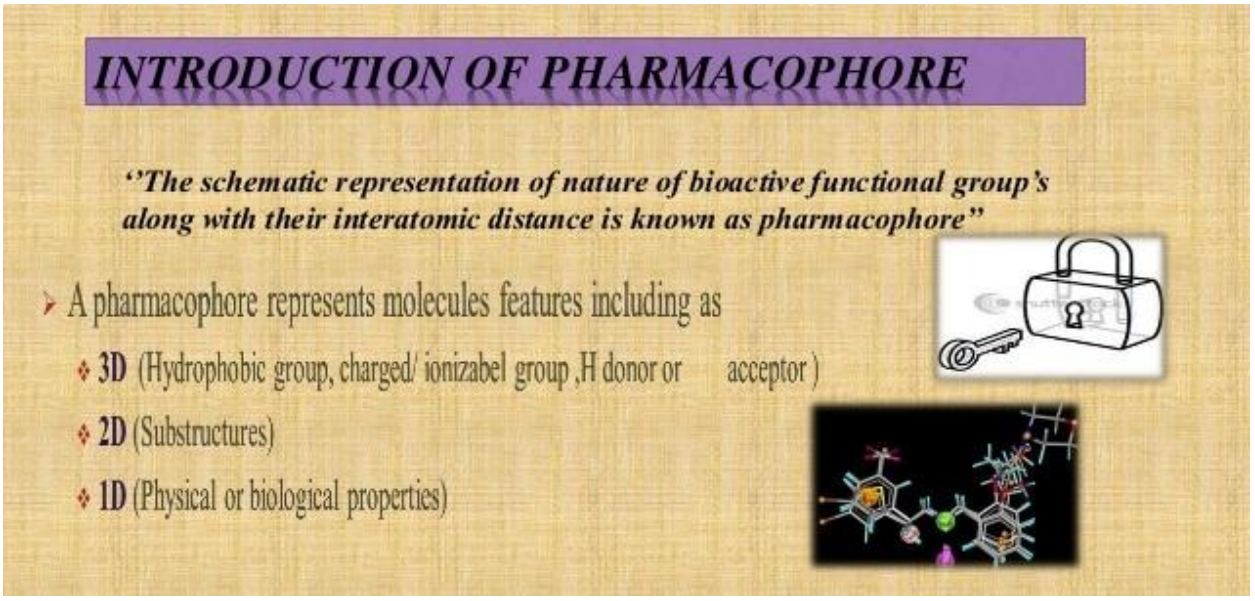
- Abstract description

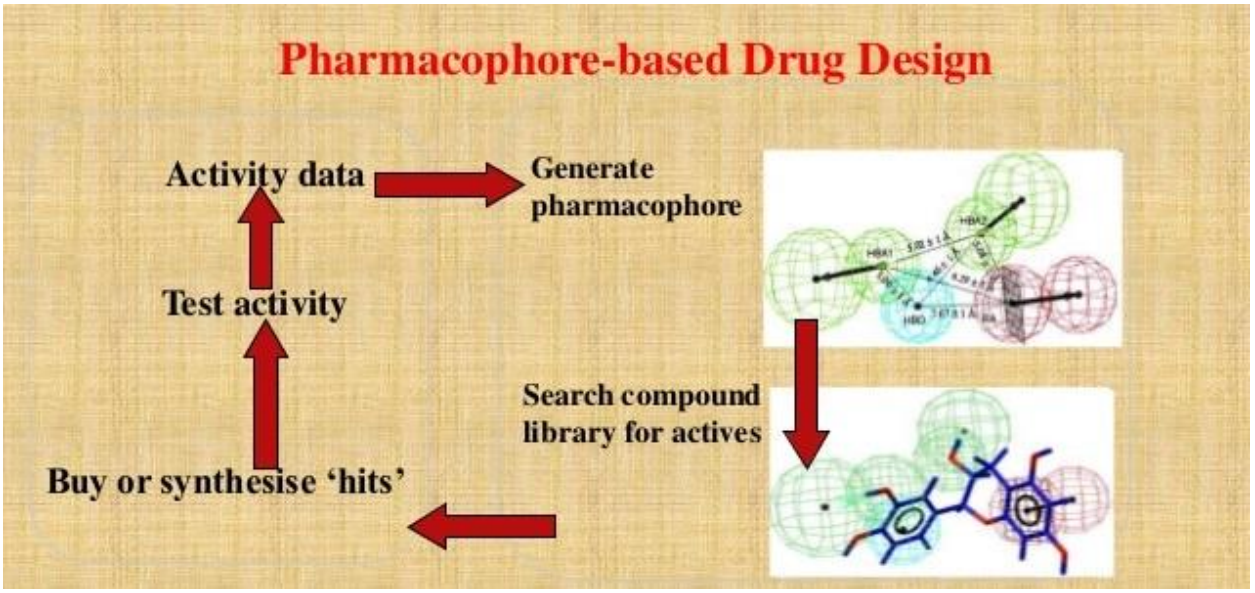
Transform the superimposed molecules into an abstract representation. For example, superimposed phenyl rings might be referred to more conceptually as an 'aromatic ring' pharmacophore element. Likewise, hydroxy groups could be designated as a 'hydrogen-bond donor/acceptor' pharmacophore element.

- Validation

A pharmacophore model is a *hypothesis* accounting for the observed biological activities of a set of molecules that bind to a common biological target. The model is only valid insofar as it is able to account for differences in biological activity of a range of molecules.

- As the biological activities of new molecules become available, the pharmacophore model can be updated to further refine it.





Molecular Docking

Molecular docking is a well established computational technique which predicts the interaction energy between two molecules. Molecular docking studies are used to determine the interaction of two molecules and to find the best orientation of ligand which would form a complex with overall minimum energy. The small molecule, known as ligand usually fits within protein's cavity which is predicted by the search algorithm. These protein cavities become active when come in contact with any external compounds and are thus called as active sites.

The results are analyzed by a statistical scoring function which converts interacting energy into numerical values called as the docking score; and also the interacting energy is calculated. The 3D pose of the bound ligand can be visualized using different visualizing tools like Pymol, Rasmol etc which could help in inference of the best fit of ligand. Predicting the mode of protein-ligand interaction can assume the active site of the protein molecule and further help in protein annotation. Moreover molecular docking has major application in drug discovery and designing.

This technique mainly incorporates algorithms like molecular dynamics, Monte Carlo stimulation, fragment based search methods.

Different types of Interactions

Interactions between particles can be defined as a consequence of forces between the molecules contained by the particles. These forces are divided into four categories:

- **Electrostatic forces** - Forces with electrostatic origin due to the charges residing in the matter. The most common interactions are charge-charge, charge-dipole and dipole-dipole.
- **Electrodynamics forces**-The most widely known is the Van der Waals interactions.
- **Steric forces** - Steric forces are generated when atoms in different molecules come into very close contact with one another and start affecting the reactivity of each other. The resulting forces can affect chemical reactions and the free energy of a system.
- **Solvent-related forces** - These are forces generated due to chemical reactions between the solvent and the protein or ligand. Examples are Hydrogen bonds (hydrophilic interactions) and hydrophobic interactions.
- A common characteristic of all these forces is their **electromagnetic** nature.
- Other physical factors - **Conformational changes** in the protein and the ligand are often necessary for successful docking.

Molecular docking

Molecular docking can be divided into two separate sections.

1) **Search algorithm** – These algorithms determine all possible optimal conformations

for a given complex (protein-protein, protein-ligand) in a environment i.e. the position and orientation of both molecules relative to each other. They can also calculate the energy of the resulting complex and of each individual interaction.

The different types of algorithms that can be used for docking analysis are given below.

- Molecular dynamics
- Monte Carlo methods
- Genetic algorithms
- Fragment-based methods
- Point complementary methods
- Distance geometry methods
- Systematic searches

2) **Scoring function** – These are mathematical methods used to predict the strength of the non-covalent interaction called as binding affinity, between two molecules after they have been docked. Scoring functions have also been developed to predict the strength of other types of intermolecular interactions, for example between two proteins or between protein and DNA or protein and drug. These configurations are evaluated using scoring functions to distinguish the experimental binding modes from all other modes explored through the searching algorithm.

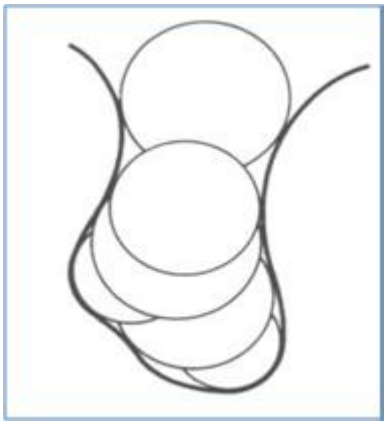
For example:

- Binding Energy

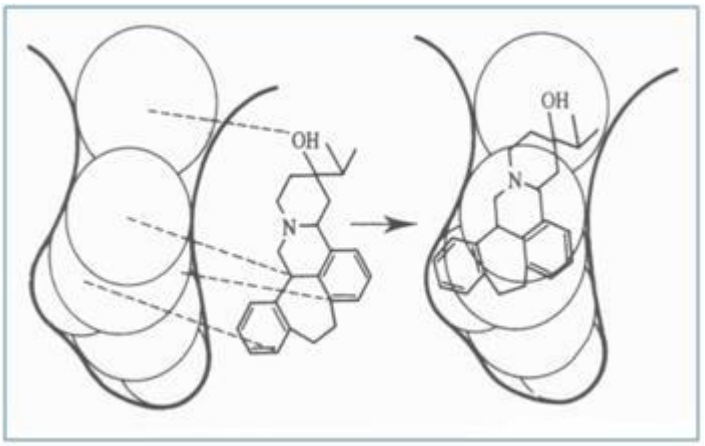
ΔG bind = ΔGvdw + ΔGhbond + ΔGelect + ΔGconform + ΔG tor + ΔG sol

General concept of the algorithm:

1) A 'negative' image of the binding site is made - a collection of spheres of varying radii, each touching the molecular surface at just 2 points.



2) Ligand atoms are then matched to sphere centers where at least four distances between ligand atoms are matched to sphere center distances.



3) Proper orientation is achieved by a least squares fit of ligand atoms to the sphere centers.
4) Orientation is checked for any steric clashes between ligand and receptor.
5) If acceptable, then interaction energy is computed as a 'score' for that binding mode
6) New orientations are obtained by matching different sets of atoms and sphere centers
7) Top-scoring orientations are retained for subsequent analysis

Types of Docking -

The following are majorly used type of docking are-

- **Lock and Key or Rigid Docking** – In rigid docking, both the internal geometry of the

receptor and ligand is kept fixed during docking.

- **Induced fit or Flexible Docking** - In this model, the ligand is kept flexible and the energy for different conformations of the ligand fitting into the protein is calculated. Though more time consuming, this method can evaluate many different possible conformations which make it more reliable.

Major steps in molecular docking:

Step I – Building the Receptor

In this step the 3D structure of the receptor should be downloaded from PDB; and modified. This should include removal of the water molecules from the cavity, stabilizing charges, filling in the missing residues, generation the side chains etc according to the parameters available. After modification the receptor should be biological active and stable.

Step II – Identification of the Active Site

After the receptor is built, the active site within the receptor should be identified. The receptor may have many active sites but the one of the interest should be selected. Most of the water molecules and heteroatoms if present should be removed.

Step III – Ligand Preparation

Ligands can be obtained from various databases like ZINC, PubChem or can be sketched using tools like Chemsketch. While selecting the ligand, the LIPINSKY'S RULE OF 5 should be applied. The rule is important for drug development where a pharmacologically active lead structure is optimized stepwise for increased activity and selectivity, as well as drug-like properties, as described.

For the selection of a ligand using LIPINSKY'S RULE:

- Not more than 5 –H bond donors.
- Molecular Weight NOT more than 500 Da.
- Log P not over 5 for octanol water partition coefficient.
- NOT more than 10 H bond acceptors.

Step IV- Docking

This is the last step, where the ligand is docked onto the receptor and the interactions are checked. The scoring function generates scores depending on which the ligand with the best fit is selected.

Software available for Molecular Docking:

SCHRODINGER, DOCK, AUTOLOCK TOOLS, DISCOVERY STUDIO, iGemDock

QSAR - **Quantitative structure–activity relationship**

Quantitative structure–activity relationship models (**QSAR** models) are regression or classification models used in the chemical and biological sciences and engineering. Like other regression models, QSAR regression models relate a set of "predictor" variables (X) to the potency of the response variable (Y), while classification QSAR models relate the predictor variables to a categorical value of the response variable.

In QSAR modeling, the predictors consist of physico-chemical properties or theoretical molecular descriptors of chemicals; the QSAR response-variable could be abiological activity of the chemicals. QSAR models first summarize a supposed relationship between chemical structures and biological activity in a data-set of chemicals. Second, QSAR models predict the activities of new chemicals.

Related terms include *quantitative structure–property relationships* (*QSPR*) when a chemical property is modeled as the response variable.

As an example, biological activity can be expressed quantitatively as the concentration of a substance required to give a certain biological response. Additionally, when physicochemical properties or structures are expressed by numbers, one can find a mathematical relationship, or quantitative structure-activity relationship, between the two. The mathematical expression, if carefully validated can then be used to predict the modeled response of other chemical structures.

A QSAR has the form of a mathematical model:

- Activity = *f*(physiochemical properties and/or structural properties) + error

The error includes model error (bias) and observational variability, that is, the variability in observations even on a correct model.

$$\textbf{Activity} = n_1x_1 + n_2x_2 + n_3x_3 + ... + \text{constant}$$

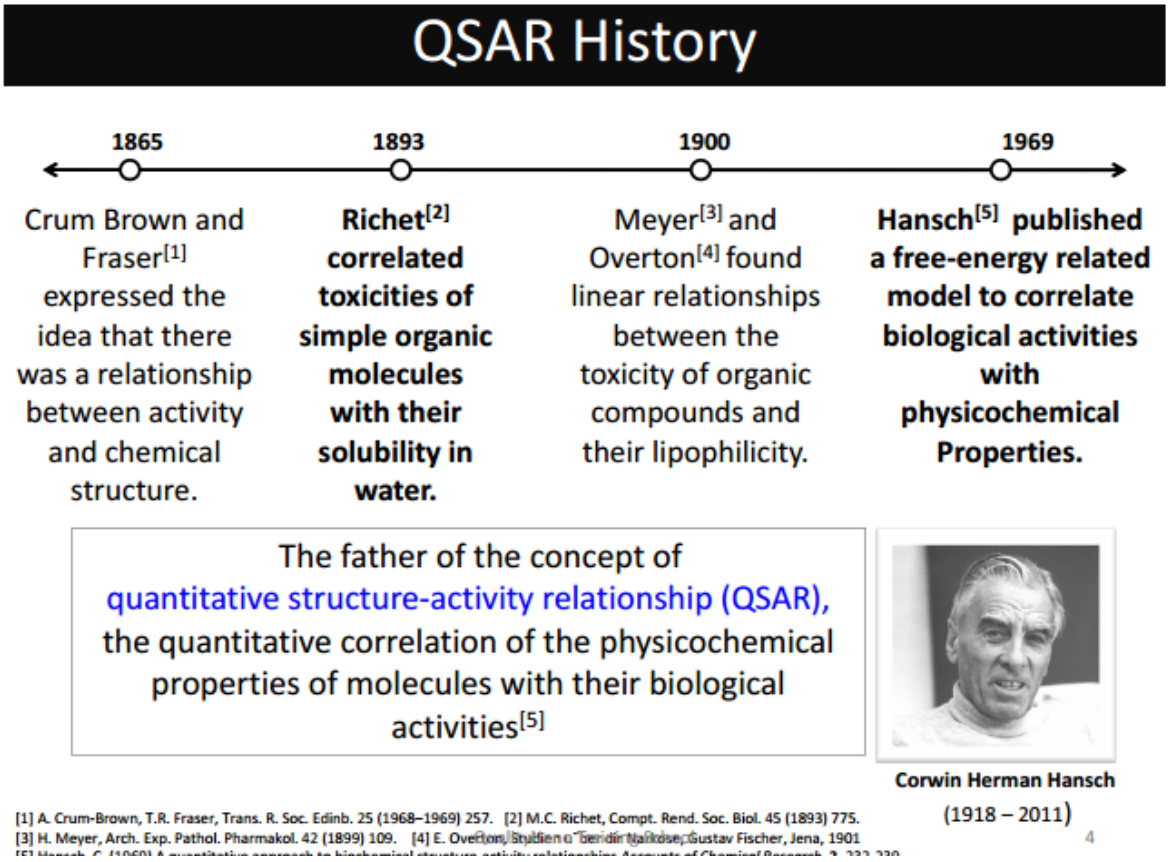
where nx = molecular descriptors

Examples of **biological activity** that can be used for QSAR studies include:

- Enzyme activity
- Minimum effective dose
- Toxicity

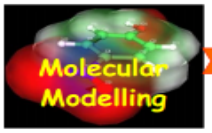
Possible **molecular descriptors** that can be used for building QSAR models may include:

- Dipole moment
- Atomic volume
- Number of carbons
- Number of aromatic moieties
- Molar volume
- Wang octanol-water partition coefficient
- Molecular weight
- Quantum chemical descriptors such as molecular orbital energies (HOMO & LUMO) and atomic net charge



Requirements for QSAR

1.Set of molecules → **2. Set of molecular descriptors** → **3. Measured biological activity or property** → **4.Statistical methods**



DESCRIPTORS
Physcochemical, biological, structural

QSARs
e.g. Neural Networks PLS, Expert Systems

Toxicity Endpoints

Daphnia magna EC50
Cancinogenicity
Mutagenicity
Rat oral LD50
Mouse inhalation LC50
Skin sensitisation
Eye irritancy

How many compounds are required to develop a QSAR?

There is no direct and simple answer → "as many as possible"

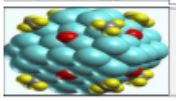
To provide some guide, it is widely accepted that between five and ten compounds are required for every descriptor in a QSAR [1]

Applications of QSAR

❑There are a large number of applications of these models within industry, academia and governmental (regulatory) agencies.

- The estimation of physico-chemical properties, biological activities and understanding the physicochemical features behind a biological response in drug desing.

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



- The prediction of a variety of physico-chemical properties of molecules.

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

- The identification of hazardous compounds at early stages of product development, the prediction of toxicity to humans and environment.

What is required for a good QSAR Model ?

The QSAR model should meet the requirements of the OECD principles [1]:

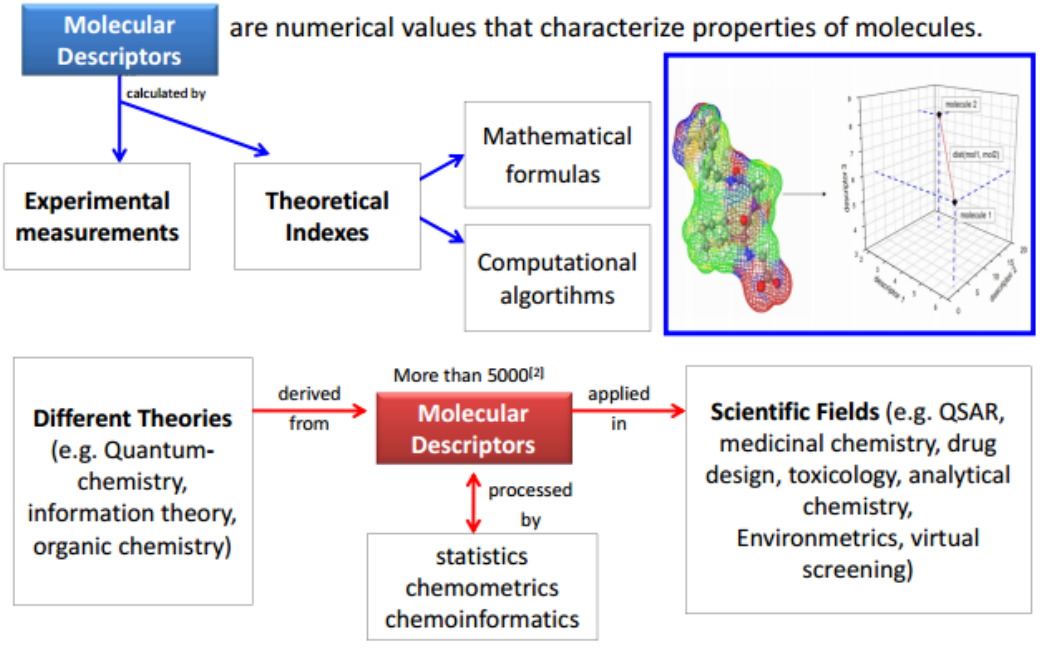
1. a defined endpoint;
2. an unambiguous algorithm;
3. a defined domain of applicability ;
4. appropriate measures of goodness-of-fit, robustness and predictivity;
5. a mechanistic interpretation if possible.

Common QSAR Modelling Errors

- Uninformative descriptors
- Poor descriptor selection and chance correlations
- Modelling complex, nonlinear structure property relationships with linear models
- Incorrectly validating QSPR models
- Not understanding the domain of applicability of models

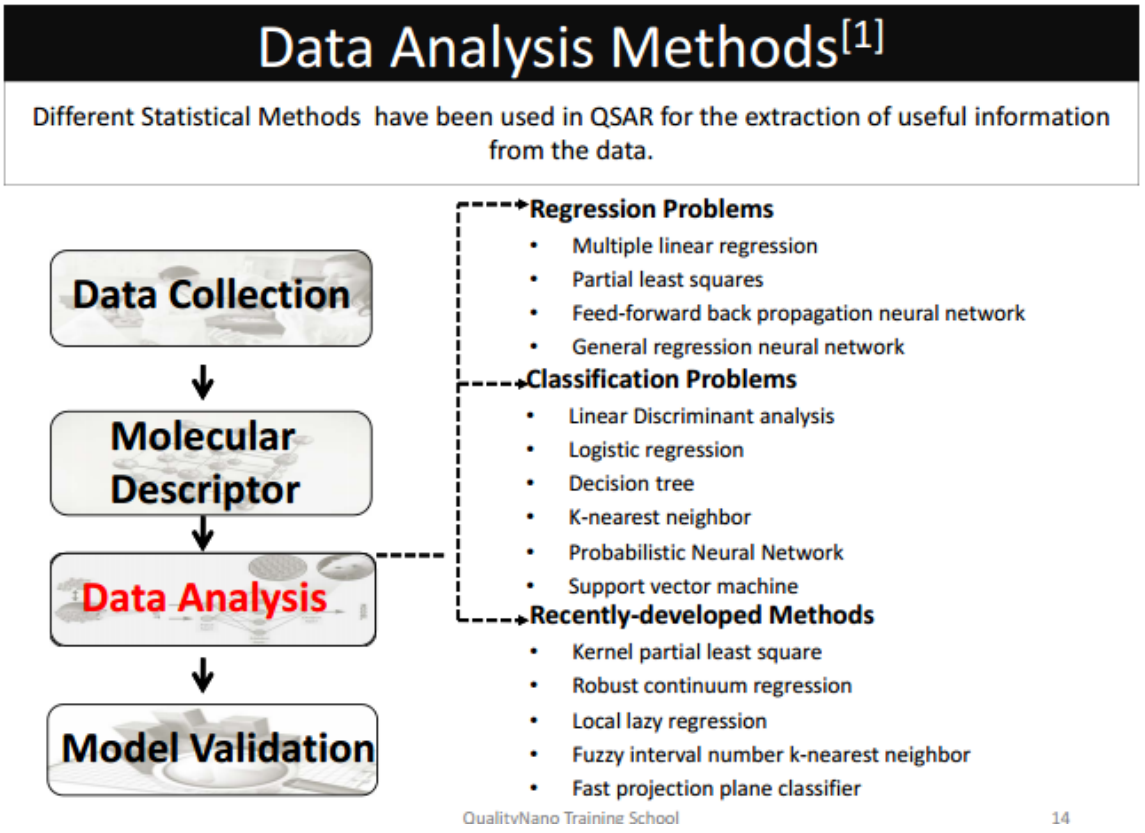
Molecular Descriptors[1]

Molecular Descriptors are numerical values that characterize properties of molecules.



[1] *T. Puzyn et al. (eds.), Recent Advances in QSAR Studies, 103–125.* DOI 10.1007/978-1-4020-9783-6_4, C Springer Science+Business Media B.V. 2010

[2] Todeschini R, Consonni V (2000) Handbook of molecular descriptors. Wiley-VCH, Weinheim



Advantages and Disadvantages of QSAR

Advantages of predicting biological activity with quantitative structure-activity relationships modelling include:

- Able to predict activities of a large number of compounds with little to no prior experimental data on activity.
- Can reveal which molecular properties may be worth investigating further.
- Regarded as a "green chemistry" approach since chemical waste is not generated when performing in silico predictions.
- In vivo and in vitro experimentation can be very expensive and time-consuming. QSAR modelling reduces the need for testing on animals and/or on cell cultures and saves time.

Disadvantages of predicting biological activity with QSAR modelling include:

- Does not provide an in-depth insight on the mechanism of biological action.
- Some risk of highly inaccurate predictions of pharmacological or biological activity.

Applications of QSAR in Pharmacology and Medicinal Chemistry

- Quantitative structure-activity relationships (QSAR) can be used during the drug design and drug discovery process. QSAR models can be used as a screening tool to test a large set of compounds or for eliminating test compounds which do not show promise in terms of predicted biological activity.
- Toxicity endpoints of compounds towards organisms can be predicted using QSAR-based methodologies. [1] For instance, the oral rat 50% lethal dose (LD50)

Molecular Descriptors used in QSAR

Molecular descriptors can be defined as a numerical representation of chemical information encoded within a molecular structure via mathematical procedure.267 This mathematical representation has to be invariant to the molecule's size and number of atoms to allow model building with statistical methods. The information content of structure descriptors depends on two major factors:

(1) The molecular representation of compounds.

(2) The algorithm which is used for the calculation of the descriptor.

The three major types of parameters initially suggested are, (1) Hydrophobic (2) Electronic (3) Steric

Table 11: Molecular Descriptors used in QSAR

Type	Descriptors
Hydrophobic Parameters	Partition coefficient ; log P
	Hansch's substitution constant; π
	Hydrophobic fragmental constant; f, f'
	Distribution coefficient; log D
	Apparent log P
	Capacity factor in HPLC; log k' , log k'_W
	Solubility parameter; log S
Electronic Parameters	Hammett constant; σ, σ^+, σ^-
	Taft's inductive (polar) constant; σ*
	Swain and Lupton field parameter
	Ionization constant; pK_a , ΔpK_a
	Chemical shifts: IR, NMR
Steric Parameters	Taft's steric parameter; Es
	Molar volume; MV
	Van der waals radius
	Van der waals volume
	Molar refractivity: MR
	Parachor
	Sterimol
Quantum chemical descriptors	Atomic net charge; Q^σ, Q^π
	Superdelocalizability
	Energy of highest occupied molecular orbital; E_{HOMO}
	Energy of lowest unoccupied molecular orbital; E_{LUMO}
Spatial Descriptor	Jurs descriptors, Shadow indices, Radius of Gyration, Principle moment of inertia

Basic

- **all ana**
- **all ana**
- **all ana**
- **the eff**
- **bindin**
- **biolog**

PARAMETERS

- The parameters used in QSAR is a measure of the potential contribution of its group to a particular property of the parent drug.
- Activity is expressed as log(1/C). C is the minimum concentration required to cause a defined biological response.
- Physicochemical property as log p.

Various parameters used in QSAR studies are

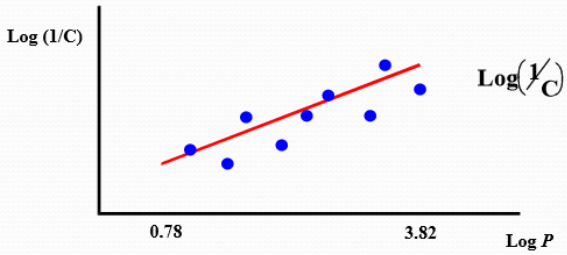
- 1.Lipophilic parameters: partition coefficient, π- substitution constant.
- 2.Electronic parameters: Hammet constant, dipole moment.
- 3.Steric parameters: Taft’s constant, molar refractivity, Verloop steric parameter

LIPOPHILIC PARAMETERS

- Lipophilicity is partitioning of the compound between an aqueous and non-aqueous phase.
- Partition coefficient: P=[drug] in octanol/[drug] in water
- High P High hydrophobicity

Linear relationship between Log p and Log 1/C

- Activity of drugs is often related to *P*
  e.g. binding of drugs to serum albumin
  (straight line - limited range of log *P*)

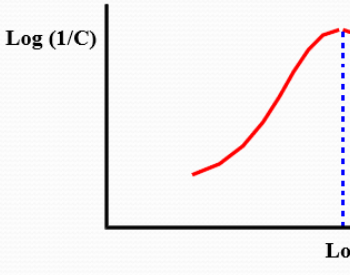


$$\text{Log}\left(\frac{1}{C}\right) = k1 \log P + k2$$

- Binding increases as log *P* increases
- Binding is greater for hydrophobic drugs

Non –linear relationship between Log P and Log 1/C

Example 2 General anaesthetic activity of ethers
(parabolic curve - larger range of log *P* values)



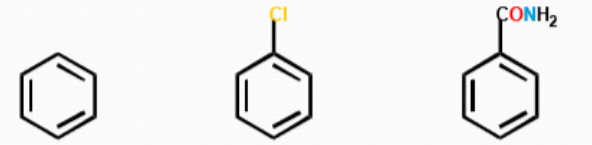
$$\text{Log}\left(\frac{1}{C}\right) = -k1 (\log P)^2 + k2 \log P + k3$$

Optimum value of log *P* for anaesthetic activity = log P^o

π-substituent constant or hydrophobic substituent constants:

- The π-substituent constant defined by Hansch and co-workers.
- Measure of how hydrophobic a substituent is,relative to H

$$\pi x = \log Px - \log PH$$

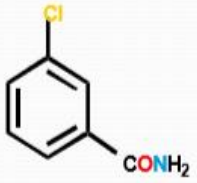


Benzene (Log *P* = 2.13) | Chlorobenzene (Log *P* = 2.84) | Benzamide (Log *P* = 0.64)

$\pi_{Cl} = 0.71$ $\pi_{CONH_2} = -1.49$

- Positive values imply substituents are more hydrophobic than H
- Negative values imply substituents are less hydrophobic than H

Example :



meta-Chlorobenzamide

$\text{Log } P_{(\text{theory})} = \log P_{(\text{benzene})} + \pi_{Cl} + \pi_{CONH_2}$
$= 2.13 + 0.71 - 1.49$
$= 1.35$

$\text{Log } P_{(\text{observed})} = 1.51$

- A QSAR equation may include both *P* and π.
- *P* measures the importance of a molecule's overall hydrophobicity (relevant to absorption, binding etc.)
- π identifies specific regions of the molecule which might interact with hydrophobic regions in the binding site

ELECTRONIC PARAMETERS

Hammett Substituent Constant (σ)

Eg. X= electron withdrawing group (e.g. NO_2)

X = electron withdrawing group

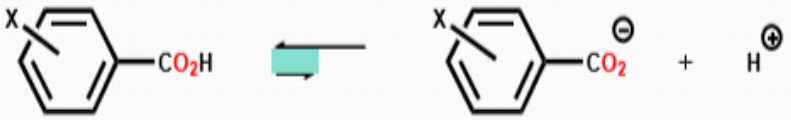
Charge is stabilised by X
Equilibrium shifts to right $K_X > K_H$

$$\sigma_x = \log \frac{K_X}{K_H} = \log K_X - \log K_H$$

Positive value

X= electron donating group (e.g. CH_3)

X = electron withdrawing group



Charge destabilised
Equilibrium shifts to left $K_X < K_H$

$$\sigma_x = \log \frac{K_X}{K_H} = \log K_X - \log K_H$$

Negative value

- σ value depends on inductive and resonance effects
- σ value depends on whether the substituent is meta or para
- ortho values are invalid due to steric factors

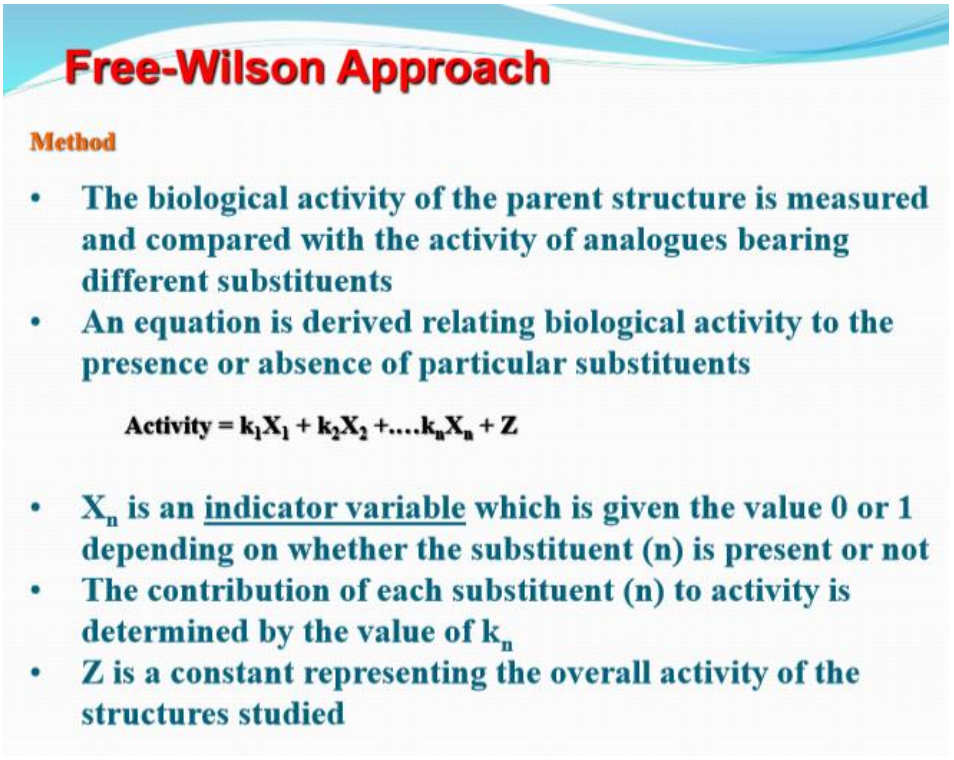
STERIC SUBSTITUTION CONSTANT

- It is a measure of the bulkiness of the group it represents and it effects on the closeness of contact between the drug and receptor site. It is much harder to quantitate.
- v Taft's steric factor (Es') •Measured by comparing the rates of hydrolysis of substituted aliphatic esters against a standard ester under acidic conditions
- Es = log kx - log ko
- kx represents the rate of hydrolysis of a substituted ester
- ko represents the rate of hydrolysis of the parent ester
- Molar refractivity (MR)--measure of the volume occupied by an atom or group--equation includes the MW, density(d), and the index of refraction(n)– $MR=(n^2-1)MW/(n^2+2)d$
- Verloop steric parameter--computer program uses bond angles, van der Waals radii, bond lengths

Hansch Equation

- **A QSAR equation relating various physicochemical properties to the biological activity of a series of compounds**
- **Usually includes log *P*, electronic and steric factors**
- **Start with simple equations and elaborate as more structures are synthesised**
- **Typical equation for a wide range of log *P* is parabolic**

$$\mathrm{Log}\left(\frac{1}{C}\right) = -k_1(\log P)^2 + k_2 \log P + k_3 \sigma + k_4 E_s + k_5$$



Advantages

- No need for physicochemical constants or tables
- Useful for structures with unusual substituents
- Useful for quantifying the biological effects of molecular features that cannot be quantified or tabulated by the Hansch method

Disadvantages

- A large number of analogues need to be synthesised to represent each different substituent and each different position of a substituent
- It is difficult to rationalise why specific substituents are good or bad for activity

3D QSAR

INTRODUCTION TO QSAR

- To relate the biological activity of a series of compounds to their physicochemical parameters in a quantitative fashion using a mathematical formula.
- The fundamental principe involved is difference in structural properties is responsible for variations in biological activities of the compound.
- Physico-chemical parameters:

  Hydrophobicity of substituents
  Electronic properties of substituents
  Hydrophobicity of the molecule
  Steric properties of substituents

- Hansch Analysis:

Corelates biological activity with physico-chemical parsmeters.

$$\text{Log}(1/c) = k_1 \text{logP} + k_2 \sigma + k_3 \text{Es} + k_4$$

- Free-Wilson Analysis:

Corelates biological activity with certain structural features of the compound.

Limitation:

Does not consider 3D structure.

No graphical output thereby making the interpretation of results in familiar chemical terms, frequently difficult if not impossible

3D QSAR

- 3D QSAR is an extension of classical QSAR which exploits the 3 dimensional properties of the ligands to predict their biological activity using robust stastical analysis like PLS, G/PLS, ANN etc.
- 3D-QSAR uses probe-based sampling within a molecular lattice to determine three-dimensional properties of molecules and can then correlate these 3D descriptors with biological activity.
- No QSAR model can replace the experimental assays, though experimental techniques are also not free from errors.
- Some of the major factors like desolvation energetics, temperature, diffusion, transport, pH, salt concentration etc. which contribute to the overall free energy of binding are difficult to handle, and thus usually ignored.
- Regardless of all such problems, QSAR becomes a useful alternative approach.

Table 2. Classification of 3D-QSAR approaches	
Classification	**Examples**
On the basis of intermolecular modeling, or the information used to develop QSAR	
Ligand-based 3D-QSAR	CoMFA, CoMSIA, COMPASS, GERM, CoMMA, SoMFA
Receptor-based 3D-QSAR	COMBINE, AFMoC, HIFA, CoRIA
On the basis of alignment criterion	
Alignment-dependent 3D-QSAR	CoMFA, CoMSIA, GERM, COMBINE, AFMoC, HIFA, CoRIA
Alignment-independent 3D-QSAR	COMPASS, CoMMA, HQSAR, WHIM, EVA/CoSA, GRIND
On the basis of the chemometric technique used for correlating structural properties and activities	
Linear 3D-QSAR	CoMFA, CoMSIA, AFMoC, GERM, CoMMA, SoMFA
Non-linear 3D-QSAR	COMPASS, QPLS

CoMFA(Comparative Molecular Field Analysis)

- In 1987, Cramer developed the predecessor of 3D approaches called Dynamic Lattice-Oriented Molecular Modeling System (DYLOMMS) that involves the use of PCA to extract vectors from the molecular interaction fields, which are then correlated with biological activities.
- Soon after he modified it by combining the two existing techniques, GRID and PLS, to develop a powerful 3D QSAR methodology, Comparative Molecular Field Analysis (CoMFA).
- The underlying idea of CoMFA is that differences in a target property, e.g., biological activity, are often closely related to equivalent changes in shapes and strengths of non-covalent interaction fields surrounding the molecules.
- Hence, the molecules are placed in a cubic grid and the interaction energies between the molecule and a defined probe are calculated for each grid point.

Protocol for CoMFA:

A standard CoMFA procedure, as implemented in the Sybyl Software, follows the following sequential steps:

- Bioactive conformations of each molecule are determined.
- All the molecules are superimposed or aligned using either manual or automated methods, in a manner defined by the supposed mode of interaction with the receptor.
- The overlaid molecules are placed in the center of a lattice grid with a spacing of 2 Å.
- The algorithm compares, in three-dimensions, the steric and electrostatic fields calculated around the molecules with different probe groups positioned at all intersections of the lattice.

- The interaction energy or field values are correlated with the biological activity data using PLS technique, which identifies and extracts the quantitative influence of specific chemical features of molecules on their biological activity.
- The results are articulated as correlation equations with the number of latent variable terms, each of which is a linear combination of original independent lattice descriptors.
- For visual understanding, the PLS output is presented in the form of an interactive graphics consisting of colored contour plots of coefficients of the corresponding field variables at each lattice intersection, and showing the imperative favorable and unfavorable regions in three dimensional space which are considerably associated with the biological activity.

DRAWBACKS AND LIMITATIONS OF CoMFA

CoMFA has several pitfalls and imperfections:

- Too many adjustable parameters like overall orientation, lattice placement, step size, probe atom type etc.
- Uncertainty in selection of compounds and variables
- Fragmented contour maps with variable selection procedures
- Hydrophobicity not well-quantified
- Cut-off limits used
- Low signal to noise ratio due to many useless field variables
- Imperfections in potential energy functions
- Various practical problems with PLS
- Applicable only to in vitro data

CoMSIA

- Comparative Molecular Similarity Indices Analysis (CoMSIA) was developed to overcome certain limitations of CoMFA.
- In CoMSIA, molecular similarity indices calculated from modified SEAL similarity fields are employed as descriptors to simultaneously consider steric, electrostatic, hydrophobic and hydrogen bonding properties.
- These indices are estimated indirectly by comparing the similarity of each molecule in the dataset with a common probe atom (having a radius of 1 Å, charge of +1 and hydrophobicity of +1) positioned at the intersections of a surrounding grid/lattice.
- For computing similarity at all grid points, the mutual distances between the probe atom and the atoms of the molecules in the aligned dataset are also taken into account.
    - To describe this distance-dependence and calculate the molecular properties, Gaussian-type functions are employed. Since the underlying Gaussian-type functional forms are 'smooth' with no singularities, their slopes are not as steep as the Coulombic and Lennard-Jones potentials in CoMFA; therefore, no arbitrary cut-off limits are required to be defined.
    - CoMSIA is provided by Tripos Inc. in the Sybyl software [33], along with CoMFA.

	CoMFA	CoMSIA
function type	Lennard-Jones potenttial, Coulomb potentials	Gaussian
descriptors	interaction energies	similarity indices
cut-off	required	no required
field	steric, electrostatic	steric, electrostatic, hydrophobic, hydrogen-bond donor and hydrogen-bond acceptor
contour map	often not contiguous	contiguous
model reproducibility	poor	good

APPLICATIONS:

1. QSAR in Chromatography: Quantitative Structure–Retention Relationships (QSRRs)
2. The Use of QSAR and Computational Methods in Drug Design.
3. In Silico Approaches for Predicting ADME Properties.
4. Prediction of Harmful Human Health Effects of Chemicals from Structure.
5. Chemometric Methods and Theoretical Molecular Descriptors in Predictive QSAR Modeling of the Environmental Behavior of Organic Pollutants
6. The Role of QSAR Methodology in the Regulatory Assessment of Chemicals
7. Nanomaterials – the Great Challenge for QSAR Modelers

CONCLUSION

CoMFA and CoMSIA are useful techniques in understanding pharmacological properties of studied compounds, and they have been successfully used in modern drug design. Despite of all the pitfalls it has now been globally used for drug discovery based on well-established principles of statistics, is intrinsically a valuable and viable medicinal chemistry tool whose application domain range from explaining the structure-activity relationships quantitatively and retrospectively, to endowing synthetic guidance leading to logical and experimentally testable hypotheses. Apart from synthetic applications it has also been used in various other fields too.

Evaluation of the quality of QSAR models

QSAR modeling produces predictive models derived from application of statistical tools correlating biological activity (including desirable therapeutic effect and undesirable side effects) or physico-chemical properties in QSPR models of chemicals (drugs/toxicants/environmental pollutants) with descriptors representative of molecular structure or properties. QSARs are being applied in many disciplines, for example: risk assessment, toxicity prediction, and regulatory decisions in addition to drug discovery and lead optimization.[30] Obtaining a good quality QSAR model depends on many factors, such as the quality of input data, the choice of descriptors and statistical methods for modeling and for validation. Any QSAR modeling should ultimately lead to statistically robust and predictive models capable of making accurate and reliable predictions of the modeled response of new compounds.

For validation of QSAR models, usually various strategies are adopted:[31]

1. internal validation or cross-validation (actually, while extracting data, cross validation is a measure of model robustness, the more a model is robust (higher q2) the less data extraction perturb the original model);
2. external validation by splitting the available data set into training set for model development and prediction set for model predictivity check;
3. blind external validation by application of model on new external data and
4. data randomization or Y-scrambling for verifying the absence of chance correlation between the response and the modeling descriptors.

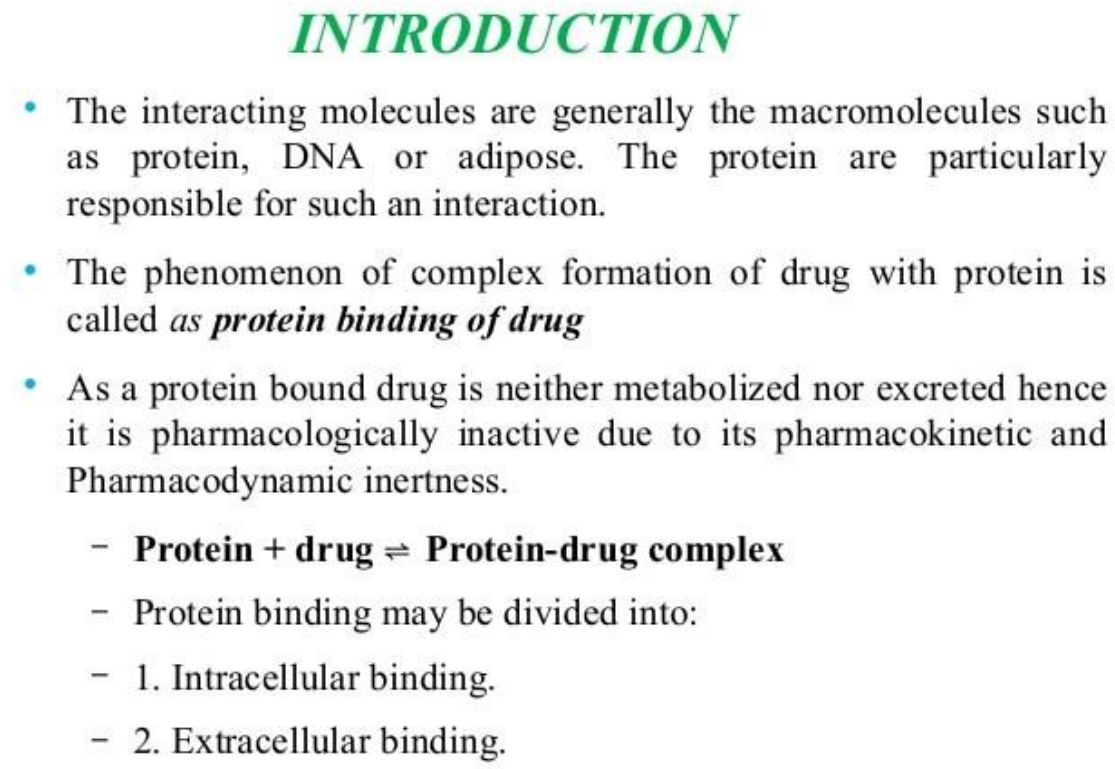
The success of any QSAR model depends on accuracy of the input data, selection of appropriate descriptors and statistical tools, and most importantly validation of the developed model. Validation is the process by which the reliability and relevance of a procedure are

established for a specific purpose; for QSAR models validation must be mainly for robustness, prediction performances and applicability domain (AD) of the models.

Some validation methodologies can be problematic. For example, *leave one-out* cross-validation generally leads to an overestimation of predictive capacity. Even with external validation, it is difficult to determine whether the selection of training and test sets was manipulated to maximize the predictive capacity of the model being published.

Different aspects of validation of QSAR models that need attention include methods of selection of training set compounds,[34] setting training set size[35] and impact of variable selection[36] for training set models for determining the quality of prediction. Development of novel validation parameters for judging quality of QSAR models is also important.

Drug binding kinetics



MECHANISMS OF PROTEIN DRUG BINDING:

- Binding of drugs to proteins is generally of reversible &irreversible.
- Reversible generally involves weak chemical bond such as:
    1. Hydrogen bonds
    2. Hydrophobic bonds
    3. Ionic bonds
    4. Van der waal's forces.

- Irreversible drug binding, though rare, arises as a result of covalent binding and is often a reason for the carcinogenicity or tissue toxicity of the drug.

KINETICS OF PROTEIN-DRUG BINDING

If P represents proteins and D the drug, then applying **law of mass** action to reversible protein-drug binding, we can write:

$$P + D \rightleftharpoons PD$$

At equilibrium,

$$K_a = \frac{[PD]}{[P][D]}$$

$$[PD] = K_a [P][D]$$

where, [P] = concentration of free protein
[D] = concentration of free drug
[PD] = concentration of protein-drug complex
K_a = association rate constant
K_d = dissociation rate constant

$K_a > K_d$ indicates forward reaction i.e. protein-drug binding is favoured. If P_T is the total concentration of protein present, bound and unbound, then:

$$P_T = [PD] + [P]$$

If r is the number of moles of drug bound to total moles of protein, then,

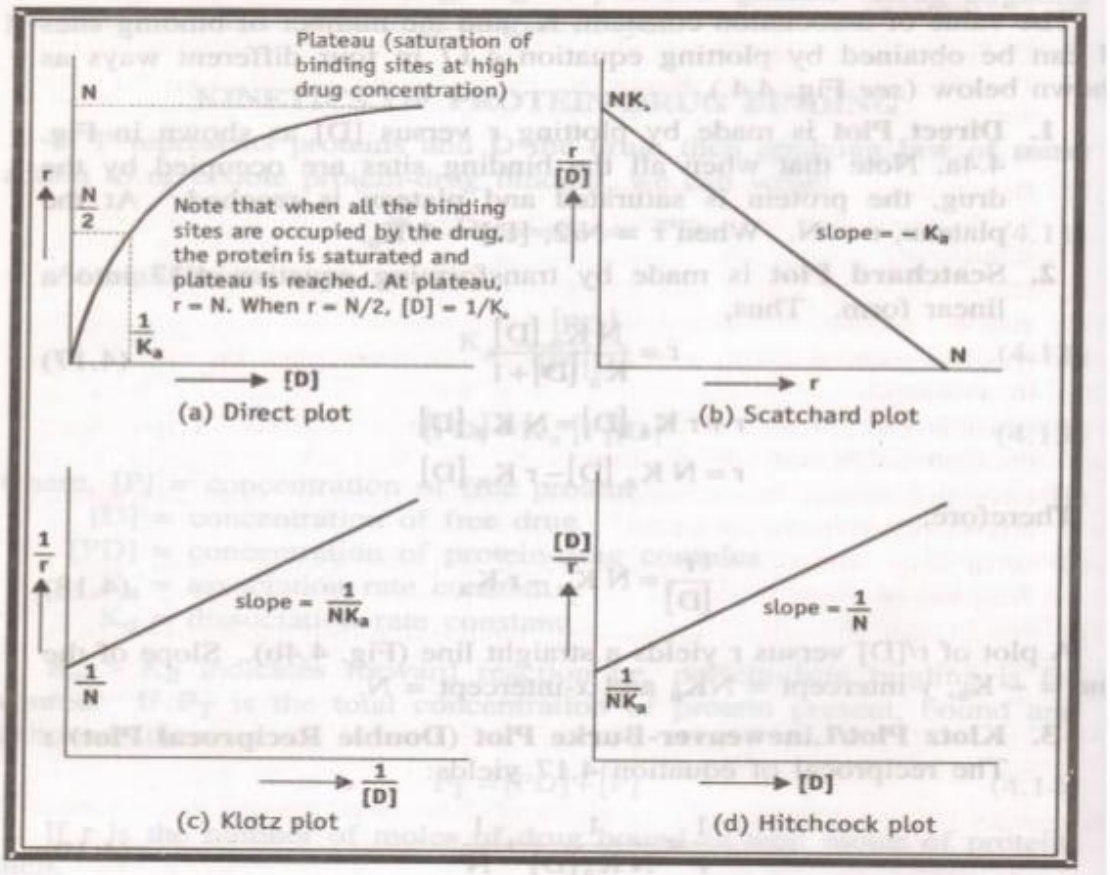
$$r = \frac{[PD]}{[P_T]} = \frac{[PD]}{[PD] + [P]}$$

Substituting the value of [PD]

$$r = \frac{K_a [P][D]}{K_a [P][D] + [P]} = \frac{K_a [D]}{K_a [D] + 1}$$

Equation holds when there is only one binding site on the protein and the protein-drug complex is a 1:1 complex. If more than one or N number of binding sites are available per mole of the protein then:

$$r = \frac{N K_a [D]}{K_a [D] + 1}$$

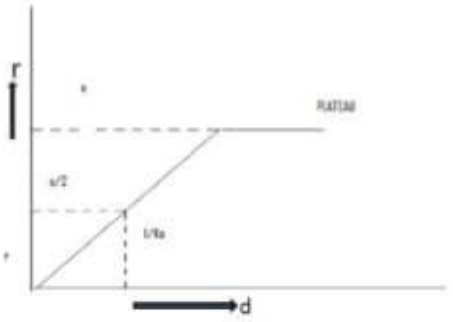


PLOTS OF DRUG DISTRIBUTION

1) DIRECT PLOT METHOD:

A direct plot of "r" Vs [D] can be used to find out the no of binding sites on protein 'n' (plateau value).

Ka is obtained by finding drug conc required to saturate the half of the total binding sites available (i.e; n/2).



2) SCATCHARD PLOT:

Obtained by rearranging an equation into linear form.

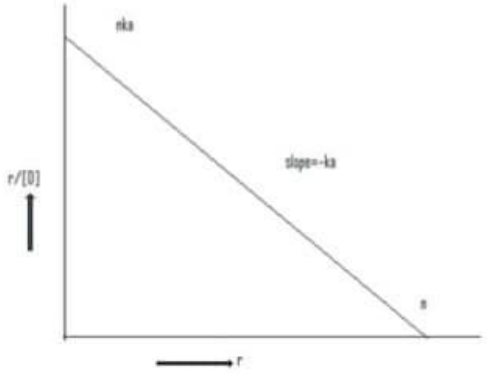
$$r = \frac{nK_a[D]}{K_a[D] + 1}$$

$$r + rK_a[D] = nK_a[D]$$

$$r = nK_a[D] - rK_a[D]$$

$$\frac{r}{[D]} = nK_a - rK_a$$

A plot of r/[D] Vs r yields a st.line with X &Y intercepts equal to 'n' & 'nKa' & the slope is equal to Ka.



Scatchard plot

3) DOUBLE RECIPROCAL PLOT OR KLOTZ PLOT:

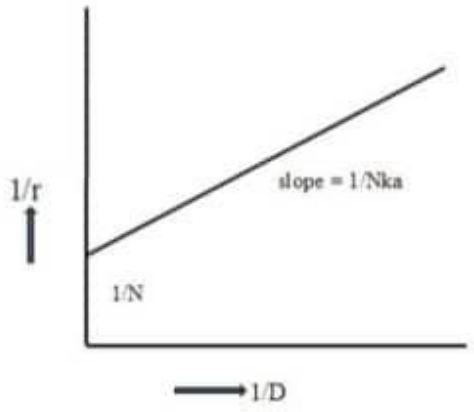
(LINE WEAVER -BURK PLOT)

Reciprocal of equation gives-

$$1/r = 1/nka(D) + 1/n$$

A plot of 1/r Vs 1/D yields a double reciprocal plot.

It is straight line with slope 1/Nka and Y-intercept 1/N



Klotz plot

4) HITCHCOCK PLOT

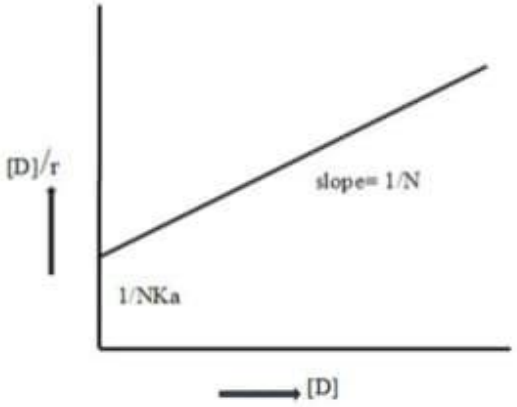
It is made by rearranging the equation as –

$$N\,Ka\,[D]/r = 1 + Ka$$

dividing both sides by Nka gives –

$$[D]/r = 1/NKa + [D]/N$$

A plot of $[D]/r$ Vs $[D]$ yields a straight line with slope $1/N$ and intercept $1/NKa$



Hitchcock plot

1. BINDING OF DRUG TO BLOOD COMPONENTS

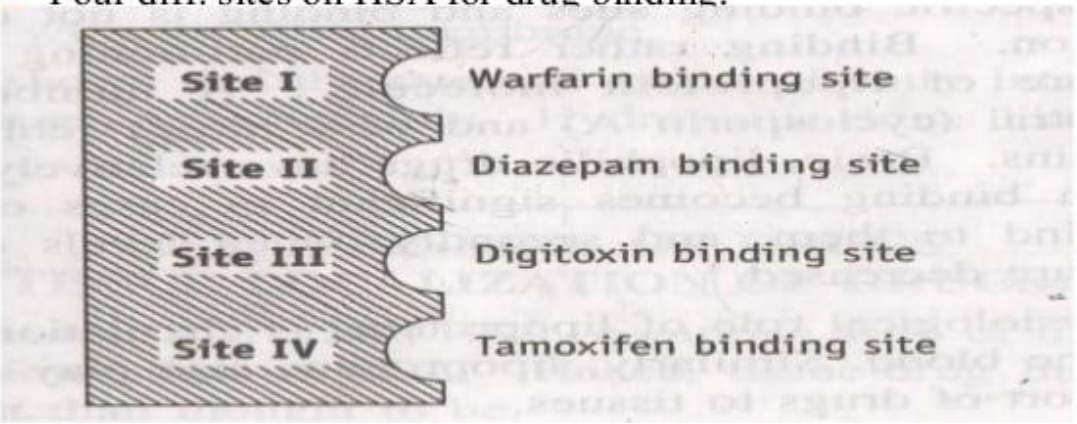
A. Plasma protein-drug binding

- The binding of drugs to plasma proteins is reversible.
- The extent or order of binding of drug to plasma proteins is:

  Albumin ›α1-Acid glycoprotein ›Lipoproteins ›Globulins.

1. Binding of drug to human serum Albumin.

- It is the most abundant plasma protein (59%), having M.W. of 65,000 with large drug binding capacity.
- Both endogenous compounds such as fatty acid, bilirubin as well as drug binds to HSA.
- Four diff. sites on HSA for drug binding.



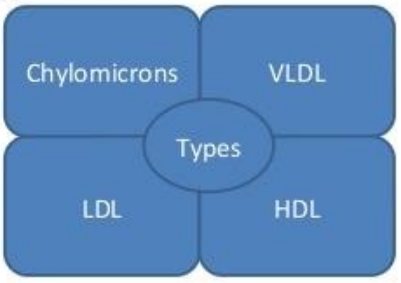
2. Binding of drug to α1-Acid glycoprotein:

(orosomucoid)

It has a M.W. 44,000 and plasma conc. range of 0.04 to 0.1 g%. It binds to no. of basic drugs like imipramine, lidocaine, propranolol, quinidine.

3. Binding of drug to Lipoproteins:

Mol wt: 2-34 Lakh dalton.
It binds to,
Acidic: Diclofenac.
Neutral: Cyclosporin A.
Basic: Chlorpromazine.



4. Binding of drug to Globulins

Globulin	Synonym	Binds to
1. α1 Globulin	Transcortine /Corticosteroid Binding globulin	Steroidal drugs, Thyroxin & Cyanocobalamine.
2. α2 Globulin	Ceruloplasmine	Vitamin A,D,E,K.
3. β1Globulin	Transferin	Ferrous ions
4. β2Globulin	---	Carotinoids
5. γ Globulin	---	Antigens

B. BINDING OF DRUG TO BLOOD CELLS

- In blood 40% of blood cells of which major component is RBC (95%). The RBC is 500 times in diameter as the albumin. The rate & extent of entry into RBC is more for lipophilic drugs.
- The RBC comprises of 3 components.

a) **Haemoglobin:** It has a M.W. of 64,500 Dal. Drugs like phenytoin, pentobarbital bind to haemoglobin.

b) **Carbonic anhydrase:** Carbonic anhydrase inhibitors drugs are bind to it like acetazolamide & chlorthalidone.

c) **Cell membrane:** Imipramine & chlorpromazine are reported to bind with the RBC membrane.

2. BINDING OF DRUG TO EXTRAVASCULAR TISSUE PROTEIN

- Importance: 1. It increases apparent volume of distribution of drug.
  2. localization of a drug at a specific site in body.

- Binding order: ***Liver > Kidney > Lung > Muscles***

Tissue	Binding of
1.Liver	Irreversible binding of Halogenated Hydrocarbon & Paracetamol.
2.Lungs	Basic drugs: Imipramine, Chlorpromazine, & AntiHistaminics.

Cont...

Tissue	Binding of
3.Kidney	Metallothionin protein binds to Heavy metals & results in Renal accumulation and toxicity.
4.Skin	Chloroquine & Phenothiazine binds to Melanin.
5.Eye	Chloroquine & Phenothiazine also binds to Eye Melanin & results in Retinopathy.
6.Hairs	Chloroquine, & Phenothiazine.
7.Bones	Tetracycline(yellow discoloration of teeth), Lead(replaces Ca & cause brittleness)
8.Fats	Lipophilic drugs (thiopental), Pesticides (DDT)
9.Nucleic Acid	Chloroquine & Quinacrine.

Comparison Between Plasma Protein-Drug Binding and Tissue-Drug Binding

Plasma protein – drug binding	Tissue-drug binding
Involves weak bonds and thus reversible	Involves strong and covalent bonds and thus irreversible
Small apparent volume of distribution	Large apparent volume of distribution
Half life is relatively short	Half life is relatively long
Does not result toxicity	Tissue toxicity is common
Displacement is possible by other drugs	Displacement by drugs generally does not occur
Competition between drugs for binding to plasma proteins can occurs	Tissue-drug binding is generally non-competitive

Determination of Protein-drug Binding

1. **Indirect technique**: Based on separation of bound form.

✓ Equilibrium dialysis

✓ Dynamic dialysis

✓ Ultrafiltration

✓ Diafiltration

✓ Gel filtration

✓ Ultracentrifugation

2. **Direct technique**: Do not required separation of bound form.

✓ UV Spectroscopy

✓ Fluorimetry

✓ Ion selective electrodes

FACTORS AFFECTING PROTEIN DRUG BINDING

1. Drug - related factors

a) Physicochemical characteristics of the drugs
b) Concentration of drugs in the body
c) Affinity of drug for a particular binding components

2. Protein / Tissue related factors

d) Physicochemical characteristics of the protein or binding agents
e) Concentration of protein or binding components
f) Number of binding sites on the binding agents

3. Drug interactions

g) Competition between drugs for the binding site
h) Competition between the drug and normal body constituents
i) Allosteric changes in protein molecule

4. Patient-related factors

j) Age
k) Intersubject variations
l) Disease states

FACTORS AFFECTING PROTEIN DRUG BINDING

1. Drug-related factors

a. Physicochemical characteristics of the drug:-

- Protein binding is directly related to the lipophilicity of drug. An increase in lipophilicity increases the extent of binding .e.g. Highly lipophilic drugs such as thiopental tend to localize in adipose tissue.

b. Concentration of drug in the body:-

- Alteration in the concentration of drug substance as well as the protein molecules or surfaces subsequently brings alteration in the protein binding process.

c. Affinity of a drug for a particular binding component:-

- This factor entirely depends upon the degree of attraction or affinity the protein molecule or tissues have towards drug moieties.
- For Digoxin has more affinity for cardiac muscles proteins as compared to that of proteins of skeletal muscles or those in the plasma like HSA.

2. Protein/ tissue related factors:

a. Physicochemical characteristics of protein or binding agent:

- Lipoproteins & adipose tissue tend to bind lipophilic drug by dissolving them in their lipid core.

b. Concentration of protein or binding component:

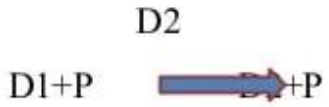
- Among the plasma protein , binding predominantly occurs with albumin, as it is present in high concentration in comparision to other plasma protein.
- The amount of several proteins and tissue components available for binding, changes during disease state.

c. Number of Binding Sites on the Protein:

- Albumin has a large number of binding sites as compared to other protein. Several drugs are capable of binding on more than one site on albumin. e.g Indomethacine is known to bind to 3 different sites.

3. Drug interactions

a. Competition between drugs for the binding sites[Displacement interactions]:-



D1: Displaced drug. D2: Displacer drug.

e.g. Administration of phenylbutazone to a patient on Warfarin therapy results in Hemorrhagic reaction.

b. Competition between drug & normal body constituents:-

The free fatty acids are known to interact with a no. of drugs that binds primarily to HSA. The free fatty acid level increase in physiological (fasting), pathological condition (Diabetes).

c. Allosteric changes in protein molecule:-

- The process involves alteration of the protein structure by the drug or it's metabolite thereby modifying its binding capacity.
- e.g. aspirin acetylates lysine fraction of albumin thereby modifying its capacity to bind NSAIDs like phenylbutazone (increased affinity) and flufenamic acid (decreased affinity).

4. Patient-related factors

a. Age:

1.Neonates: Low albumin content: More free drug.

2.Young infants: High dose of Digoxin due to large renal clearance.

3.Elderly:Low albumin: So more free drug.

b. Intersubject variability:

Due to genetics & environmental factors.

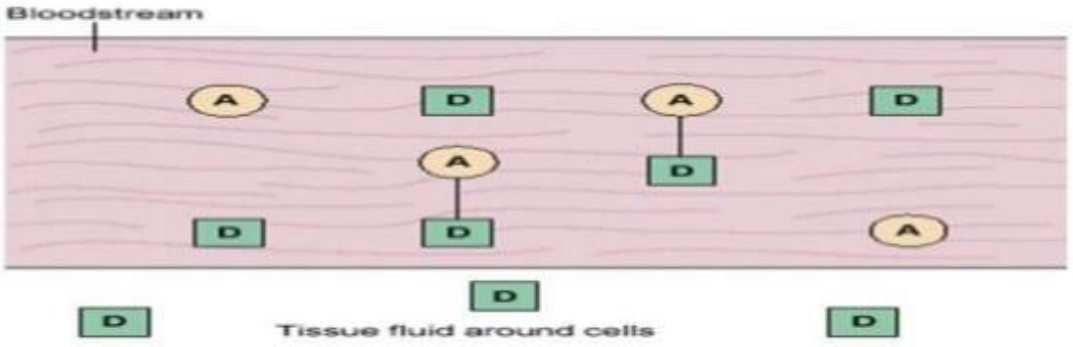
c. Disease states:

Disease	Influence on plasma protein	Influence on protein drug binding
Renal failure	↓ Albumin content	↓ binding of acidic drugs; neutral and basic drugs are un affected
Hepatic failure	↓ Albumin synthesis	↓ binding of acidic drugs; and binding of basic drugs is normal or ↓ depending on AAG levels
Inflamatory states i.e,truama surgery etc…	↑AAG levels	↑ binding of basic drugs; neutral and acidic drugs are un affected

SIGNIFICANCE OF PROTEIN/TISSUE BINDING OF DRUG

a. Absorption

- The absorption equilibrim is attained by transfer of free drug from the site of administration to the systemic circulation. Following the equilibrium the process may stop.
- However, binding of the absorbed drug to plasma proteins decreases free drug concentration thus sink condition and concentration gradient are re-established which now act as driving force for further absorption.



b. Distribution

- A protein bound drug in particular does not cross the BBB, the placental barrier, the glomerulus.
- Thus protein binding decreases the distribution of drugs.

c. Metabolism

- Protein binding decreases the metabolism of drugs & enhances the biological half life.
- Only unbound fraction get metabolized.
- e.g. Phenylbutazone & Sulfonamide.

d. Elimination

- Only the unbound drug is capable of being eliminated.
- Protein binding prevent the entry of drug to the metabolizing organ (liver) & to glomerulus filtration.
- e.g. Tetracycline is eliminated mainly by glomerular filtration.

e. Systemic solubility of drug

- Lipoprotein act as vehicle for hydrophobic drugs like steroids, heparin, oil soluble vit.

f. Drug action

- Protein binding inactivates the drugs so sufficient concentration of drug can not be build up in the receptor site for action.
- e.g. Naphthoquinone

g. Sustain release

- The complex of drug protein in the blood act as a reservoir & continuously supply the free drug.
- e.g. Suramin sodium-protein binding for antitrypanosomal action.

h. Diagnosis

- The chlorine atom of chloroquine replaced with radiolabeled I-131 can be used to visualize-melanomas of eye & disorders of thyroid gland.

Binding affinity

Three important notes about binding affinity

- Binding affinity is the concentration of a drug required to occupy 50% of the target molecules at equilibrium. This value quantifies the extent of target occupancy, which is used to predict *in-vivo* efficacy.1
- Residence time is a quantitative value representing the time a drug-receptor interaction takes to reach target occupancy. It is usually calculated as dissociation half time: t1/2 = 0.693/koff, where koff is the dissociation rate of the interaction (Hoare et al., 2019). An interaction is widely considered to be at equilibrium when five dissociation half times have passed.1
- In simplified terms, very slow dissociation (off) rates frequently result in erroneous affinity measurements due to not actually reaching equilibrium. This is just as critical when designing SPR experiments to measure steady-state affinity, but can be circumvented by measuring kinetics directly.

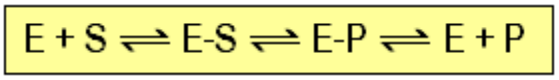
Enzyme inhibitors

Various compounds can reduce the activity of enzymes. They may act in a variety of different ways, and indeed may be reversible or irreversible inhibitors of the enzyme.

On this page there are notes about:

An irreversible inhibitor causes covalent modification of the enzyme, so that its activity is permanently reduced. Compounds that act as irreversible inhibitors are often useful as drugs that need be taken only every few days, although adjusting the dose to suit the patient's response is a lengthy process with such compounds. By contrast, the effect of a reversible inhibitor can be reversed by removing the inhibitor, e.g. by dialysis or gel filtration.

The normal sequence of an enzyme reaction can be represented as:



where:

E	=	enzyme
S	=	substrate
E-S	=	enzyme-substrate complex
E-P	=	enzyme-product complex

P = product

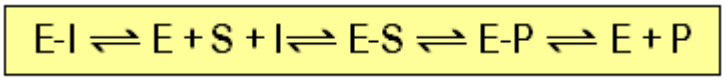
There are three main types of reversible inhibitor:

- competitive inhibitor
- non-competitive inhibitor
- uncompetitive inhibitor

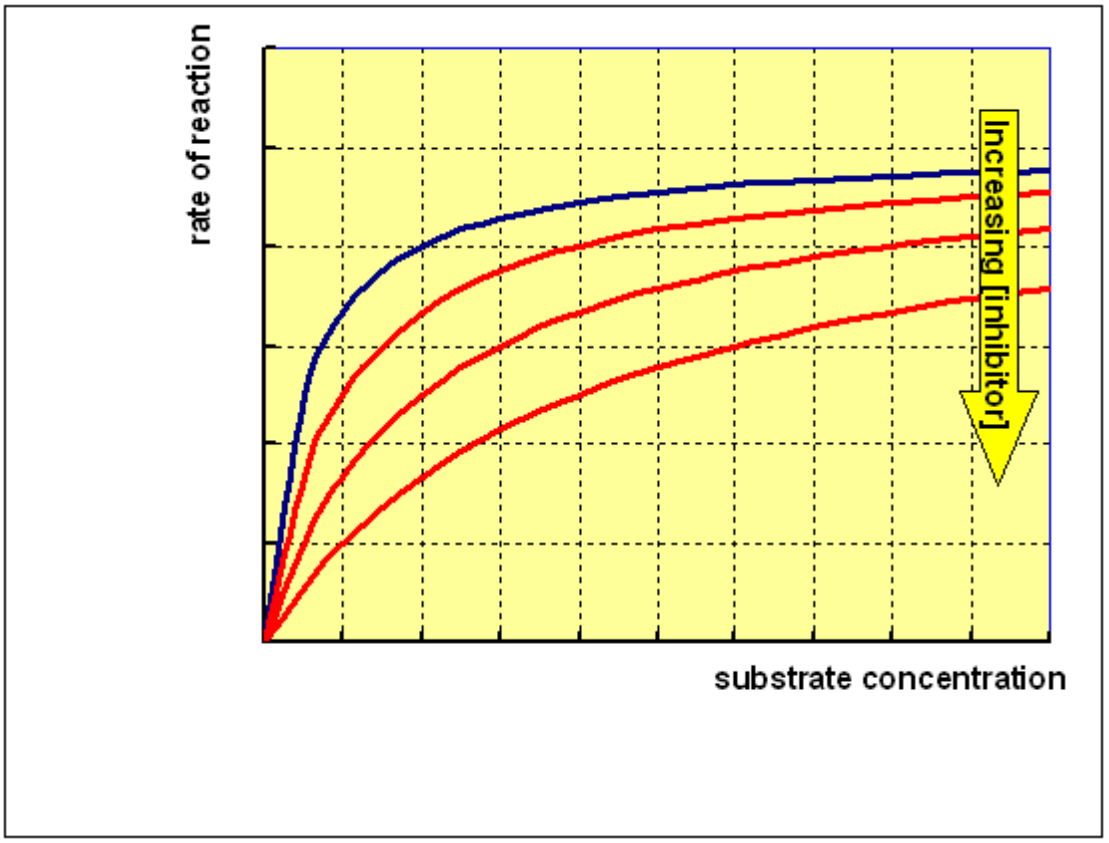
They interact with the enzyme or enzyme-substrate complex at different stages in the sequence

Competitive inhibition

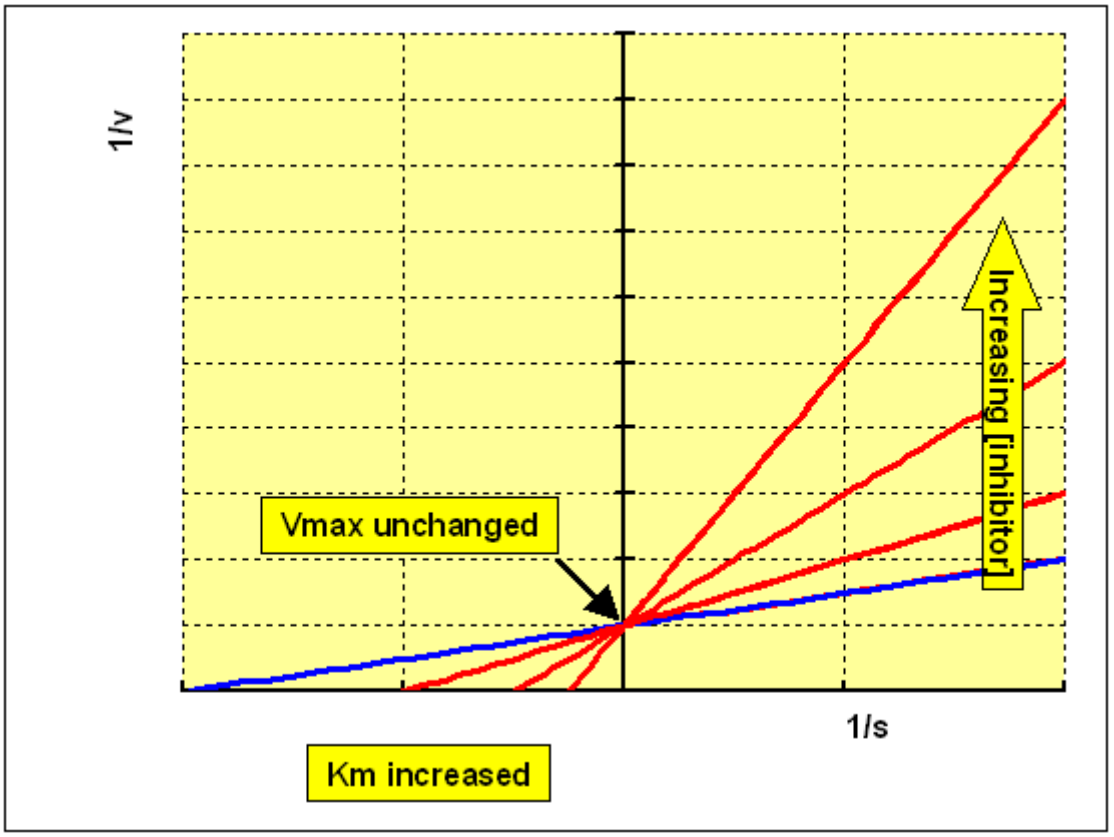
A competitive inhibitor competes with the substrate for the active site of the enzyme:



This means that increasing the concentration of substrate will decrease the chance of inhibitor binding to the enzyme. Hence, if the substrate concentration is high enough the enzyme will reach the same Vmax as without the inhibitor. However, it will require a higher concentration of substrate to achieve this and so the Km of the enzyme will also be higher. Reacting the enzyme with a range of concentrations of substrate at different concentrations of a competitive inhibitor will give a family of curves as shown below:



The Lineweaver-Burk double reciprocal plot for this set of data shows a series of lines crossing the y (1/v) axis at the same point - i.e. Vmax is unchanged, but with a decreasing value of 1/Km (and hence a higher Km) in the presence of the inhibitor:

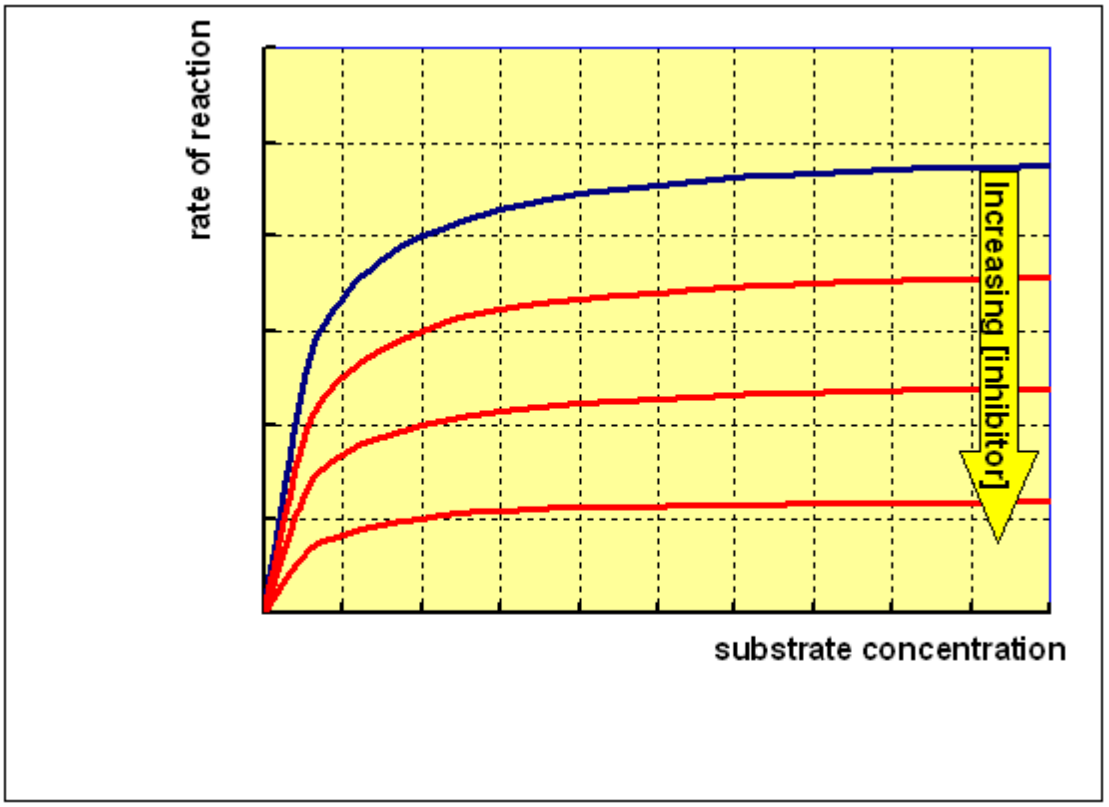


Non-competitive inhibition

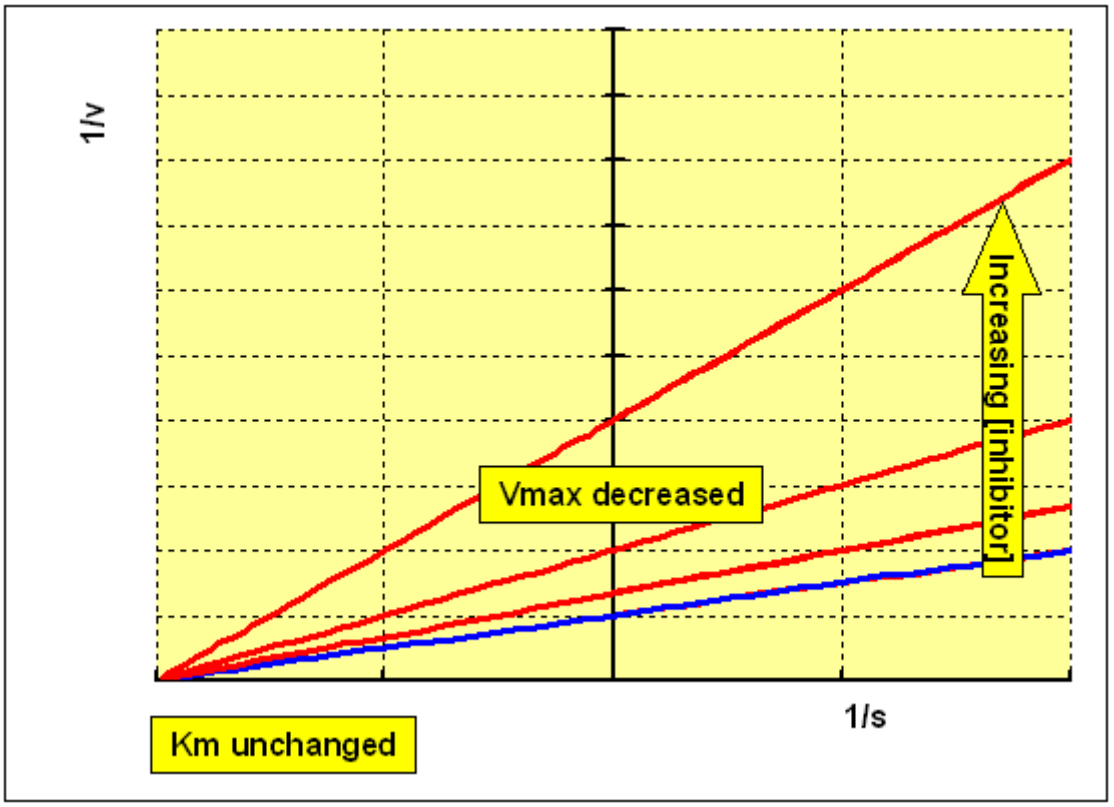
A non-competitive inhibitor reacts with the enzyme-substrate complex, and slows the rate of reaction to form the enzyme-product complex.

$$E + S + I \rightleftharpoons E\text{-}S + I \rightleftharpoons E\text{-}S\text{-}I \rightleftharpoons E\text{-}P\text{-}I \rightleftharpoons E + P + I$$

This means that increasing the concentration of substrate will not relieve the inhibition, since the inhibitor reacts with the enzyme-substrate complex. Reacting the enzyme with a range of concentrations of substrate at different concentrations of a non-competitive inhibitor will give a family of curves as shown below:



The Lineweaver-Burk double reciprocal plot for this set of data shows a series of lines converging on the same point on the X (1/S) axis - i,.e. Km is unchanged, but Vmax is reduced:

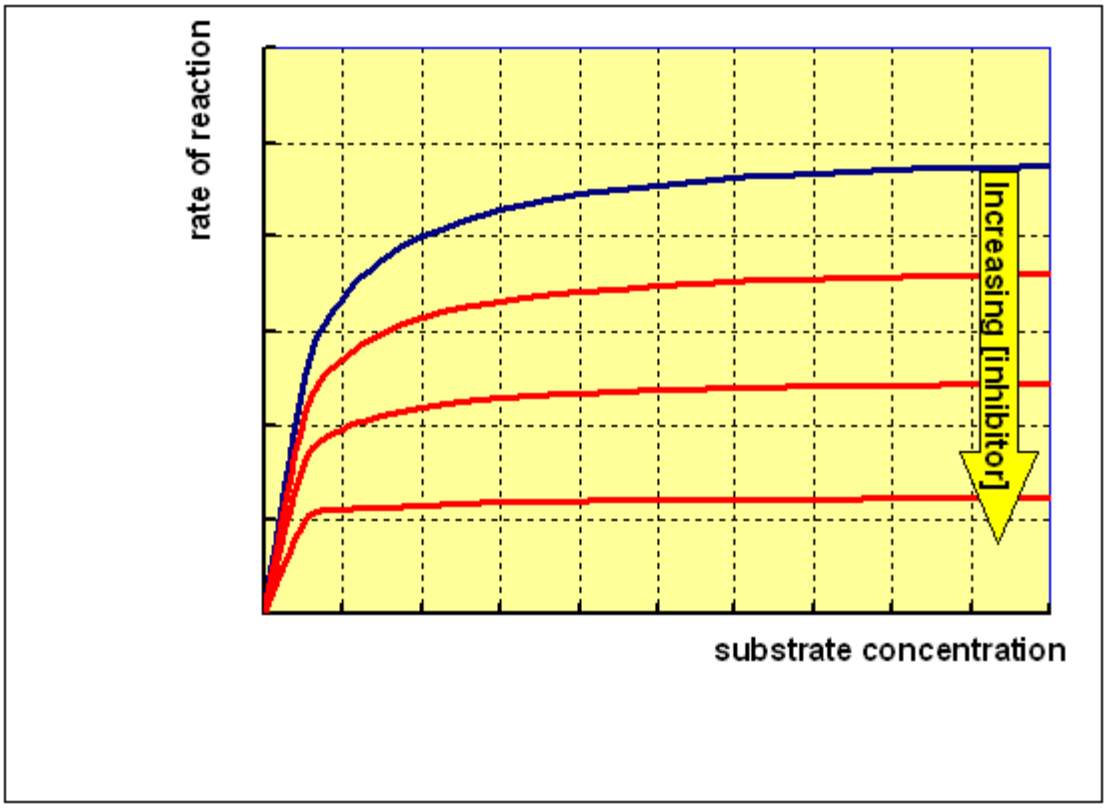


Uncompetitive inhibition

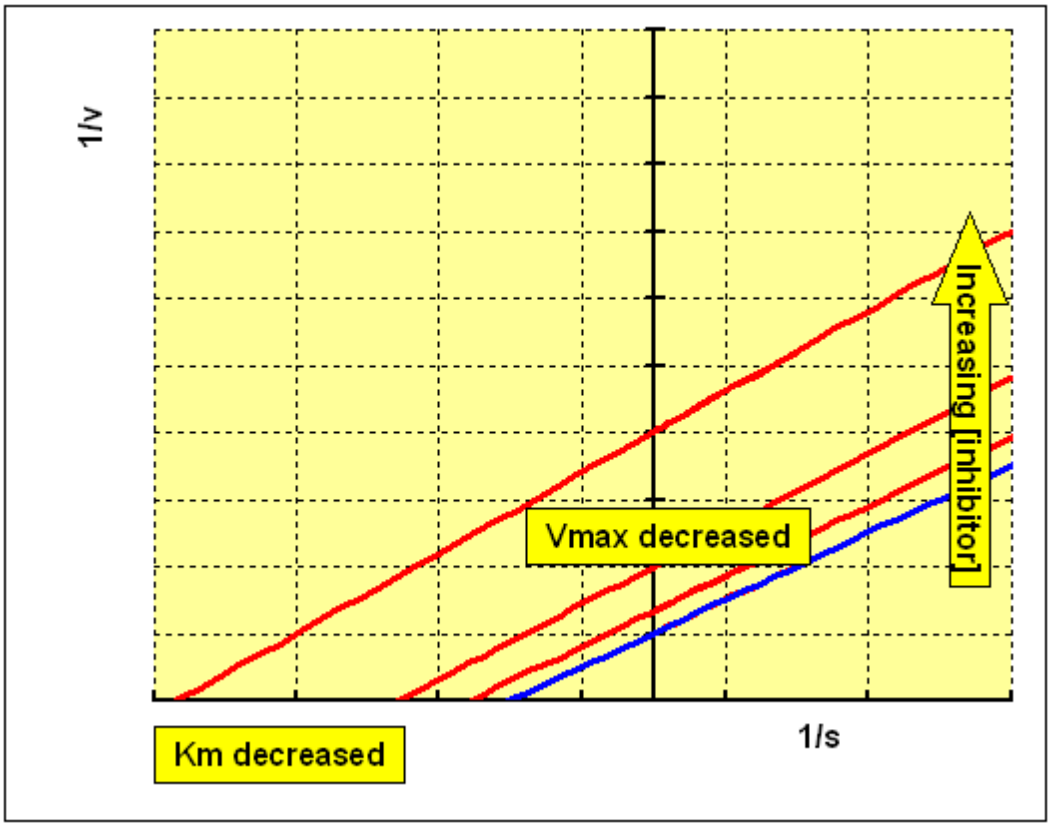
This is a very rare class of inhibition. An uncompetitive inhibitor binds to the enzyme and enhances the binding of substrate (so reducing Km), but the resultant enzyme-inhibitor-substrate complex only undergoes reaction to form the product slowly, so that Vmax is also reduced:

$$E + S + I \rightleftharpoons E\text{-}I + S \rightleftharpoons E\text{-}I\text{-}S \rightleftharpoons E\text{-}I\text{-}P \rightleftharpoons E + P + I$$

Reacting the enzyme with a range of concentrations of substrate at different concentrations of an uncompetitive inhibitor will give a family of curves as shown below:



The Lineweaver-Burk double reciprocal plot for this set of data shows a series of parallel lines - both Km and Vmax are reduced:



The choice of a competitive or non-competitive inhibitor as a drug

If the requirement is to increase the intracellular concentration of the substrate, then either a competitive or non-competitive inhibitor will serve, since both will inhibit the utilisation of substrate, so that it accumulates.

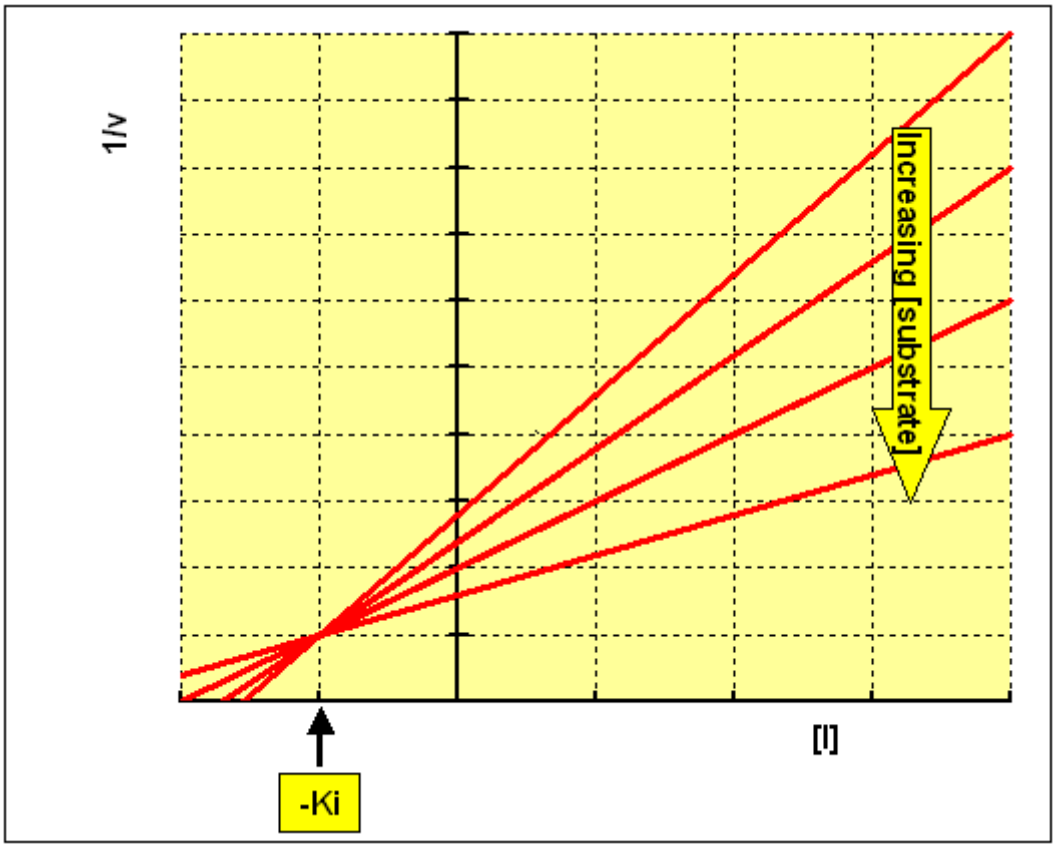
However, if the requirement is to decrease the intracellular concentration of the product, then the inhibitor must be non-competitive. As unused substrate accumulates, so it will compete with a competitive inhibitor, and the final result will be a more or less normal rate of formation of product, but with a larger pool of substrate. Increasing the concentration of substrate does not affect a non-competitive inhibitor.

Ki, the inhibitor constant

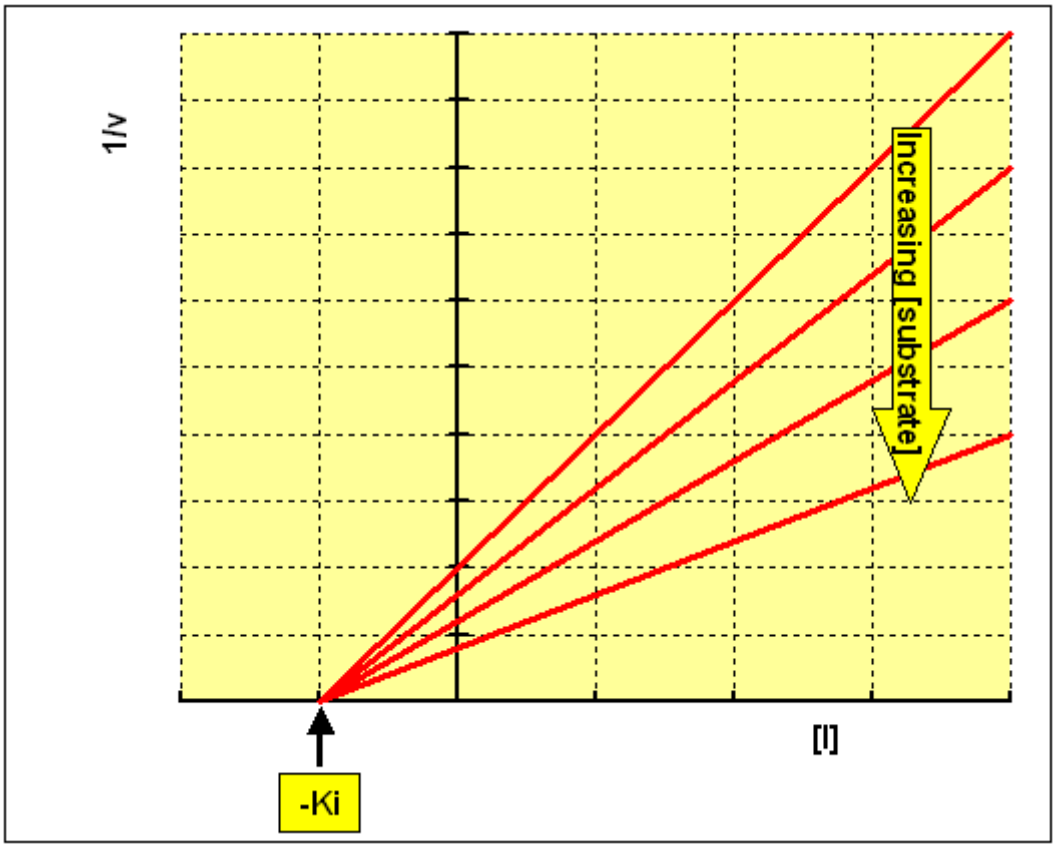
The inhibitor constant, Ki, is an indication of how potent an inhibitor is; it is the concentration required to produce half maximum inhibition.

Plotting 1/v against concentration of inhibitor at each concentration of substrate (the Dixon plot) gives a family of intersecting lines.

For a competitive inhibitor, the lines converge above the x axis, and the value of [I] where they intersect is -Ki



For a non-competitive inhibitor, the lines converge on x axis, and the value of [I] where they intersect is -Ki



UNIT – 4- SBIA5302 – Computer aided drug design

Molecular modelling encompasses all methods, theoretical and computational, used to model or mimic the behaviour of molecules. The methods are used in the fields of computational chemistry, drug design, computational biology and materials science to study molecular systems ranging from small chemical systems to large biological molecules and material assemblies. The simplest calculations can be performed by hand, but inevitably computers are required to perform molecular modelling of any reasonably sized system. The common feature of molecular modelling methods is the atomistic level description of the molecular systems. This may include treating atoms as the smallest individual unit (a molecular mechanics approach), or explicitly modelling protons and neutrons with its quarks, anti-quarks and gluons and electrons with its photons (a quantum chemistry approach).

Molecules can be modelled either in vacuum, or in the presence of a solvent such as water. Simulations of systems in vacuum are referred to as gas-phase simulations, while those that include the presence of solvent molecules are referred to as explicit solvent simulations. In another type of simulation, the effect of solvent is estimated using an empirical mathematical expression; these are termed implicit solvation simulations.

Most force fields are distance-dependent, making the most convenient expression for these Cartesian coordinates. Yet the comparatively rigid nature of bonds which occur between specific atoms, and in essence, defines what is meant by the designation molecule, make an internal coordinate system the most logical representation. In some fields the IC representation (bond length, angle between bonds, and twist angle of the bond as shown in the figure) is termed the Z-matrix or torsion angle representation. Unfortunately, continuous motions in Cartesian space often require discontinuous angular branches in internal coordinates, making it relatively hard to work with force fields in the internal coordinate representation, and conversely a simple displacement of an atom in Cartesian space may not be a straight line trajectory due to the prohibitions of the interconnected bonds. Thus, it is very common for computational optimizing programs to flip back and forth between representations during their iterations. This can dominate the calculation time of the potential itself and in long chain molecules introduce cumulative numerical inaccuracy. While all conversion algorithms produce mathematically identical results, they differ in speed and numerical accuracy.[3] Currently, the fastest and most accurate torsion to Cartesian conversion is the Natural Extension Reference Frame (NERF) method.

Molecular mechanics

Molecular Mechanics Force Field

The "mechanical" molecular model was developed out of a need to describe molecular structures and properties in as practical a manner as possible. The range of applicability of molecular mechanics includes:

Molecules containing thousands of atoms.

Organics, oligonucleotides, peptides, and saccharides (metallo-organics and inorganics in some cases).

Vacuum, implicit, or explicit solvent environments.

Ground state only.

Thermodynamic and kinetic (via molecular dynamics) properties.

The great computational speed of molecular mechanics allows for its use in procedures such as molecular dynamics, conformational energy searching, and docking. All the procedures require large numbers of energy evaluations.

Molecular mechanics methods are based on the following principles:

Nuclei and electrons are lumped into atom-like particles.

Atom-like particles are spherical (radii obtained from measurements or theory) and have a net charge (obtained from theory).

Interactions are based on springs and classical potentials.

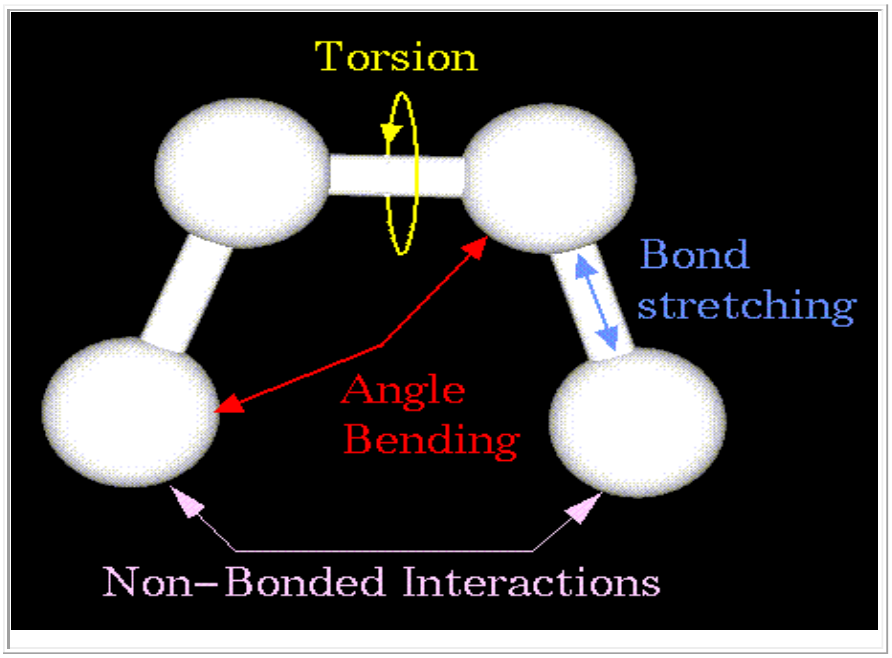
Interactions must be preassigned to specific sets of atoms.

Interactions determine the spatial distribution of atom-like particles and their energies

To define a force field one must specify not only the functional form but also the parameters (i.e.the various constants). Two force fields may use an identical functional form yet have very different parameters. A force field should be considered as a single entity; it is

not strictly correct to divide the energy into its individual components, let alone to take some of the parameters from one forcefield and mix them with parameters from another force field. The forcefields used in molecular modelling are primarily designed to reproduce structural properties but they can also be used to predict other properties, such as molecular spectra. However, molecular mechanics force fields can rarely predict spectra with great accuracy (although the more recent molecular. mechanics force fields are much better in this regard). A force field is generally designed to predict certain properties and will be parametrised accordingly. While it is useful to try to predict other quantities which have not been included in the parametrisation process it is not necessarily a failing if a force field is unable to do so. Transferability of the functional form and parameters is an important feature of a forcefield. Transferability means that the same set of parameters can be used to model a series of related molecules, rather than having to define a new set of parameters for each individual molecule. A concept that is common to most force fields is that of an atom type. When preparing the input for a quantum mechanics calculation it is usually necessary to specify the atomic numbers of the nuclei present, together with the geometry of the system and the overall charge and spin multiplicity. For a force field the overall charge and spin multiplicity are not explicitly required, but it is usually necessary to assign an atom type to each atom in the system. The atom type is more than just the atomic number of an atom; it usually con• tains information about its hybridisation state and sometimes the local environment. For example, it is necessary in most force fields to distinguish between sp3 - hybridised carbon atoms (which adopt a tetrahedral geometry), sp2-hybridised carbons (which are trigonal) and sp-hybridised carbons (which are linear).

The mechanical molecular model considers atoms as spheres and bonds as springs. The mathematics of spring deformation can be used to describe the ability of bonds to stretch, bend, and twist:



Non-bonded atoms (greater than two bonds apart) interact through van der Waals attraction, steric repulsion, and electrostatic attraction/repulsion. These properties are easiest to describe mathematically when atoms are considered as spheres of characteristic radii.

The object of molecular mechanics is to predict the energy associated with a given conformation of a molecule. However, molecular mechanics energies have no meaning as absolute quantities. Only differences in energy between two or more conformations have meaning. A simple molecular mechanics energy equation is given by:

Energy = Stretching Energy + Bending Energy + Torsion Energy + Non-Bonded Interaction Energy

- A **force field** refers to the form and parameters of mathematical functions used to describe the potential energy of a system of particles (typically molecules and atoms).
- calculates the molecular system's potential energy (E) in a given conformation as a sum of individual energy terms.
- where the components of the covalent and noncovalent contributions are given by the following summations:

$$E_{\text{noncovalent}} = E_{\text{electrostatic}} + E_{\text{van der Waals}}$$

- where the components of the covalent and noncovalent contributions are given by the following summations

$$E_{\text{covalent}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{dihedral}}$$

$$E_{\text{noncovalent}} = E_{\text{electrostatic}} + E_{\text{van der Waals}}$$

- FF is a mathematical function which returns the energy of the system as a function of the conformation of the system.

$$\mathscr{V}(\mathbf{r}^N) = \sum_{\text{bonds}} \frac{k_i}{2}(l_i - l_{i,0})^2 + \sum_{\text{angles}} \frac{k_i}{2}(\theta_i - \theta_{i,0})^2 + \sum_{\text{torsions}} \frac{V_n}{2}(1 + \cos(n\omega - \gamma)) + \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right)$$

$\mathscr{V}(\mathbf{r}^N)$ Potential energy as a function of position r of N particles

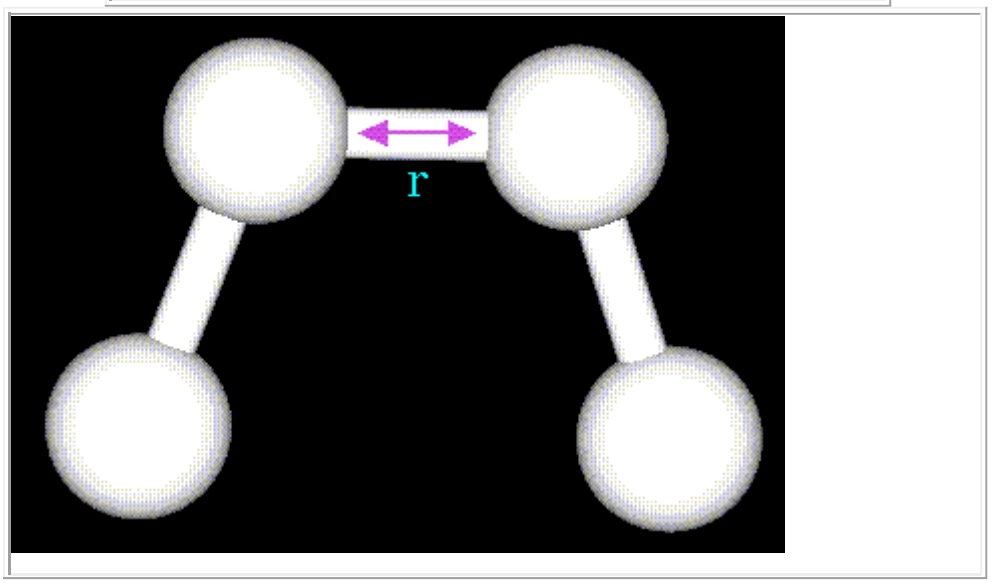
- Reproduce the structural properties such as molecular spectra
- Transferability

These equations together with the data (parameters) required to describe the behavior of different kinds of atoms and bonds, is called a force-field. Many different kinds of force-fields have been developed over the years. Some include additional energy terms that describe other kinds of deformations. Some force-fields account for coupling between bending and stretching in adjacent bonds in order to improve the accuracy of the mechanical model.

The mathematical form of the energy terms varies from force-field to force-field. The more common forms will be described.

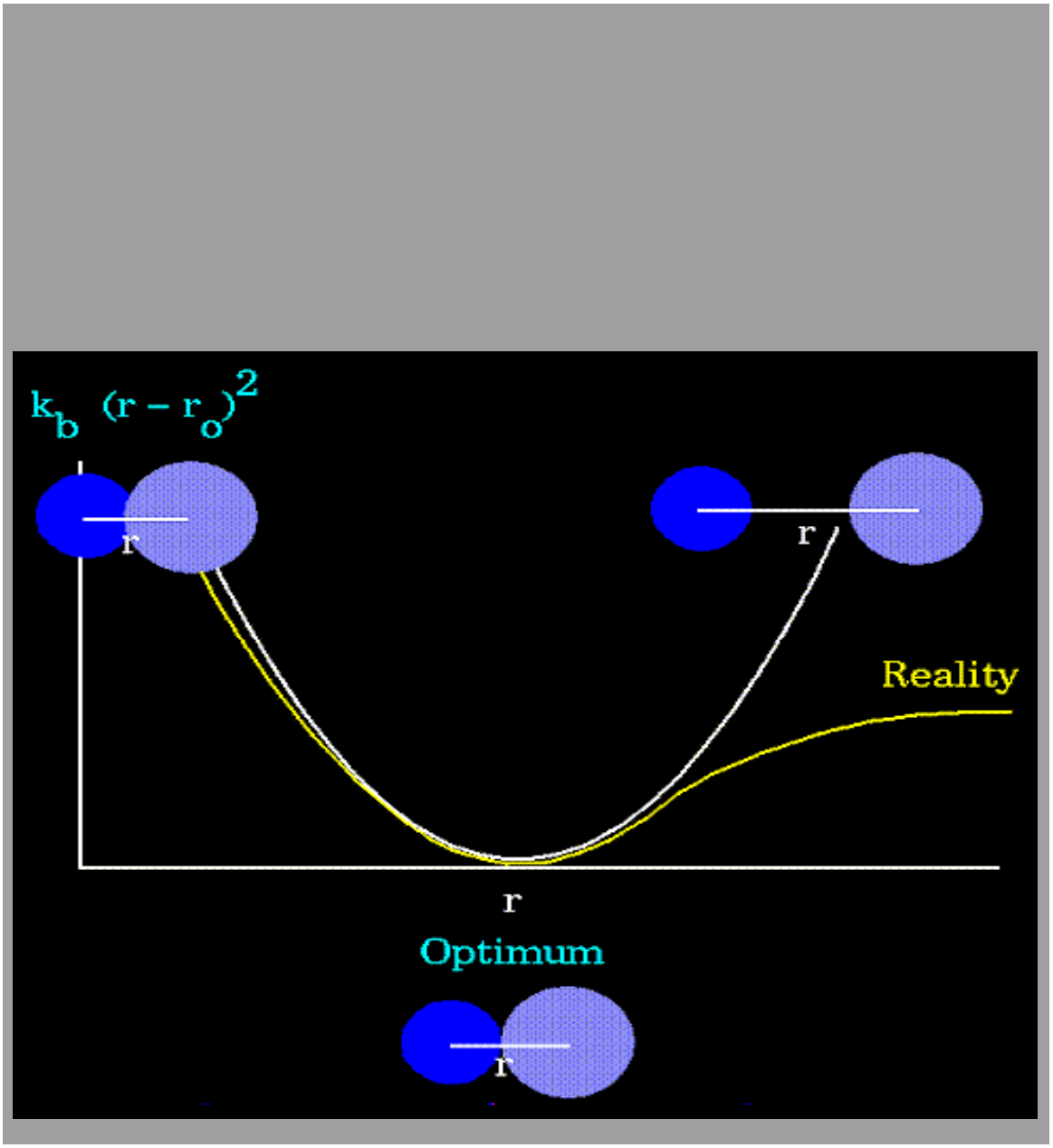
Stretching Energy

$$E = \sum_{bonds} k_b (r - r_o)^2$$



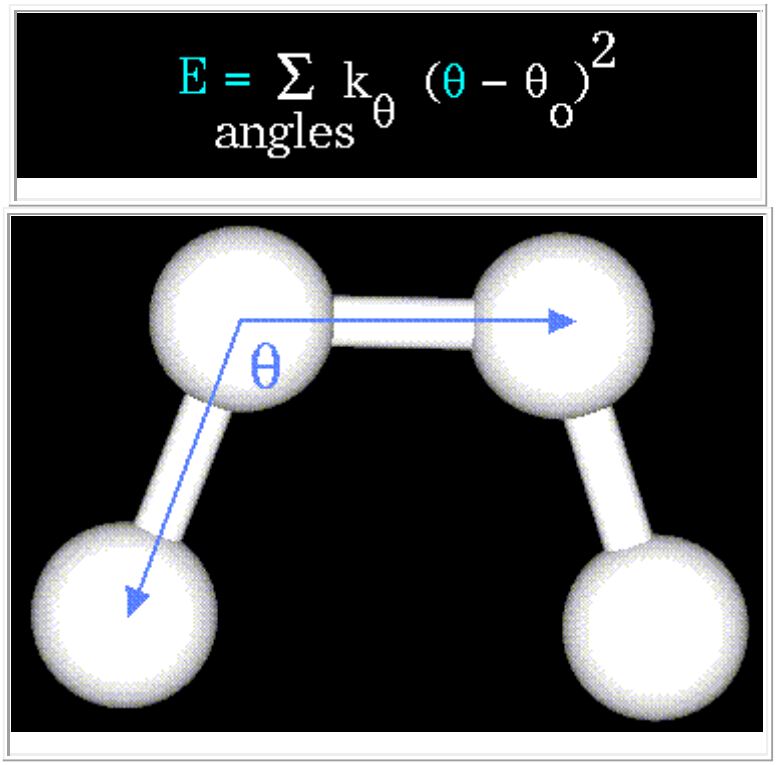
The stretching energy equation is based on Hooke's law. The "kb" parameter controls

the stiffness of the bond spring, while "ro" defines its equilibrium length. Unique "kb" and "ro" parameters are assigned to each pair of bonded atoms based on their types (e.g. C-C, C-H, O-C, etc.). This equation estimates the energy associated with vibration about the equilibrium bond length. This is the equation of a parabola, as can be seen in the following plot

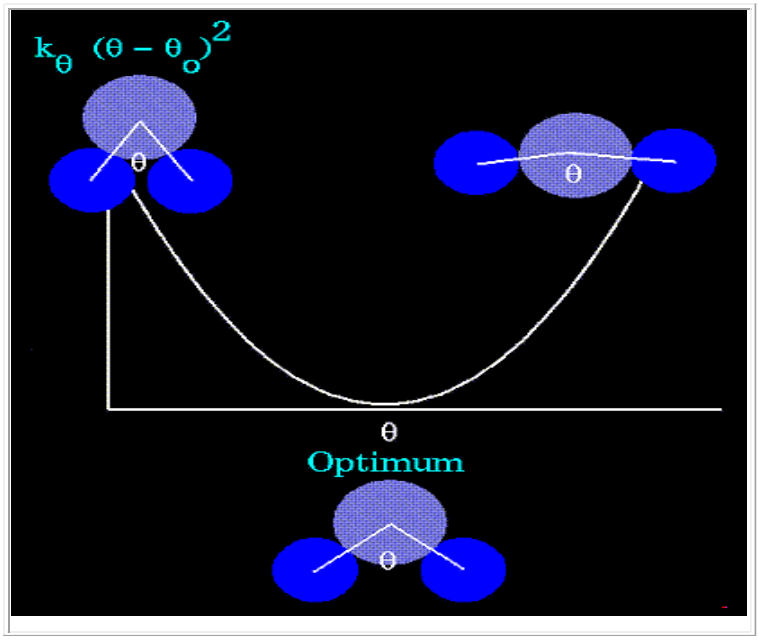


Notice that the model tends to break down as a bond is stretched toward the point of dissociation.

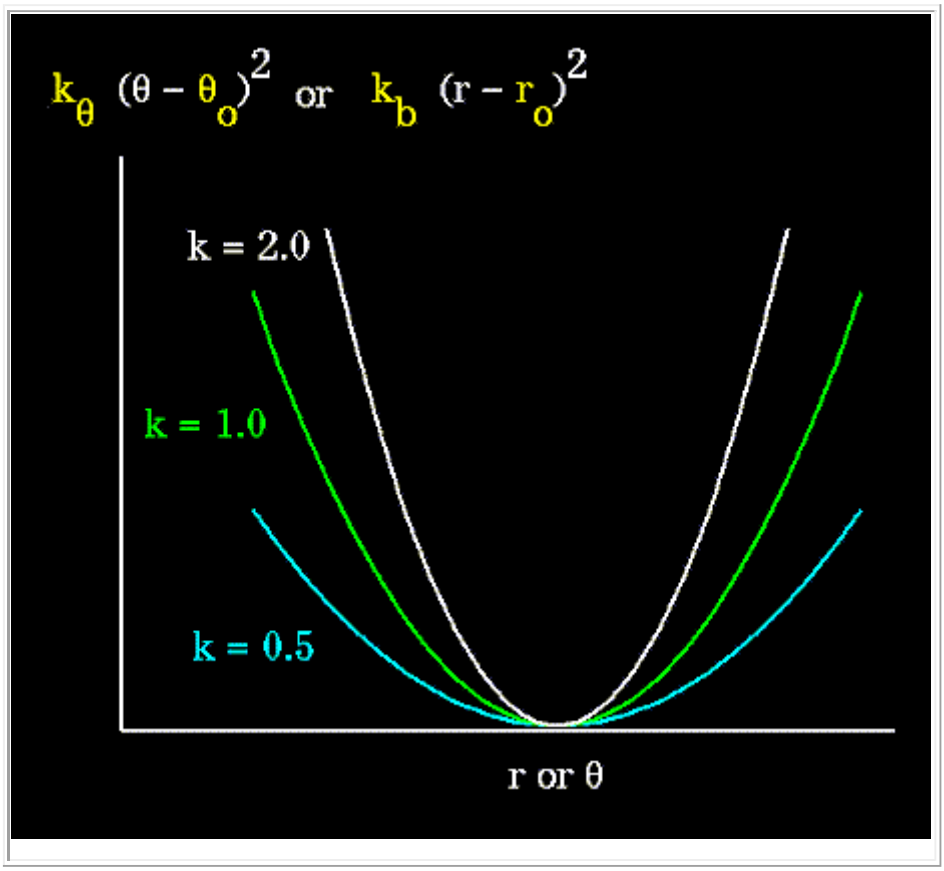
Bending Energy



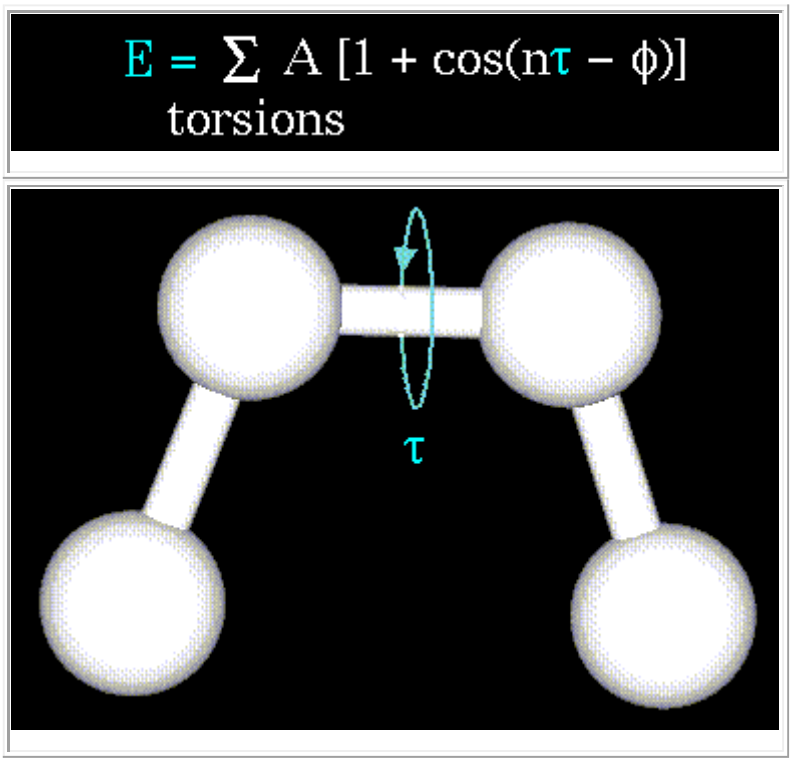
The bending energy equation is also based on Hooke's law. The "ktheta" parameter controls the stiffness of the angle spring, while "thetao" defines its equilibrium angle. This equation estimates the energy associated with vibration about the equilibrium bond angle:



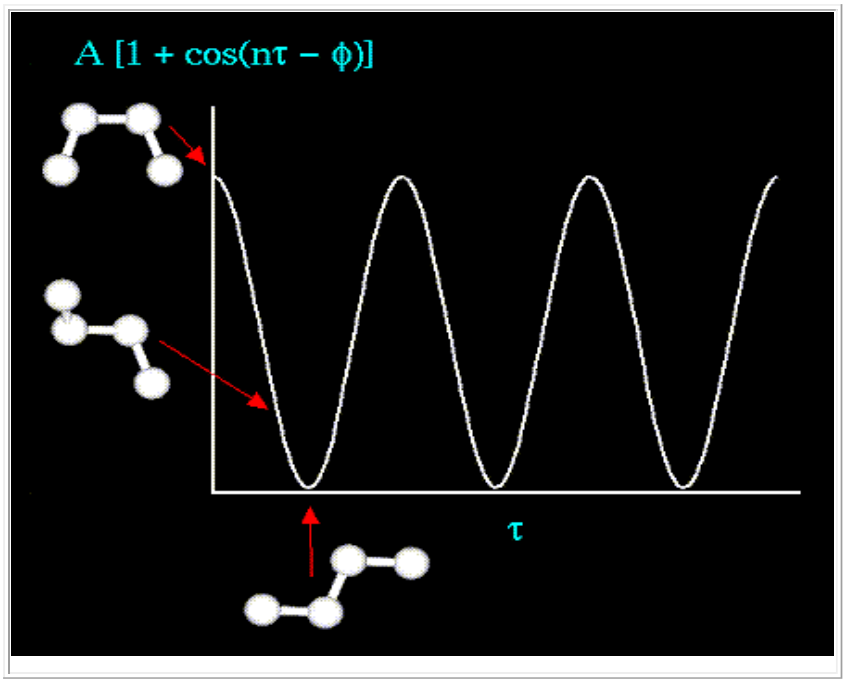
Unique parameters for angle bending are assigned to each bonded triplet of atoms based on their types (e.g. C-C-C, C-O-C, C-C-H, etc.). The effect of the "kb" and "ktheta" parameters is to broaden or steepen the slope of the parabola. The larger the value of "k", the more energy is required to deform an angle (or bond) from its equilibrium value. Shallow potentials are achieved for "k" values between 0.0 and 1.0. The Hookeian potential is shown in the following plot for three values of "k":



Torsion Energy



The torsion energy is modeled by a simple periodic function, as can be seen in the following plot:

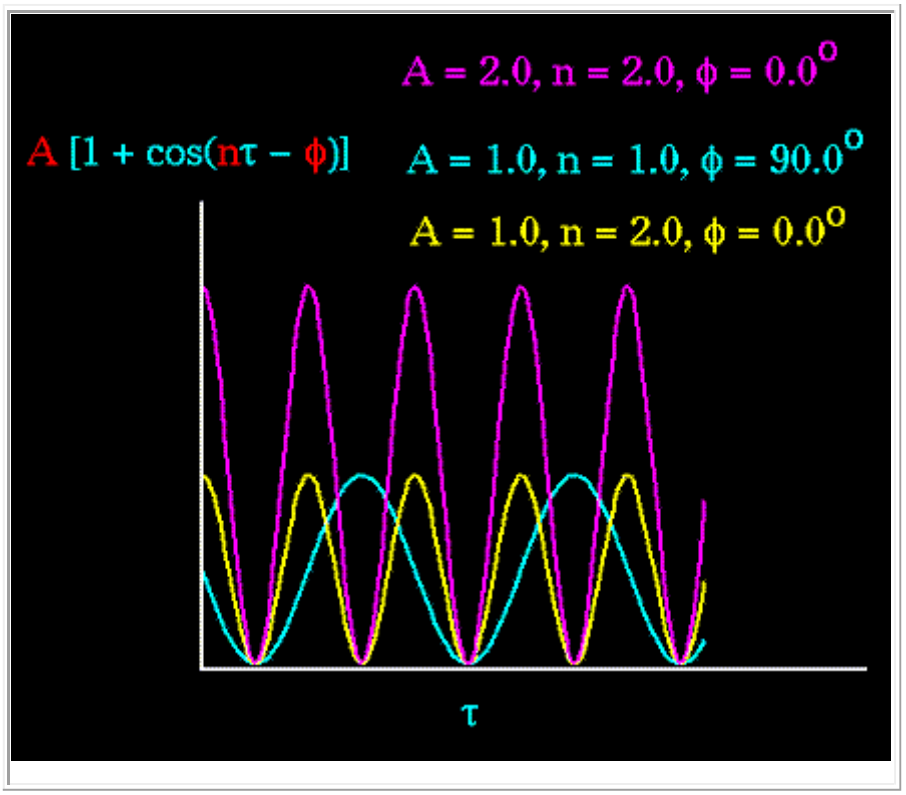


The torsion energy in molecular mechanics is primarily used to correct the remaining energy terms rather than to represent a physical process. The torsional energy represents the amount of energy that must be added to or subtracted from the Stretching Energy + Bending Energy

+ Non-Bonded Interaction Energy terms to make the total energy agree with experiment or rigorous quantum mechanical calculation for a model dihedral angle (ethane, for example

might be used a a model for any H-C-C-H bond).

The "A" parameter controls the amplitude of the curve, the n parameter controls its periodicity, and "phi" shifts the entire curve along the rotation angle axis (tau). The parameters are determined from curve fitting. Unique parameters for torsional rotation are assigned to each bonded quartet of atoms based on their types (e.g. C-C-C-C, C-O-C-N, H-C-C-H, etc.). Torsion potentials with three combinations of "A", "n", and "phi" are shown in the following plot:



Notice that "n" reflects the type symmetry in the dihedral angle. A CH3-CH3 bond, for example, ought to repeat its energy every 120 degrees. The cis conformation of a dihedral

angle is assumed to be the zero torsional angle by convention. The parameter phi can be used to synchronize the torsional potential to the initial rotameric state of the molecule whose energy is being computed.

Cross terms

The presence of cross terms in a forcefield reflects coupling between the internal coordinates. For example, as a bond angle is decreased it is found that the adjacent bonds stretch to reduce the interaction between the 1,3 atoms, as illustrated in Figure.

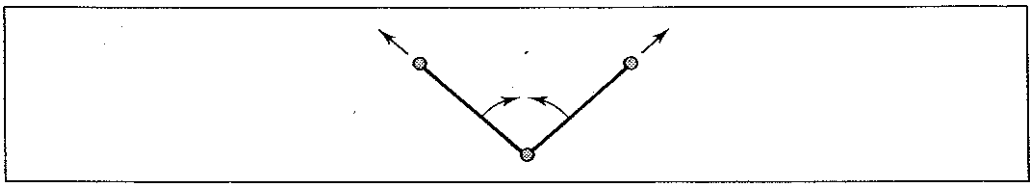
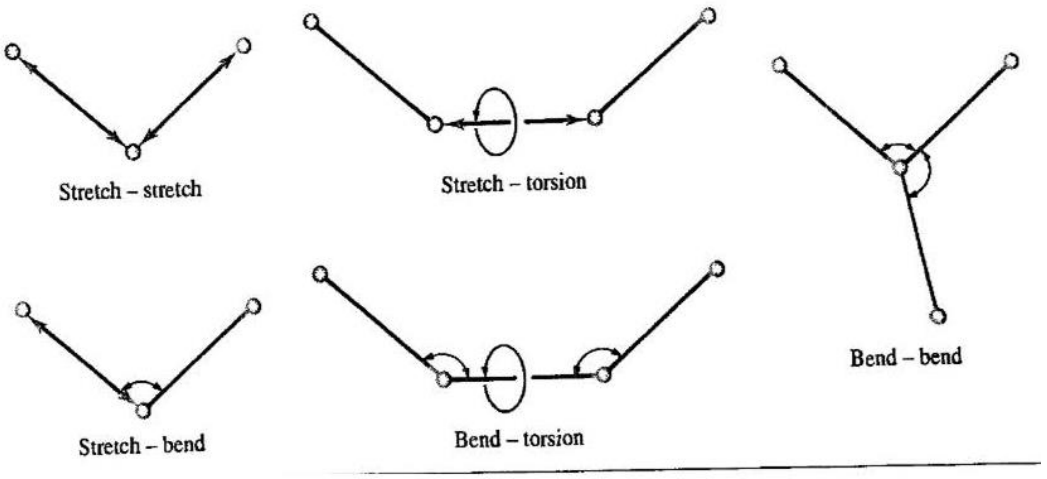


Fig. 4.12: Coupling between the stretching of the bonds as an angle closes.

One should in principle include cross terms between all contributions to a force field. However, only a few cross terms are generally found to be necessary in order to reproduce structural properties accurately; more may be needed to reproduce other properties such as vibrational frequencies, which are more sensitive to the presence of such terms. In general, any interactions involving motions that are far apart in a molecule can usually be set to zero. Most cross terms are functions of two internal coordinates, such as stretch-stretch, stretch-bend and stretch-torsion terms, but cross terms involving more than two internal coordinates such as the bend- bend- torsion have also been used.

Cross terms



Various functional forms are possible for the cross terms. For example,the stretch- stretch cross term between two bonds 1 and 2 can be modelled as:

$$v(l_1, l_2) = \frac{k_{l_1,l_2}}{2}[(l_1 - l_{1,0})(l_2 - l_{2,0})] \quad (4.13)$$

The stretching of the two bonds adjoining an angle could be modelled using an equation of the following form (as in MM2, MM3 and MM4):

$$v(l_1, l_2, \theta) = \frac{k_{l_1,l_2,\theta}}{2}[(l_1 - l_{1,0}) + (l_2 - l_{2,0})](\theta - \theta_0) \quad (4.14)$$

Non-Bonded Energy

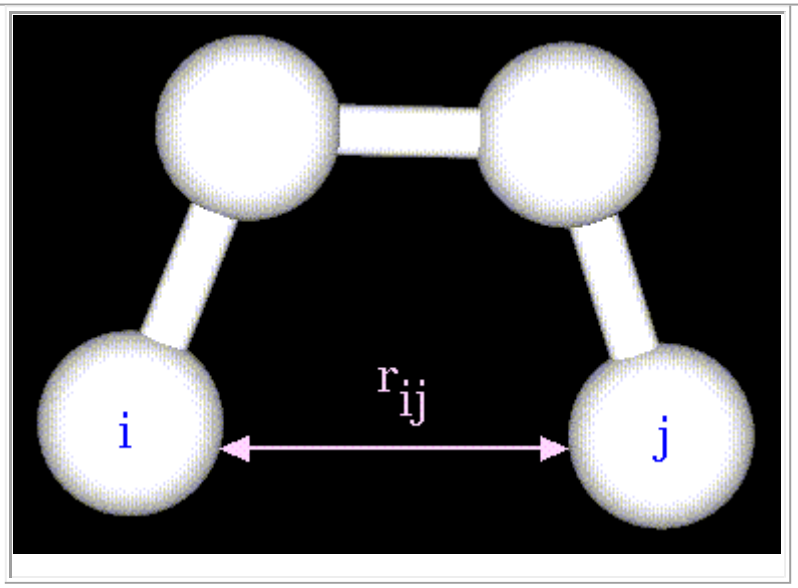
Independent molecules and atoms interact through non-bonded forces, which also play an important role in determining the structure of individual molecular species. The non-bonded interactions do not depend upon a specific bonding relationship between atoms. They are 'through-space' interactions and are usually modelled as a function of some inversepower of the distance. The non-bonded terms in a forcefield

are usually considered in two groups, one comprising electrostatic interactions and the other van der Waals interactions.

The non-bonded energy represents the pair-wise sum of the energies of all possible interacting non-bonded atoms i and j:

$$E = \sum_i \sum_j \frac{-A_{ij}}{r_{ij}^6} + \frac{B_{ij}}{r_{ij}^{12}} + \sum_i \sum_j \frac{q_i q_j}{r_{ij}}$$

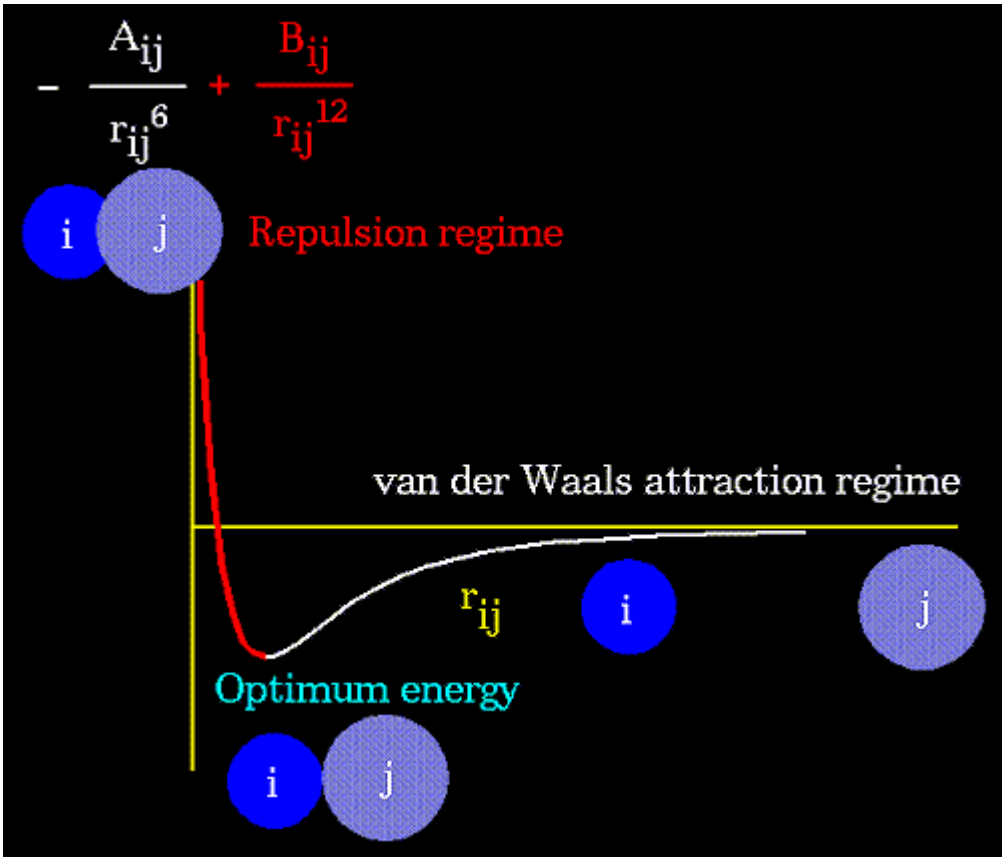
van der Waals term Electrostatic term



The non-bonded energy accounts for repulsion, van der Waals attraction, and electrostatic interactions.

Van der Waals attraction occurs at short range, and rapidly dies off as the interacting atoms move apart by a few Angstroms. Repulsion occurs when the distance between interacting atoms becomes even slightly less than the sum of their contact radii. Repulsion is modeled by an equation that is designed to rapidly blow up at close distances. The energy term that describes attraction/repulsion provides for a smooth transition between these two regimes. These effects are often modeled using a 6-12 equation, as shown in the following plot:

The "A" and "B" parameters control the depth and position (interatomic distance) of the potential energy well for a given pair of non-bonded interacting atoms (e.g. C:C, O:C, O:H, etc.). In effect, "A" determines the degree of "stickiness" of the van der Waals attraction and "B" determines the degree of "hardness" of the atoms (e.g marshmallow-like, billiard ball-like, etc.).


$-\frac{A_{ij}}{r_{ij}^{6}}+\frac{B_{ij}}{r_{ij}^{12}}$
i
j
Repulsion regime
van der Waals attraction regime
r_{ij}
i
j
Optimum energy
i
j

Vanderwaals interaction

- Dispersive interactions- long range attractive forces
- Due to instantaneous dipoles which arise due to fluctuation in electron clouds
- This can induce a dipole in neighboring atoms – giving rise to an attractive inductive effect

A simple model to explain the dispersive interaction was proposed by Drude. This model consists of 'molecules' with two charges, $+q$ and $-q$, separated by a distance r. The negative charge performs simple harmonic motion with angular frequency ω along the z axis about the stationary positive charge (Figure 4.33). If the force constant for the oscillator is k and if the mass of the oscillating charge is m, then the potential energy of an isolated Drude molecule is $\frac{1}{2}kz^2$, where z is the separation of the two charges. ω is related to the force constant by $\omega = \sqrt{k/m}$. The Schrödinger equation for a Drude molecule is:

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial z^2} + \frac{1}{2}kz^2\psi = E\psi \qquad (4.59)$$

This is the Schrödinger equation for a simple harmonic oscillator. The energies of the system are given by $E_\nu = (\nu + \frac{1}{2}) \times \hbar\omega$ and the zero-point energy is $\frac{1}{2}\hbar\omega$.

Electrostatic interactions

Electrostatic interactions also arise from changes in the charge distribution of a molecule or atom caused by an external field, a process called polarisation. The primary effect of the external electric field (which in our case will be caused by neighbouring molecules) is to induce a dipole in the molecule. The magnitude of the induced dipole moment μind is proportional to the electric field E, with the constant of proportionality being the polarisability a:

$$\boldsymbol{\mu}_{\text{ind}} = \alpha\mathbf{E} \qquad (4.51)$$

The energy of interaction between a dipole $\boldsymbol{\mu}_{\text{ind}}$ and an electric field $\mathbf{E}$ (the induction energy) is determined by calculating the work done in charging the field from zero to E, using the following integral:

$$v(\alpha, E) = -\int_0^E d\mathbf{E}\,\boldsymbol{\mu}_{\text{ind}} = -\int_0^E d\mathbf{E}\,\alpha\mathbf{E} = -\tfrac{1}{2}\alpha E^2 \qquad (4.52)$$

In strong electric fields contributions to the induced dipole moment that are proportional to E^2 or E^3 can also be important, and higher-order moments such as quadrupoles can also be induced. We will not be concerned with such contributions.

The electrostatic contribution is modeled using a Coulombic potential. The electrostatic energy is a function of the charge on the non-bonded atoms, their interatomic distance, and a molecular dielectric expression that accounts for the attenuation of electrostatic interaction by the environment (e.g. solvent or the molecule itself). Often, the molecular dielectric is set to a constant value between 1.0 and 5.0. A linearly varying distance-dependent dielectric (i.e. 1/r) is

sometimes used to account for the increase in environmental bulk as the separation distance between interacting atoms increases.

- Central multipole expansion
  - Electronegative elements attract electrons
  - Unequal charge distribution – fractional point charges through out the mol
  - Charges – produce the electrostatic potential
  - Charges restricted to nuclear centres – partial atomic charges

often referred to as *partial atomic charges* or *net atomic charges*. The electrostatic interaction between two molecules (or between different parts of the same molecule) is then calculated as a sum of interactions between pairs of point charges, using Coulomb's law:

$$\mathscr{V} = \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \qquad (4.19)$$

N_A and N_B are the numbers of point charges in the two molecules. This approach to the

Conformational analysis

The most important concerns in Medicinal chemistry and pharmaceutical research are structure elucidation, conformational analysis, physicochemical characterization and biological activity determination. The determination of molecular structure is essential as the structure of the molecule predicts the physical, chemical, and biological properties of the molecule.

Conformational search methods find applications in the design of targeted chemical hosts and drug discovery [2]. Conformations are different 3D spatial arrangements of the atoms in a molecule are interconvertible by free rotation of single bonds [3].

The major objective of conformational analysis is to gain insight on conformational characteristic of flexible biomolecules and drugs but to also identify the relation between the role of conformational flexibility and their activity. Therefore, it plays a significant role in computer aided design as well. The significance of conformational analysis not just extends to computational docking and screening but also for lead optimization

Conformational Analysis

- Properties of molecules depend on their three-dimensional structures (i.e. conformations)
- Conformational analysis is the study of the conformations of a molecule and their influence on its properties
- Conformational analysis is used in drug design to search conformations of small molecules (putative drugs)
- In protein folding this is used to find protein 3D structure with minimal energy that usually corresponds to biologically active structure
- A key component of conformational analysis is the conformational search, the objective of which is to identify 'preferred' conformations, i.e. conformations with low energies

Conformation Searching Methods

- The following algorithms are used to find global minima on conformational potential surface:
  - Systematic search (also called 'grid search', 'exhaustive search' or 'brute force search')
  - Random search (Monte Carlo search)
  - Simulated annealing
  - Genetic algorithms
  - Distance-geometry algorithms
  - The fragment approach
  - Chain growth
  - Rule-based systems

Conformational Analysis

- Conformation generally means structural arrangement
- Conformational analysis is needed to identify the ideal conformation of a molecule

$$N = \left(\frac{360}{\delta}\right)^{nbonds}$$

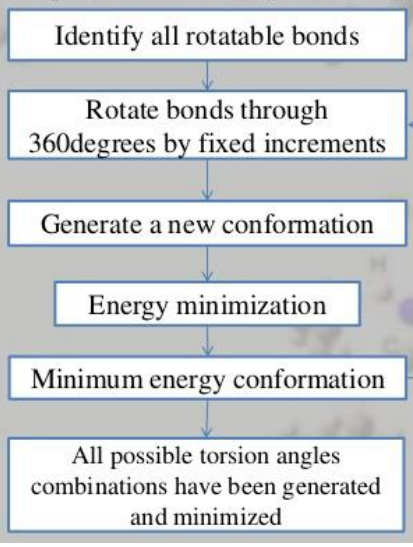
N = # conformations
δ = rotation increment in degrees
nbonds = # of rotatable bonds (degrees of freedom)

- The biological activity of molecules is strongly dependent on their conformation
- Done by exploring the energy surface of a molecule and determining the conformation with minimum energy
- Needed:
  - Conformational space
  - Search method
  - An energy determination method

Systematic Method

- Simplest
- Deterministic
- *Grid search*

Assumption: All bond lengths and angles remain fixed throughout the calculation

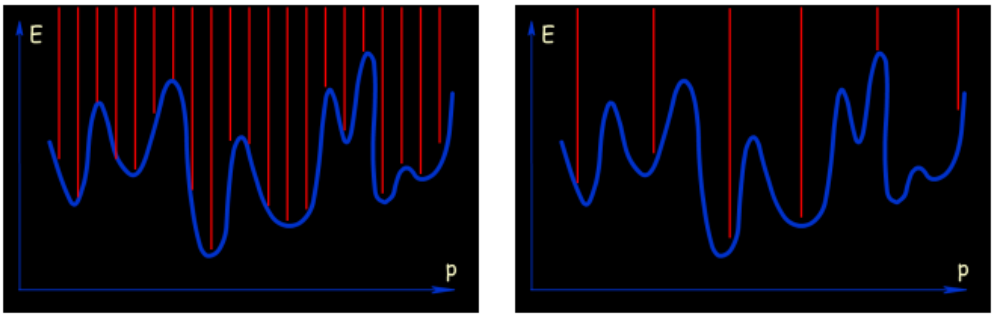


Advantages

- A deterministic approach, hence a systematic exploration of conformational space
- Can be used for small molecules or systems with 10-15 bonds

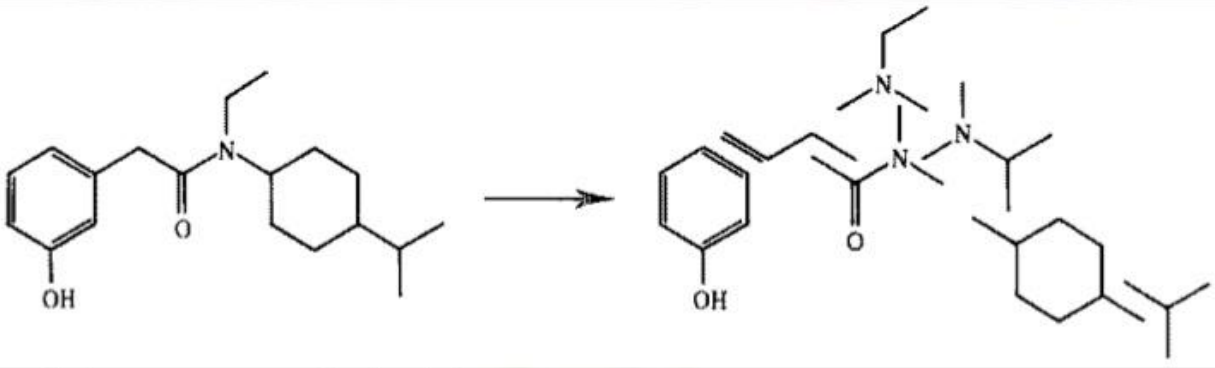
Systematic Search

- A simple method for exploring the conformational potential surface is systematic scanning of all geometries of the molecule
- Within this systematic searches dihedral angles are changed systematically by specified increment value
- Too big increment values can result in missing the global minima while small increment values significantly increase algorithm complexity
- Complexity of systematic search grows rapidly with number of rotatable bonds (~$(360/m)^n$ where n is number of rotatable bonds and m is increment value)



Model-building method

- Construct conformations of molecules by joining together three-dimensional structures of molecular fragments



- Substructure search algorithm
  - Substructure searching is finding a mapping for a query to a target molecule.
  - In other words, no bonds are broken and no new bonds are formed.

Energy minimization

In the field of computational chemistry, energy minimization (also called energy optimization, geometry minimization, or geometry optimization) is the process of finding an arrangement in space of a collection of atoms where, according to some computational model of chemical bonding, the net inter-atomic force on each atom is acceptably close to zero and the

position on the potential energy surface (PES) is a stationary point. The collection of atoms might be a single molecule, an ion, a condensed phase, a transition state or even a collection of any of these. The computational model of chemical bonding might, for example, be quantum mechanics.

The motivation for performing a geometry optimization is the physical significance of the obtained structure: optimized structures often correspond to a substance as it is found in nature and the geometry of such a structure can be used in a variety of experimental and theoretical investigations in the fields of chemical structure, thermodynamics, chemical kinetics, spectroscopy and others.

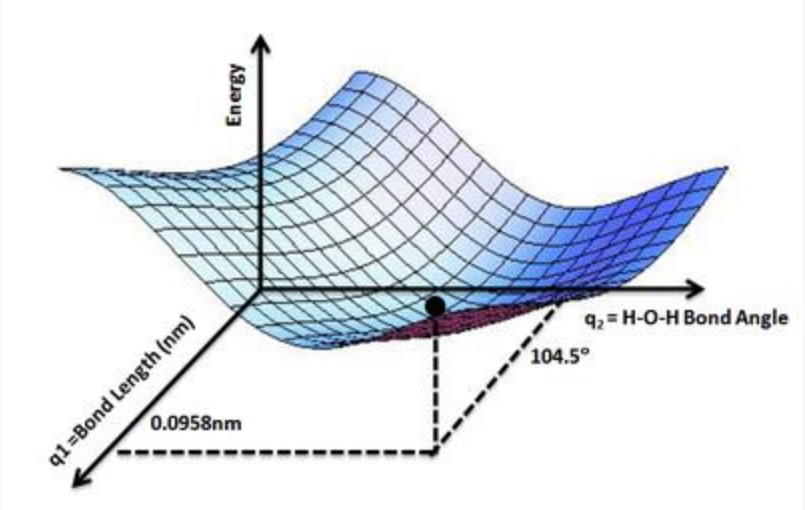
Typically, but not always, the process seeks to find the geometry of a particular arrangement of the atoms that represents a local or global energy minimum. Instead of searching for global energy minimum, it might be desirable to optimize to a transition state, that is, a saddle point on the potential energy surface. Additionally, certain coordinates (such as a chemical bond length) might be fixed during the optimization.

- Energy minimization methods can precisely locate minimum energy conformations by mathematically "homing in" on the energy function minima (one at a time).
- The goal of energy minimization is to find a route (consisting of variation of the intramolecular degrees of freedom) from an initial conformation to the nearest minimum energy conformation using the smallest number of calculations possible.
- The way in which the energy varies with the coordinates is usually referred to as PES or hyper surface
- Energy of any conformation is a function of its internal or cartesian coordinates
- N atoms – energy is a function of 3N-6 internal coordinates or 3N cartesian coordinates
- Changes in the energy are a function of its nuclear coordinates.

Potential energy Surface

A **potential energy surface** (**PES**) describes the energy of a system, especially a collection of atoms, in terms of certain parameters, normally the positions of the atoms. The surface might define the energy as a function of one or more coordinates; if there is only one coordinate, the surface is called a *potential energy curve* or energy profile. An example is the Morse/Long-range potential.

It is helpful to use the analogy of a landscape: for a system with two degrees of freedom (e.g. two bond lengths), the value of the energy (analogy: the height of the land) is a function of two bond lengths (analogy: the coordinates of the position on the ground).[1]



PES for water molecule: Shows the energy minimum corresponding to optimized molecular structure for water- O-H bond length of 0.0958nm and H-O-H bond angle of 104.5°

The PES concept finds application in fields such as chemistry and physics, especially in the theoretical sub-branches of these subjects. It can be used to theoretically explore properties of structures composed of atoms, for example, finding the minimum energy shape of a molecule or computing the rates of a chemical reaction.

The geometry of a set of atoms can be described by a vector, **r**, whose elements represent the atom positions. The vector **r** could be the set of the Cartesian coordinates of the atoms, or could also be a set of inter-atomic distances and angles.

Given **r**, the energy as a function of the positions, $E(\mathbf{r})$, is the value of $E(\mathbf{r})$ for all **r** of interest. Using the landscape analogy from the introduction, *E* gives the height on the "energy landscape" so that the concept of a potential energy *surface* arises.

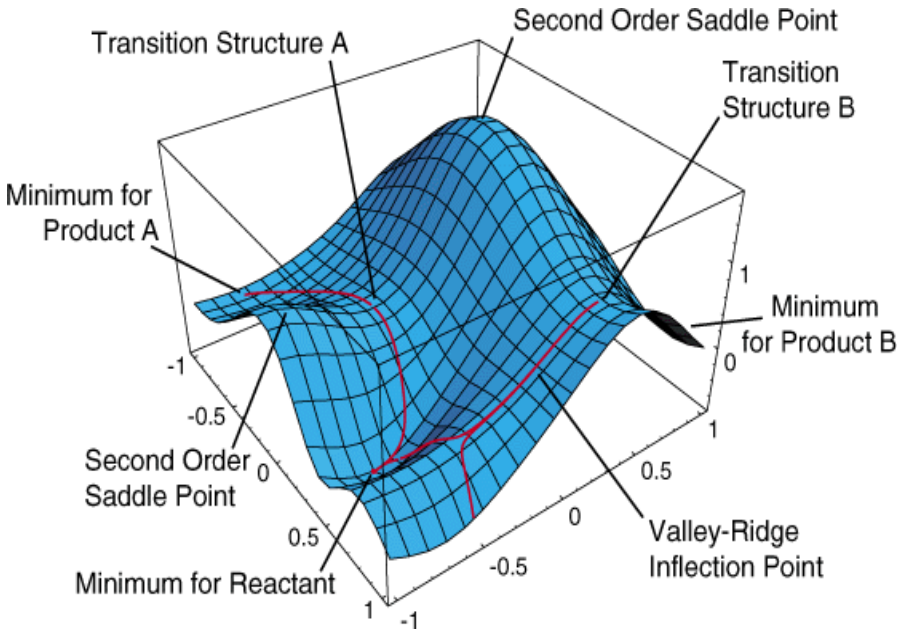
To study a chemical reaction using the PES as a function of atomic positions, it is necessary to calculate the energy for every atomic arrangement of interest. Methods of calculating the energy of a particular atomic arrangement of atoms are well described in the computational chemistry article, and the emphasis here will be on finding approximations of $E(\mathbf{r})$ to yield fine-grained energy-position information.

For very simple chemical systems or when simplifying approximations are made about inter-atomic interactions, it is sometimes possible to use an analytically derived expression for the energy as a function of the atomic positions.

- Changes in the energy of a system can be considered as movements on

  a multidimensional surface called energy surface.
- Changes in the energy funtion of its nuclear coordinates.
- Movement of the nuclei influences change in energy

- Mathematical function that gives the energy of a molecule as a function of its geometry
- Energy is plotted on the vertical axis, geometric coordinates (e.g bond lengths, valence angles, etc.) are plotted on the horizontal axes
- A PES can be thought of it as a hilly landscape, with valleys, mountain passes and peaks
- Real PES have many dimensions, but key feature can be represented by a 3 dimensional PES



- Equilibrium molecular structures correspond to the positions of the minima in the valleys on a PES
- Energetics of reactions can be calculated from the energies or altitudes of the minima for reactants and products
- A reaction path connects reactants and products through a mountain pass
- A transition structure is the highest point on the lowest energy path
- Reaction rates can be obtained from the height and profile of the potential energy surface around the transition structure
- The shape of the valley around a minimum determines the vibrational spectrum
- Each electronic state of a molecule has a separate potential energy surface, and the separation between these surfaces yields the electronic spectrum
- Properties of molecules such as dipole moment, polarizability, NMR shielding, etc. depend on the response of the energy to applied electric and magnetic fields

- Minima, lowest – global energy minima
- Minimization algorithms
- Highest point in the pathway between 2 minima is saddle point represents the transition state
- Minima and saddle points are stationary states on PES where the first derivative of energy function is 0
- E = f(x)
- E is a function of coordinates either cartesian or internal
- At minimum the first derivatives are zero and the second derivatives are all positive

5.1.1 Energy Minimisation: Statement of the Problem

The minimisation problem can be formally stated as follows: given a function f which depends on one or more independent variables $x_1, x_2, \ldots, x_i$, find the values of those variables where f has a minimum value. At a minimum point the first derivative of the function with respect to each of the variables is zero and the second derivatives are all positive:

$$\frac{\partial f}{\partial x_i} = 0; \qquad \frac{\partial^2 f}{\partial x_i^2} > 0 \tag{5.1}$$

The functions of most interest to us will be the quantum mechanics or molecular mechanics energy with the variables x_i being the Cartesian or the internal coordinates of the atoms.

- Minimization algorithm can go down hill on the energy surface and hence locate minima that is nearest to starting point

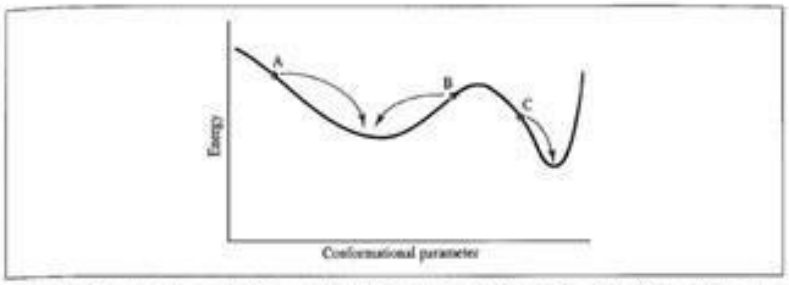


Fig. 3.3: A schematic one-dimensional energy surface. Minimisation methods move downhill to the nearest minimum. The statistical weight of the narrow, deep minimum may be less than a broad minimum which is higher in energy

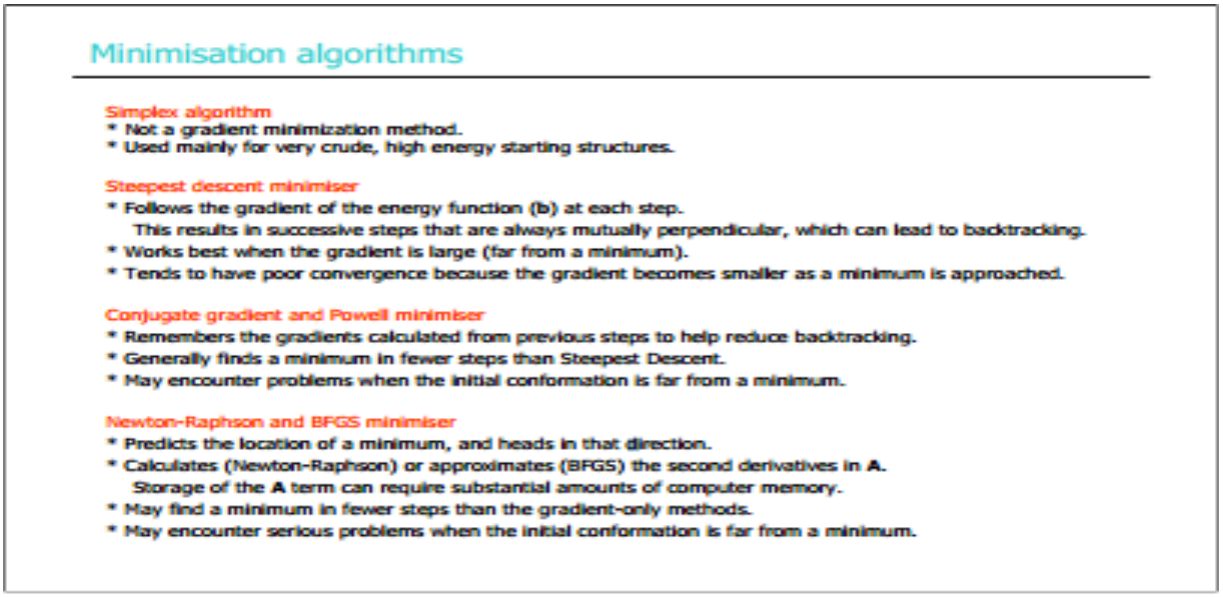
The input to a minimisation program consists of a set of initial coordinates for the system. The initial coordinates may come from a variety of sources. They may be obtained from an experimental technique, such as X-ray crystallography or NMR. In other cases a theoretical method is employed, such as a conformational search algorithm. A combination of experimental and theoretical approaches may also be used. For example, to study the

- $x_{new} = x_{old}$ + correction .
- In the equation, x_{new} refers to the value of the geometry at the next step (for example, moving from step 1 to 2 in the figure),
- x_{old} refers to the geometry at the current step, and correction is some adjustment made to the geometry.
- In all these methods, a numerical test is applied to the new geometry (xnew) to decide if a minimum is reached.

Minimization Methods

Several methods exist for finding a minimum of an arbitrary continuous function. One way to classify a minimization method is based on what kind of derivatives are used to guide the minimization. In this classification, we can distinguish between:

- methods that use no derivatives (function values, such as the energy)
- methods that use only first derivatives (slope, or force)
- methods that use second derivatives (curvature, or force constants)



Minimisation algorithms

The **steepest descent** minimiser uses the numerically calculated first derivative of the energy function to approach the energy minimum. The energy is calculcated for the initial geometry and then again when one of the atoms has been moved in a small increment. This process will be repeated for all atoms which finally are moved to new positions downhill on the energy surface. The optimisation process is slow near the minimum. Usually used as a first run (e.g. start of crystallographic refinement).

The **conjugate gradient** method accumulates the information about the function from one iteration to the next. With this proceeding, the reverse of the progress made in an earlier iteration can be avoided. Computational effort and storage requirements are greater than for steepest descent, but conjugate gradient is the method of choice for larger systems.

The **Powell** method is very similar to the conjugate gradient approach. It is faster in finding convergence and suitable for a variety of problems. However, torsion angles may sometimes be modified dramatically.

The **Newton-Raphson** minimiser also uses the curvature of the energy function to identify the search direction. Its efficiency increases as convergenc eis approached. Main disadvantage is the computational effort and large storage requirements for calculating larger systems. Also, for structures with high starin, the minimisation process can become instable. This method is thus not recommended as the first method in a refinement procedure.

Computer simulation

Computer simulation is the process of mathematical modelling, performed on a computer, which is designed to predict the behaviour of or the outcome of a real-world or physical system. Since they allow to check the reliability of chosen mathematical models, computer simulations have become a useful tool for the mathematical modeling of many natural systems in physics (computationalphysics), astrophysics, climatology, chemistry, biology and manufactur ing, as well as human systems in economics, psychology, social science, health care and engineering. Simulation of a system is represented as the running of the system's model.

It can be used to explore and gain new insights into new technology and to estimate the performance of systems too complex for analytical solutions.

A computer model is the algorithms and equations used to capture the behavior of the system being modeled. By contrast, computer simulation is the actual running of the program that contains these equations or algorithms. Simulation, therefore, is the process of running a model. Thus one would not "build a simulation"; instead, one would "build a model", and then either "run the model" or equivalently "run a simulation"

Benefits

- Gain greater understanding of a process
- Identify problem areas or bottlenecks in processes
- Evaluate effect of systems or process changes such as demand, resources, supply, and constraints
- Identify actions needed upstream or downstream relative to a given operation, organization, or activity to either improve or mitigate processes or events
- Evaluate impact of changes in policy prior to implementation

Types

- Discrete Models – Changes to the system occur at specific times
- Continuous Models – The state of the system changes continuously over time
- Mixed Models – Contains both discrete and continuous elements

Types of Data/Information Needed to Develop a Simulation Model:

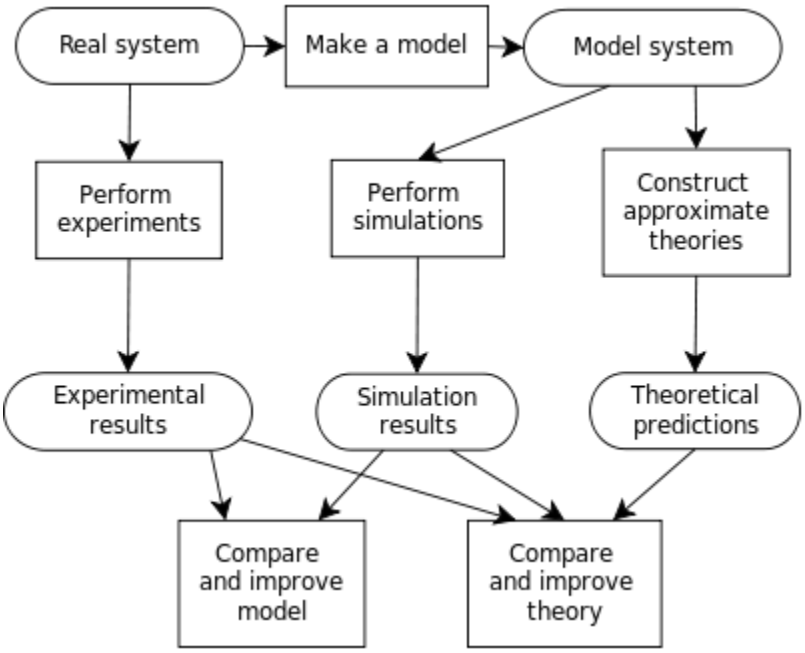
- The overall process flow and its associated resources
- What is being produced, served, or acted upon by the process (entities)
- Frequency at which the entities arrive in the process
- How long do individual steps in the process take
- Probability distributions that characterize real life uncertainties and variations in the process

- Computer simulation is the use of a computer to represent the dynamic responses of one system by the behavior of another system modeled after it.
- A simulation uses a mathematical description, or model, of a real system in the form of a computer program.
- This model is composed of equations that duplicate the functional relationships within the real system.
- When the program is run, the resulting mathematical dynamics form an analog of the behavior of the real system, with the results presented in the form of data.
- A simulation can also take the form of a computer-graphics image that represents dynamic processes in an animated sequence.

- Computer simulations have become a useful part of mathematical modeling of many natural systems in physics, astrophysics, chemistry, biology, climatology, psychology, social science, etc

USES

- Computer simulations are used to study the dynamic behavior of objects or systems in response to conditions that cannot be easily or safely applied in real life.
- Simulations are especially useful in enabling observers to measure and predict how the functioning of an entire system may be affected by altering individual components within that system.
- Simulations have great military applications also. Many uses for a computer simulation can be found within various scientific fields of study such as meteorology, physical sciences, etc



Process of building a computer model, and the interplay between experiment, simulation, and theory.

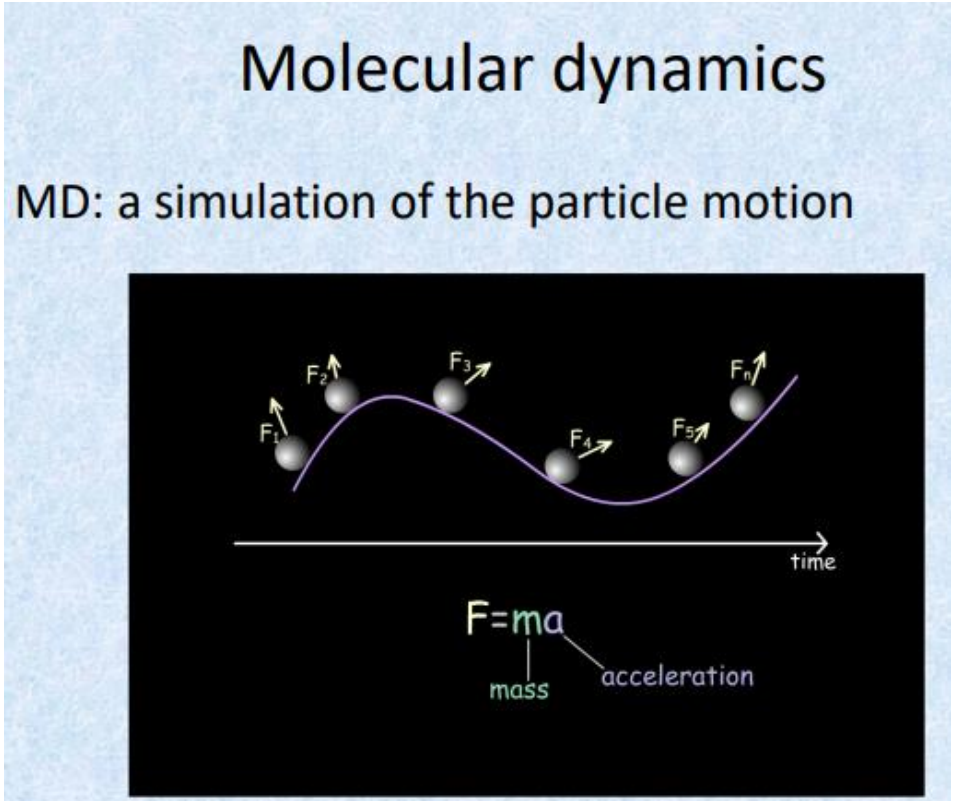
Basic Simulation Techniques

To explore the energy landscape described by the molecular mechanics force field, *i.e.* to sample molecular conformations, a simulation is required. This is also the route to relate the microscopic movements and positions of the atoms to the macroscopic or thermodynamic quantities that can

be measured experimentally. There are two major simulation methods to sample biomolecular systems: molecular dynamics (MD) and Monte Carlo (MC)

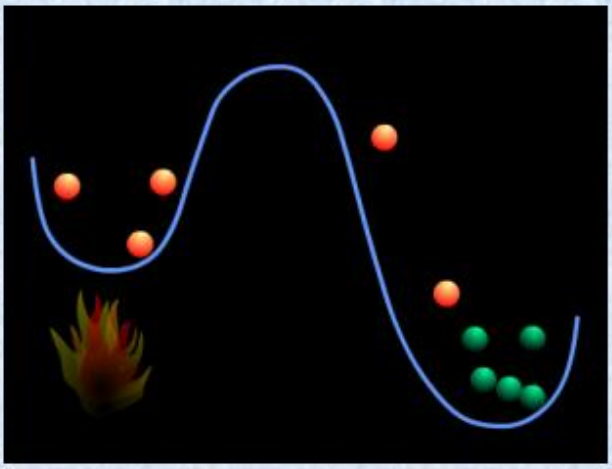
Molecular dynamics (MD) is a computer simulation method for analyzing the physical movements of atoms and molecules. The atoms and molecules are allowed to interact for a fixed period of time, giving a view of the dynamic "evolution" of the system.

Molecular dynamics provide an alternative approach to determine the preferred conformers and the global minimum of a molecule. This is achieved by the simulation of the dynamical motions of the molecule as it vibrates and undergoes internal rotation.



Molecular dynamics

The motion (determined by the temperature) allows conformational changes



Molecular dynamics

- Calculates the time dependent behaviour of a molecular system
- Provides detailed information on the fluctuations and conformational changes of macromolecules
- Routinely used to investigate the structure, dynamics and thermodynamics of biological molecules
- Used in the determination of structures from x-ray and NMR experiments

In a molecular dynamics (MD) simulation it is possible to explore the macroscopic properties of a system

The connection between microscopic simulation and macroscopic properties is made through statistical mechanics

Allows to study both thermodynamic properties and time dependent (kinetic) phenomenon

A MD simulation is practically carried out through the application of the Newton law:

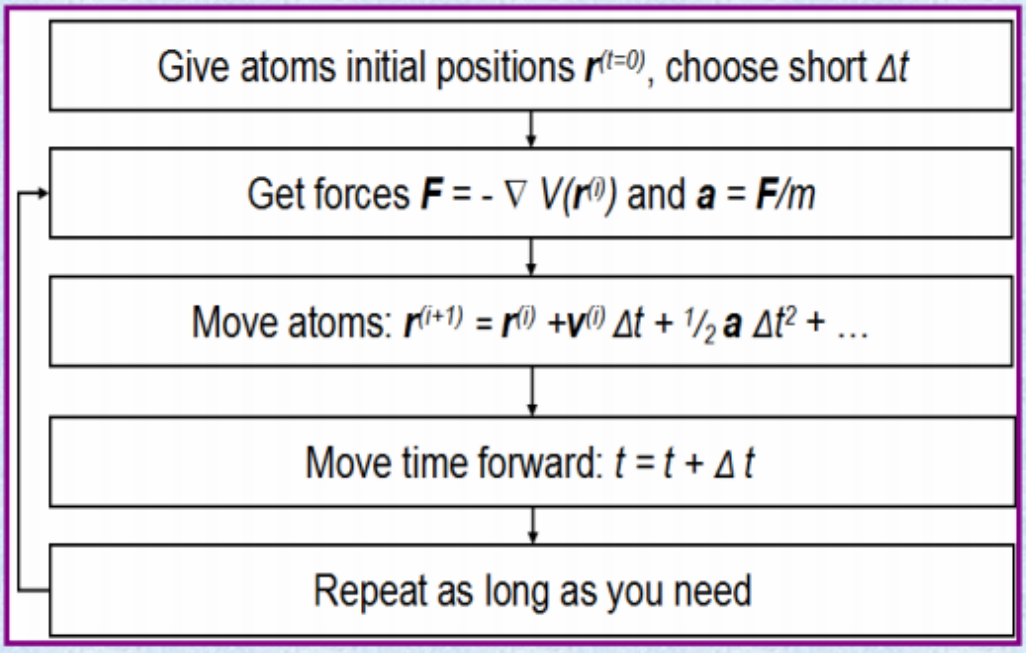
$$f = m \times a$$

The motion of each particle of the system is calculated from a

a is calculated from f

f is calculated from the potential V

A very simplified MD run



Molecular dynamics

- The potential ***V*** can be calculated at different accuracy level (from MM to QM)
- In biology the potential ***V*** is generally obtained by a MM force field
- This is a classical treatment allowing the calculation of conformational changes but usually it is not able to reproduce chemical reactions

Δt cannot be longer than the fastest atomic motion, therefore:

$$\Delta t = 10^{-15}$$

consequently a simulation of a microsecond needs one billion steps

Molecular dynamics

Temperature is directly correlated with kynetic energy:

$$K = \frac{3}{2} N k_B T$$

Generally a "free" evolution of the system is not allowed. Constraints on temperature and/or pressure are imposed in order to reproduce a particular ensemble.

Molecular dynamics

Environment simulation

The solvent can be simulated in an implicit and in an explicit manner.

Implicit solvent (in most cases the *continuum* approximation is used): fast calculation but poor results

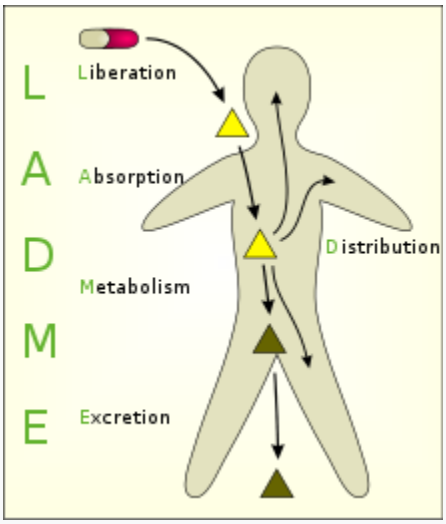
Explicit solvent (*periodic boundary conditions* are generally used): accurate results but time consuming

UNIT – 5- SBIA5302 – Computer aided drug design

Pharmacokinetics

Pharmacokinetics (from Ancient Greek *pharmakon* "drug" and *kinetikos* "moving, putting in motion"; see chemical kinetics), sometimes abbreviated as **PK**, is a branch of pharmacology dedicated to determine the fate of substances administered to a living organism. The substances of interest include any chemical xenobiotic such as: pharmaceutical drugs, pesticides, food additives, cosmetics, etc. It attempts to analyze chemical metabolism and to discover the fate of a chemical from the moment that it is administered up to the point at which it is completely eliminated from the body. Pharmacokinetics is the study of how an organism affects a drug, whereas pharmacodynamics (PD) is the study of how the drug affects the organism. Both together influence dosing, benefit, and adverse effects, as seen in PK/PD models.

Pharmacokinetics describes how the body affects a specific xenobiotic/chemical after administration through the mechanisms of absorption and distribution, as well as the metabolic changes of the substance in the body (e.g. by metabolic enzymes such as cytochrome P450 or glucuronosyltransferase enzymes), and the effects and routes of excretion of the metabolites of the drug.[2] Pharmacokinetic properties of chemicals are affected by the route of administration and the dose of administered drug. These may affect the absorption rate.[3]



Topics of Pharmacokinetics

Models have been developed to simplify conceptualization of the many processes that take place in the interaction between an organism and a chemical substance. One of these, the multi-compartmental model, is the most commonly used approximations to reality; however, the

complexity involved in adding parameters with that modelling approach means that *monocompartmental models* and above all *two compartmental models* are the most-frequently used. The various compartments that the model is divided into are commonly referred to as the ADME scheme (also referred to as LADME if liberation is included as a separate step from absorption):

- **L**iberation – the process of release of a drug from the pharmaceutical formulation.[4][5] See also IVIVC.
- **A**bsorption – the process of a substance entering the blood circulation.
- **D**istribution – the dispersion or dissemination of substances throughout the fluids and tissues of the body.
- **M**etabolism (or biotransformation, or inactivation) – the recognition by the organism that a foreign substance is present and the irreversible transformation of parent compounds into daughter metabolites.
- **E**xcretion – the removal of the substances from the body. In rare cases, some drugs irreversibly accumulate in body tissue.[*citation needed*]

The two phases of metabolism and excretion can also be grouped together under the title elimination. The study of these distinct phases involves the use and manipulation of basic concepts in order to understand the process dynamics. For this reason, in order to fully comprehend the *kinetics* of a drug it is necessary to have detailed knowledge of a number of factors such as: the properties of the substances that act as excipients, the characteristics of the appropriate biological membranes and the way that substances can cross them, or the characteristics of the enzyme reactions that inactivate the drug.

All these concepts can be represented through mathematical formulas that have a corresponding graphical representation. The use of these models allows an understanding of the characteristics of a molecule, as well as how a particular drug will behave given information regarding some of its basic characteristics such as its acid dissociation constant (pKa), bioavailability and solubility, absorption capacity and distribution in the organism.

The model outputs for a drug can be used in industry (for example, in calculating bioequivalence when designing generic drugs) or in the clinical application of

pharmacokinetic concepts. Clinical pharmacokinetics provides many performance guidelines for effective and efficient use of drugs for human-health professionals and in veterinary medicine

Metrics

The following are the most commonly measured pharmacokinetic metrics:[6] The units of the dose in the table are expressed in moles (mol) and molar (M). To express the metrics of the table in units of mass, instead of Amount of substance, simply replace 'mol' with 'g' and 'M' with 'g/dm3'. Similarly, other units in the table may be expressed in units of an equivalent dimension by scaling.

In pharmacokinetics, *steady state* refers to the situation where the overall intake of a drug is fairly in dynamic equilibrium with its elimination. In practice, it is generally considered that steady state is reached when a time of 3 to 5 times the half-life for a drug after regular dosing is started.

v · t · e **Pharmacokinetic metrics**

Characteristic	Description	Symbol	Unit	Formula	Worked example value
Dose	Amount of drug administered.	D	mol	Design parameter	500 mmol
Dosing interval	Time between drug dose administrations.	τ	s	Design parameter	24 h
C_{max}	The peak plasma concentration of a drug after administration.	C_{max}	M	Direct measurement	80.9 mmol/L
t_{max}	Time to reach C_{max}.	t_{max}	s	Direct measurement	3.9 h
C_{min}	The lowest (trough) concentration that a drug reaches before the next dose is administered.	$C_{\mathrm{min,ss}}$	M	Direct measurement	27.7 mmol/L
Volume of distribution	The apparent volume in which a drug is distributed (i.e., the parameter relating drug concentration in plasma to drug amount in the body).	V_d	$\mathrm{m^3}$	$\frac{D}{C_0}$	6.0 L
Concentration	Amount of drug in a given volume of plasma.	C_0, C_{ss}	M	$\frac{D}{V_\mathrm{d}}$	83.3 mmol/L
Absorption half-life	The time required for 50% of a given dose of drug to be absorbed into the systemic circulation.[citation needed]	$t_{\frac{1}{2}a}$	s	$\frac{\ln(2)}{k_a}$	1.0 h
Absorption rate constant	The rate at which a drug enters into the body for oral and other extravascular routes.	k_a	$\mathrm{s^{-1}}$	$\frac{\ln(2)}{t_{\frac{1}{2}a}}$	0.693 $\mathrm{h^{-1}}$
Elimination half-life	The time required for the concentration of the drug to reach half of its original value.	$t_{\frac{1}{2}b}$	s	$\frac{\ln(2)}{k_e}$	12 h
Elimination rate constant	The rate at which a drug is removed from the body.	k_e	$\mathrm{s^{-1}}$	$\frac{\ln(2)}{t_{\frac{1}{2}b}} = \frac{CL}{V_\mathrm{d}}$	0.0578 $\mathrm{h^{-1}}$
Infusion rate	Rate of infusion required to balance elimination.	k_{in}	$\mathrm{mol/s}$	$C_{ss} \cdot CL$	50 mmol/h
Area under the curve	The integral of the concentration-time curve (after a single dose or in steady state).	$AUC_{0-\infty}$	$\mathrm{M \cdot s}$	$\int_0^\infty C \,\mathrm{d}t$	1,320 mmol/L·h
		$AUC_{\tau,ss}$	$\mathrm{M \cdot s}$	$\int_t^{t+\tau} C \,\mathrm{d}t$	
Clearance	The volume of plasma cleared of the drug per unit time.	CL	$\mathrm{m^3/s}$	$V_\mathrm{d} \cdot k_a - \frac{D}{AUC}$	0.38 L/h
Bioavailability	The systemically available fraction of a drug.	f	Unitless	$\frac{AUC_{\mathrm{po}} \cdot D_{\mathrm{iv}}}{AUC_{\mathrm{iv}} \cdot D_{\mathrm{po}}}$	0.8
Fluctuation	Peak trough fluctuation within one dosing interval at steady state.	$\%PTF$	%	$\frac{C_{\mathrm{max,ss}} - C_{\mathrm{min,ss}}}{C_{\mathrm{av,ss}}} \cdot 100\%$ where $C_{\mathrm{av,ss}} = \frac{1}{\tau} AUC_{\tau,ss}$	41.8%

Pharmacokinetic models

Pharmacokinetic modelling is performed by noncompartmental or compartmental methods. Noncompartmental methods estimate the exposure to a drug by estimating the area under the curve of a concentration-time graph. Compartmental methods estimate the concentration-time graph using kinetic models. Noncompartmental methods are often more versatile in that they do not assume any specific compartmental model and produce accurate results also acceptable for bioequivalence studies. The final outcome of the transformations that a drug undergoes in an organism and the rules that determine this fate depend on a number of interrelated factors. A number of functional models have been developed in order to simplify the study of pharmacokinetics. These models are based on a consideration of an organism as a number of related compartments. The simplest idea is to think of an organism as only one homogenous compartment. This *monocompartmental model* presupposes that blood plasma concentrations of the drug are a true reflection of the drug's concentration in other fluids or tissues and that the elimination of the drug is directly proportional to the drug's concentration in the organism (first order kinetics).

However, these models do not always truly reflect the real situation within an organism. For example, not all body tissues have the same blood supply, so the distribution of the drug will be slower in these tissues than in others with a better blood supply. In addition, there are some tissues (such as the brain tissue) that present a real barrier to the distribution of drugs, that can be breached with greater or lesser ease depending on the drug's characteristics. If these relative conditions for the different tissue types are considered along with the rate of elimination, the organism can be considered to be acting like two compartments: one that we can call the *central compartment* that has a more rapid distribution, comprising organs and systems with a well-developed blood supply; and a *peripheral compartment* made up of organs with a lower blood flow. Other tissues, such as the brain, can occupy a variable position depending on a drug's ability to cross the barrier that separates the organ from the blood supply.

This *two compartment model* will vary depending on which compartment elimination occurs in. The most common situation is that elimination occurs in the central compartment as the liver and kidneys are organs with a good blood supply. However, in some situations it may be that elimination occurs in the peripheral compartment or even in both. This can mean that

there are three possible variations in the two compartment model, which still do not cover all possibilities.

This model may not be applicable in situations where some of the enzymes responsible for metabolizing the drug become saturated, or where an active elimination mechanism is present that is independent of the drug's plasma concentration. In the real world each tissue will have its own distribution characteristics and none of them will be strictly linear. If we label the drug's volume of distribution within the organism $\mathbf{Vd_F}$ and its volume of distribution in a tissue $\mathbf{Vd_T}$ the former will be described by an equation that takes into account all the tissues that act in different ways, that is:

$$Vd_F = Vd_{T1} + Vd_{T2} + Vd_{T3} + \cdots + Vd_{Tn}$$

This represents the *multi-compartment model* with a number of curves that express complicated equations in order to obtain an overall curve. A number of computer programs have been developed to plot these equations.[8] However complicated and precise this model may be, it still does not truly represent reality despite the effort involved in obtaining various distribution values for a drug. This is because the concept of distribution volume is a relative concept that is not a true reflection of reality. The choice of model therefore comes down to deciding which one offers the lowest margin of error for the drug involved.

Non compartmental model

Noncompartmental PK analysis is highly dependent on estimation of total drug exposure. Total drug exposure is most often estimated by area under the curve (AUC) methods, with the trapezoidal rule (numerical integration) the most common method. Due to the dependence on the length of x in the trapezoidal rule, the area estimation is highly dependent on the blood/plasma sampling schedule. That is, the closer time points are, the closer the trapezoids reflect the actual shape of the concentration-time curve. The number of time points available in order to perform a successful NCA analysis should be enough to cover the absorption, distribution and elimination phase to accurately characterize the drug. Beyond AUC exposure

measures, parameters such as Cmax (maximum concentration), Tmax(time at maximum concentration), CL and Vd can also be reported using NCA methods.

Compartmental analysis

Compartmental PK analysis uses kinetic models to describe and predict the concentration-time curve. PK compartmental models are often similar to kinetic models used in other scientific disciplines such as chemical kinetics and thermodynamics. The advantage of compartmental over some noncompartmental analyses is the ability to predict the concentration at any time. The disadvantage is the difficulty in developing and validating the proper model. Compartment-free modelling based on curve stripping does not suffer this limitation. The simplest PK compartmental model is the one-compartmental PK model with IV bolus administration and first-order elimination. The most complex PK models (called PBPK models) rely on the use of physiological information to ease development and validation.

Gráfica de absorción de un modelo monocompartimental

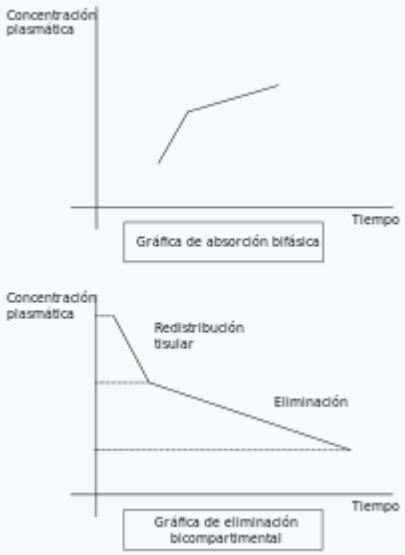
Gráfica de eliminación en un modelo monocompartimental

Single-compartment model

Linear pharmacokinetics is so-called because the graph of the relationship between the various factors involved (dose, blood plasma concentrations, elimination, etc.) gives a straight line or an approximation to one. For drugs to be effective they need to be able to move rapidly from blood plasma to other body fluids and tissues.

The change in concentration over time can be expressed as

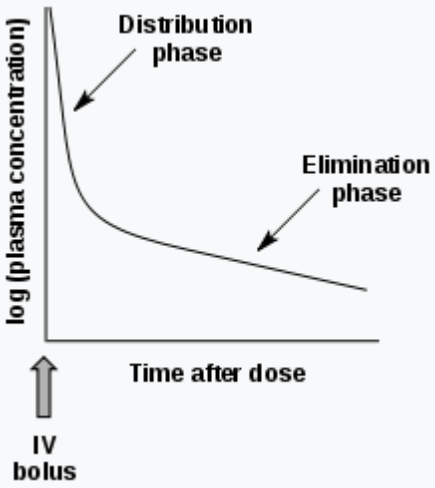
Multi-compartmental models[edit]



Graphs for absorption and elimination for a non-linear pharmacokinetic model.

The graph for the non-linear relationship between the various factors is represented by a curve; the relationships between the factors can then be found by calculating the dimensions of different areas under the curve. The models used in *non-linear pharmacokinetics* are largely based on Michaelis–Menten kinetics. A reaction's factors of non-linearity include the following:

- Multiphasic absorption: Drugs injected intravenously are removed from the plasma through two primary mechanisms: (1) Distribution to body tissues and (2) metabolism + excretion of the drugs. The resulting decrease of the drug's plasma concentration follows a biphasic pattern (see figure).



Plasma drug concentration vs time after an IV dose

- Alpha phase: An initial phase of rapid decrease in plasma concentration. The decrease is primarily attributed to drug distribution from the central compartment (circulation) into the peripheral compartments (body tissues). This phase ends when a pseudo-equilibrium of drug concentration is established between the central and peripheral compartments.
- Beta phase: A phase of gradual decrease in plasma concentration after the alpha phase. The decrease is primarily attributed to drug elimination, that is, metabolism and excretion.[9]
- Additional phases (gamma, delta, etc.) are sometimes seen.[10]

- A drug's characteristics make a clear distinction between tissues with high and low blood flow.
- Enzymatic saturation: When the dose of a drug whose elimination depends on biotransformation is increased above a certain threshold the enzymes responsible for its metabolism become saturated. The drug's plasma concentration will then increase disproportionately and its elimination will no longer be constant.
- Induction or enzymatic inhibition: Some drugs have the capacity to inhibit or stimulate their own metabolism, in negative or positive feedback reactions. As occurs with fluvoxamine, fluoxetine and phenytoin. As larger doses of these pharmaceuticals are administered the plasma concentrations of the unmetabolized drug increases and the elimination half-life increases. It is therefore necessary to adjust the dose or other treatment parameters when a high dosage is required.

- The kidneys can also establish active elimination mechanisms for some drugs, independent of plasma concentrations.

It can therefore be seen that non-linearity can occur because of reasons that affect the entire pharmacokinetic sequence: absorption, distribution, metabolism and elimination.

Bioavailabilty

In pharmacology, **bioavailability** is a subcategory of absorption and is the fraction (%) of an administered drug that reaches the systemic circulation.[1]

By definition, when a medication is administered intravenously, its bioavailability is 100%.[2][3] However, when a medication is administered via routes other than intravenous, its bioavailability is generally[TH] lower than that of intravenous due to intestinal endothelium absorption and first-pass metabolism. Thereby, mathematically, bioavailability equals the ratio of comparing the area under the plasma drug concentration curve versus time (AUC) for the extravascular formulation to the AUC for the intravascular formulation.[4] AUC is used because AUC is proportional to the dose that has entered the systemic circulation.[5]

Bioavailability of a drug is an average value; to take population variability into account, deviation range is shown as ±.[4] To ensure that the drug taker who has poor absorption is dosed appropriately, the bottom value of the deviation range is employed to represent real bioavailability and to calculate the drug dose needed for the drug taker to achieve systemic concentrations similar to the intravenous formulation.[4] To dose without knowing the drug taker's absorption rate, the bottom value of the deviation range is used in order to ensure the intended efficacy, unless the drug is associated with a narrow therapeutic window.[4]

For dietary supplements, herbs and other nutrients in which the route of administration is nearly always oral, bioavailability generally designates simply the quantity or fraction of the ingested dose that is absorbed.

Absolute bioavailability

Absolute bioavailability compares the bioavailability of the active drug in systemic circulation following non-intravenous administration (i.e., after oral, buccal, ocular, nasal, rectal, transdermal, subcutaneous, or sublingual administration), with the bioavailability of the same drug following intravenous administration. It is the fraction of the drug absorbed through

non-intravenous administration compared with the corresponding intravenous administration of the same drug. The comparison must be dose normalized (e.g., account for different doses or varying weights of the subjects); consequently, the amount absorbed is corrected by dividing the corresponding dose administered.

In pharmacology, in order to determine absolute bioavailability of a drug, a pharmacokinetic study must be done to obtain a *plasma drug concentration vs time* plot for the drug after both intravenous (iv) and extravascular (non-intravenous, i.e., oral) administration. The absolute bioavailability is the dose-corrected area under curve (*AUC*) non-intravenous divided by *AUC* intravenous. The formula for calculating the absolute bioavailability, *F*, of a drug administered orally (po) is given below (where *D* is dose administered).

Therefore, a drug given by the intravenous route will have an absolute bioavailability of 100% ($f = 1$), whereas drugs given by other routes usually have an absolute bioavailability of *less* than one. If we compare the two different dosage forms having same active ingredients and compare the two drug bioavailability is called comparative bioavailability.[*citation needed*]

Although knowing the true extent of systemic absorption (referred to as absolute bioavailability) is clearly useful, in practice it is not determined as frequently as one may think. The reason for this is that its assessment requires an *intravenous reference*; that is, a route of administration that guarantees all of the administered drug reaches systemic circulation. Such studies come at considerable cost, not least of which is the necessity to conduct preclinical toxicity tests to ensure adequate safety, as well as potential problems due to solubility limitations. These limitations may be overcome, however, by administering a very low dose (typically a few micrograms) of an isotopically labelled drug concomitantly with a therapeutic non-isotopically labelled oral dose (the isotopically-labelled intravenous dose is sufficiently low so as not to perturb the systemic drug concentrations achieved from the non-labelled oral dose).

The intravenous and oral concentrations can then be deconvoluted by virtue of their different isotopic constitution, and can thus be used to determine the oral and intravenous pharmacokinetics from the same dose administration. This technique eliminates pharmacokinetic issues with non-equivalent clearance as well as enabling the intravenous dose to be administered with a minimum of toxicology and formulation. The technique was first applied using stable-isotopes such as ^{13}C and mass-spectrometry to distinguish the isotopes by mass difference. More

recently, ^{14}C labelled drugs are administered intravenously and accelerator mass spectrometry (AMS) used to measure the isotopically labelled drug along with mass spectrometry for the unlabelled drug.

There is no regulatory requirement to define the intravenous pharmacokinetics or absolute bioavailability however regulatory authorities do sometimes ask for absolute bioavailability information of the extravascular route in cases in which the bioavailability is apparently low or variable and there is a proven relationship between the pharmacodynamics and the pharmacokinetics at therapeutic doses. In all such cases, to conduct an absolute bioavailability study requires that the drug be given intravenously.[18]

Intravenous administration of a developmental drug can provide valuable information on the fundamental pharmacokinetic parameters of volume of distribution (*V*) and clearance (*CL*)

Relative bioavailability and equivalence

In pharmacology, relative bioavailability measures the bioavailability (estimated as the *AUC*) of a formulation (A) of a certain drug when compared with another formulation (B) of the same drug, usually an established standard, or through administration via a different route. When the standard consists of intravenously administered drug, this is known as absolute bioavailability (see above).

$$F_{\mathrm{rel}} = 100 \cdot \frac{AUC_{\mathrm{A}} \cdot D_{\mathrm{B}}}{AUC_{\mathrm{B}} \cdot D_{\mathrm{A}}}$$

Relative bioavailability is one of the measures used to assess bioequivalence (*BE*) between two drug products. For FDA approval, a generic manufacturer must demonstrate that the 90% confidence interval for the ratio of the mean responses (usually of *AUC* and the maximum concentration, C_{max}) of its product to that of the "brand name drug"[OB] is within the limits of 80% to 125%. Where *AUC* refers to the concentration of the drug in the blood over time $t = 0$ to $t = \infty$, C_{max} refers to the maximum concentration of the drug in the blood. When T_{max} is given, it refers to the time it takes for a drug to reach C_{max}.

While the mechanisms by which a formulation affects bioavailability and bioequivalence have been extensively studied in drugs, formulation factors that influence bioavailability and bioequivalence in nutritional supplements are largely unknown.[19] As a result, in nutritional sciences, relative bioavailability or bioequivalence is the most common measure of bioavailability, comparing the bioavailability of one formulation of the same dietary ingredient to another.

Factors affecting bioavailability

The absolute bioavailability of a drug, when administered by an extravascular route, is usually less than one (i.e., $F < 100\%$). Various physiological factors reduce the availability of drugs prior to their entry into the systemic circulation. Whether a drug is taken with or without food will also affect absorption, other drugs taken concurrently may alter absorption and first-pass metabolism, intestinal motility alters the dissolution of the drug and may affect the degree of chemical degradation of the drug by intestinal microflora. Disease states affecting liver metabolism or gastrointestinal function will also have an effect.

Other factors may include, but are not limited to:

- Physical properties of the drug (hydrophobicity, pKa, solubility)
- The drug formulation (immediate release, excipients used, manufacturing methods, modified release – delayed release, extended release, sustained release, etc.)
- Whether the formulation is administered in a fed or fasted state
- Gastric emptying rate
- Circadian differences
- Interactions with other drugs/foods:
    - Interactions with other drugs (e.g., antacids, alcohol, nicotine)
    - Interactions with other foods (e.g., grapefruit juice, pomello, cranberry juice, brassica vegetables
- Transporters: Substrate of efflux transporters (e.g. P-glycoprotein)
- Health of the gastrointestinal tract
- Enzyme induction/inhibition by other drugs/foods:

    - Enzyme induction (increased rate of metabolism), e.g., Phenytoin induces CYP1A2, CYP2C9, CYP2C19, and CYP3A4
    - Enzyme inhibition (decreased rate of metabolism), e.g., grapefruit juice inhibits CYP3A → higher nifedipine concentrations
- Individual variation in metabolic differences
    - Age: In general, drugs are metabolized more slowly in fetal, neonatal, and geriatric populations
    - Phenotypic differences, enterohepatic circulation, diet, gender
- Disease state
    - E.g., hepatic insufficiency, poor renal function

Each of these factors may vary from patient to patient (inter-individual variation), and indeed in the same patient over time (intra-individual variation). In clinical trials, inter-individual variation is a critical measurement used to assess the bioavailability differences from patient to patient in order to ensure predictable dosing.

Bioavailability of drugs vs dietary supplements

In comparison to drugs, there are significant differences in dietary supplements that impact the evaluation of their bioavailability. These differences include the following: the fact that nutritional supplements provide benefits that are variable and often qualitative in nature; the measurement of nutrient absorption lacks the precision; nutritional supplements are consumed for prevention and well-being; nutritional supplements do not exhibit characteristic dose-response curves; and dosing intervals of nutritional supplements, therefore, are not critical in contrast to drug therapy.[11]

In addition, the lack of defined methodology and regulations surrounding the consumption of dietary supplements hinders the application of bioavailability measures in comparison to drugs. In clinical trials with dietary supplements, bioavailability primarily focuses on statistical descriptions of mean or average AUC differences between treatment groups, while often failing to compare or discuss their standard deviations or inter-individual variation. This failure leaves open the question of whether or not an individual in a group is likely to experience the benefits described by the mean-difference comparisons. Further, even if this issue were

discussed, it would be difficult to communicate meaning of these inter-subject variances to consumers and/or their physicians.

LADME

A number of phases occur once the drug enters into contact with the organism, these are described using the acronym LADME:

- Liberation of the active substance from the delivery system,
- Absorption of the active substance by the organism,
- Distribution through the blood plasma and different body tissues,
- Metabolism that is inactivation of the xenobiotic substance, and finally
- Excretion or elimination of the substance or the products of its metabolism.

Some textbooks combine the first two phases as the drug is often administered in an active form, which means that there is no liberation phase. Others include a phase that combines distribution, metabolism and excretion into a disposition phase. Other authors include the drug's toxicological aspect in what is known as *ADME-Tox* or *ADMET*.

Each of the phases is subject to physico-chemical interactions between a drug and an organism, which can be expressed mathematically. Pharmacokinetics is therefore based on mathematical equations that allow the prediction of a drug's behavior and which place great emphasis on the relationships between drug plasma concentrations and the time elapsed since the drug's administration.

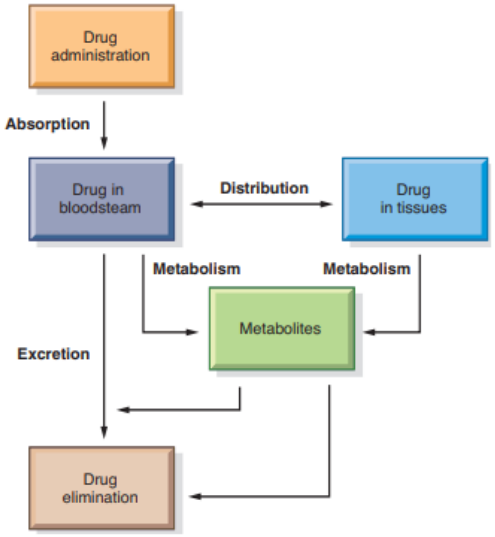


Figure 3.1 A representation of the four fundamental pathways of drug movement and modification in the body, the ADME processes.

Knowledge of these processes and the ways that they can vary between individuals is an important part of understanding how and why a drug is selected for a patient. To investigate the pharmacokinetic characteristics of a study drug (drug X), researchers will give a group of healthy adults a standard dose of drug X intravenously (IV) or orally at the start of the study. Blood is drawn from the study subjects repeatedly, at predetermined times, and analyzed for the amount of drug per volume of blood at each point in time. The value obtained is the serum or plasma concentration of the drug at the time the blood was drawn. When serum drug concentrations are graphed versus time, the result is the serum concentration versus time curve illustrated in Figure 3.2.

Urine is also collected from the study subjects to monitor for the appearance of drug x or related metabolites. Immediately after intravenous administration and at a later time after oral or other routes of administration, the amount of drug in the blood will reach a peak (peak serum concentration) and then begin to fall off, eventually disappearing from the blood completely. The time it takes for the serum concentration to fall by one half is called the t1 /2, or half-life, of elimination. At some point after the drug is no longer detectable in the blood, the last of the drug and its metabolites, which have been filtered out by the kidneys, will also disappear from the urine. Plasma concentration data collected from this type of study is plotted against time and

analyzed in order to understand the behavior of a specific drug in the body. This type of pharmacokinetic data, collected from average adults, is the basis for determining dose, dosing intervals, and limitations on the safe use of a drug. However, it is important for the technician student to remember that individuals do not always behave the way the average adult does. It is individual differences in ADME processes that create the need to modify doses or select different drugs in order to prevent poor treatment outcomes and adverse effects.

Absorption

Absorption is the transfer of a drug from its site of administration to the bloodstream. The rate and extent of absorption depends on the route of administration, the formulation and chemical properties of the drug, and physiologic factors that can impact the site of absorp-

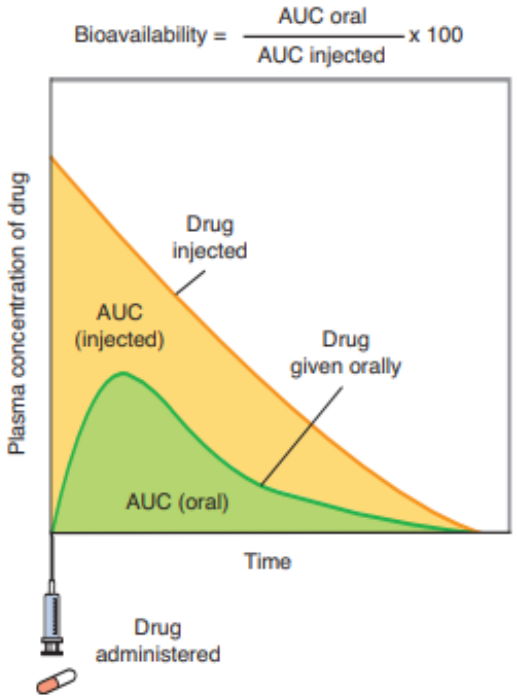


Figure 3.2 Serum concentration versus time curves for oral and IV administration of a drug. The time it takes for the serum concentration to fall by one half is called the $t_{1/2}$, or half-life of elimination. The bioavailability of the drug is determined by comparing the area under the two curves. AUC, area under curve. (Adapted with permission from Harvey RA, Champe PC, Howland RD, et al. *Lippincott's Illustrated Reviews: Pharmacology,* 3rd ed. Baltimore: Lippincott Williams & Wilkins, 2006.)

tion. When a drug is administered intravenously, absorption is not required because the drug is transferred from the administration device directly into the bloodstream. In the case of intravenous administration, the entire dose of the drug is available to move to the sites of drug action. Administration by other routes may result in less availability due to incomplete absorption. When this occurs, less of the drug is delivered by the bloodstream to the site of

action. When a tablet or capsule is swallowed it must dissolve before it can be absorbed. The dissolving of a tablet or capsule is referred to as dissolution. Manufacturing processes and the water solubility of the drug affect dissolution rates. Highly water-soluble medications dissolve more readily in the gastrointestinal (GI) tract, while fat-soluble drugs dissolve more slowly. Drugs with smaller particle sizes go into solution more readily. The inert ingredients added to formulations can also affect their dissolution. Manufacturers must avoid producing tablets so compacted that they pass through the GI tract without ever dissolving. Tablets that dissolve too early are also problematic, because they taste bad and are difficult to swallow. Special formulations or coatings can be used to delay dissolution, thereby protecting the drug from stomach acid or allowing the gradual release of the drug to intentionally lengthen the absorption process. These are referred to as delayed or sustained release formulations. Liquid preparations do not require the step of dissolution. This explains the more rapid onset of action seen with liquid formulations as compared with the same drug given in tablet or capsule form.

Once dissolution has occurred, the drug molecules must pass through the selectively permeable membranes of the cells lining the gastrointestinal tract to reach the bloodstream. Depending on their chemical and physical properties, drugs will be absorbed either by passive diffusion or carrier-mediated transport across these membranes. Passive diffusion occurs when there is a high concentration of the drug on one side of the membrane and a low concentration on the other side. This difference from one side of the membrane to the other is called a concentration gradient. It is the natural tendency of substances to move from a region of higher concentration to a region of lower concentration; in other words, substances move down the concentration gradient. Drug molecules move across membranes or move through pores between the epithelial cells (Fig. 3.3). Diffusion is most efficient with drugs that are small molecules. This movement is solely driven by the kinetic energy within molecules, and continues until concentrations reach equilibrium. When equilibrium exists, the concentration of the substance is approximately equal on both sides of the membrane.

Passive diffusion does not involve a carrier molecule and the process is not a saturable process. Saturable processes are limited to a certain rate of activity by some aspect of the process. The vast majority of drugs gain access to the blood stream by diffusion. Drugs, which are usually somewhat lipid-soluble (fat-soluble), readily move across most biological membranes. Those drugs that are highly water-soluble penetrate the cell membrane through

aqueous channels. A few drugs that closely resemble naturally occurring compounds are absorbed via carrier-mediated transport. This process requires carrier proteins that attach to and actively carry the drug molecules across the membrane, utilizing a natural "pump" mechanism. This method of absorption is limited by the availability of the carrier protein and is therefore, saturable. Carrier-mediated transport requires energy and can move molecules against the concentration gradient. For an illustration of passive diffusion and carrier-mediated transport see Figure 3.4.

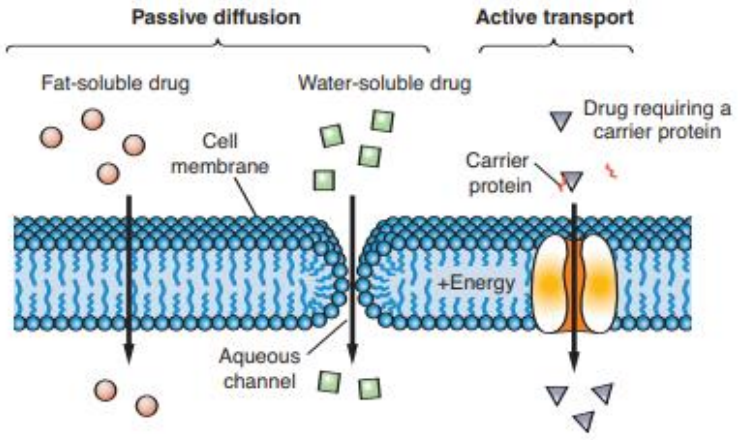


Figure 3.4 Drugs are absorbed through the epithelium of the GI tract by both passive diffusion and carrier-mediated transport. Fat-soluble drugs pass through membranes easily, highly water soluble drugs pass through aqueous channels, and carrier-mediated transport requires a pump process and the expenditure of energy.

Bioavailability

The relationship between the drug dose and the amount ultimately delivered to the bloodstream is defined as bioavailability and is generally expressed as a percentage. If a 1 gram dose of a drug is administered by mouth, and half of that reaches the systemic circulation, the drug is 50% bioavailable. Bioavailability is calculated, not measured directly. Previously, the half-life of elimination and how it is determined was discussed. The same graph of serum concentrations against time also provides the data necessary to derive bioavailability. The area under the plasma concentration versus time curve represents the total amount of the drug reaching the circulatory system (see Fig. 3.2). This curve will have a different shape depending on the route of drug administration. The curve obtained from plotting values after intravenous administration of a drug serves as a reference for complete bioavailability. To determine bioavailability for nonintravenous formulations, the area under the curve obtained after drug administration is compared with the area achieved when the same dose is given intravenously.

The ratio of the two is the bioavailability of the formulation tested. The acid environment or presence of food in the stomach, the solubility and other chemical properties of the drug, and the effect of the initial exposure to metabolic processes in the liver may all reduce the amount of drug that reaches the systemic circulation after oral administration, thereby reducing the bioavailability of the drug. When a drug is absorbed through the GI tract, it must travel through the liver before entering the systemic circulation (see Fig. 3.5). If the drug is subject to metabolism by the liver, the amount of drug that reaches the systemic circulation is decreased. Some drugs, such as propranolol or enalapril, undergo significant metabolism during a single passage through the liver. This is called the first-pass effect. When drugs are highly susceptible to the first-pass effect, the oral dose needed to cause a response will be significantly higher than the intravenous dose used to cause the same response.

Bioavailability becomes important in drug product selection. While two generic products may contain the same active ingredients, they may not have the same dissolution or absorption characteristics and therefore cannot be considered bioequivalent. In the case of extended-release products, the change in dissolution characteristics is intentional. In some cases, however, products are simply poorly manufactured and should be avoided. Generic equivalent products approved by the Food and Drug Administration (FDA) must meet a standard of less than 20% variation from the comparison product, an amount that does not significantly affect therapeutic efficacy or safety in most cases. In practice, however, most generic products typically vary from the original by less than 5%. Bioequivalency data is published by the FDA's Center for Drug Evaluation and Research in the "Approved Drug Products with Therapeutic Equivalence Evaluations" publication, universally referred to as the "Orange Book" and available on the web.

Factors Affecting Absorption

A number of patient-specific factors can affect absorption. Absorption from any site of administration requires blood flow. A patient who is in shock or cardiopulmonary arrest will need to have medications administered intravenously to achieve the desired response because of the reduced blood circulation in these situations. Most absorption after oral administration occurs in the small intestine because of its larger surface area and therefore greater blood flow.

Significant impairment of absorption can result when sections of the small intestine are removed for medical reasons. Contact time with the epithelial lining of the GI tract is also an important factor in drug absorption. In people with a very rapid transit time through the GI tract, due to severe diarrhea for example, medication cannot be effectively absorbed. Conversely, anything that delays stomach emptying (for instance, a large meal) will also delay and potentially reduce absorption.

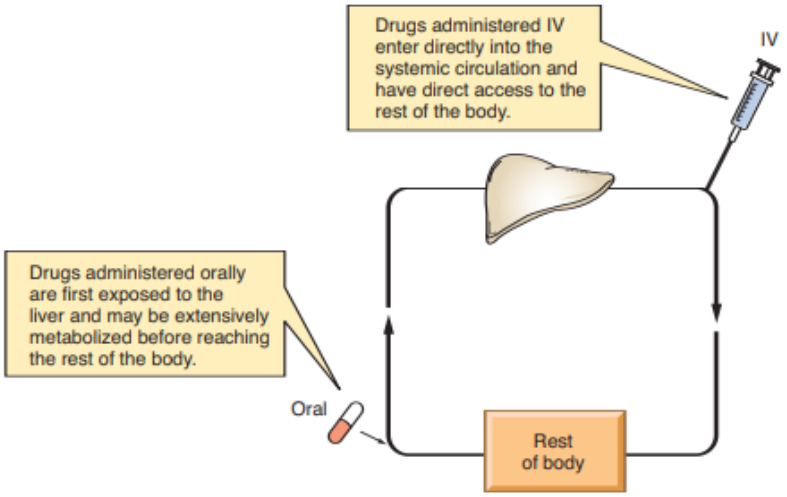


Figure 3.5 Orally administered drugs are exposed to metabolic processes in the liver before they reach the systemic circulation, which may significantly reduce bioavailability. (Adapted with permission from Harvey RA, Champe PC, Howland RD, et al. *Lippincott's Illustrated Reviews: Pharmacology,* 3rd ed. Baltimore: Lippincott Williams & Wilkins, 2006.)

Some medications exhibit drug-food or drug-drug interactions with other compounds present in the GI tract. Specific foods or other drugs may bind a drug and prevent absorption. The interaction between tetracycline and dairy products or antacids is a good example of this effect. Although excess exposure to stomach acid may negatively impact some drugs, patients with very low levels of stomach acid (achlorhydria) may experience inadequate tablet dissolution and therefore poor drug absorption. Achlorhydria is most common in the elderly population.

Distribution

Once a drug is absorbed into the bloodstream it can be carried throughout the body. This process is called distribution, and is a reversible process; while some molecules may be interacting with receptors on cell membranes or inside of cells, other molecules may move back into the bloodstream. The delivery of a drug from the bloodstream to the site of drug action primarily depends on blood flow, capillary permeability, the degree of binding (attachment) of the drug to blood and tissue proteins, and the relative lipid-solubility of the drug molecule. Blood flow to different organs of the body is not equal. The most vitally important organs of the body receive the greatest supply of blood. These organs include the brain, liver, and kidneys. Skeletal muscle and bone receive less blood, and adipose tissue (fat) receives the least. If blood flow were the only factor affecting distribution, it would be reasonable to expect that high concentrations of administered medications would always appear in the brain and liver. In reality, few drugs exhibit good penetration of the central nervous system.

The anatomical structure of the capillary network in the brain creates a significant barrier to the passage of many drugs and is commonly referred to as the blood-brain barrier. This barrier is an adaptation that for the most part protects brain tissue from invasion by foreign substances. To readily penetrate into the brain, drugs must be fairly small and lipidsoluble or must be picked up by the carrier-mediated transport mechanism in the central nervous system. This explains why the small and highly fat-soluble anesthetic gases quickly and easily penetrate the brain to cause anesthesia, while other larger and water soluble molecules like penicillin antibiotics penetrate the central nervous system to a much lesser degree. Compare the fairly impermeable capillaries of the brain to the highly permeable capillary walls in the liver and spleen (Fig. 3.6). These capillaries have gaps between their cells that allow large proteins to pass through to the capillary basement membrane. Capillary structure here is well adapted to the function of the liver, the key protein producer in the body and a center for chemical change of other compounds. In order to function, the liver must have access to amino acids, sugars, and other large molecules from the bloodstream. These molecules undergo chemical processing in the liver, and then must be moved out of the hepatic cells and back to the bloodstream.

Factors Affecting Distribution

The blood is composed of a number of elements, including plasma, red and white blood cells, and plasma proteins. Most drugs reversibly bind to plasma proteins in varying degrees. Albumin is the plasma protein with the greatest capacity for binding drugs. Binding to plasma proteins affects drug distribution into tissues, because only drug that is not bound is available to penetrate tissues, bind to receptors, and exert activity. As free drug leaves the bloodstream, more bound drug is released from binding sites. In this way, drugs maintain a balance between free and bound drug that is unique to each compound, based on its affinity for plasma proteins. Albumin, then, acts as a reservoir of an administered drug.

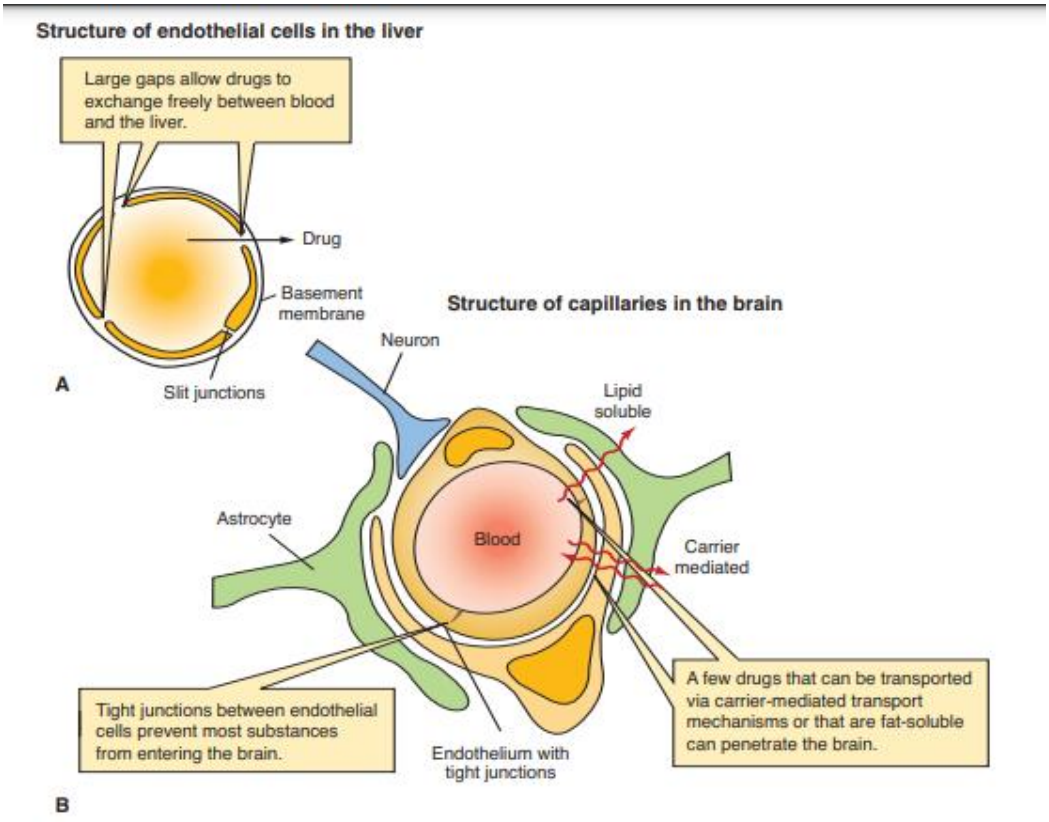


Figure 3.6 Diagram demonstrating the difference in the anatomy of brain and liver capillaries. The tight junctions in brain capillaries prevent passage of most drugs into the brain. (Adapted with permission from Harvey RA, Champe PC, Howland RD, et al. *Lippincott's Illustrated Reviews: Pharmacology,* 3rd ed. Baltimore: Lippincott Williams & Wilkins, 2006.)

Some drugs have a high affinity for binding to serum proteins and may be 95% to 98% protein bound. With highly protein bound drugs, low albumin levels (as in protein-calorie malnutrition, or chronic illness) may lead to toxicity because there are fewer than the normal sites for the drug to bind. The amount of free drug is significantly increased in that case. The physician or

pharmacist must consider the patient's serum albumin level when the dose of a highly protein bound medication is selected. Competition for binding sites is one important way that drugs might interact. If a patient is using two highly protein bound drugs at the same time, there will be competition for binding sites on the albumin. The drug with the greatest affinity for the albumin will bind, and is thought to disrupt the normal ratio of free to bound drug for the second medication. As a result, the second medication will be more available to distribute to the site of action and potentially cause side effects.

Other patient variables that can affect distribution include body composition, cardiac decompensation (heart failure), and age of the patient. These factors all affect the apparent volume of distribution (Vd) of a drug and ultimately play a role in determining the appropriate dose of a drug. Volume of distribution is the hypothetical volume needed to account for all of the drug in the body based on the serum concentration in a blood sample. For example, assume that a 1-gram dose of drug X is given by IV injection to a patient. Thirty minutes later the serum concentration (Cp) is 25mcg/ml. If the drug were evenly distributed through the body, the apparent volume of distribution would have to be 40 liters (L) to account for the entire dose of the drug.

Cp = Dose (amount in the body)/Vd

The adult human body is about 60% water. Therefore, the body of an average adult who weighs 70 kg (about 150 lbs) contains 42 L, or 10 gallons, of water. Total body water can be conceptually divided into three spaces or compartments. The fluid contained in the bloodstream makes up about 5 L, or about 9% of the volume of an average sized adult. The total water outside of the cells (extracellular fluid) includes the plasma volume plus the fluid in the interstitial space and is about 14 L. Intracellular fluid makes up the remaining 28 L. Drugs with high molecular weights or drugs that are extremely hydrophilic (with a strong affinity for water) tend to stay within the circulatory system and organs with a rich blood supply, and have a smaller apparent volume of distribution. Small, highly lipophilic drugs have a large apparent volume of distribution.

Volume of distribution often becomes important when dosing calculations are made based on weight. A short, obese woman who weighs 250 lbs cannot handle the same dose of an aminoglycoside antibiotic that a tall, muscular man of the same weight would require. This is because a greater portion of the woman's body is made up of fat. Aminoglycoside drugs are

water-soluble and stay mainly in extracellular fluid, therefore dosing must be based on adjusted body weight, which will more correctly reflect the true volume of extracellular fluid in the body. Dosing of medications in infants and children requires special consideration. It is the often-repeated wisdom of pediatric health care specialists that “children are not simply small adults.” This means that dosing cannot simply be adjusted based on the lower weight of children. The body composition of children is very different from adults. Their bodies contain a much higher percentage of water and a lower percentage of muscle and fat. Albumin levels may also be lower, especially in neonates. These variations result in different values for volume of distribution and significantly affect drug dosing.

Metabolism

Drugs are eliminated from the body either unchanged through the kidneys and bile, or they may undergo chemical changes that allow them to be more easily excreted. The process of undergoing chemical changes is called biotransformation, or metabolism. As previously noted, anything absorbed through the GI tract goes directly into the portal circulation that feeds into the liver. The liver is adapted to clear toxins from the body and is the major site for drug metabolism, but specific drugs may undergo biotransformation in other tissues. The kidneys cannot efficiently excrete highly fat-soluble drugs that readily cross cell membranes because they are reabsorbed in the last stages of filtration. These compounds must first be metabolized in the liver to more water-soluble compounds and then removed. There are two types of metabolic processes drugs undergo in the liver. Most undergo one or both types of reactions. In the first type of reaction drugs are made more polar through oxidation-reduction reactions or hydrolysis.

These reactions use metabolic enzymes, most often those of the cytochrome P450 enzyme system, to catalyze the biotransformation. In enzyme-catalyzed reactions, the rate of the reaction is accelerated by the presence of enzymes. A limited amount of enzyme is present at any given time in the liver. Since the rate of enzyme-catalyzed drug metabolism is limited by the quantity of available enzyme, metabolism in these cases is considered a saturable process. This means that the rate of conversion will only continue at the normal pace until the available supply of enzyme is used. At that point, metabolism is slowed until enzyme becomes available again. For the usual doses of most drugs, these reactions never reach saturation. There are a few drugs where doses may reach the saturation point of the enzymes. Once enzymes become saturated,

blood levels increase exponentially toward toxicity. Examples include metabolism of alcohol and phenytoin.

The second type of metabolism involves conjugation reactions. In this type of reaction the drug undergoing change is joined with another substance, such as glucuronic acid, sulfuric acid, acetic acid, or an amino acid. Glucuronidation is the most common conjugation reaction. The result of conjugation is a more water-soluble compound that is easier for the kidneys to excrete. These metabolites are most often therapeutically inactive. Some agents are initially administered as an inactive compound (prodrug) in order to improve availability or reduce side effects. Metabolism converts the prodrug to the active form. Fosphenytoin, for example, is a prodrug of phenytoin, a drug used for seizure disorders. Fosphenytoin is more completely and quickly absorbed when given by IM injection than phenytoin and can be used in critical situations with greater ease because it dose not require insertion of an intravenous catheter.

Factors Affecting Metabolism

Metabolism of drugs can vary widely between population groups. Deficiency of some drug metabolizing enzymes is genetic and will result in poor tolerance of certain drugs. For example, many Asians and Native Americans have difficulty metabolizing drugs that require acetylation, such as ethanol. These individuals will exhibit a low tolerance of such drugs, and can suffer adverse drug reactions at a much higher rate than the average population. Age is another important variable that has a bearing on metabolism. Organ function gradually declines with age and the elderly may poorly tolerate drugs that require metabolism. The very young require special consideration of drug dosing because of immaturity of their organ systems. This subject matter is worthy of further study for those technicians who will be serving these special populations.

Drug interactions may occur between two drugs that are metabolized by the same enzyme systems in the liver. Because there is a limit on available enzymes for metabolism, excess drug will remain active and free to exert an effect elsewhere in the body. Usually, of two drugs that are metabolized by the same enzyme system, one has a higher affinity for the enzyme, and levels of the second drug build up. In some cases, the drug being metabolized will induce the production of more of the enzymes. Enzyme induction sets the stage for another type of drug

interaction, because the increased production of metabolizing enzymes may result in higher rates of removal and the need for an increased dose of the second drug. A good example of a drug that stimulates production of metabolizing enzymes is phenobarbital, a drug used to treat some forms of epilepsy.

Excretion

When a drug is taken into and distributed throughout the body, it must be subsequently removed, or concentrations of the drug would continue to rise with each successive dose. The complete removal of the drug from the body is referred to as elimination. Elimination of the drug encompasses both the metabolism of the drug, and excretion of the drug through the kidneys, and to a much lesser degree into the bile. Excretion into the urine is one of the most important mechanisms of drug removal.

The kidneys act as a filter for the blood and create urine as a vehicle for removal of waste. Blood enters the kidney through renal arteries and then is filtered by the glomerulus. The glomerular filtrate becomes concentrated and substances are removed as it passes through the renal tubule and eventually becomes urine. Drug molecules in the bloodstream that are not bound to albumin are also filtered out into the glomerular filtrate. When drugs have not been converted to water soluble compounds in the liver, they are likely to be reabsorbed back into the bloodstream at the end of the filtration process, and will cycle through the body again. If they are water soluble, they will end up in the urine and be excreted.

When a medication is given repeatedly, as most are in real patients, the total amount of drug in the body will increase up to a point and then stabilize. At this point, the amount being taken in by the patient is equal to the amount being removed by the liver and kidneys (Fig. 3.7). This state of equilibrium is called steady state, and drug levels will remain fairly constant unless there is a dose change, an interruption in treatment, or failure of the organs of elimination. The therapeutic effects of many drugs are closely correlated to a specific range of steady state serum drug levels, and physicians or clinical pharmacists will monitor these levels and adjust doses when necessary so that patients obtain the appropriate drug response

Factors Affecting Excretion

The complete elimination of a drug from the body is dependent on normal liver and kidney function. The kidney is the major organ of excretion; however, the liver also contributes to elimination through metabolism and excretion into the feces via the bile. When a patient has reduced kidney function or another problem that lengthens the half-life of a drug, dosage adjustment is required. If the dosage is not adjusted, the drug will accumulate in the body.

Kidney or liver failure, or conditions where blood flow to these organs is reduced, complicate drug and dose selection. Drugs that are dependent upon excretion through the kidneys are not the best choice for patients with renal failure. Patients with liver disease will better tolerate drugs that can be cleared exclusively through the kidneys. Age must be considered in a discussion of drug excretion. The very young and very old will have lower rates of excretion; the old because of deterioration in organ function and the very young, because the kidneys have not reached full maturity. Doses often require reduction in these patients. Drug interactions, such as when multiple drugs compete for metabolic processes, can also reduce drug removal. Clearly, the interactions between a drug and the human body are incredibly complex. This makes choosing the most appropriate medication and dose complicated, and that choice becomes more obscure when a patient is taking many other medications. This complexity makes it all the more essential that prescriptions be filled carefully and accurately at all times.

Pharmacodynamics

Pharmacodynamics (**PD**) is the study of the biochemical and physiologic effects of drugs (especially pharmaceutical drugs). The effects can include those manifested within animals (including humans), microorganisms, or combinations of organisms (for example, infection).

Pharmacodynamics and pharmacokinetics are the main branches of pharmacology, being itself a topic of biology interested in the study of the interactions between both endogenous and exogenous chemical substances with living organisms.

In particular, pharmacodynamics is the study of how a drug affects an organism, whereas pharmacokinetics is the study of how the organism affects the drug. Both together influence dosing, benefit, and adverse effects. Pharmacodynamics is sometimes abbreviated as PD and pharmacokinetics as PK, especially in combined reference (for example, when speaking of PK/PD models).

Pharmacodynamics places particular emphasis on dose–response relationships, that is, the relationships between drug concentration and effect.[1] One dominant example is drug-receptor interactions as modeled by

where *L*, *R*, and *LR* represent ligand (drug), receptor, and ligand-receptor complex concentrations, respectively. This equation represents a simplified model of reaction dynamics that can be studied mathematically through tools such as free energy maps.

Effects on the body

Desired activity

The desired activity of a drug is mainly due to successful targeting of one of the following:

- Cellular membrane disruption
- Chemical reaction with downstream effects
- Interaction with enzyme proteins
- Interaction with structural proteins
- Interaction with carrier proteins
- Interaction with ion channels
- Ligand binding to receptors:
  - Hormone receptors
  - Neuromodulator receptors
  - Neurotransmitter receptors

General anesthetics were once thought to work by disordering the neural membranes, thereby altering the Na^+ influx. Antacids and chelating agents combine chemically in the body. Enzyme-substrate binding is a way to alter the production or metabolism of key endogenous chemicals, for example aspirin irreversibly inhibits the enzyme prostaglandin synthetase (cyclooxygenase) thereby preventing inflammatory response. Colchicine, a drug for gout, interferes with the function of the structural protein tubulin, while Digitalis, a drug still used in heart failure, inhibits the activity of the carrier molecule, Na-K-ATPase pump. The widest class of drugs act as ligands that bind to receptors that determine cellular effects. Upon drug binding,

receptors can elicit their normal action (agonist), blocked action (antagonist), or even action opposite to normal (inverse agonist).

In principle, a pharmacologist would aim for a target plasma concentration of the drug for a desired level of response. In reality, there are many factors affecting this goal. Pharmacokinetic factors determine peak concentrations, and concentrations cannot be maintained with absolute consistency because of metabolic breakdown and excretory clearance. Genetic factors may exist which would alter metabolism or drug action itself, and a patient's immediate status may also affect indicated dosage.

Undesirable effects

Undesirable effects of a drug include:

- Increased probability of cell mutation (carcinogenic activity)
- A multitude of simultaneous assorted actions which may be deleterious
- Interaction (additive, multiplicative, or metabolic)
- Induced physiological damage, or abnormal chronic conditions

Therapeutic Window

The therapeutic window is the amount of a medication between the amount that gives an effect (effective dose) and the amount that gives more adverse effects than desired effects. For instance, medication with a small pharmaceutical window must be administered with care and control, e.g. by frequently measuring blood concentration of the drug, since it easily loses effects or gives adverse effects.

Duration of action

The *duration of action* of a drug is the length of time that particular drug is effective.[4] Duration of action is a function of several parameters including plasma half-life, the time to equilibrate between plasma and target compartments, and the off rate of the drug from its biological target

Therapeutic Index

The therapeutic index (TI; also referred to as therapeutic ratio) is a quantitative measurement of the relative safety of a drug. It is a comparison of the amount of a therapeutic

agent that causes the therapeutic effect to the amount that causes toxicity.[1] The related terms therapeutic window or safety window refer to a range of doses which optimize between efficacy and toxicity, achieving the greatest therapeutic benefit without resulting in unacceptable side-effects or toxicity.

Classically, in an established clinical indication setting of an approved drug, TI refers to the ratio of the dose of drug that causes adverse effects at an incidence/severity not compatible with the targeted indication (e.g. toxic dose in 50% of subjects, TD50) to the dose that leads to the desired pharmacological effect (e.g. efficacious dose in 50% of subjects, ED50). In contrast, in a drug development setting TI is calculated based on plasma exposure levels.

In the early days of pharmaceutical toxicology, TI was frequently determined in animals as lethal dose of a drug for 50% of the population (LD50) divided by the minimum effective dose for 50% of the population (ED50). Today, more sophisticated toxicity endpoints are used.

$$\text{Therapeutic Index} = \frac{\text{LD}_{50}}{\text{ED}_{50}} \text{ in animal studies, or for humans, } \text{Therapeutic Index} = \frac{\text{TD}_{50}}{\text{ED}_{50}}$$

For many drugs, there are severe toxicities that occur at sublethal doses in humans, and these toxicities often limit the maximum dose of a drug. A higher therapeutic index is preferable to a lower one: a patient would have to take a much higher dose of such a drug to reach the toxic threshold than the dose taken to elicit the therapeutic effect.

Generally, a drug or other therapeutic agent with a narrow therapeutic range (i.e. having little difference between toxic and therapeutic doses) may have its dosage adjusted according to measurements of the actual blood levels achieved in the person taking it. This may be achieved through therapeutic drug monitoring (TDM) protocols. TDM is recommended for use in the treatment of psychiatric disorders with lithium due to its narrow therapeutic range.

Term	Meaning
ED	Effective Dose
TD	Toxic Dose
LD	Lethal Dose
TI	Therapeutic Index

TR	Therapeutic Ratio