Intro

- Draggability: Interacts with target, readily absorbed by gut (LogP, Acid/enzyme stability), avoid reactive groups such as aldehyde (oxidative instability)
- Ideal drug: Max 4 steps for synthesis, No heavy metal catalyst, No problematic wastes, Stable up to 70C, Solid state properties, Solubility, Oral bioavailability > 90%, High potency (5-10mg dose per day)

Production steps

- Discovery → identification of active substances (4 methods)
- Optimization → improvement of an active compound (decrease toxicity by increasing affinity and selectivity)
- Development → preparation of better absorbed compounds

Drug activity phases (Pharmaceutical Phases

- Pharmaceutical phase → crystalline for tablets and soluble for injections
- Pharmacokinetics Phase → ADME properties for solubility and lipophilicity to allow both absorption and passage through membranes, what the body does to the drug`
- Pharmacodynamic → How the drug interacts with the target and the effects, what the drug does to the body

Classification

- By target → simultaneous interference of several receptors, need for system biology
- By mechanism of action → used for static targets only
- ATC system → organ to act upon, therapeutic properties, chemical properties

Binding targets

- Binding Groups: Functional groups involved in forming intermolecular bonds with target, binding regions are part of the carbon skeleton which are important in binding.
- Water → H2O molecules interacting with the target must be stripped away. If the energy required is more than the energy gained from the binding the drug may be ineffective. Hydrophobic effect plays minor role in drug binding
    - Polar groups can be added to make the drug more soluble, but should be solvent exposed, to exposed to the target
- Non-covalent bonds:
    - Electrostatic or ionic bonds (20-40 kj/mol): Strength in inversely proportional with distance (coulomb law), increases in hydrophobic environments. Drop-off not strong (compared to other non-covalent), so Ionic interactions are likely to be the most important initial interaction.
    - Hydrogen bonds (16-60 kj/mol) - alcohols, amines, esters, amides, carb acids, phenols, ketones. Takes place between an electron-rich heteroatom (most likely O or N) and an electron-deficient hydrogen (usually linked to an electron-rich atom). The electron distribution will be weighted towards the electronegative atom → H gains a small positive charge → formation of 180° weak σ bond. HBD (H bond donor) provides the hydrogen, HBA (H bond acceptor) provides the electron-rich atom. Some groups may be both donors and acceptors → hydrogen bond flip-flop.
    - Van-der-Waals interactions (2-4 kj/mol) - alkenes and aromatic rings, Takes place between hydrophobic regions of different molecules. Electronic distribution is never neutral or symmetrical → temporary dipoles →interaction fall off quickly
        - Lennard-Jones potential: Mathematical relationship between two pairs of neutral atoms, equilibrium is called Van-Der-Waals radius, balance between attract and repulsive forces
    - Dipole-dipole interactions: Result from different electronegativities on functional groups of the interaction → formation of dipole moments which are parallel and opposite in direction. The alignment is beneficial for binding and for better activity. Fall-off is slower than in VDW but quicker than in electrostatic interactions.
    - Ion-Dipole interactions - ammonium ions, quaternary ammonium salts, and carboxylate groups Stronger than a dipole-dipole (falls-off less rapidly). Induced dipole moment takes place when a ionic group distorts an aromatic ring which changes its electron distribution.

Enzymes Targets

- Ehrlich: Hypothesized something on the cell (receptors) are complementary to drugs/dye, magic bullet idea
- Reversible inhibition when the drug binds through intermolecular interactions.
- Irreversible inhibition results if the drug reacts with the enzyme and forms a covalent bond.
- Competitive inhibitors bind to the active site and compete with either the substrate or the cofactor.
- Allosteric inhibitors bind to an allosteric binding site, which is different from the active site. They alter the shape of the enzyme such that the active site is no longer recognizable.
- Transition-state analogues are enzyme inhibitors designed to mimic the transition state of an enzyme-catalyzed reaction mechanism. They bind more strongly than either the substrate or the product.
- Suicide substrates are molecules that act as substrates for a target enzyme, but which are converted into highly reactive species forming covalent bond

Receptor Targets

- Agonists are compounds that mimic the natural ligand for the receptor. They may have a similar structure to the natural ligand. Partial agonists induce a weaker effect than a full agonist.
- Inverse agonists act as antagonists, but also eliminate any resting activity associated with a receptor, safer than antagonists
- Antagonists are agents that bind to the receptor, but which do not activate it. They block binding of the natural ligand. They bind differently from the natural ligand such that the receptor is not activated. They can bind to regions of the receptor that are not involved in binding the natural ligand. In general, antagonists tend to have more binding interactions than agonists and bind more strongly.

Pharmacokinetics

- If a drug won't reach its target, it doesn't matter how well it interacts, Ferrari with no gas
- ADME: Main concerns of pharmacokinetics are absorption, distribution, metabolism, and excretion

Absorption and Distribution

- BioPhase: The site of drug action, from site of administration goes through absorption into circulation, distribution via blood plasma. At all points some portion is eliminated via biotransformation's or excretion
- Drug absorption →passage from its site of administration (and type) to the systematic circulation
- First Pass Effect: orally administered drugs go directly to the kidney where a certain percent is transformed before it reache s its target, some of the drug can be taken up in route, injections and inhalations avoide it

Types of administration:

- Enteral: drug enters in contact with the mucosa or by swallowing or rectal.
    - Intravenous bolus→ no absorption, immediate effect. Complications may have serious consequences.
    - Intravenous infusion → dose injected slowly at constant rate which allows control
    - Intramuscular injections → absorbed through muscular blood, easy to perform. Rate of abs. can vary
- Parenteral: GI tract is bypassed; can be I.v. or intramuscular.
    - Buccal or Sublingual →through mucosa, passive diffusion. X may be swallowed. No first pass
    - Oral → absorption from GI tract, safe and easy.
    - Rectal → first-pass metabolism is partially avoided, absorption varies.

Transcellular drug transport:

1. Passive diffusion → molecules spontaneously move from a region of higher concentration to a region of lower concentration. Polarity and size are hindrance factors → Fick's law A:Area, P: LogP, C:concentration, D:diffusion coefficient, h:membrane thickness

$$\text{Rate of diffusion} = \frac{\text{DAP}(C_{\text{out}} - C_{\text{in}})}{h}$$

2. Carrier-mediated transport → thanks to the transmembrane protein's molecules can either go along a concentration gradient (facilitated diffusion) or against it requiring ATP (active transport). At high drug concentrations the carrier sites become saturated, competitive inhibition can occur
3. Vesicular transport → internalization of molecules (endocytosis) or release of a molecule (exocytosis).

- Paracellular drug transport happens through small acquis contact points (cell junctions). Initiated by a concentration gradient over the cell layer or by a hydrostatic pressure gradient across the cell layer. Size and characteristics vary a lot, wide and open in liver, very tight in brain, no paracellular transport

Acid-base properties

- Directly affect absorption, excretion, and compatibility with other drugs. It is possible to predict the degree of ionization of a molecule at a particular pH. Blood (pH 7.4), stomach (1-3)
- pKa > 7.5 → essentially unionized while pKa from 3-7.5 fraction changes with pH, lower than 2.5 almost all ionized → slow diffusion, drugs with pKa between 6-8 are approximately 50% ionized in blood
- Aspirin is a weak acid, pKa 3.5, Amphetamine is a weak base pKa 9.8

Henderson-Hasselbach equation:

- % Ionization = (ionized) / (Ionized + Unionized) x 100
- Ion trapping → The degree of ionization depends on their pKa and pH of the solution. At diffusion equilibrium, concentration of unionized and ionized molecules is equal on both sides of the biological barrier. If there's any difference in pH the ionized one will be concentrated on one side

For a weak acid

$$pH - pK_a = \log \frac{[Ionized]}{[Un-ionized]}$$

For a weak base

$$pH - pK_a = \log \frac{[Un-ionized]}{[Ionized]}$$

- How do we achieve the right balance of lipid/water soluble?
    - Amine functional group because are partially ionized at blood pH (7.4) → pH = pKa when 50% of amino is ionized allowing them to cross membranes as non-ionized, while ionized form gives good water solubility and good target binding. That's why many drugs contain amine

LRO5

- Parameters indicating if a drug is likely to be orally active
- A small number of therapeutic categories fall outside our parameter cutoffs: antibiotics, antifungals, vitamins and cardiac glycosides.

1. LogP < 5 → Partition coefficient (P) is the ratio of the drug's concentration in octanol (mimic lipids) to that in water. It determines water solubility.
2. Molecular weight < 500 no direct implications, but they're more likely to have more than 5 HBD
3. HBD<5
4. HBA<10

Veber's parameters

1. [RB < 10 + polar surface are < 140A] OR [HDB+HDA < 12]

- PSA: The sum of surfaces of polar atoms in a molecule. The only input needed is the molecular topology.
- RB: flexible molecules are less likely to be orally active, and harder to predict binding conformation, sometimes structures can be reified to improve pharmacodynamic properties
- Exceptions: Large molecule weight by high jacking transporter (levodopa, phenylalanine) or polar and low MW

Drug Metabolism

- Metabolites: Metabolic products of drugs, often inactive, some toxic. Mostly in the liver
- Prodrugs: become active after some metabolic process
- Drug metabolism is unpredictable and different among individuals. For a drug to be able to perform its function and be excreted biotransformation is necessary.

Phase I transformations

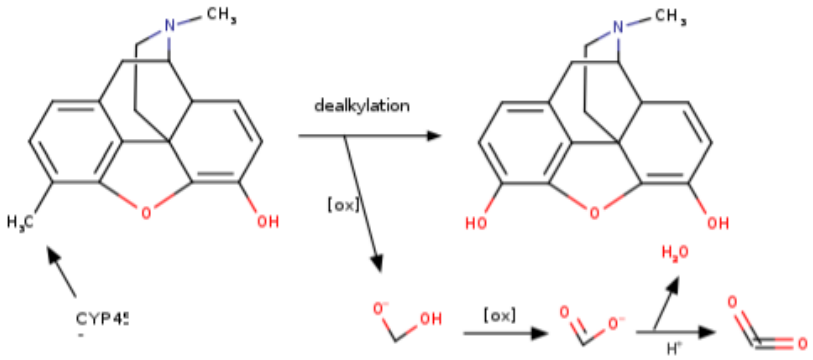
- Oxidation, reduction, and hydrolysis
- Too polar they will be excreted before serving their functions; too nonpolar, the body won't be able to excrete
- A process is needed to either convert them to more polar by adding polar functional groups or by "unmasking" polar groups that are already present in the drug. This reaction can also take place to create more reactive species for phase II reactions.

- Cytochrome P450 enzymes are the biggest family of oxidative enzymes, each drug reacts with its specific CYP450, and this is determinant to establish drugs compatibility (antagonist binding). CYP2C9, CYP2D6 and CYP3A4 make up for 70% of drug metabolism, depends on population, always changing
- Flavin Containing Monooxygenases: mainly oxidation of nucleophilic N, S and P, similarly to P450
- Monoamine Oxidases involved in deamination, alcohol and aldehyde dehydrogenases also present

Oxidation

What is oxidized? **Exposed/Activates C, N, S atoms**

|| Ex: codeine → morphine (more reactive species)



1) **CYP450** is a monooxygenase located in liver cells → splits molecular **oxygen** and yield one to form H_2O
2) Oxidized form CYP450 (Fe3+) binds **substrate** → CYP450 enters its **high-spin form**
3) **NADPH** donates **e**$^-$ to the complex→ **reduction** of the enzyme (Fe2+ state)
4) Ferrous form has high **affinity** for diatomic gasses (**CO**) → O strips e' from Fe2+ second **e**$^-$
5) **Reduced** complex (Fe3+ state) **picks** a second **e-** and passes it to an O atom→ formation of **peroxide** (O^{2-}_2)
6) The peroxide anion is **split** by the addition of two **H**+ to stabilize the O → **H_2O** liberated
7) Remaining O is stabilized by the electron density of Fe→ binary enzyme-product **dissociation**

Reduction

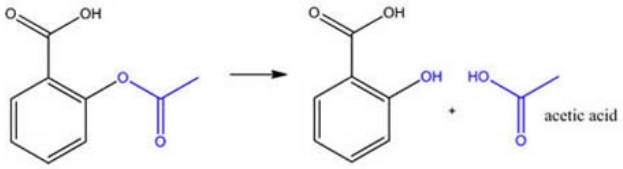
(Not frequent and always reversible) What is reduced?

Nitro, azo and carbonyl groups

Hydrolysis

(Operated by esterase to give amides and esters)

|| Ex. Aspirin → salicylic acid |



Phase II transformations

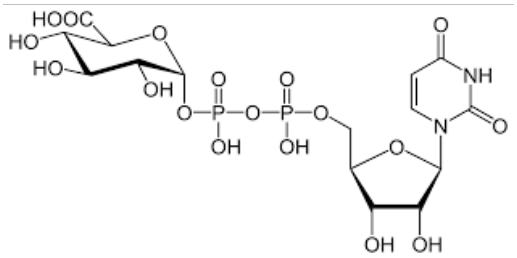
Reactions catalyzed by transferase enzyme; resulting **conjugates** are usually inactive. Are "soft" reactions because they add polar molecules without changing the original molecule. The resulting molecule has increased polarity and is easily excreted.

Glucuronic acid conjugation:

Usually involve phenols, alcohols, and carboxylic acids,

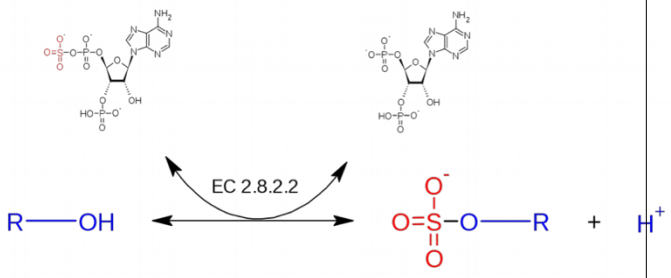
|| Ex. Paracetamol on -OH

The sp2 carbon atom gets a nucleophilic attack by O atom of alcohol



Sulfate conjugation:

Less common conjugation, Ex: Paracetamol on OH



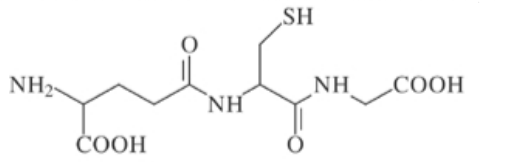
Glutathione conjugation:

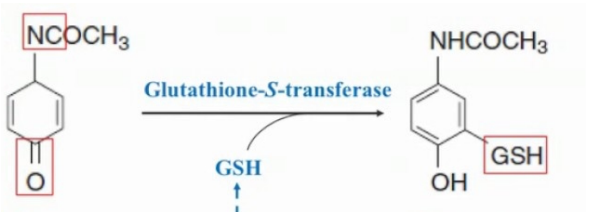
Usually takes place with epoxides, alkyl halides and radicals (electrophiles formed in phase I).

Important in **detoxification** (stabilize dangerous oxidative compounds).

|| Ex. Paracetamol toxic intermediate

Glutathione attacks delta (-) with its SH (thiol) group (later peptidase break cys conjugate and acetylation of N leads to a highly polar compound)



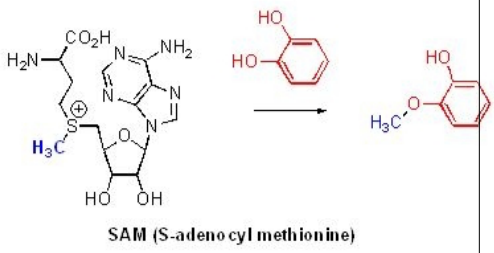


Methylation and acetylation

decrease in polarity of the drug. Enzymes involved AcCoA and SAM.

Methylation: Amines, thiols and phenols: $R_2NH \rightarrow R_2NMe$ (catalyzed by SAM)

Acetylation: Primary amines, hydrazine: $RNH_2 \rightarrow RNHAc$

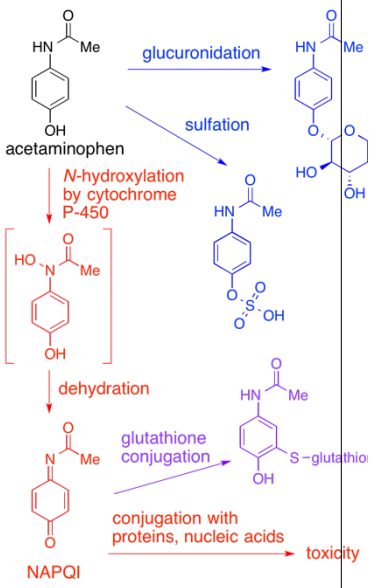


- **Metabolic stability** → production of metabolites complicates drug therapy, either blocking it or by producing toxic substances. Different populations have different level of enzymes (pharmacogenetics). Some external factors, as food, may have some influence on P450.

Biotransformation

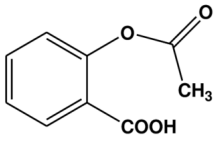
- In vivo bioactivation may lead to the formation of electrophilic entities that are capable of linking with biological macromolecules inducing alternate functioning.
- **Toxification**: Inactivate drugs, toxic metabolites could be formed

Ex. **Oxidation of paracetamol to NAPQI**

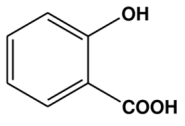


- **Indirect toxicity** → Biotransformation usually has an enzymatic nature, either to activate prodrugs or produce toxic metabolites
  - **(**a) Transient formation of a reactive intermediate (Reactive intermediate, NAPQI)
  - (b) Primary metabolites accumulate
  - (c) Excess of final metabolites.
- Inactivate drugs, called **prodrugs** (usually drugs with a long half life)
- **Drug latentation:** is the chemical modification of a biologically active compound to form a new compound, which in vivo will liberate the parent compound. Drug latentiation is synonymous with prodrug design.

  || Ex. **Aspirin is a prodrug for salicylic acid**



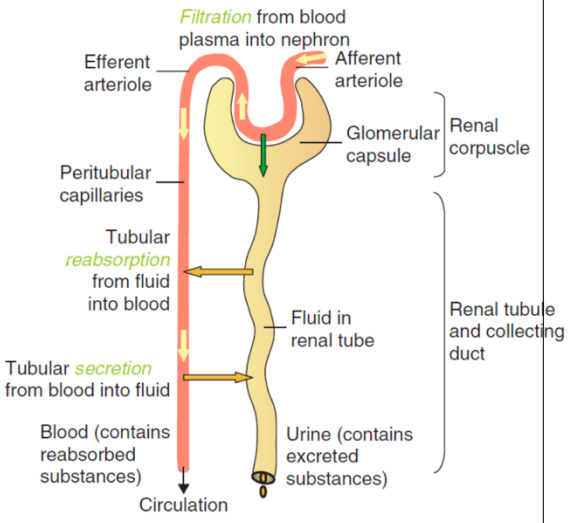
Aspirin (prodrug)



Salicylic acid

Excretion:

- Volatile and gaseous drugs excreted through the lungs, Ex gaseous general anesthetics
- The **kidneys** are the principal way through which drugs are excreted.
    - Pathway: renal artery → capillary with knotted structure → duct (nephron) →pressure → plasma passes through pores → drugs and metabolites pass through → much of the content of the nephron is reabsorbed through aquaporins → excretion.
- Up to 10-15% of the drug is lost through skin in sweat, can also be excreted through saliva and breast milk



Classical Drug Discovery

Intro

- Used to be trial and error process, often killing patients (low purity) in 1920s organic chemistry contributed
- First drugs were alkaloids, natural compounds containing a basic nitrogen, example Nicotine
- Drug Discovery Steps: Choose disease → Choose Target → Identify Bioassay → Find a lead → Isolate/purify → determine structure (Last two steps only needed from natural sources)
- Drug Design Steps: Identify SARs → Identify pharmacophore → Improve PD → improve PK
- Drug Development steps: Patent → Preclinical → manufacturing process → Clinical trials → register → $
- In the past, drug targets were discovered after drugs, now we have more knowledge of receptors than drugs
    - Orphan receptor → unknown endogenous ligand, ex morphine until 10 years ago
    - Recently discovered: Caspases can be activated or inhibited to increased or decreased cell death
- Orphan Drug → less than 200k people in US, or won't be profitable for 7 years, publicly funded
- Selectivity: Example penicillin, cell wall not present in Eukaryotes
- Specificity: Example B-Adrenergic receptors in heart (B1, propranolol) vs lungs (B2)
- Subtype Selectivity: Antipsychotic agents need to be more selective and not bind D2 → Parkinson's symptoms
- Pitfalls: There is often no one target for a disease, hypertension for examples, roadblock example
- Multi-Target Drugs: Activate a range of specific and similar targets

Bioassays

- Fail Fast, fail cheap: Minimizes risk of investing money in a failed drug, refining library
- Bioassays: Should be simple, quick and relevant
- Combinational Synthesis: Thousands of related samples synthesized, processed, screened
- Parallel Synthesis: Synthesis of small amount of many different compounds, each vial contains different compound, along with combinational builds a library (synthetic compound archives) for HTS
- In Vitro Test: Culture of cells measures therapeutic index (Therapeutic dose/Toxic dose), and PK properties via microsomes and hepatocytes from liver containing P450, identifies likely metabolism of drug
    - Antibacterial drugs can be tested how effectively they inhibit/kill colonies
    - Determines if inhibition is competitive, non-competitive, and determine IC50 values
    - Caco-2 cell monolayer used to predict absorption in GIT
- In Vivo Test: Transgenic animals (containing human genes) with induced disease tested, many invalid due to large differences between test animal and humans, and different ways to produce a specific symptom
- HTS: Automated testing of 1000s (HTS), 100s (MTS) or 10s (LTS) (depending on $) compounds on large number of targets. Effects should be observable, such as cell growth, enzyme reaction
    - Positive hits have an activity range of 30-100 uM, unfortunately many false positives
    - Promiscuous Inhibitors: Main case of false hits, agents that inhibit range of protein targets, very poor selectivity

    - Other False Hits: Chemically reactive agents which alkylate or acylate, irreversibly inhibiting protein, generally irreversible inhibitors are not included in HTS and not considered a potential lead. Ex: Epoxides, alkyl halides
    - HTS also used for lead optimization, but focus is on lead discovery
- Screening methods: NMR (Structure), Affinity screening (Binding ability), Virtual screening (lead hits the target), SPR (If target binds drug), SPA (Visual method if ligand binds), ITC (Thermodynamic properties of binding)

Hits and Leads

- Hit: an active substance having a preferential activity for the target which satisfies some criteria:
    - 1) Reproducible activity 2) Confirmed structure and high purity 3) Specificity 4) Confirmed novelty 5) Chemically tractable structures (not too aggressive)
- Lead→ Validation of a hit:
    - 1) active in vivo 2) no hERG 3) Analogs of the hit must display SAR 4) not reactive 5) patentable
- Preclinical Candidate → Undergo toxicology, lead optimization, cross interactions, ADME
- Clinical Candidate → passes preclinical, undergoes human trials
- Ex: Nevirapine series

Hit and Lead finding Strategies

- Four main methods: Analogue design, Systematic screening, exploitation of biological information, planned research and rational approach

Screening of Natural Compounds

- Cragg and Newman: most drugs are from natural sources, only 37% are totally synthetic
- Radical Chemistry: Artemisinin (antimalarial) has trioxane rings, looks very unstable
- Natural screening: has been carries out by evolution to find stable molecules
- Plants: Morphine, Cocaine, Taxol (anticancer, from yew)
- Microorganisms: Especially antibacterial therapy such as Penicillin
- Marine sources: interesting inflammatory, antiviral, and anticancer (Curacin A, antitumor)
- Animal Sources: Magainin's- family of antibiotics extracted from frog skin
- Venom and Toxins: Very potent and specific interactions, teprotide from venom (antihypertensive) or nonpeptide venoms such as tetrodotoxin (pufferfish)
- Medical Folklore: Often attributed to placebo effect. Rhubarb has been a purgative for centuries, active ingredient Dantron used for lead compound for laxatives

Screening synthetic compound libraries

- Pharmaceutical companies have vast libraries of molecules that never made it to market, or were produced for synthetic chemistry, they can be searched for possible lead compounds
- It can also be worth testing synthetic intermediates, such as quinoline 3-carboxainide intermediates showing antiviral activates

Existing drugs- Analogue Design

- Me-too and Me-better: companies use their competitors' drugs as lead compounds, either creating a me too (such as the antihypertensive captopril) or me better (such as modern penicillin's)
- Most popular strategy for drug design is synthesis of existing active molecules, main job of a medicinal chemist
- Analogues can be of three types:
    - Structural: Only chemical similarities (Ex. Testosterone and progesterone)
    - Functional: similar pharmacological properties, but not structures (Ex. Fentanyl and Morphine)
    - True/Full: Present chemical and pharmacological similarities (Ex. Morphine and Codeine)

SOSA

- Many drugs have a minor undesirable side effect, SOSA aims to enhance this side effect and reduce original

- Choosing a known drug for a lead compound has advantage that lead is already drug-like, unlike many Hits
- Ex. Sulfonamides (antibacterial) had hypoglycemia side effect, used to create tolbutamide for diabetes

Starting from Natural Ligand

- Adrenaline and noradrenaline were used as leads for adrenergic p-antagonists such as salbutamol
- The natural ligand can also be a lead for antagonists, ex histamine used as lead for antihistamine Cimetidine, turning an agonist into antagonist is often by adding extra binding group to lead structure
- In the case of orphan receptor, finding natural ligand can be a lot of work, but very beneficial Ex. Receptor for morphine led to discovery of endorphins, natural pain killers
- Discovery of cannabinoid receptor lead to discovery of anandamide, useful in suppressing nausea in chemo
- A natural substrate for an enzyme can also be used as a lead for an enzyme inhibitor, Ex. Natural substrate for HIV protease was used as a lead for first protease inhibitors in HIV
- Enzyme products can also be used as a lead, because reactions are reversable

Natural Modulators

- Receptors and enzymes under allosteric control can be studied and use their modulators as leads
- Ex. Benzodiazepine was known to inhibit the then called benzodiazepine receptor (GABA) as the natural modulator was unknown, we now know the natural modulator is y-aminobutyric acid (GABA)

Other methods

- Computer aided design of lead compounds from computer simulations that study the binding of compounds to analogous receptors or the receptor if it has been discovered (SBDD)
- Serendipity and prepared mind: Some leads found through chance, but the medicinal chemist has to know what to look for
    - Ex. Propranolol the B-blocker is lipophilic, meaning it cannot enter CNS. In an attempt to further cut down CNS entry (and side effects) a hydrophilic amide group was added, making a compound called practolol which did reduce CNS entry but also was highly selective for B-receptors in heart instead of in other organs

Drug Design

- Drug Design Steps: Identify SARs → Identify pharmacophore → Improve PD → improve PK
- Often the most potent lead doesn't have favorable PK properties, so we select a less potent one

SARs

- SARS is defined as the structural features of the molecule contributing (or taking away) to the biological activities of the molecule of interest, similar molecules exert similar biological actions in a qualitative sense
- Crum-Brown-Fraser (1869): The first SAR, observed tertiary amine containing compounds (morphine, nicotine) became muscle relaxants when converted to quaternary ammonium compounds. Later proven wrong, showed the link between chemical structures and activities, they likely blocked acetylcholine's (muscle constricting)
- The thinking became: One chemical group gives one biological action
- Ehrlich: Developed the concept of drug receipts, postulated "side-chains" on cells were "complementary" to drugs or dyes, and introduced the magic bullet idea, also developed arsenicals which were toxic to trypanosomes but were not toxic to the host
- Albert: Developed the concept of selective toxicity from the magic bullet idea
- Paradox: How could one chemical group be a muscle reactant (N-Methyl nicotine) and a and contractant (acetylcholine)? Molecules much bigger than the antagonist bind competitively and block the action (agonist effect), so muscle relaxants will contain quaternary amine AND be larger than acetylcholine
- Different functional groups yield compounds with similar physicochemical properties:
    - Ex. Sulfanilamide and p-aminobenzoic acid, similar steric electric properties, Sulfanilamide acts as an agonist to PABA metabolism in bacteria

SAR properties

- Solubility (described by acid/base properties), P describes the ratio of drug in octanol:water

- pKa value (HH equation), calculating percent ionization of a drug at a specific pH
- π of functional groups
- **Lemke Rules:** Empirical approach to predict water solubility of molecules based on carbon solubilizing potential of several functional groups. If solubilizing potential > #C molecule is soluble, if < it is not. Functional groups interacting either through intramolecular H bonds, or Ion-Ion inside the same molecule decrease the solubility
- ClogP**:** Analytical approach to predicting LogP, by using the π values of all present functional groups which can have a positive (ex. Aliphatic carbon- nonaromatic) or negative (N, Amine) contribution to ClogP

Stereochemistry

- Stereoisomers: Enantiomers (mirror images, hard to separate) or diastereomers (non-enantiomers, easy)
- Eutomer: The eutomer of a drug is the more important enantiomer, distomer is less important
- Racemic Mixture: Equal parts of both enantiomers, FDA requiring marketing of one specific enantiomer, one in every 4 drugs in the market is an isomeric mixture
- **Easson-Stedman Hypothesis:** Drugs containing chiral centers administered as enantiomers, there must always be one eutomer which presents a minimum of three intermolecular interactions with receptor and the distomer only two.
- Stereoselectivity between enantiomers can also be attributed to the enantiomers ability to reach the target, because biological systems are chiral, absorption, metabolism, and excretion depend on chirality
- Conformational isomers can be rotamers which spatially rotate without breaking any bonds

Refinement of Lead Structure

Alteration of alkyl chains:

Chain length, branching, and rings can all be used to attempt to optimize a leads interaction.

- Adding methylene groups increases lipophilicity of molecule (increasing activity) and also decrease solubility (decreases activity). Varying chain length can also change receptor interactions, positively or negatively.
- Adding branches of methylene groups decreases flexibility of molecule, sometimes producing very different interactions with target
- Replacing CH3 with CH2CH2-Benzene increases potency of morphine 14x
- Positional isomers of substituents on aromatic rings often obtain different properties, especially ring substituents located orth to flexible side chains where they can participate in steric/electronic intramolecular interactions

Most popular strategy of drug design is synthesizing analogues of existing active molecules; the main drug of a medicinal chemist is creating analogues. The whole thing revolves around $$

Isosterism

TABLE 15.1 Groups of Isosteres as Identified by Langmuir

Groups	Isosteres
1	H^-, He, Li^+
2	O^{2-}, F^-, Ne, Na^+, Mg^{2+}, Al^{3+}
3	S^{2-}, Cl^-, Ar, K^+, Ca^{2+}
↓	↓
8	N_2, CO, CN^-
9	CH_4, NH_4^+
10	CO, N_2O, N_3^+, CNO^-

- **Allen** (1918): observed two molecules with the same molecular number shared many parameters
- **Langmuir** (1919): Introduces isosterism, focused on similarities of electronic and steric arrangement of atoms, groups, and radicles. Isosteres are those molecules that contain the same number and arrangement of electrons -most importantly in outer shell. Identified 21 isosteric groups
- **Grimm's Hydride Displacement Law:** From carbon, we can create isosteres for any next atom on the periodic table by adding a H, and in some way have properties like each other (not always), or even analogues with favorable properties

Table 1: Grimm's Hydride Displacement Law

C	N	O	F	Ne	Na
	CH	NH	OH	FH	-
		CH_2	NH_2	OH_2	FH_2^+
			CH_3	NH_3	OH_3^+
				CH_4	NH_4^+

- **Erlenmeyer:** Extended Grimm's law defining isosteres as atoms, ions, molecules which the peripheral layer of electrons can be considered identical
- **Classical Isoster:** Compounds or groups of atoms having the same number and arrangement of electrons

- **Heinsberg:** applied the concept of isosterism to entire molecules, and developed ring equivalents
- **Bioisosterism:** by Friedman: All molecules fitting the broadest definition of isosteres and eliciting similar biological activity. Used in analogue design, broad application and unpredictable depending on process and organism
- **Burger:** Expanded definition of Bioisosters, said "Bioisosters are compounds or groups that possess NEAR equal molecular shape and volume, approximately the same distribution of electrons, and similar physiochemical properties, transforming a chemical concept into a medicinal chemistry concept and statistical application

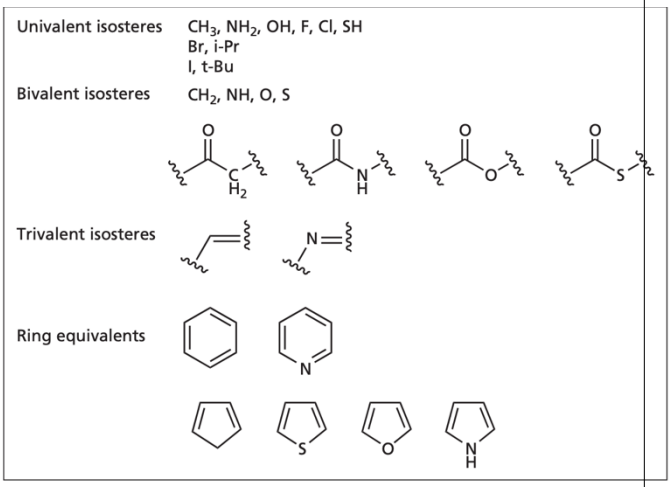
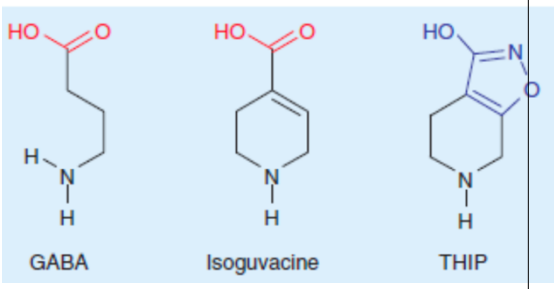


FIGURE 13.29 Examples of classic isosteres.

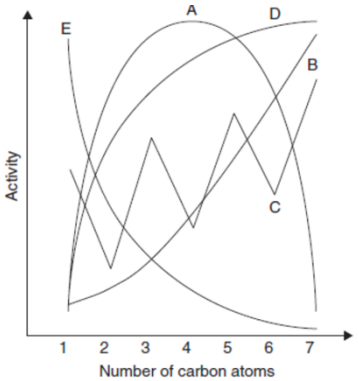
- Non-classical Bioisosters: Any Bioisosters not conformed to classical definitions but shown in experimental literature to work, trial and error process.



- Scaffold Hopping: Illustrated by diazepam, zolpidem and others acting on GABA-A, used recursivly to discober structurally novel ($) componenets from known bilogically active compounds.
    - Prodedure: Hopping is through HTS, SOSA, and serendipity and the prepared mind.
    - Virtual screening: Very important in hopping, through shape matching, pharmacophore searching, fragment replacement, and similarity search.
- **Iso/biosteric replacement:** From a lead compound with undesirable potency, toxicology, or PK parameters,
- the medicinal chemistry can replace or modify functional groups to achieve the desired properties

Homologous Series

- **Gerhardt:** Introduced the concept in medicinal chemistry that describes molecules differing only by a methylene group
- Monoalkylated Derivatives: X-R → X-CH2-R → X-CH2-CH2-R etc.
    - Ex. Neuraminidase inhibition based on methylene series
- Cyclopolymethylenic compounds: Increasing ring size
    - Ex. Optimal ring size of cyclonol carbamates
- Open, difunctional polymethylenic series: X—(CH2)n —Y → X—(CH2)n+1 —Y, where X and Y represent diverse functional elements
- Substituted cation heads: With cationic head groups, homology achieves simultaneously a progressive increase in bulkiness and in lipophilicity.



Biological Response Curves of homologous series

1. Most common are bell shaped (A) having a highest value around a certain number of carbons
2. The activity can increase, without any particular rule, with the number of carbon atoms (B).
3. 2. The biological activity can alternate with the number of carbon atoms, resulting in a zig-zag pattern (C).
4. 3. In other series, the activity increases first with the number of carbon atoms and then reaches a plateau (D).
5. 4. The activity can also decrease regularly, starting with the first member of the series (curve E). This was found for the toxicity of aliphatic nitriles or for the antiseptic properties of aliphatic aldehydes.
6. 5. A last possibility resides in inversion of the pharmacological activity accompanying the increase in the number of carbon atoms

Vinylogues and Benzologues

- Vinylogy formulated by Claisen, observed formyl-acetone has acidic properties similar to acetic acid
- The vinyl group plays the role of electron conducting channel between carbonyl and hydroxyl groups
- The vinylogy principle is explained by the Mesomeric effect
- Due to important changes in the geometry, vinylogues often have unpredictable activity. For this reason, vinylogues play a minor role in medicinal chemistry.

Twin Drug Approach

- During SAR studies of a lead compound, one approach is to combine two pharmacological entities into a twin drug
- Can be identical (homodimers) or non-identical (heterodimers),
    - Homodimers aim to be twice as potent especially for multimeric targets
        - Ex. Disprin, two aspirin molecules linked by acetyl region (no-linker)
    - Heterodimers target two related proteins in the same pathway, called symbiotic approach
        - Ex. Acetaminosalol (overlap)
- Can be classified by combination mode: Linker mode, no-linker mode, and overlap mode
- Preferred to taking two different pills
- Anchimeric assistance: when the binding of one group assists the binding of the other

Similarity

- Molecular similarities: structural features of a compound (shared structures, ring systems, topologies)
- Chemical Similarities: Based on the physiochemical similarities of the compounds (MW, logP, boiling point)
- Similar property principle: States by Johnson and Maggiora that structurally similar molecules tend to have similar properties
- Advantages: Important role in predicting the properties of chemical compounds, designing chemicals with a predefined set of properties, and in conducting drug design studies by screening large databases of structures
- Neighborhood Principle: Structurally similar molecules to a known biologically active molecule are likely to exhibit the same activity → example Morphine and Codeine. Many exceptions, but rule of thumb

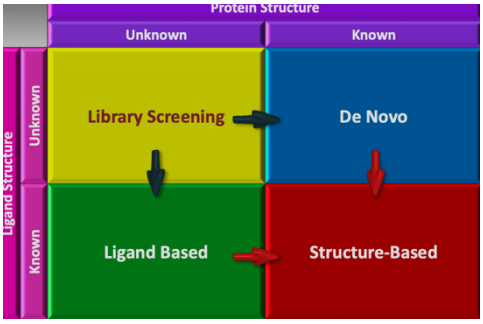
Similarity Measures:

- Main Difficulty: To quantify similarity between two molecules we need:
    - 1) A set of numerical descriptors used to compare the molecules
    - 2) a similarity coefficient which quantifies the degree of similarity based on the descriptors
- Chemical Descriptors: 0D(Bond count, MW) 1D(HBA/HBD) 2D(Topology) 3D(Geometry) 4D(3D coordinates)
- Structural Descriptors: Fingerprints
- Maximum Substance Substructure: Searching method for finding analogies using neighborhood approach
- Distance Measures:
    - Basic bit counts (only a, only b, both a and b, neither)
    - *Euclidean (√(both+neither)/total) → most common for chemical similarity, range from 0-1
    - Manhattan (√(only a + only b)/ total)
    - *Tanimoto (both/(only a + only b + both)) → most common for structural similarities, range 0-1
    - Tversky (both/total)
- Molecular recognition depends on 3D structure and properties (Electrostatic, steric) of the molecule rather than the underlying substructures, Ex Morphine and Methadone are only .20 similar using fingerprints

Computational Medicinal Chemistry

- Purpose: To speed up the hit lead finding, increase accuracy, reduce possible failures, and animal cruelty, lower cost, and identify novel targets
- Major advantages: Virtual screening and de novo design, in silico PK predictions, improved determining of PD

- Drug Discovery process: Target identification (Draggability, drug like, activity) → Hit to lead (Virtual screening, HTS analysis, ADMET prediction, SBDD, LBDD) → lead optimization → (QSAR, adverse in vivo effects prediction)
- Main Methods: Conformation search, Docking, Simulate annealing, Pharmacophore, 3DQSAR, Molecular dynamics, QSAR, COMBINE, Scoring functions, Homology modeling)
- Drug Design: Computational and synthesis tandem (Dry and wet lab)



LBDD

- Main Methods: QSAR, Pharmacophore, 3DQSAR
- Indirect, can only talk about the ligands because the target structure is unknown, applicability domain
- Rational Drug Design: Aims to generate an extremely active and selective compound that binds the only the active site of malfunction enzyme, inhibiting the progression of a disease.
- In the absence of 3D structure of drug target, LBDD is a popular approach to drug discovery and design, most important tools are 3DQSAR and pharmacophore modeling

Pharmacophore

- "A pharmacophore is the ensemble of steric and electric features that are necessary to insure optimal supramolecular interactions with a specific biological target structure and to trigger (or block) a biological response"
    - Errors: 1) we don't know the "interactions" 2) we don't know the "specific biological target"
- A Pharmacophore does not represent a real molecule, or association but is a purely abstract concept that accounts for the common molecular interaction capacities of a group of compounds toward their target. It can be considered the largest common denominator shared by a set of active molecules
- Result: Array of points connected in space, no chemistry, geometry of potentially biologically similar molecules
- Components:
    - 1) Rings, hydrophobic components
    - 2) HBD: H on O or N
    - 3) HBA: accessible lone pair on O or N
    - 4) Charged center
- Active analogue approach: Molecules fitting a pharmacophore might be active, filtering of active vs not
- BDZ Cook Model: Benzodiazepines on GABA, agonists, inverse agonists, antagonists defined in multiclass model for a single pharmacophore

QSAR History

- QSAR assumes that a correlation between calculated properties of the molecule and its experimentally determined biological activity exists. QSAR is used to predict the activity of new molecules
- First QSAR observed toxicity of alcohols increasing as water solubility (Cros), Later toxicity of organic compounds related to their lipophilicity (Meyer, Overton)
- Electronic parameter (Hammett, σ), Steric parameter (Taft, Es), Hydrophobic parameter (Hansch, LogP, π)

Hansch Equation

- Linear free energy approach, the ΔG of the reaction can be decomposed into smaller ΔG's (Steric, electric ect..)
- Not real QSAR because no structure involved, but able to relate calculated parameters to Ki and (1/C)
- Used 0D or 1D parameters about the chemistry of the ligand with no structural information

Free-Wilson Approach

- The first real QSAR, based on a fingerprint like matrix which quantifies the presence or absence of a particular feature to the biological activity
- There exist some Hansch/Free-Wilson combinations but they did not become popular

QSAR

- Steps Needed: Identification of active ligands and activity data → Identify molecular descriptors (fingerprints) → Establish mathematical expression relating descriptors to biological activity → Construction and validation of QSAR model
- Limitations: Must have congeneric series, same binding mode, lack of 3D structural info, statistical limitations
- Topless: Says we need at least 3 (ideally 5) experiments for each parameter to avoid chance correlation
- Statistical tools for model development and validation:
    1. Multivariable Linear Regression Analysis (MLR): used to assess the association between two or more independent variables and a single continuous dependent variable
    2. Principle component analysis (PCA): Dimensional reduction method to extract almost all the important information from the matrix into a non-colinear one that still fits the original. We can have multiple PC's with a maximum of 1/3 of the variables of the experiments, and if the Q-squared is less for that principle component, this principle component Is used
    3. Projection of latent structure (PLS): PCA taking also into account the Y matrix, allowing us to work and digest the data considering the Y variable, the latent structures refer to principle components
- Validation methods:
    1. Squared Correlation coefficient (R^2): Defines the fitness of the model, ideally >0.7
    2. Cross-Validation of R^2: Robustness of the model, internal prediction of Y-predict, Ideally > 0.5
        - Leave one out: Remove one, calculate Q^2
        - Leave some out: Randomly remove some, needs to repeat a lot of times
        - K-Fold: Divide into groups, remove one group
    3. Y-Scrambling: Determine the chance probability, good value less than original R and Q values
        - Randomize the Y values 50-100 times, then calculate Q^2 and R^2, if the scrambled values are always lower than the real one, model is good
    4. SDEP: Standard deviation error prediction, good value < 1
        - where we use a few experiments taken from the same dataset (removed before creating the model) and test how well the model predicts their activity
    5. Test on external data set
- Descriptors: Topological, Geometrical, Electronic, 3D- descriptors, Fingerprints

Equation

$$r^2 = 1 - \frac{\sum_{i=1}^{N}(Y_{\exp,i} - Y_{calc,i})^2}{\sum_{i=1}^{N}(Y_{\exp,i} - \bar{Y})^2}$$

$$q^2 = 1 - \frac{\sum_{i=1}^{N}(Y_{\exp,i} - Y_{pred,i})^2}{\sum_{i=1}^{N}(Y_{\exp,i} - \bar{Y})^2}$$

$$SDEP = \sqrt{\frac{\sum_{i=1}^{N}(Y_{\exp,i} - Y_{pred,i})^2}{N}}$$

3DQSAR

Grid approach to calculate MIF (Towards 3DQSAR)

- Goodfords grid: goodform was a crystallographer, there are some holes in PDB files. Good ford imagined to put a atom in a node on a grid, a probe atom was used to calculate van der Waals and electrostatic. Each point is assigned a number, and we can calculate which atom is in the cloud based on energy based approach
- Cramer applied a new method (DLOM), this same grid approach for small molecules, to see if a particular atom is stable in a particular position: Cramer invented the field based pharmacophore

Wolds PLS (towards 3dQSAR)

- From PLS we can reduce the dimensions into a few new highly informative entities called principal components

CoMFA/QSAR

- First 3DQSAR, comparative molecular field analysis
- Molecular properties are calculated using computer
- Parameters: Electrostatic field, and steric field
- Advantages over classical QSAR:
    - No reliance on experimental data
    - Can be applied to molecules with unusual substituents

    - Not restricted to molecules of the same structural class
    - Improved predictive capabilities
- QSAR procedure: Training Set → Alignment → MIF → Model generation → Internal/external validation
- Limits of CoMFA: Conformation of training set and superimposition is a long and slow process
    - Cramer used the trick: all molecules in training set are rigid, (steroid scaffold) so alignment is easier
- 4D (multiple conformations, orientations, and protonated states), 5D (inclusion of induced fit) and 6D (with solvents) also exist, but take much longer and are not significantly better than 3DQSAR

SBDD

- Must know the structure of the target, if we don't, we cannot do it
- Main Methodologies:
    1. Scoring Function: Sort of QSAR, but instead of putting the molecules in the grid, we put them in the receptor, and define the parameters from the molecule and the receptor, he called it hybrid QSAR
        - Used in molecular docking, to evaluate the bioactive conformation
    2. COMBINE: Comparative binding energy analysis, full structure based 3DQSAR method, instead of grid, molecules are placed in receptor and a MIF is generated and the receptors residues are all assigned a number based on importance – can be used to study mutations in receptor or compare similar receptors to each other allowing polypharmacophore. multidimensional method, allows multiple molecules compared to different receptors
    3. 3DQSAR: When molecules are aligned in the receptor space
    4. Proteochemometrics: Most potent for polypharmacophore; we need to know the fingerprints of the molecule and protein, which can be used to calculate cross descriptors between millions of parameters, and generate regions of the receptor important for ligand binding
    5. Molecular dynamics: Used to calculate the Ki of an inhibitor but takes a long time, used to refine leads, turbation method is used to calculate ΔG of different substituents in the receptor to increase potency
    6. Structure Based Pharmacophore Approach: Building pharmacophore with receptor information
    7. Homology Modeling: Build the structure of the target protein
- LBDD: methods allow screening of more compounds than SBDD, it's an important to refine before SBDD

Structure based Alignment:

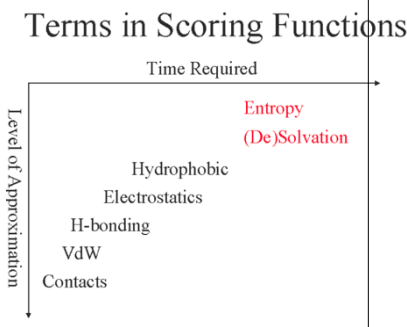
- Molecular alignment rules of Cramer: Rigid scaffold, very similar molecules → ease alignment
- Proteins and ligand complex alignment used RMSD of proteins, then proteins are removed leaving just the ligands aligned in a structure-based method
- This alignment can be validated using LB methods, seeing if the alignment is the same
- The RMSD of the aligned ligand to the original one, is the test of cross validation
- Another method of validation is cross docking, we remove a complex, and try to ask the program to redock the ligand into the other receptors
- Consensus scoring: Average agreement between various scoring functions of molecules aligned by LB and other by SB, the consensus is the one that has the lowest score in both because alignment of SB and LB are slightly different

Docking

- Uses the lock and key hypothesis
- Docking is the way to predict binding conformation of molecules but only predict, because we are using models, we can only say "This is likely how it binds"
- Needed for docking: A forcefield (to make conformational analysis) and scoring function (to quantify the score of the conformation based on the receptor)
- Scoring function is the weakness of any docking procedure, as they are generated only based on the training set and different parameters

- Molecular recognition is through electrostatic, Van der Waals, π-π/OH/cation, hydrophobic, and electromagnetic as well as metal ion coordination in some cases(poorly defined because of quantum mechanism cannot talk about the electron state in a precise way)
- Docking should be done on multiple software's to verify, as well as performing docking validations
- Binding mode: position, orientation, and conformation of the ligand in the protein
- Redocked molecules can be considered well docked if the RMSA is < 2 A, half docked if 2-3, misdocked 3+
- Docking Scoring Functions:
    - Physical based scoring functions: Free energy simulations, linear interaction energy
    - Empirical-based scoring functions: Regression based approaches
    - Knowledge based approach: Potential of mean force
- Used in virtual screening, de novo design, lead optimization
- Limitations: Rigid protein side chains, binding relevance for biological activity is not guaranteed, scoring functions
- Docking assessment: Best docked (Lowest energy), best cluster (Best docked in most populated cluster), or best fit (can software find the right confirmation
    - DA = # well docked + .5(# Half docked - #Well docked

Docking algorithms



- By ligand conformation: Rigid (protein-protein), fixed sample (ligand-protein), flexible ((molecular dynamics)
- By constraints: Residues and pockets, pharmacophore
- Charges and tautomer's
- Water molecules: Can be essential (mediate binding), optional (for some ligands required)
- Solvation and desolvation free energy critical for scoring functions
- Scoring Functions: Knowledge based, energy based, emery + parameterized solvation, free energy perturbation

COMBINE

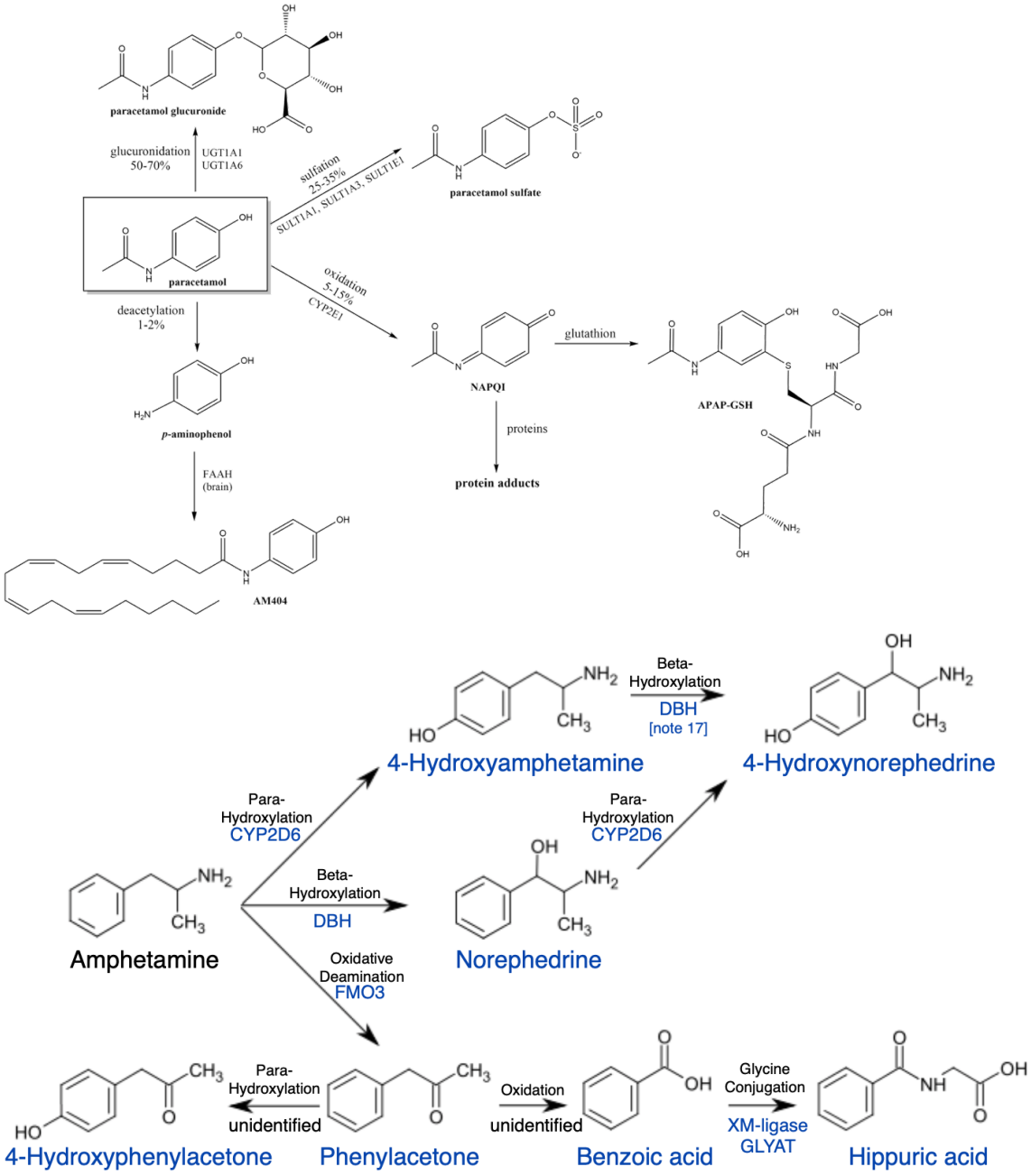
- For each residue and each point in the ligand we calculate the interactions inside the receptor
- We rank the residue in the receptor based on importance, transform the matrix using PLS method to build a COMBINE model with just the ligands ranked on binding affinity
- We then analyze the energy by histogram, only analyzed using plots
- COMBINEr: A way to calculate the interaction energy on the fly

Activity Contribution Plots

- Different residues contribute differently either decreasing or increase, based on electronic, Dry, or steric

Random vocabulary:

- Xenobiotics: A chemical substance found in an organism that is not naturally produced or expected to be there
- Blood Brain Barrier: Only Non-polar drugs can enter the CNS
- pKa: Constant for any molecule, the ratio of acid/base determined by the pH. Ionization determines absorption into different membranes


paracetamol glucuronide
glucuronidation
50-70%
UGT1A1
UGT1A6
sulfation
25-35%
SULT1A1, SULT1A3, SULT1E1
paracetamol sulfate
paracetamol
deacetylation
1-2%
oxidation
5-15%
CYP2E1
NAPQI
glutathion
APAP-GSH
proteins
protein adducts
p-aminophenol
FAAH
(brain)
AM404
Beta-
Hydroxylation
DBH
[note 17]
4-Hydroxyamphetamine
4-Hydroxynorephedrine
Para-
Hydroxylation
CYP2D6
Para-
Hydroxylation
CYP2D6
Amphetamine
Beta-
Hydroxylation
DBH
Norephedrine
Oxidative
Deamination
FMO3
Para-
Hydroxylation
unidentified
Oxidation
unidentified
Glycine
Conjugation
XM-ligase
GLYAT
4-Hydroxyphenylacetone
Phenylacetone
Benzoic acid
Hippuric acid