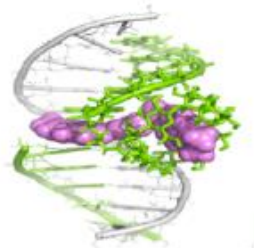
Medicinal chemistry- structure activity relationship

Prof M Kaur

Stages of Drug Discovery

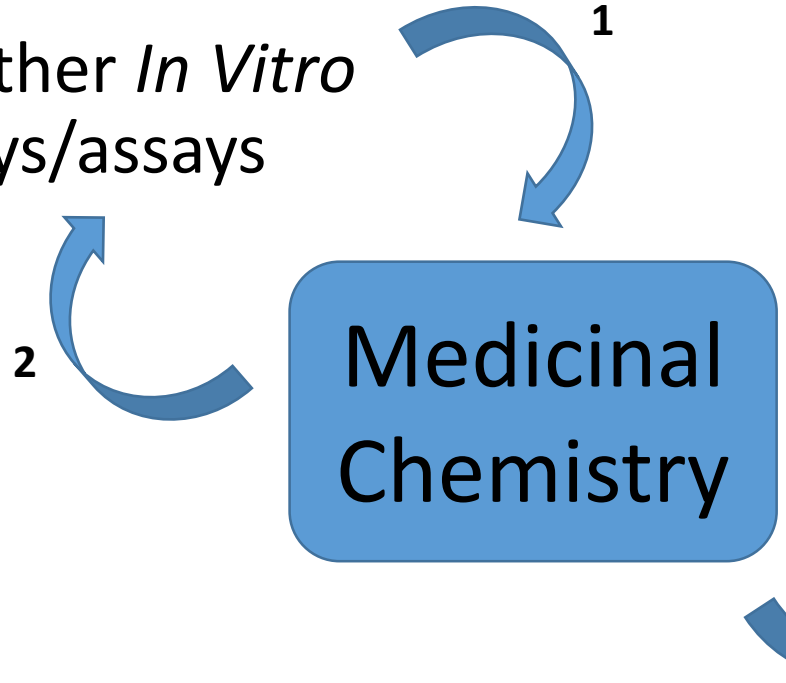
1	2	3	4	5	6
					
Disease identification	Therapeutic Target	Lead Identification -High throughput Screening -*In vitro* and *in vivo* pharmacology ...	Preclinical trials -Animal test -HTS	Clinical Trials	FDA Approval

High-throughput screening (HTS):
Lead identification and Preclinical toxicology

http://www.ibd.utm.my/wellness2012/pdf/presentation/ICBWI2010_D1P03_Prof_Dr_Shang_Tian_Yang.pd

Primary Bioassays

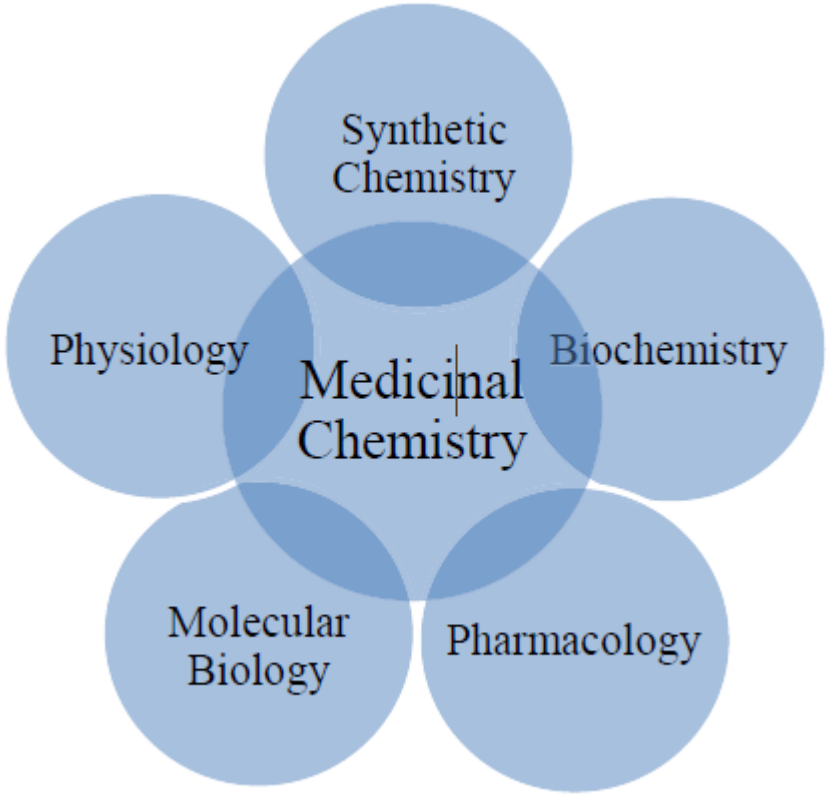
- Cell-based Bioassays
- Many other *In Vitro* bioassays/assays



Secondary Bioassays

- **Animal-based assays (*In Vivo*)**
- Toxicological Assessments in whole animals
- ADME Studies
- Behavioral Studies
- Preclinical Studies

Medicinal chemistry is a field focused on understanding the chemical basis of the biological effects of compounds by integrating fundamental concepts from different fields such as synthetic chemistry, biochemistry, pharmacology, physiology, and molecular biology



What are the design aspects of medicinal chemistry and how are they applied as a program moves forward?

Structure-Activity relationship
Structure-Property relationship

- a compound's biological properties are a function of its chemical

- Interaction between macromolecule and ligands
- Binding site- increase or decrease in strength
- Macromolecules are not static entities

solubility, membrane permeability, lipophilicity, and total polar surface area

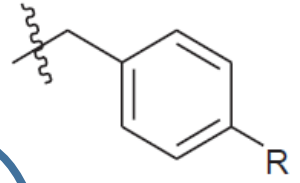
functional activity, and selectivity

The relationship between a series of compounds and their physicochemical properties is referred to as **structure–property relationships.**

Factors affecting potency- based on SAR

Hypothetical Example- SAR

Potency at the target is strongly influenced by the electronic nature of the R-group. Electron-donating substituents, such as methyl and methoxy, decrease potency, while electron-withdrawing groups, such as chlorine, bromine, or trifluoromethyl, increase potency.



Electron withdrawing groups

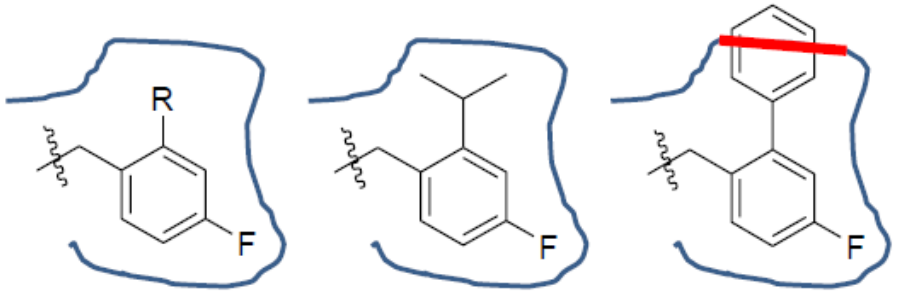
Entry	R	IC_{50} (nM)	Entry	R	IC_{50} (nM)
1	Methoxy	10000	4	Fluoro	50
2	Methyl	1000	5	CN	20
3	Chloro	100	6	NO_2	2

FIGURE 5.2 Altering the substituents on an aromatic ring can have a significant impact on biological activity. In this hypothetical set of compounds, potency at the biological target increases as the benzene ring becomes increasingly electron poor. These data suggest that additional electron-donating R-groups such as an NH_2 would have lower potency than R-groups that are electron withdrawing, such as a CF_3.

If benzene ring is directed into solvent space (outside of binding pocket), then substitutions will have little impact on potency

Conclusion- new compounds should contain aromatic rings with electron-withdrawing groups in the position shown

Physical restrictions of a ligand-binding site can be elucidated by examining binding data



Entry	R	IC_{50} (nM)	Entry	R	IC_{50} (nM)
1	H	1100	4	Isopropyl	20
2	Methyl	500	5	t-Butyl	50
3	Ethyl	200	6	Phenyl	10000

FIGURE 5.3 In any given ligand-binding site, the "walls" of the binding site are comprised of the backbone and side chain features of the macromolecular target. In this example, the outer perimeter of the biding cavity is designated by the blue line. Binding of the hypothetical ligand becomes more efficient as the R-group fills the available space and hydrophobic interaction increases (entries 2–5). Exceeding the size of the cavity (entry 6), however, will cause a dramatic loss in binding efficiency, as the candidate compound no longer fits within the allowed space (as indicated by the red line).

Conclusion: Predict that compounds in which the R-group is an aromatic ring will be significantly less potent than those where smaller substituents are employed, irrespective of the nature of the appended aromatic ring. There is not enough room in the binding site to accommodate larger groups in this position.

Other considerations

- Not two but three-dimensional space
- Concept of chirality introduced by Lord Kelvin 100 years ago.
- In this scenario, the substituents can be attached to the central carbon atom in one of two arrangements, which are non-superimposable mirror images of each other.

Role of Chirality

Chiral molecules usually contain at least one carbon atom with four non-identical substituents.

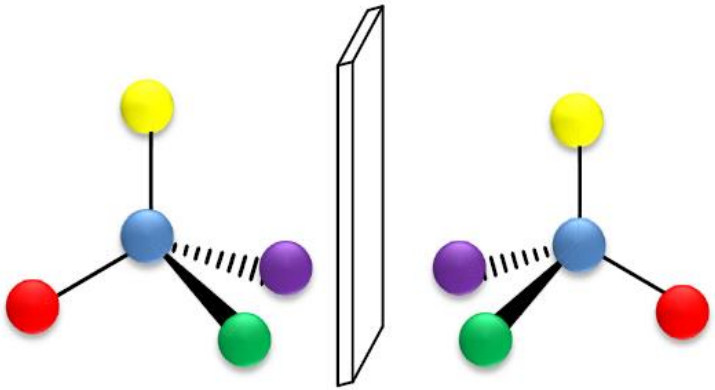


FIGURE 5.5 Chiral compounds cannot be superimposed on their mirror image. The two isomers are referred to as enantiomers (R and S). Chirality often plays a key role in biological activity, as nature is a chiral environment.

Biological macromolecules, being chiral, offer an environment that predictably discriminates between enantiomers or optical isomers, leading to dramatic changes in the functional activity between pairs of enantiomers.

Example

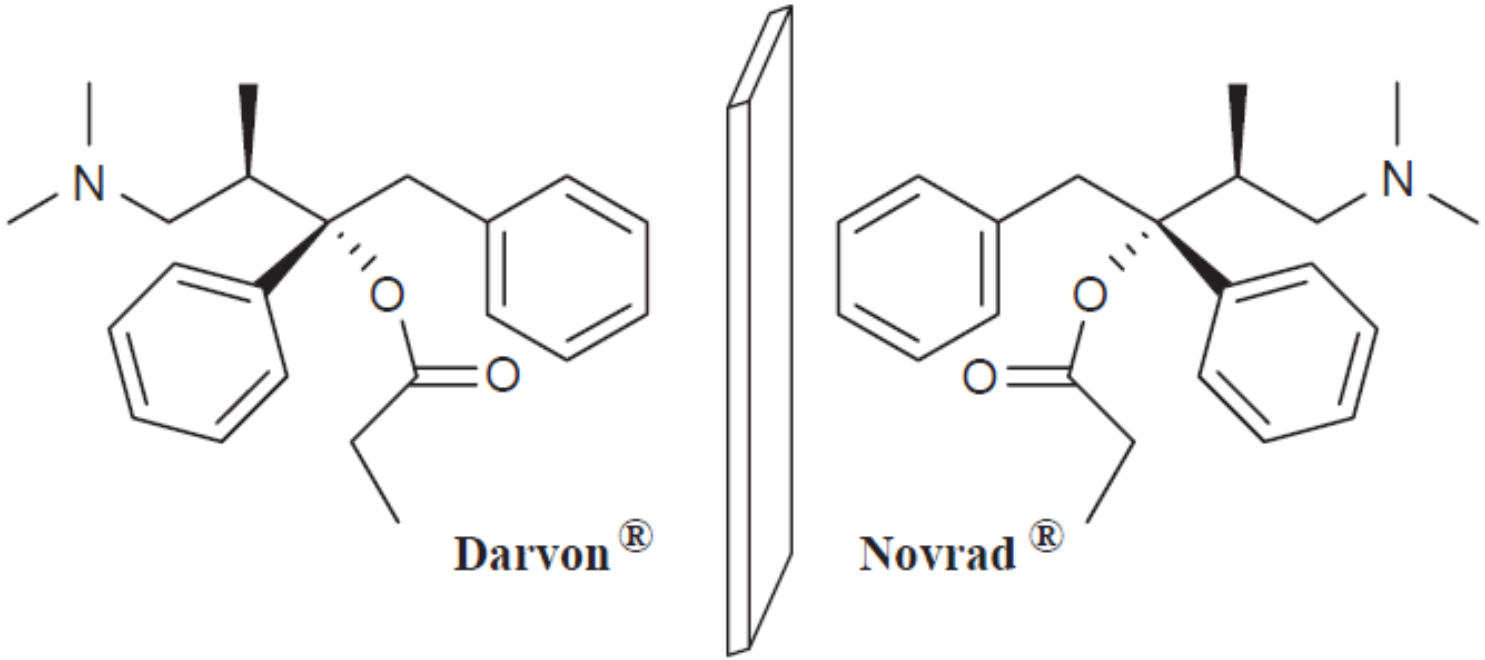


FIGURE 5.6 Darvon® and Novrad® are identical, save for the nature of the chiral center present in each. The enantiomers have very different properties in a biological setting. Darvon® is an analgesic that activates the μ-opioid receptor, while Novrad® is an antitussive agent with minimal efficacy as an analgesic. Unlike its enantiomer, Novrad® has little affinity for the μ-opioid receptor.

Antitussive- cough reliever

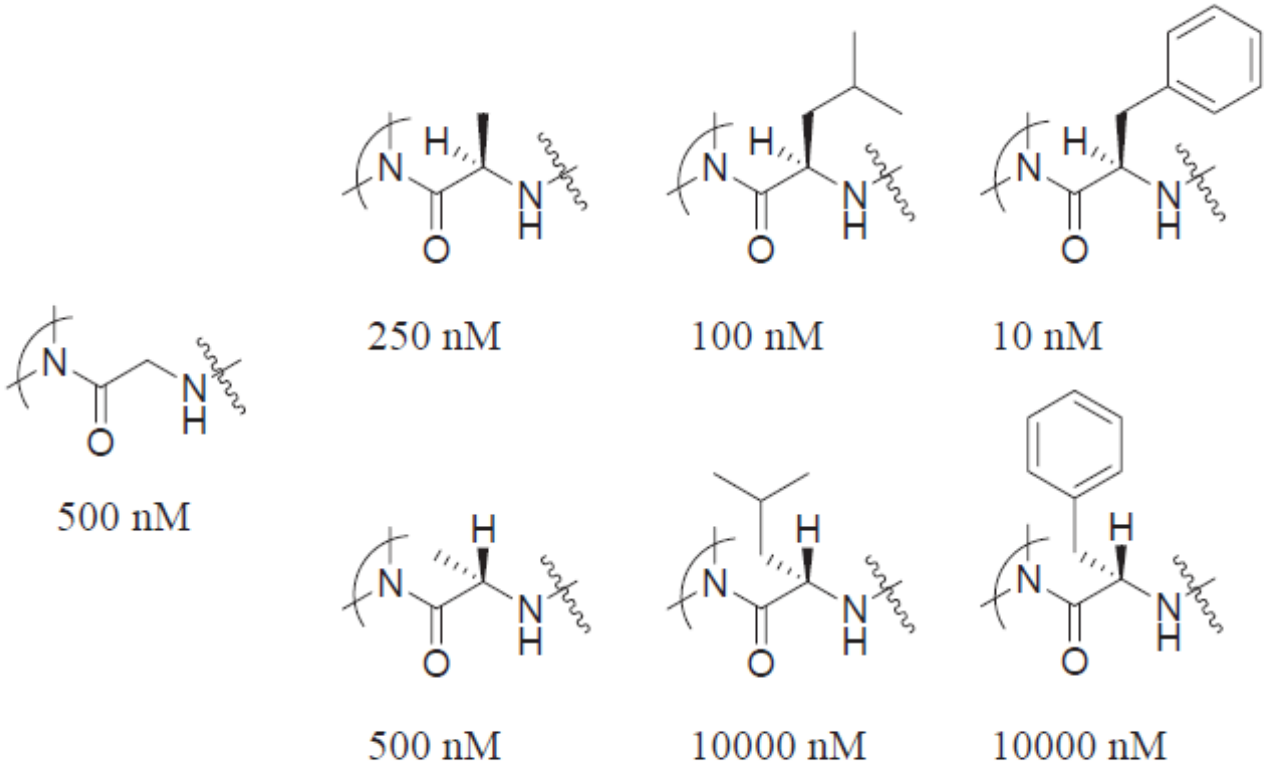


FIGURE 5.8 In a series of otherwise identical compounds, differences in chirality can have a significant impact on biological activity. Although an X-ray crystal is a valuable tool in identifying areas where changes in chirality will have the largest impact, it is not required. In this hypothetical example, potency differences between the R-series (top) and the S-series (bottom) clearly demonstrate that the R-series is heavily favored.

Ligand-binding site strongly favours one of the two possible enantiomers. *In vitro* potency improves with increased steric bulk in one enantiomeric series (the R-series, top), but the same changes in the opposite enantiomeric series (the S-series, bottom) lead to significant loss of potency.

Conclusion: predict that compounds with the S-configuration at this chiral center would be significantly less potent than the R-series and should be deprioritized when considering potential candidate compounds.

THE PHARMACOPHORE

- International Union of Pure and Applied Chemistry (IUPAC) defines pharmacophore as “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.”

- The remainder of the compound, portions of which may provide structural support for the position of the functionality within the pharmacophore but play no direct role in binding events, is often referred to as the auxophore.

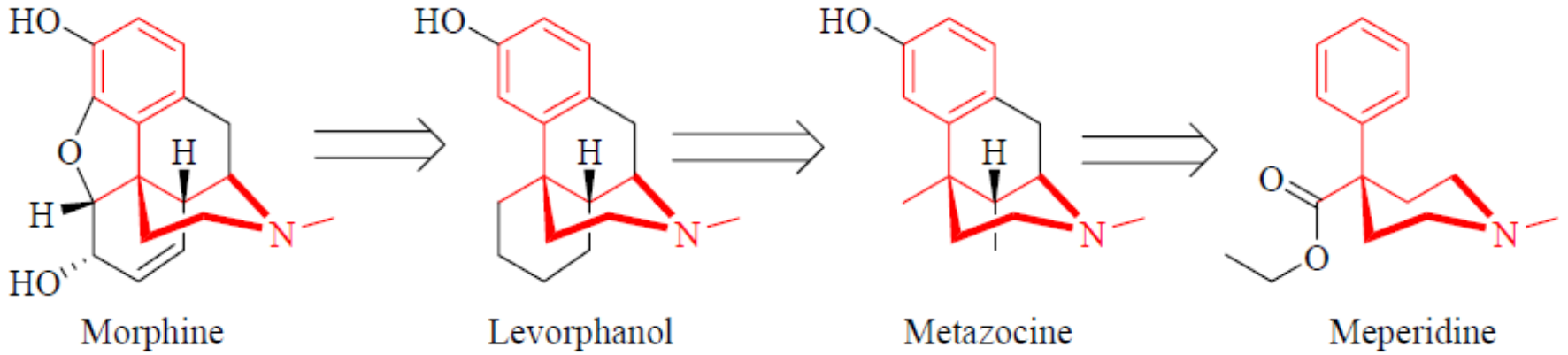
Red- Pharmacophore (binding)

Black- Auxophore (structural support)

How changes in pharmacophore or auxophore affect biological activity?

three to four times *more* potent than morphine

molecular elements responsible for rigidity and shape have been removed



Has 10–12% of the analgesic potency of morphine

FIGURE 5.12 A pharmacophore can be described as the minimum structural requirements that must be present in order for a compound to have biological activity at a given target. The μ-opioid receptor family of analgesics, typified by morphine, levorphanol, metazocine, and meperidine, can be described by the minimum pharmacophore outlined in red in each structure. Features outlined in black can have a significant impact on biological activity, but can also be modified (levorphanol), or even removed entirely (metazocine and meperidine), while still maintaining the ability to bind to the μ-opioid receptor. Loss of the features in red, however, leads to a loss of μ-opioid receptor binding.

maintaining the overall three-dimensional shape of candidate compounds is important

How does one begin the process of developing an SAR database that can be used as a basis for design strategy and project decisions?

- HTS- either physical or virtual

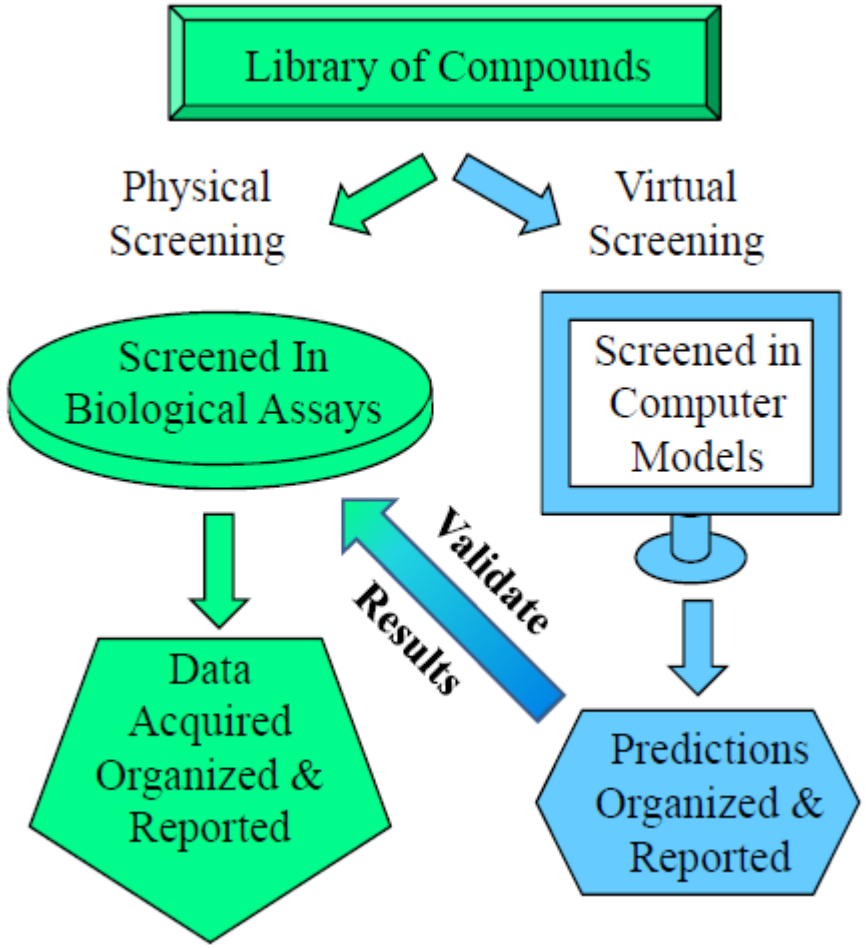


FIGURE 5.14 An initial data set for a drug discovery program can be derived from a physical screening of a compound library or a virtual screening of the same library if an appropriate computer model is available. Each method has advantages and disadvantages that should be considered, and neither method is exclusionary of the other. It is, however, very important to validate the *in silico* screening results with physical screening methods to ensure that the virtual screening is truly predictive of real-world results.

Which screening method?

PHYSICAL SCREENING

- 100,000 compounds
- 12-point dilution curve to determine IC50 values at a biological target, this would produce
- 2.4 million data points when run in duplicate, not including positive and negative controls used to ensure assay integrity.

7500 microtiter plates, thousands of disposable pipette tips for reagents – **COST??????????**

VIRTUAL SCREENING

ZINC Database

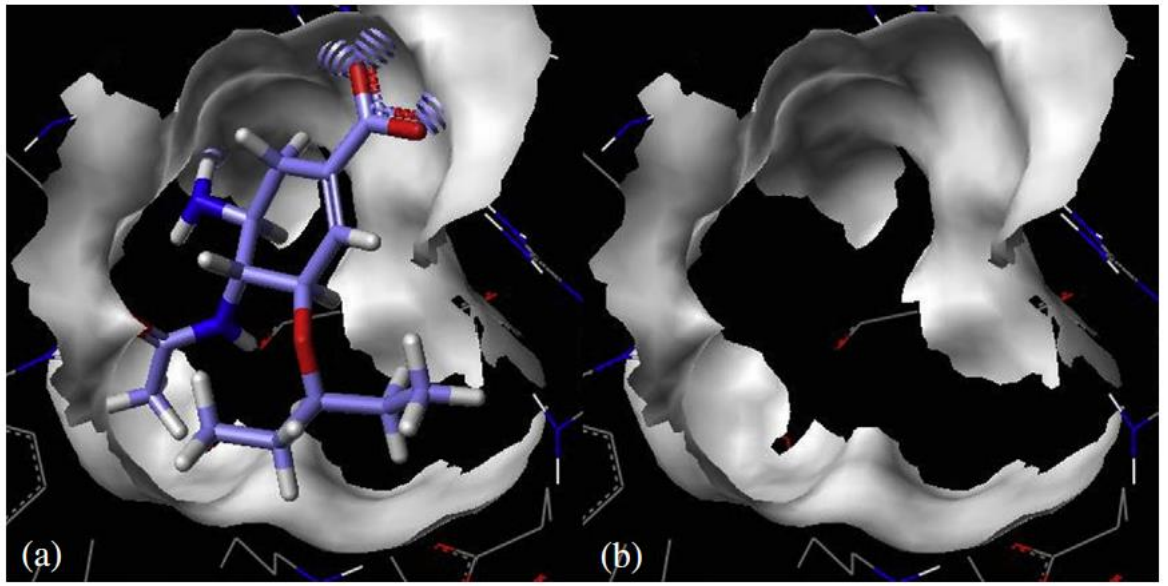


FIGURE 5.15 (a) An X-ray crystallographic structure of influenza virus neuraminidase with an inhibitor bound to the active site provides information that can be used in the design of additional inhibitors. The boundaries of the binding site (gray) and important interactions are displayed, including hydrogen bonds (barrels). (b) *In silico* excision of the inhibitor from the binding site using protein modeling software provides a template for virtual examination candidate compounds. RCSB 2QWK.

DOCKING- Binding Score

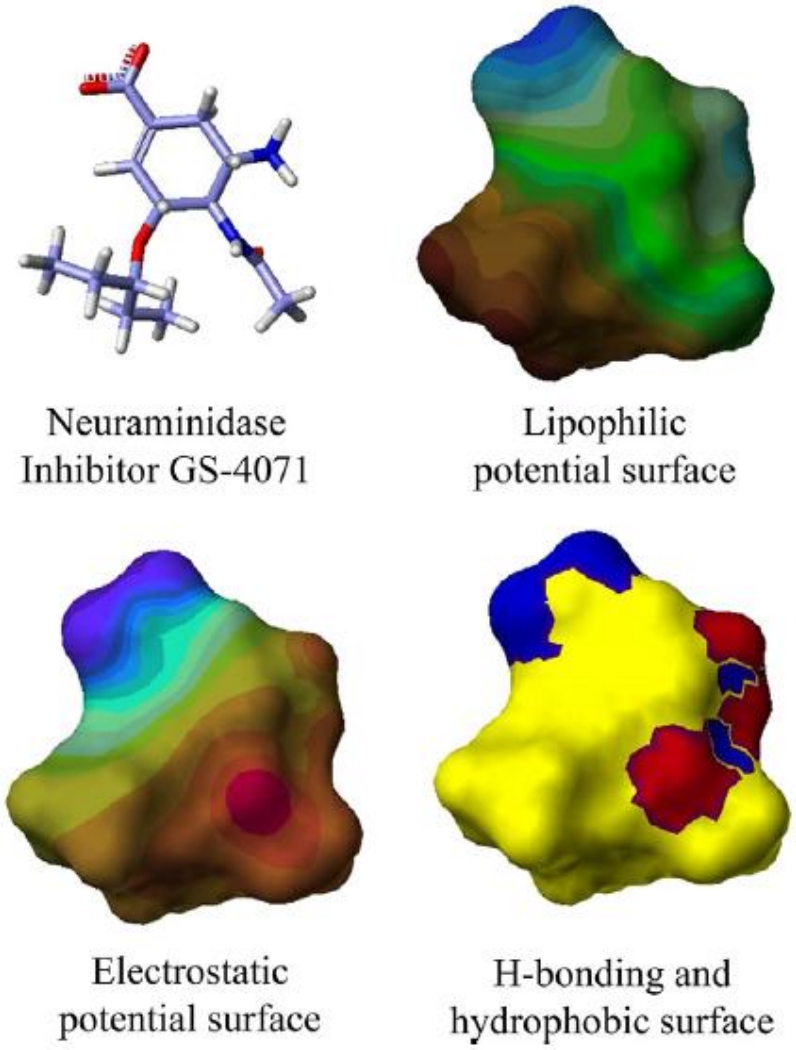


FIGURE 5.17 The neuraminidase inhibitor GS-4071 is structurally described by the classical stick figure (top left), but computer modeling can provide improved visualization of molecular properties such as lipophilic potential surfaces (top right), electrostatic potential surface (bottom left), and hydrogen bonding/hydrophobic surfaces (bottom right). These "maps" can be combined to provide a model system useful for the identification of compounds that are similar to GS-4071.

property sets are determined for all of the members of the compound collection, the individual members of the compound collections are compared with the lead compound, and each is assigned a score based on its overall similarity to the lead compound. Compounds with higher scores (those that more closely resemble the lead compound) can be identified in this manner and similar compounds can be grouped together to select compounds for physical screening, once again significantly decreasing the overall cost of physical screening.

Correlating scores and inhibitory concentration of compounds

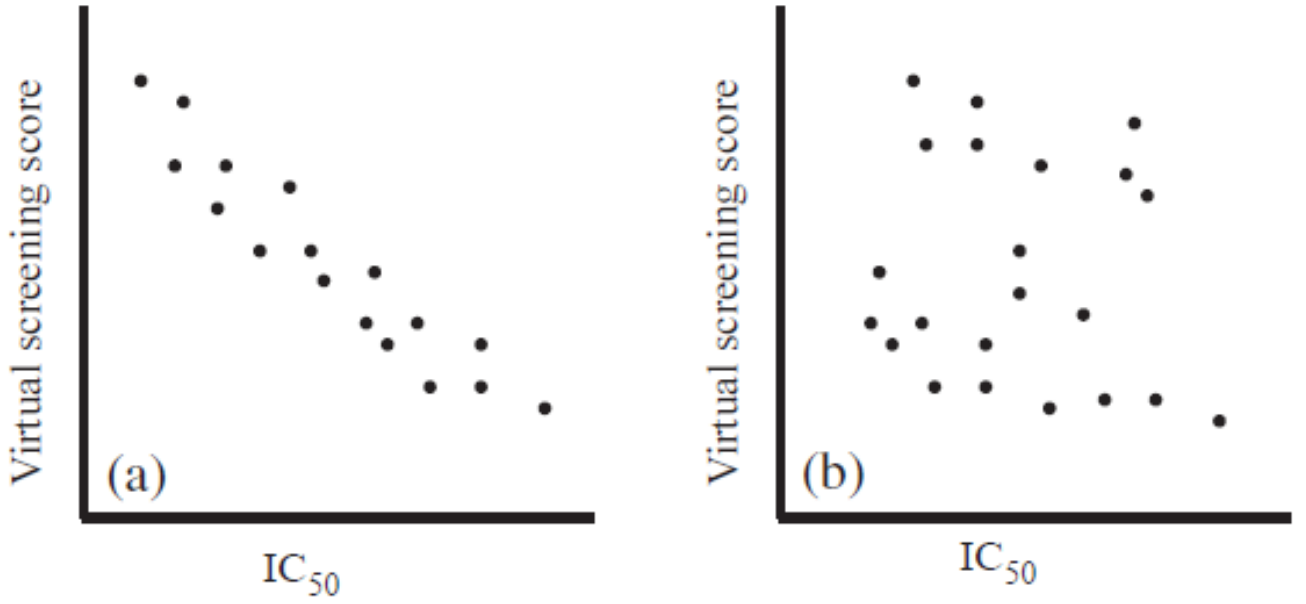


FIGURE 5.18 In order for a virtual screen to be useful, it must correlate with real-world results (a). In the absence of this correlation (b), data provided by the virtual screen cannot be used to predict biological activity of candidate compounds.

Fragment-based Screening

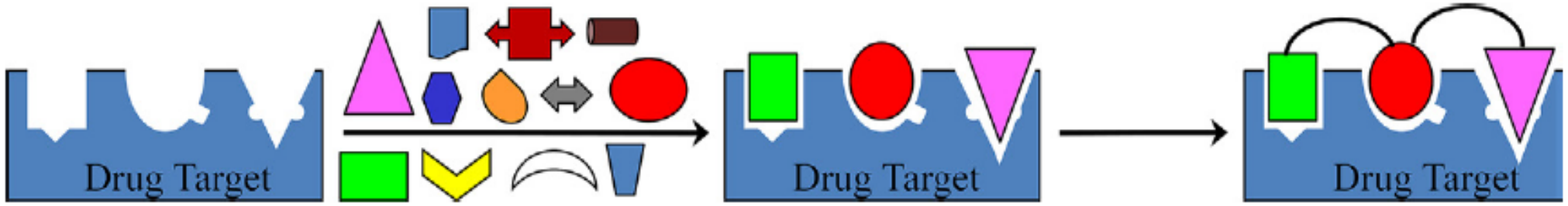


FIGURE 5.20 In one approach to fragment-based screening, low-molecular weight compounds are screened to determine binding affinities using sensitive screening techniques such as NMR, SPR, and X-ray crystallography. Low-molecular weight compounds that demonstrate binding are then stitched together with linker units designed to allow the individual components to access their individual binding sites. The tethered individual units act collectively to create a single compound with improved biological activity based on synergistic binding.

Just as the individually weak interaction in protein structure can collectively form very stable three-dimensional structures, joining together multiple fragment structures can have a synergistic effect, leading to tighter binding than is observed for the individual fragments, possibly greater than the sum of the individual components

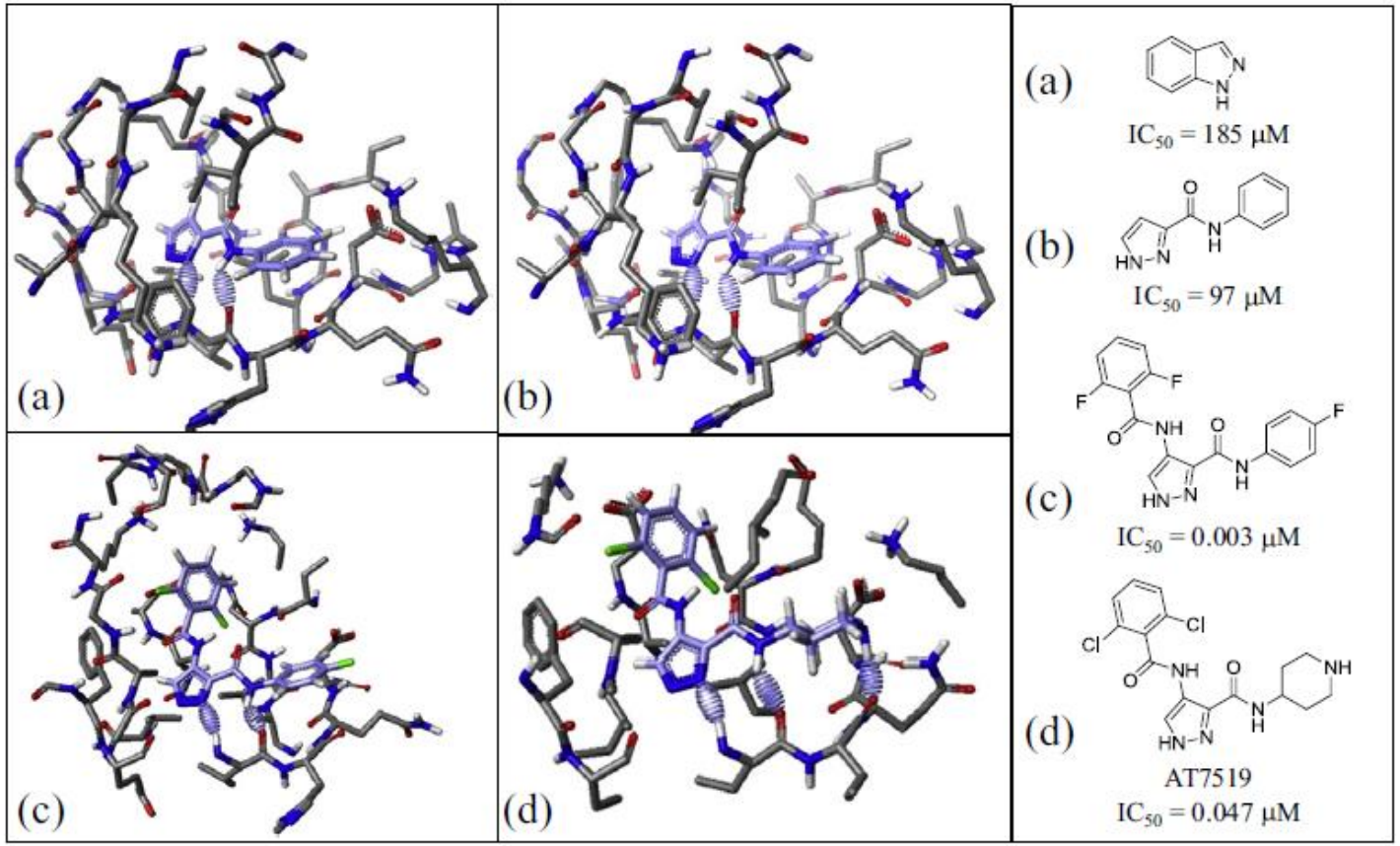


FIGURE 5.21 (a) Screening of a fragment library via X-ray crystallography identified indazole as a weak CD-2 binder (RCSB 2VTA) anchored into the binding site with a hydrogen bond (cyan barrel). (b) Removal of the benzene ring and growth of the fragment via addition of an N-phenyl-acetamide provided access to a neighboring binding site and an additional hydrogen bond interaction (RCSB 2VTL). (c) Growing the structure in the opposite direction by adding a benzamide created additional opportunities for hydrophobic interactions in a proximal section of the binding site, leading to a dramatic improvement in potency (RCSB 2VTP). (d) Incorporation of a piperidine to increase solubility and cellular activity led to the identification of the clinical candidate AT7519 (RCSB 2VU3).

Indazole was identified as a fragment “hit” with very low binding potency (185 μM) for cyclin-dependent kinase 2 (CDK-2) and eventually transformed into AT7519, a significantly more potent CDK-2 inhibitor

Structure-Activity Relationship Cycle

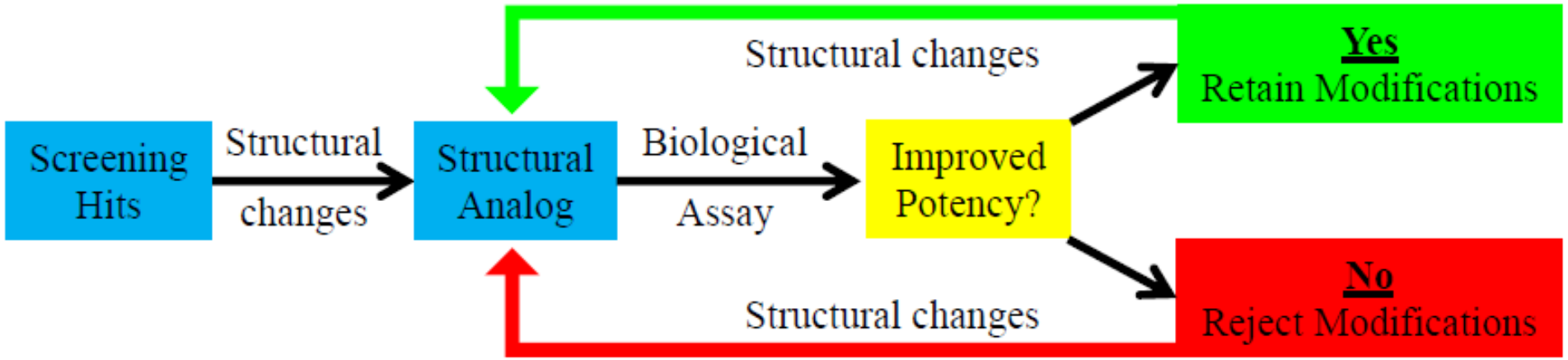


FIGURE 5.22 Identifying an initial set of hits in a drug discovery program, irrespective of their source and methods employed, is generally the beginning of an iterative process. Analogs containing specific structural changes are prepared and assessed for biological activity. Favorable changes are maintained, unfavorable changes are discarded, and a new round of compounds is prepared. Each successive round of synthesis and biological assessment builds on the previous round with the ultimate goal of optimizing biological activity.

Importance of SAR- an example

- While targeting Kv1.5, a voltage-gated potassium channel that has clinical relevance in atrial arrhythmia, optimizing this activity would be necessary. It would also be necessary, however, to establish that candidate compounds did not negatively impact the hERG channel, a voltage-gated potassium channel associated with Torsades de pointes and sudden cardiac death.
- In this instance, a structure– activity relationship could be established for activity at the hERG channel, focusing on minimizing the potency at this channel. Thus, structural changes that *decreased* hERG activity would be maintained, while those that increased hERG activity would be dropped. Successful compounds would be identified by analyzing the SAR of *both* Kv1.5 and hERG in an effort to maximize one while minimizing the other.

“DRUGLIKE” GUIDELINES

- It is estimated that there are over 10^60 synthetically accessible small molecules
- Christopher Lipinski and his colleagues examined the physicochemical properties of over 2000 drugs and drug candidates in clinical trials at the time (mid-1990’s).
- Lipinski and his colleagues hypothesized that successful drugs and clinical candidates occupied specific portions of the larger chemical universe and that compounds that failed as a result of poor pharmacokinetic profiles existed outside of this “druglike” chemical space.

Lipinski Rule of Five- "druglike"

1) Molecular weight <500 AMU
2) LogP below 5
3) Fewer than 5 hydrogen bond donors
4) Fewer than 10 hydrogen bond acceptors
5) Fewer than 10 rotatable bonds

Assumption- compounds are absorbed into cells via passive diffusion

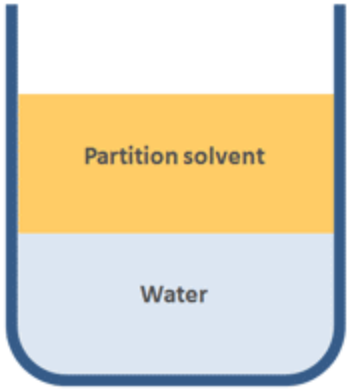
LogP- the logarithm of the partition coefficient between water and 1-octanol; a measure of lipophilicity

Assessing Lipophilicity

- The lipophilicity of a compound can be assessed by determining its partition ratio in an 1-octanol/water system.
- Compounds that are more lipophilic will be more heavily concentrated in the 1-octanol layer of the system, while less lipophilic (more polar) compounds will preferentially occupy the water layer.

- Partition Ratio is easily determined with software packages and expressed in a logarithmic scale. It is most often designated as LogP (or cLogP), and refers to the calculated value under Neutral (un-ionized) conditions (pH = 7).

$$\text{P = Partition Coefficient} = \frac{\text{Concentration dissolved in partition solvent}}{\text{Concentration dissolved in water}}$$



Conditions:
The solvents are "immiscible"
The system must be at equilibrium
All the solute must be dissolved
Temperature should be constant

high logP values cause poor absorption or permeation

- A negative value for logP means the compound has a higher affinity for the aqueous phase (it is more hydrophilic);
- when logP = 0 the compound is equally partitioned between the lipid and aqueous phases;
- a positive value for logP denotes a higher concentration in the lipid phase (i.e., the compound is more lipophilic).

LogP = 1 means there is a 10:1 ratio Organic : Aqueous

A desired logP value (octanol-water partition coefficient) is no more than 5
logP 5 = 100,000:1 concentration difference between octanol and water phases).

Medicinal chemistry in combination with screening science provides the tools to maximize the positive traits and minimize the negative characteristics of candidate compounds.

Selected compounds are suitable for further study (e.g., pharmacokinetic studies, *in vivo* animal models of efficacy, and safety studies)