

PERSPECTIVE ON DRUG DISCOVERY & DESIGN

What is a drug ?

FDA Definition of a Drug

“ An active ingredient that is intended to furnish pharmacological activity or other direct effect in the diagnosis, cure, mitigation, treatment, or prevention of a disease, or to affect the structure of any function of the human body, but does not include intermediates used in the synthesis of such ingredient. ”

Over-the-counter (OTC) drugs: sold without a doctor's prescription

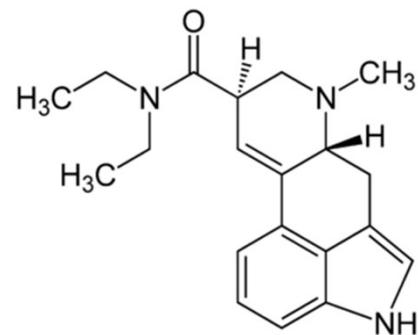
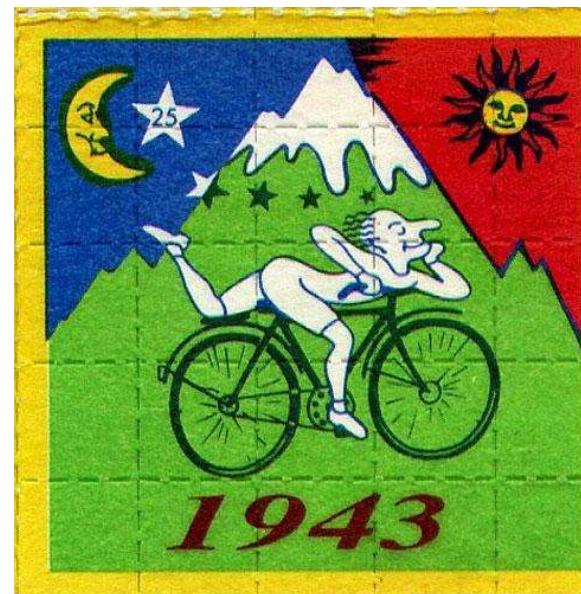
Ethical drugs: prescription drugs

Biologics: drugs that are biomolecules like antibodies, proteins, peptides, nucleic acids, etc...

Illegal drugs: possession, use or commerce may be restricted or forbidden

FDA: Food and Drug Administration of the USA

What is a drug ?



LSD 25

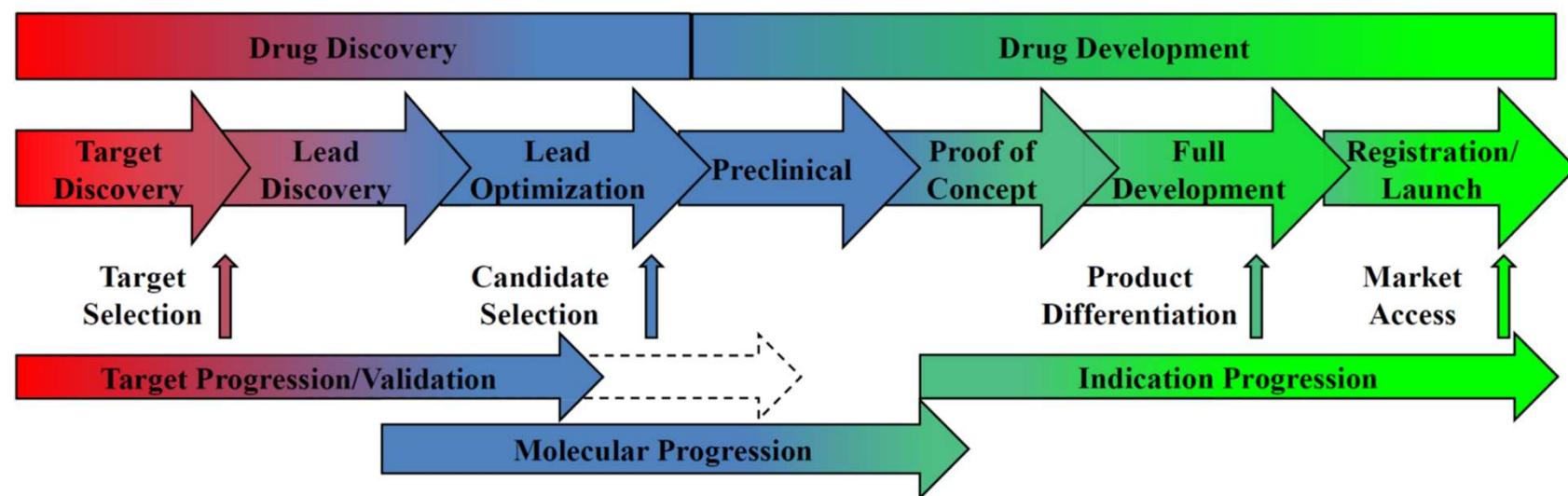
What is drug design & discovery ?

- **Drug discovery** – a generic term that encompasses all activities leading to new substances with pharmacological activity, be it natural substances, synthetic compounds, found by chance, search or design
- **Drug design** – a more specific term that refers to the process of creating new drugs, through a combination of biological, chemical and computational techniques

Drug Discovery *versus* Drug Development

Drug discovery - all of the experimentation and studies designed to move a program from the initial identification of a biological target and associated disease state to the identification of single compound with the potential to be clinically relevant.

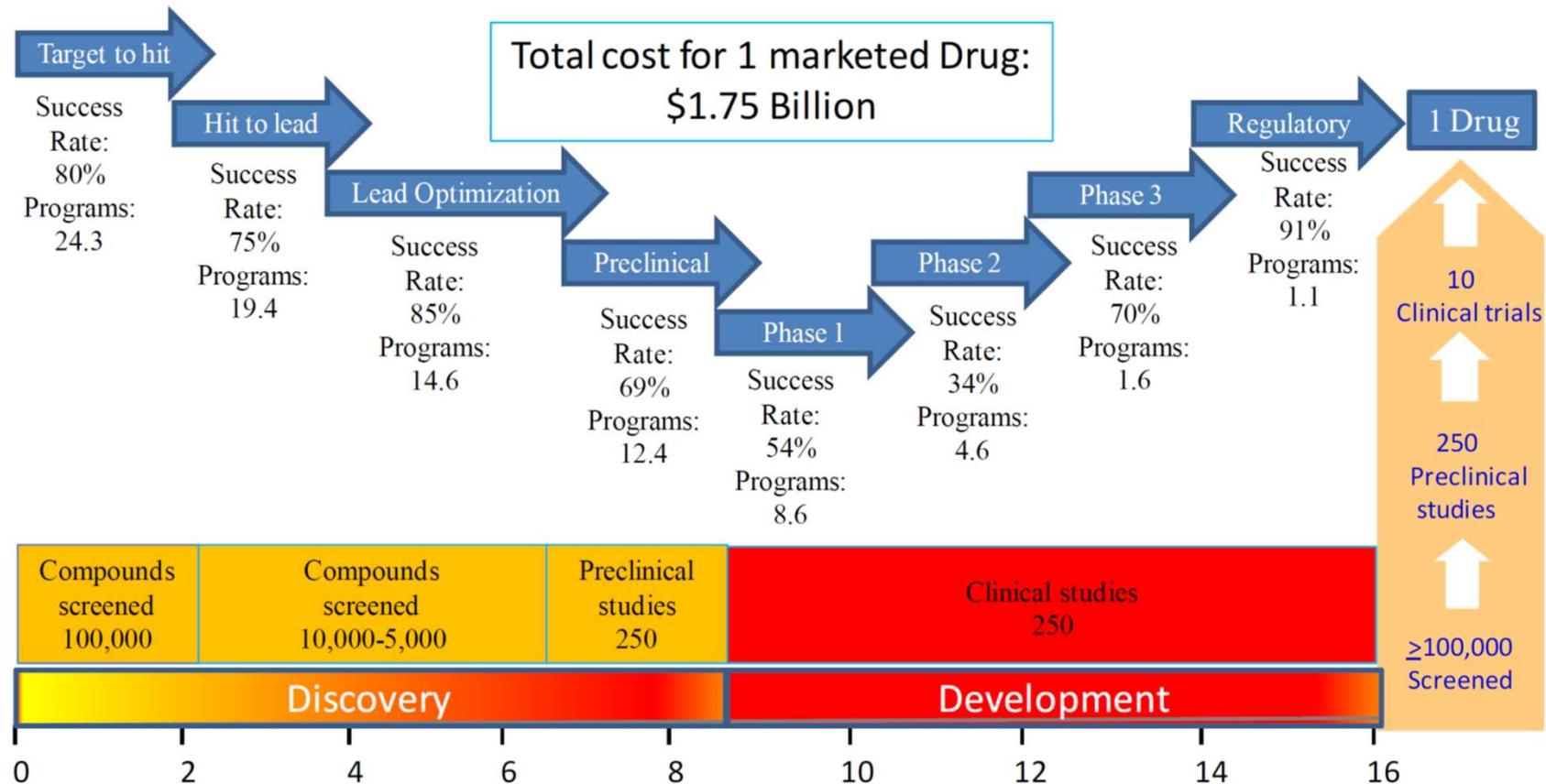
Drug Development - typically begins once a single compound has been identified, which is then progressed through various studies designed to support its approval for sale by the appropriate regulatory bodies.



The challenge of DD

- The task of discovering new drugs is hard, expensive, lengthy and dependent on a very large number of scientific disciplines, techniques and expertise.
- Millions of compounds may have to be screened in activity tests to select but a few candidates (hits), of which only a few show promise as drug candidate (leads).
- Lengthy and thorough clinical testing in both animals and humans is required, without guarantee of approval by the regulatory entities.
- Millions (or billions) of dollars and ~5-15 years are required for the whole process.
- A large share of the profit generate by the pharmaceutical industries comes from only a few drugs.
- Patent expiry reduces narrows the profitability range of drugs and pushes the “me too” drug concept

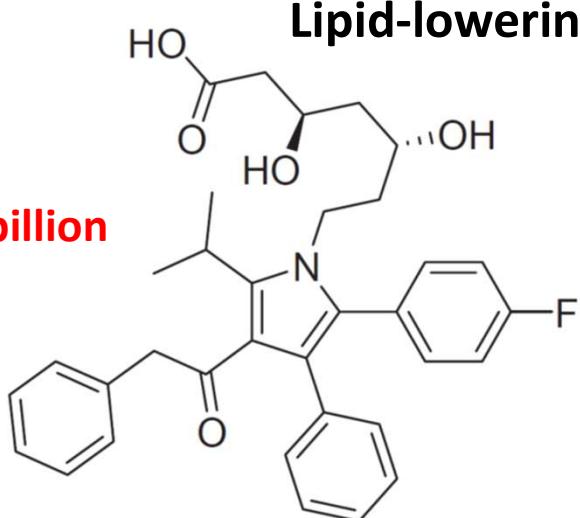
The path to a new drug



Commercially successful drugs

Lipid-lowering

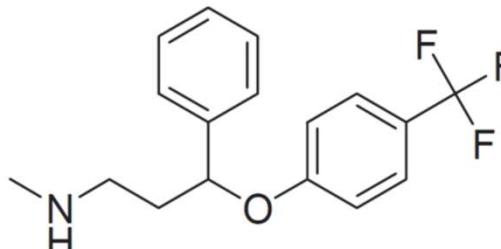
\$13 billion



Lipitor®
(Atorvastatin)

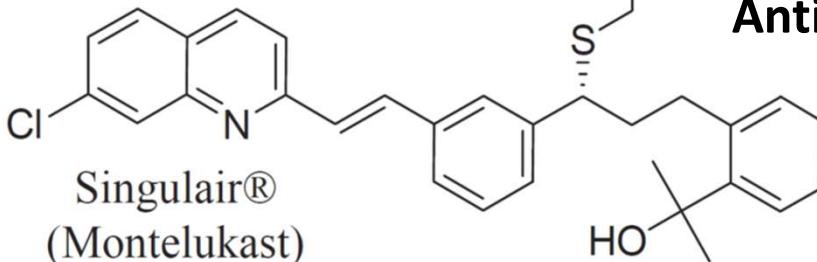
Anti-depressant

\$2.8 billion



Prozac®
(Fluoxetine)

Anti-asthma

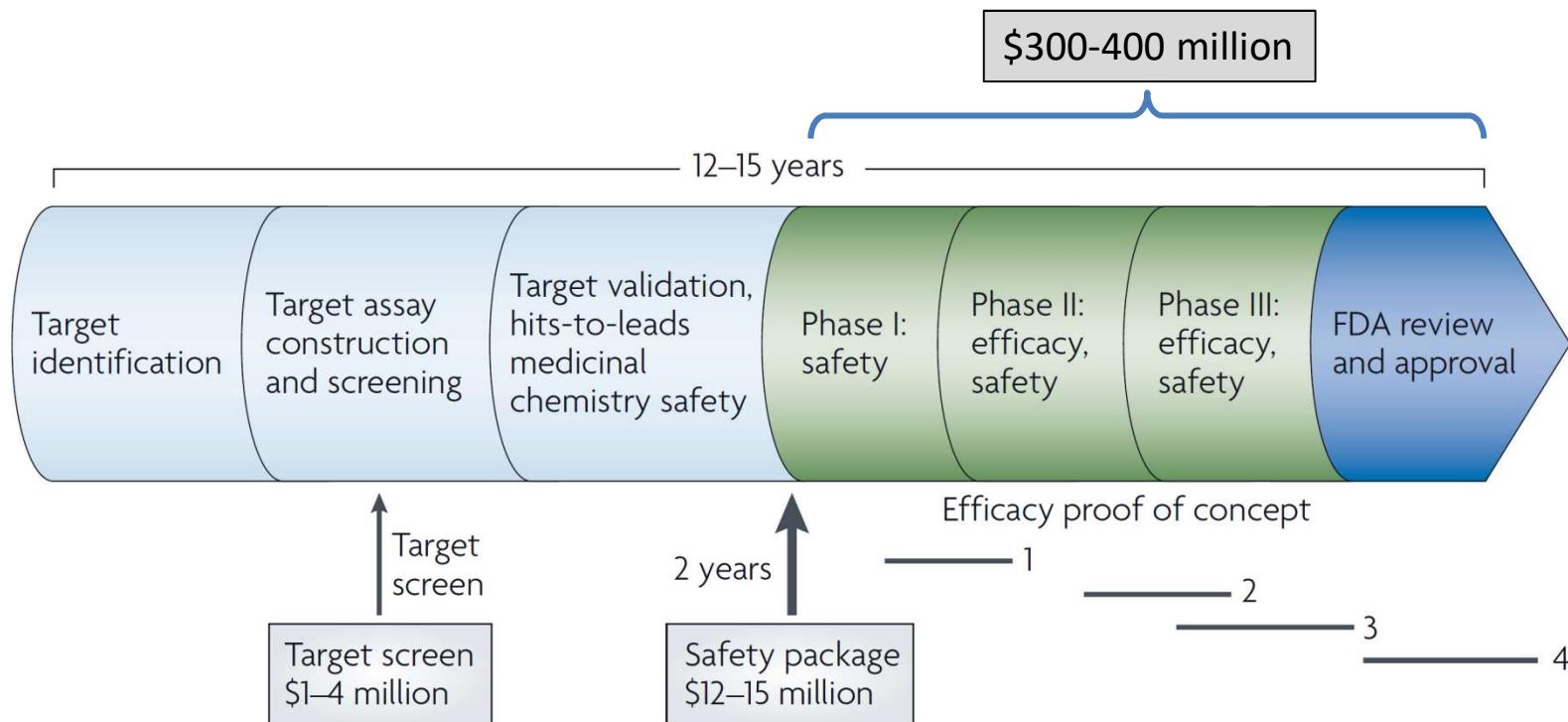


Singulair®
(Montelukast)

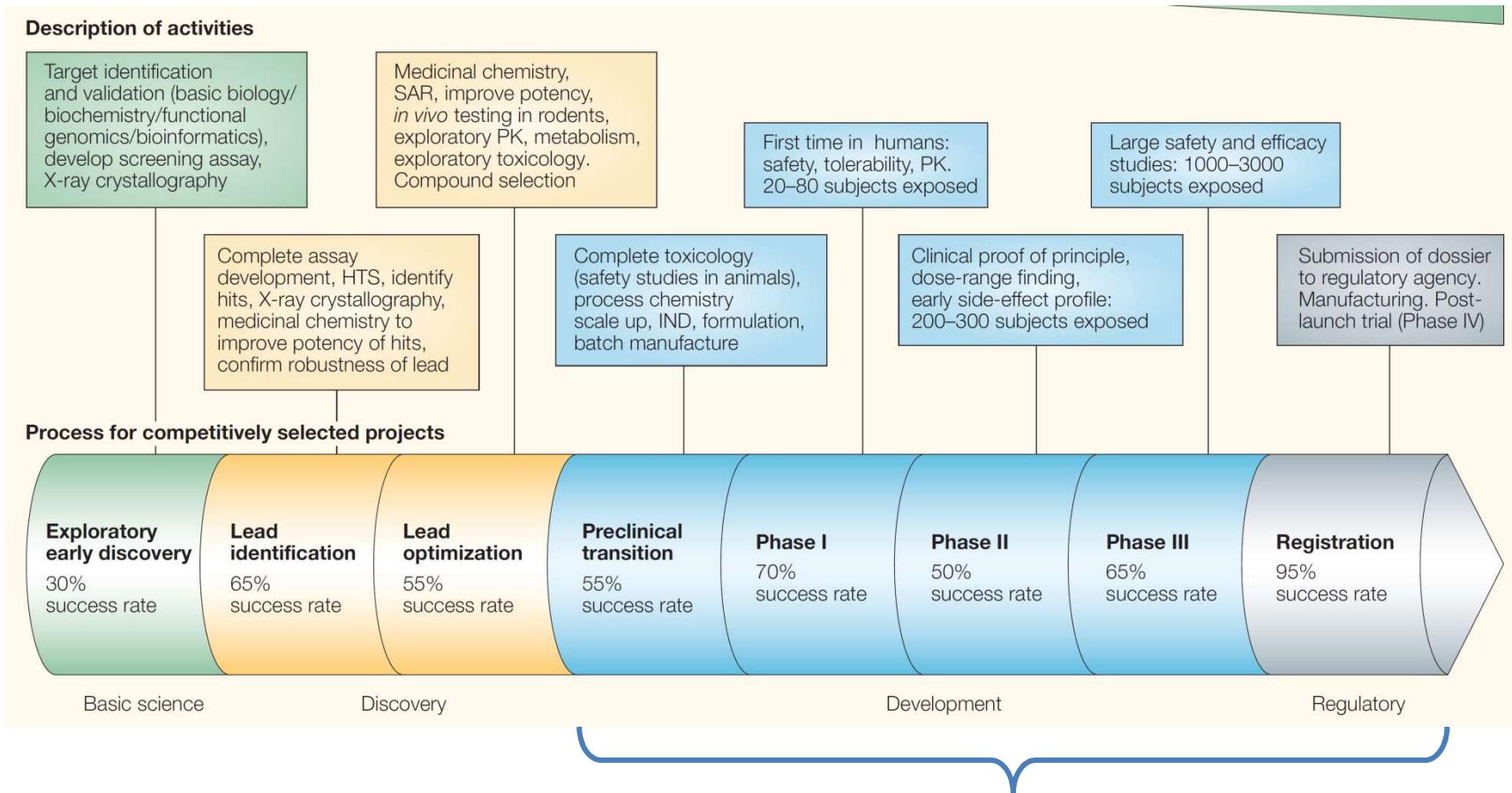
\$5.5 billion

Peak annual sales

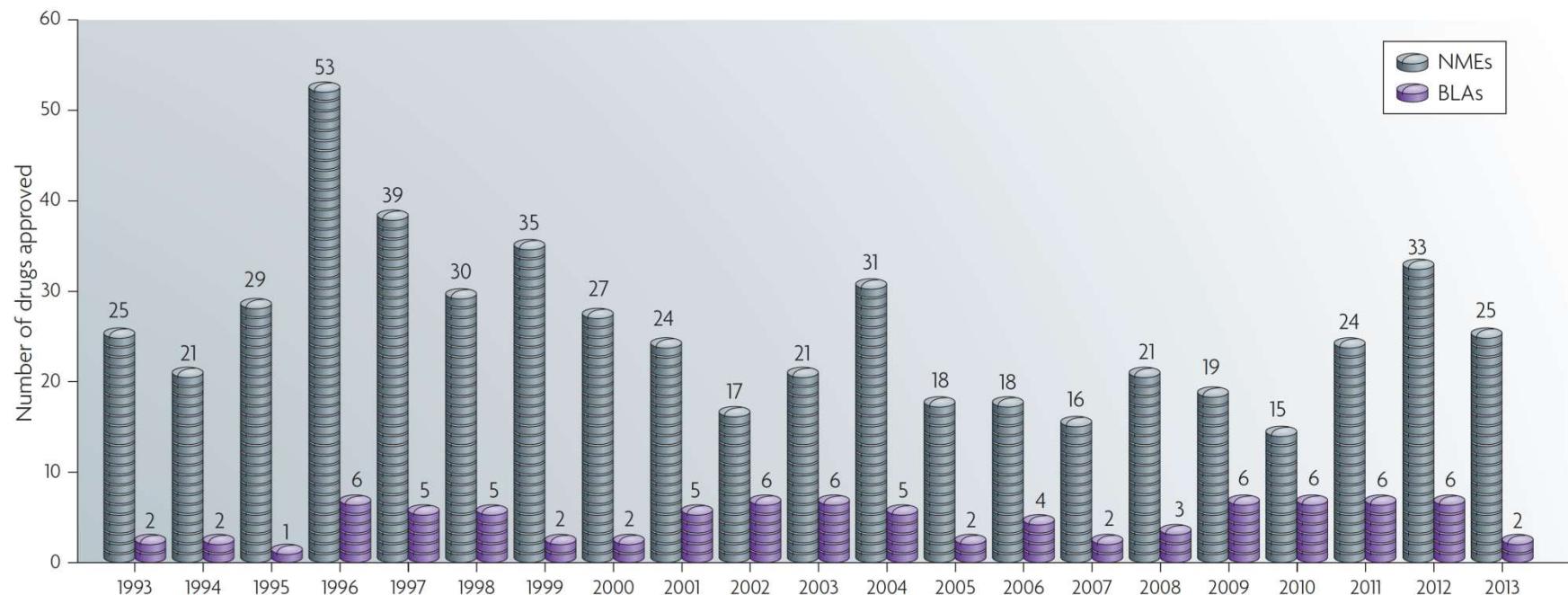
The Drug Discovery Pipeline



Only ~10% of the drugs that start phase I trials are eventually approved for marketing.
The total cost of developing a single drug can surpass \$1 billion.



FDA drug approvals 1993-2013



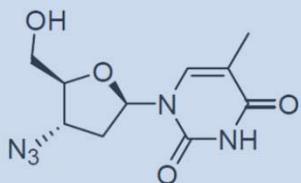
NME- new molecular entities

BLA – Biologics license applications

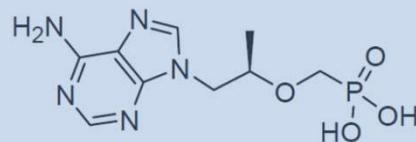
NME – a drug that contains an active moiety that has never been approved by the FDA or marketed in the US.

Anti-HIV drugs

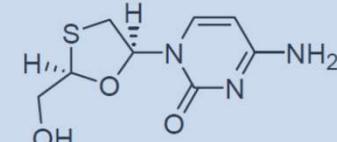
Reverse transcriptase inhibitors



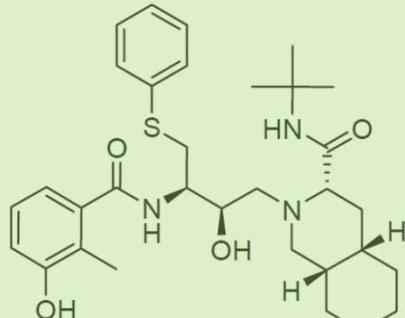
AZT®
(Retrovir)



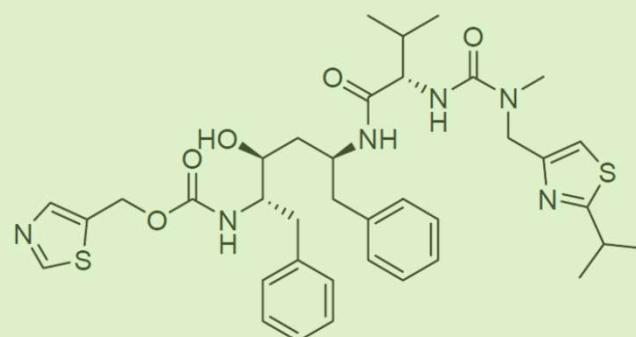
Viread®
(Tenofovir)



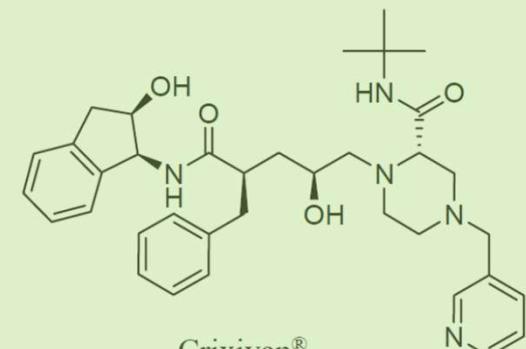
Zeffix®
(Lamivudine)



Viracept®
(Nelfinavir)



Norvir®
(Ritonavir)

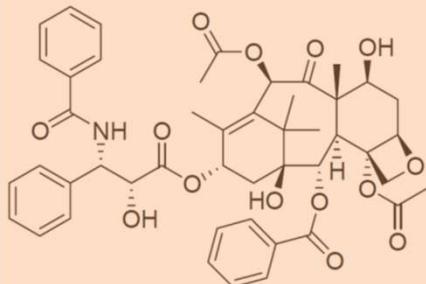


Crixivan®
(Indinavir)

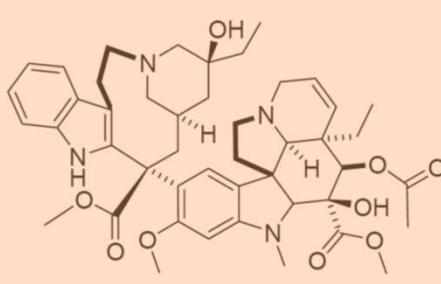
Protease inhibitors

Anti-cancer drugs

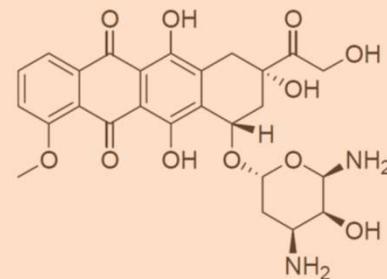
Natural Products (classical DD)



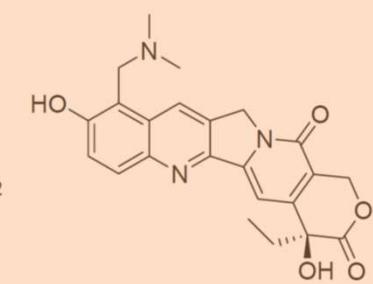
Taxol®
(Paclitaxel)



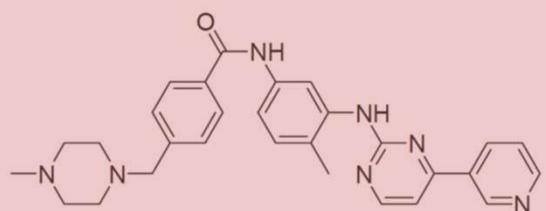
Velban®
(Vinblastine)



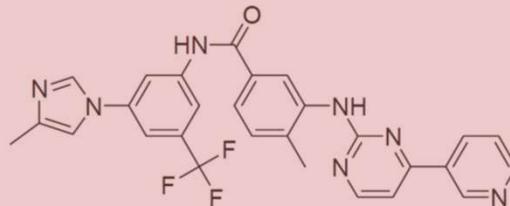
Adriamycin®
(Doxorubicin)



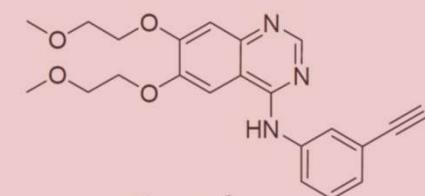
Hycamtin®
(Topotecan)



Gleevac® (Imatinib)



Tasigna®
(Nilotinib)

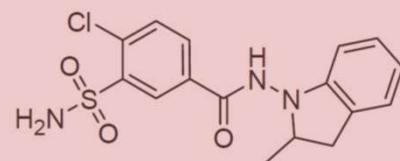
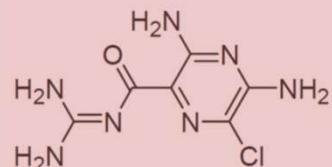


Tarceva®
(Erlotinib)

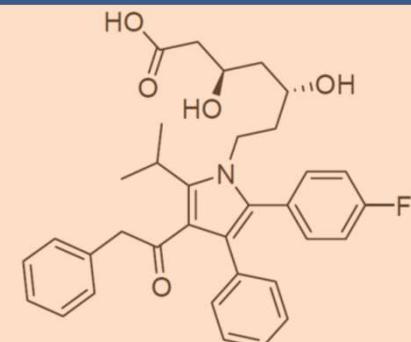
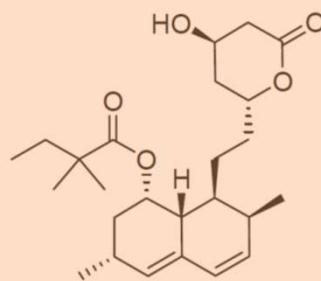
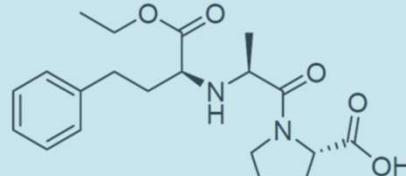
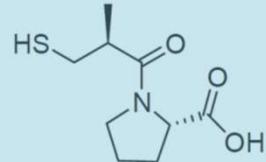
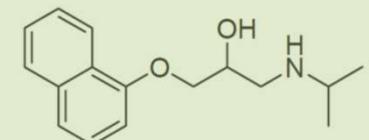
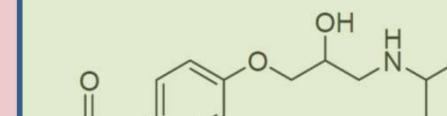
Synthetic Drugs (modern DD)

Cardiovascular drugs

Diuretics



β-blockers



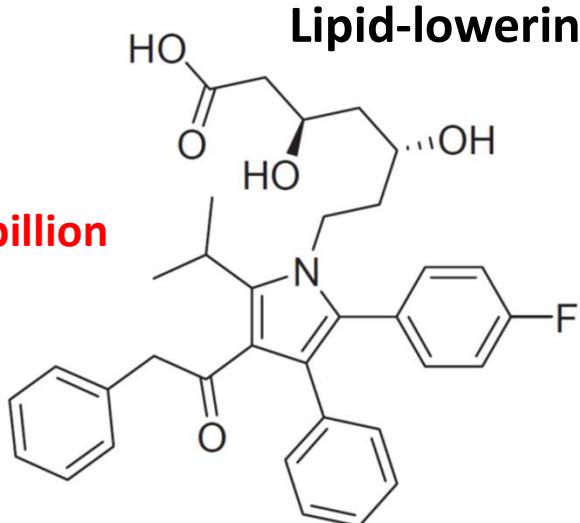
ACE inhibitors

HMG-CoA reductase inhibitors

Commercially successful drugs

Lipid-lowering

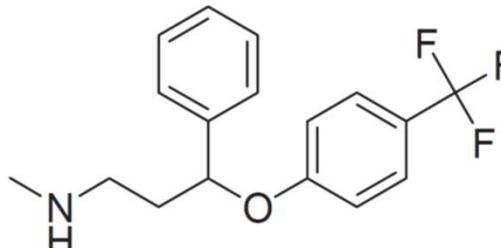
\$13 billion



Lipitor®
(Atorvastatin)

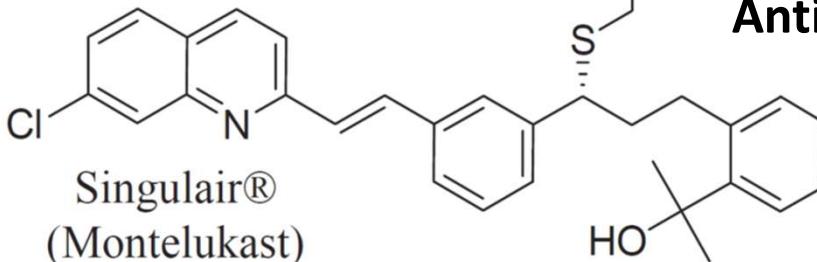
Anti-depressant

\$2.8 billion



Prozac®
(Fluoxetine)

Anti-asthma



Singulair®
(Montelukast)

\$5.5 billion

Peak annual sales

Tight regulation and approval

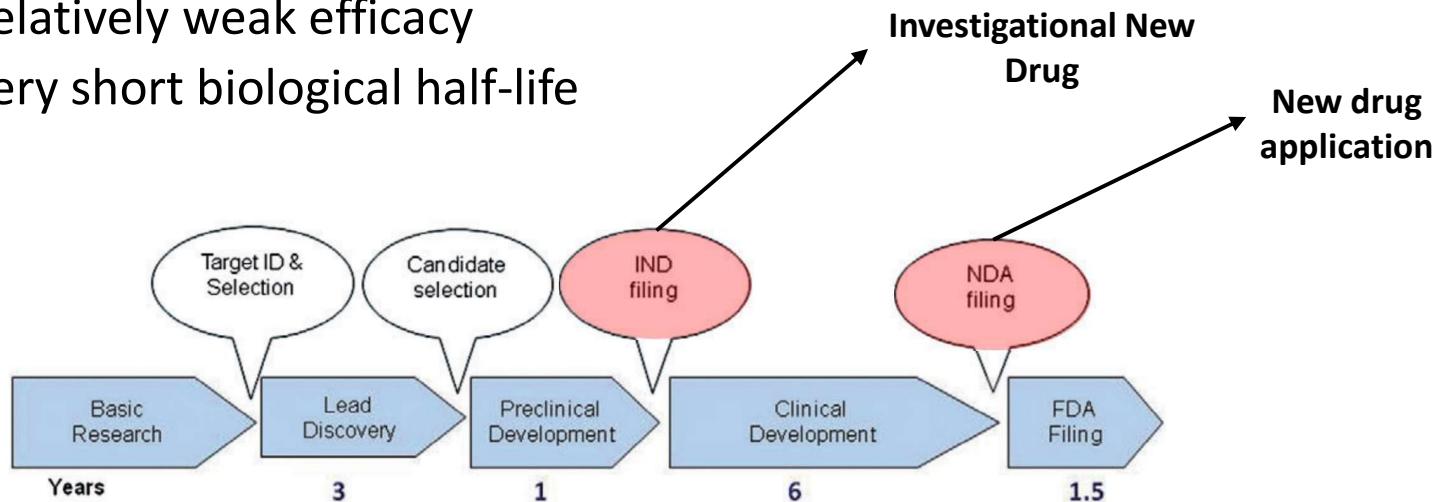
- Elixir of Sulfanilamide disaster (1937)
- Thalidomid disaster (1956-62)
- The Vioxx (rofecoxib) case (2004)
- ... and others...

Pure Food and Drug Act (1906)
Food, Drug and Cosmetic Act (1938)
Durham–Humphrey Amendment (1951)
Kefauver–Harris Amendment (1962)
Hatch–Waxman Act (1984)

Due to various mishappenings, drug manufacturing is probably the most regulated human activity!

By today's standards, Aspirin[©] wouldn't make it into the market:

- Causes gastric bleeding
- It is an irreversible inhibitor
- Relatively weak efficacy
- Very short biological half-life

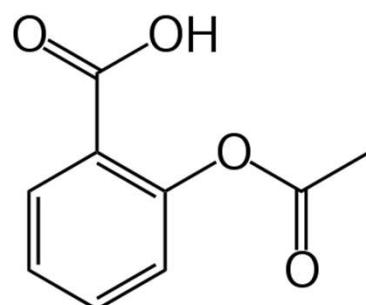


The Evolution of Drug Research

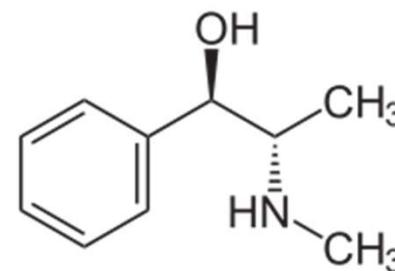
1. Empirical methods were the only source of medicines
2. Targeted isolation of active compounds from plants
3. Beginning of a systematic search for new synthetic materials with biological effects and the introduction of **animal models**
4. Use of molecular and other *in vitro* test systems as precise models replacing animal experiments (**screening**).
5. Introduction of theoretical and experimental methods: X-ray crystallography, QSAR, molecular modelling for the targeted **structure-based and computer-assisted** design of drugs
6. Discovery and therapeutic validation of targets through genomic, proteomic and transcriptomic analysis, knock-in and knock-out animal models and siRNA gene silencing

Drugs known by the end of the XIX century

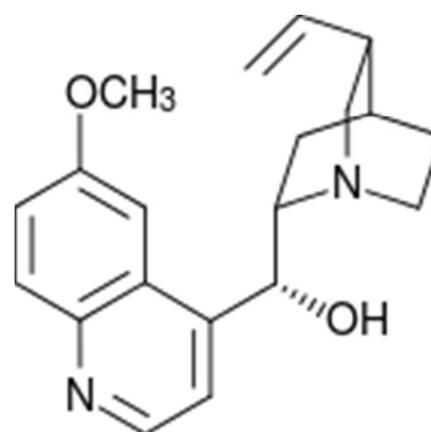
- Digitalin (heart stimulant)
- Quinine (anti-malarial)
- Ipecac (emetic)
- Aspirin (anti-inflammatory)
- Ephedrine (antiasthmatic and stimulant)
- Mercury (syphilis)



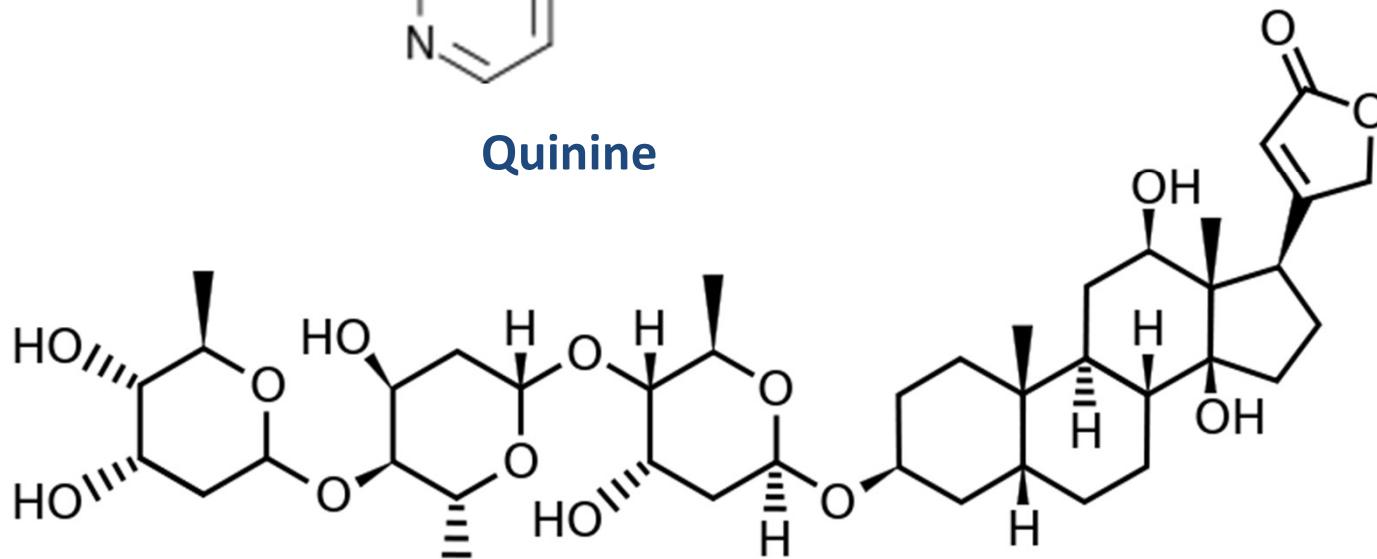
Aspirin



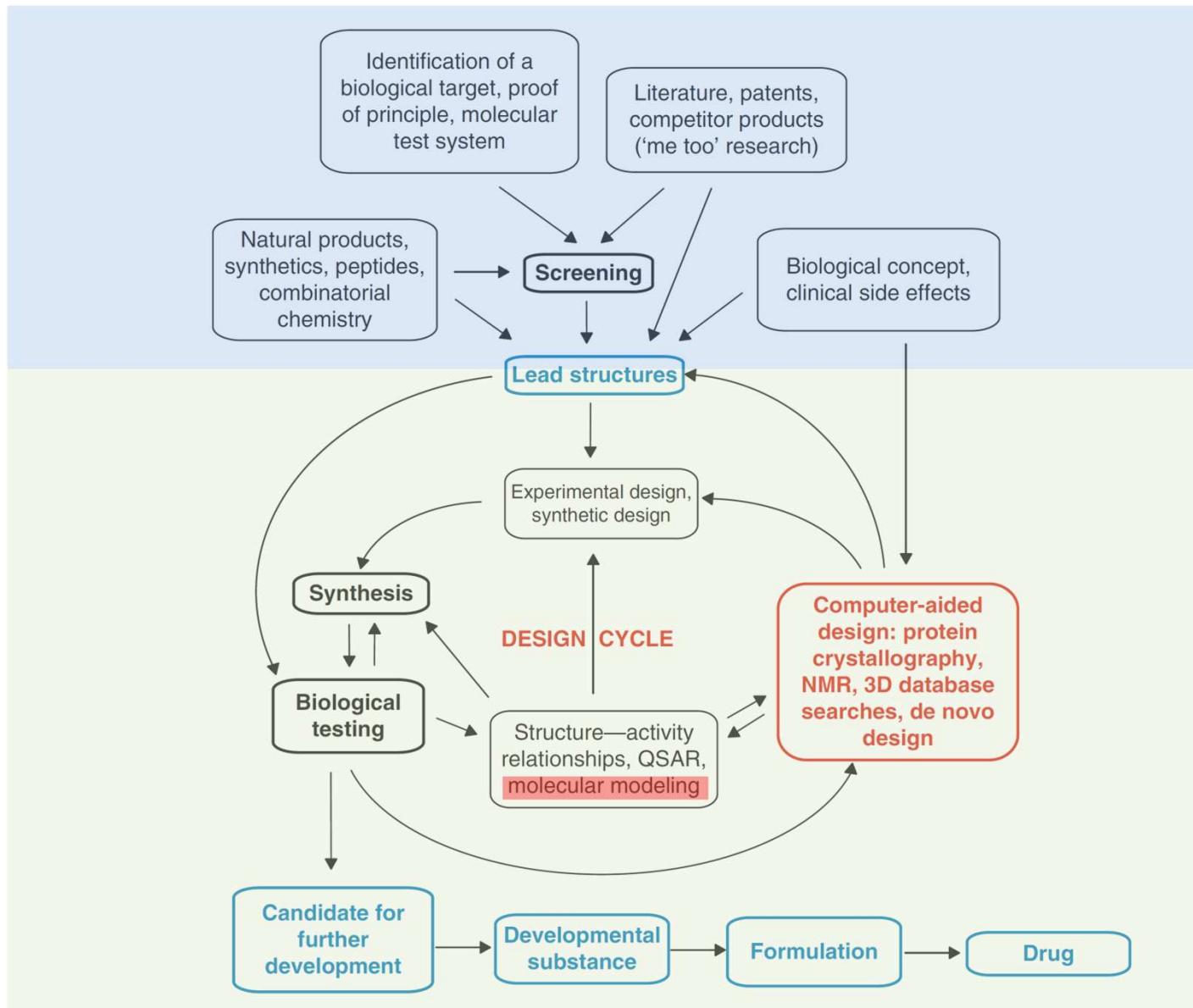
Ephedrine



Quinine



The drug design cycle

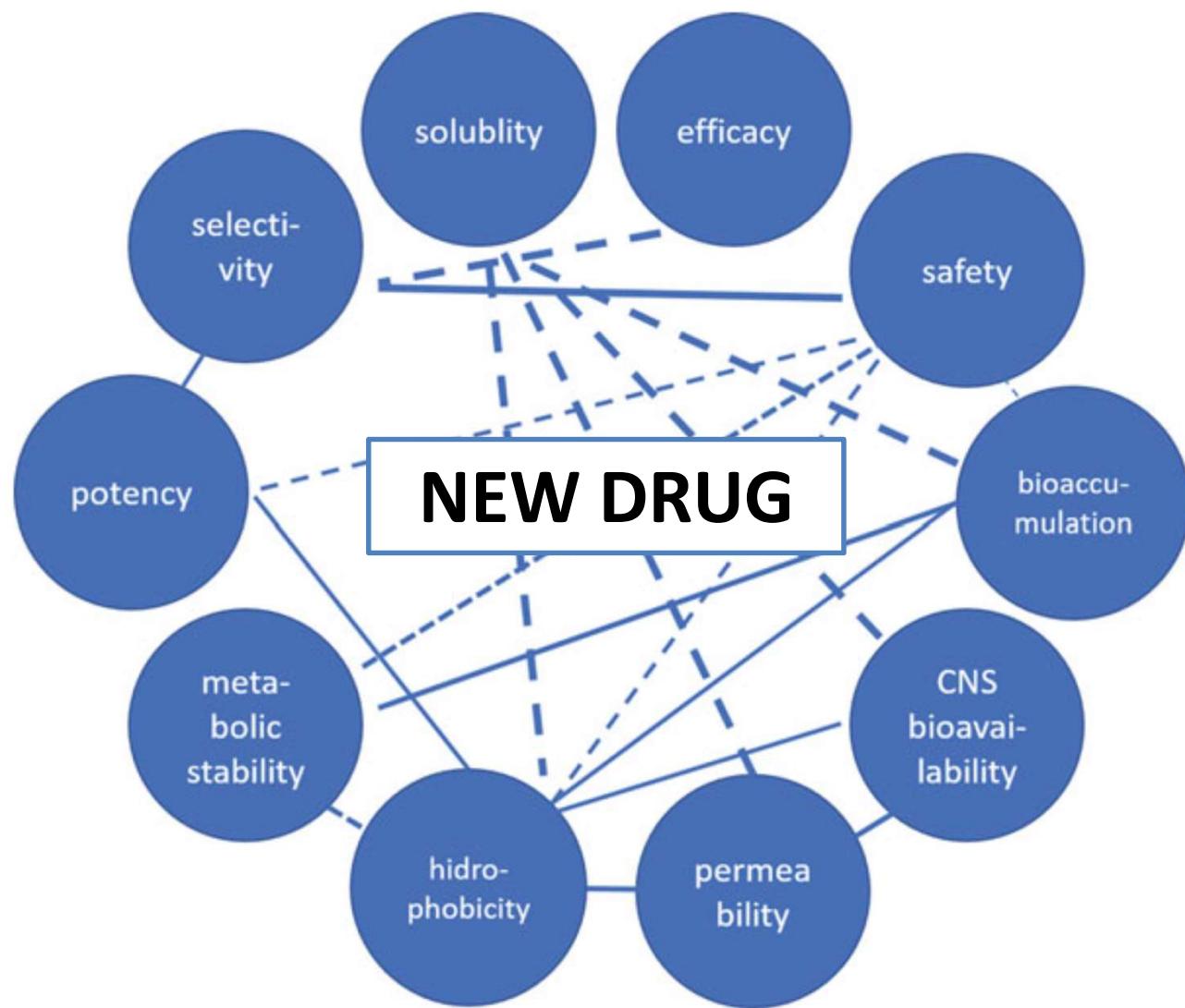


Disciplines for DD

- Biochemistry
- Molecular Biology
- Medicinal Chemistry
- Pharmacology
- Genetics
- Physiology
- Biophysics
- Molecular Modelling
- Computational Biochemistry
- Bioinformatics
- Systems Biology

What makes a good drug?

- Potency
- Selectivity
- Few side effects
- Good bioavailability
- Ease of synthesis
- No drug-drug interactions
- High therapeutic index



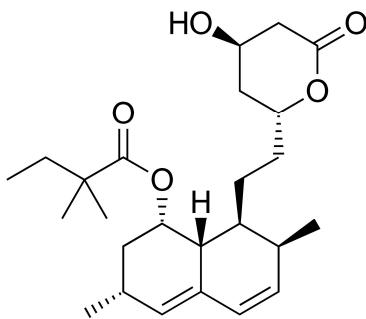
Requirements for drug candidates

- Efficacy data:
 - Enzyme activity
 - Whole organism activity
 - Animal models
- Metabolism:
 - *In vitro* metabolism
 - *In vivo* pharmacokinetics
- Safety:
 - *In vitro* selectivity
 - *In vitro* mutagenicity
 - *In vitro* cardiac
 - Animal toxicology
- Chemistry:
 - Physical form
 - Manufacture related
 - Back-up strategy
 - Objectives

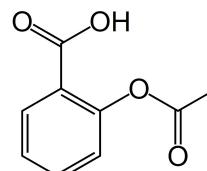
How are new drugs found ?

- Natural products (e.g. Aspirin)
- Screening assays
- Synthetic chemistry
- Combinatorial chemistry
- Similarity with known drugs (“Me too” drugs)
- Re-purposing (searching known drugs for a new effect)
- Serendipity:
 - drugs found by chance (e.g. Penicillin)
 - Unforeseen side-effect of a drug or candidate (e.g. Viagra)

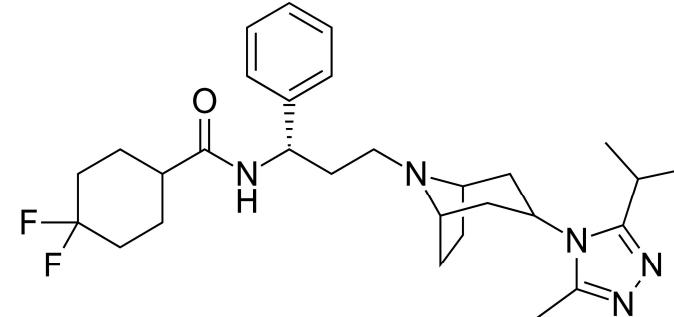
Drugs found by different methods



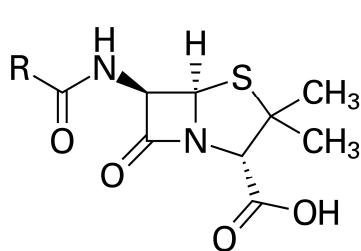
Simvastatin
("me too")



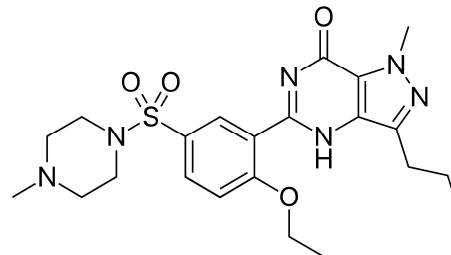
Aspirin
(natural product)



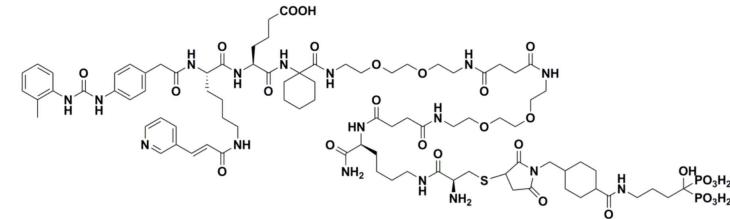
Maraviroc
(HTS assay)



Penicillin
(serendipity)

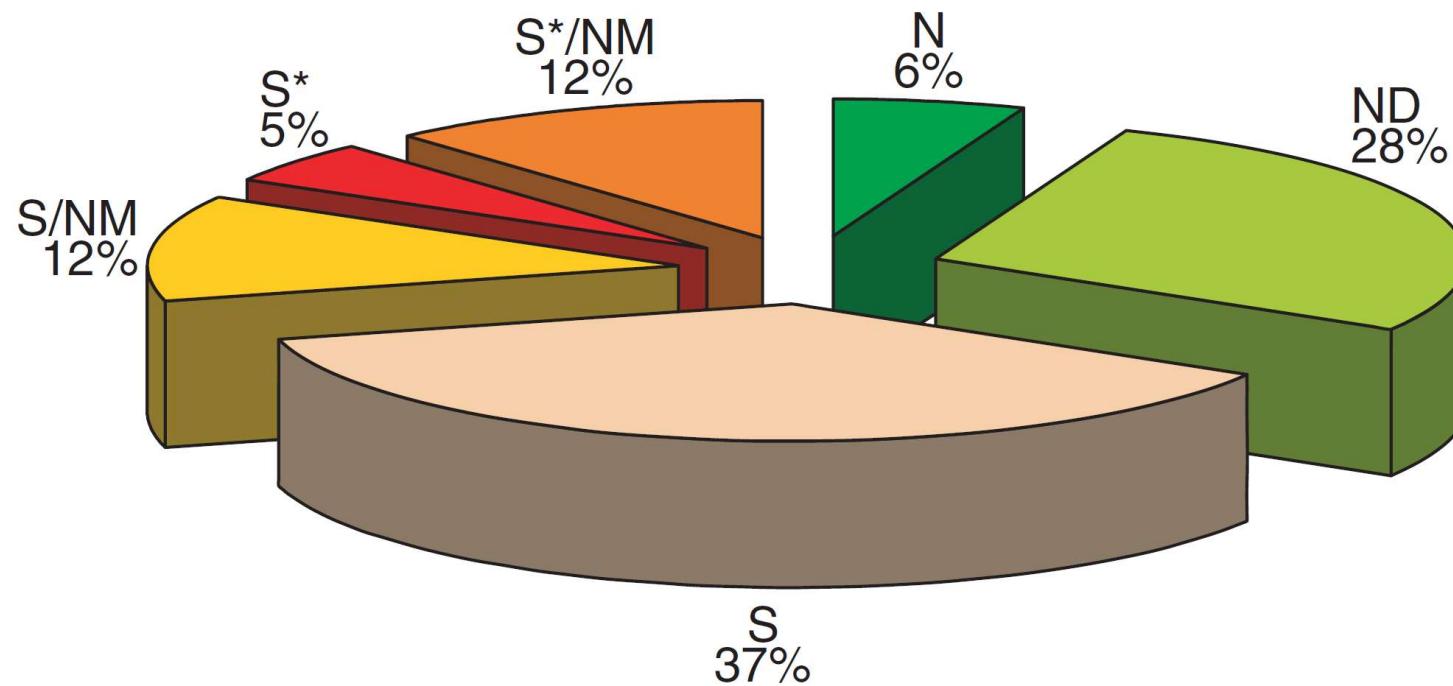


Sildenafil
(repurposing)



LLP2A-Ale
(combinatorial chemistry)

Where do drugs come from ?



N – unmodified natural product

Only 37% truly synthetic

ND – modified natural product

S – synthetic compound

S* - synthetic compound with natural product pharmacophore

S/NM – synthetic compound showing competitive inhibition of the natural product substrate

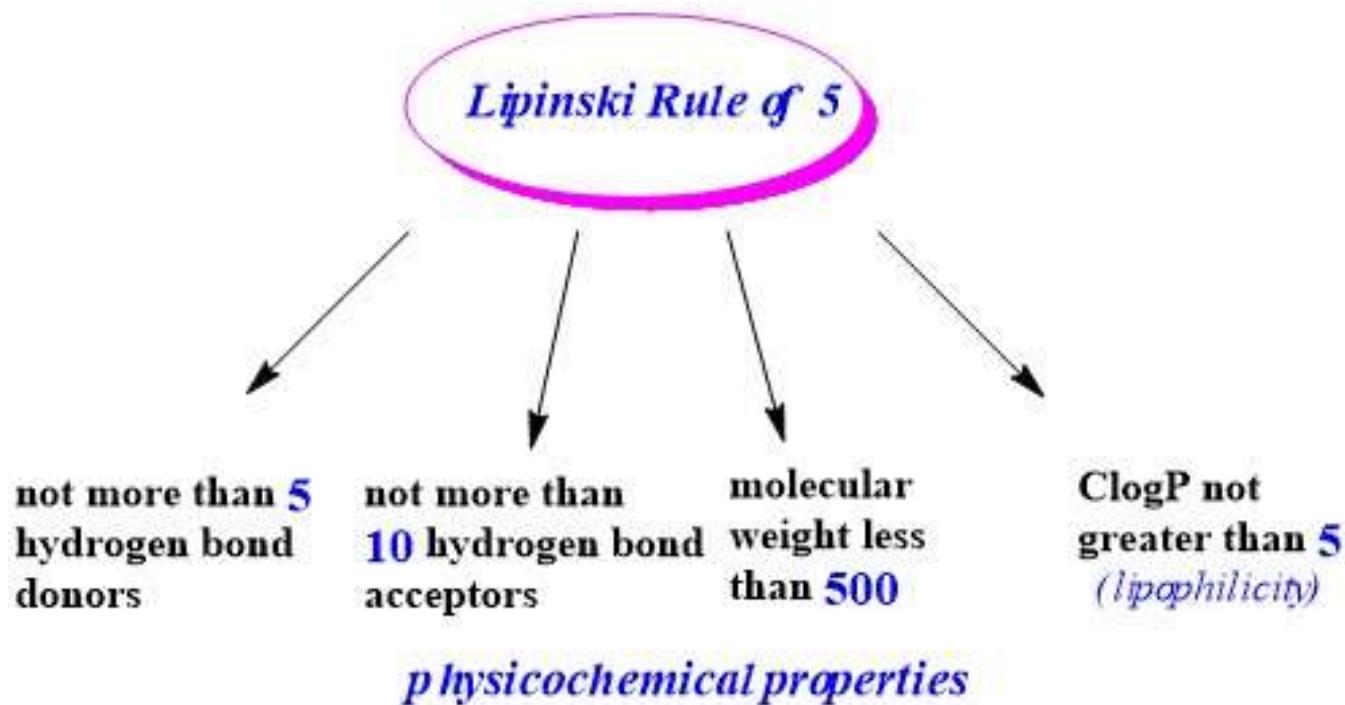
What makes a compound drug-like ?

Typical pharmaceutical compounds have:

- Molecular weight in the range $160 < MW < 480$
- Number of atoms between 20 and 70
- Lipophilicity in the range $-0.4 < \log P < +.56$
- Molar refractivity in the range $40 < MR < 130$
- Few H-bond donors (<5)
- Few H-bond acceptors (<10)
- At least one –OH group (except CNS-active drugs)

Lipinski's rule of 5

Christopher Lipinski formulated this rule of thumb to determine if a pharmacologically active substance is likely to work as an *oral drug*.



Lipinski, CA (2000) "Drug-like properties and the causes of poor solubility and permeability"
J Pharm Tox Meth 44:235-239

Rule of 5 in PubChem

0:500[mw] 0:5[hbdc] x PubChem PC3D View x www.ncbi.nlm.nih.gov/pccompound?term=0%3A500%5Bmw%5D+0%3A5%5Bhbdc%5D+0%3A10%5Bhbac%5D+-5%3A5%5Blogp%5D

NCBI Resources How To Sign in to NCBI

PubChem Compound 0:500[mw] 0:5[hbdc] 0:10[hbac] -5:5[logp] Search Help

Display Settings: Summary, 20 per page, Sorted by Default order Send to: Filters: Manage Filters

Results: 1 to 20 of 34559871 Actions on your results

Lipinski's rule of 5

<< First < Prev Page 1 of 1727994 Next > Last >>

1. Methyl 4-ethoxy-3-oxobutanoate; AK141825; 415678-65-8

MW: 160.167780 g/mol MF: C₇H₁₂O₄
IUPAC name: methyl 4-ethoxy-3-oxobutanoate
CID: 54303951
[Summary](#)

2. 6-bromo-3-iodopyridin-2-amine; AK142103; 1245643-34-8

MW: 298.907130 g/mol MF: C₅H₄BrI N₂
IUPAC name: 6-bromo-3-iodopyridin-2-amine
CID: 52987942
[Summary](#)

3. AK138368; 4-(2,2,2-Trifluoroethoxy)pyridin-2-amine; 1379361-82-6

MW: 192.138490 g/mol MF: C₇H₇F₃N₂O
IUPAC name: 4-(2,2,2-trifluoroethoxy)pyridin-2-amine
CID: 15724964
[Summary](#)

BioActivity Analysis Analyze the BioActivities of the compounds

Structure Clustering Cluster structures based on structural similarity

Structure Download Download the structures in various formats

Pathways Analyze pathways containing the compounds

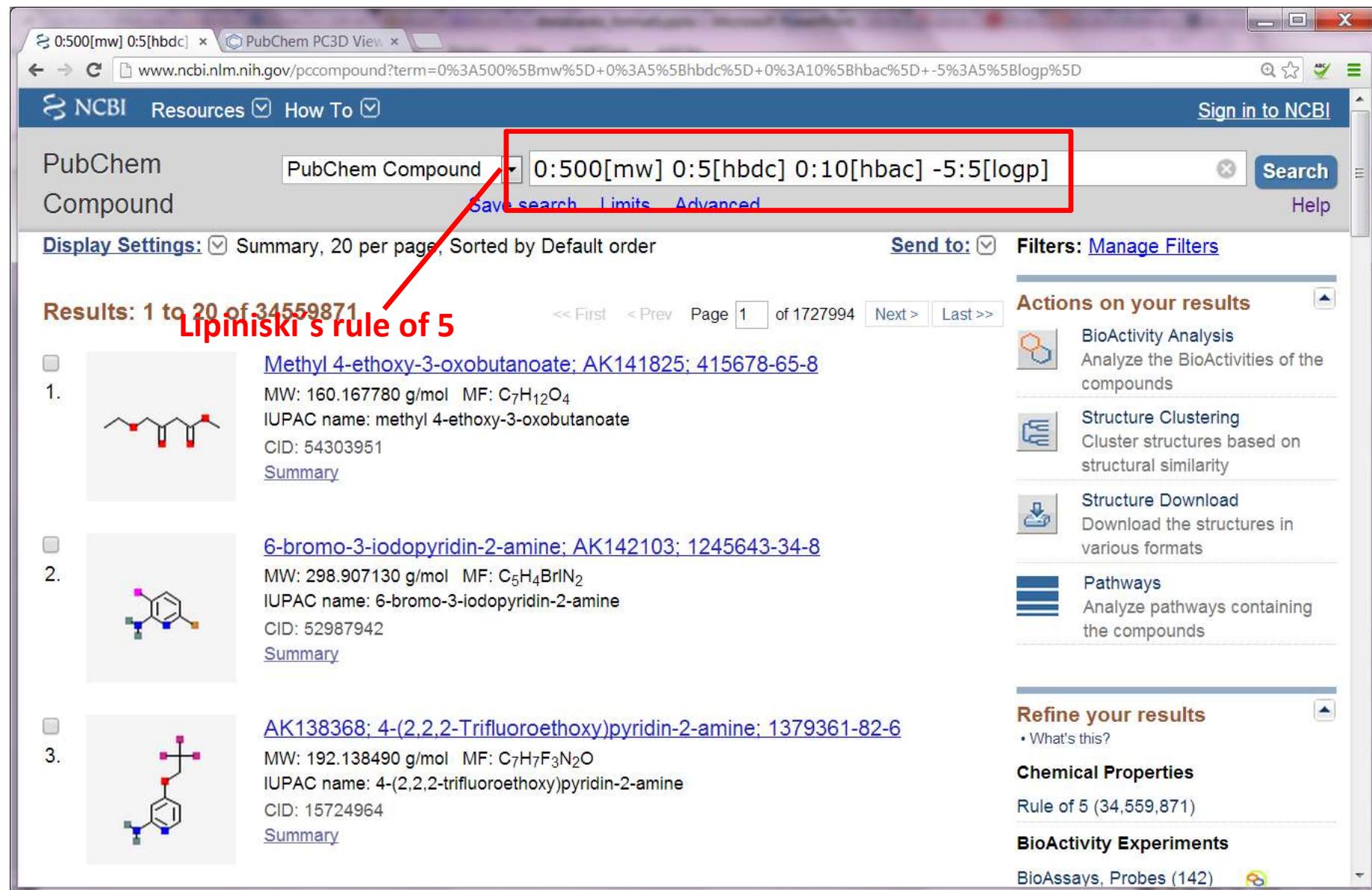
Refine your results

- What's this?

Chemical Properties Rule of 5 (34,559,871)

BioActivity Experiments

BioAssays, Probes (142)



logP

logP is the logarithm of the partition coefficient of a substance between octanol and aqueous phases. It is a measure of the **lipophilicity**.

A drug must be lipophilic enough to cross cell membranes, but not so much it can't dissolve in the plasma.

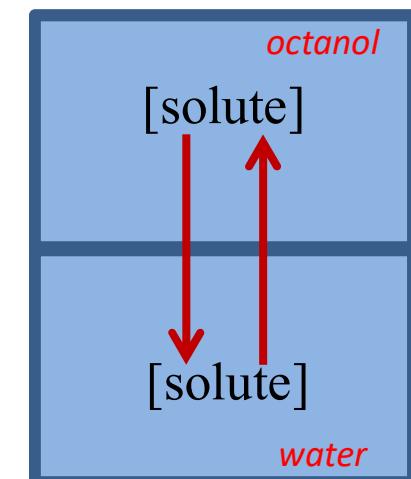
$$\log P = \log \left(\frac{[\text{solute}]_{\text{octanol}}}{[\text{solute}]_{\text{water}}} \right)$$

hydrophilic $-4.0 < \log P < +8.0$ lipophilic

Citric acid -1.72

Iodobenzene $+3.25$

Typical drugs < 5.0



ClogP, XlogP – theoretical estimates of logP based on structure

Computational prediction of drug likeness

molsoft
molecules *in silico*

About Products Downloads Technology Services News Events Support [t](#) [f](#)

Knowledge Base | [What's New](#) | Contact Us |

Drug-Likeness Molecular property prediction.

High speed Molecular properties calculator can be licensed from Molsoft for the local use in the batch mode.
For more information mail us at info@molsoft.com

Draw the structure and click the button below.

MolEdit © 2017 MolSoft

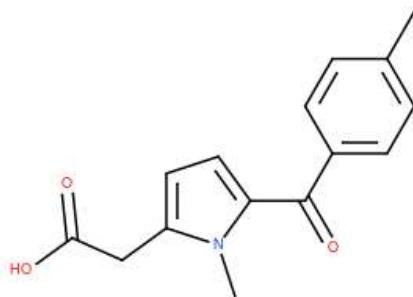
Calculate Properties

[Read more about molecular property prediction](#)

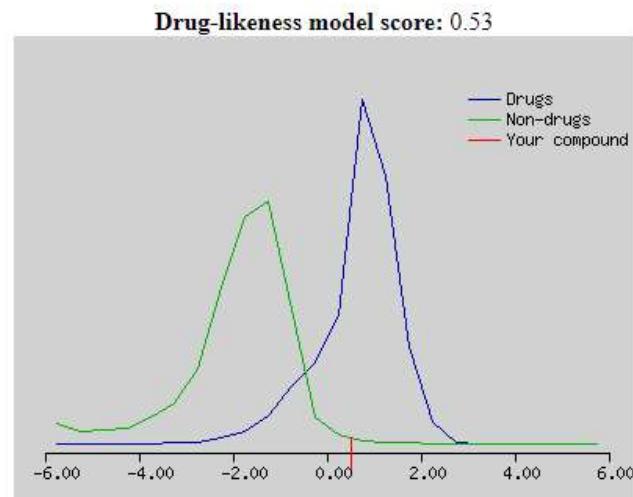
<http://molsoft.com/mprop>

Computational prediction of drug likeness

Molecular Properties and Drug-likeness.



Molecular formula: C₁₅ H₁₅ N O₃
Molecular weight: 257.11
Number of HBA: 3
Number of HBD: 1
MolLogP : 1.93
MolLogS : -3.09 (in Log(moles/L)) 210.05 (in mg/L)
MolPSA : 43.86 Å²
MolVol : 262.08 Å³
Number of stereo centers: 0

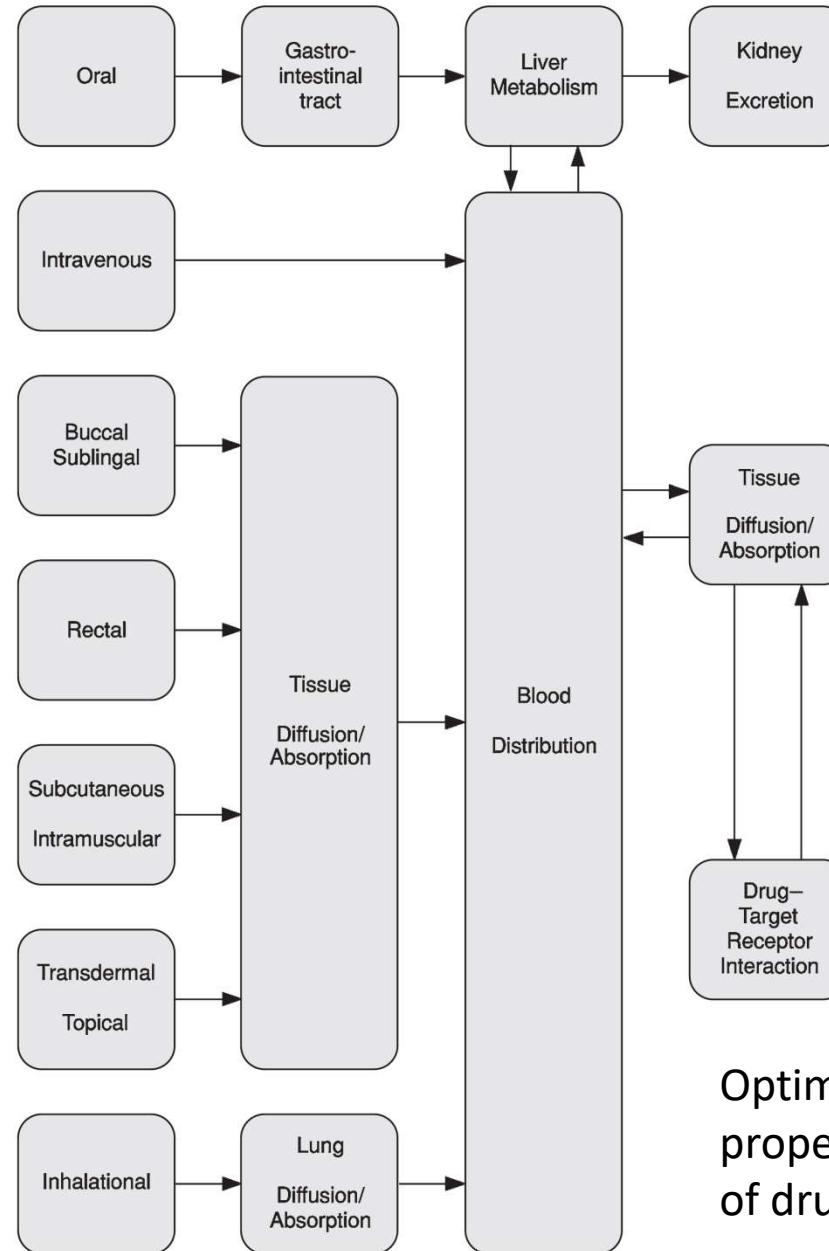


[New molecule](#) [Modify molecule](#) [Search molecule](#)

<http://molsoft.com/mprop>

ADMET

Absorption
Distribution
Metabolism
Excretion
Toxicity



Optimization of ADMET properties is a crucial aspect of drug design!

Structure modification strategies for solubility improvement

Structure modification

Add ionizable group

Reduce Log P

Add hydrogen bonding

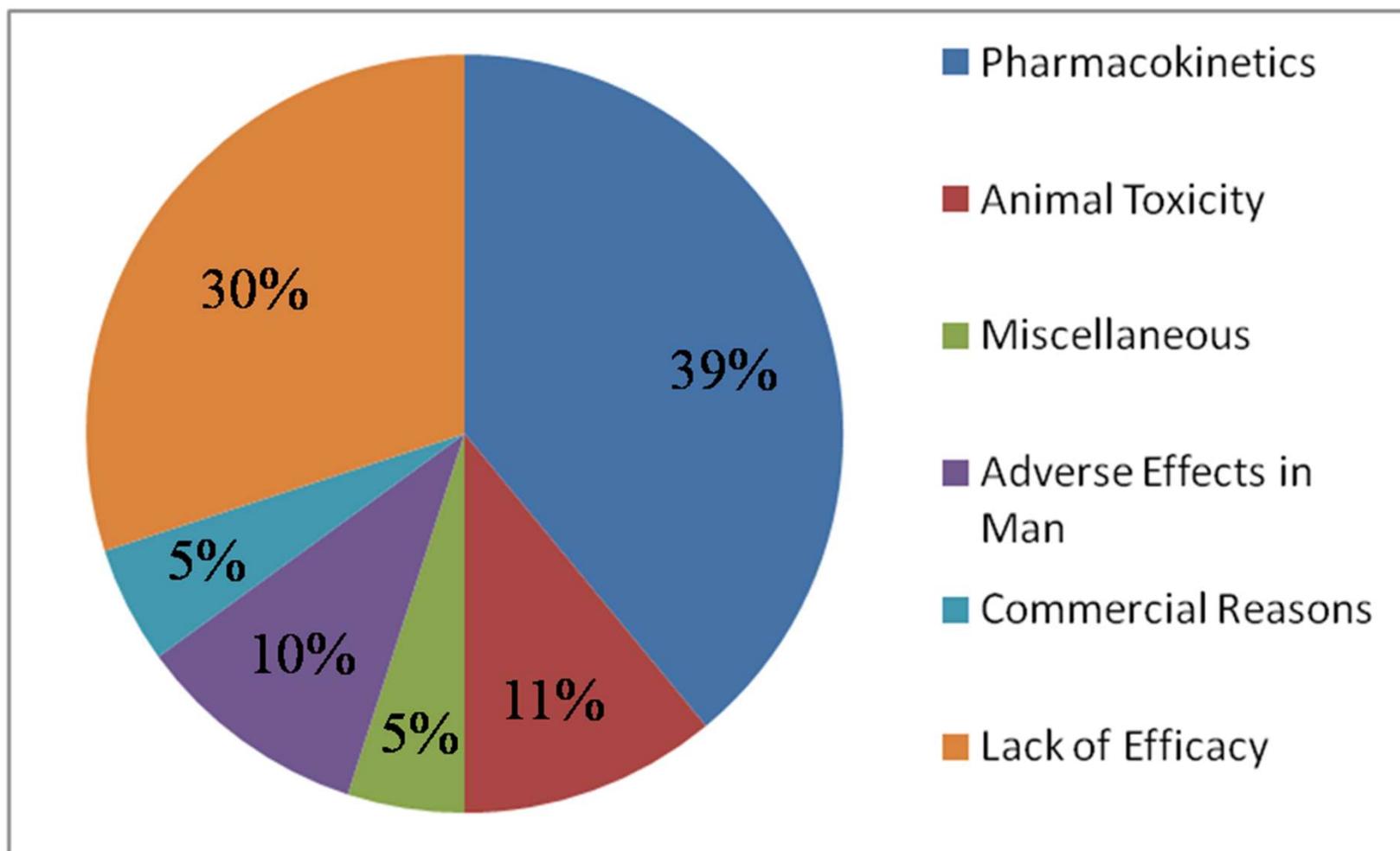
Add polar group

Reduce molecular weight

Out-of-plane substitution to reduce crystal packing

Construct a prodrug

Why drugs fail – the importance of pharmakocinetics



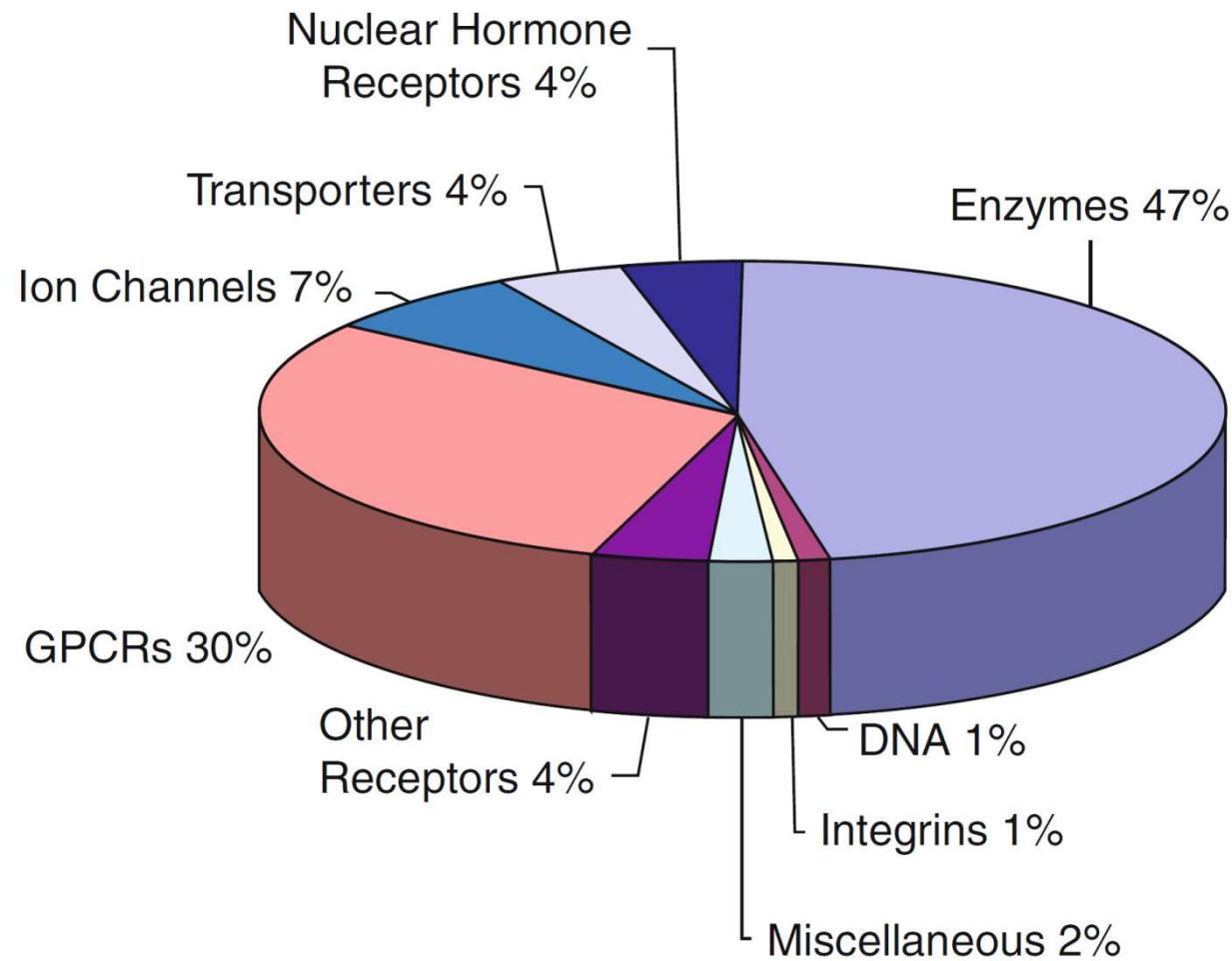
Pharmokinetics and Pharmacodynamics

- **Pharmacodynamics** or “what the drug does to the body” :
 - Mechanisms of drug action, interaction with target receptor or enzyme, mode of inhibition, allosteric. Concepts such as affinity, selectivity, agonist, antagonist,...
- **Pharmacokinetics** or “what the body does to the drug”:
 - Processes of drug absorption, transport and metabolism. Concepts such as half-life, solubility, permeability, therapeutic index...

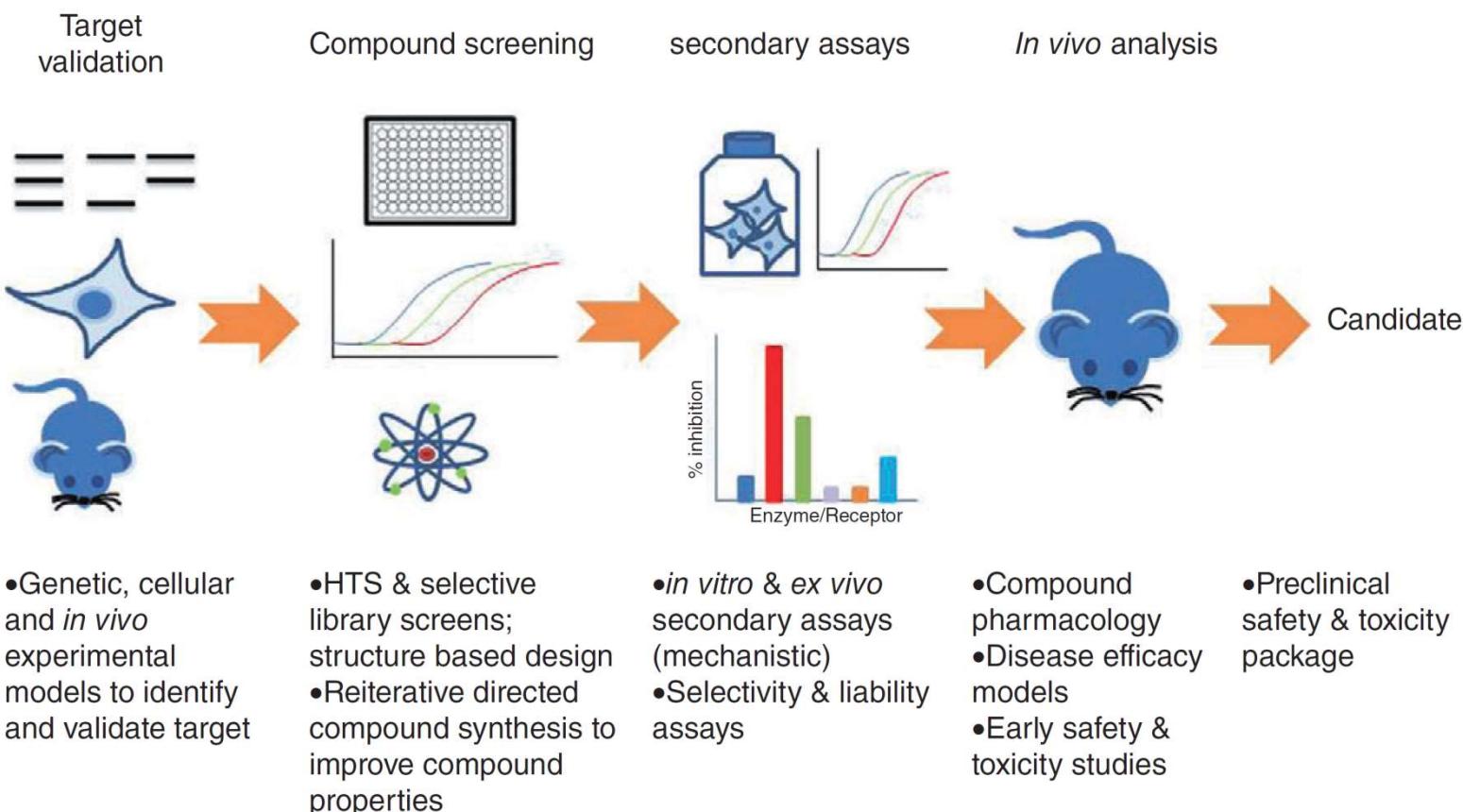
Drug targets

- **Enzymes:** There are many different types of enzymes in the human body. They are required for a variety of functions. Drugs can interact with enzymes to modulate their enzymatic activities.
- **Intracellular Receptors:** These receptors are in the cytoplasm or nucleus. Drugs or endogenous ligand molecules have to pass through the cell membrane (a lipid bilayer) to interact with these receptors. The molecules must be hydrophobic or coupled to a hydrophobic carrier to cross the cell membrane.
- **Cell Surface Receptors:** These receptors are on the cell surface and have an affinity for hydrophilic binding molecules. Signals are transduced from external stimuli to the cytoplasm and affect cellular pathways via the surface receptors. There are three main super families (groups) of cell surface receptors: G-protein coupled receptors, ion channel receptors, and catalytic receptors using enzymatic activities.
- **Nucleic Acids:** DNA and RNA support genetic information and its replication and translation. NA drugs can be groove binders, intercalators, chain terminators or alkylating agents.

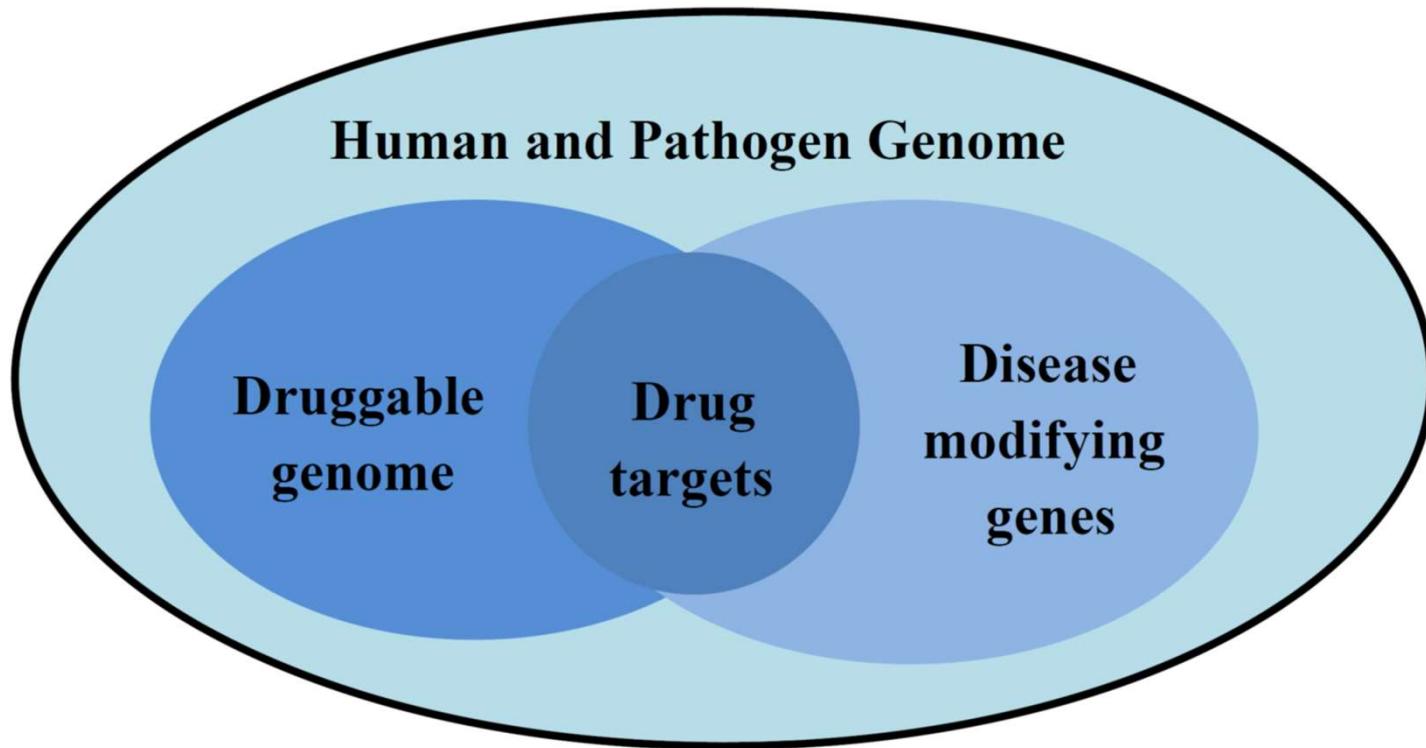
The drug targets



The early DD phase



The “druggable” genome



~5000 druggable genes ?

Evaluating druggability

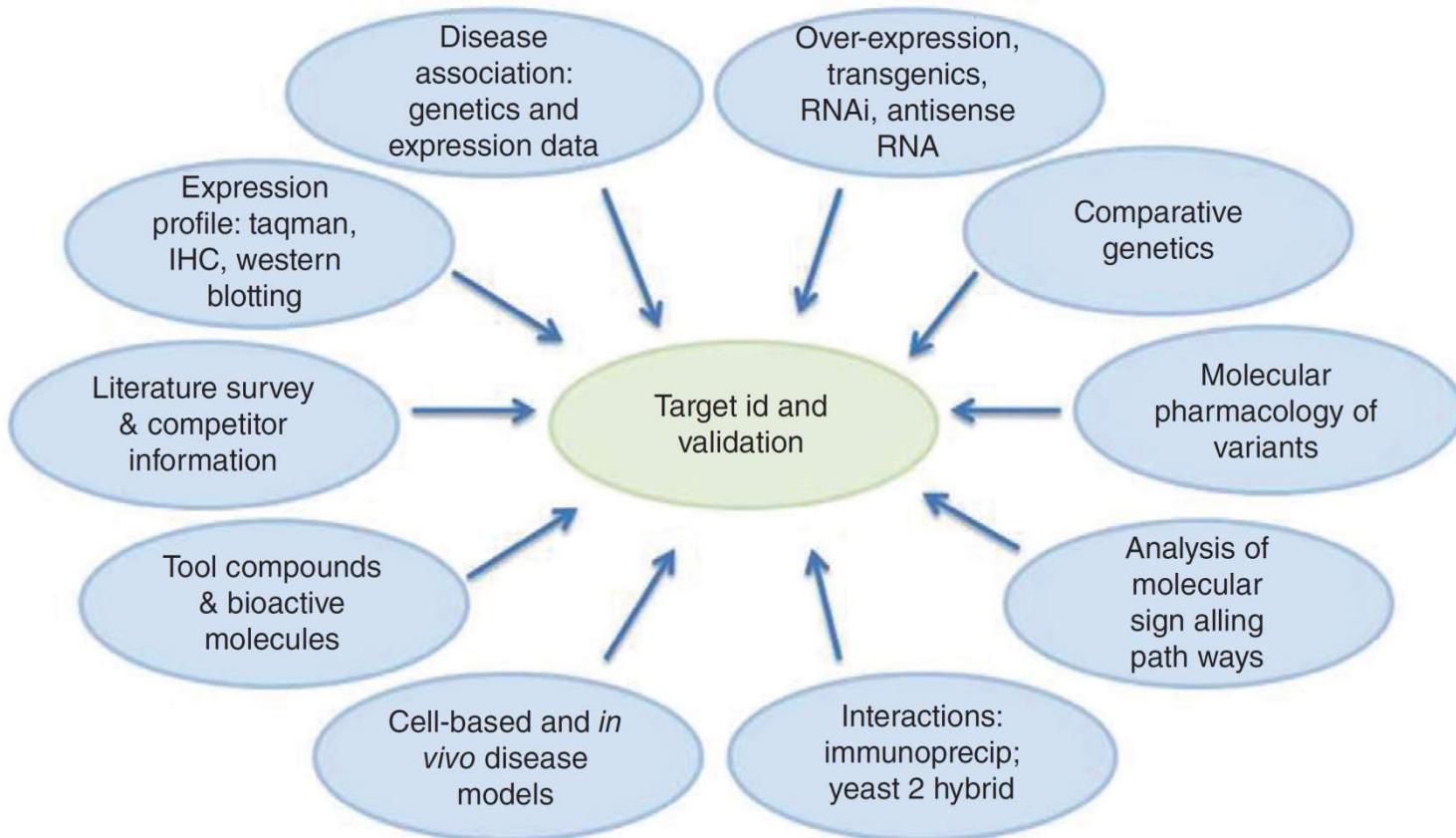
Druggability – the ability of a macromolecular target to bind small drug-like molecules

Ways to evaluate druggability:

- **Precedence** – target is a member of a family known to bind small molecules
- **Structural analysis** – 3D structural analysis of the target aiming at the identification of structural features relevant to binding:
 - Identification of cavities or pockets in the structure
 - Calculation of physico-chemical properties of pockets
 - Assessing fitness of structural properties to a training set of druggable targets (machine-learning methods)
- **Feature based** – using other properties of the target, like those that can be derived from the aminoacid sequence,

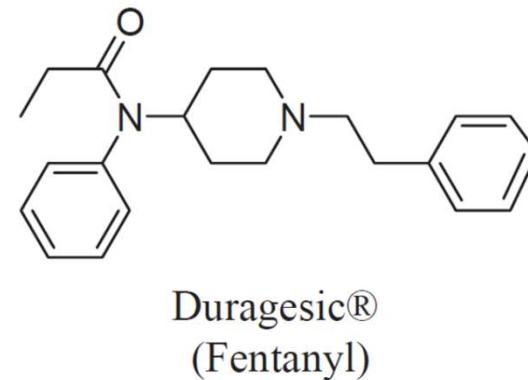
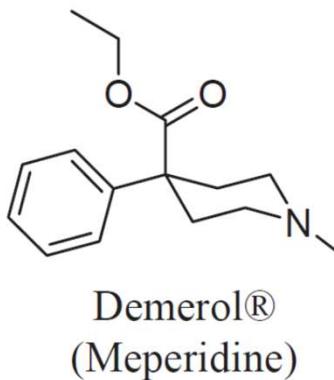
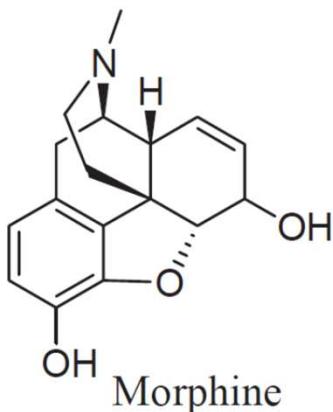
“Out of the nearly 20,000 protein-coding genes in the human genome, approximately 3,000 are estimated to be part of the druggable genome, the subset of genes expressing proteins with the ability to bind drug-like molecules. Yet, less than ten percent of the druggable proteins are currently targeted by FDA-approved drugs” - NIH, Illuminating the Druggable Genome

Target selection



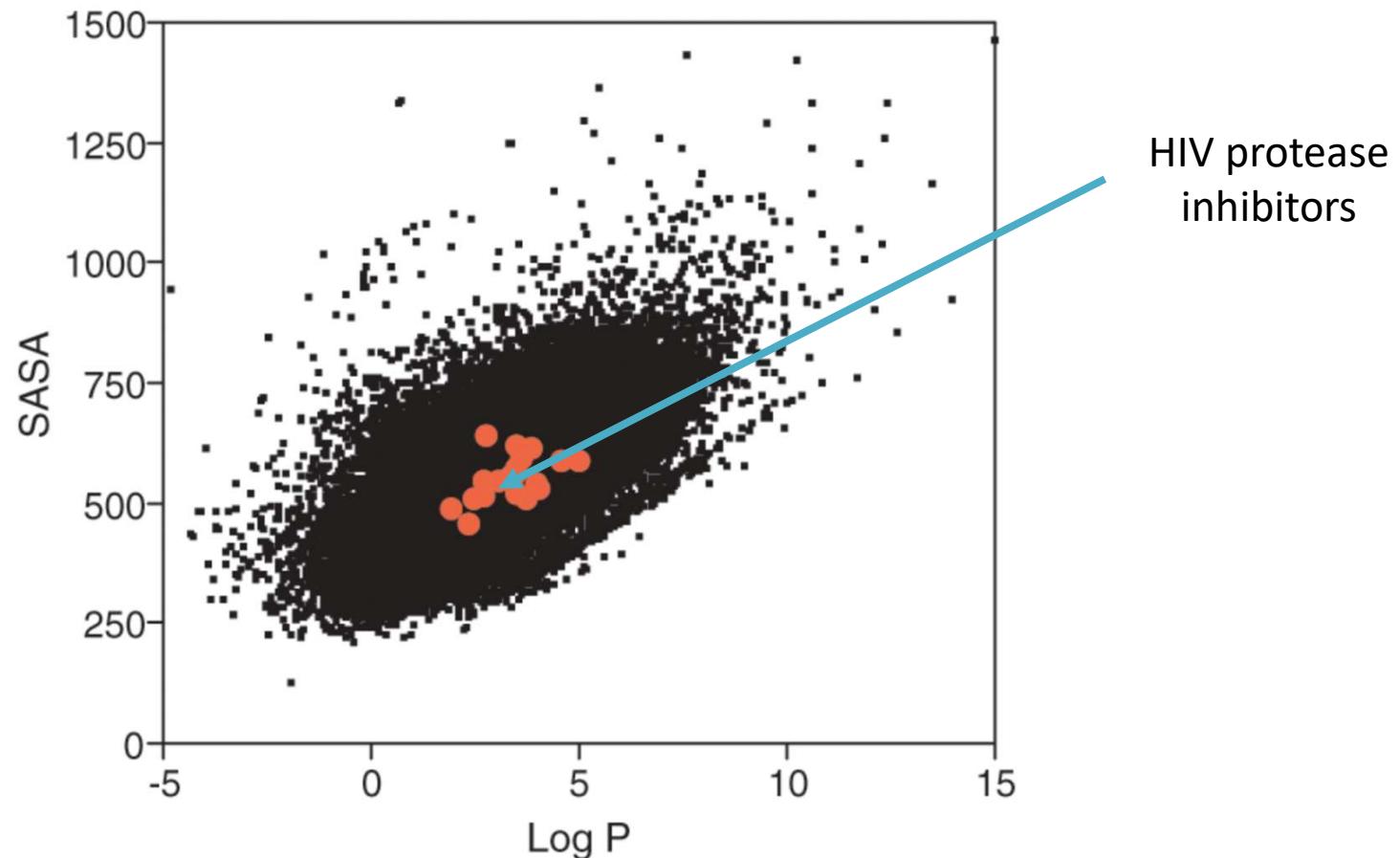
Choosing the test compounds

- There are over 70 million compounds registered in the Chemical Abstracts Service (CAS)
- Filtering this compounds for drug-likeness will still leave a big number
- Structural analogy is often not required for binding the same target
- First-in-class drugs are more profitable, but much harder to discover
- High Throughput Screening (HTS), real or virtual (*in silico*) may be used to deal with a large subset of the chemical space



Structurally diverse μ -opioid receptor agonists

In Silico filtering by descriptor



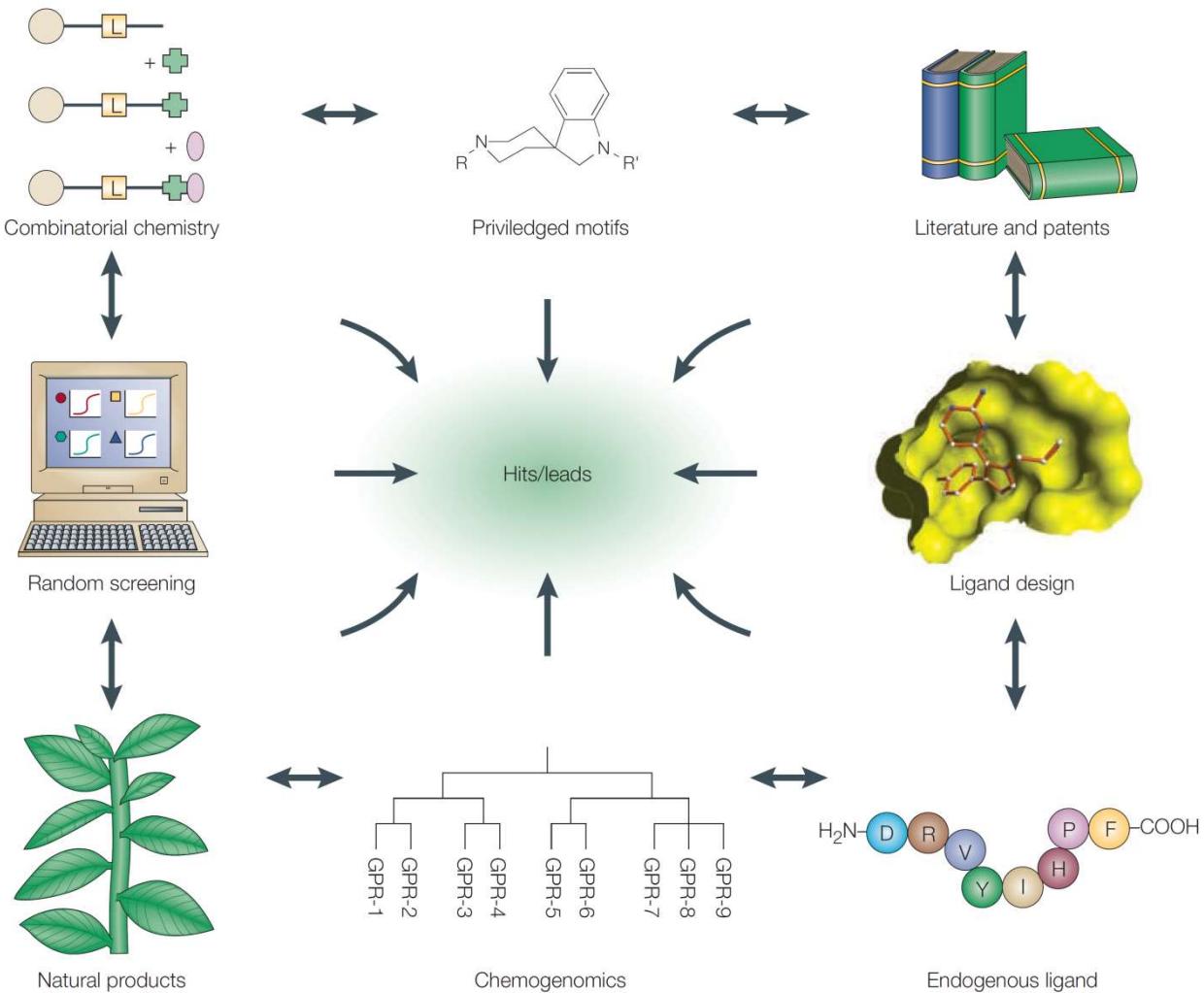
The 70k compounds in the Maybridge catalog are represented as dots of logP versus solvent accessible area. Known HIV-1 (red dots) protease inhibitors cluster on a narrow region of SASA and logP.

Jorgensen(2004) *Science* **303**:1813

Hits and leads

- Hit – compound which has the desired activity in a compound screen and whose activity is confirmed upon retesting
- Lead – a hit compound with sufficient potency, selectivity, drug-likeness, bioavailability, and *in vivo* effect to be selected as drug candidate

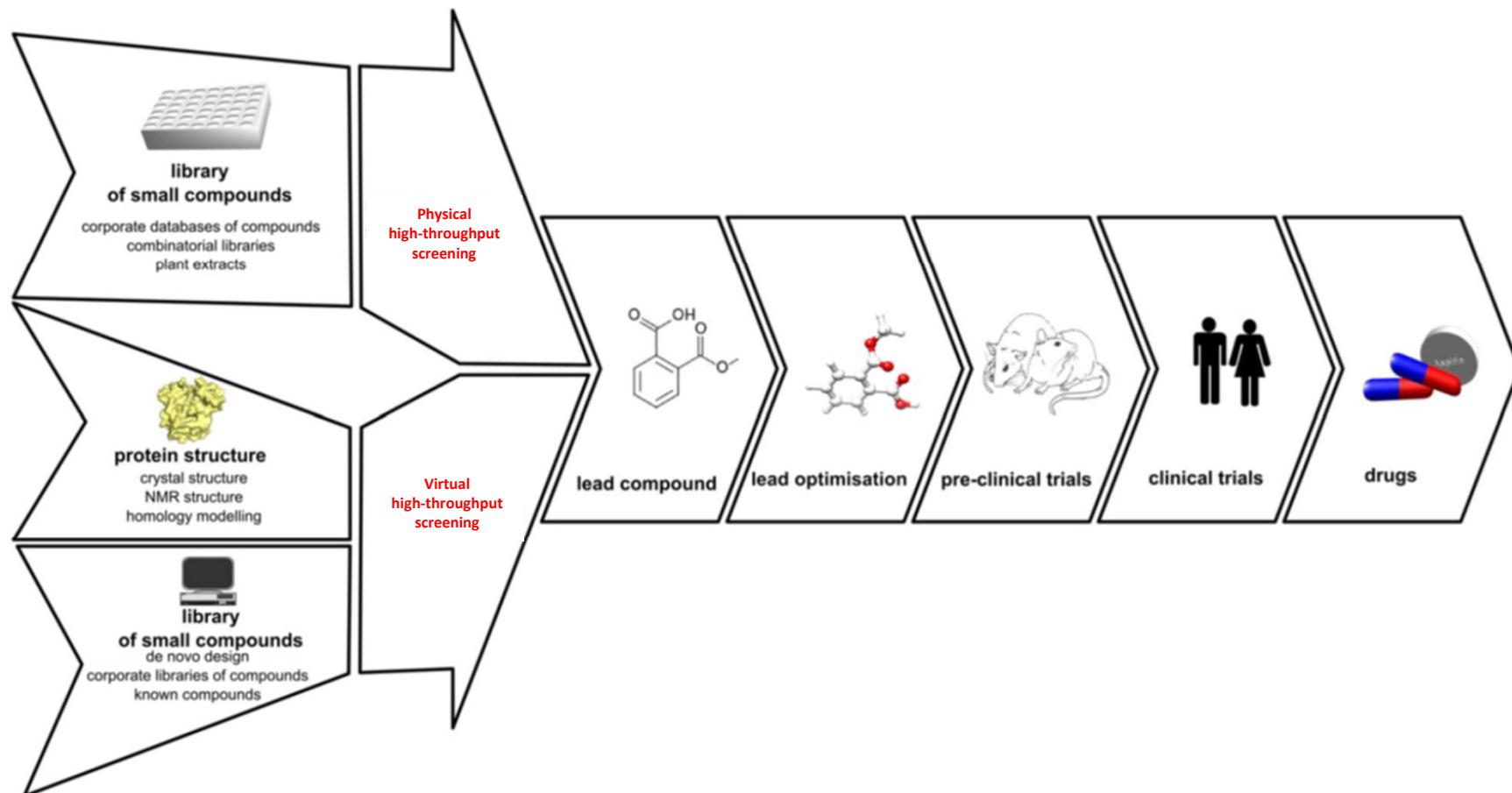
Generation of Hits/Leads



Screening strategies

Screen	Description	Comments
High throughput	Large numbers of compounds analysed in a assay generally designed to run in plates of 384 wells and above	Large compound collections often run by big pharma but smaller compound banks can also be run in either pharma or academia which can help reduce costs. Companies also now trying to provide coverage across a wide chemical space using computer assisted analysis to reduce the numbers of compounds screened.
Focused screen	Compounds previously identified as hitting specific classes of targets (e.g. kinases) and compounds with similar structures	Can provide a cheaper avenue to finding a hit molecule but completely novel structures may not be discovered and there may be difficulties obtaining a patent position in a well-covered IP area
Fragment screen	Soak small compounds into crystals to obtain compounds with low mM activity which can then be used as building blocks for larger molecules	Can join selected fragments together to fit into the chemical space to increase potency. Requires a crystal structure to be available
Structural aided drug design	Use of crystal structures to help design molecules	Often used as an adjunct to other screening strategies within big pharma. In this case usually have docked a compound into the crystal and use this to help predict where modifications could be added to provide increased potency or selectivity
Virtual screen	Docking models: interrogation of a virtual compound library with the X-ray structure of the protein or, if have a known ligand, as a base to develop further compounds on	Can provide the starting structures for a focused screen without the need to use expensive large library screens. Can also be used to look for novel patent space around existing compound structures
Physiological screen	A tissue-based approach for determination of the effects of a drug at the tissue rather than the cellular or subcellular level, for example, muscle contractility	Bespoke screens of lower throughput. Aim to more closely mimic the complexity of tissue rather than just looking at single readouts. May appeal to academic experts in disease area to screen smaller number of compounds to give a more disease relevant readout
NMR screen	Screen small compounds (fragments) by soaking into protein targets of known crystal or NMR structure to look for hits with low mM activity which can then be used as building blocks for larger molecules	Use of NMR as a structure determining tool

Finding Hits with HTS



Finding Hits with Physical HTS

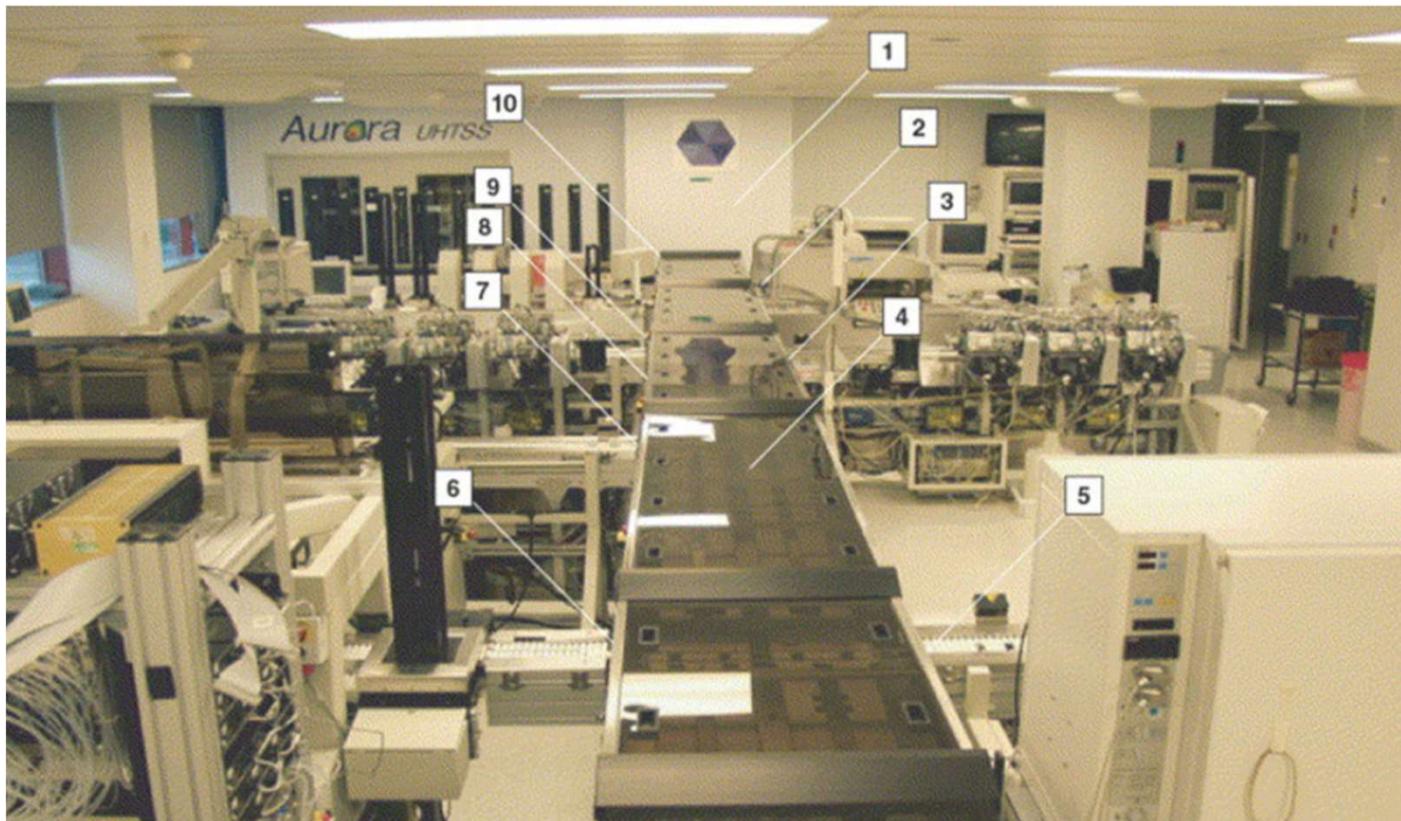
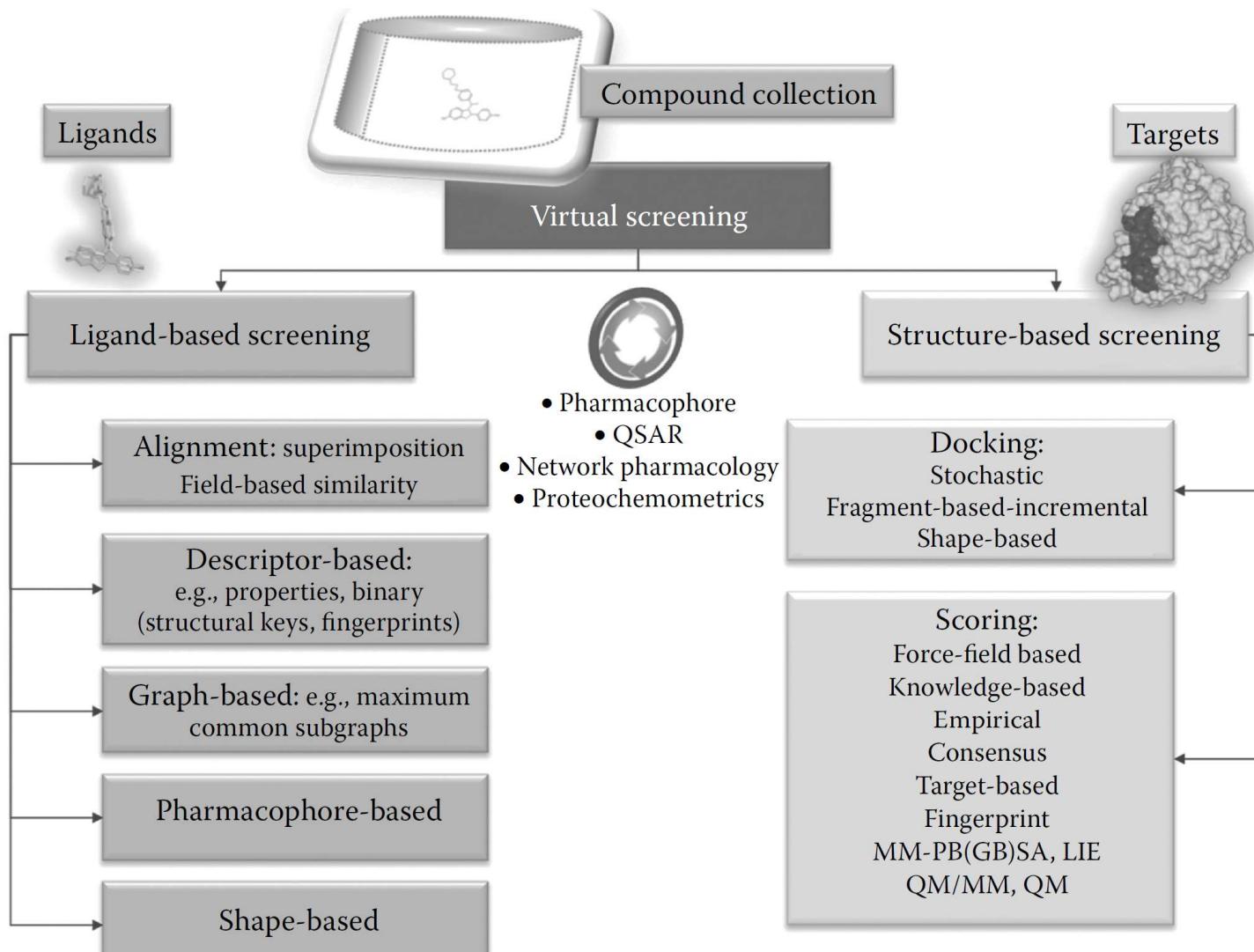
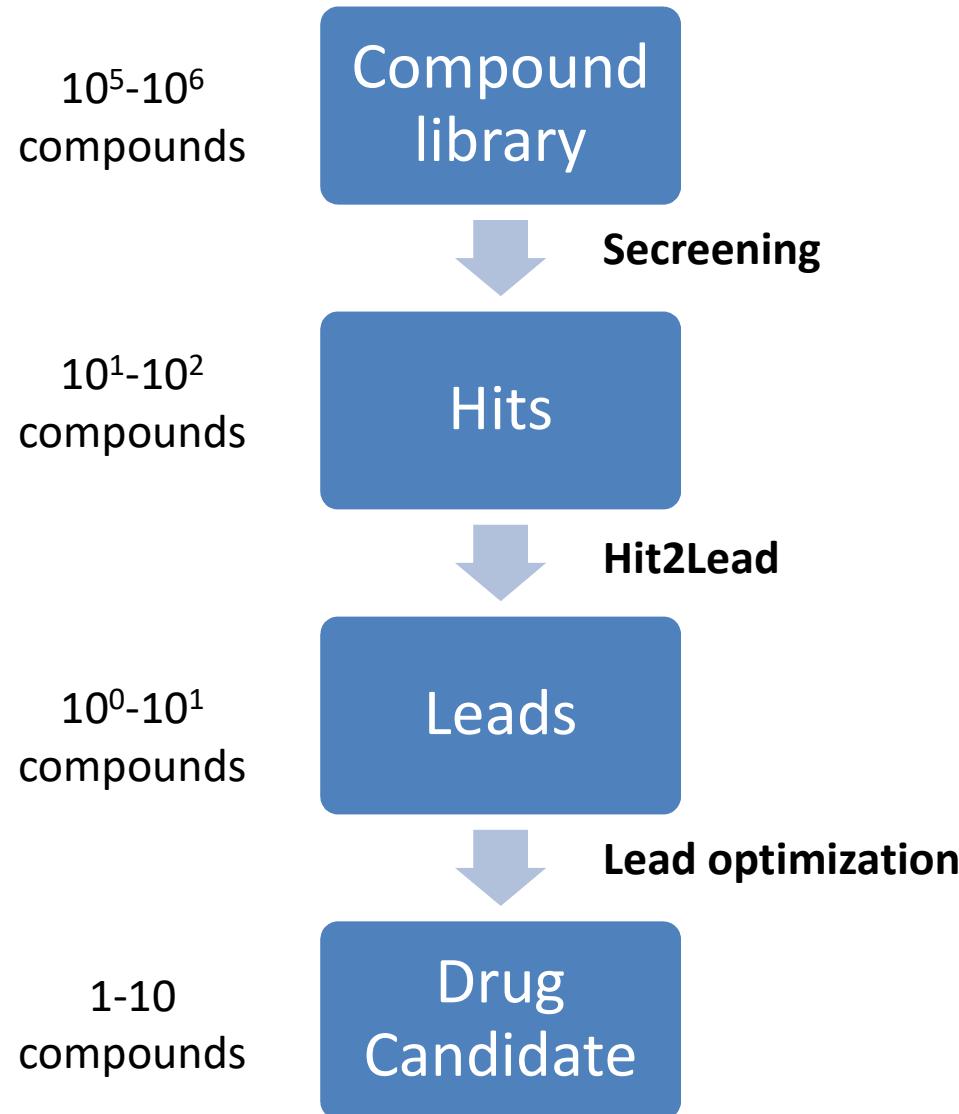


FIGURE 2.17 The automated uHTS system at Bristol-Myers Squibb. Integral components and subsystems are shown; (1) Compound store, (2) Hit-picking robot, (3) 3456 reagent dispensing robot, (4) Transport, (5) Incubators, (6) Piezo-electric distribution robot, (7) Topology compensating plate reader, (8) 1536 reagent dispensing robot, (9) Automated plate replicating system, (10) High-capacity stacking system. Source: Reprinted from Cacace, A.; Banks, M.; Spicer, T.; Civoli, F.; Watson, J. An ultra-HTS process for the identification of small molecule modulators of orphan G-protein-coupled receptors. *Drug Discovery Today*, 8 (17), 785–792, copyright 2003, with permission from Elsevier.

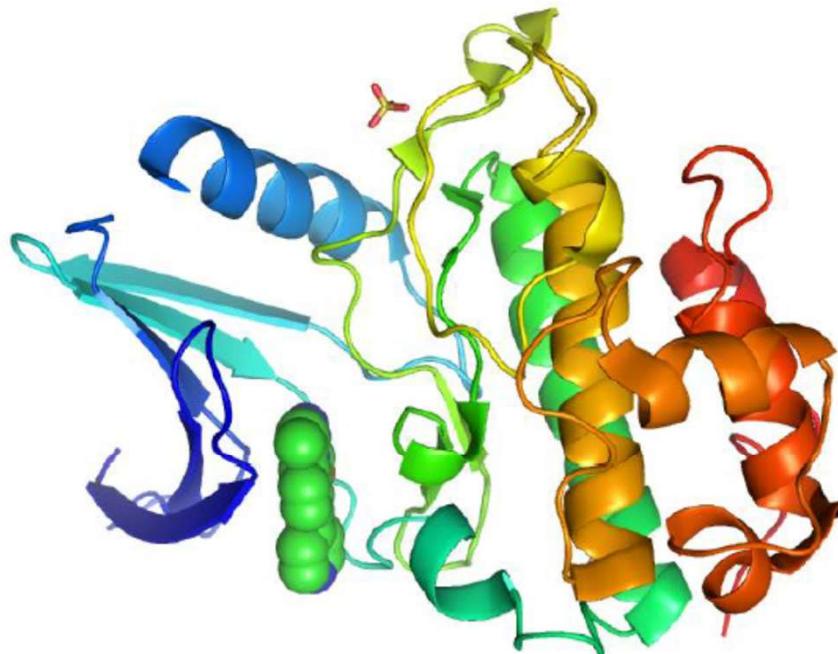
Finding Hits with Virtual HTS





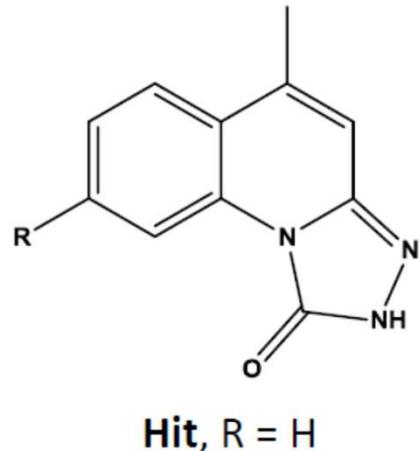
Hit2lead by example: Checkpoint kinase inhibitor

Checkpoint kinase (ChK) is a protein kinase that is activated in response to DNA damage and is involved in cell cycle arrest. This target is highly perspective as drug target for curing the cancer disease (esp. breast cancer, Li-Fraumeni syndrome and other type of cancers).

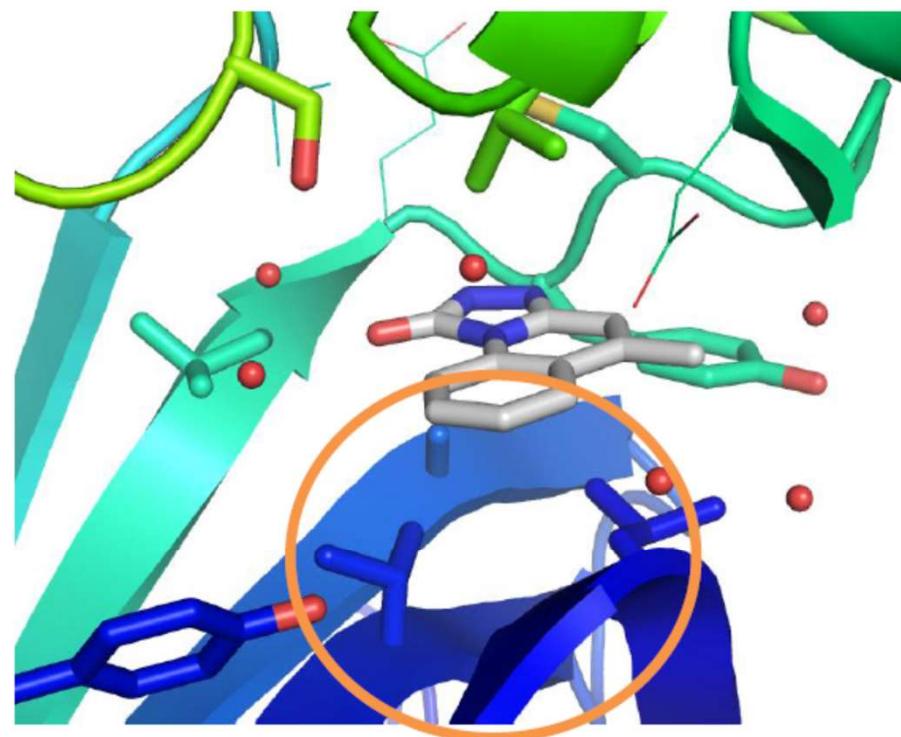


Discovery of a novel class of triazolones as Checkpoint Kinase inhibitors—Hit to lead exploration (AstraZeneca). DOI: [10.1016/j.bmcl.2010.07.015](https://doi.org/10.1016/j.bmcl.2010.07.015). PDB: 2BXD, 2BXI, 2BXE.

Hit2lead by example: Checkpoint kinase inhibitor



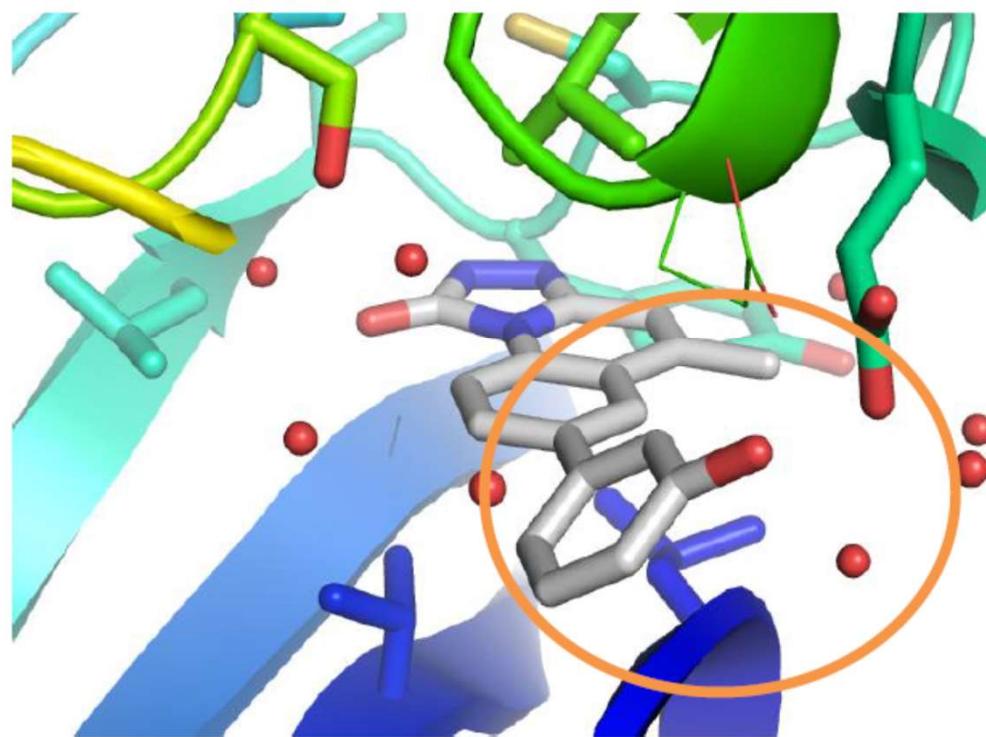
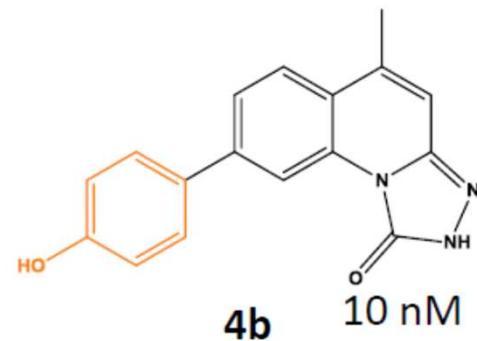
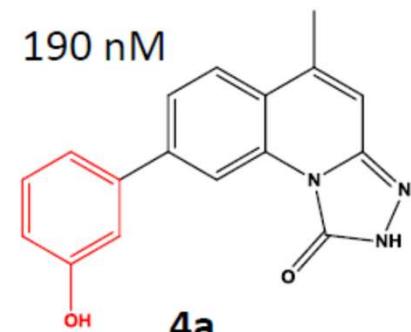
It was found that hit is selective to the Checkpoint kinase (ChK) among the others found by HTS.



Synthesis of the derivates was made to fill the so-called sugar pocket.

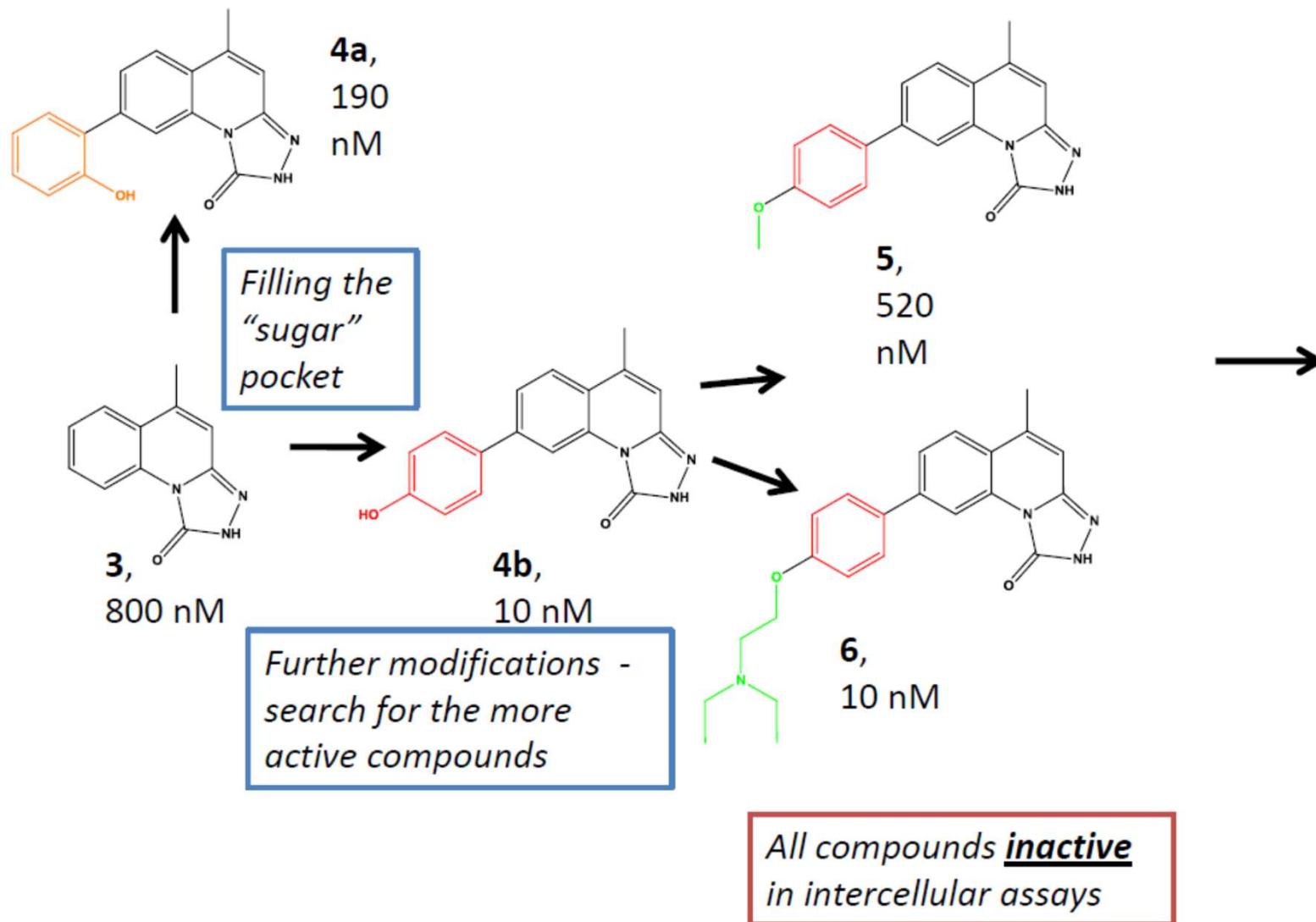
Discovery of a novel class of triazolones as Checkpoint Kinase inhibitors—Hit to lead exploration (AstraZeneca). DOI: [10.1016/j.bmcl.2010.07.015](https://doi.org/10.1016/j.bmcl.2010.07.015). PDB: 2BXD, 2BXI, 2BXE.

Hit2lead by example: Checkpoint kinase inhibitor

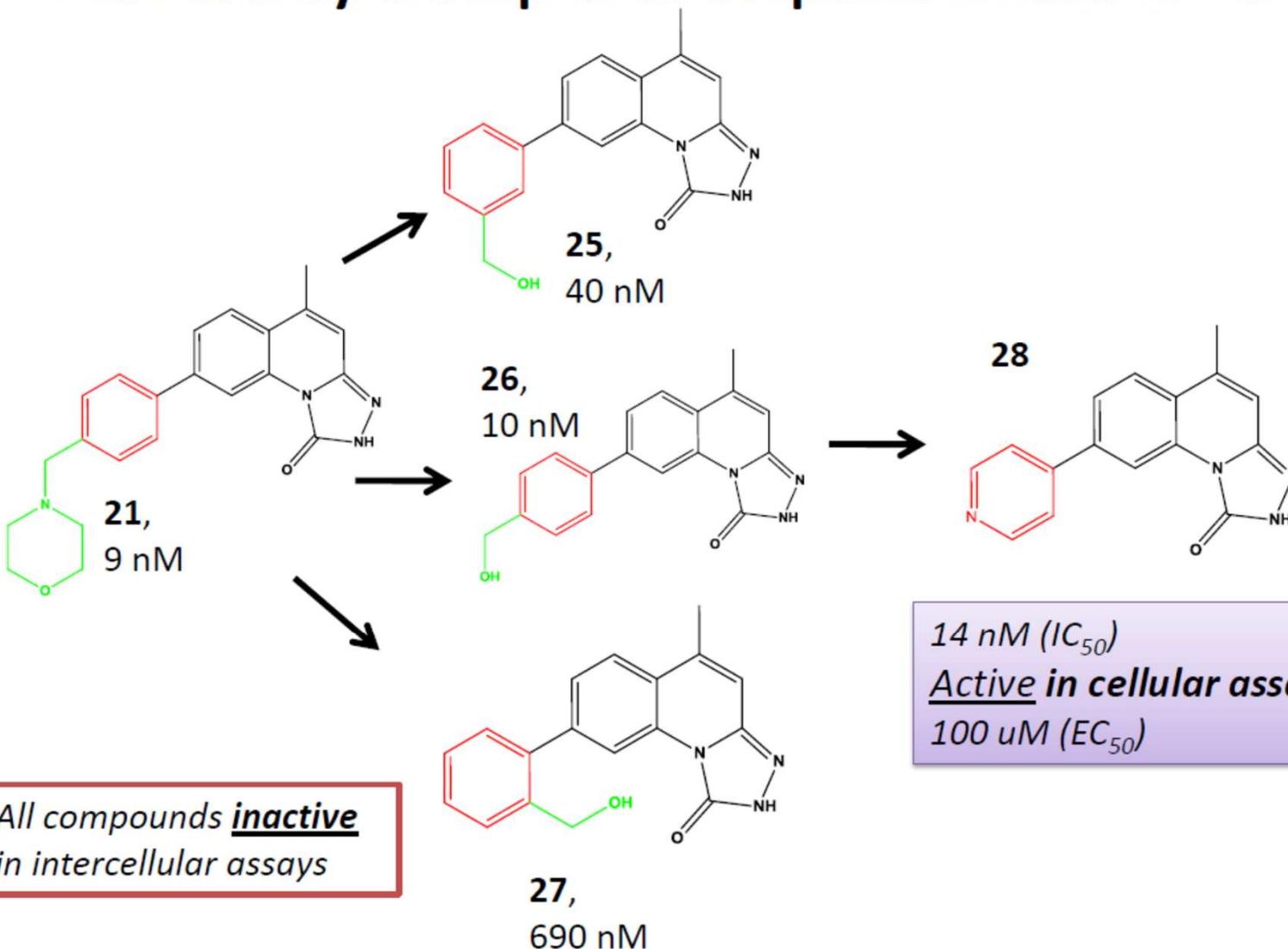


4a and **4b** analogues
filled the sugar pocket.

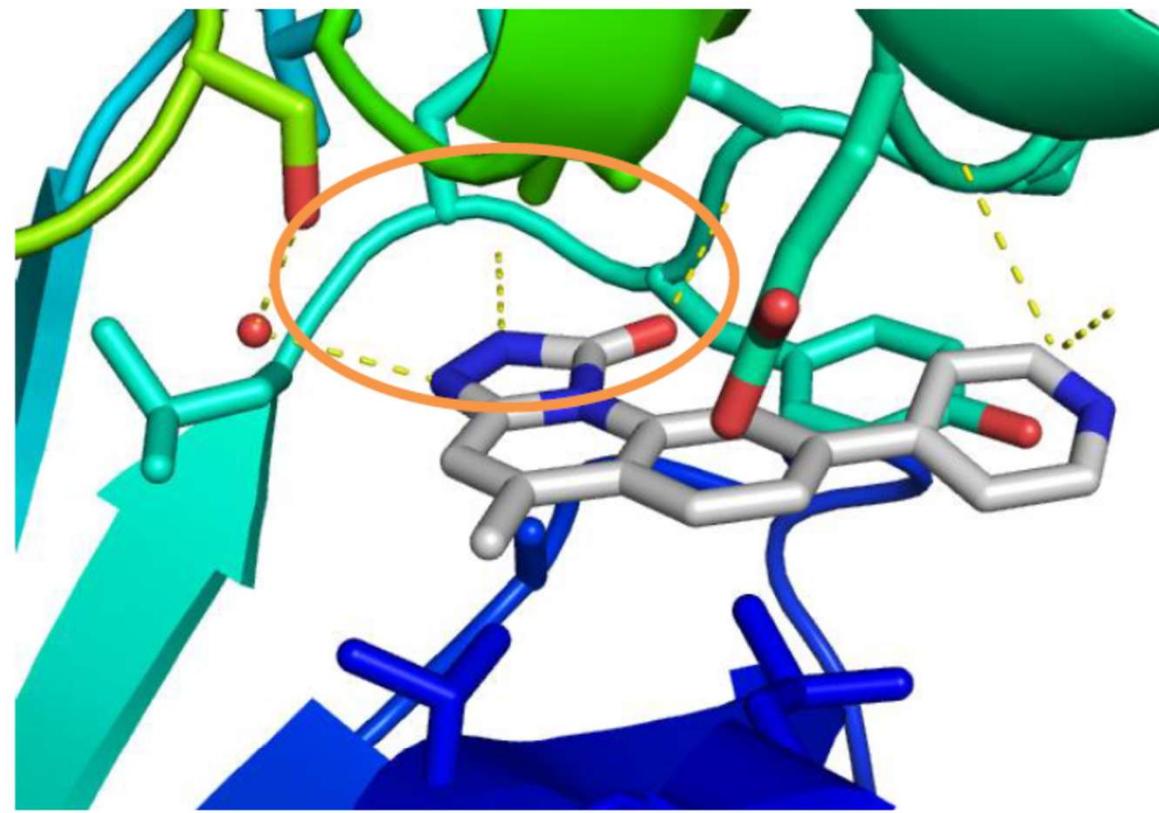
Hit2lead by example: Checkpoint kinase inhibitor



Hit2lead by example: Checkpoint kinase inhibitor

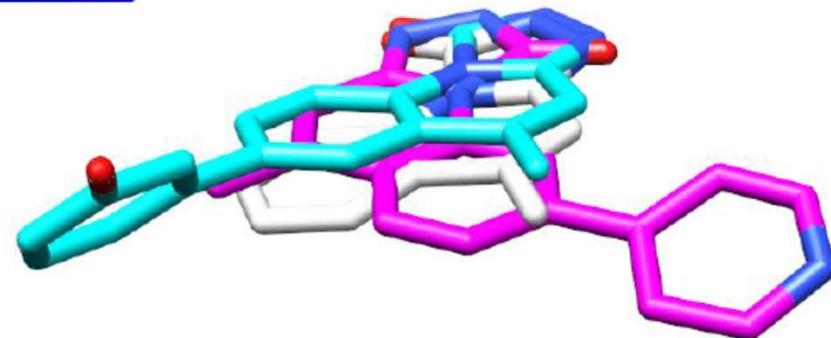


Hit2lead by example: Checkpoint kinase inhibitor



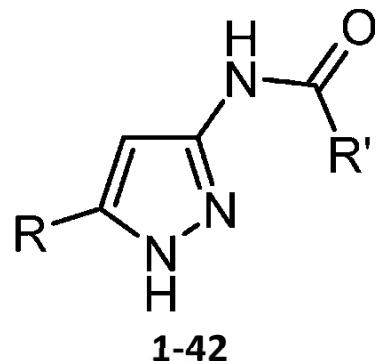
PDB: 2BXE

Superposition of the ligands from the three PDB structures. The resultant derivate binds in different way but forms almost the same pattern of hydrogen bonds.

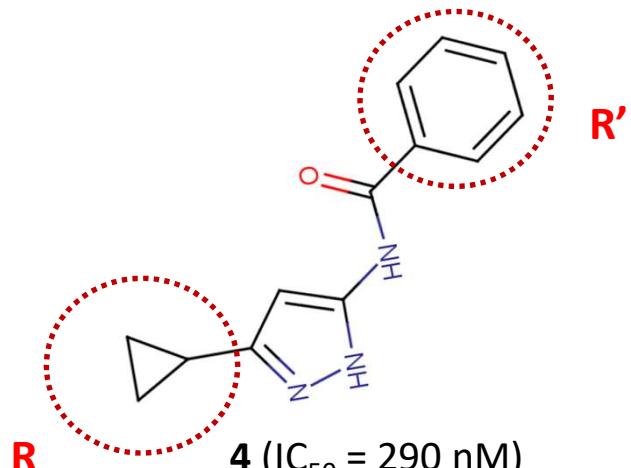
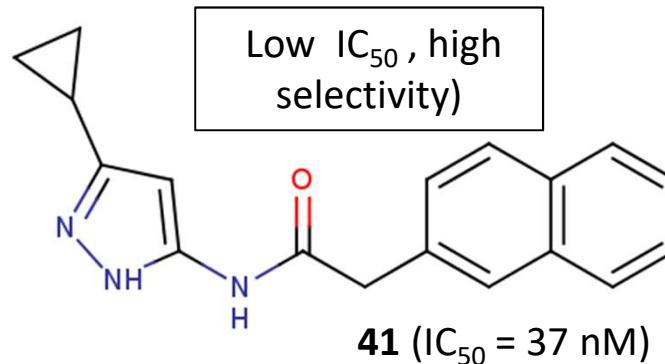


Lead Optimization Example: CDK2 inhibitors

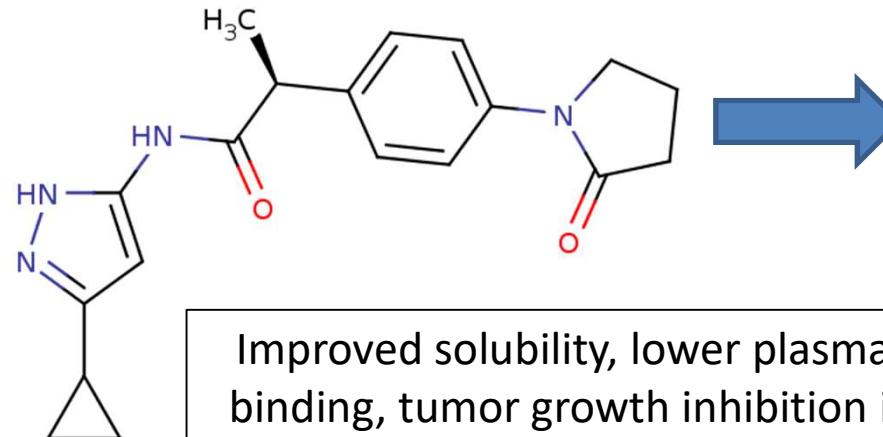
42 hit compounds



Select lead



Optimize lead



Improved solubility, lower plasma protein binding, tumor growth inhibition in *in vivo* mouse model