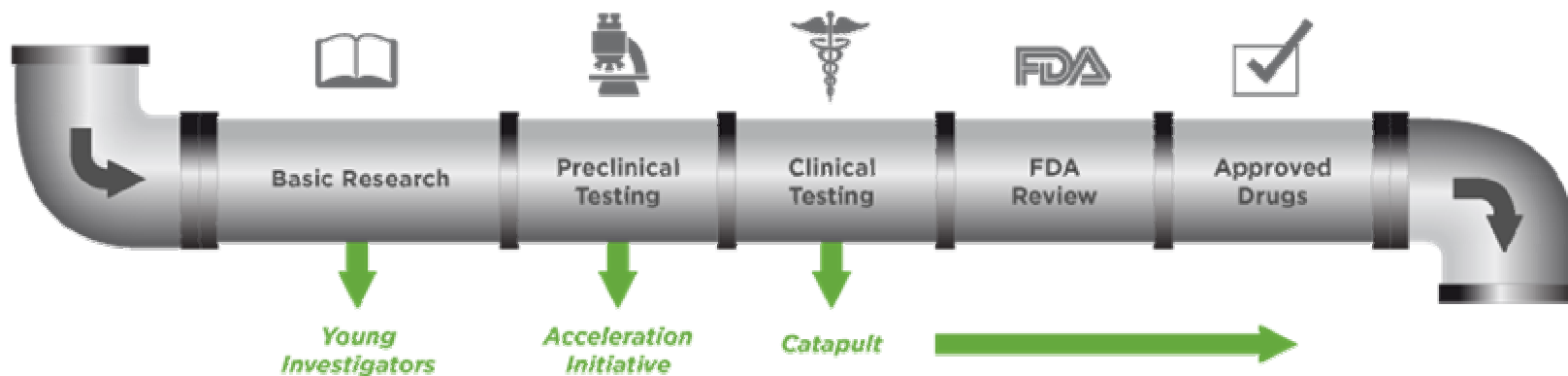


Modern Methods in Drug Discovery

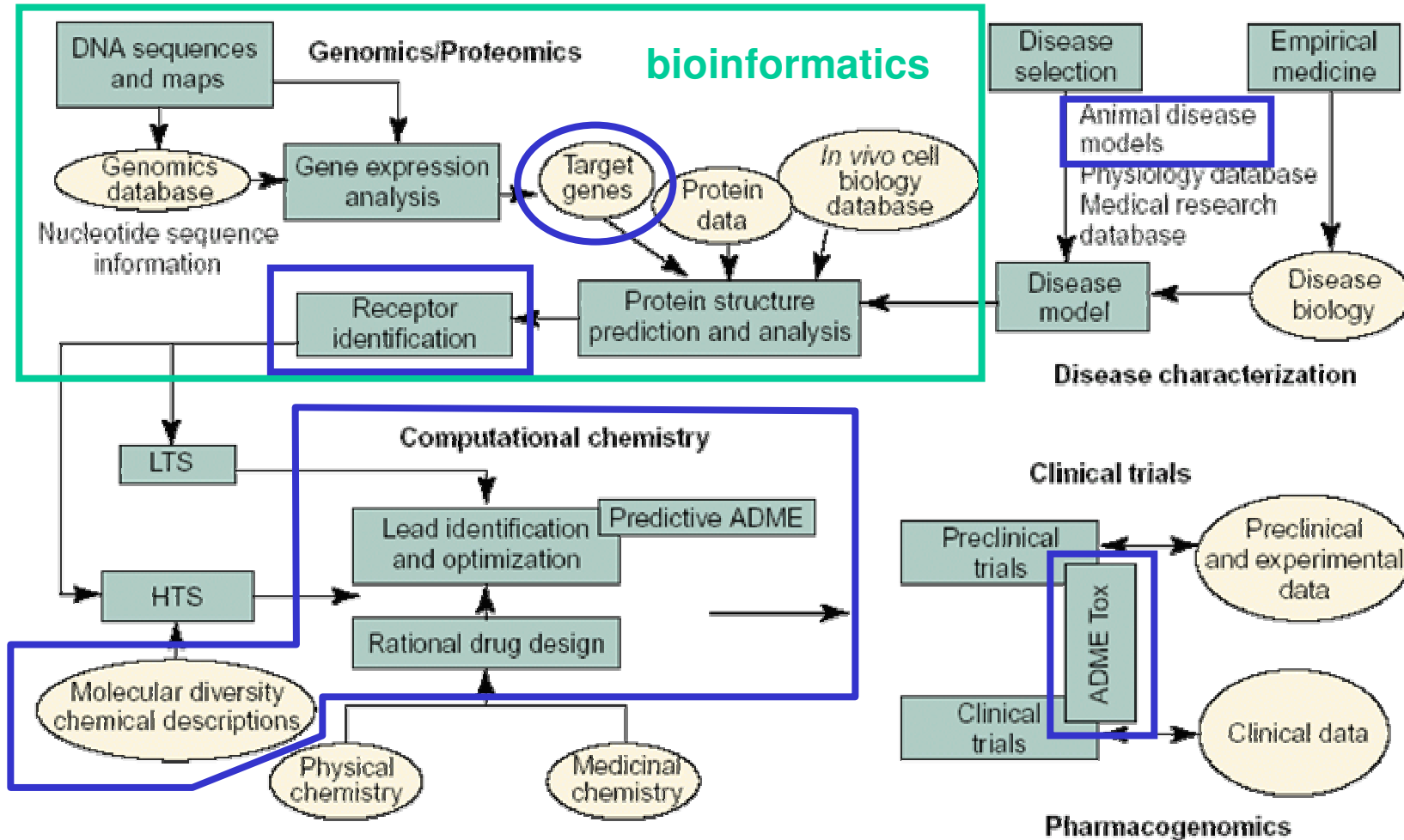
Aims of this course:

- comprehensive knowledge about all processes in the *drug discovery pipeline*
- in particular *in silico* methods of *drug design*



Picture source: <https://curesearch.org/Impact-Report-Winter-2016/images/research-pipeline.png>

Flow of information in a *drug discovery pipeline*



Covered by this lecture

Drug Discovery Today

Related topics not covered by this lecture

medicinal chemistry

organic synthesis

biopharmaceutical aspects

(tissue models, non-oral administration)

clinical aspects

molecular modelling theory

homology modelling theory

docking basics and applications

computational chemistry

genome, proteome, metabolome

bioethics, legal aspects, and patent law



Required knowledge

Use of tools for sequence analysis, e.g. BLAST, CLUSTALW

Use of visualizing tools, e.g. BALL, VMD, SPDBV, JMOL

recommended courses:

Softwarewerkzeuge der Bioinformatik

Computational Chemistry

Bioinformatics I + II

Actual applications during the exercises and homework:

multiple sequence alignment, homology in sequences

analyzing protein-ligand interactions

SMILES and SMARTS notation of chemical structures

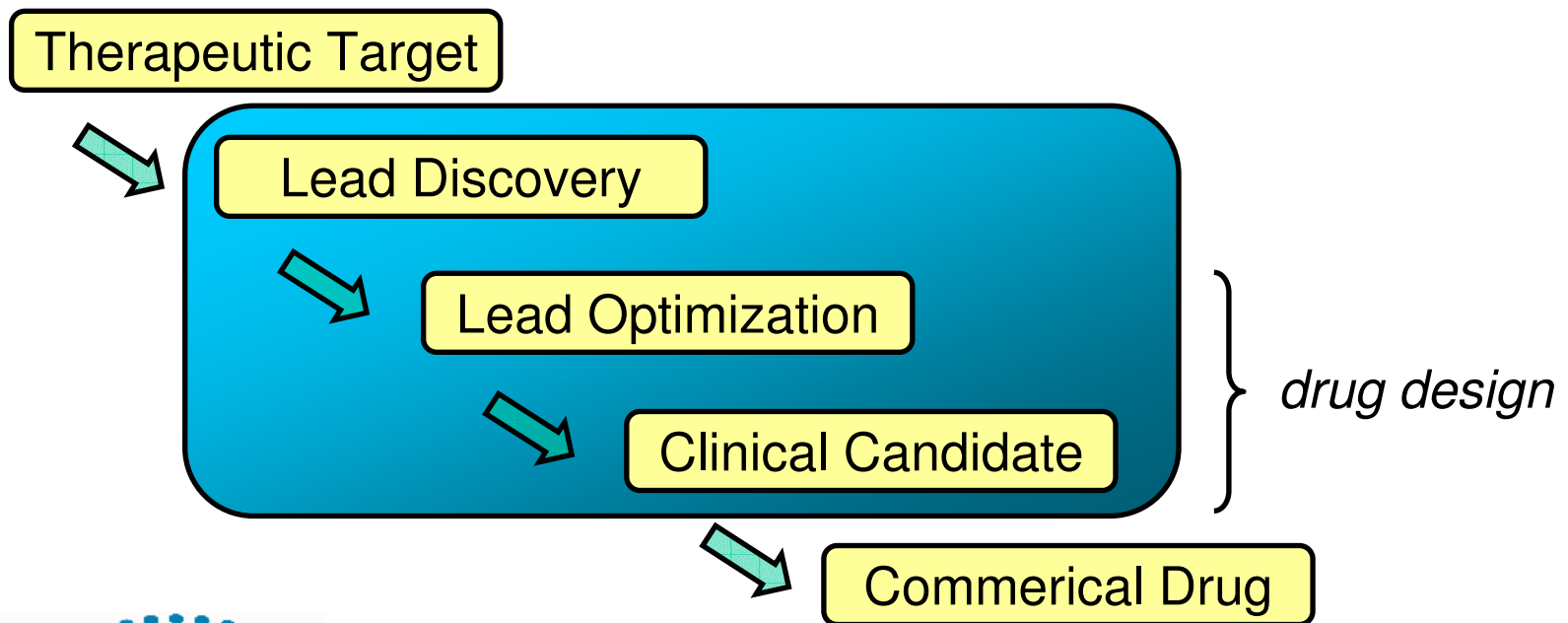
using SMARTS with Open Babel

substance database queries (PubChem, ChEMBL,...)

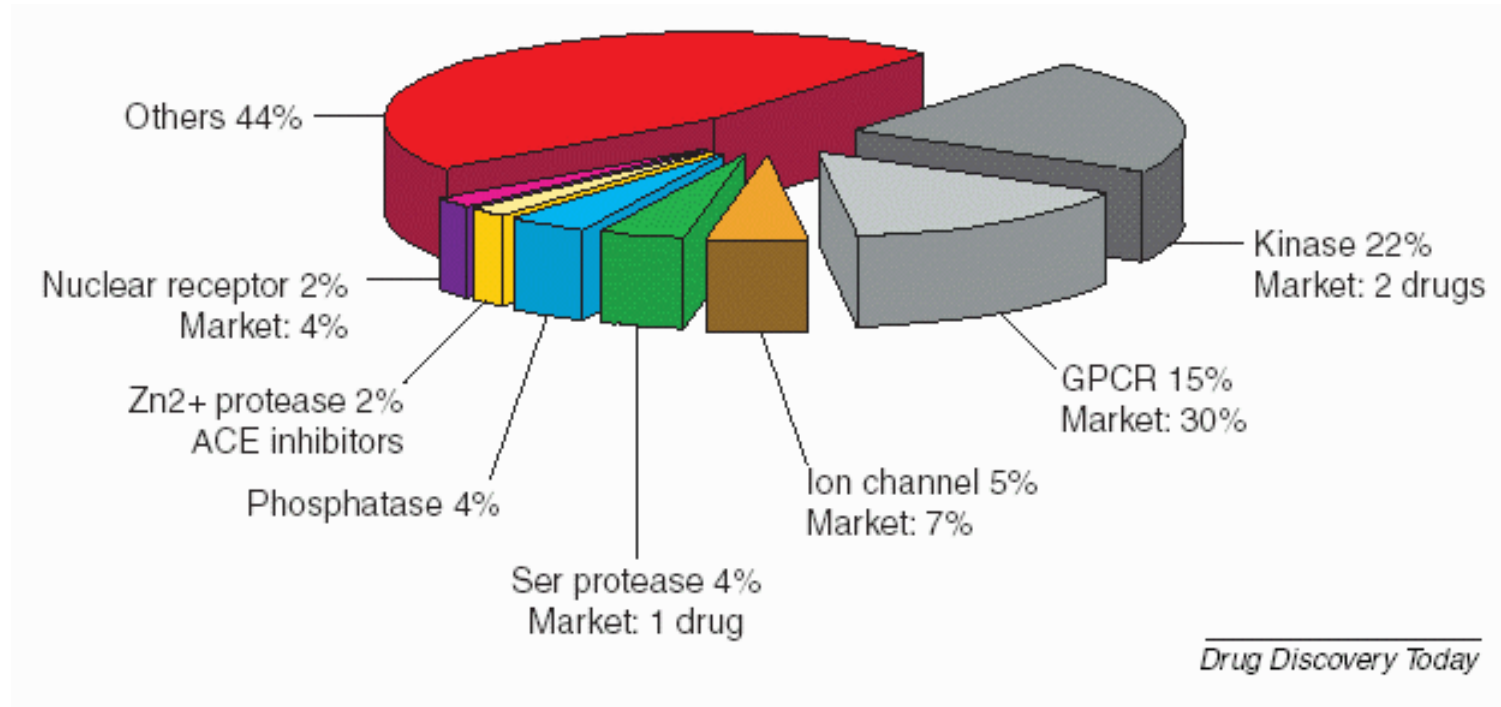


What is *drug discovery* ?

rational and targeted search for new drugs



Typical targets (I)

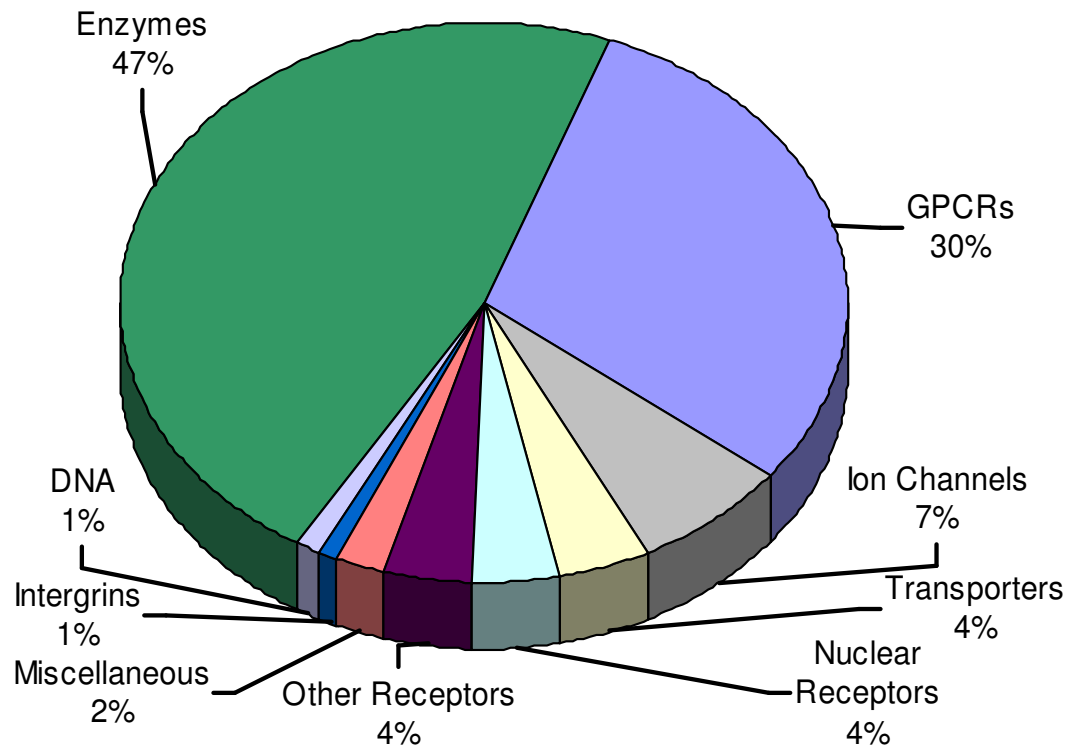


contribution to the human genome and marketed drugs



Typical targets (II)

drug targets by biochemical class



Fractional content of marketed drugs according to their biochemical targets

data: Hopkins & Groom, *Nat.Rev.Drug.Disc.* **1** (2002) 727



Preliminary schedule (lectures)

1. Introduction, overview, recap of chemical structures
2. typical diseases
3. properties of drugs and their mode of action
4. Substance databases and bioisosteric compounds
5. QSAR, statistics and descriptors
6. ADME models
7. metabolism and toxicology
8. target identification, animal models
9. cytochrome P450, polymorphisms, transporters
10. more complex diseases malaria, obesity
11. in silico prediction of molecular properties
12. current trends, disease vs. lifestyle drugs, doping



Preliminary schedule (exercises)

Biweekly in the CIP-Pool (building E 2.1 room 003)
computer account and access card required

- chemical structures of drugs
- enzyme-ligand interactions, analysis of .pdb files
- substance databases and SMARTS queries
- QSAR, statistics and descriptor handling
- ortholog targets in model organisms
- further online tools

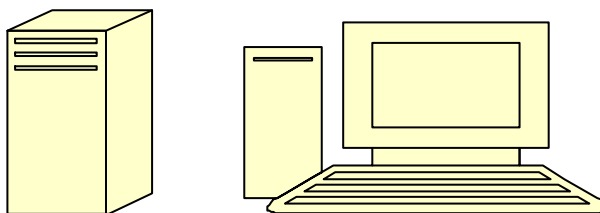


Compound Databases

existing substance libraries

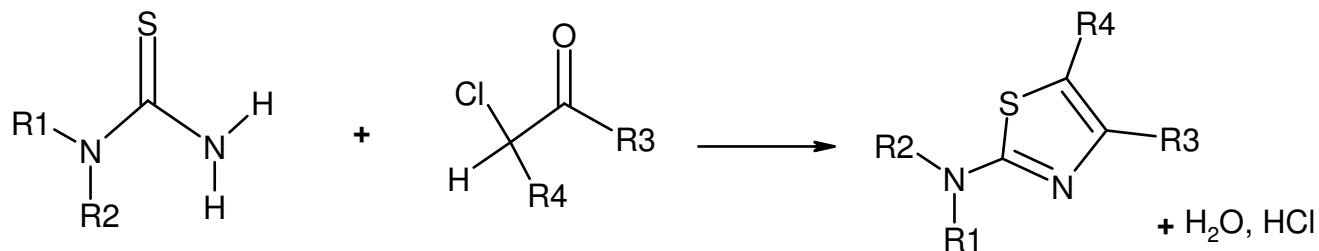
ACD	>100,000 chemicals	} commercial company, in house
World Drug Index	58,000 compounds	
USAN	<10,000 in clinical trial	
virtual library	≈100,000 compounds	

PubChem	> 3,000,000 compounds	NCBI	} academic
ChEMBL	> 1,213,000 compounds	EMBL	
ZINC	>73,126,243 compounds	UCSF	

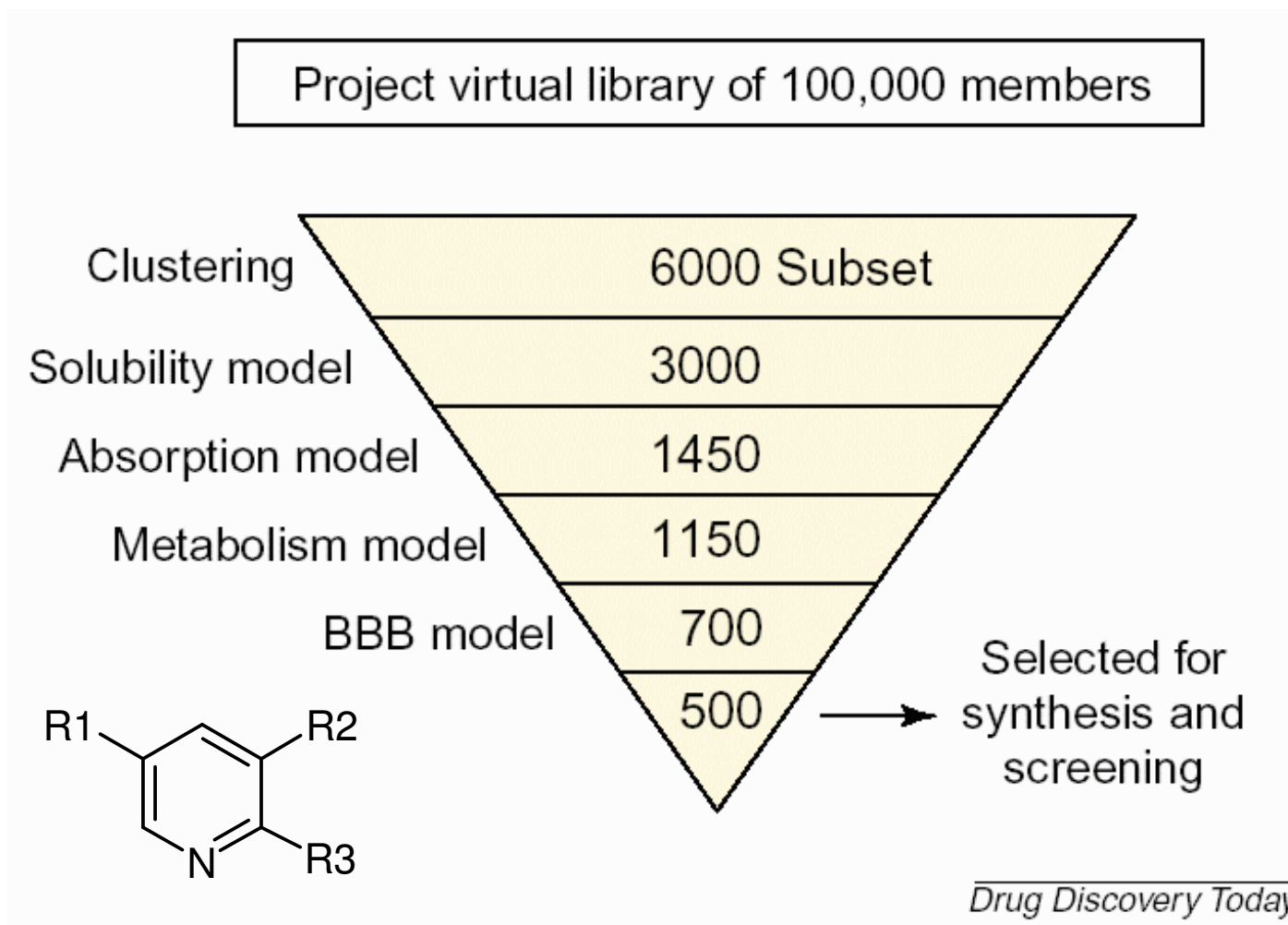


Investment per new chemical entity: >500,000 \$
New chemical entities per year: ca. 15

Methods of Combinatorial Synthesis for High Throughput Screening (HTS)



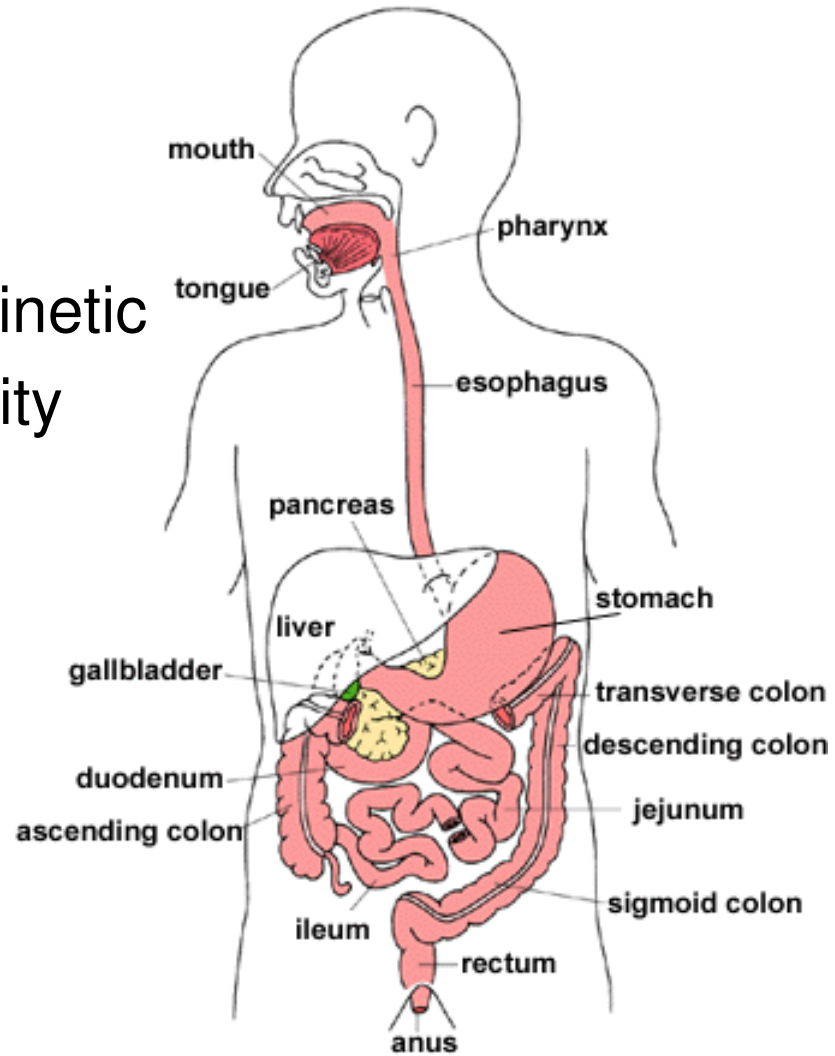
Selection of compounds for High Throughput Screening (HTS)



Predictive ADME

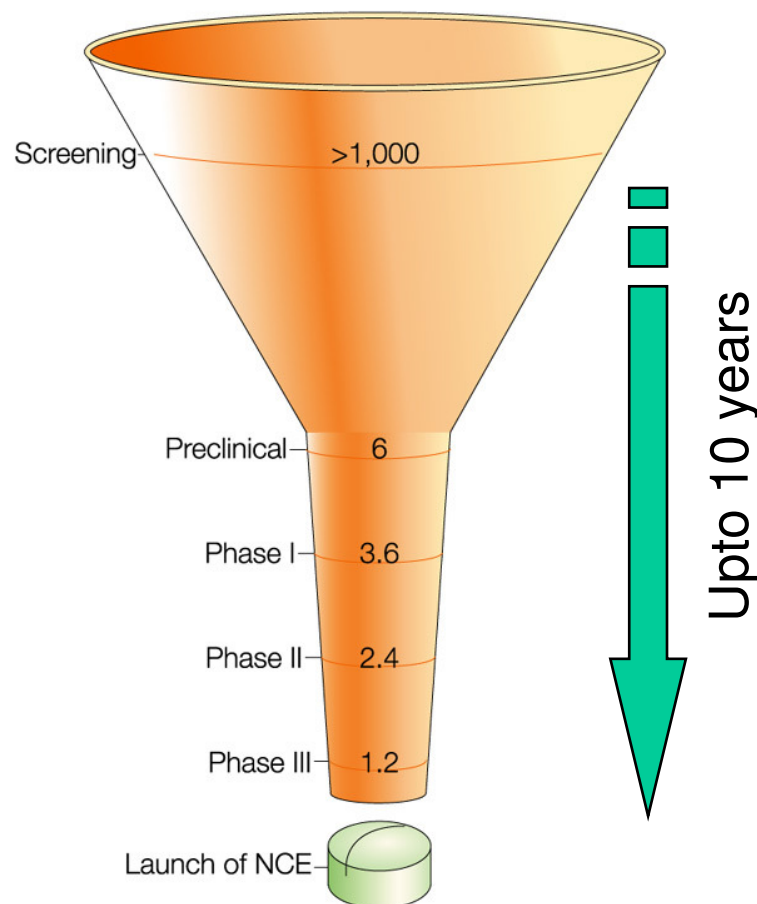
Absorption
Distribution
Metabolism
Elimination

Pharmacokinetic
Bioavailability



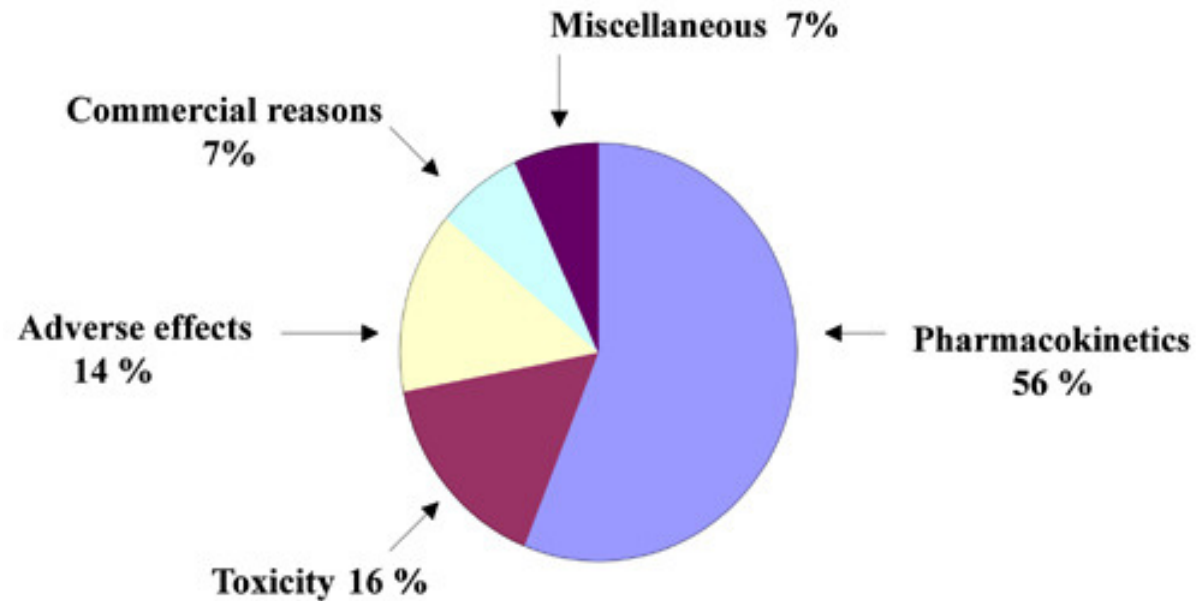
From the pipeline until the commercial launch

For each actual marketed drug (*new chemical entity, NCE*) there have been more than 1000 substances that underwent screening *in vitro*. Without the use of available computer-based ADMET filters, this number would be even larger.



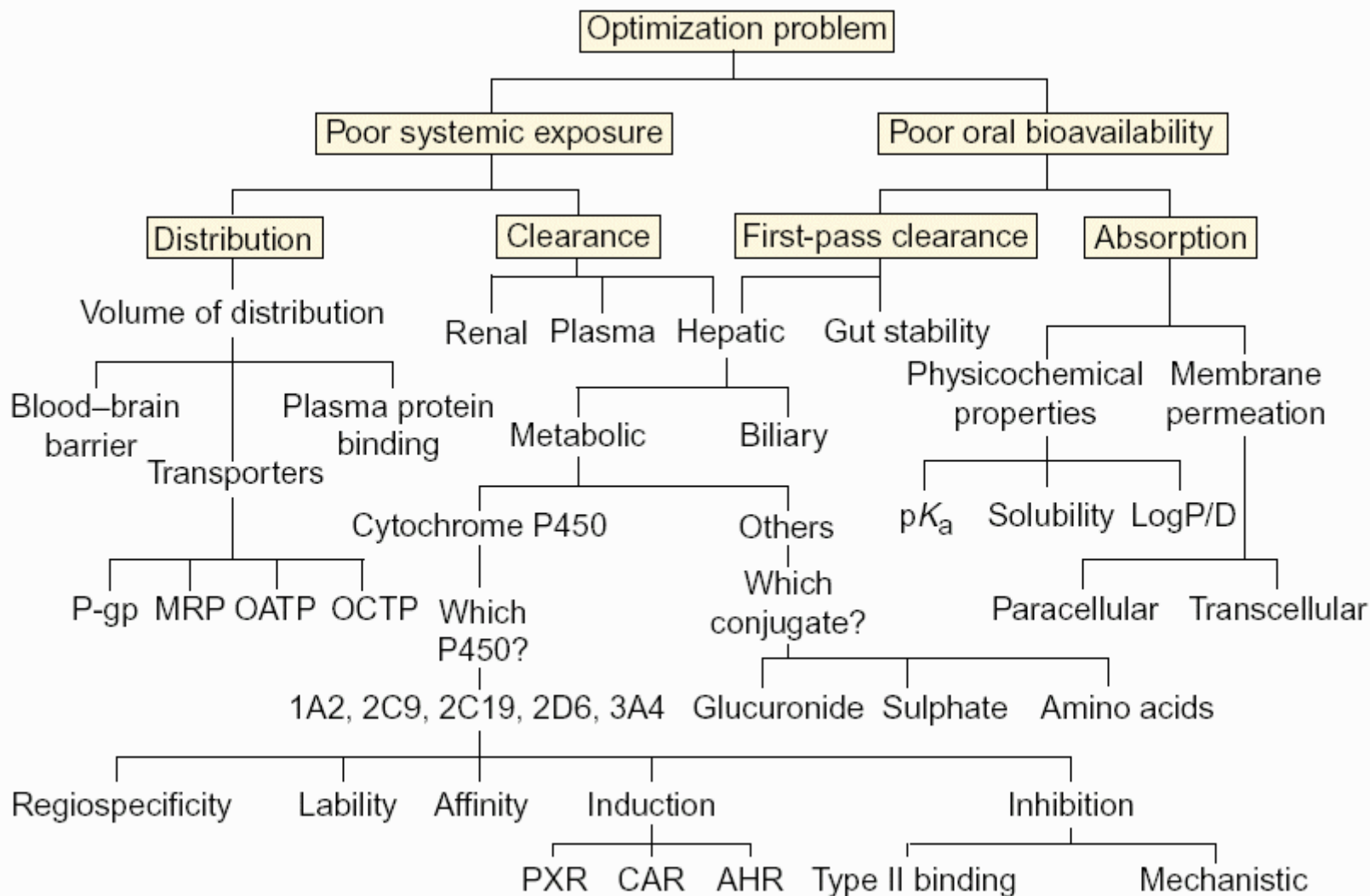
Nature Reviews | Drug Discovery

Why is the prediction of ADME parameters that important ?



Reasons that lead to failure or withdrawal of a potential drug by the mid 1990's

Pharmacokinetics and Bioavailability



Drug Discovery Today

Metabolism

(bio-)chemical reactions of xenobiotics in the body

First pass effect:

Extensive metabolization of mainly lipophilic molecules, such with $MW > 500$, or those that have a specific affinity to certain transporters, during the first passage through the liver

Phase I:

Oxidation, reduction and hydrolysis →
esp. cytochrome P450 enzymes

Phase II:

Conjugation with small molecules (e.g. glutamine)

Phase III:

elimination by transporters



Cytochrome P450 Enzymes (I)

flavin monooxygenase isoenzyme (FMO)

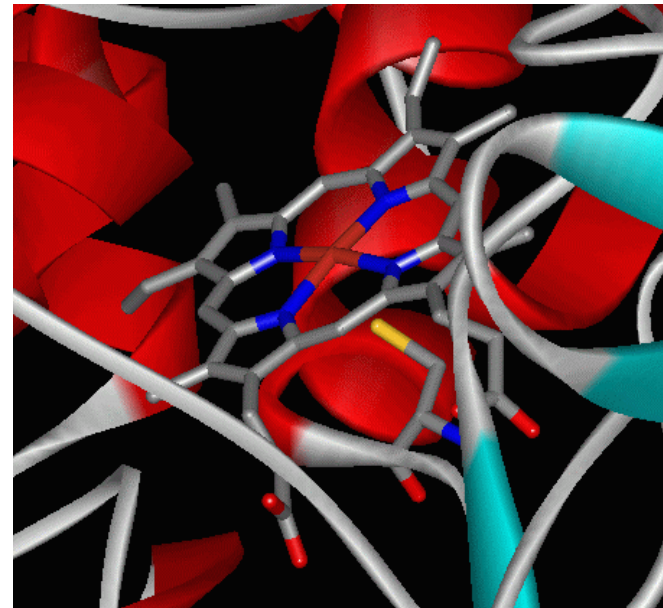
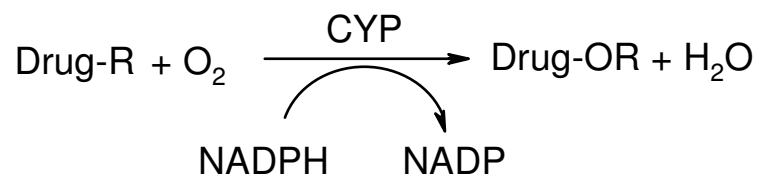
monoamine dehydrogenase (MAO)

aldo-keto reductase (AKR)

alcohol dehydrogenase

aldehyde oxidase

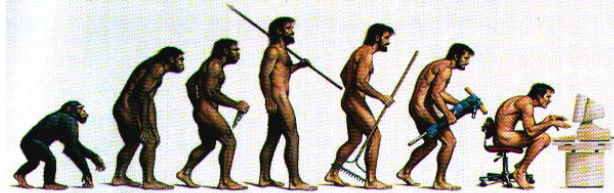
} Further
phase I
enzymes



Cytochrome P450 gene families



Plants 22



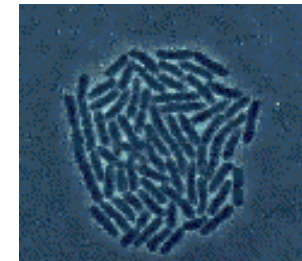
Human 17+



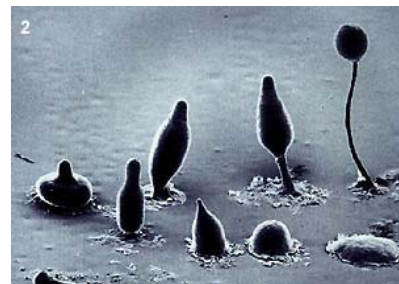
Molluscs 1



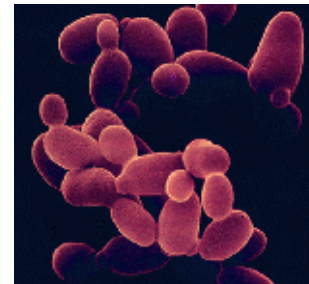
Insects 3



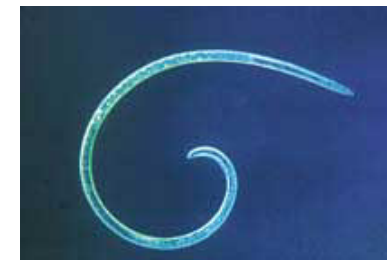
Bacteria 18



Fungi 11



Yeasts 2



Nematodes 3



Cytochrome P450 polymorphism

„Every human is (more or less) different“

Determination of the phenotype by the actual activity or the amount of the expressed enzyme.

In contrast, the genotype is determined by the individual DNA sequence.

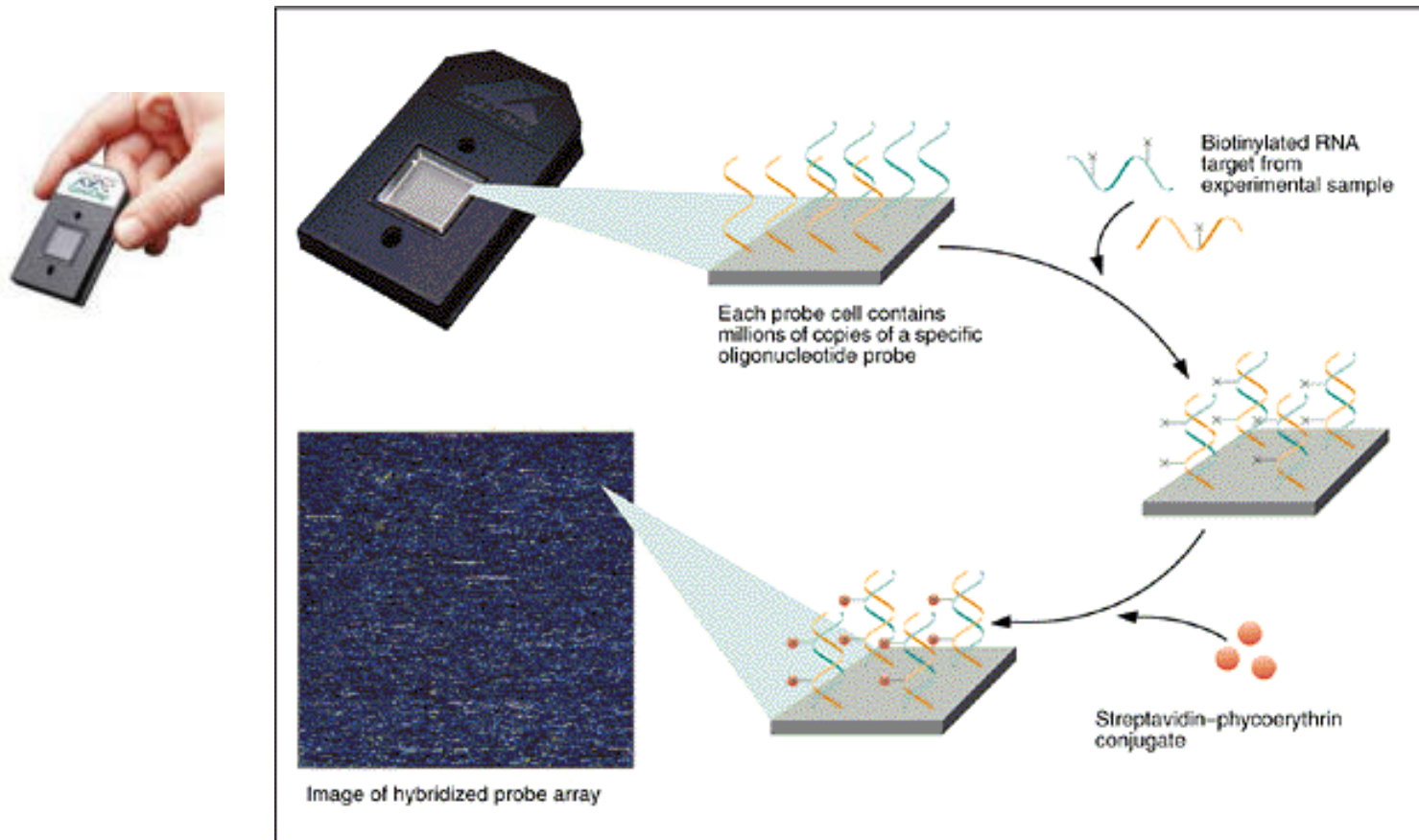
Thus, the same genotype enables several different phenotypes

According to their metabolic activity of CYP there is a classification into normal (*extensive metabolizer*), weak (*poor metabolizer*), und accelerated (*ultra-rapid metabolizer*) metabolism.

Lit: K. Nagata et al. *Drug Metabol. Pharmacokin* **3** (2002) 167

Genotyping of CYP P450 alleles

By using immobilized, synthetic copies of P450 nucleotides, the Affymetrix company (USA) has developed microarrays (gene chips) that allow the identification of all clinically relevant alleles.



Prediction of molecular properties (I)

The keynote of *rational drug design*

The general question is:

What is the connection between the biological space (activity) and the chemical space (structure) ?

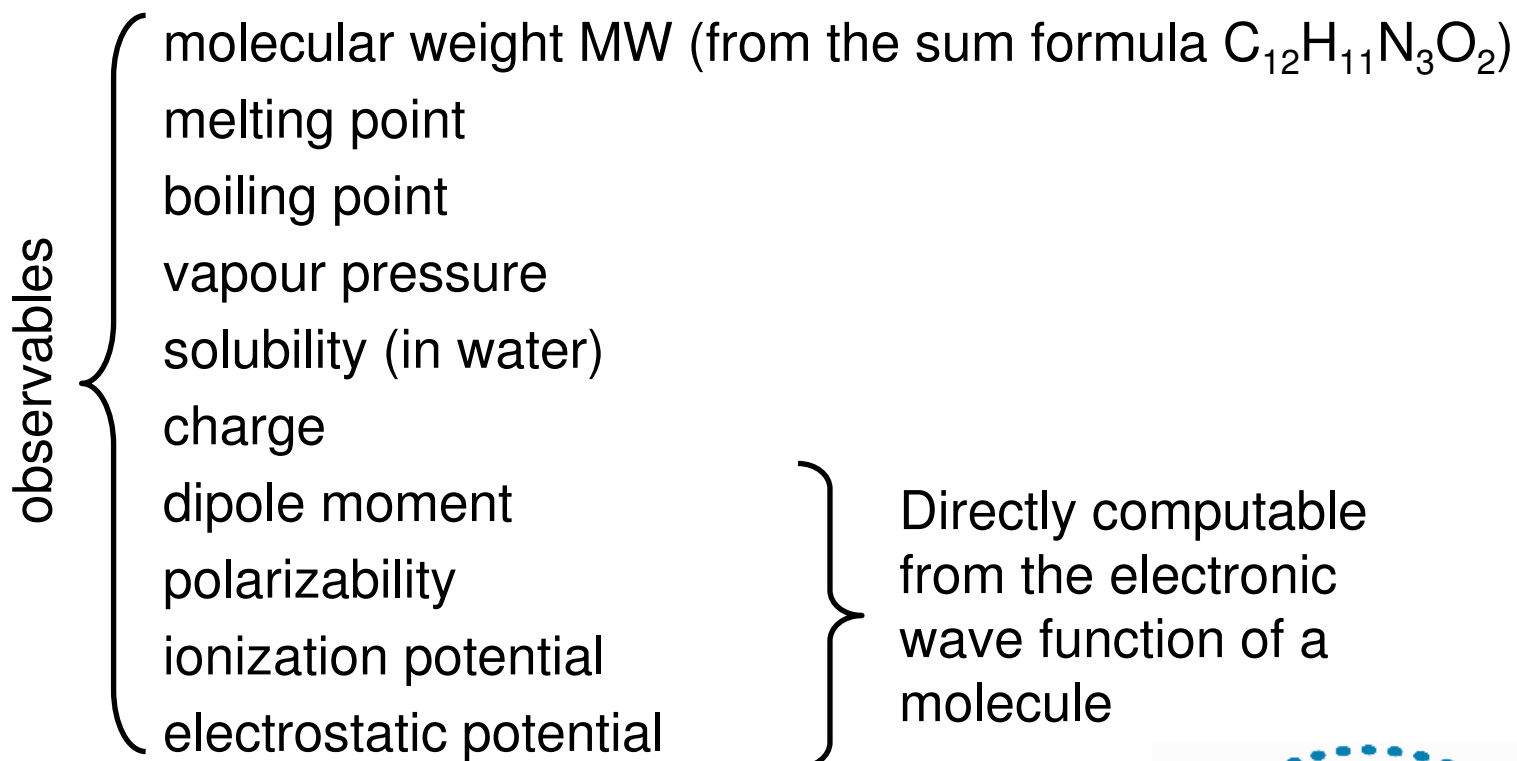
How are we able to make structure-based prediction ?

- QSAR and QSRP, regression analysis
- decision trees, machine learning algorithms
- other statistical methods



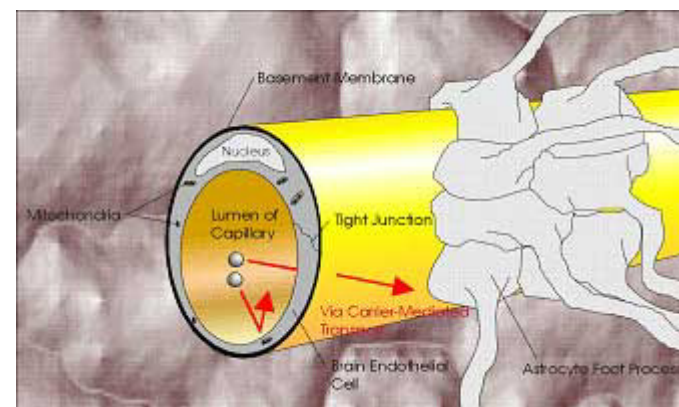
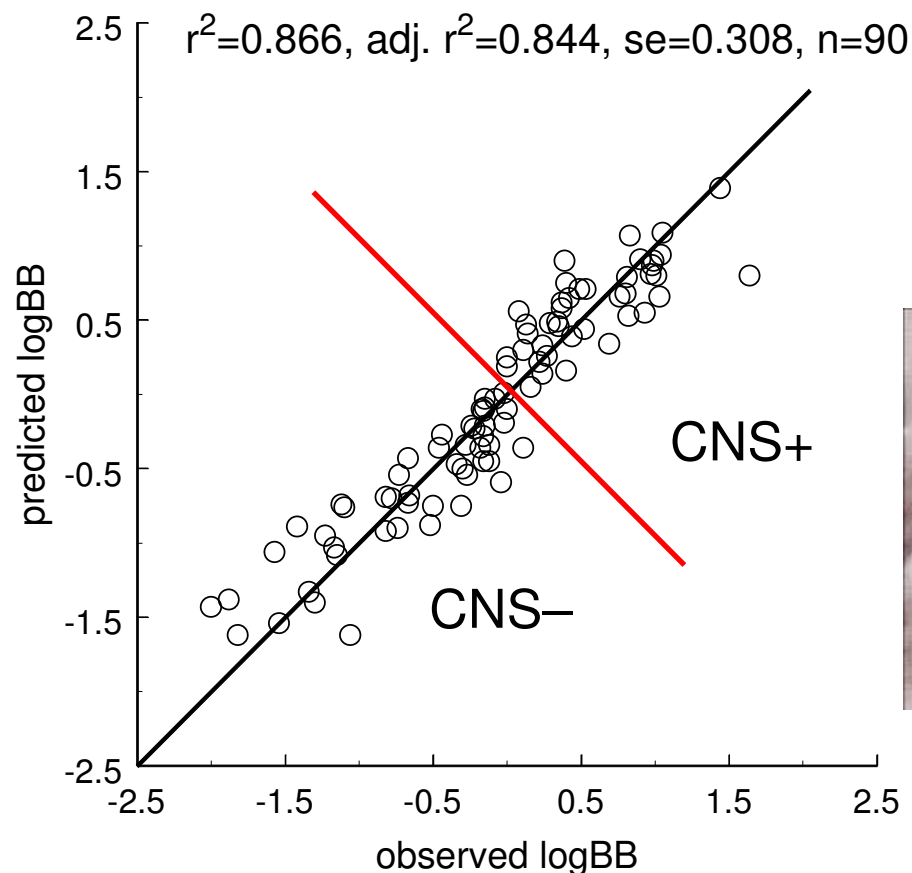
Prediction of molecular properties (II)

What are molecular properties?



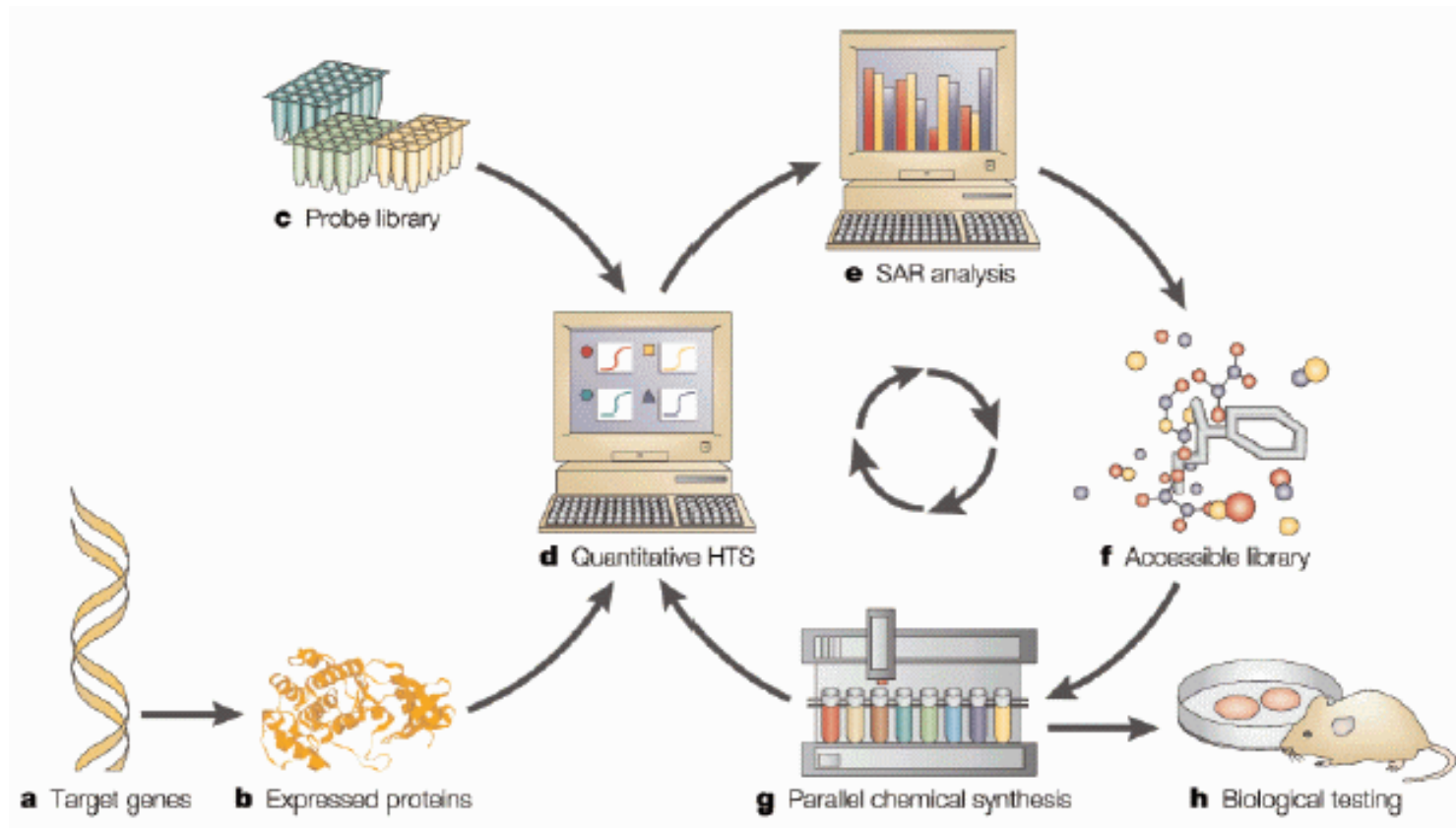
BBB-model with 12 descriptors

Descriptors mainly from QM calculations: electrostatic surface, principal components of molecular geometry, H-bond properties



Lit: M. Hutter *J.Comput.-Aided.Mol.Des.* **17** (2003) 415.

Cycle of optimization in the drug discovery pipeline

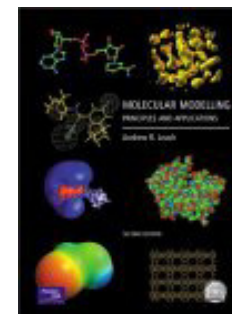


Source: D.K. Agrafiotis et al. *Nature.Rev.Drug.Discov.* **1** (2002) 337.

Accompanying books and further reading (I)

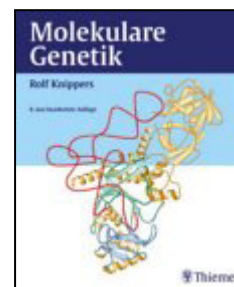
Andrew R. Leach*

Molecular Modelling. Principles and Applications
2nd edition, Prentice Hall, 2001



Rolf Knippers*

Molekulare Genetik
8. Auflage, Thieme, 2001



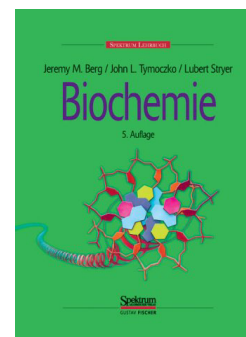
The Merck Index*

13th edition, Merck & CO., Inc., 2001



J.M. Berg, L. Stryer*

Biochemie, Spektrum Verlag
Biochemistry, W.H. Freeman & Co Ltd.



*Available in the „Semesterapparat“

Accompanying books and further reading (II)

Gerhard Klebe*

Wirkstoffdesign

2. Auflage, Spektrum Akad. Verlag, 2009



C.A. Orengo, D.T. Jones, J.M. Thornton*

Bioinformatics

Genes, Proteins & Computers

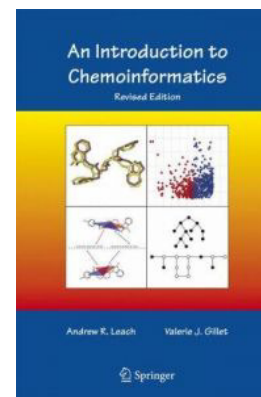
1st ed., Bios Scientific Publishers, 2003



A.R. Leach, V. Gillet*

An Introduction to Chemoinformatics

revised ed., Springer, 2007

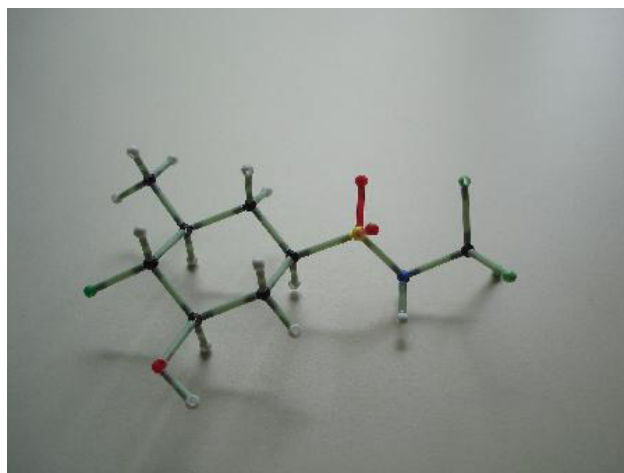
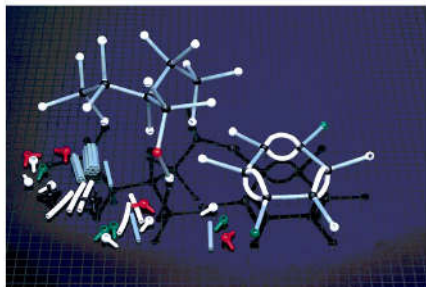


*Available in the „Semesterapparat“



Further hands-on tools

Molecular model sets / Molekülbaukasten



Commerically available at various price ranges

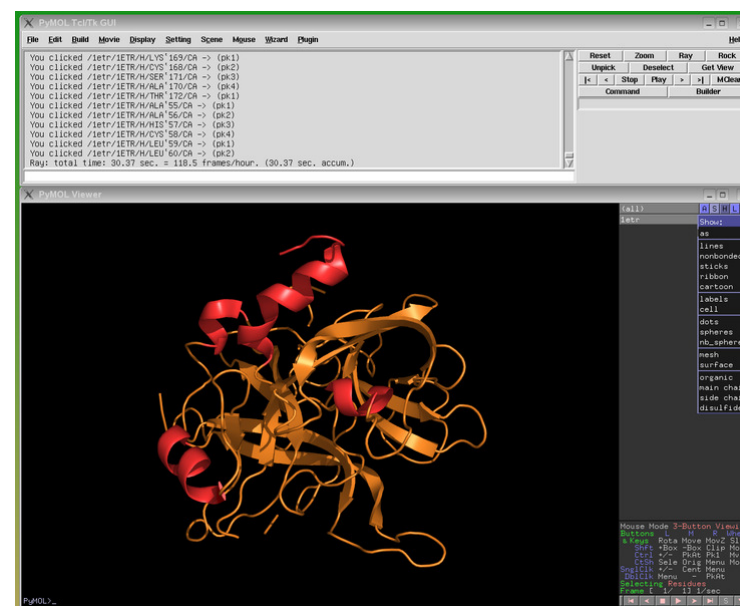
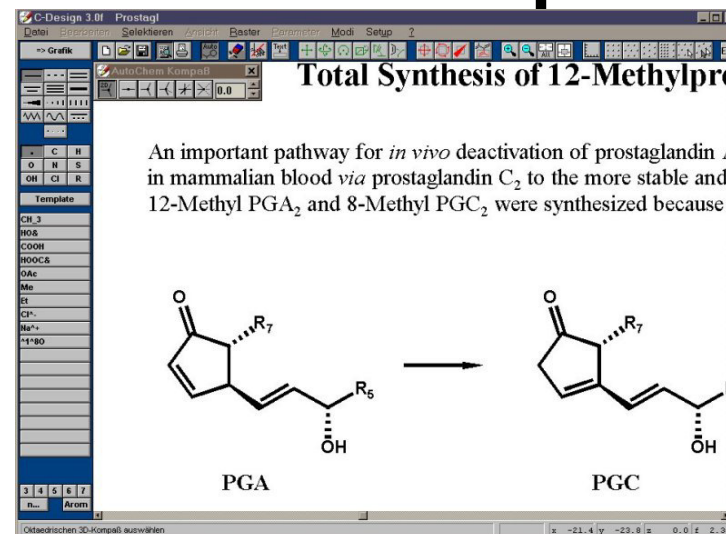
General remark: The lecturer does not endorse any of the mentioned books/software/products. Enquiries are welcome.



Other useful software to make nice pictures

Chemical structures
and other objects:
C-Design 3.0f
Windows-Platform

Protein structures:
PyMOL
www.pymol.org
Linux, Mac OS X, Windows



Requirements to obtain the certificate and the credit points

1. 50% of all accomplishable points from the home work. Two thirds (66.7%) of all assignments (ca. 5-6) must be returned. The assignments have to be handed in until the beginning of the next exercise unit.
2. 50% of all accomplishable points from the final exam taking place at the end of the lecture period. If necessary, repeated (written) exam or oral exam (subject to the corresponding study regulation).



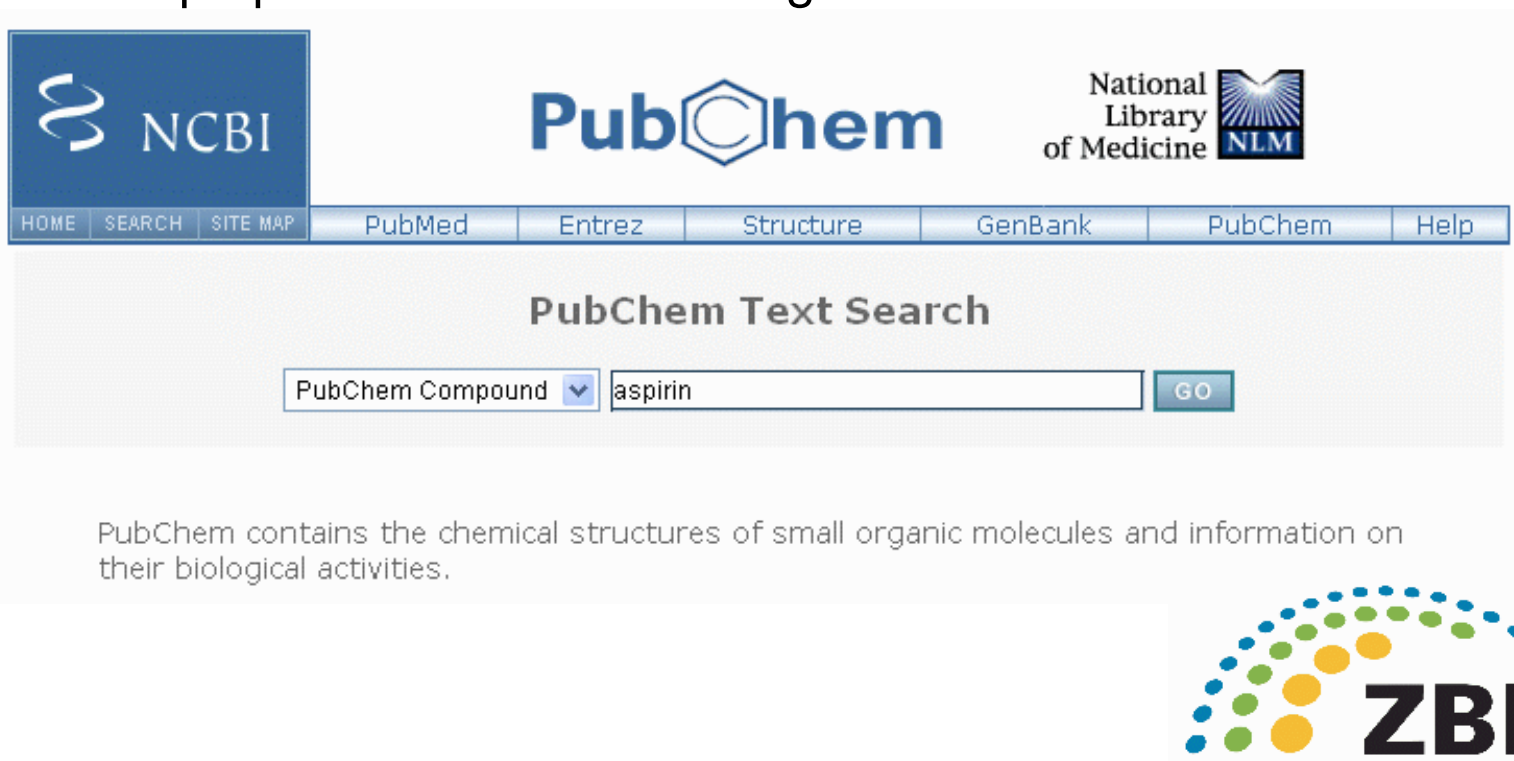
1st assignment (I)

Refer to a prescription medicine of your own choice

Write down the active ingredient

Try to find out its molecular structure:

<http://pubchem.ncbi.nlm.nih.gov/>



The screenshot displays the PubChem website interface. At the top left is the NCBI logo. In the center is the PubChem logo, and on the right is the National Library of Medicine (NLM) logo. Below these logos is a navigation bar with links for HOME, SEARCH, SITE MAP, PubMed, Entrez, Structure, GenBank, PubChem, and Help. The main content area features a "PubChem Text Search" section with a dropdown menu set to "PubChem Compound" and a search input field containing the text "aspirin". A "GO" button is positioned to the right of the search field. Below the search area, a paragraph states: "PubChem contains the chemical structures of small organic molecules and information on their biological activities." In the bottom right corner of the page, there is a logo for ZBI, consisting of a stylized arch of colored dots (blue, green, yellow) above the letters "ZBI".

1st assignment (II)



HOME SEARCH SITE MAP

PubMed

Entrez

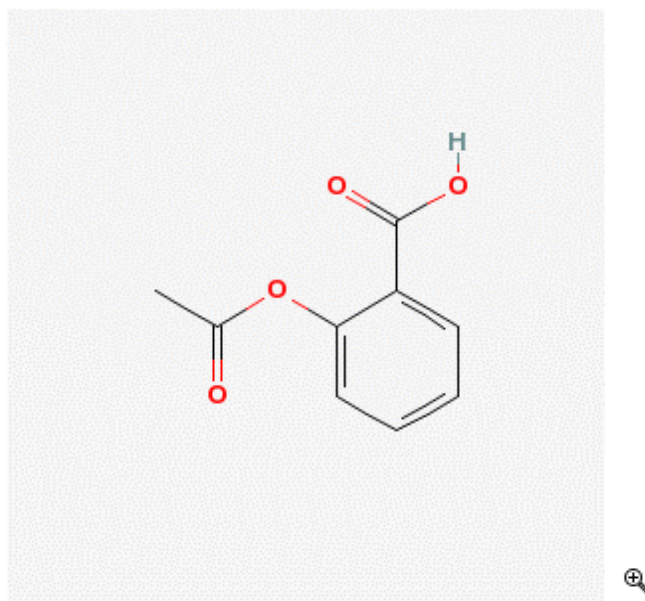
Structure

GenBank

PubChem

Help

Compound Summary:



CID: 2244 [?](#)



Substances: [?](#)

All: [52 Links](#)

Same: [10 Links](#)

Mixture: [42 Links](#)



BioActivity: [66 Links](#) [?](#)



Protein Structures: [2 Links](#) [?](#)



NLM Toxicology: [Link](#) [?](#)



Related Compounds: [?](#)

Same, Connectivity: [2 Links](#)



Similar Compounds: [30 Links](#) [?](#)



Structure Search [?](#)

1st assignment (III)



Medical Subject Annotations: (Total:11) [?](#)

Display: [Next 1](#) | [All](#)



Aspirin

The prototypical analgesic used in the treatment of mild to moderate pain. It has anti-inflammatory and antipyretic properties and acts as an inhibitor of cyclooxygenase which results in the inhibition of the biosynthesis of prostaglandins. Aspirin also inhibits platelet aggregation and is used in the prevention of arterial and venous thrombosis. (From Martindale, The Extra Pharmacopoeia, 30th ed, p5)

[Show MeSH Tree Structure](#)

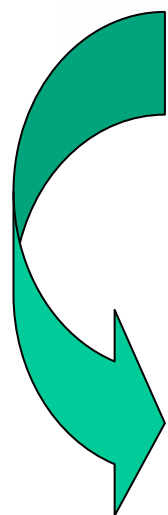
Pharmacological Action:

[Anti-Inflammatory Agents, Non-Steroidal](#)

[Fibrinolytic Agents](#)

[Platelet Aggregation Inhibitors](#)

[Cyclooxygenase Inhibitors](#)

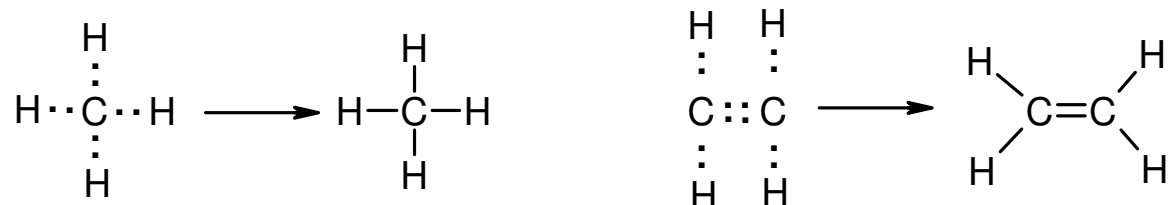


Explain why the medicine has a completely different name compared to the actual substance.

Try to find out some information about its actual *molecular target* (here: Cyclooxygenase) e.g. using Wikipedia

Representation of chemical structures (I)

The valence electrons of the atoms are pairwise grouped together

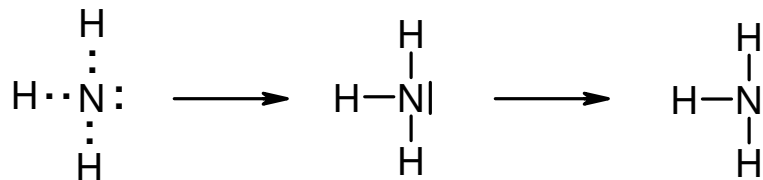


Lewis structures reflect covalent bonds between atoms in a molecule

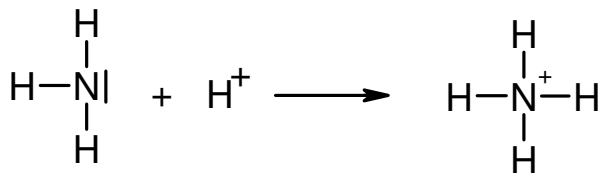


Representation of chemical structures (II)

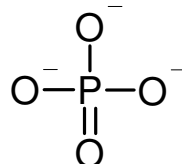
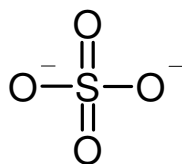
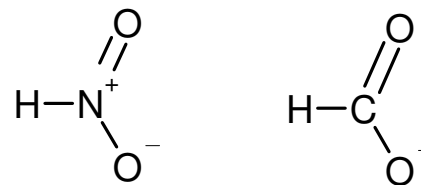
(electron) lone pairs are often not shown for clarity



octet rule and hypervalent atoms

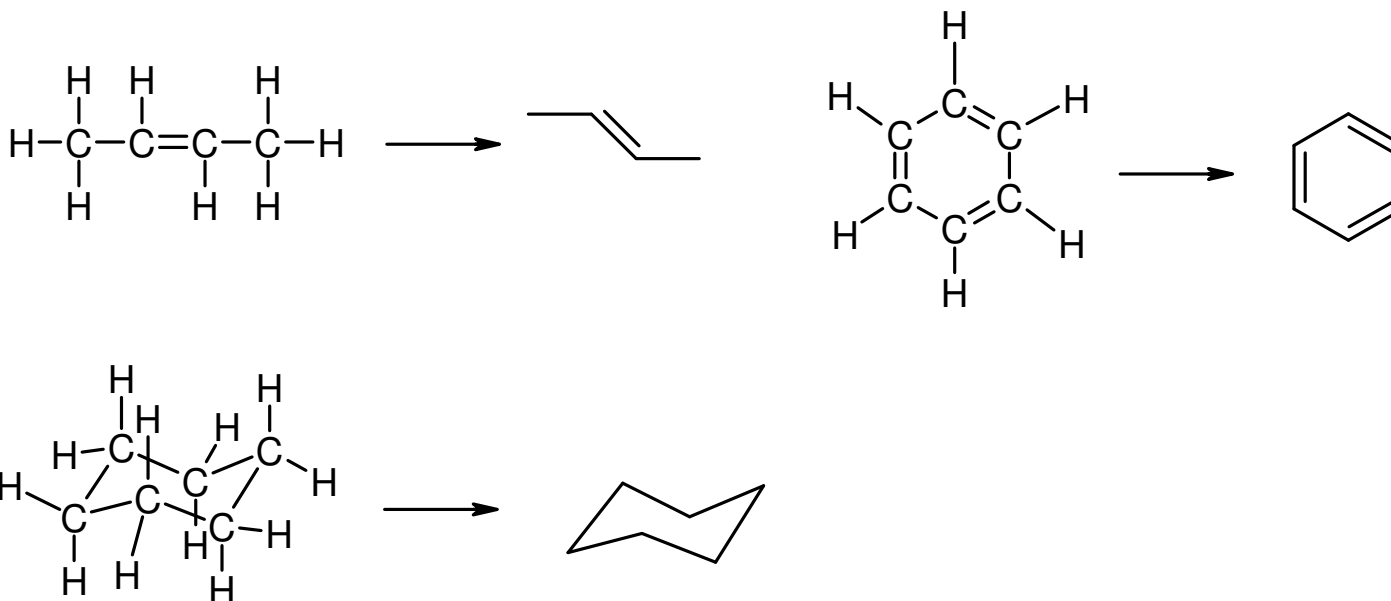


Equal bond lengths !



Representation of chemical structures (III)

Also carbon atoms are often omitted

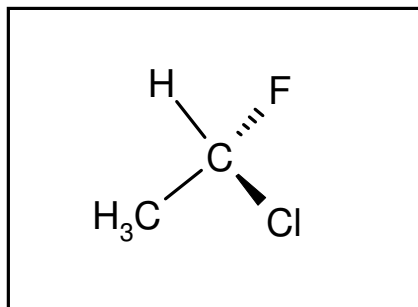


Corners and end of lines denote carbon atoms saturated with the appropriate number of hydrogen atoms



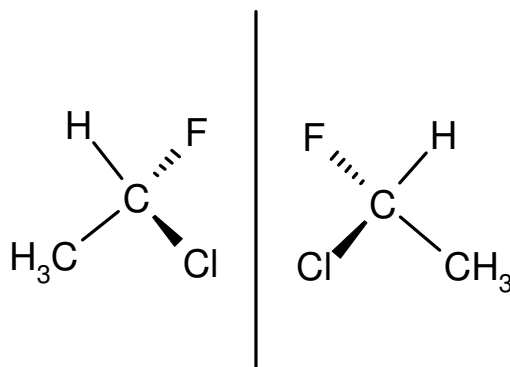
Representation of chemical structures (IV)

Stereochemistry



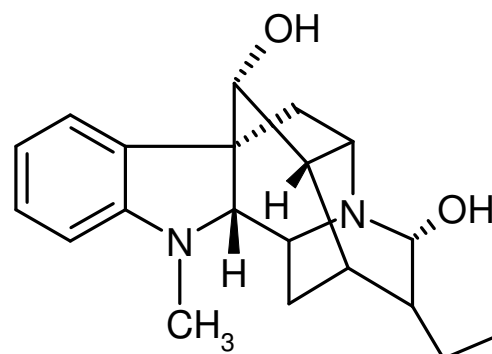
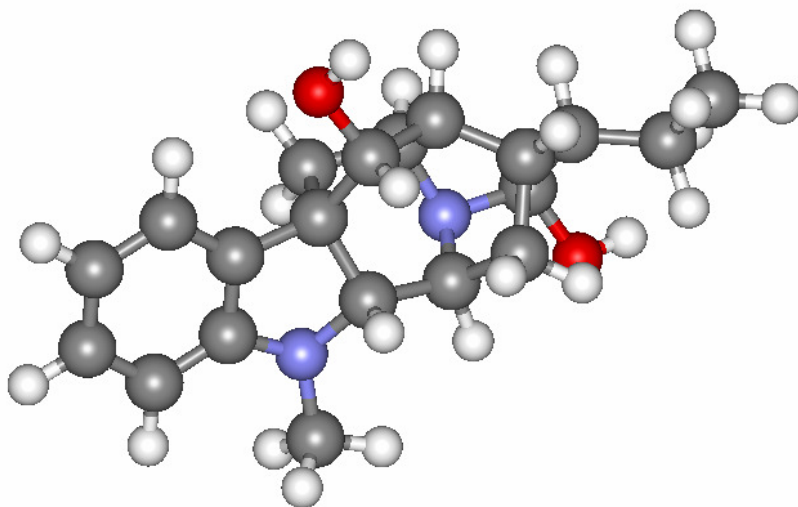
Solid wedges denote atoms in front of the plane, dashed wedges denote atoms behind

Four different substituents at a carbon atom cause chirality



Representation of chemical structures (V)

Particular for more complex molecules, these structural drawings provide more clarity than a picture of an actual 3D representation does.



Exercise: Construct this molecule using a molecular model set. Specify the chiral carbon atoms.



Bond distances and bond dissociation energies (I)

bond	distance [Å]	D_0 [kJ/mol] (homolytic cleavage)
H–H	0.742	432
C–H	1.09 ± 0.01	411 ± 7
C–C	1.54	345
C=C	1.34 - 1.40*	602 ± 21 *aromatic bond
C≡C	1.20	835
C–N	1.47	305
C=N	1.35	615
C≡N	1.16	887
C–O	1.43	358
C=O	1.20	526
C–Si	1.85	318
C–P	1.84	264
C–S	1.82	272
C=S	1.60	577 ± 21

longer
←

H							He
Li	Be	B	C	N	O	F	Ne
Na	Mg	Al	Si	P	S	Cl	Ar
K	Ca	Ga	Ge	As	Se	Br	Kr
Rb	Sr	In	Sn	Sb	Te	I	Xn
Cs	Ba	Tl	Pb	Bi	Po	At	Rn

↓ longer, weaker

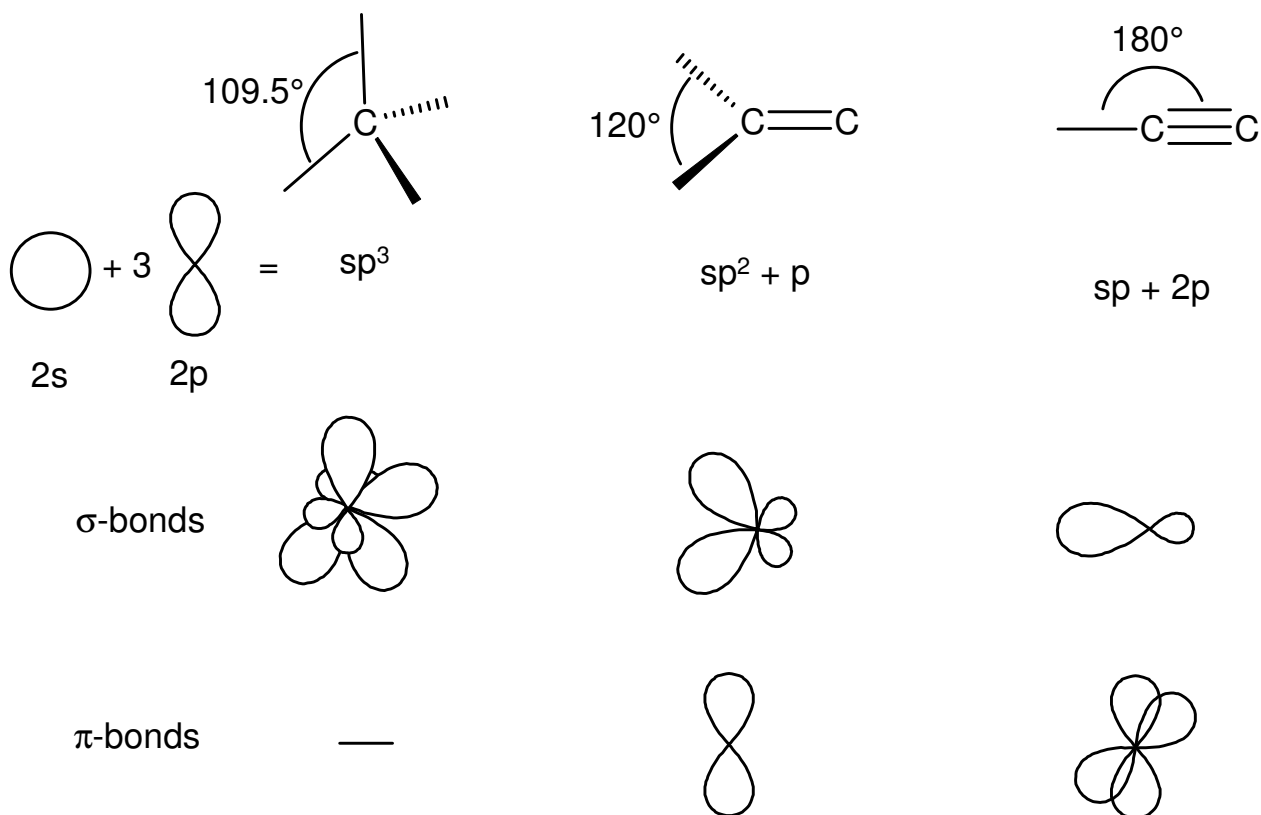
Adapted from: J.E.Huheey
Inorganic Chemistry, Wiley.

Bond distances and bond dissociation energies (II)

bond	distance [Å]	D_0 [kJ/mol]	
C–F	1.35	485	
C–Cl	1.77	327	
C–Br	1.94	285	
C–I	2.14	213	
C–H	1.09	411	non-polar hydrogen
O–H	0.96	459	} polar hydrogens, exchangable in polar solvents
N–H	1.01	386 ± 8	
S–H	1.34	363 ± 5	
N–N	1.45	247 ± 13	reason:
N=N	1.25	418	N, O, and S are more
N–O	1.40	201	electronegative than C;
N=O	1.21	607	heterolytic cleavage
P–O	1.63	≈ 335	that leads to ions
P=O	≈ 1.50	≈ 544	

Bond angles (I)

Strongly dependend on the hybridization



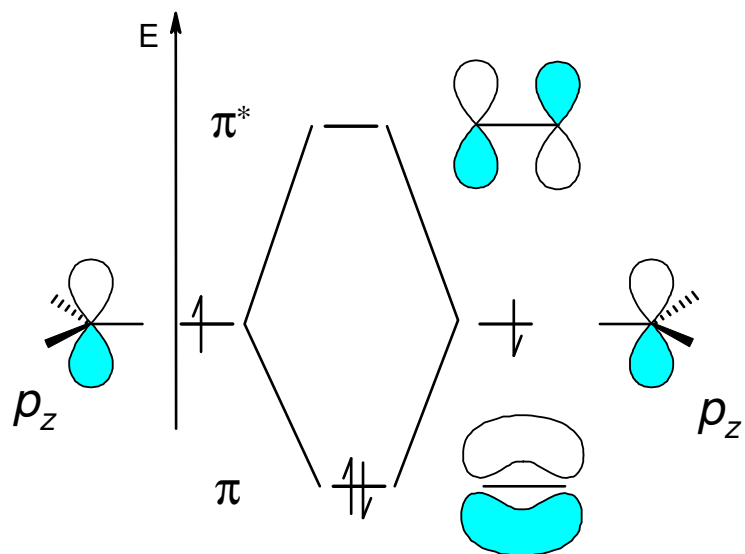
The C–C σ -bond is formed by overlap of the 1s orbitals

These are hybridized atomic orbitals. Do not confuse with molecular orbitals (=linear combination of atomic orbitals)

Molecular Orbitals

MO = linear combination of atomic orbitals (LCAO)

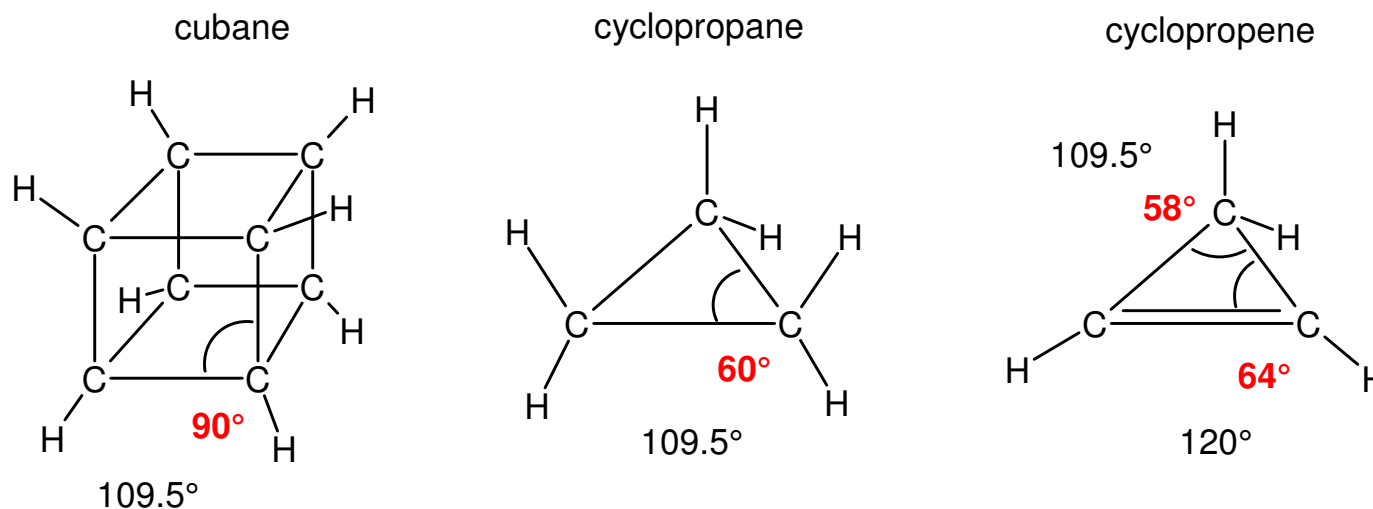
π -bond of ethylene $\text{H}_2\text{C}=\text{CH}_2$



The two combinations usually result in one bonding and one anti-bonding MO

Bond angles (II)

Extreme deviations from ideal bond angles

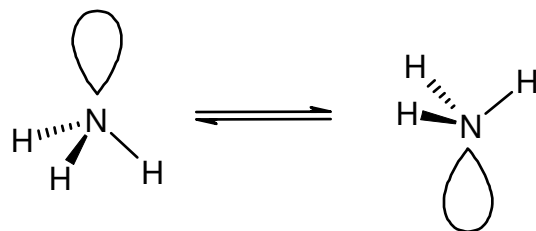


gives rise to strain energy in small rings

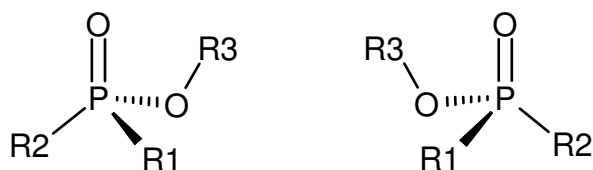
→ problems in force fields. More than one atom type for each hybridization state necessary.

Chiral atoms

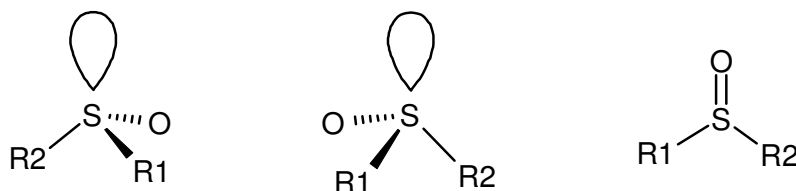
Further elements showing chirality/stereochemistry
(the lone electron pair behaves like a substituent)



Fast exchange at room temperature,
but slow at 77K



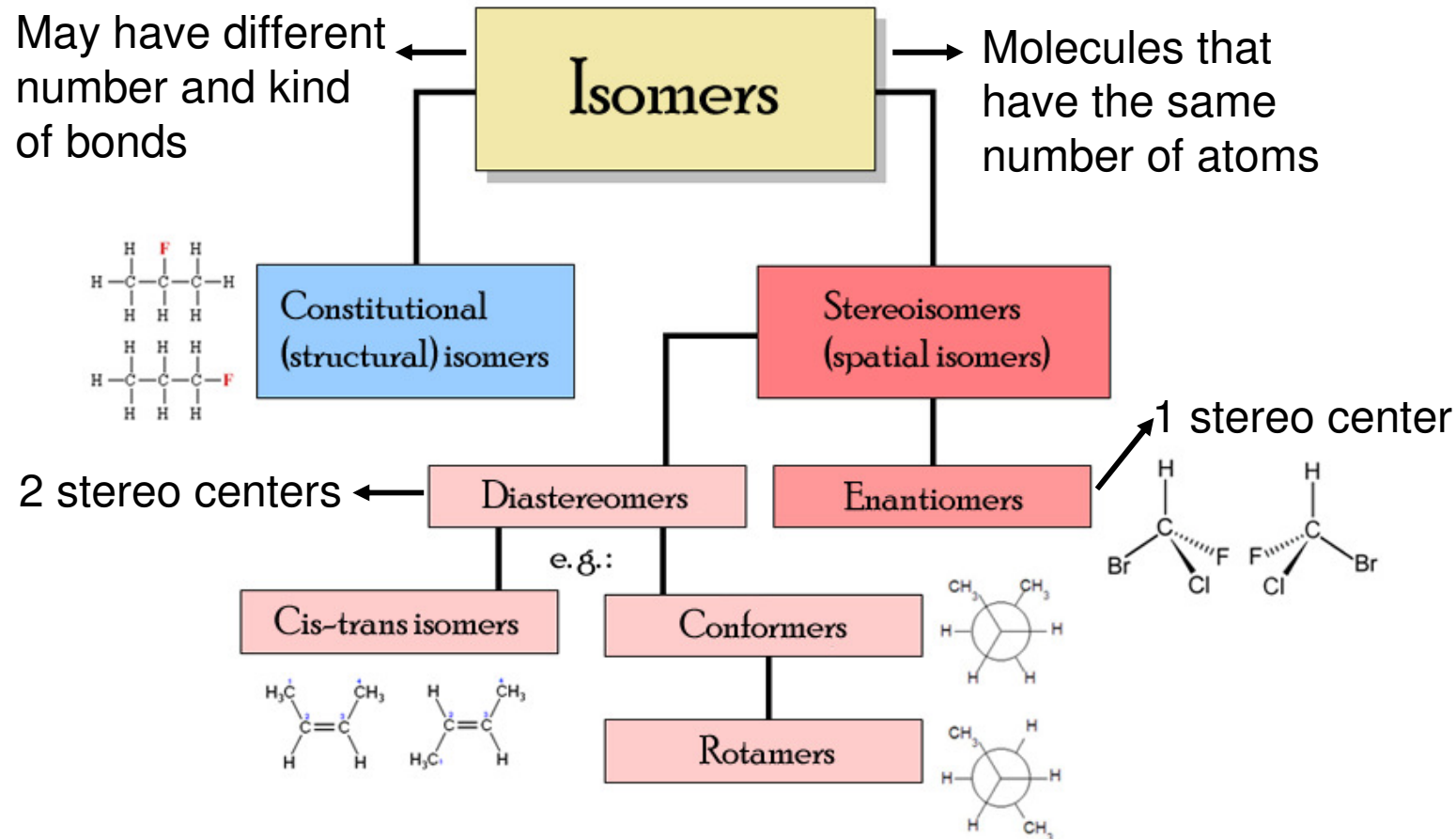
phosphorus inverts even slower



sulfoxides, sulfinic esters, etc
→ Different compounds at
room temperature

Furthermore: As, Si, ..., compounds with transition elements,
esp. octahedral and square planar metal complexes e.g. Pt

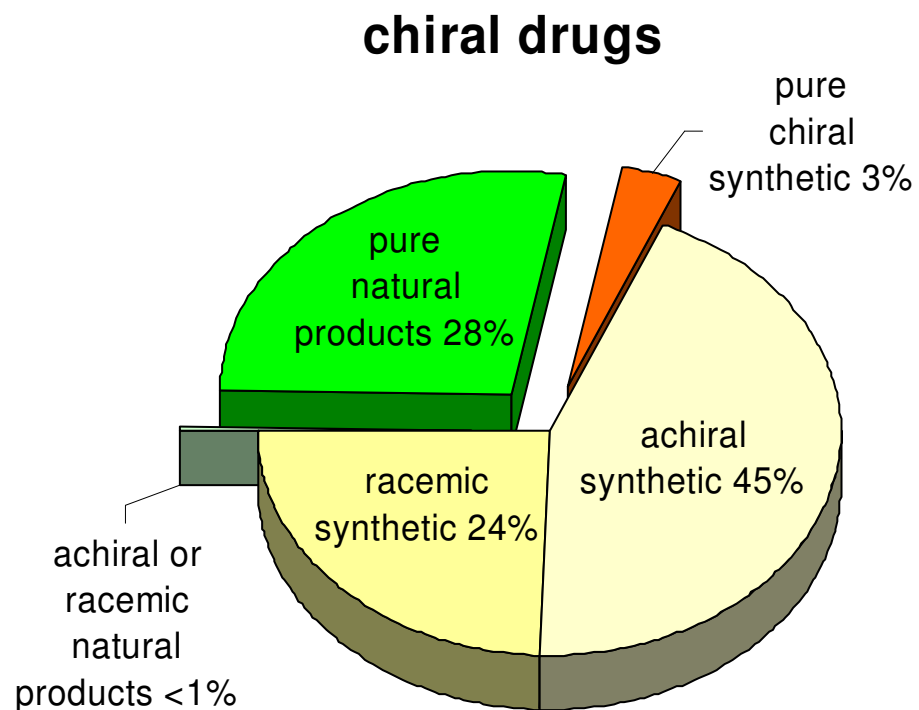
Isomers



Source: enhanced from wikipedia

Exercise: Which kind of computational method(s) allow(s) to calculate differences in energy between the respective isomers ?

Is stereochemistry important ?



Data from 1982: Böhm, Klebe & Kubinyi, *Wirkstoffdesign*