Properties of Drugs

What makes a chemical compound acting as pharmaceutically active agent?

• high affinity towards the target:

High binding constant (the drug should bind to the enzyme in concentrations as low as micro to nano molar)

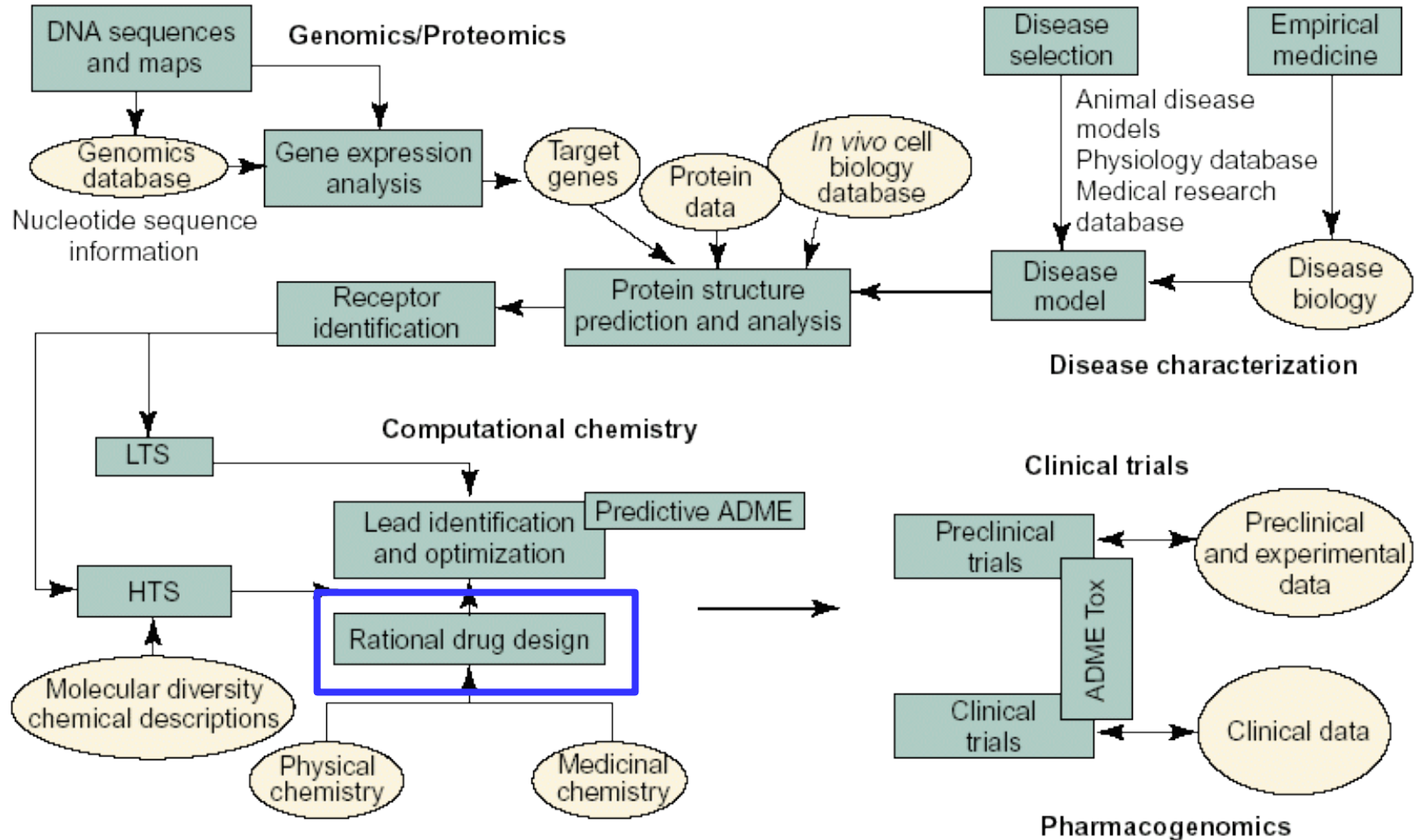• selectivity with respect to the target:

The drugs should bind preferably to the target, but not to other enzymes (so-called off-targets)

• high bioavailability und low toxicity:

Sufficient concentration in the body and a broad therapeutic range (dosage) along a minimum of adverse side effects

Flow of information in a drag discovery pipeline

Drug Discovery Today

Rational drug design

Basic principles:

- Improving the affinity
- Improving the selectivity
- Improving the bioavailability
- Reducing toxicity and adverse side effects

specificity

allows lower dosage

Frequently only possible by testing on animals and clinical trials

What are rational strategies? → create and test similar compounds

- systematic modification of the lead structure
- High Throughput Screening
- Combinatorial Synthesis
- bioisosteric exchange of fragments ⟶ lecture 4

Lit: H.Gohlke & G.Klebe, *Angew.Chem.* **114** (2002) 2764.
Angew.Chem.Int.Ed. **41** (2002) 2644.

Improving Affinity/Specificity (I)

How to increase the affinity of a molecule to its receptor?

Non-covalent interaction between the ligand and the receptor is an equilibrium process.

binding constant K_i (association constant of the complex)

$$K_i = \frac{[\text{ligand}] \cdot [\text{enzyme}]}{[\text{ligand-enzyme complex}]} \cdot \frac{[\text{mol/l}]\,[\text{mol/l}]}{[\text{mol/l}]}$$

dimension: K_i [mol/l = molar] ; e.g. $K_i = 10^{-9}$ M = 1 nM

The binding constant is associated with the change in free energy upon binding: $-\,RT\,\ln K = \Delta G = \Delta H - T\Delta S$

suitable values of K_i are in the range of 10^{-6} to 10^{-12} M
(micro to pico molar range).

This confers to values for ΔG of -4 to -17 kcal / mol
(≈ -17 to -71 kJ /mol)

Improving Affinity/Specificity (II)

The binding constant K_i can be determined experimentally by microcalorimetric measurements, such as isothermal titration calorimetry (ITC).
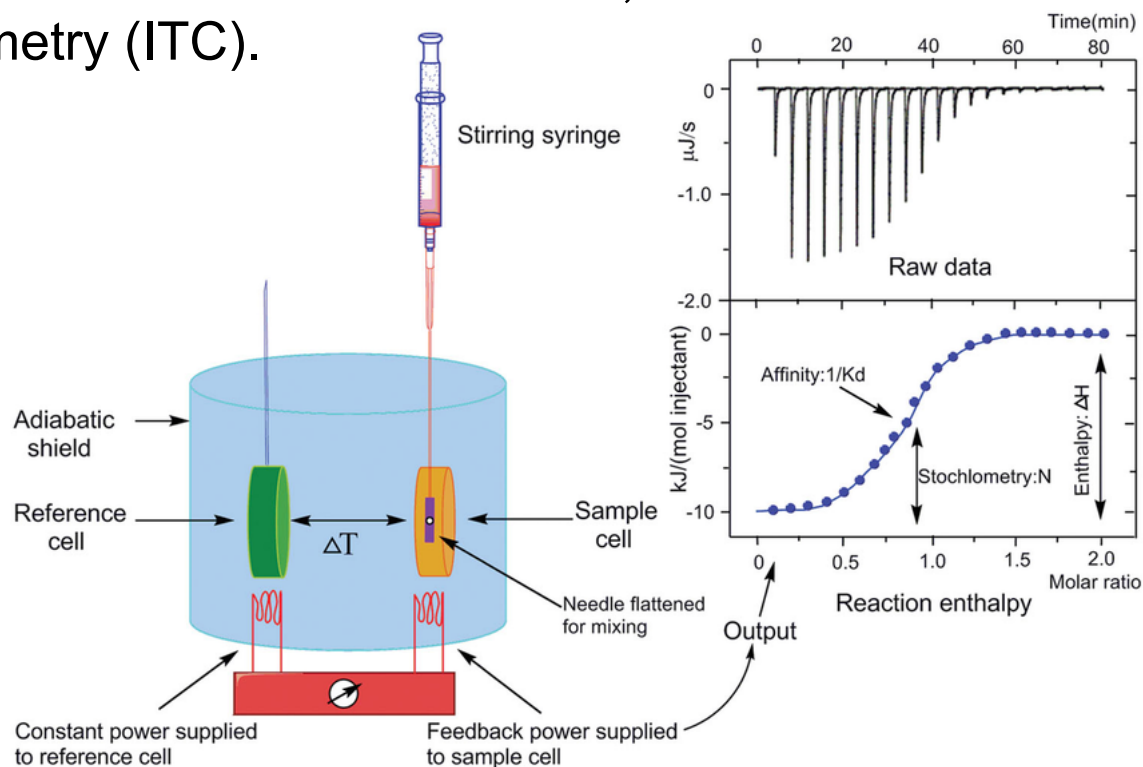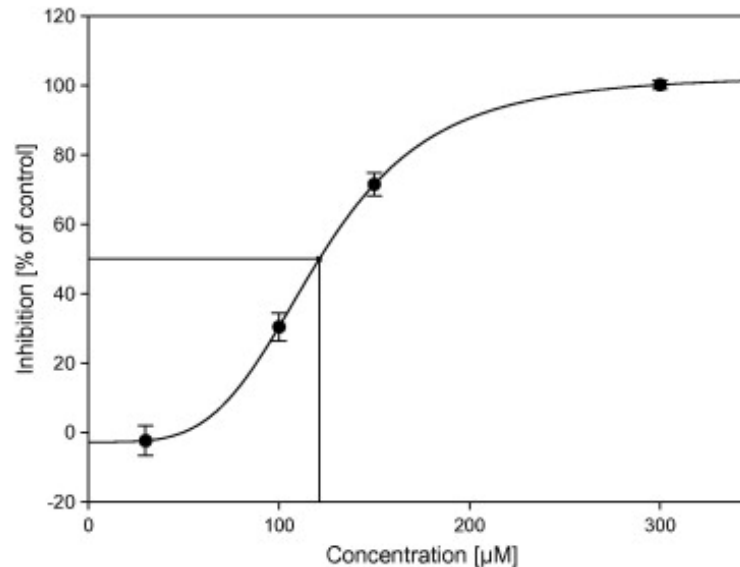

Furthermore the mechanism and kinetics of binding can be determined.

Picture source: www.researchgate.net

Improving Affinity/Specificity (III)

More often IC_{50} values are reported, which can be determined more easily.

IC_{50} : added amount or concentration of the ligand that produces an effect of 50%. e.g. reduces the enzymatic activity by 50%.

Testing of the enzymatic assay with different concentrations of the ligand and interpolation to 50%.

Picture source: www.researchgate.net

Improving Affinity/Specificity (III)

How to increase the affinity of a ligand to its receptor ?

Energy of binding ΔH must become more negative.

The energetic interactions between ligand and receptor have to become more favorable
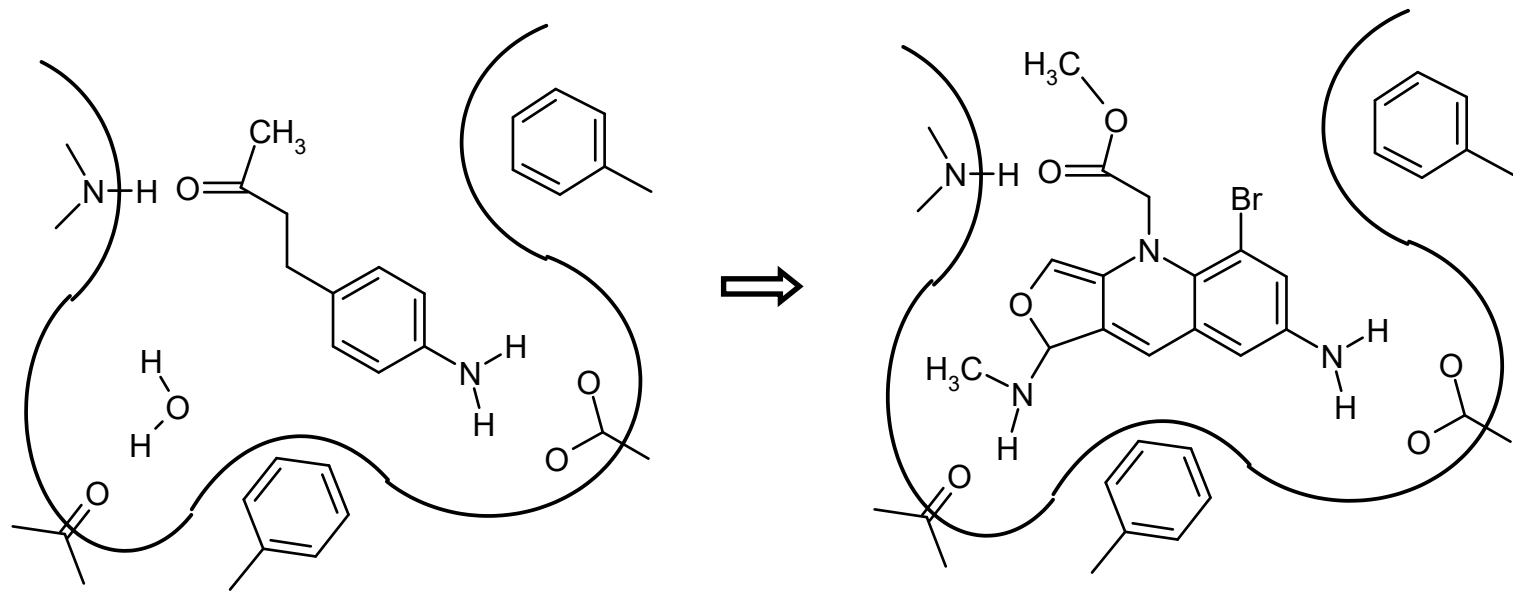
Improving Affinity/Specificity (IV)

The energy terms can be calculated according to force fields:

Interactions between ligand and receptor

$$E = E_{stretch} + E_{bend} + E_{tors} + E_{vdW} + E_{ES}$$

$$= \sum_{bonds\,(ij)} \frac{k^{(ij)}}{2}\left(r_{ij} - r_{0(ij)}\right)^2 + \sum_{angles\,(ijk)} \frac{k^{(ijk)}}{2}\left(\varphi_{ij} - \varphi_{0(ijk)}\right)^2$$

$$+ \sum_{torsions\,(ijkl)} \frac{k^{(ijkl)}}{2}\left(1 + \cos\left(n^{(ijkl)}\tau - \tau_{0(ijkl)}\right)\right)^2$$

$$+ \sum_{pairs\,(ij)} \left(\frac{A_{(ij)}}{r_{ij}^{12}} - \frac{B_{(ij)}}{r_{ij}^{6}}\right) + \frac{1}{4\pi\varepsilon\varepsilon_0} \sum_{pairs\,(ij)} \frac{q_i q_j}{r_{ij}}$$
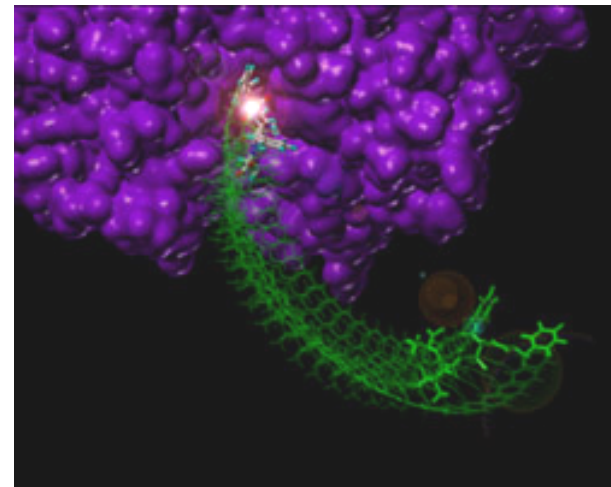
Most docking programs apply this concept.

Furthermore, a high resolution X-ray structure or an appropriate homology model of the target are necessary.

Enzyme-Ligand Interactions (I)

Which do exist ?

range in energy
upto:

electrostatic interactions: salt bridges ≈250 kcal/mol

coordinative binding of metals (complexes) ≈200 kcal/mol

hydrogen bonds: also to charged groups 1-8 kcal/mol (neutral)

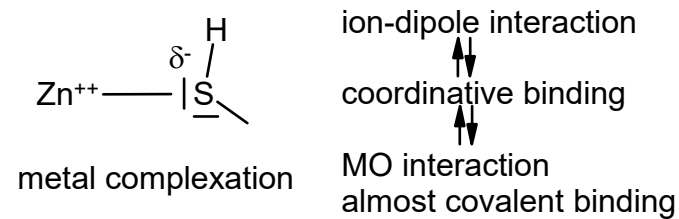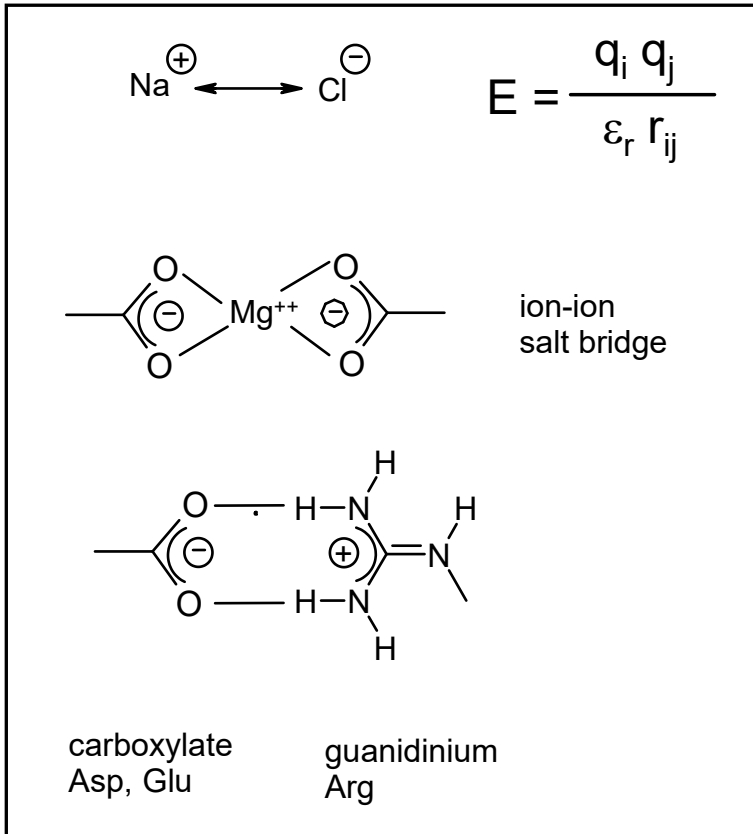van der Waals interactions 0.5 kcal/mol (per atom pair)
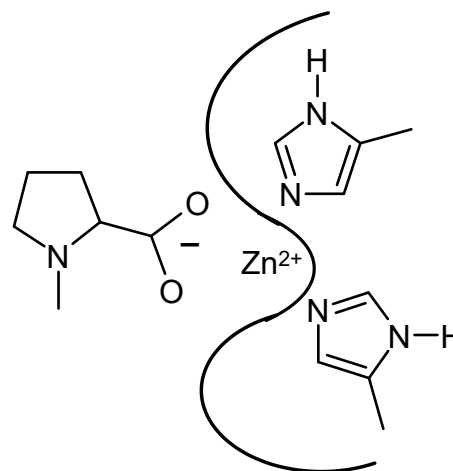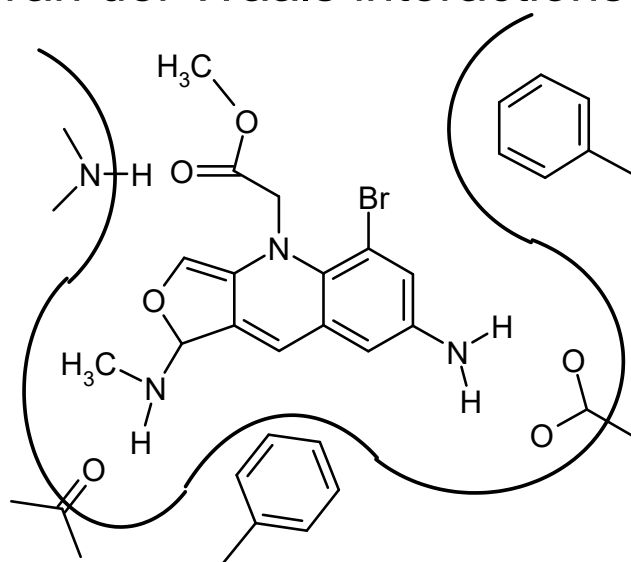
Enzyme-Ligand Interactions (II)

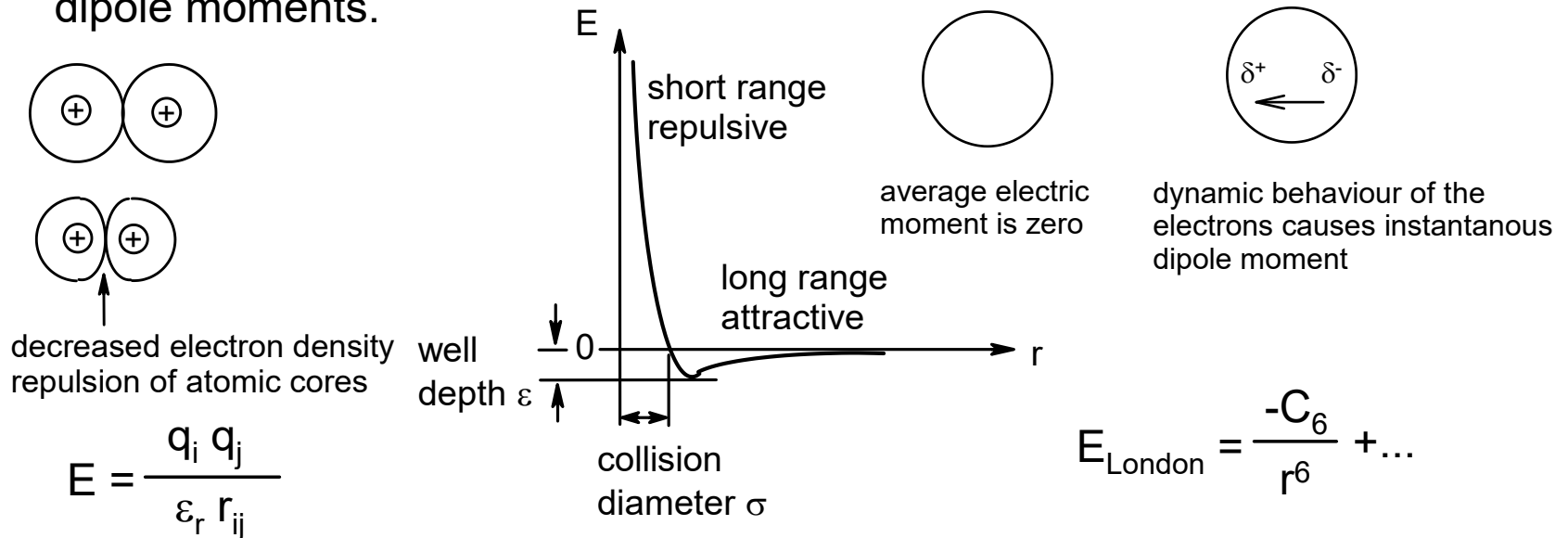## Strong and medium electrostatic interactions (static)

$$E = \frac{q_i\, q_j}{\varepsilon_r\, r_{ij}}$$

ion-ion
salt bridge

metal complexation

ion-dipole interaction

coordinative binding

MO interaction
almost covalent binding

carboxylate
Asp, Glu

guanidinium
Arg

ion-dipole
ionic hydrogen bond

dipole-dipole
hydrogen bond
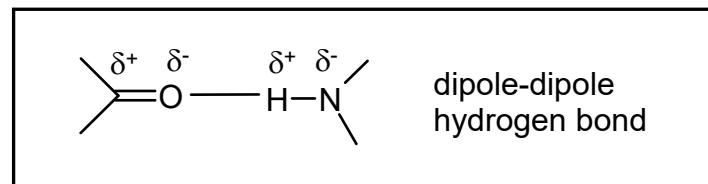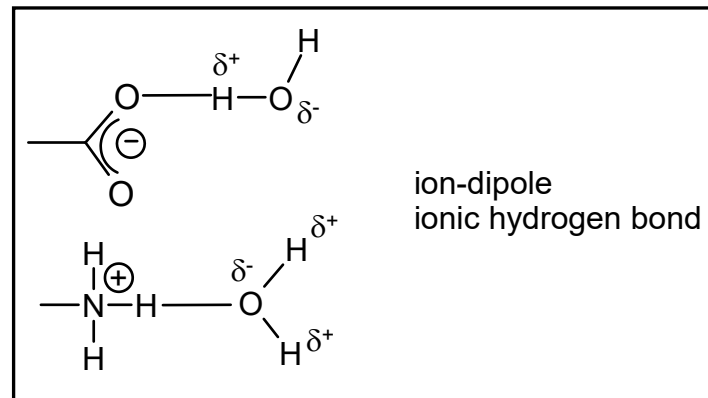
weaker

Enzyme-Ligand Interactions (III)

weak electrostatic interactions (induced, dynamic)

ion-quadrupole
cation-π

quadrupole-quadrupole
π-π stacking

octupole-quadrupole

octupole-octupole

weaker

Hydrophobic
interactions
van der Waals
(vdW)

delocalized
π-system

localized
π-bond

σ-bond

Lower polarizability
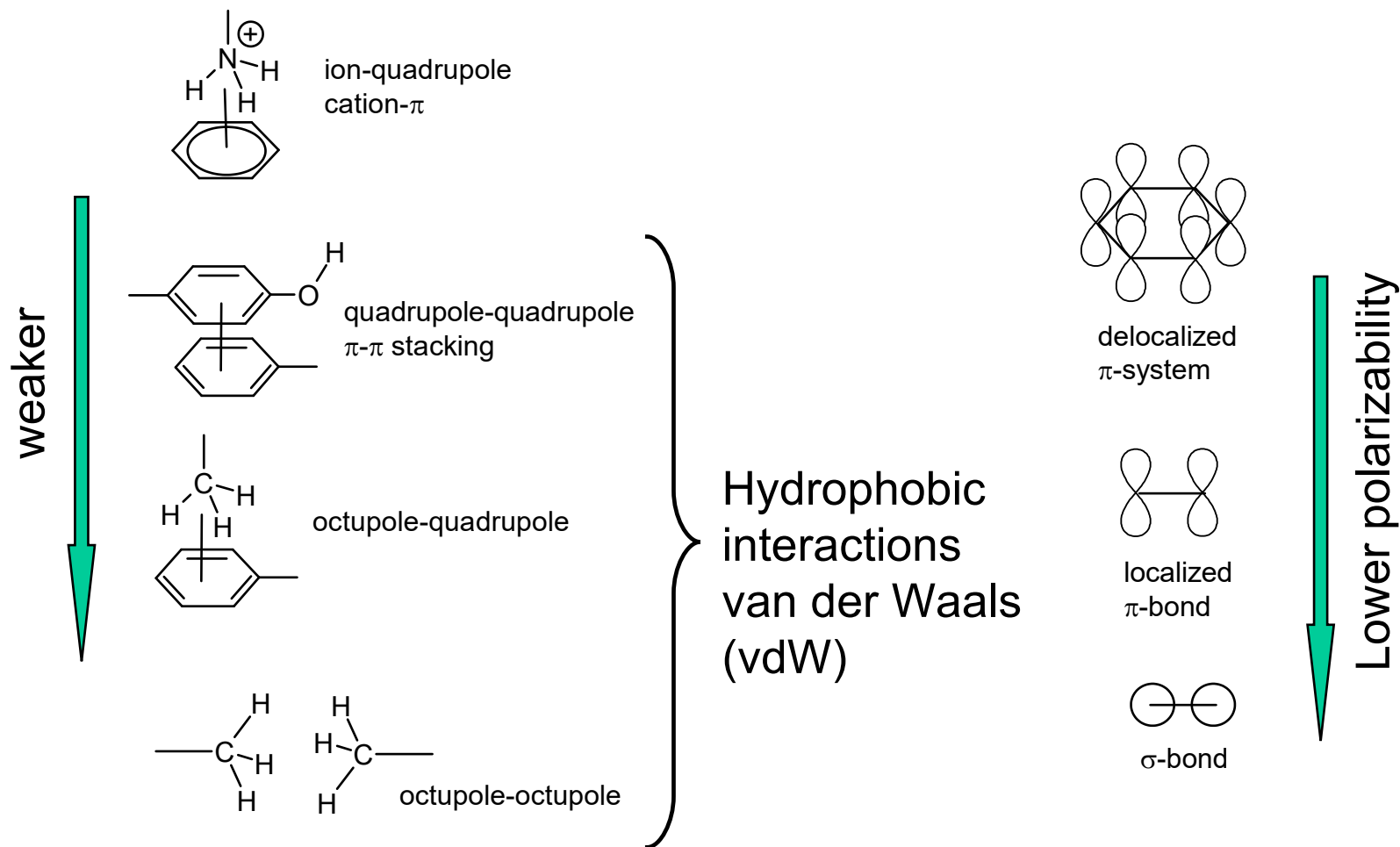
Enzyme-Ligand Interactions (IV)

Dispersive interactions: London forces and van der Waals

The attractive force is due to instantaneous dipols which arise
from fluctuations in the electron clouds. These induce mutual
dipole moments.

E

short range
repulsive

average electric
moment is zero

δ^+ δ^-

dynamic behaviour of the
electrons causes instantanous
dipole moment

long range
attractive

decreased electron density
repulsion of atomic cores

well
depth ε

0

r

collision
diameter σ

$$E = \frac{q_i \, q_j}{\varepsilon_r \, r_{ij}}$$

$$E_{London} = \frac{-C_6}{r^6} + ...$$

Lennard-Jones potential

$$E_{vdW} = 4\,\varepsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right)$$

Enzyme-Ligand Interactions (V)

Hydrophobic Interactions are characterized by the absence of polar hydrogens and low differences in electronegativity between the atoms.

Examples of non-polar groups:

| methyl
Me | isopropyl
ipr | tert-butyl
t-butyl | cyclohexyl | phenyl | naphthyl |
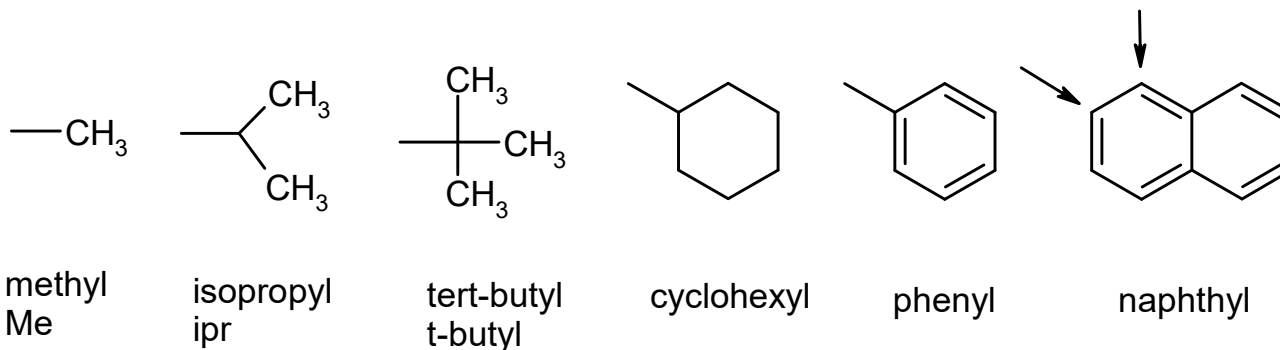
Examples of non-polar substituents:

$(-F)$ $-Cl$ $-Br$ $-I$

electronegativity (electron pulling)

polarizability (larger electron cloud)
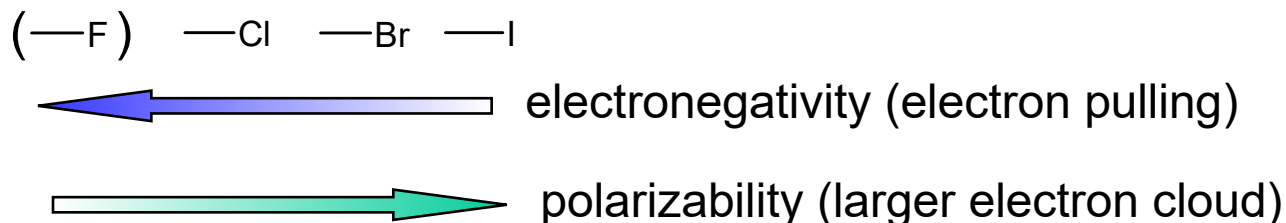
Electronegativity (EN)

The EN is a measure of the ability of an atom (or group) to attract electrons in the context of a chemical bond.

Concepts and definitions (not comprehensive!)

R.S. Mulliken: $EN = \dfrac{E_{\text{Ionization}} + E_{\text{ElectronAffinity}}}{2}$

L. Pauling: using the bond dissociation energies D of diatomic molecules containing the elements A and B

$$D_{AB} - \sqrt{D_{AA} - D_{BB}} = 96.48\,\frac{kJ}{mol}\cdot\left(EN_A - EN_B\right)^2$$

Element	H	C	N	O	F	Cl	Br	I	Si	P	S
Mulliken	2.2	2.5	2.9	3.5	3.9	3.3	2.7	2.2	1.7	2.1	2.4
Pauling	2.2	2.5	3.0	3.4	4.0	3.2	3.0	2.7	1.9	2.2	2.6

Improving Affinity/Specifity (V)

Favorable intermolecular interactions lower the energy of binding:
Many side chains of amino acids can change their protonation state,
depending on the local environment and pH! (which ones?)

• hydrogen bonds   1-8 kcal mol^{-1} (average ≈3 kcal mol^{-1})

electrostatic interactions
• salt bridges   up to 250 kcal mol^{-1}

coordinative binding of metals (complexes)

• van der Waals   max. 0.5 kcal mol^{-1} per atom pair
burying of hydrophobic fragments

Sign of interaction energies : positiv = repulsive; negative = attractive

Improving Affinity/Specificity (VI)

enzyme-ligand interactions that are energetically unfavorable upon binding:

Burying of polar or charged fragments (amino acid side chains) up to 7 kcal mol^{-1}. Reason:

Transition from a medium of high dielectric constant (physiological solution ≈78) into an environment of much lower ε (hydrophobic pocket $\varepsilon \approx$ 2-20)

Desolvation:
Displacement of water molecules involved in hydrogen-bonds from the binding pocket. Breaking of H-bonds and formation of an empty cavity which allows the ligand to enter. However, these water molecules can now form more possible H-bonds in the solvent and therefore gain entropy. $\Delta G = \Delta H - T \cdot \Delta S$

Improving Affinity/Specificity (VII)

Entropically (ΔS) unfavorable during binding are :

• Loss of all translational degrees of freedom (x,y,z direction)

• Loss of rotational degrees of freedom
about 1 kcal mol^{-1} per rotatable bond (single bonds) between two non-hydrogen atoms.
Terminal $-CH_3$ bonds are not considered, because these rotate even in the complex.
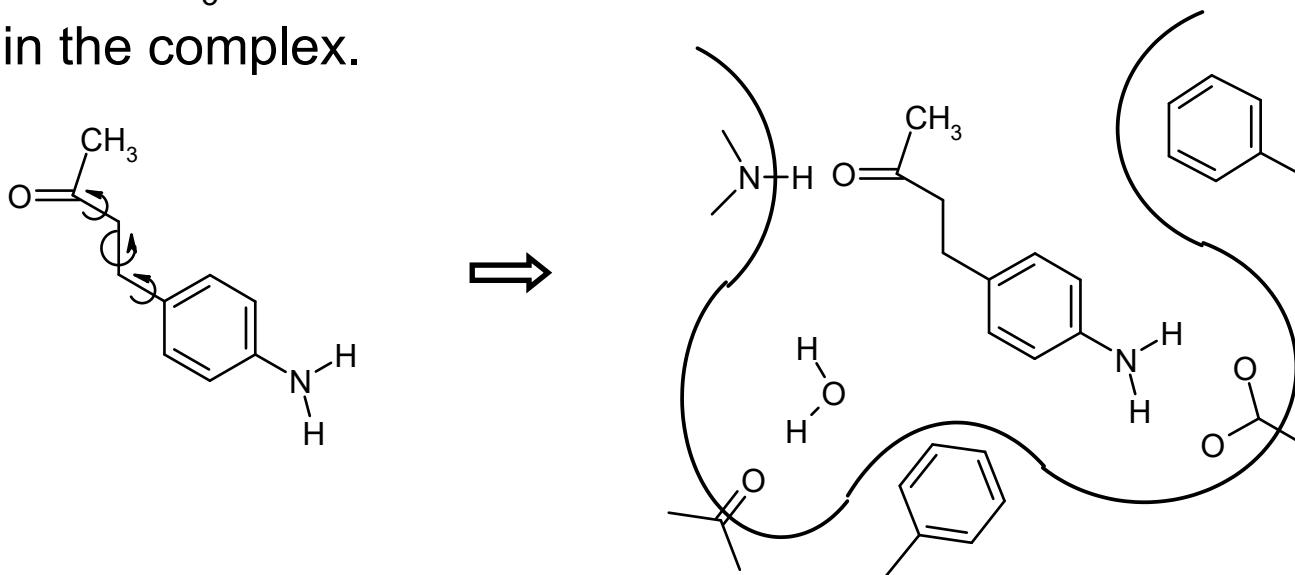
Improving Affinity/Specificity (VIII)

Entropic (ΔS) considerations:

Displaced water molecules can form usually more hydrogen bonds (with other waters) outside the binding pocket. Likewise the dynamic exchange of H-bonds is simplified in bulk solution.

Thus: The ligand should fit more precisely and thoroughly into the binding pocket.
Simultaneously, the selectivity is improved (ligand fits only in one special binding pocket)
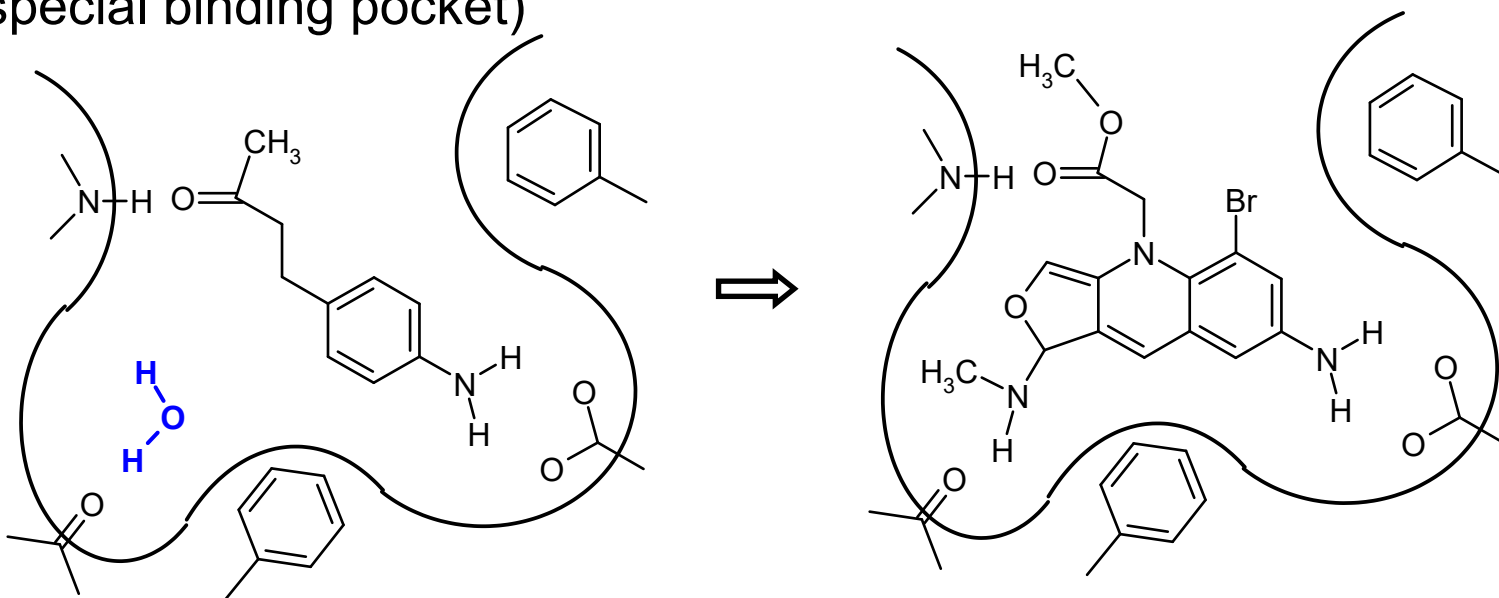
Improving Affinity/Specificity (IX)

Experience in *rational drug design* shows:

• binding pockets are predominately hydrophobic, so are the ligands

• hydrogen-bonds are important for selectivity

• energy - entropy compensation:

Adding one OH-group to the ligand in order to form an additional H-bond in the binding pocket will lead to displacement of a water molecule, but this water will be solvated in the surrounding bulk water. Thus no additional H-bonding energy is gained.

Therewith, all possibilities of *ligand design* by *docking* are exploited.

Remaining are issues related to bioavailability, toxicology, and adverse effects.

Bioavailablity & ADME prediction

Absorption

Distribution

Metabolism

Elimination

Pharmacokinetic

Bioavailability

mouth
pharynx
tongue
esophagus
pancreas
liver
stomach
gallbladder
transverse colon
descending colon
duodenum
jejunum
ascending colon
sigmoid colon
ileum
rectum
anus
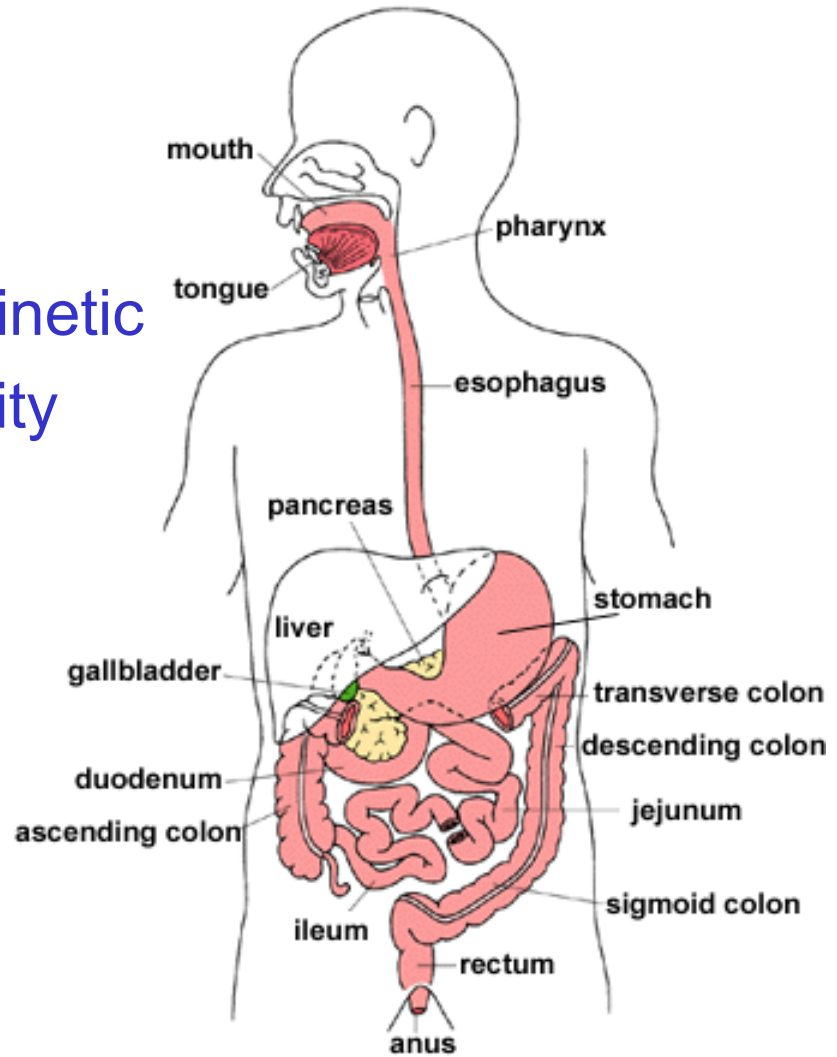
Why is AMDE prediction so important ?
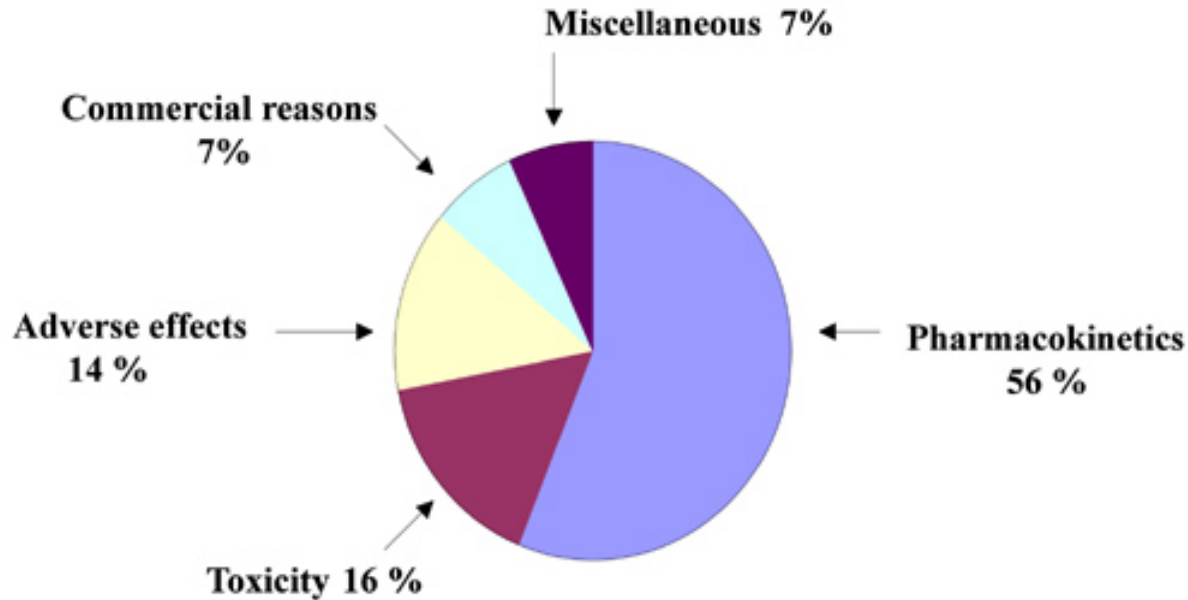
Reasons that lead to the failure of a potential drug
as of the mid 1990's

In silico ADME filter

Project virtual library of 100,000 members

Clustering	6000 Subset
Solubility model	3000
Absorption model	1450
Metabolism model	1150
BBB model	700
	500

Selected for synthesis and screening
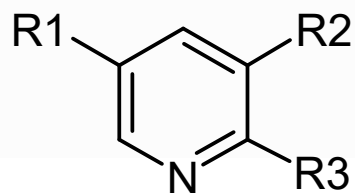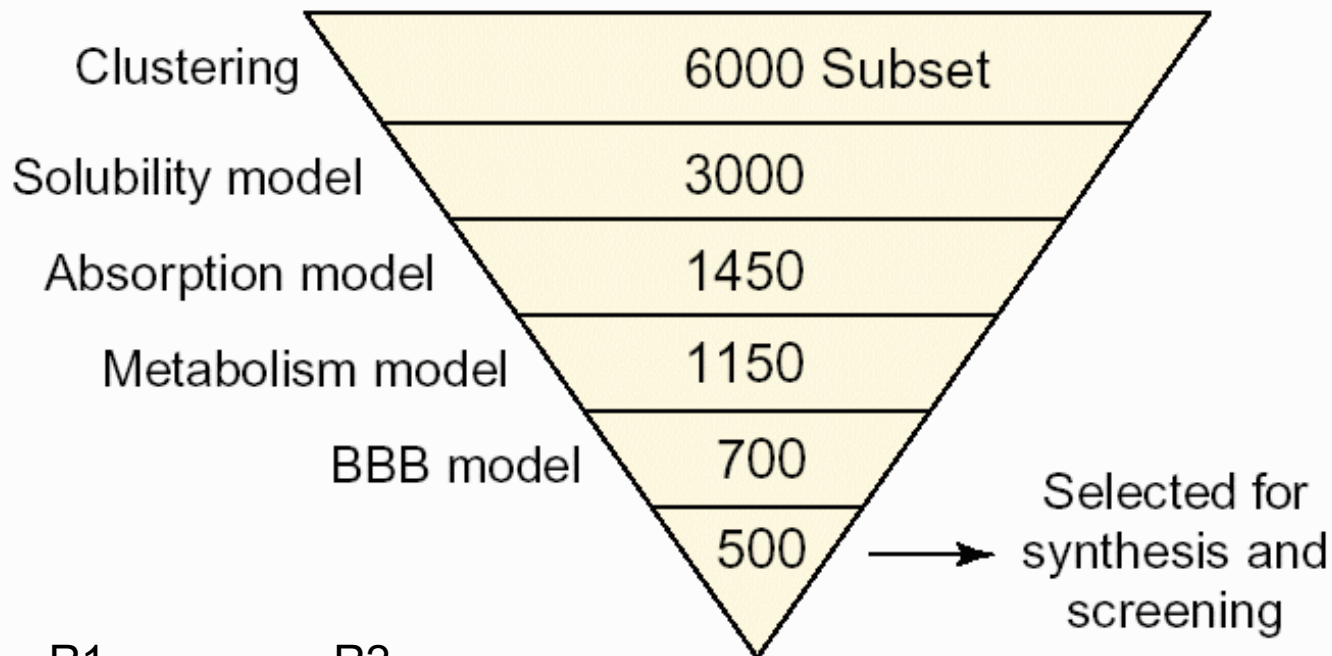
Drug Discovery Today

More about ADME-models in lecture 7

Which physico-chemical properties are recommended for drugs ?

Solubility and absorption: A hardly soluble compound is hardly transfered into the systemic blood flow.

C. Lipinski's rule of five: Orally administered substances,
                                                  but not neccessarly *drug-like*

Molecular weight < 500
logP < 5
H-bond donors (N-H, O-H) < 5
H-bond acceptors (N, O) < 10

Influence on the membrane passage

Less than 8 rotatable bonds

polar surface area < 140 Å^2

not part of the original rule

→ *drug-like* compounds

From the lead compound to the drug (I)

Therapeutic Target

Lead Discovery

Lead Optimization

Clinical Candidate

drug design

Commerical Drug
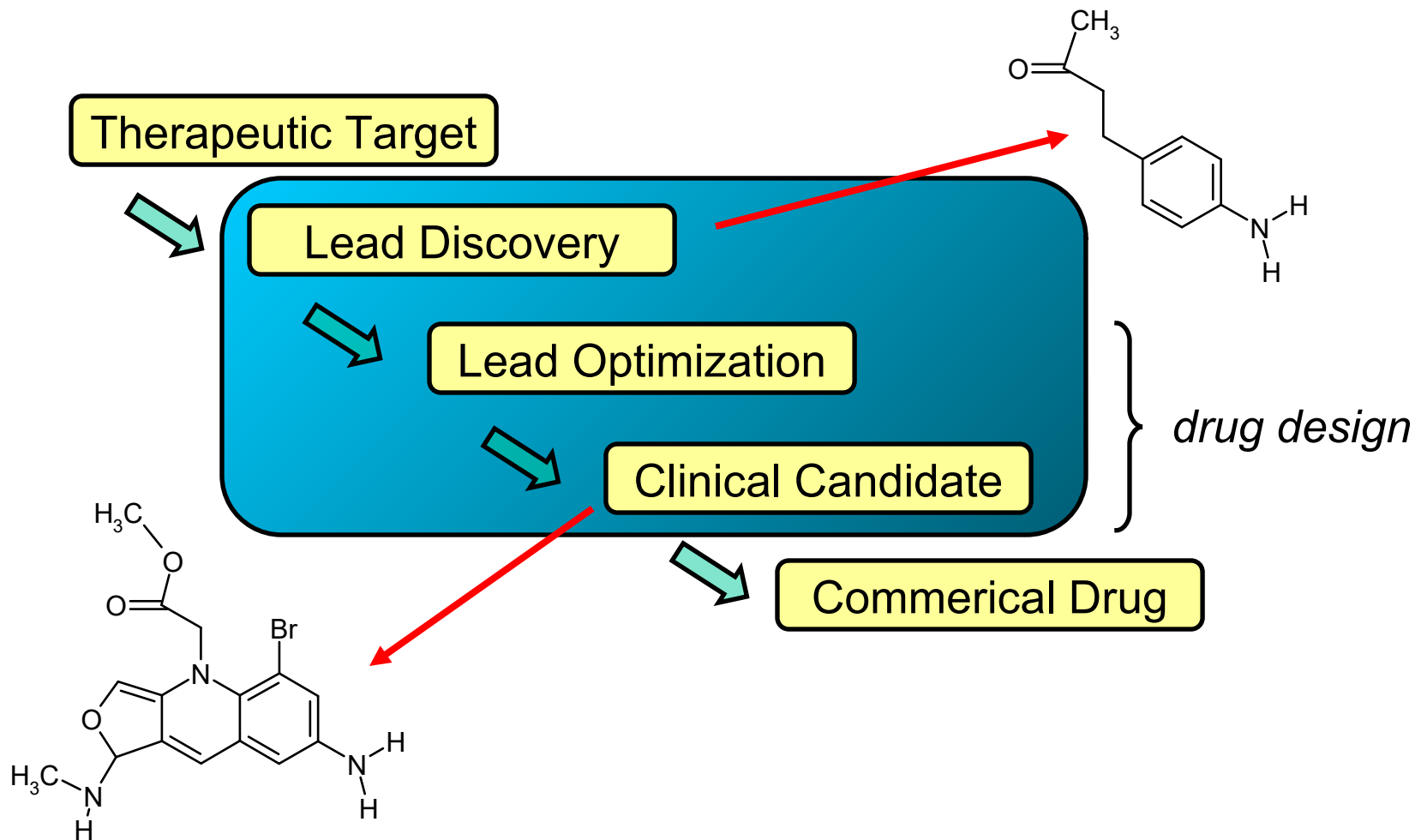
From the lead compound to the drug (II)

During the optimization from the lead compound to the clinical candidate, molecules are usually becoming larger and more lipophilic (binding pocket is filled better).

Thus, following properties are desirable for *lead-like* compounds:

- molecular weight < 250
- low lipophilicity (logP<3) for oral administration
- enough possibilities for side chains
- already sufficient affinity and selectivity

MW 164
logP 1.84

MW 366
logP 2.58
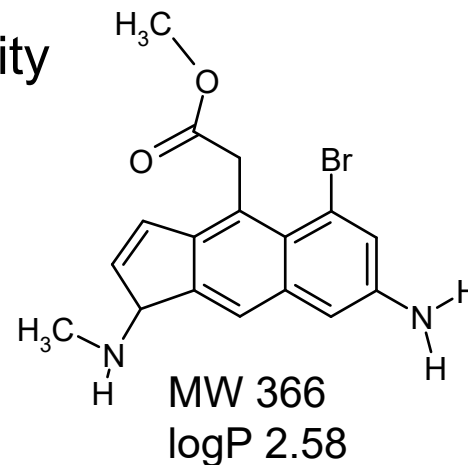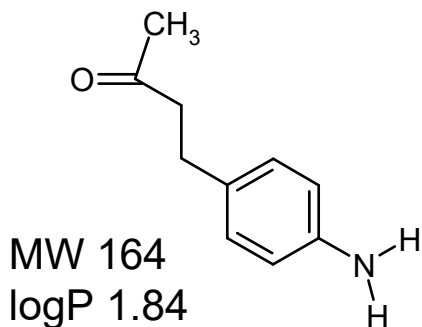
More about substance libraries in lecture 4

What makes a compound *drug-like* ?

„typical" pharmaceutic compounds show following properties:

• Molecular weight in the range of 160 < MW < 480

• Number of atoms between 20 and 70

• lipophily in the range of –0.4 < logP < +5.6

• Molar refractivity in the range of 40 < MR < 130

• few H-bond donors (< 5)

• few H-bond acceptors (< 10)

• At least one OH-group (exception: CNS-active substances)

But: numerous exceptions! (up to 50% of actual drugs)

Lit: A.K.Ghose et al. *J.Comb.Chem.* **1** (1999) 55.

More about *in silico* drug/non-drug prediction in later lectures

From the lead compound to the drug (III)

Example: Inhibitors of the Angiotensin Converting Enzyme

$$\text{Angiotensin I} \xrightarrow{\text{ACE}} \text{Angiotensin II} \quad + \text{HL}$$

DRVYIHPFHL   DRVYIHPF
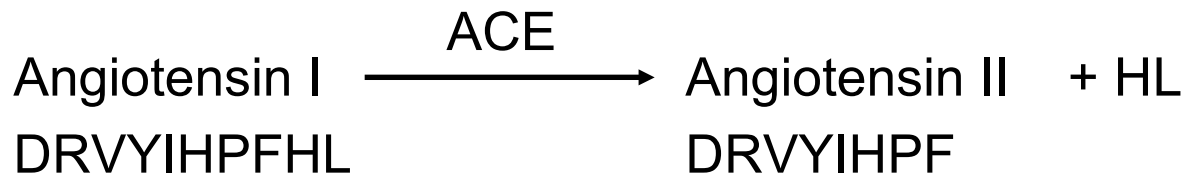
Lead compound: Phe-Ala-Pro
K_i in μM range
Discovered in snake venoms that lead to
drastic drop of the blood pressure

Captopril (1977) X-Ala-Pro
IC_{50} = 23 nM; K_i = 1.7 nM
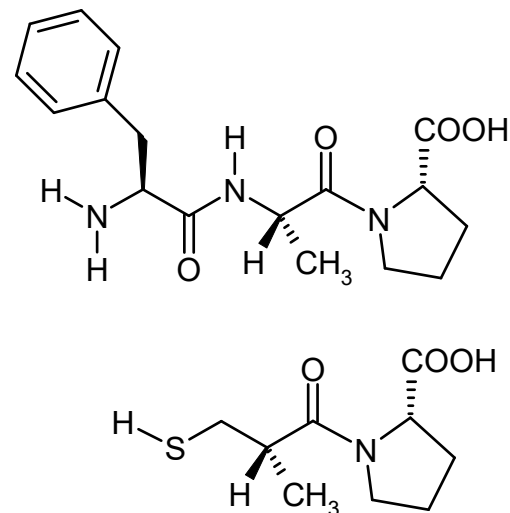
From the lead compound to the drug (IV)

The somatic ACE (sACE) is a membrane bound protein.
The X-Ray structure of the N-terminal domain (2C6F.pdb) is known since 2006.

Germinal ACE (tACE) which is soluble shows a high sequence similarity and was used in modified form for crystallization with known inhibitors. Furthermore, structure-based design of new inhibitors is possible as the shape of the binding pocket around the catalytic zinc-ion is known.
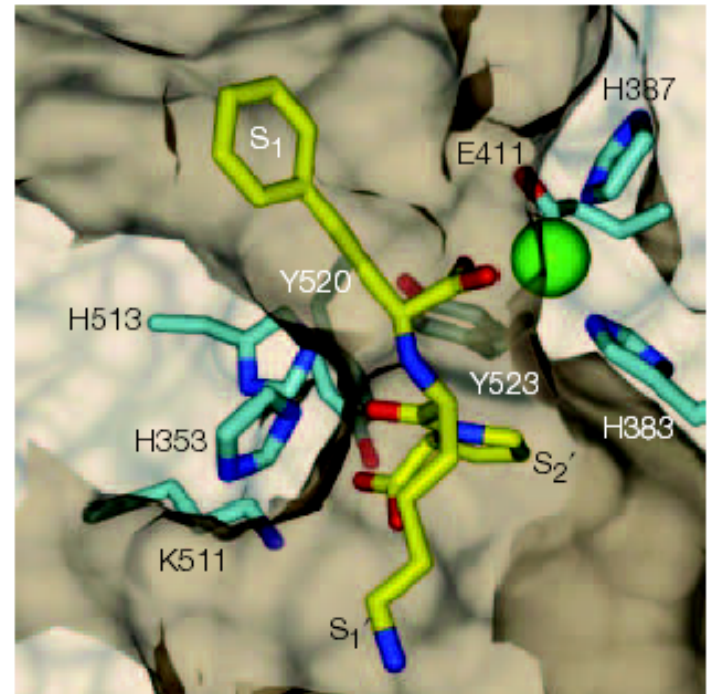
Lit: K.R.Acharya *Nature Rev. Drug Discov.* **2** (2003) 891.

From the lead compounds to the drug (V)

X-Ray structures of tACE (selection)

inhibitor (patent as of year)

1UZF.pdb   Captopril (1977)

1O86.pdb   Lisinopril (1980)

1UZE.pdb   Enalapril (1980)
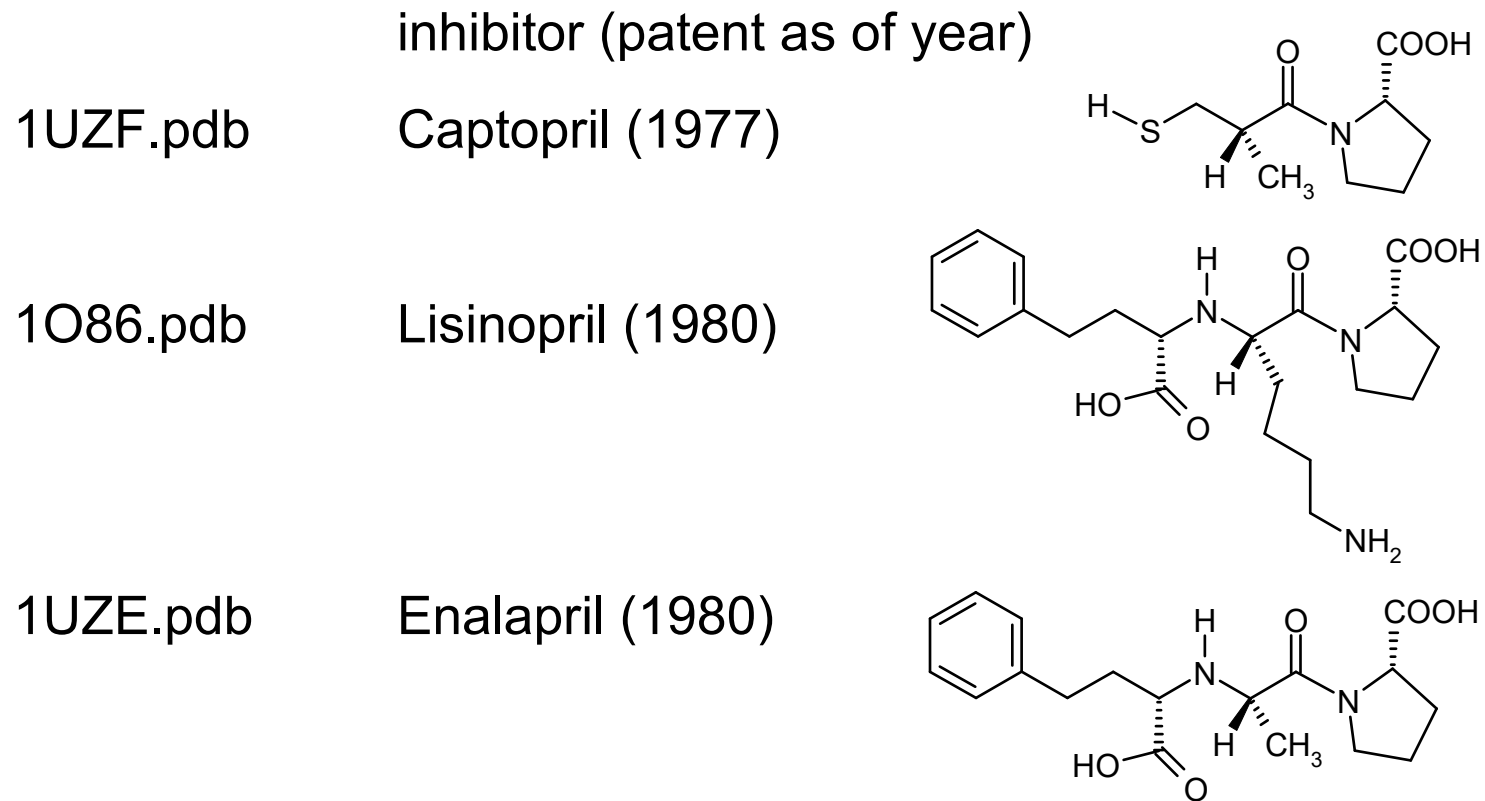
From the lead compound to the drug (VI)

More recent ACE-Inhibitors (patent as of year)

Trandolapril (1980)
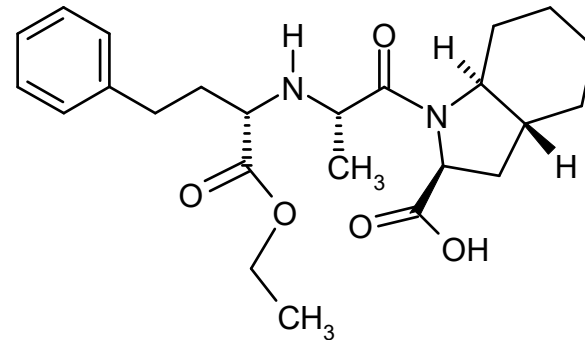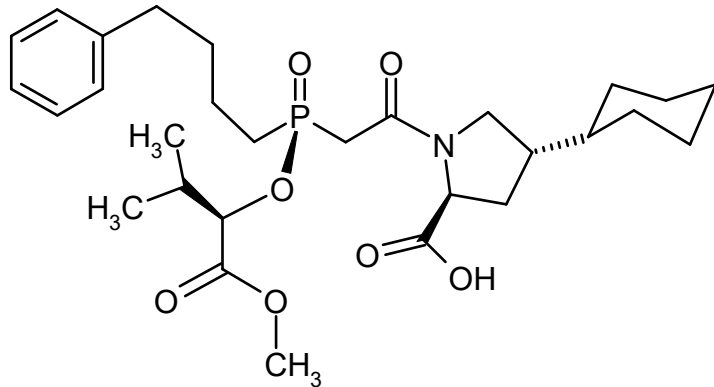
Fosinopril (1982)
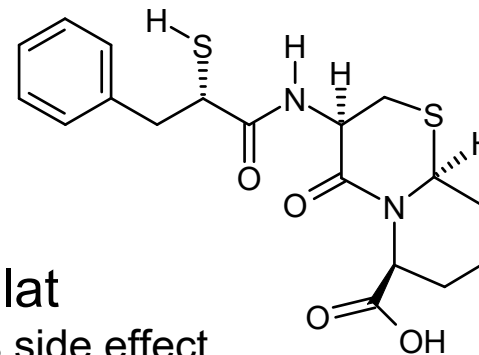
Omapatrilat
also inhibits the Neutral Endopeptidase as side effect

From the lead compound to the drug (VII)

Another possibility to obtain information about the structure is to crystallize homolog enzymes from model organisms followed by homology modelling.

In the case of human tACE (E.C. 3.4.15.1) an ortholog protein of *Drosophila melanogaster* (ANCE) is present, from which another X-Ray structure is available.

In vivo screening of inhibitors is possible with according animal models that possess orthologue enzymes (mouse, rat). For hypertension the rat is establish as animal model.

Lit: K.R.Acharya *Nature Rev. Drug Discov.* **2** (2003) 891.

2nd assignment

Scope:

Ligand-enzyme interactions

Considered systems:

Comparison of lisinopril and captopril bound to tACE

biotin – streptavidin complex

Searching Compound Databases

Problem: How to encode structural information of
   chemical compounds in a digital way?

Solution 1: Depicted structure is used directly as query,
e.g. in the CAS-online (SciFinder) database.

Assignment of a so-called CAS-registry number
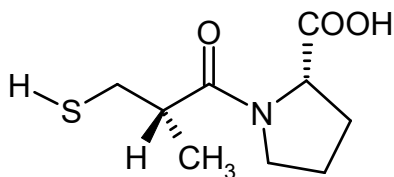
Captopril [62571-86-2]

checksum digit

Solution 2: Using ASCII code as so-called SMILES.

This format is supported by most substance databases
and can be processed using Open Babel and RDKit.

SMILES and SMARTS

Simplified **M**olecular **I**nput **L**ine **E**ntry **S**pecification

Depiction of molecular 2D-structures (configuration) in 1D-form as an alphanumerical string

CC	H_3C-CH_3
C=C	$H_2C=CH_2$
C#C	$HC≡CH$
CCO	H_3C-CH_2OH

rules:

1) Atoms are given by their element names

C B N O P S Cl Br I H   organic subset, others: e.g. [Si] [Fe] [Co]

2) Hydrogens are added automatically: C becomes CH_4

SMILES tutorials and references see
http://www.daylight.com/dayhtml/doc/theory/theory.smiles.html
D. Weininger *J. Chem. Inf. Comput. Sci.* **28** (1988) 31.
http://opensmiles.org

SMILES (II)

2) atoms and bonds

CC   single bonds are not needed to be specified

C=C   bouble bonds

C#C   triple bonds

c:c   aromatic bond between aromatic carbons
(no need to specify)

C@C   any kind of bond in any ring

C~C   any kind of bond (single, double, ring, etc.)

SMILES (III)

3) Parenthesis denote branching

CCN(CC)CCC
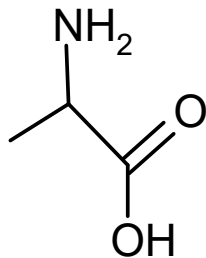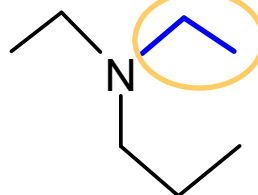
CC(N)C(=O)O

Hint: Determine the longest possible chain in the molecule, first

SMILES (IV)

4) Cyclic compounds: Cutting through a bond yields a chain

Also find the longest chain, first.
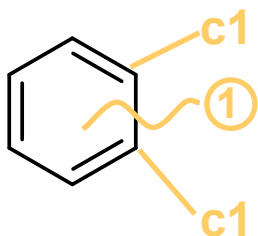
c1ccccc1
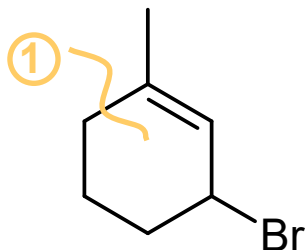
CC1=CC(Br)CCC1

Typical mistake while setting up rings manually: One atom too much upon closing the ring. Hint: Count the number of ring atoms.
Number of atoms in SMILES = number of atoms in molecule

SMILES (V)

polycyclic compounds
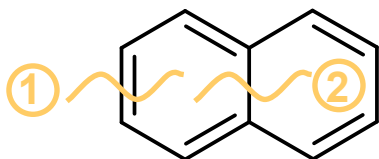
c1cc2ccccc2cc1

There can be more than one ring closures at one atom:
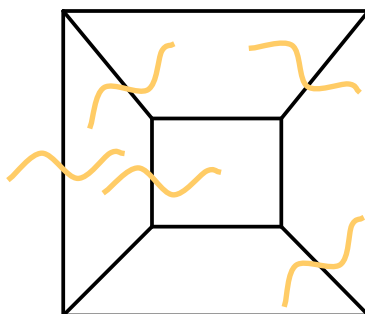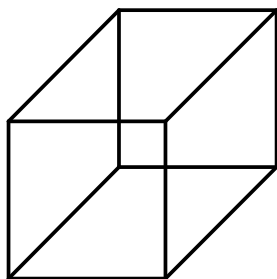
c12c3c4c1c5c4c3c25

Numbers larger than 9 are denoted by a preceeding % : c%11

SMILES (VI)

5) non-covalently bonded fragments are separated by a .
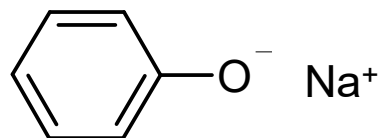
 —O⁻ Na⁺   [Na+].[O-]c1ccccc1

6) isotopes

^{13}C   [13C]

^{13}CH$_4$ [13CH4]   specify the hydrogens !

D$_2$O   [2H]O[2H]

SMILES (VII)

7) Configuration at double bonds

F/C=C\F   above, above
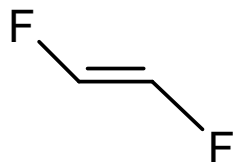
F/C=C/F   below, below

FC=CF   unspecified

matches both
possibilities

Does not (yet) work with Open Babel:

Any difference in configuration will be ignored.

SMILES (VIII)

8) chirality
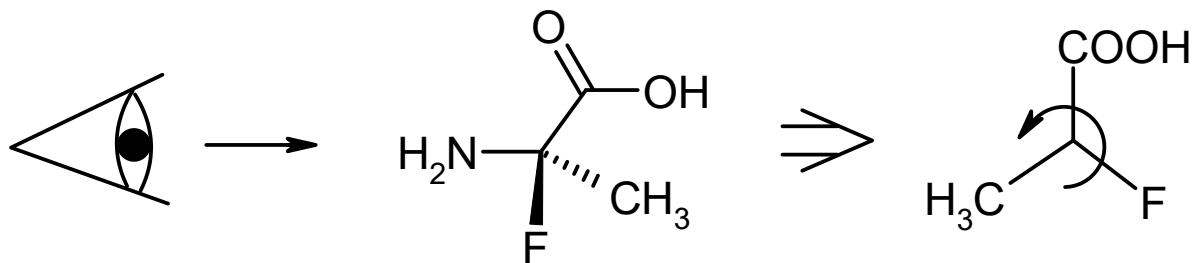
N[C@] (C)(F)C(=O)O

@ anti-clockwise sequence of substituents

@@ clockwise sequence of substituents (anti-anti-clockwise)

Caution: Not conform with the IUPAC R/S nomenclature at stereo centers.

SMILES & SMARTS (IX)

9) Explicit hydrogen atoms

Since hydrogens are added automatically, they only have to be specified if a certain number of hydrogens is required:

C[OH1] H_3C—OH ✓ H_3C—CH_2—OH ✓ H_3C—O—CH_3 ✗

C[NH2] H_3C—NH_2 ✓ H_3C—NH—CH_3 ✗ H_3C—N(CH_3)—CH_3 ✗

C[NH0] H_3C—NH_2 ✗ H_3C—NH—CH_3 ✗ H_3C—N(CH_3)—CH_3 ✓

H^+ [H+]   proton

H_2 [H][H]
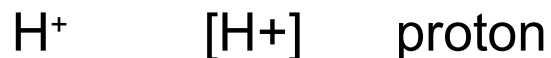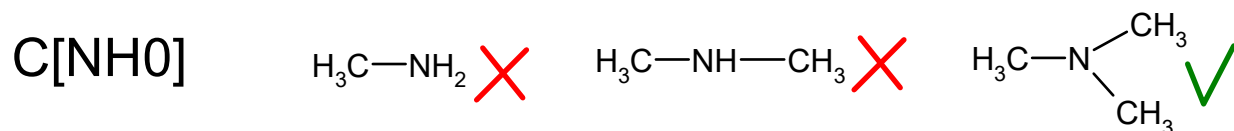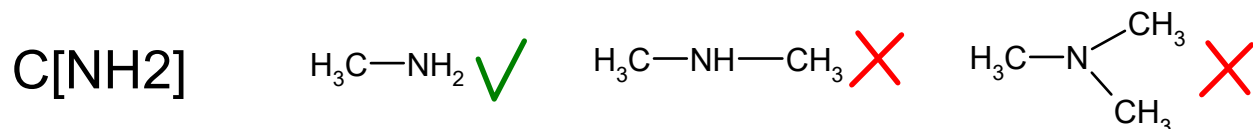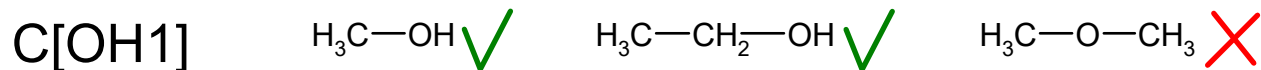
SMARTS (I)

Description of possible substructures and „patterns"

SMARTS are a superset of SMILES with molecular patterns.
A pattern ist grouped by [ ]

example:

      [F,Cl,Br,I]   one atom being either F or Cl or Br or I

1) atoms

c	aromatic carbon
a	aromatic atom (C, N, O, S,...)
A	aliphatic atom (= not aromatic)
*	any atom (including no atom)
[#16]	element no.16 (any kind of sulfur)
[rn]	atom in a n-membered ring
[SX2]	sulfur with two substituents —S— but not —S— or ＝S
[Fe]	iron atom of arbitary charge (0, +1, +2, +3, ...)

SMARTS (II)

2) logical (boolean) operators

A,B	A or B
A&B	A and B (high priority)
A;B	A and B (low priority)
!A	not A

examples:

[F,Cl,Br,I]   F or Cl or Br or I

[!C;R]   non-aliphatic carbon and in a ring (c, N, O,...)

[CH2]   aliphatic carbon with 2 Hs (methylene group)
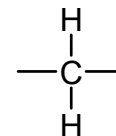
[c,n&H1] aromatic carbon or aromatic NH   [A or (B and C)]

[c,n;H1]   aromatic C or N, and exactly one H   [(A or B) and C]

[#7;r5]   any nitrogen in a 5-membered ring

[CH2;r0] aliphatic carbon with 2 Hs that is not in a ring

SMARTS (III)

3) configuration of substituents. Examples:

     [CaaO]   C ortho to O

     [CaaaN]   C meta to N

     [Caa(O)aN]

     [Ca(aO)aaN]
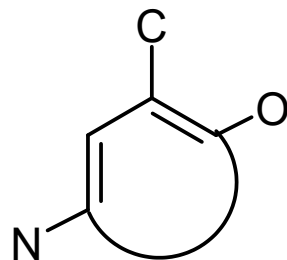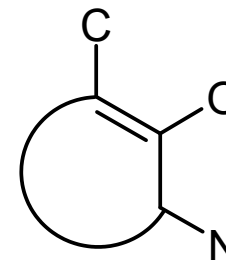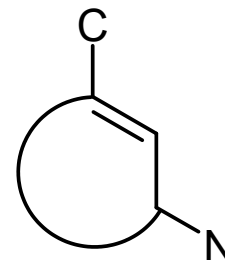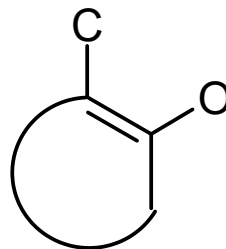
If the patterns consists of more than one atom, they must be grouped using $()
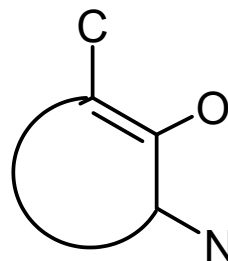
   C[$(aaO),$(aaaN)]   as well as
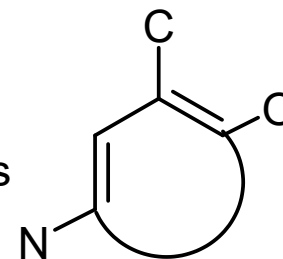
SMARTS (IV)

typical datebase queries

[s,o]1cccc1   thiophenes and furanes
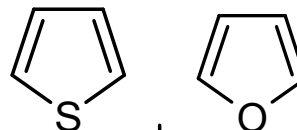
[CX4][NH2]   primary aliphatic amine
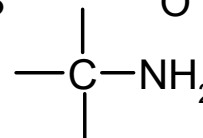
[C1OC1]   epoxides

C(=O)[OH,O-,O-.+]   carbonic acid, carboxylate, or with a cation

C(=O)[NH1]   peptide linkage

=[OH]   acids and enoles

F.F.F.F.F   a total of 5 fluorine atoms in the molecule
                   (does not work with Open Babel)

further examples: E.J. Martin *J. Comb. Chem.* **1** (1999) 32.

Converting different formats of molecule files with Open Babel:
http://openbabel.sourceforge.net

SMILES database entries (I)

typical database entries:

dexamethasone 21-phoshate disodium salt
CC1CC2C3CCC4=CC(=O)C=CC4(C3(C(CC2(C1(C(=O)C
OP(=O)([O-])[O-])O)C)O)F)C.[Na+].[Na+]

text processing
i.e. using Perl-scripts
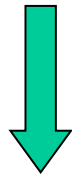
Actual compound in neutral form:

dexamethasone 21-phosphate

CC1CC2C3CCC4=CC(=O)C=CC4(C3(C(CC2(C1(C(=O)C
OP(=O)(O)O)O)C)O)F)C

Virtual screening, QSAR,

3D-Structure conversion, docking: protonation state may be different to
that in solution
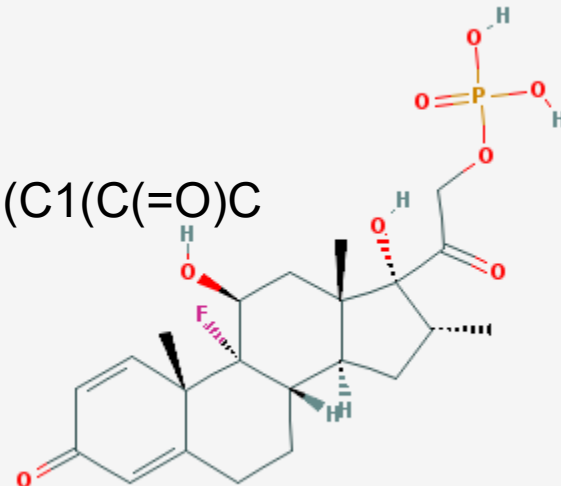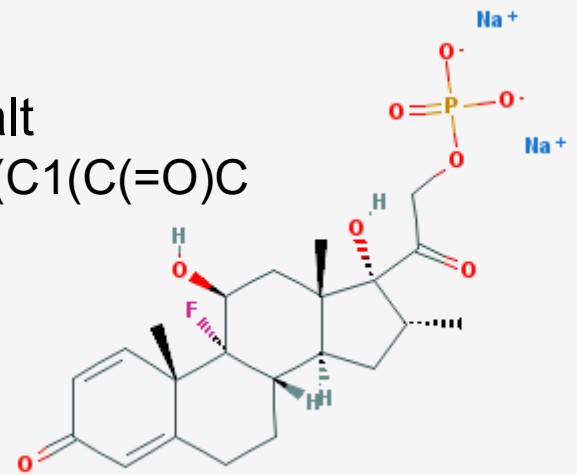
Excursion: Protonation states (I)

The protonation state of chemical groups depends on the pH of the surrounding media and the pK_a of the respective group.

Gasphase
(ions are energetically disfavored in vacuo, if no stabilizing interactions can be formed)
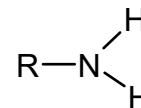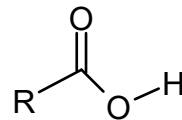
Aqueous solution (pH=7)
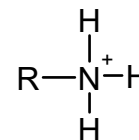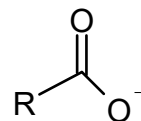
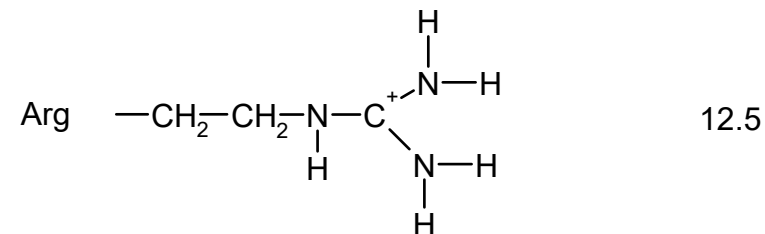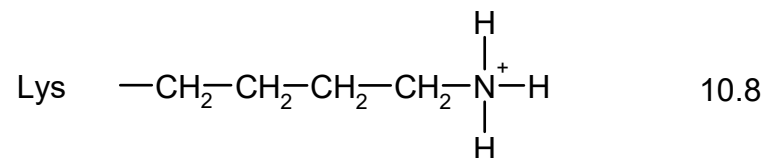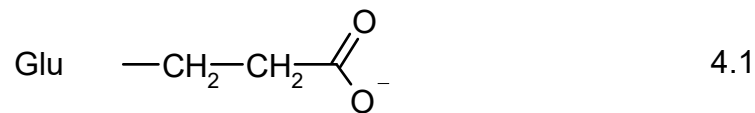Stomach (pH=0.8 – 1.5 empty)
(0.1 mol/l HCl)
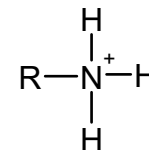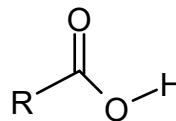
Blood (pH=7.4)
(buffered by H_2CO_3/CO_2, haemoglobin, $H_2PO_4^-/HPO_4^{2-}$)

Excursion: Protonation states (II)

The protonation state is a pH-dependent equilibrium:

$$H_3C\text{-}COOH + H_2O \rightleftharpoons H_3C\text{-}COO^- + H_3O^+$$

HAc   Ac$^-$

$$HAc + NaOH \longrightarrow Na^+Ac^- + H_2O$$
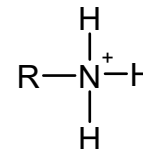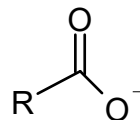
Protonation states of amino acids (I)

typical pK$_a$ values of amino acids side chains.
Protonation state depicted as present in proteins

Asp   —CH$_2$—COO$^-$ 3.9

Glu   —CH$_2$—CH$_2$—COO$^-$ 4.1

Lys   —CH$_2$—CH$_2$—CH$_2$—CH$_2$—N$^+$H$_3$ 10.8

Arg   —CH$_2$—CH$_2$—N(H)—C$^+$(NH$_2$)(NH$_2$)   12.5

His   —CH$_2$—(imidazole ring, N$^+$—H)   6.04   protonation state may vary!
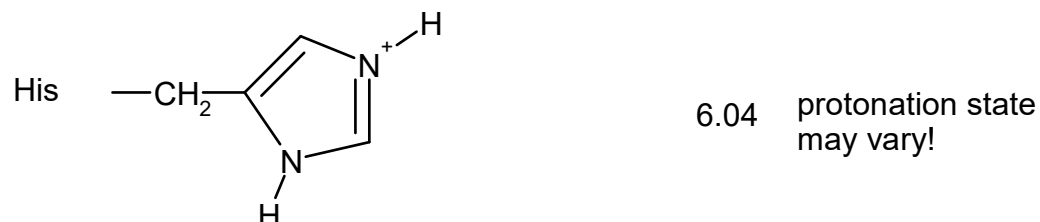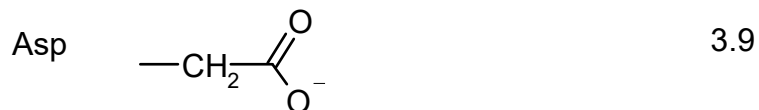
Histidine: depending on possible hydrogen-bonding either one or both nitrogens have a hydrogen attached

Protonation states of amino acids (II)

typical pK_a values of amino acids side chains.
Protonation state depicted as present in proteins

Ser   —CH_2—OH   13
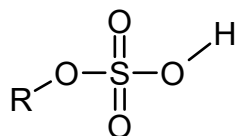
Thr   —CH—OH   13
             CH_3

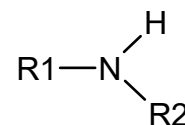Tyr   —CH_2——⟨benzene⟩——OH   10.1
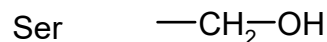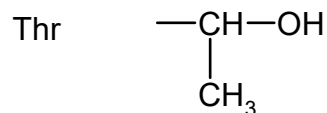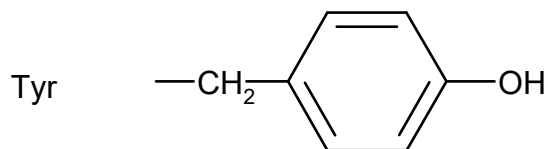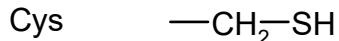
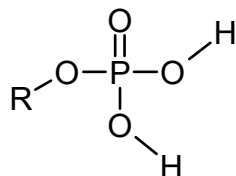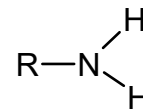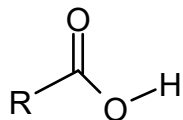Cys   —CH_2—SH   8.3   oxidation to disulfide bridges may occur

Shown pK_a values are those of isolated amino acids in water (*bulk properties*). Depending on the local environment in the protein these may differ.

Frequently found acidic and basic chemical groups in drugs
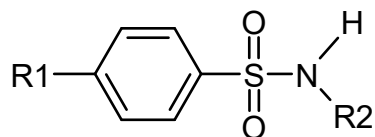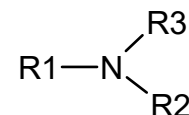
Will be deprotonated
in solution

Will be protonated
in solution

R, R1, R2, R2, R3 = carbon

SMILES database entries (II)

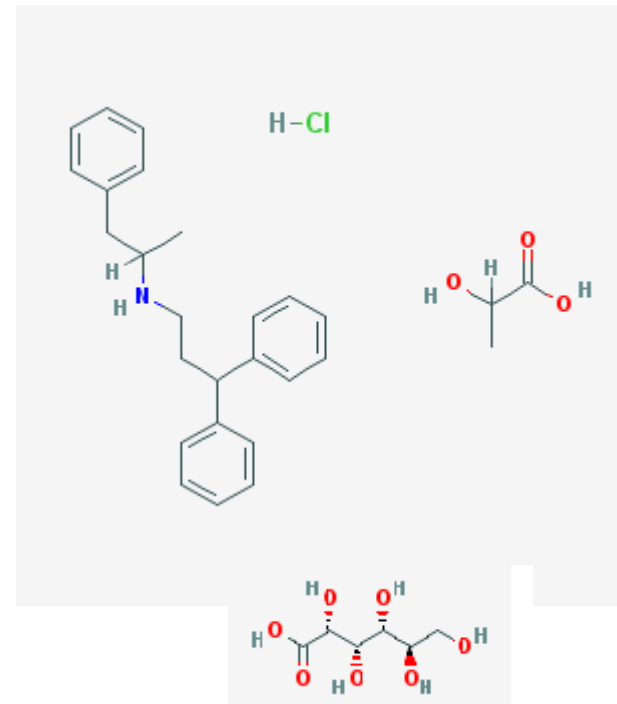Database entries of compound mixtures:

prenylamine hydrochloride

CC(CC1=CC=CC=C1)NCCC(C2=CC=CC=C2)
C3=CC=CC=C3.Cl

prenylamine lactate

CC(CC1=CC=CC=C1)NCCC(C2=CC=CC=C2)
C3=CC=CC=C3.CC(C(=O)O)O

prenylamine gluconate

CC(CC1=CC=CC=C1)NCCC(C2=CC=CC=C2)
C3=CC=CC=C3.C([C@H]([C@H]([C@@H]([C@H](C(=O)O)O)O)O)O)O

Usually the larger molecule is the actually active compound.

Also mixtures of more than two compounds may occur.

→ Intelligent text processing needed!