Pharmacophore Analyses

Hugo Kubinyi

Germany

E-Mail kubinyi@t-online.de
HomePage www.kubinyi.de

Similarity and Dissimilarity

2D similarity based on groups + connectivity

e.g., Daylight fingerprints or MDL keys

2D similarity = Tanimoto index

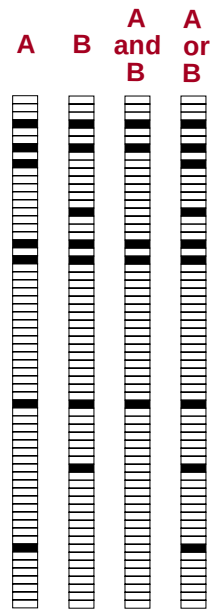
$$\frac{N_{AB}}{N_A + N_B - N_{AB}} = \frac{\text{\# bits set in A and B}}{\text{\# bits set in A or B}} =$$

$$\frac{\text{\# keys common in A and B}}{(\text{\# keys in A})\ (\text{\# keys in B}) - (\text{\# keys common in A and B})}$$

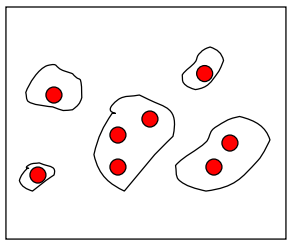
0 <= Tanimoto Index (i, j) <= 1

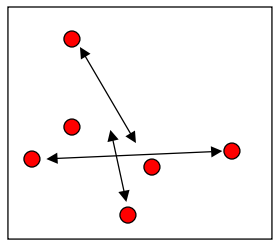
e.g. (example): T(A,B) = 5 / 9 = 0.555

2D dissimilarity = 1 - Tanimoto Index



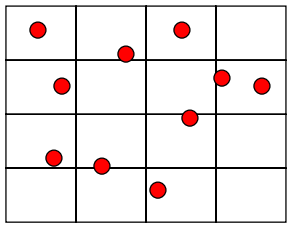
Diversity selections

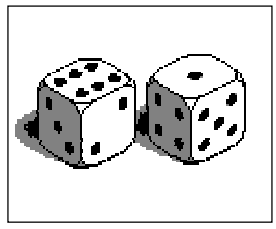




Cluster-based methods

Dissimilarity-based methods

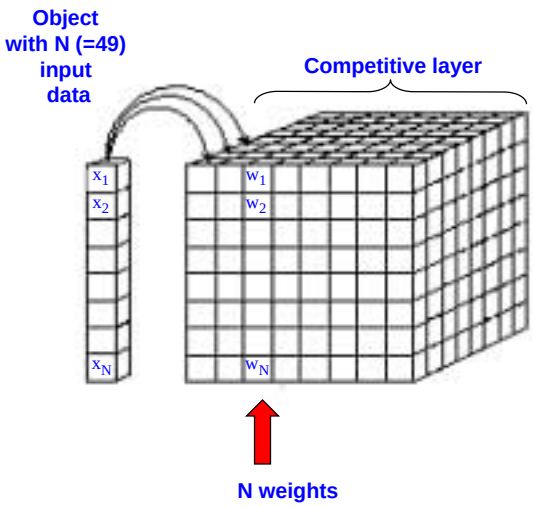




Cell-based methods

Stochastic methods

Self-organizing Maps (SOM, Kohonen Maps; (© J. Gasteiger)



Step 1: initialization of weights with random values
Step 2: comparison of weights with all input data vectors
Step 3: most similar weight vector is associated with input vector which influences its neighborhood within a given radius

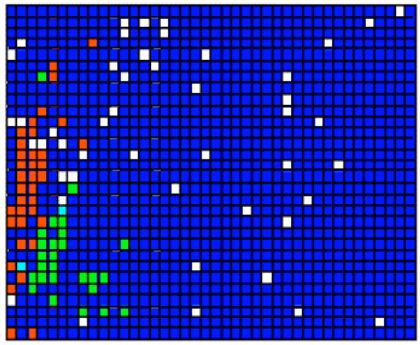
Loop many times with decreasing neighbour-hood radius until chart stabilizes

Self-organizing Maps (SOM, Kohonen Maps)

Classification of large datasets (© J. Gasteiger)

112 dopamine and 60 benzodiazepine agonists in a much larger set of 8323 structures of unknown biological activity

Kohonen map (40x30):



- dopamine agonists
- benzodiazepine agonists
- compounds of unknown activity (Janssen Chimica catalog)
- collisions
- empty neurons

Pharmacophore (pharmacophoric pattern)

A **pharmacophore** is the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target structure and to trigger (or to block) its biological response.

A pharmacophore **does not represent a real molecule** or a real association of functional groups, but a purely abstract concept that accounts for the common molecular interaction capacities of a group of compounds towards their target structure. The pharmacophore can be considered as the largest common denominator shared by a set of active molecules.

This definition discards a **misuse often found** in the medicinal chemistry literature which consists of naming as pharmacophores simple chemical functionalities such as guanidines, sulfonamides or dihydroimidazoles (formerly imidazolines), or typical structural skeletons such as flavones, phenothiazines, prostaglandins or steroids.

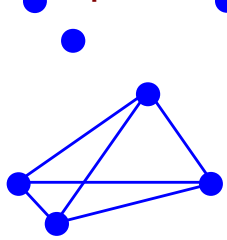
Pharmacophoric descriptors are used to define a pharmacophore, including H-bonding, hydrophobic and electrostatic interaction sites, defined by atoms, ring centers and virtual points.

C. G. Wermuth et al., Pure Appl. Chem. 70,1129-1 143 (1998)

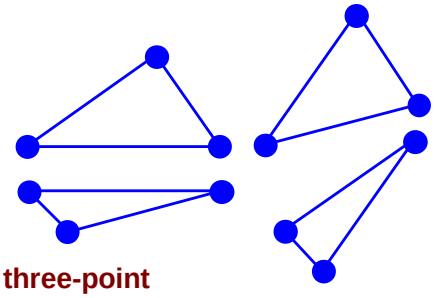
Pharmacophore Definitions

pharmacophore points

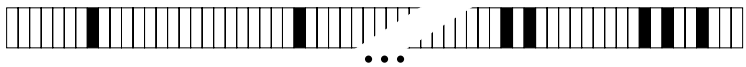
properties: A, D, L
(A, D, P, N, Ar, Al)
distance bins



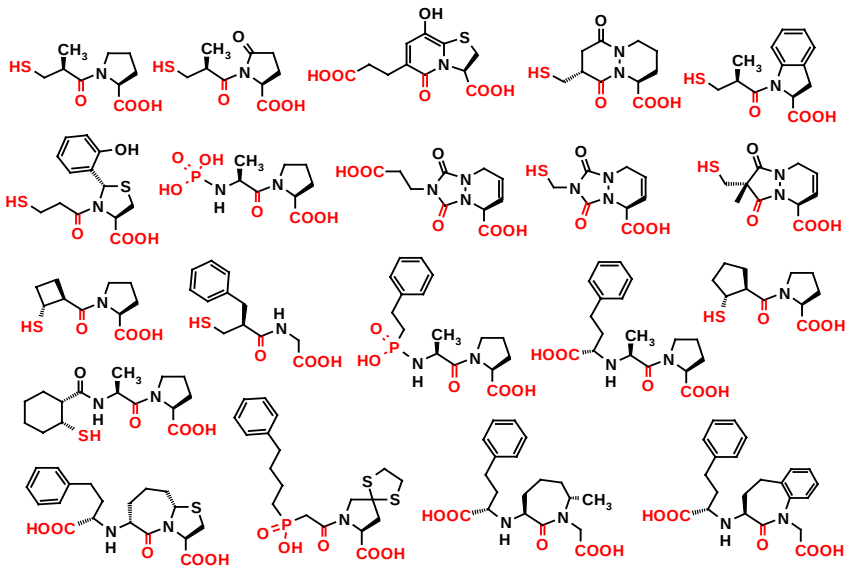
four-point pharmacophore



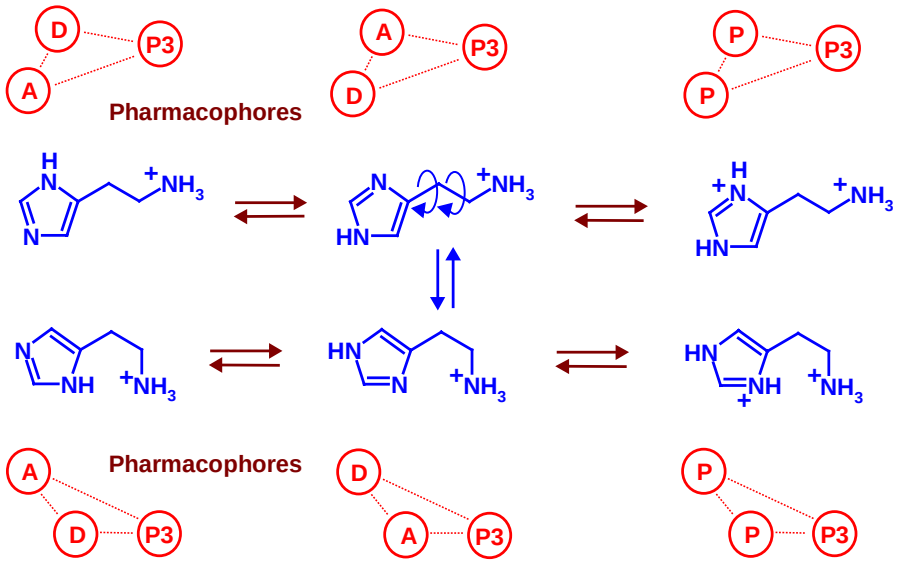
three-point pharmacophores



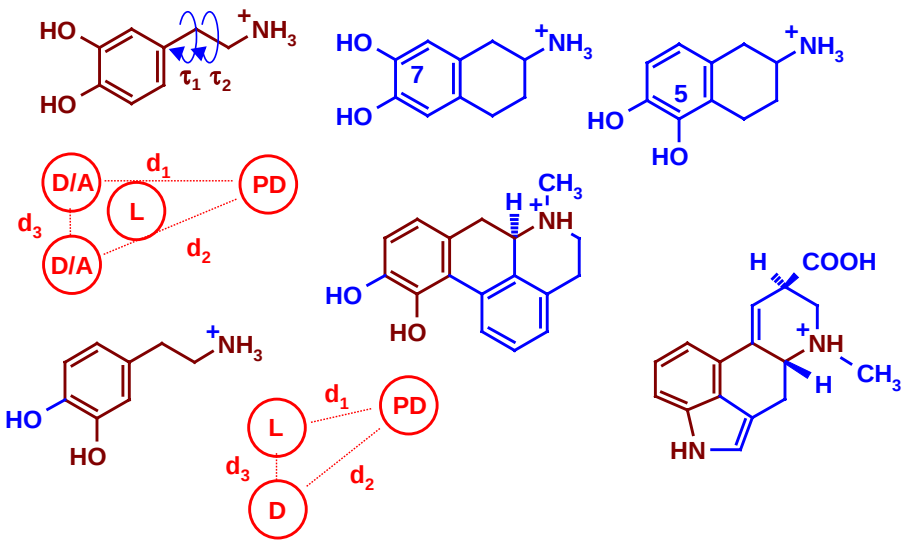
Pharmacophore Hypotheses - ACE Inhibitors



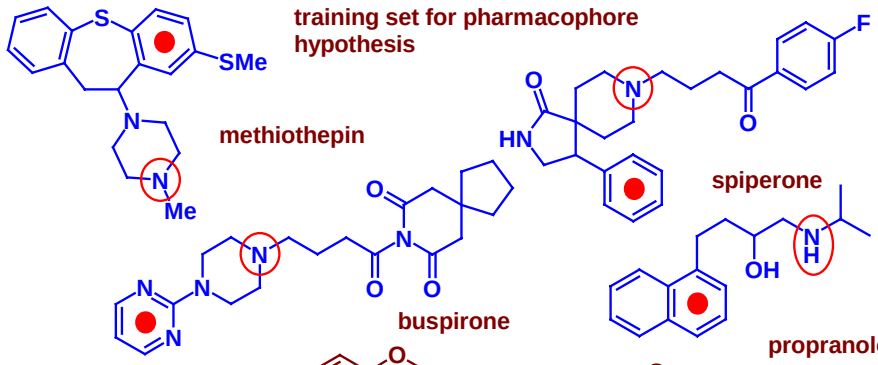
Pharmacophore Hypotheses - Histamine



Pharmacophore Hypotheses - Dopamine

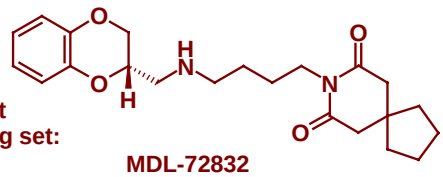


Pharmacophore Hypothesis for 5-HT$_{1A}$ Ligands



proposed molecule: pIC$_{50}$ = 9.1

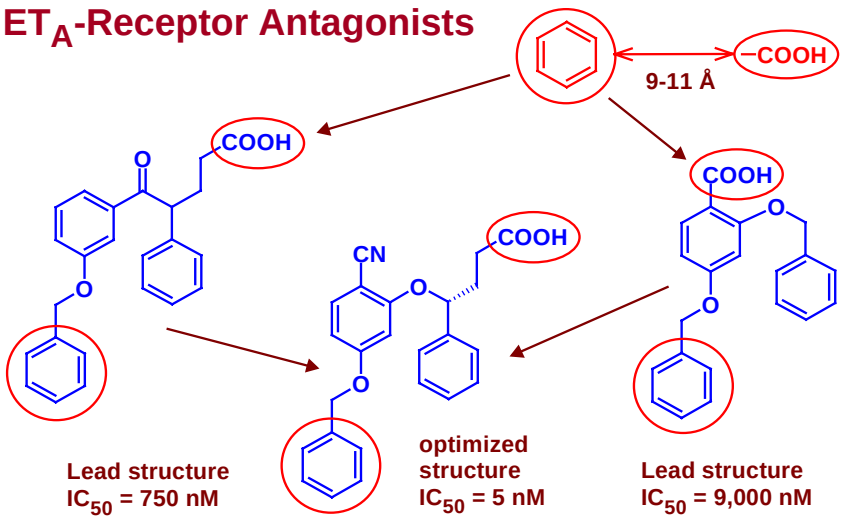
(25x more active than most active analog in the training set: buspirone, pIC$_{50}$ = 7.7)



MDL-72832

M. F. Hibert et al., J. Med. Chem. 31, 1087-1093 (1988)

Pharmacophore Hypotheses for ET$_A$-Receptor Antagonists



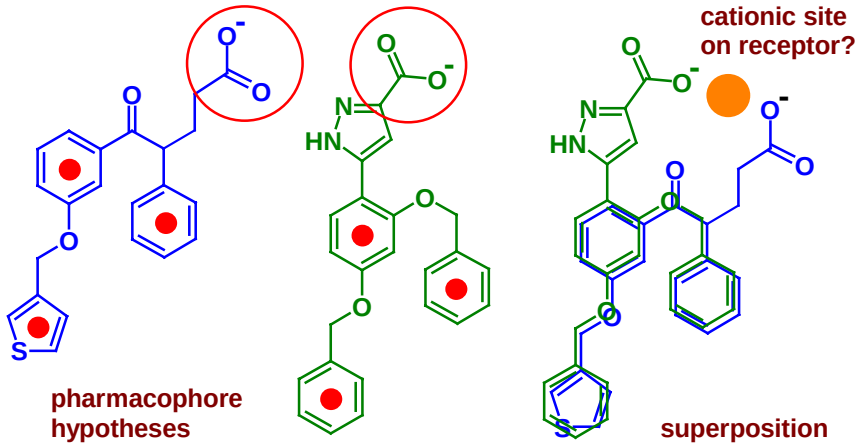
Lead structure
IC$_{50}$ = 750 nM

optimized structure
IC$_{50}$ = 5 nM

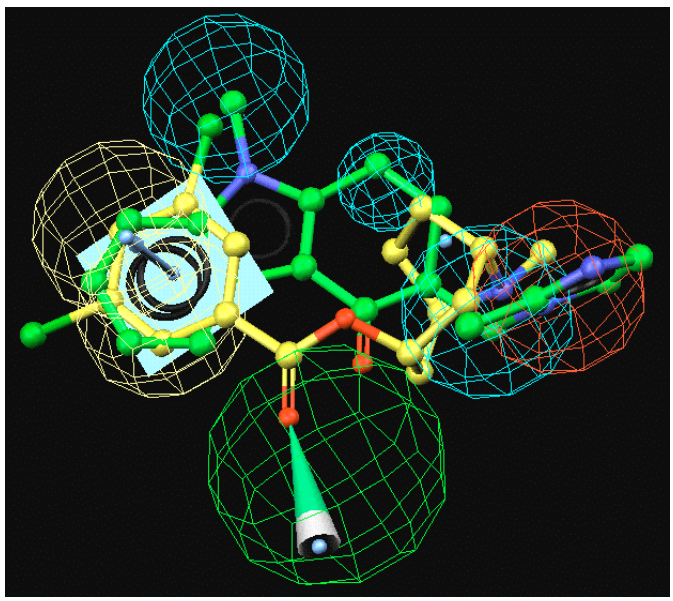
Lead structure
IC$_{50}$ = 9,000 nM

P. C. Astles et al., Eur. J. Med. Chem. 32, 409-423 (1997)

A Unique Pharmacophore Hypotheses for ET_A-Receptor Antagonists



P. C. Astles, J. Med. Chem. 41, 2732-2744 (1998)

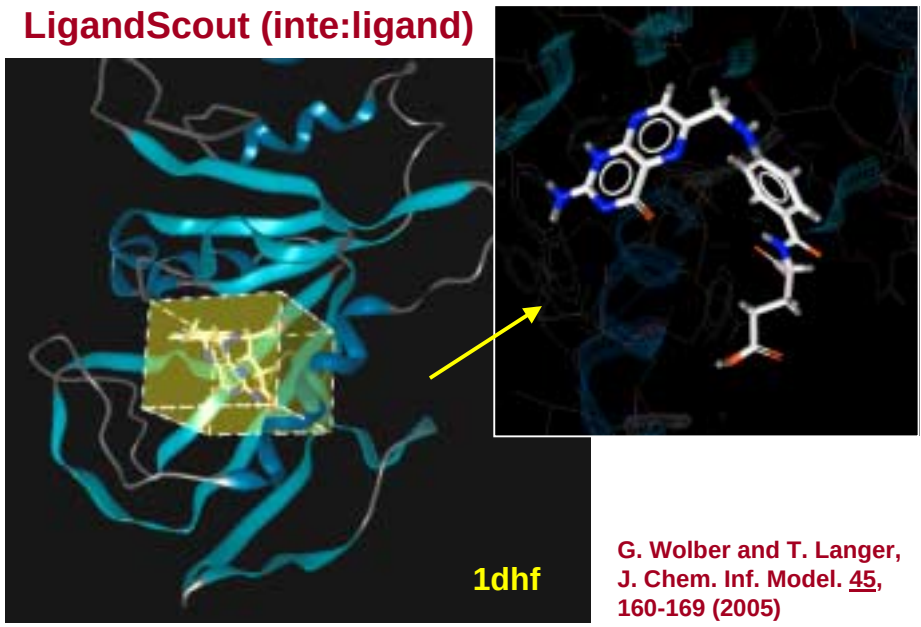


CATALYST (Accelrys)

pharmaco-phore hypothesis of 5-HT_3 ligands for 3D database searches

www.accelrys.com

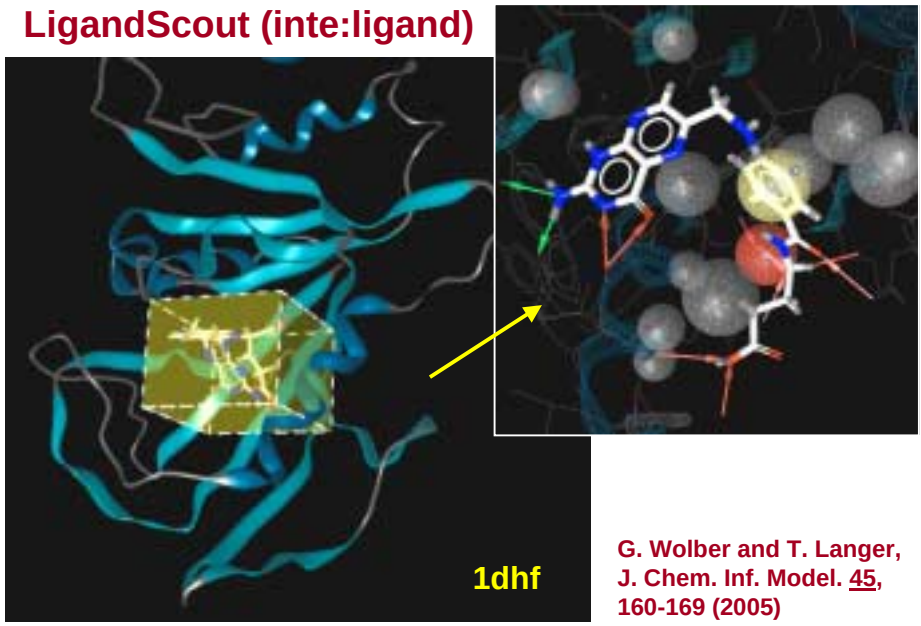
LigandScout (inte:ligand)



G. Wolber and T. Langer, J. Chem. Inf. Model. 45, 160-169 (2005)

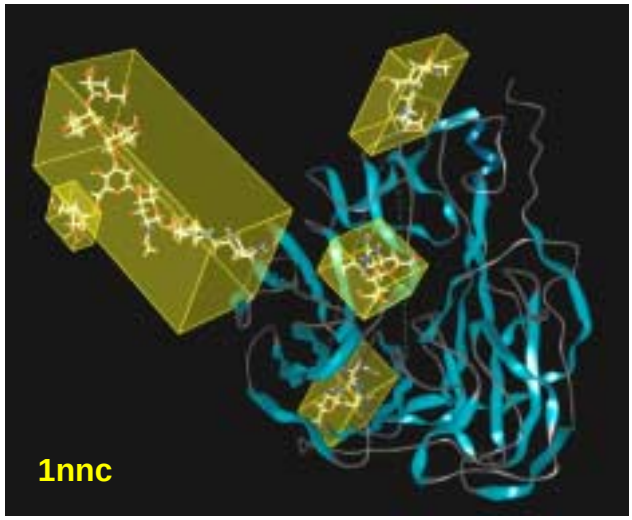
Hugo Kubinyi, www.kubinyi.de

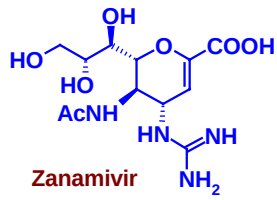
LigandScout (inte:ligand)



G. Wolber and T. Langer, J. Chem. Inf. Model. 45, 160-169 (2005)

LigandScout Superposition: Zanamivir vs. GS 4071

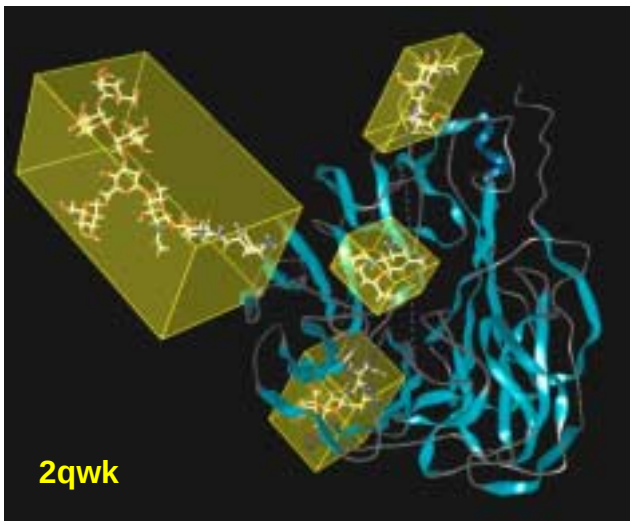


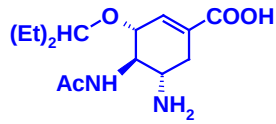


Zanamivir

Hugo Kubinyi, www.kubinyi.de

LigandScout Superposition: Zanamivir vs. GS 4071

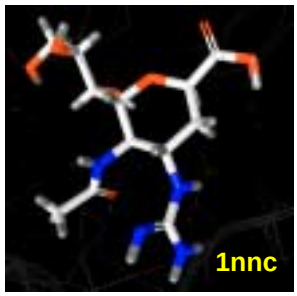


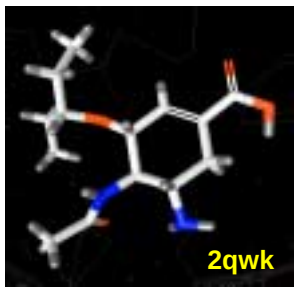


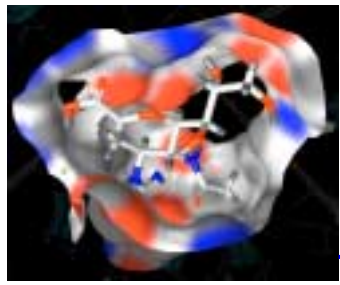
GS 4071 (active form of oseltamivir)

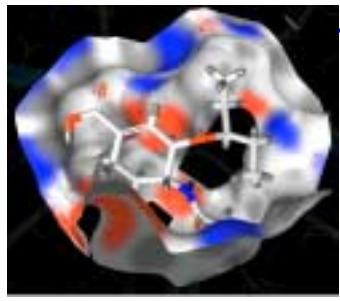
Hugo Kubinyi, www.kubinyi.de

LigandScout Superposition: Zanamivir vs. GS 4071







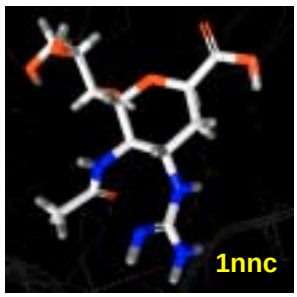


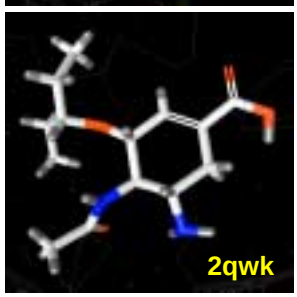
the images of the binding sites are turned around by 180° to show the differences in the glycerol- vs. alkyl-binding pockets

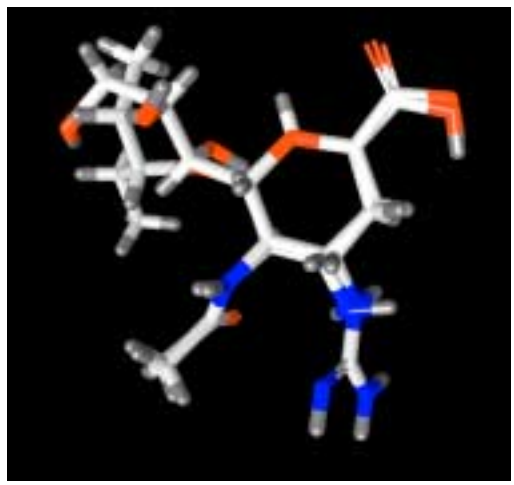
graphics: LigandScout (inte:ligand, Innsbruck)

Hugo Kubinyi, www.kubinyi.de

LigandScout Superposition: Zanamivir vs. GS 4071

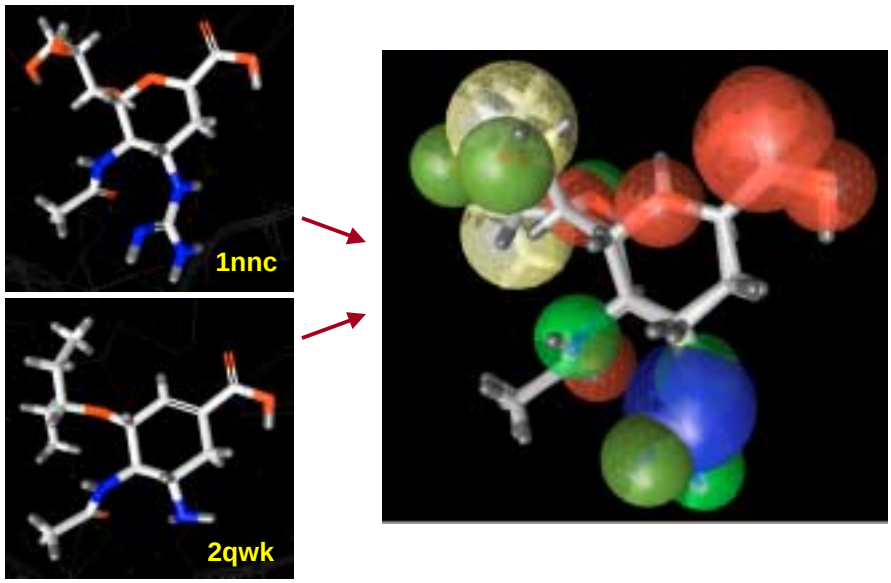






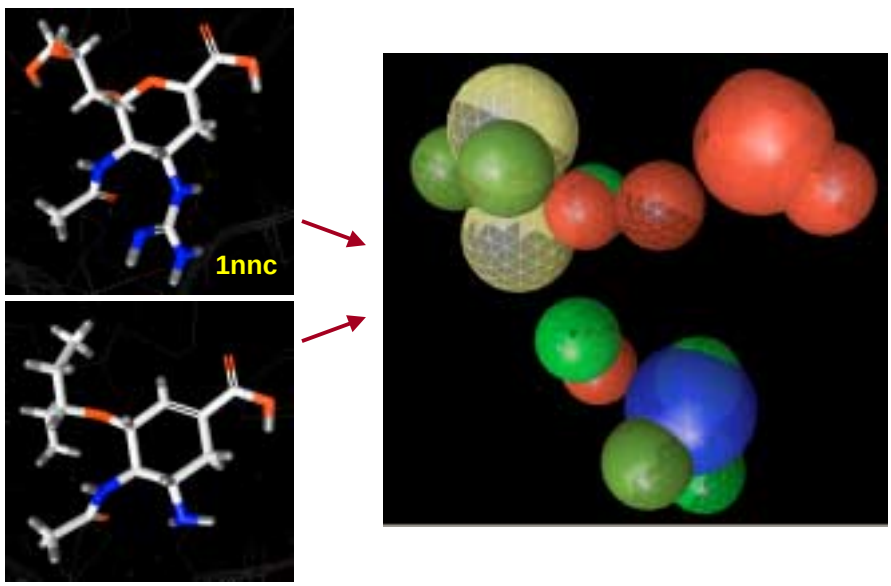
Hugo Kubinyi, www.kubinyi.de

LigandScout Superposition: Zanamivir vs. GS 4071



Hugo Kubinyi, www.kubinyi.de

LigandScout Superposition: Zanamivir vs. GS 4071



Problems in Pharmacophore Definition

Ionisation and Dissoziation
(Sadowski rules, ACS Boston, 2002)

Tautomeric and protomeric forms
(program AGENT, ETH Zurich;
ChemoSoft tautomer recognition, ChemDiv)

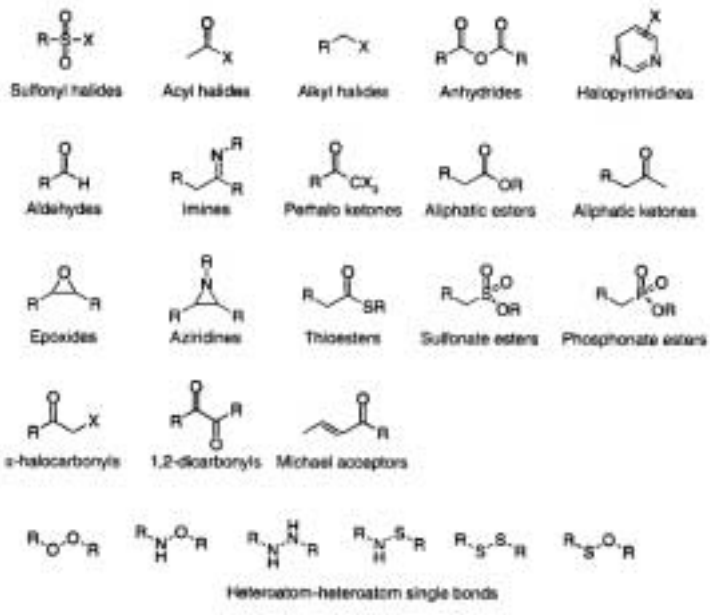
Acceptor properties of oxygen and sulfur atoms
(esters, aromatic ethers, oxazoles,
isoxazoles, thiazoles, etc.)

Pre-Processing of Compound Databases, I

Removal of duplicates
Elimination of counterions
Garbage filter: chemically reactive groups
(e.g. electrophiles, metal chelators, Michael acceptors), undesirable atoms (e.g. organo-metallic complexes); certain groups (option)
Dissociation / protonation equilibria of acids and bases
Protomeric equilibria (e.g. imidazole)
Tautomeric equilibria (or predominant tautomers)

Pre-Processing of Compound Databases, II

- Property filters, e.g. MW, lipophilicity, solubility, PSA, ... (option)
- Bioavailability filters, e.g. Lipinski ROF (option)
- Lead-like / drug-like character (option)
- Selection by chemical diversity (option)
- Generation of correct or alternative configurations, enantiomers, diastereomers
- Generation of „reliable“ 3D structures, by e.g. CORINA, CONCORD, and/or multiple 3D structures, by CATALYST, Mimumba (option)
- Definition or elimination of certain pharmacophores (option)

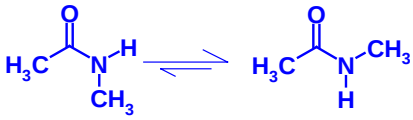


Garbage Filter

reactive functional groups which produce *in vitro* false positive screening hits

The Importance of Correct Conformations

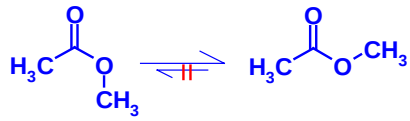
acyclic amides



cis (about 3%) trans (about 97%)

very slow interconversion (<1 sec^{-1})

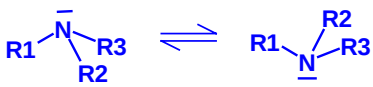
acyclic esters



cis (< 0.01%) trans (> 99.99%)

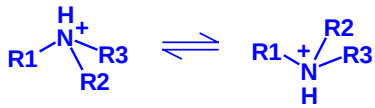
no interconversion

amines



equilibrium, rapid interconversion

protonated amines



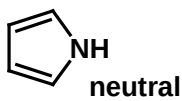
two different configurations (enantiomers)

Hugo Kubinyi, www.kubinyi.de

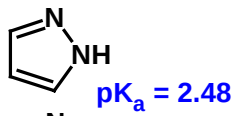
Dissociation of Acids and Protonation of Bases

strong acids	CF_3COOH
acids	arom. + aliph. COOH, $CF_3SO_2NH_2$, tetrazole
weak acids	arom. OH, arom. SO_2NH_2
neutral	aliph. -OH, $-CONH_2$
weak bases	arom. NH_2, imidazole
bases	aliph. NH_2
strong bases	amidines, guanidines

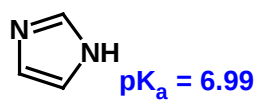
pK_a Values of Selected Organic Compounds



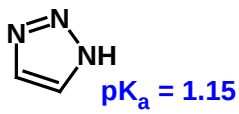
neutral



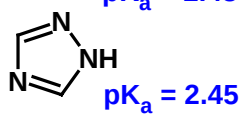
pK_a = 2.48



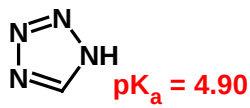
pK_a = 6.99



pK_a = 1.15



pK_a = 2.45



pK_a = 4.90

J. Catalan et al., Adv. Heterocyclic Chem. 41, 187-274 (1987)

Dissociation of Acids and Protonation of Bases

Sadowski (AstraZeneca) rules

(ACS Meeting August 2002, Boston)

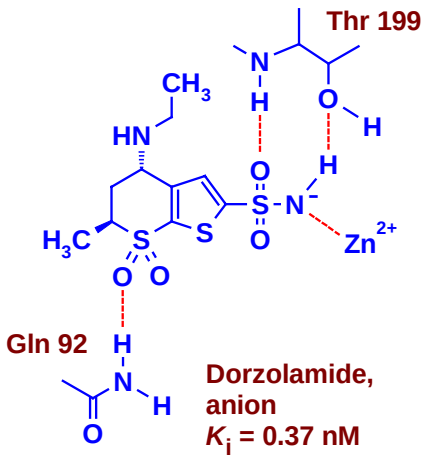
permanently charged: acids, amidines, guanidines, quart. N, ...

negative charges: tetrazole, thiols, hydroxamic acids,
acidic nitrogen (e.g. activated sulfonamides), ...

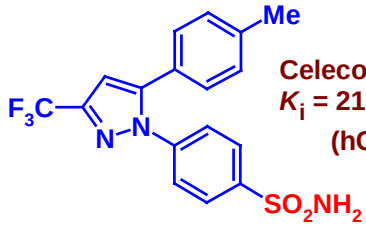
positive charges: basic amines, imidazoles, pyridines, ...

protonation restrictions
- **maximum number of permanent charges per molecule**
- **maximum number of „chargable“ atoms**
- **maximum total charge**
- **maximum number of charges in the same ring**
- **no identical charges in adjacent positions**

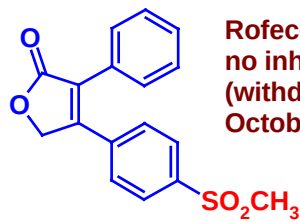
Dissociation of Carbonic Anhydrase Inhibitors



Dorzolamide, anion
K_i = 0.37 nM



Celecoxib
K_i = 21 nM
(hCA II)

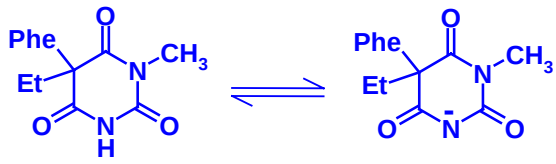


Rofecoxib
no inhibition
(withdrawn in October 2004)

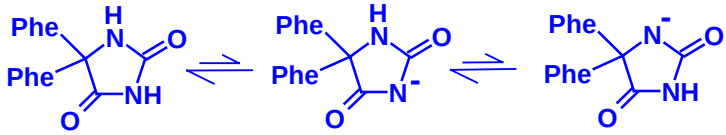
A. Weber et al., J. Med. Chem. 47, 550-557 (2004)

Dissociation of Selected Organic Compounds

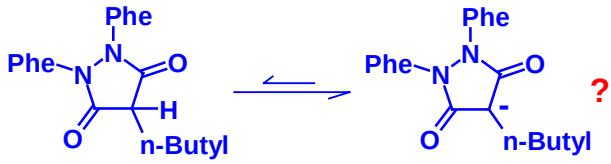
N-Methyl-phenobarbital
pK_a = 7.4



Diphenyl-hydantoin
pK_a = 8.3

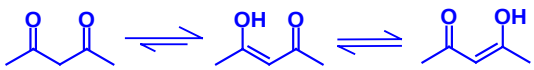


Phenyl-butazone
pK_a = 4.5

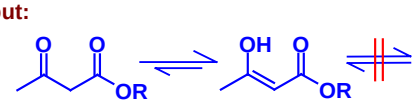


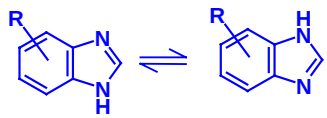
?

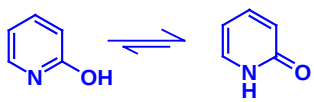
Some Typical Tautomeric Equilibria

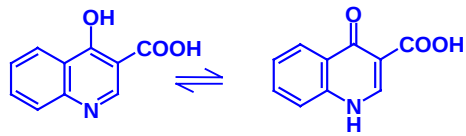


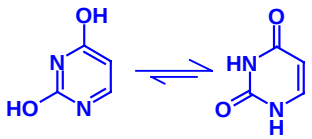
but:

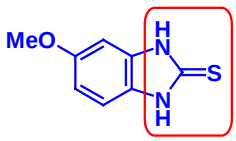




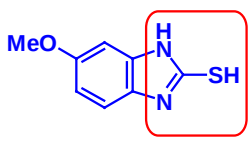








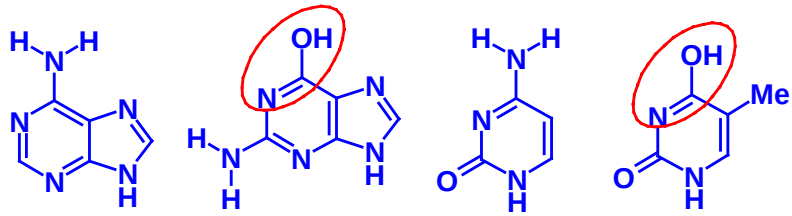
ACD # 126 388
Maybridge RH 00232



ACD # 40 405
Maybridge KM 07202

The Discovery of the DNA Double Helix

Summer 1952: Erwin Chargaff critisizes that Francis Crick and James Watson are ignorant about the structures of the bases



adenine guanine cytosine thymine

J. N. Davidson, The Biochemistry of Nucleic Acids, London, 1950

early 1953: Pauling publishes a DNA model with a phosphate core
February 27, 1953: Jerry Donohue corrects the formulas of the bases
February 28, 1953: Watson and Crick derive the correct DNA model
April 02, 1953: Manuscript sent to Nature; published April 25, 1953

cited from: J. Watson and A. Berry, DNA. The Secret of Life, 2003

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three inter-twined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This

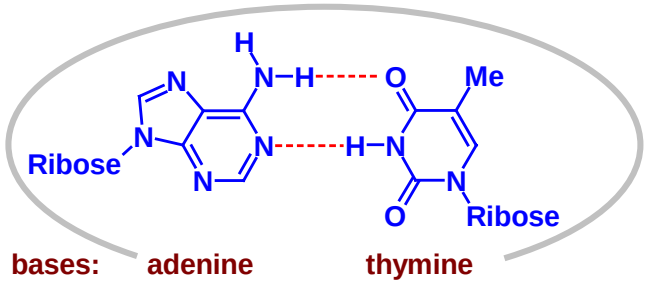
The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical z-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

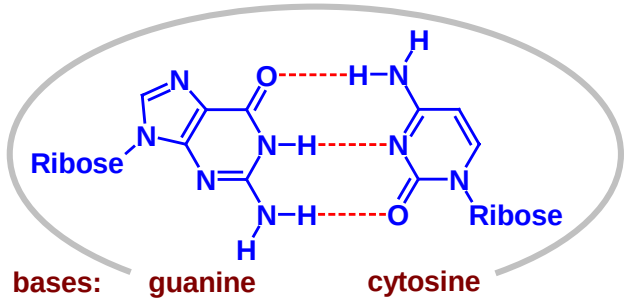
In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

J. D. Watson and F. H. C. Crick, Nature 171, 737-738 (April 25, 1953)

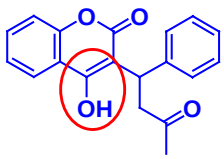
A-T and G-C Pairs in DNA



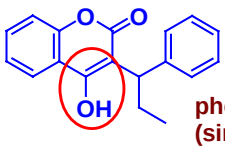
(Watson and Crick, 1953)



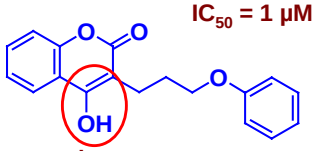
HIV-Protease Inhibitors from Anticoagulants



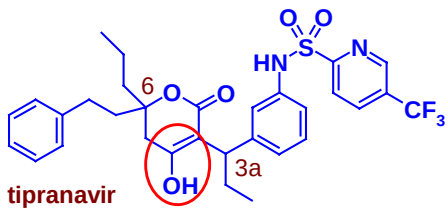
warfarin (screening at Upjohn) IC_{50} = 30 µM



phenprocoumon (similarity search at Upjohn) IC_{50} = 1 µM



screening at Parke/Davis K_i = 2.3 µM

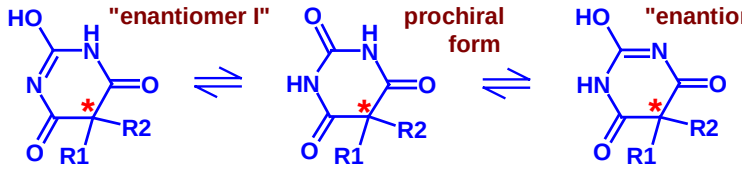


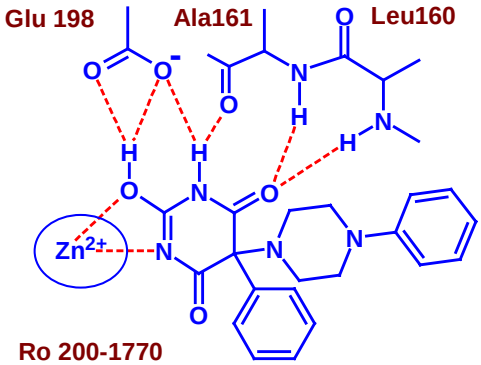
tipranavir (PNU 140 690)

diastereo-mer	K_i pM	IC_{50} µM	IC_{90} µM
R,R	8	0.03	0.10
R,S	18	0.14	0.84
S,R	32	0.41	1.8
S,S	220	1.7	3.0

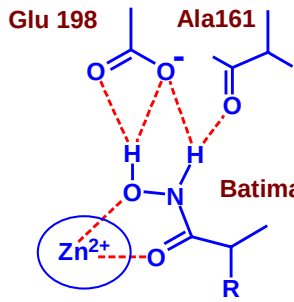
S. R. Turner et al., J. Med. Chem. 41, 3467-3476 (1998)

Tautomeric Forms of an MMP-8 Inhibitor (1jj9)





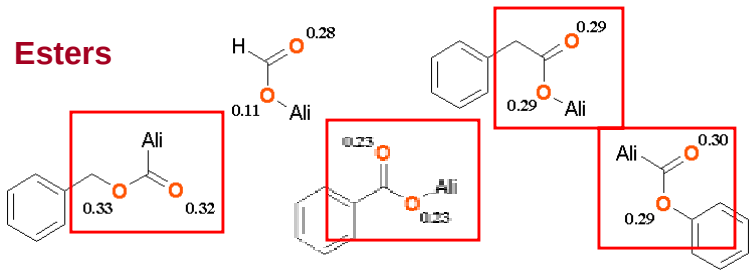
Ro 200-1770



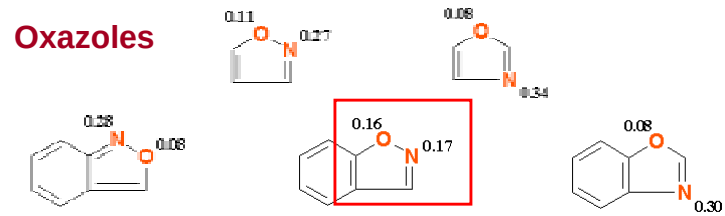
H. Brandstetter et al., J. Biol. Chem. 276, 17405-17412 (2001)

Acceptor Potentials of Esters and Oxazoles

Esters



Oxazoles



?

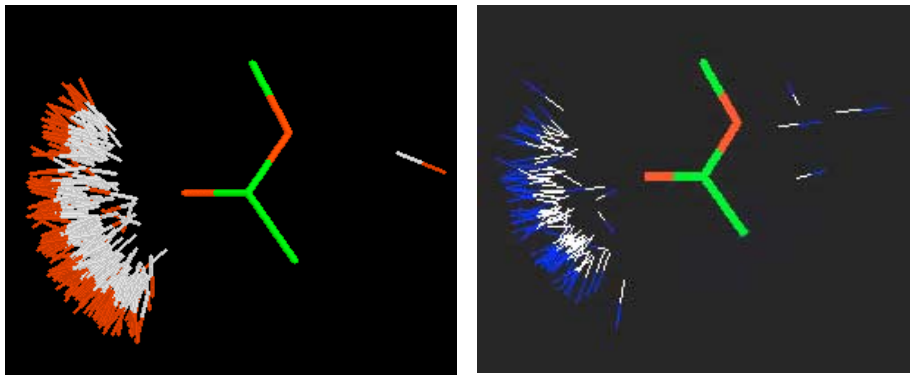
© http://ictsg10.unil.ch/ict/systahl/frag2/frag2_index.htm

Pharmacophore Analyses Must Consider Correct Donor and Acceptor Properties of Ligands

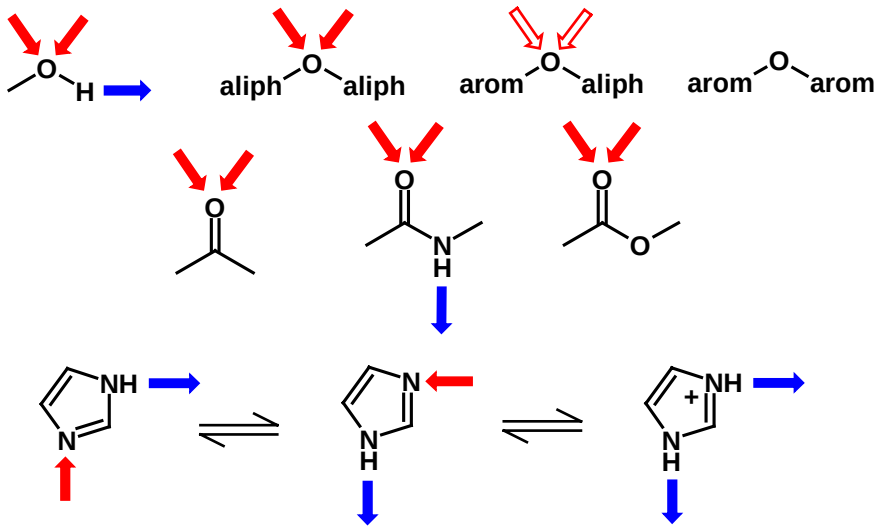
The billion dollar question:

how many acceptor positions has an ester group ?

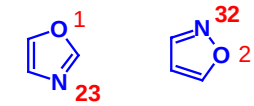
Correct answer: Two ...., but why?



Donor and Acceptor Properties of O and N

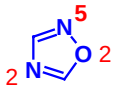


Acceptor Properties of O, N and S



oxazole

isoxazole

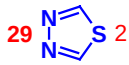


1,2,4-oxadiazole

1,2,5-oxadiazole

thiazole

isothiazole



1,3,4-thiadiazole

isoxazole

thiazole

hydrogen bonding contacts observed in the Cambridge Crystallographic Database are indicated as red numbers (www.ccdc.cam.ac.uk/prods/isostar/appnot1.html)

Hugo Kubinyi, www.kubinyi.de

References: Acceptor Properties of O and N

P. Murray-Rust, J. P. Glusker, Directional hydrogen bonding to sp2- and sp3-hybridized oxygen atoms and its relevance to ligand-macromolecule interactions, J. Am. Chem. Soc. 106, 1018-1025 (1984).

H.-J. Böhm, S. Brode, U. Hesse, G. Klebe, Oxygen and nitrogen in competitive situations: Which is the hydrogen-bond acceptor? Chem. Eur. J. 2, 1509-1513 (1996).

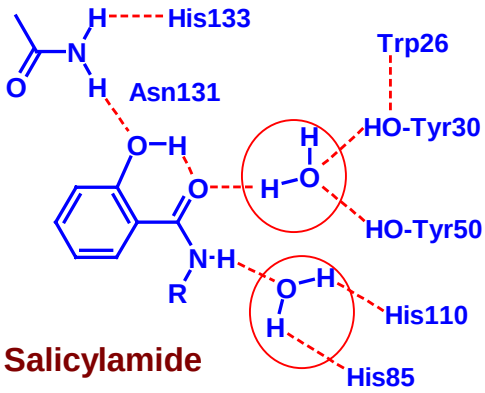
I. J. Bruno, J. C. Cole, J. P. M. Lommerse, R. S. Rowland, R. Taylor, M. L. Verdonk, IsoStar: A library of information about nonbonded interactions, J. Comput.-Aided Mol. Design 11, 525-37 (1997).

J. P. M. Lommerse, S. L. Price, R. Taylor, Hydrogen bonding of carbonyl, ether, and ester oxygen atoms with alkanol hydroxyl groups, J. Comput. Chem. 18, 757-774 (1997)

R. Taylor, Life-science applications of the Cambridge Structural Database, Acta Cryst. D58, 879-888 (2002).

Isostar (CCDC): www.ccdc.cam.ac.uk/prods/isostar/index.html.

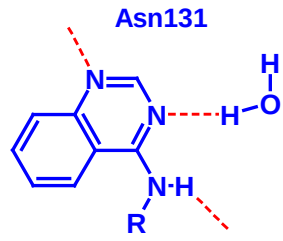
Scytalone Dehydratase Inhibitors



Salicylamide

R = -CH(CH$_3$)C$_6$H$_4$-p-Br

***K*$_i$ = 0.14 nM**



Quinazoline

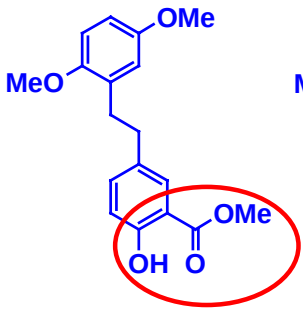
R = -CH$_2$CH$_2$CH(C$_6$H$_5$)$_2$

***K*$_i$ = 0.15 nM**

J. M. Chen et al., Biochemistry 37, 17735-17744 (1998)

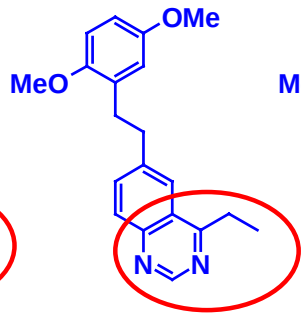
Bioisosterism of Salicylates and Quinazolines

SDZ LAP 977

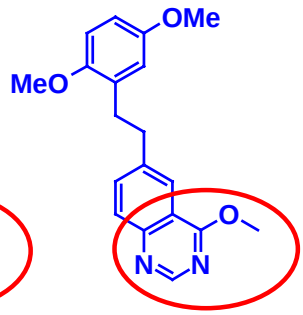


IC$_{50}$ = 47 nM

SDZ LAV 694



7 nM



4 nM

(inhibition of tubulin polymerisation; antiproliferative activity in a keratinocyte cell line)

P. Nussbaumer, Novartis, 17th Int. Symp. Med. Chem., Sept. 2002