Virtual Screening - The Road to Success

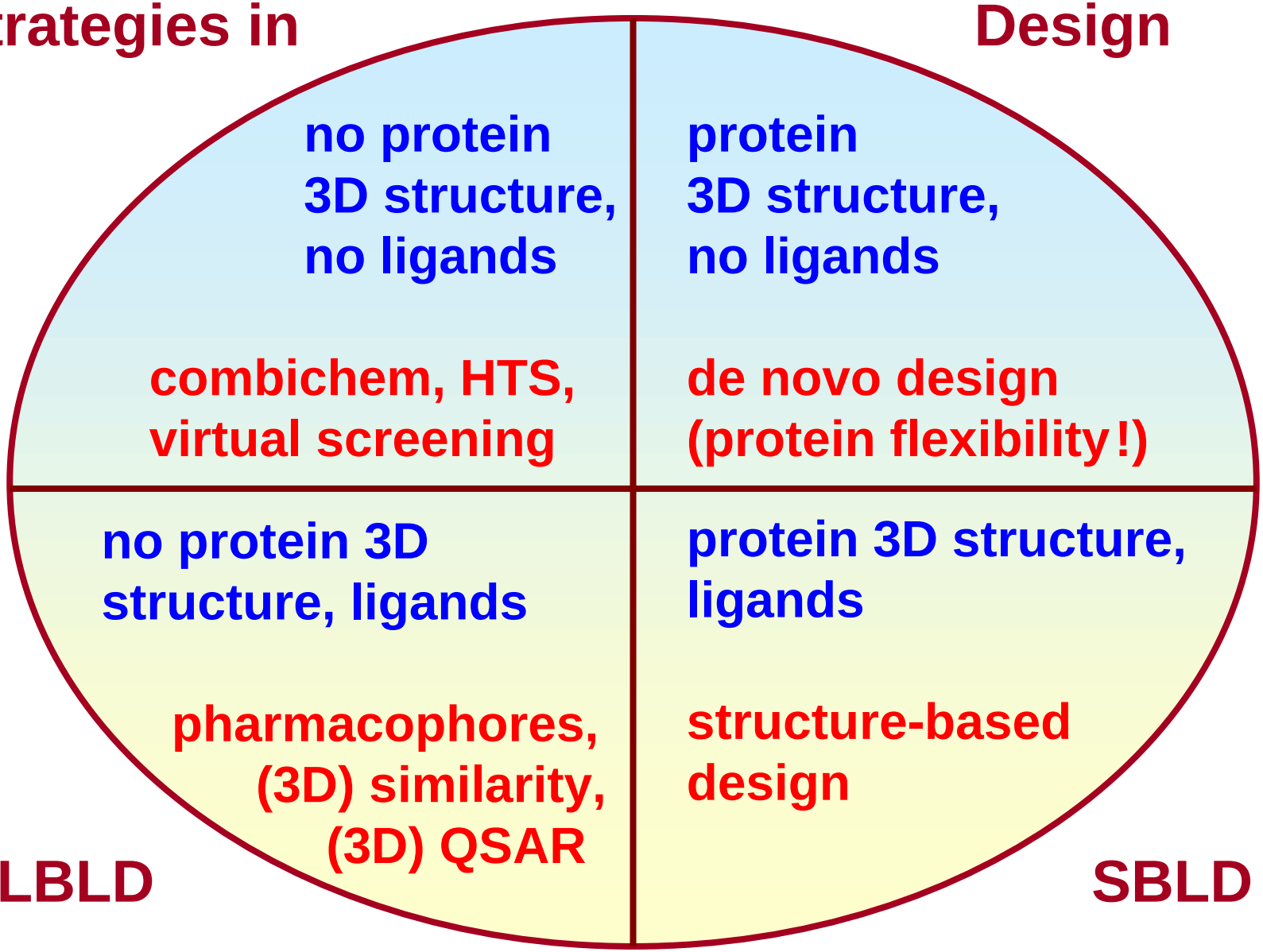
Hugo Kubinyi

Germany

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

XIXth ISMC, Istanbul, Turkey,
August 29-September 02, 2006

Strategies in Design



Drug Research is ....

the Search for a Needle in a Haystack

Virtual Screening Reduces the Size of the Haystack by Selecting:

Compounds or libraries that are either
lead-like, or
drug-like, or have the
potential of oral bioavailability, or are
similar to a lead, or
fit the binding site of a certain protein

by rules (e.g. Lipinski bioavailability rules),
neural nets (e.g. drug-like character),
similarity analyses,
pharmacophore analyses,
scaffold hopping, or
docking and scoring

Chris Rescues CombiChem and Screening Collections

Advanced Drug Delivery Reviews 23 (1997) 3–25

advanced drug delivery reviews

Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings

Christopher A. Lipinski*, Franco Lombardo, Beryl W. Dominy, Paul J. Feeney

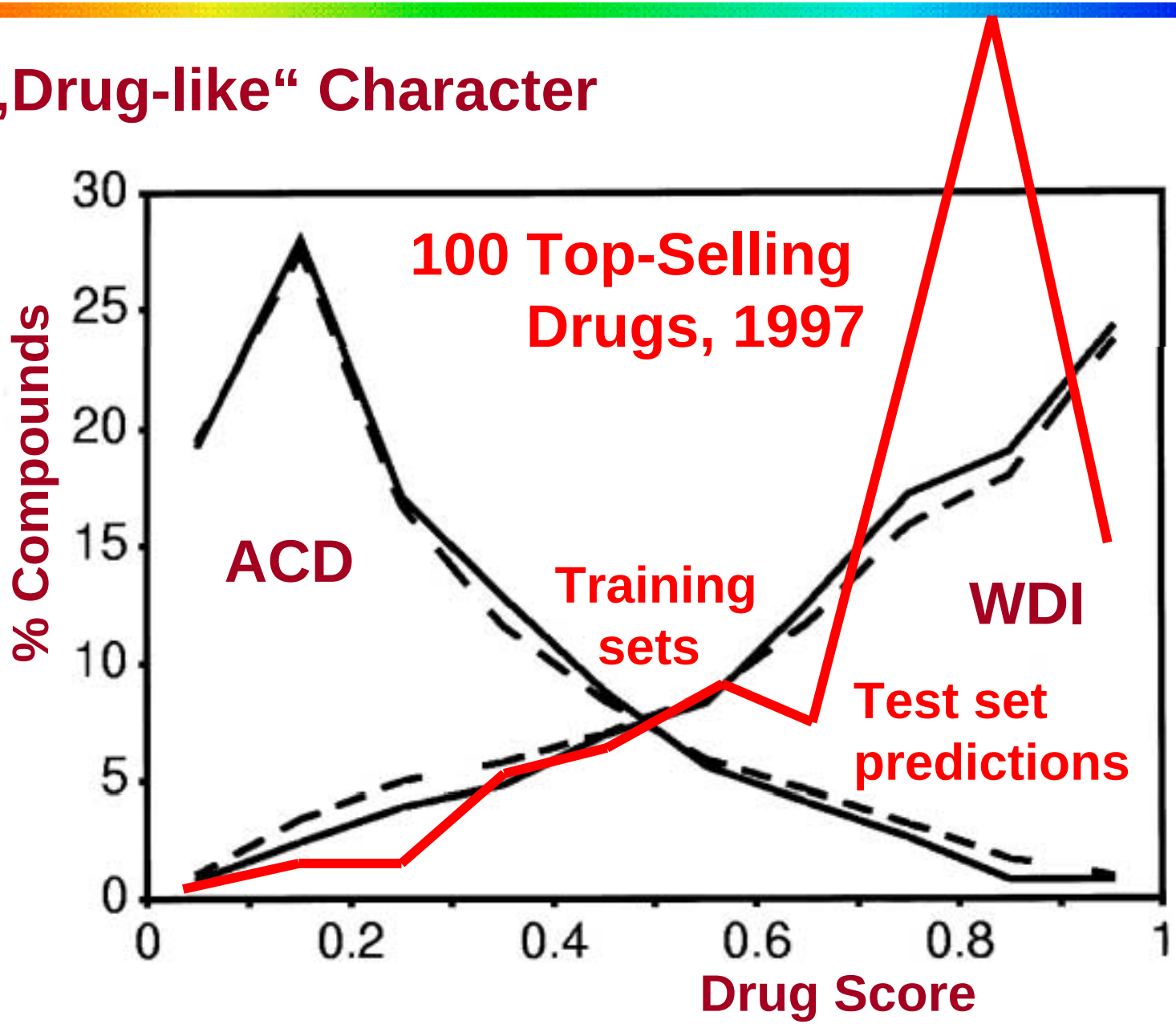
Central Research Division, Pfizer Inc., Groton, CT 06340, USA

Received 9 August 1996; accepted 14 August 1996

Abstract

Experimental and computational approaches to estimate solubility and permeability in discovery and development settings are described. In the discovery setting 'the rule of 5' predicts that poor absorption or permeation is more likely when there are more than 5 H-bond donors, 10 H-bond acceptors, the molecular weight (MWT) is greater than 500 and the calculated Log P (CLogP) is greater than 5 (or MlogP > 4.15). Computational methodology for the rule-based Moriguchi Log P (MLogP) calculation is described. Turbidimetric solubility measurement is described and applied to known drugs. High throughput screening (HTS) leads tend to have higher MWT and Log P and lower turbidimetric solubility than leads in the pre-HTS era. In the development setting, solubility calculations focus on exact value prediction and are difficult because of polymorphism. Recent work on linear free energy relationships and Log P approaches are critically reviewed. Useful predictions are possible in closely related analog series when coupled with experimental thermodynamic solubility measurements.

„Drug-like“ Character



Filters for Virtual Screening

	remaining
Garbage filter	90%
Druglike / Non-druglike	60%
Bioavailability	40%
Cytotoxicity	:
hERG channel inhibiton	:
Antitargets	:
α1a (orthostatic hypotension)	:
D2 (extrapyramidal syndrome)	:
5-HT2c (obesity)	:
musc. M1 (hallucinations, memory)	:
CYP inhibition (3A4, 2C9, 2D6)	0% ?

Pharmacophore Generation and Searches

Catalyst (Accelrys)

established tool for hypothesis generation and 3D searches

CATS topological pharmacophores (Roche)

no 3D structures required

FTree (feature trees; BioSolveIT)

no 3D searches required, ultrafast searches

LigandScout (inte:ligand)

automated generation of bioactive pharmacophores from protein 3D structures

Problems in Pharmacophore Generation

Isomers, enantiomers, diastereomers

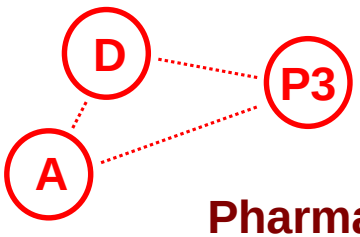
Superposition of flexible molecules

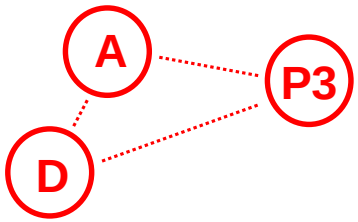
Ionisation and Dissoziation
(Sadowski rules, ACS Boston, 2002)

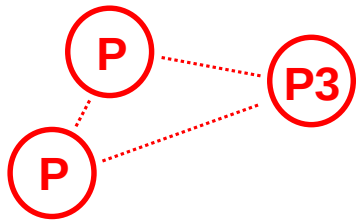
Tautomeric and protomeric forms
(program AGENT, ETH Zurich; ChemoSoft tautomer recognition, ChemDiv)

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

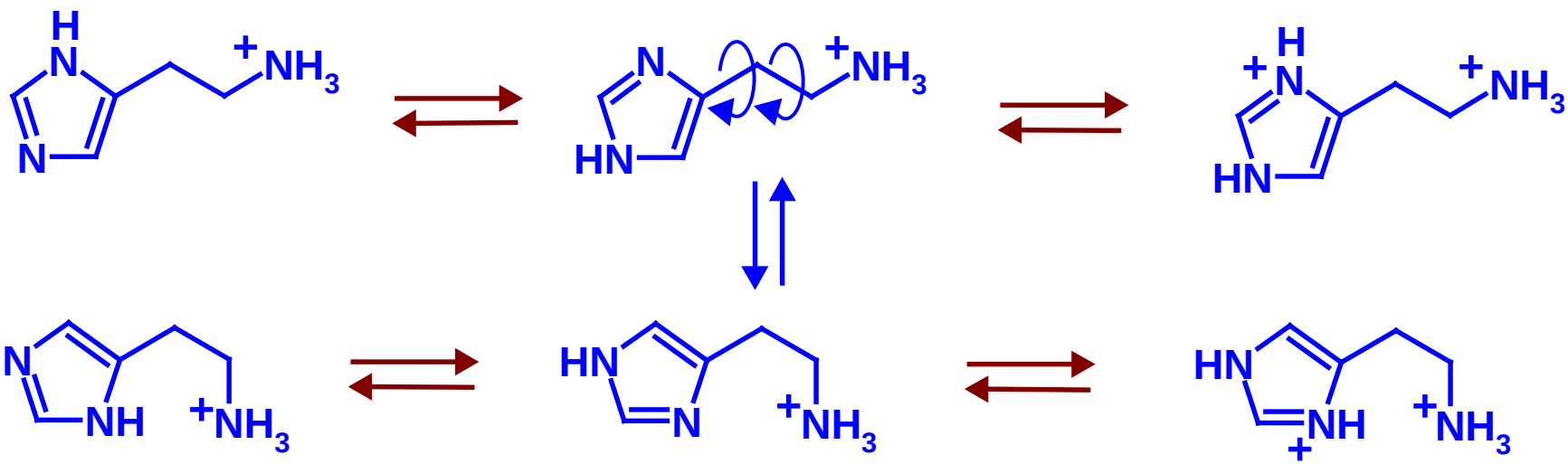
Pharmacophore Hypotheses - Histamine



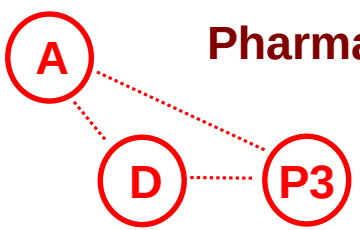


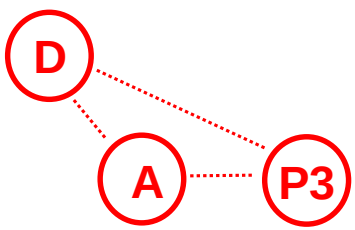


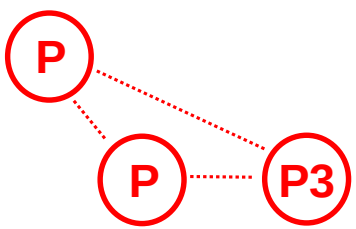
Pharmacophores



Pharmacophores



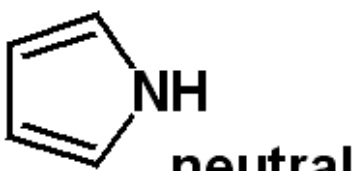




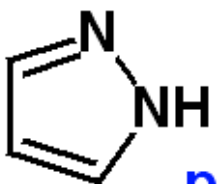
Dissociation of Acids and Protonation of Bases

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

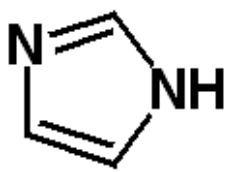
pK_a Values of Selected Organic Compounds



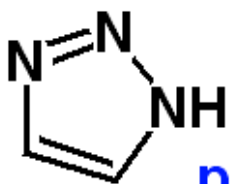
neutral



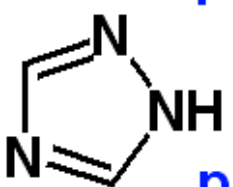
pK_a = 2.48



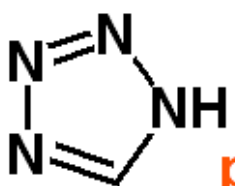
pK_a = 6.99



pK_a = 1.15



pK_a = 2.45

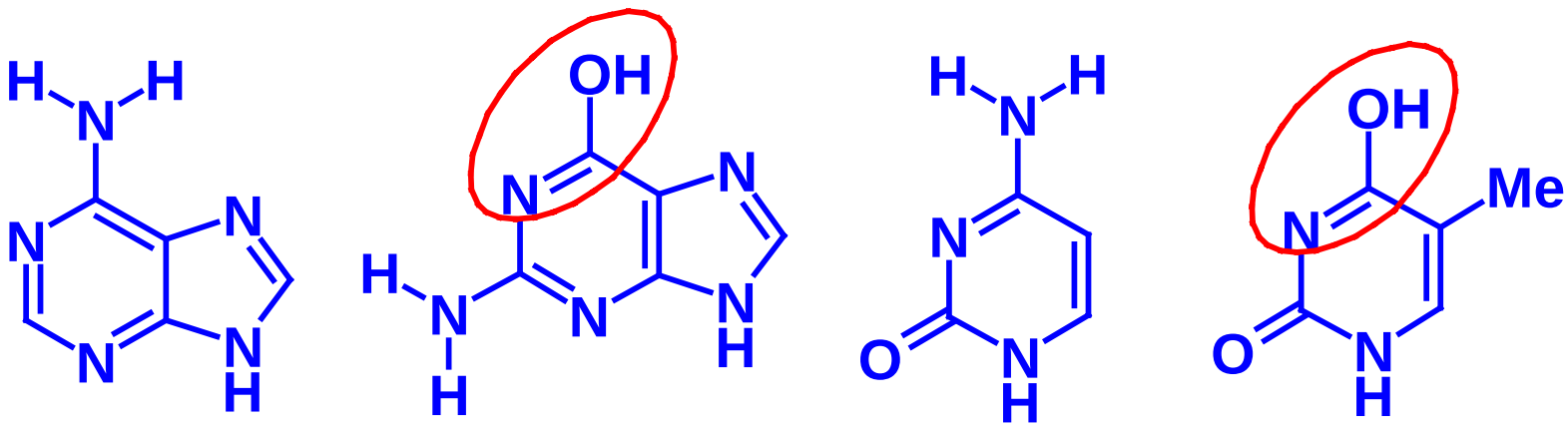


pK_a = 4.90

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

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



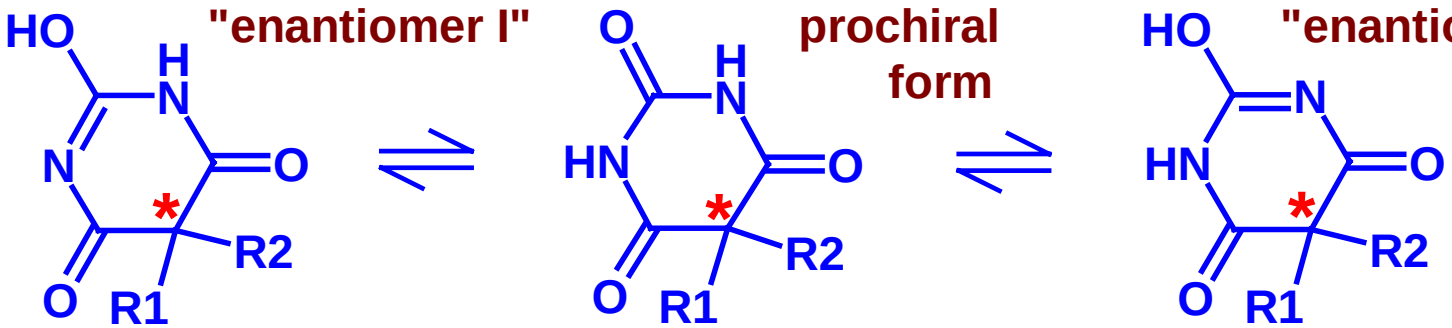
adenine guanine cytosine thymine

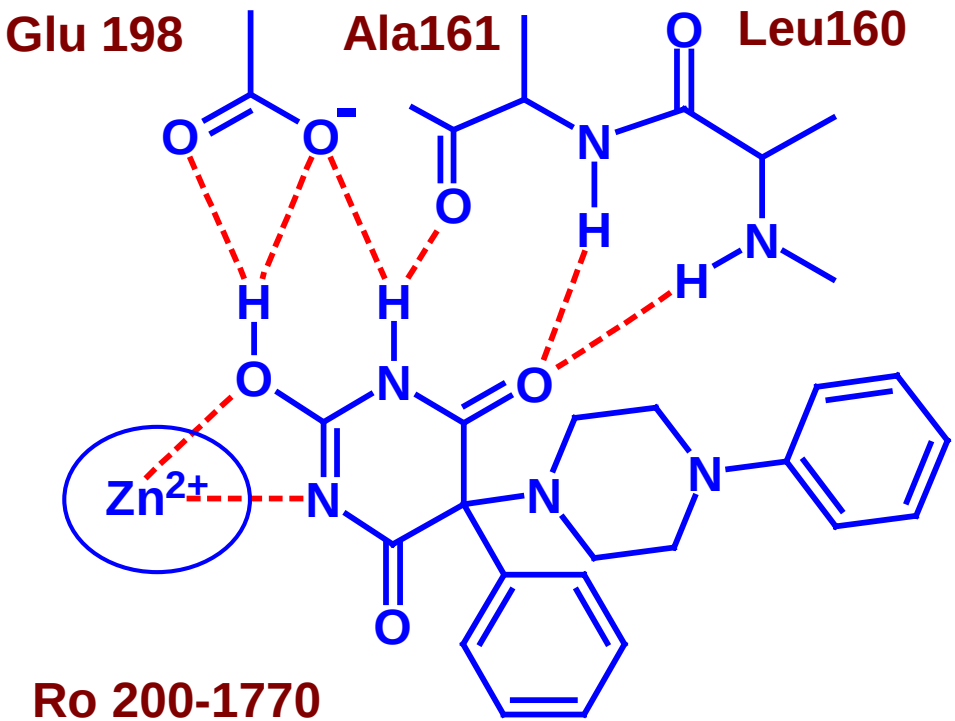
J. N. Davidson, The Biochemistry of Nucleic Acids, London, 1950

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**

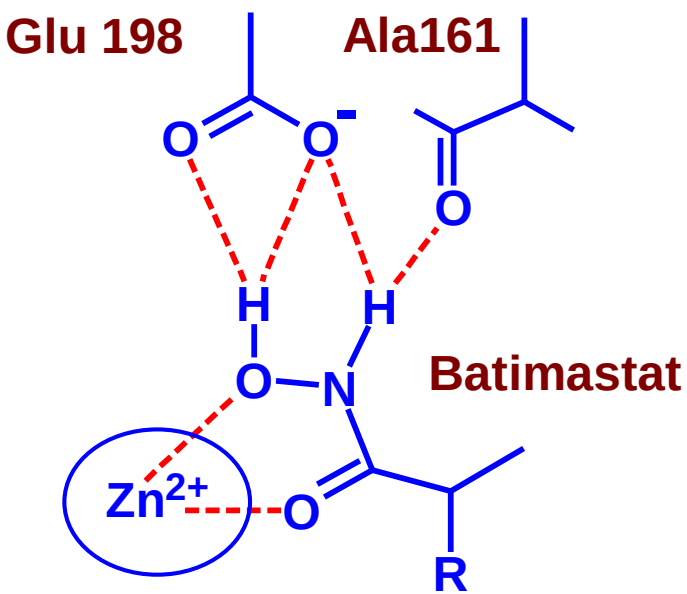
cited from: J. Watson and A. Berry, DNA. The Secret of Life, 2003

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





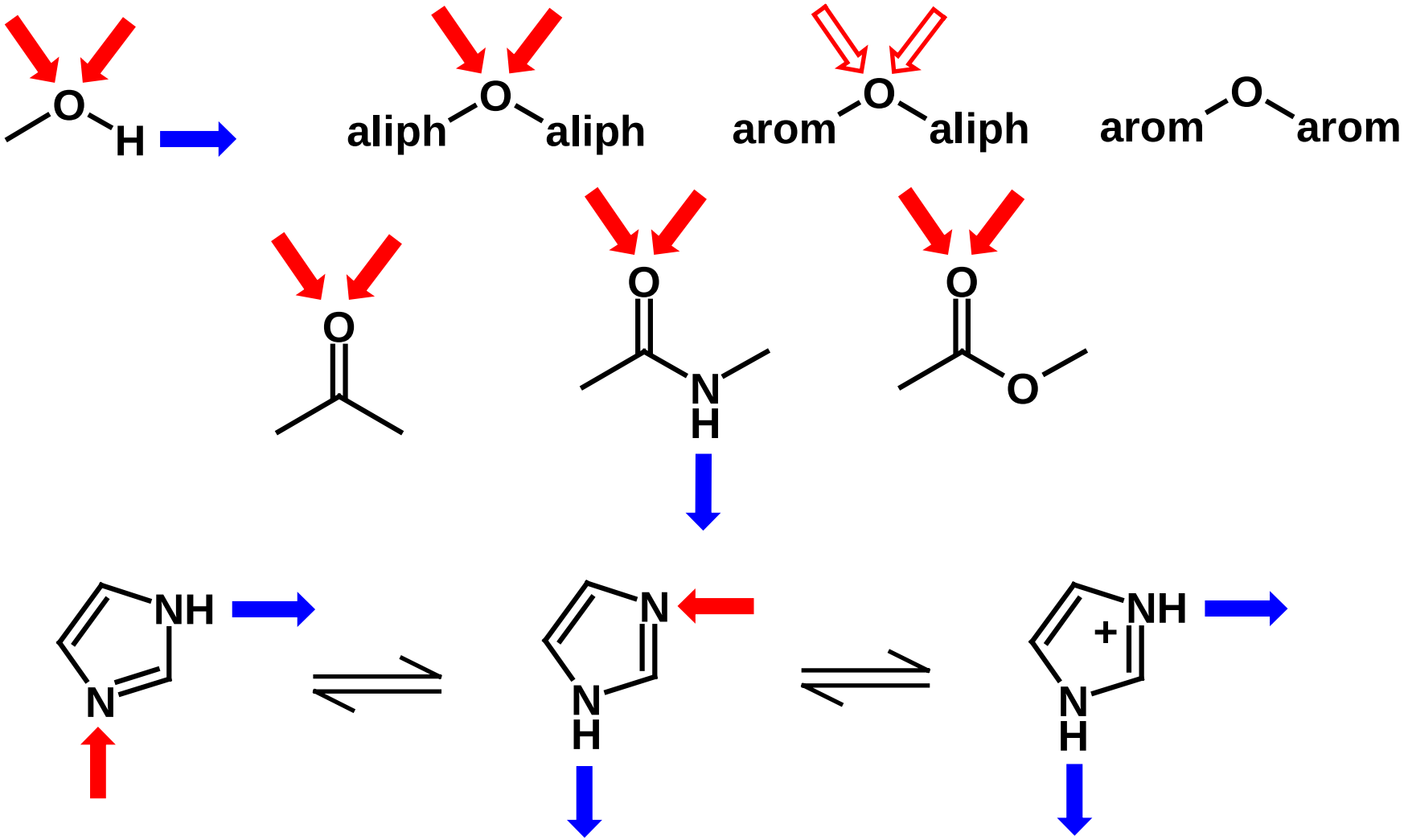
Ro 200-1770



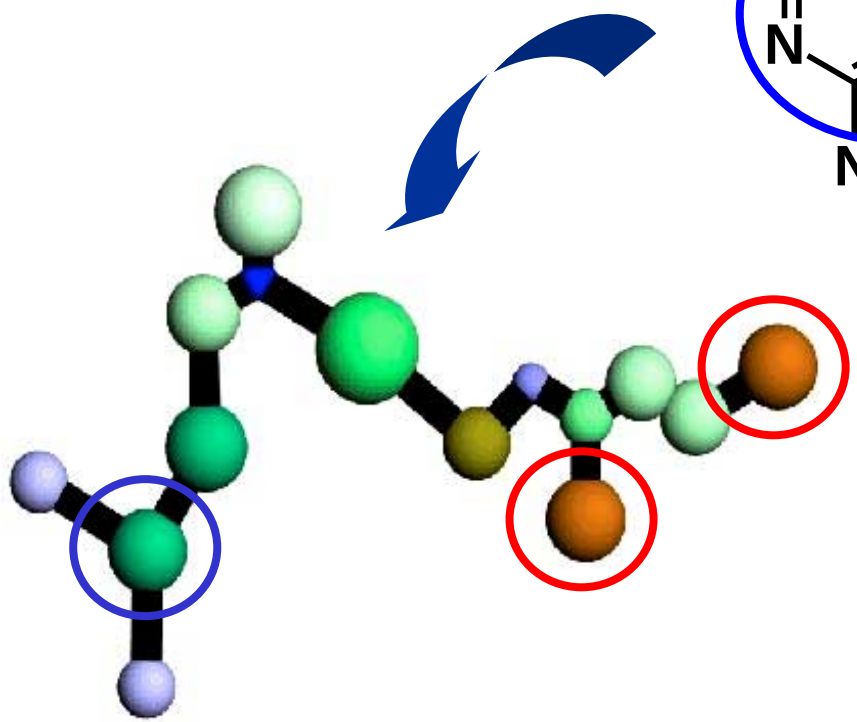
Batimastat

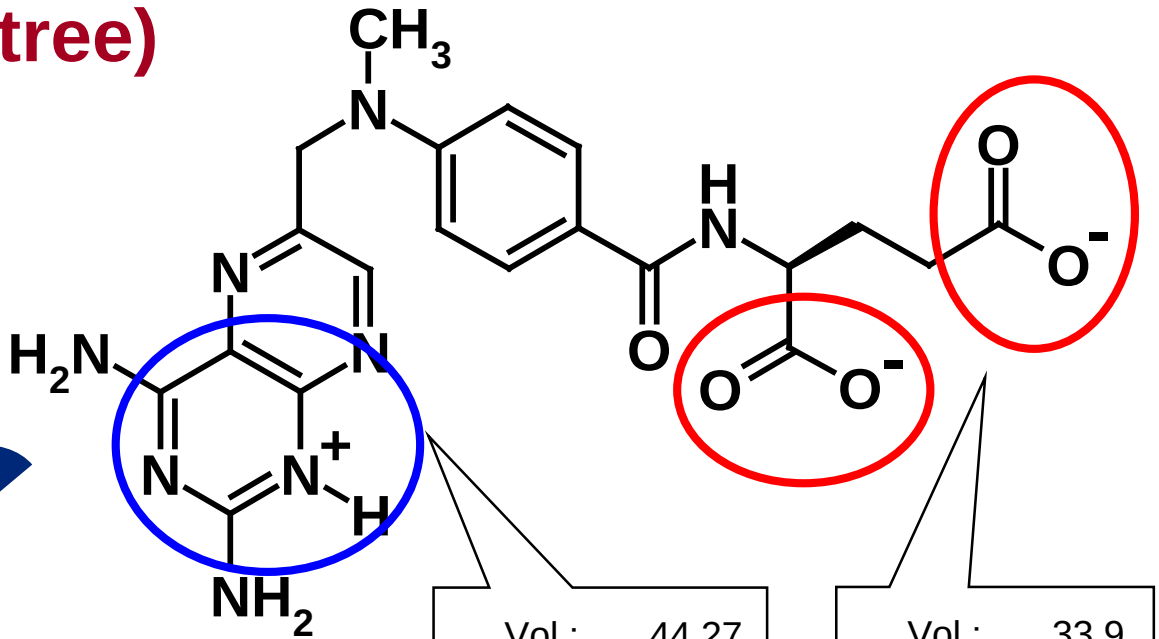
H. Brandstetter et al., J. Biol. Chem. 276, 17405-17412 (2001)

Donor and Acceptor Properties of O and N



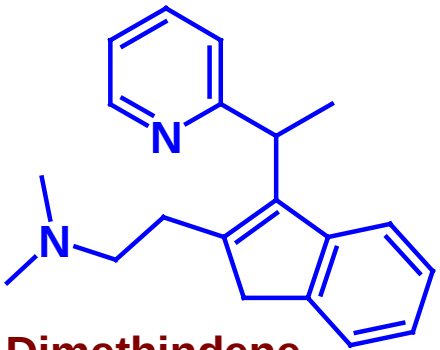
FTree (feature tree) Generation for Fast Similarity Searches



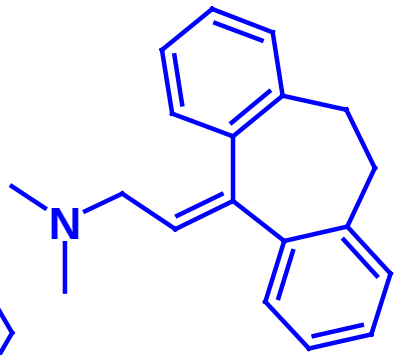


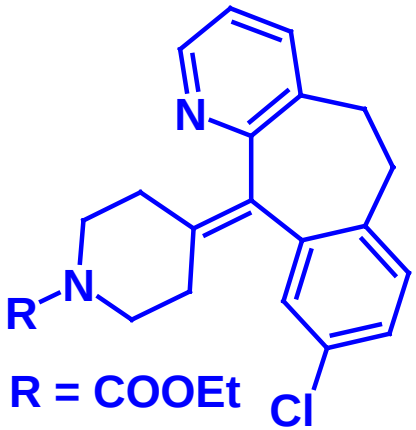
M. Rarey and J. S. Dixon, J. Comput.-Aided Mol. Design 12, 471-490 (1998);
M. Rarey and M. Stahl, J. Comput.-Aided Mol. Design 15, 497-520 (2001)

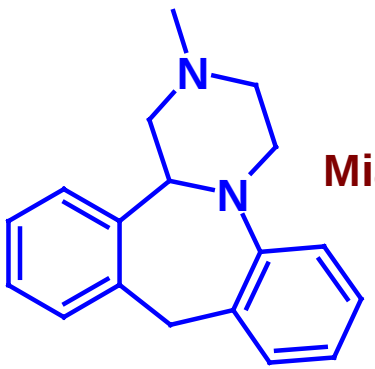
FTree Query Results for H1 Antagonists and Antidepressants



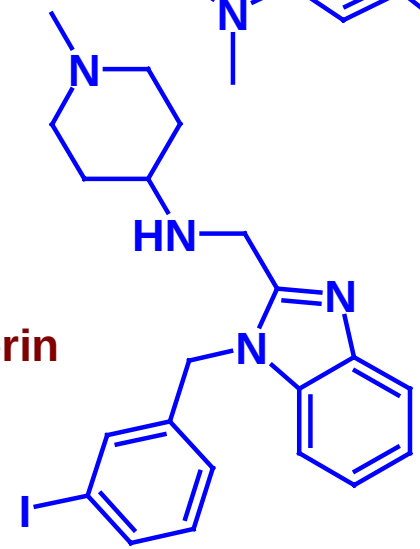
Dimethindene

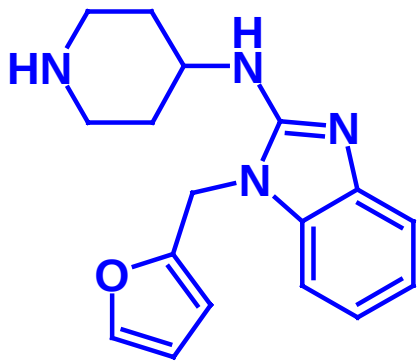






Mianserin





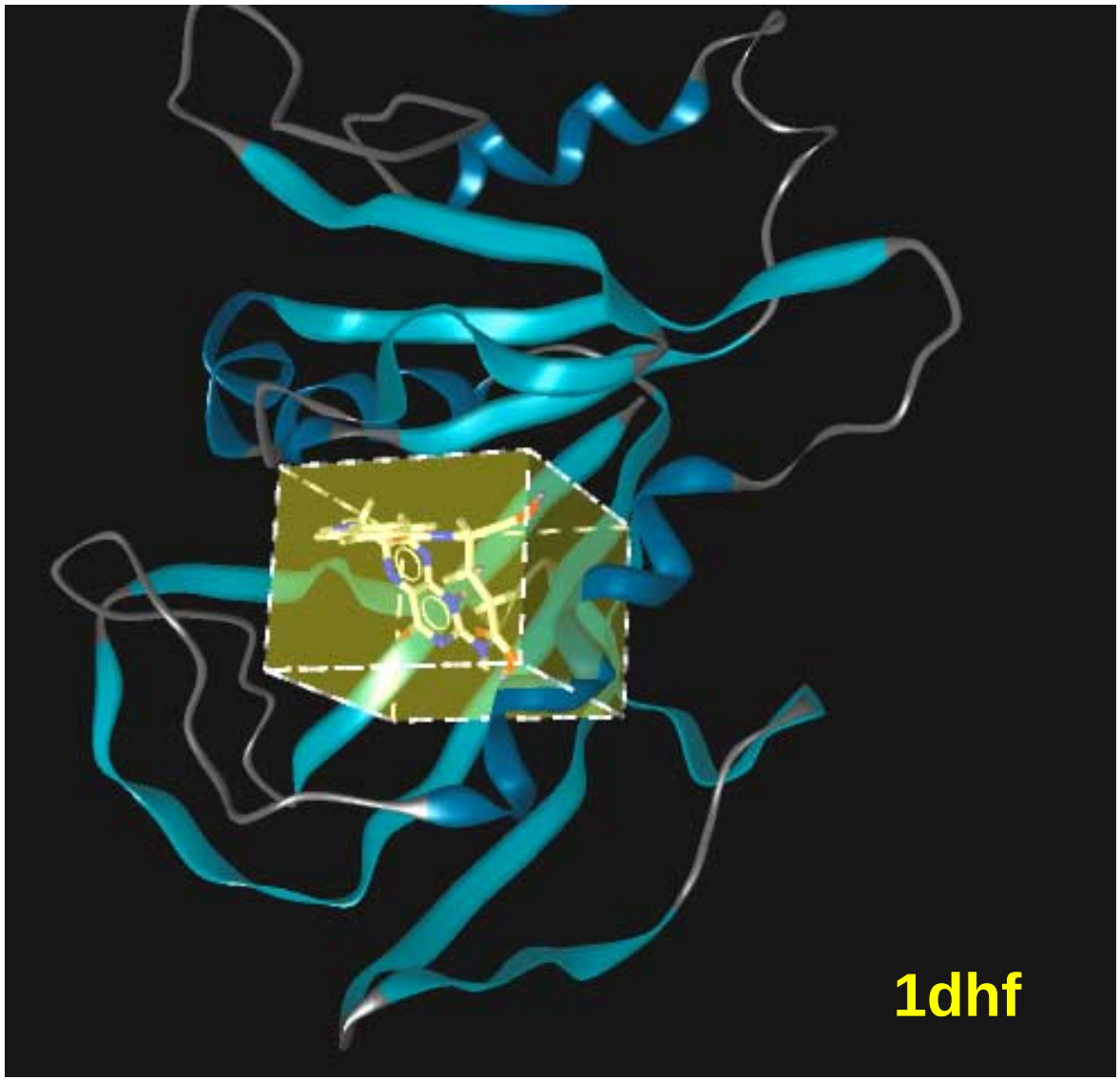
known actives

plausible hits

known actives related to the plausible hit

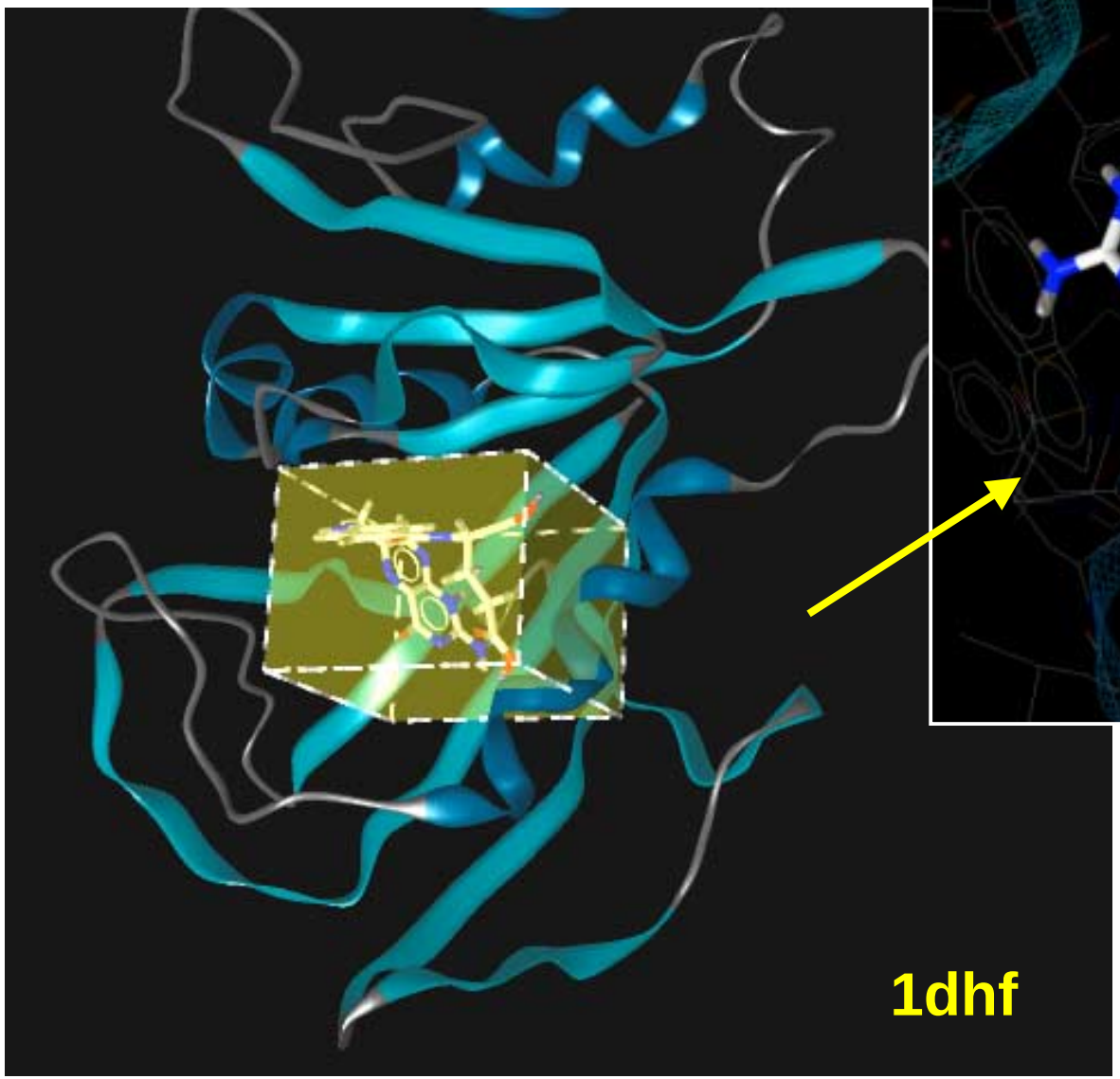
M. Rarey and M. Stahl, J. Comput.-Aided Mol. Design 15, 497-520 (2001)

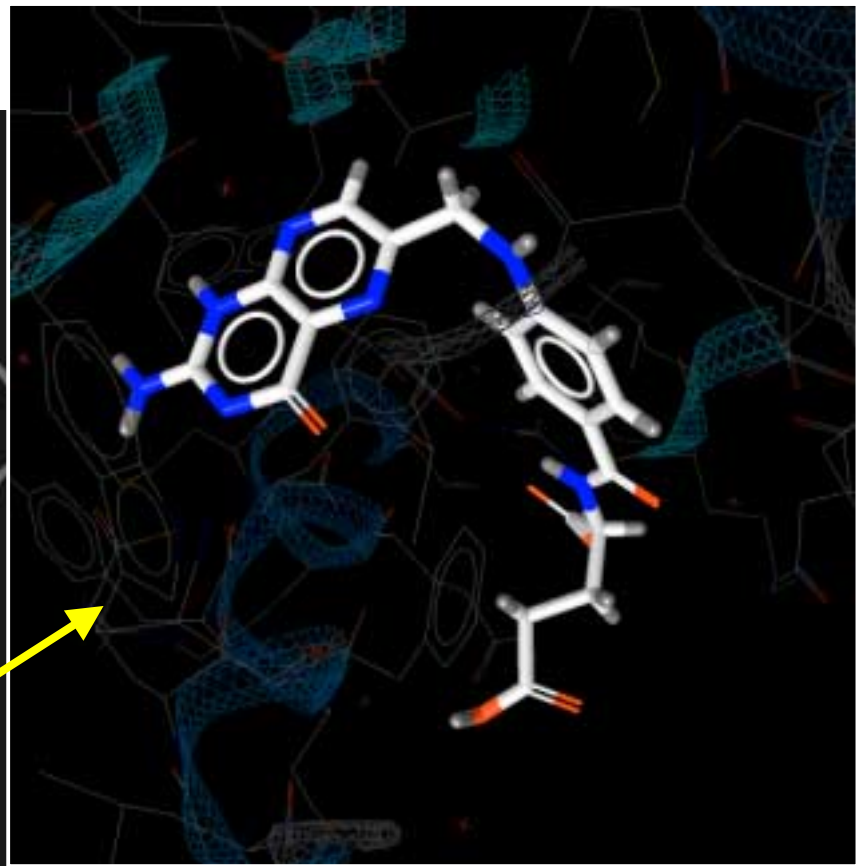
LigandScout (inte:ligand)



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

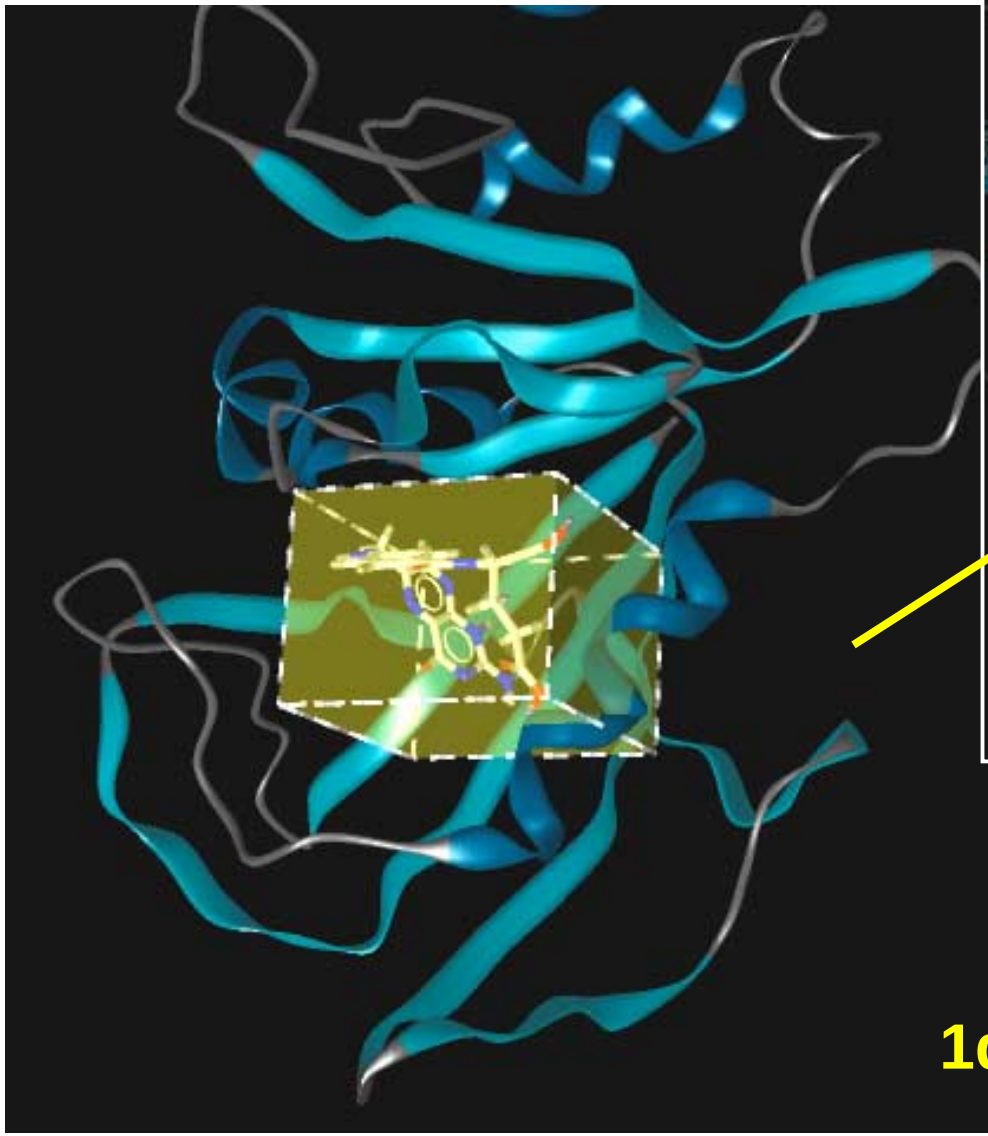
LigandScout (inte:ligand)

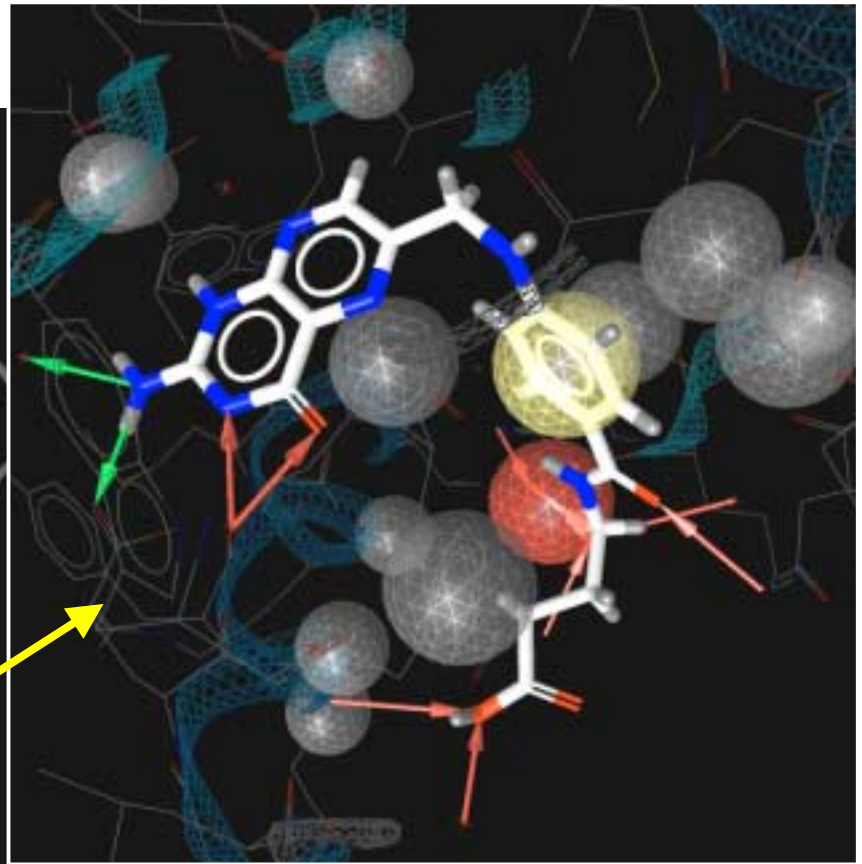




G. Wolber and T. Langer, J. Chem. Inf. Model. 45, 160-169 (2005)

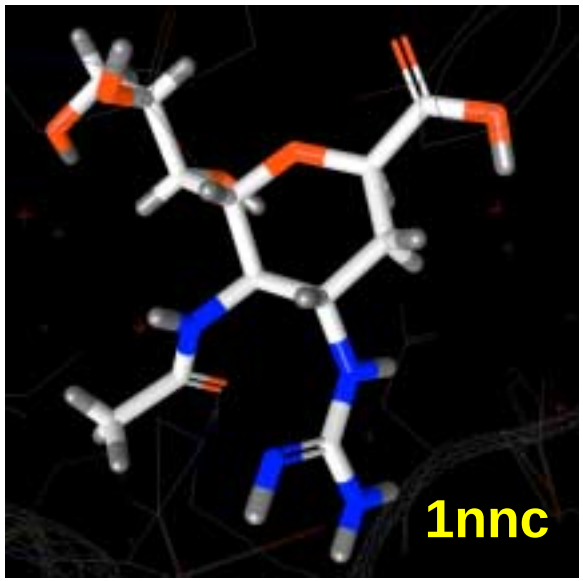
LigandScout (inte:ligand)

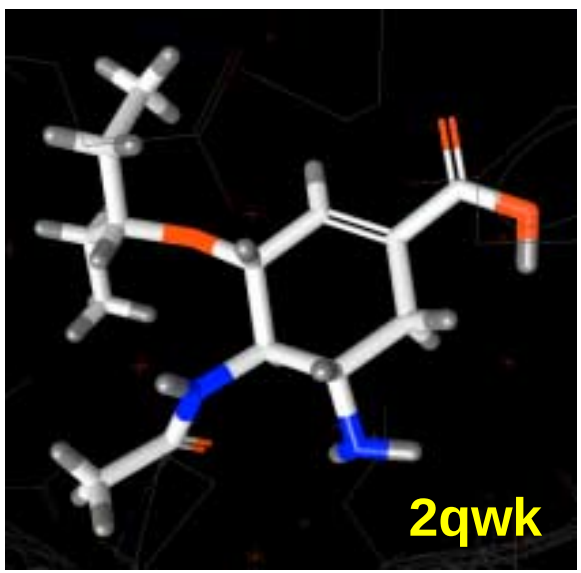


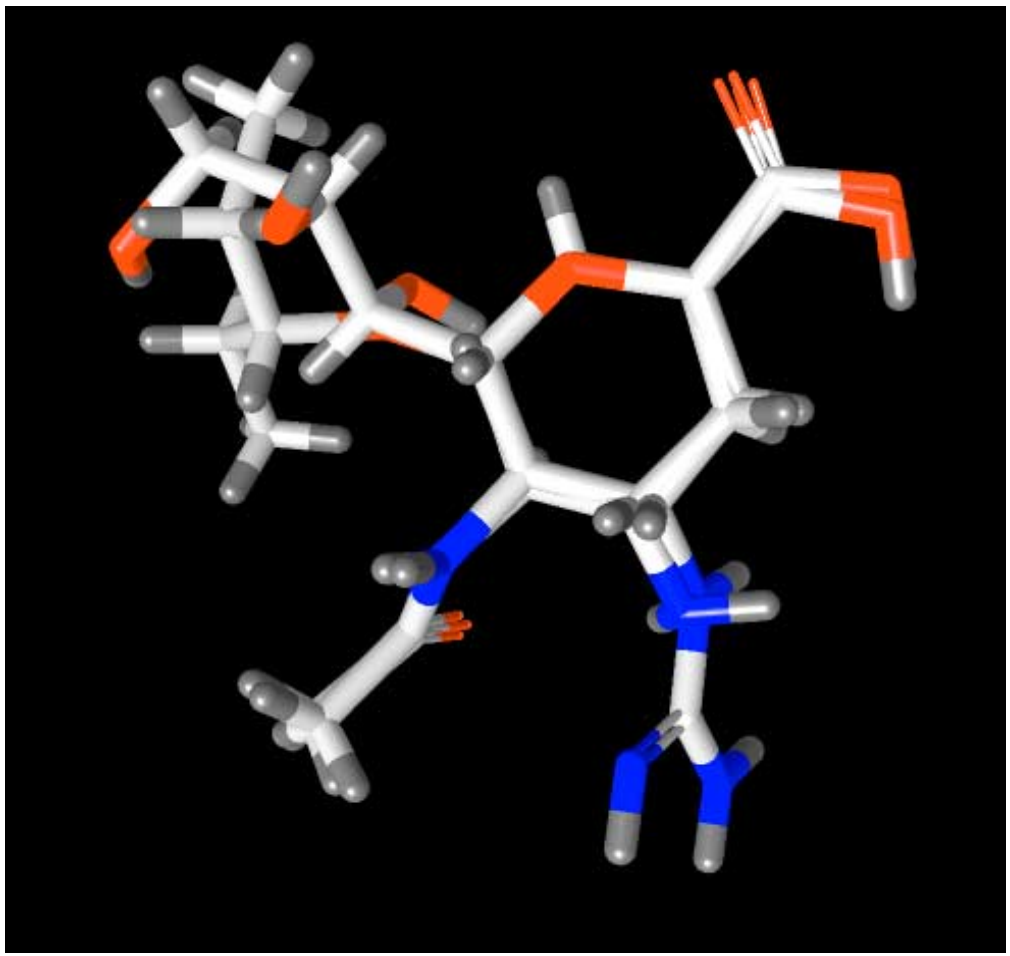


G. Wolber and T. Langer, J. Chem. Inf. Model. 45, 160-169 (2005)

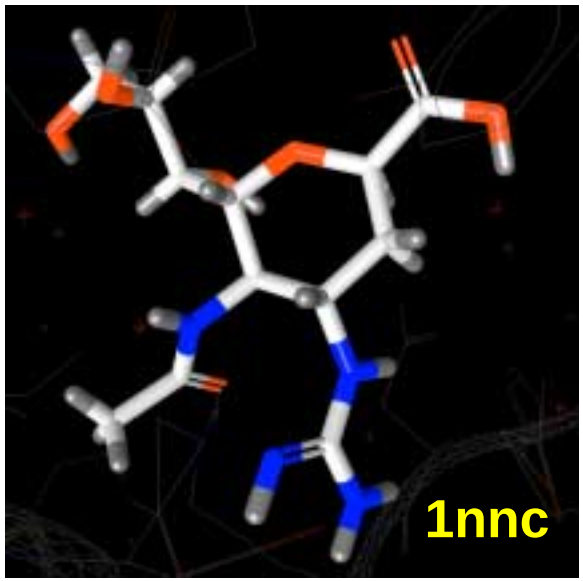
LigandScout Superposition: Zanamivir vs. GS 4071

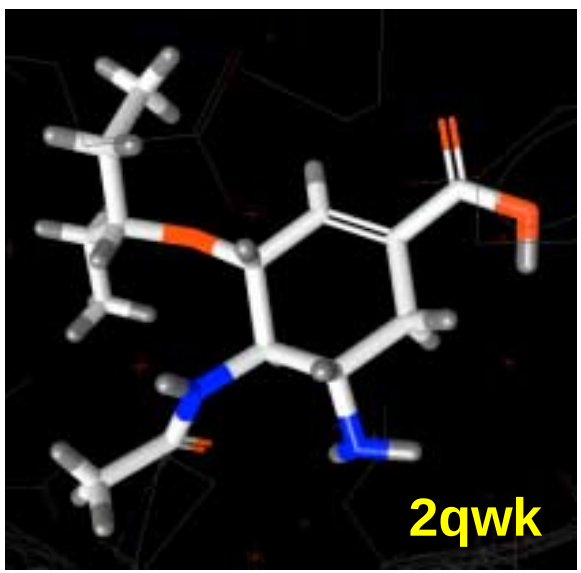


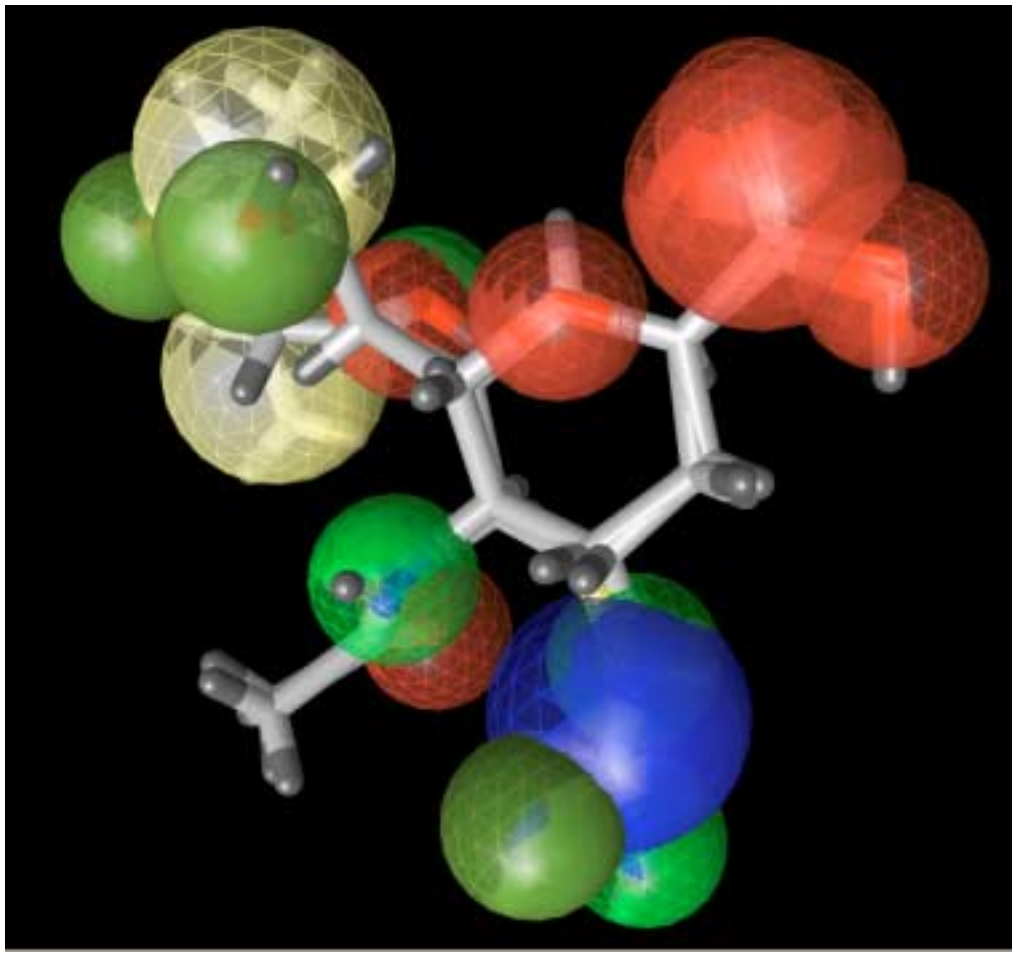




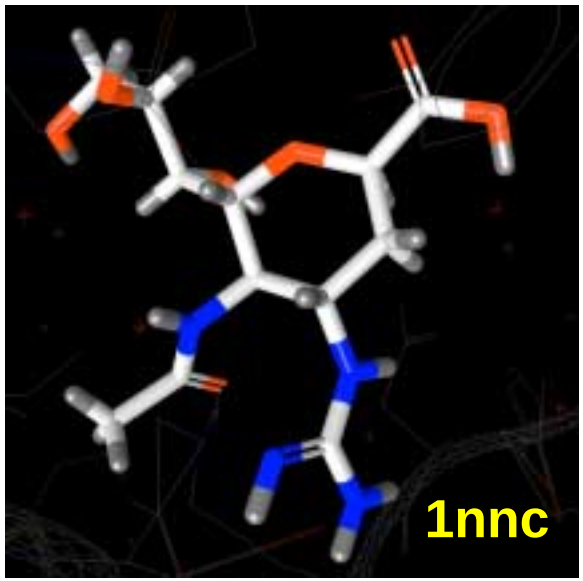
LigandScout Superposition: Zanamivir vs. GS 4071

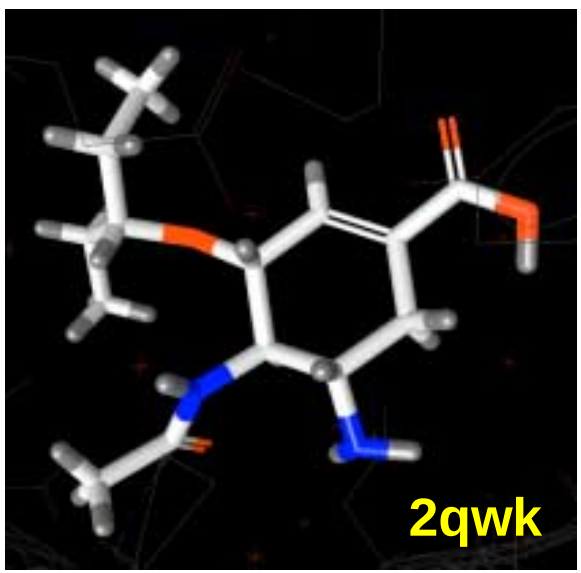


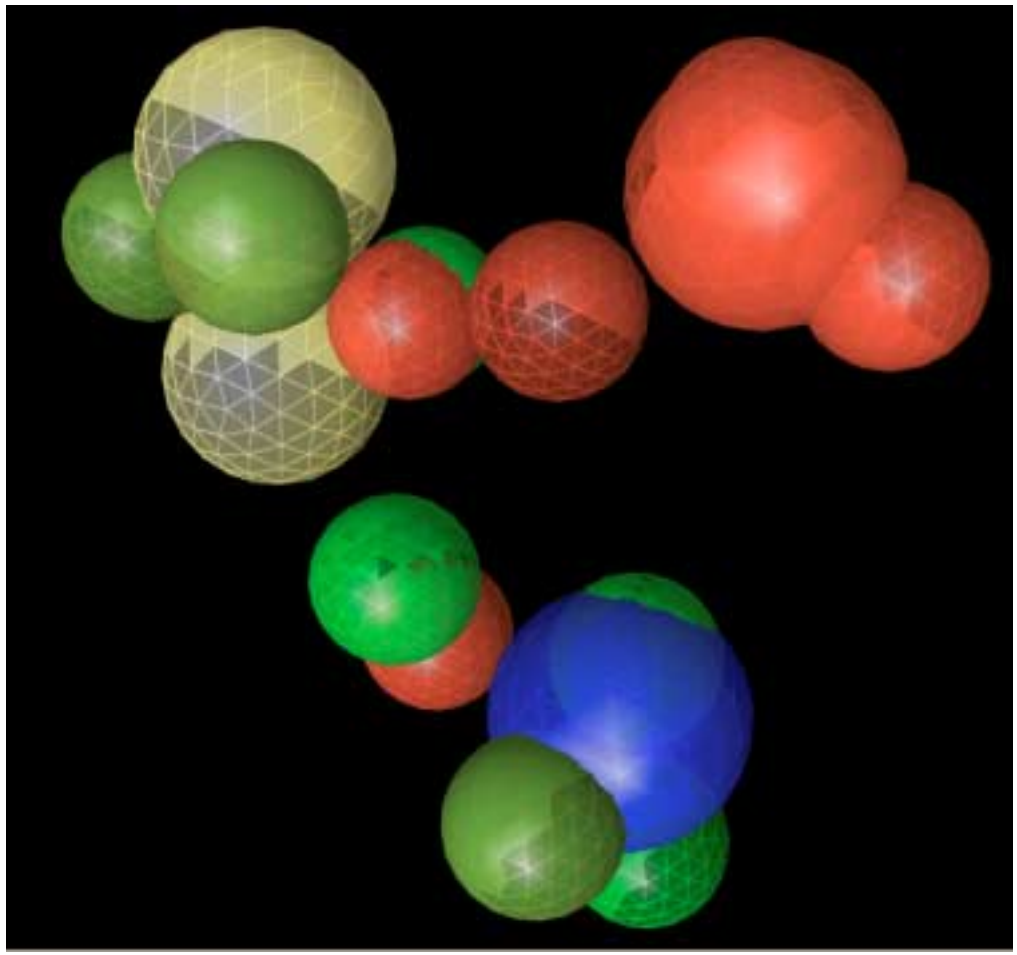




LigandScout Superposition: Zanamivir vs. GS 4071



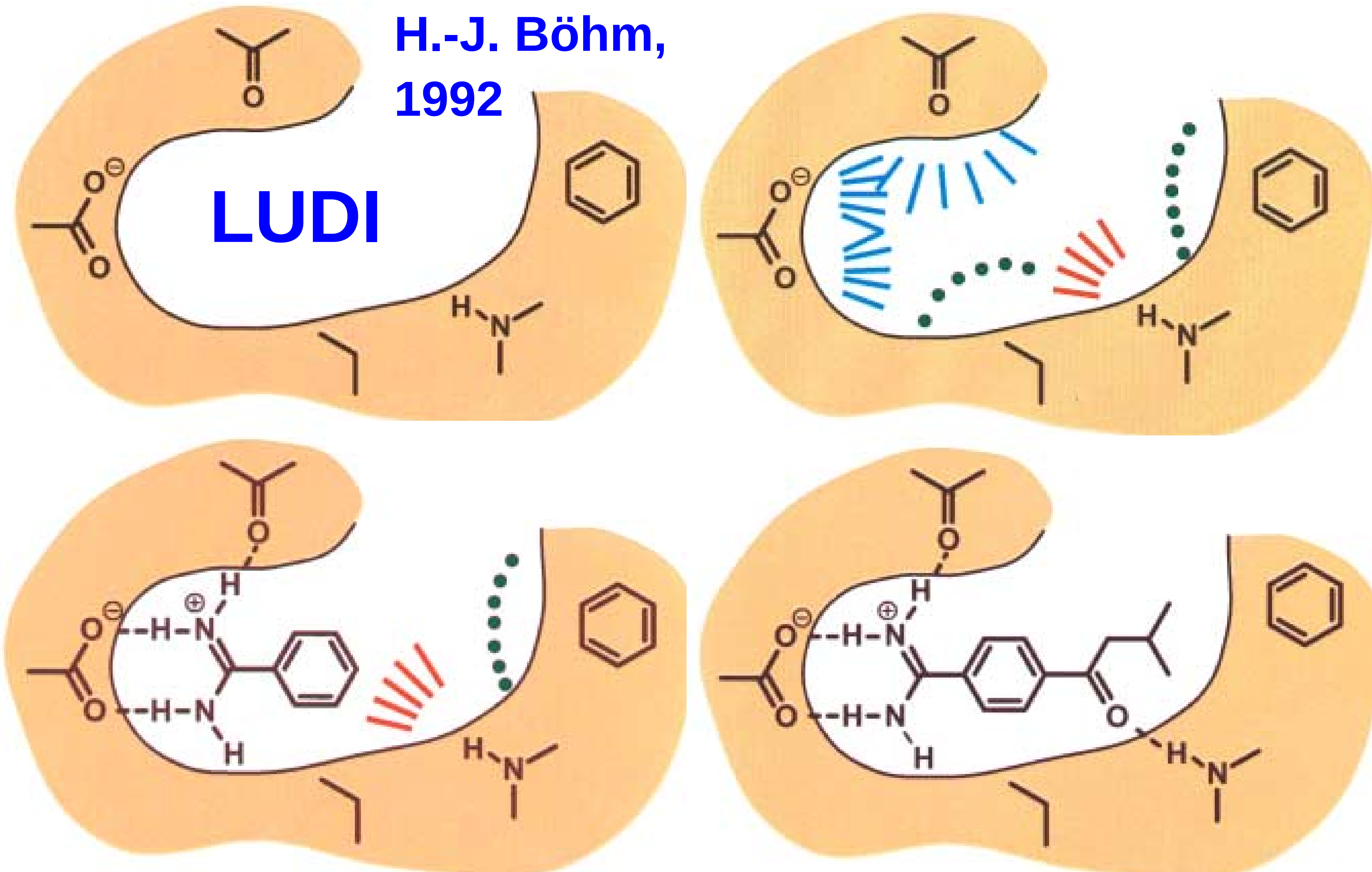




Computer-Aided Drug Design

H.-J. Böhm, 1992

LUDI



Problems in Docking and Scoring

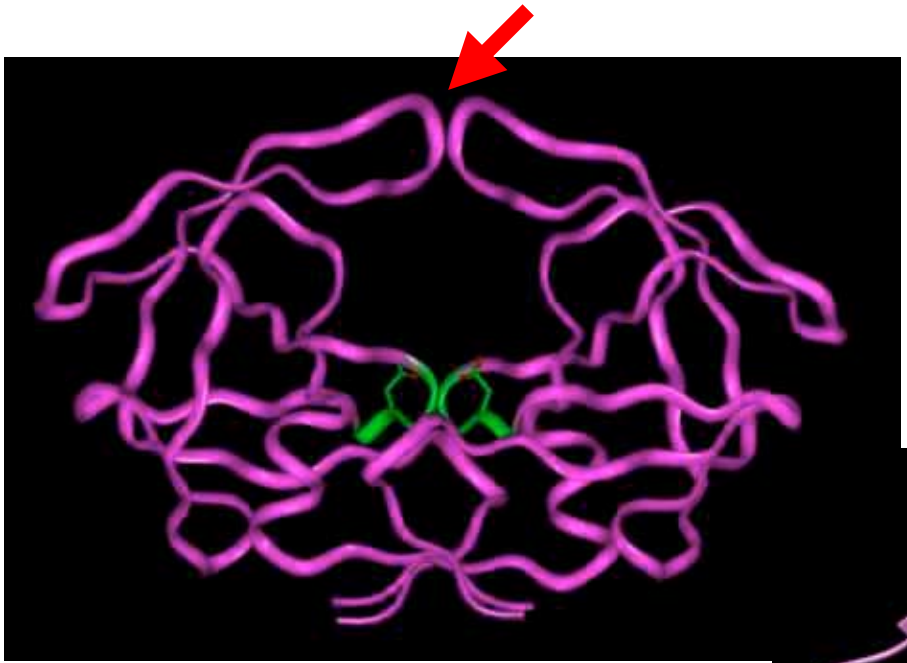
Pre-processing of the protein
lacking hydrogens, hydrogen bonds network, protonation states of his, lys, asp, glu

Pre-processing of the ligands
protonation states, tautomers

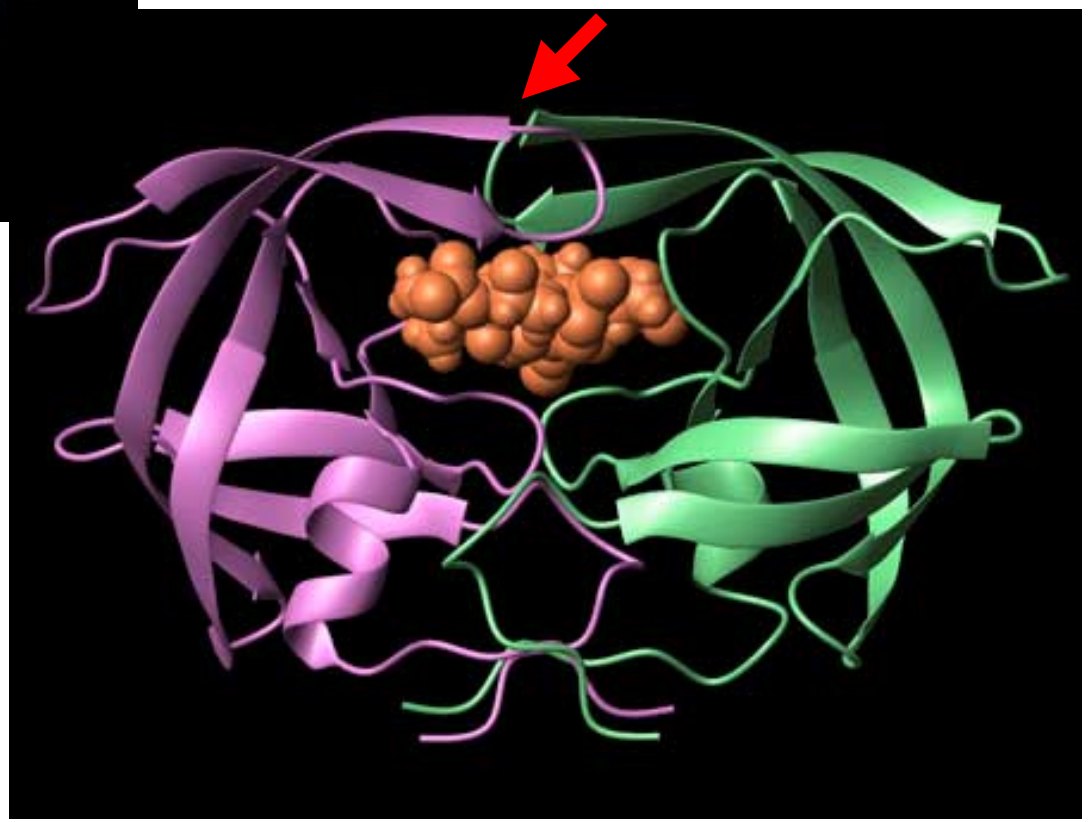
Flexibility of the ligand (no serious problem)

Flexibility of the protein / binding site (the real problem)

Fuzzy scoring functions (the biggest problem)

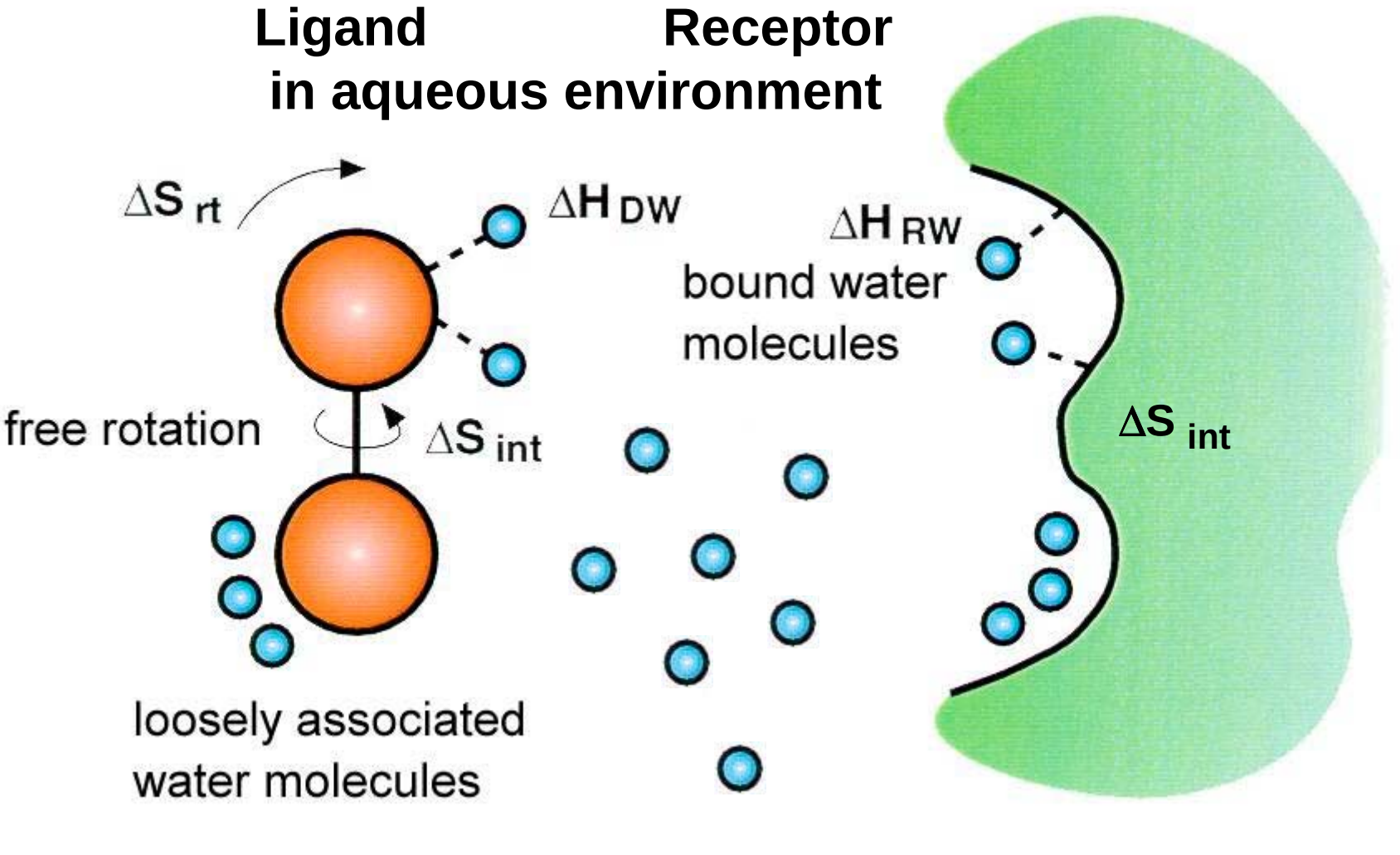


HIV Protease, without a ligand



HIV Protease, with a ligand

Consideration of Water, Flexibility and Mobility



Factors to be Considered in Scoring Functions

Desolvation enthalpy and entropy (ligand and protein)

Protonation state of the ligand and the binding site

Distortion energy of the ligand and its binding site

Loss of translational and rotational degrees of freedom of the ligand

MEP + dielectric constant at the binding site

Dipole moment of the ligand and local dipole moment at the binding site

Binding enthalpy of the ligand-protein complex

Repulsive effects (e.g. -O····O-)

Inserted water molecules

Solvation enthalpy and entropy of the complex

Virtual Screening vs. High-Throughput Screening

Comparison of the performance of high-throughput screening and virtual screening of potential leads of protein tyrosine phosphatase 1B (PTP1B):

a) High throughput screening of 400,000 compounds from a corporate collection → 300 hits < 300 µM,
85 validated hits with IC_{50} < 100 µM
= 0.021 % hit rate (many violate Lipinski rules)

b) Virtual screening of 235,000 commercially available compounds, using DOCK, version 3.5
→ 365 high-scoring molecules are tested
127 with IC_{50} < 100 µM
= 34.8% hit rate (hits are more drug-like?)

T. N. Doman et al., J. Med. Chem. 45, 2213-2221 (2002)

Stepwise Virtual Screening

560,000 compounds (subsection of AstraZeneca repository)

MW, rot-bond filter, presence of hinge region binding motif

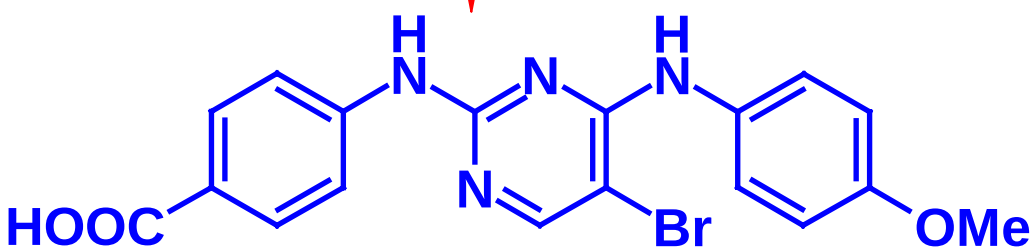
199,000 hits

FlexX-Pharm docking into ATP binding site

250 highest-scoring hits

visual inspection for unrealistic conformations

103 compounds tested, 36 hits in the range 110 nM to 68 μM



Checkpoint kinase 1 (Chk-1) inhibitor

IC_{50} = 450 nM

P. D. Lyne et al., J. Med. Chem. 47, 1962-1968 (2004)

Stepwise Virtual Screening

Aventis in-house compound repository

↓ MW, rot-bond filter, 3D pharmacophore search

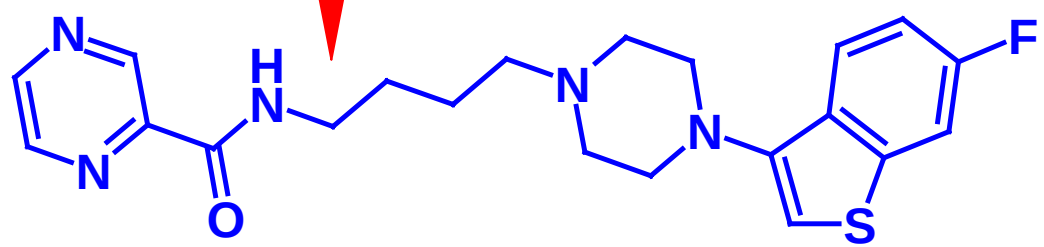
22,950 compounds

↓ docking into an α_{1A} receptor model (GOLD, PMF)

300 top-scoring compounds

↓ clustering, diversity selection

80 compounds tested, 37 hits with K_i < 10 μM



α_{1A} adrenergic receptor antagonist, K_i = 1.4 nM

A. Evers and T. Klabunde, J. Med. Chem. 48, 1088-1097 (2005)

Stepwise Virtual Screening

250,251 NCI compounds (3D database)

↓ 3D pharmacophore search

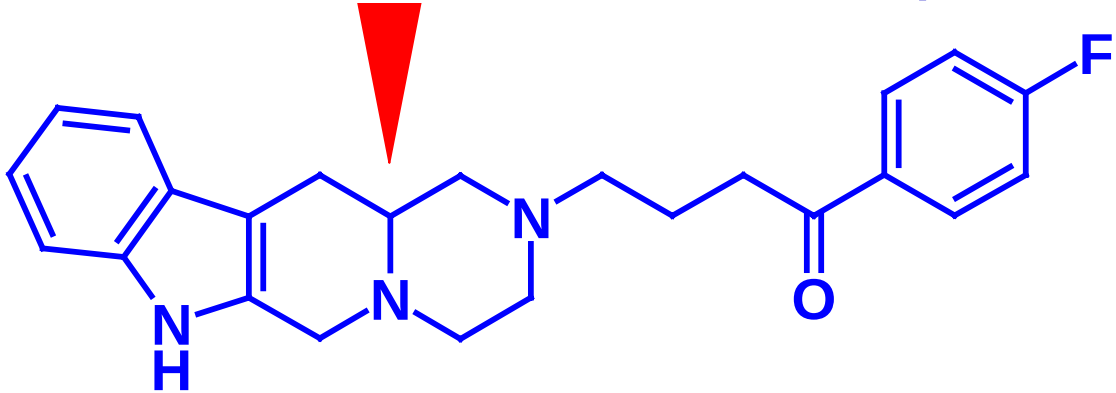
6,727 hits

↓ docking into four conformational clusters of a D_3 receptor homology model

2,478 potential ligands

↓ elimination of known chemotypes by similarity

20 compounds tested, 8 hits with K_i < 0.5 μM



dopamine D_3 receptor antagonist, K_i = 11 nM

J. Varady et al., J. Med. Chem. 46, 4377-4392 (2003)

Stepwise Virtual Screening

259,747 ACD compounds

Ro5 filter with MW < 350 and rot-bond < 9, presence of -COO^- or equivalent

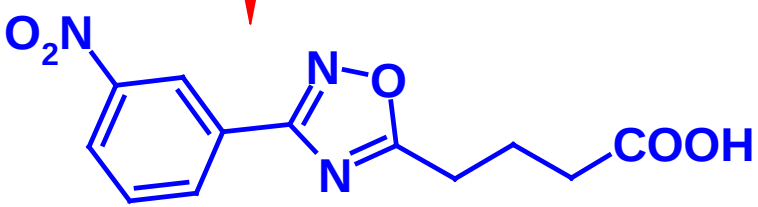
12,545 candidates

3D pharmacophore search (derived from binding site analysis)

1,261 hits

FlexX docking into 0.66 Å aldose reductase 3D structure

216 highest-scoring compounds, after clustering and visual inspection:
9 hits for biological testing



aldose reductase inhibitor, IC_{50} = 2.4 μM

O. Krämer et al., Proteins Struct. Funct. Genet. 55, 814–823 (2004)

Virtual Screening of Carbonic Anhydrase Inhibitors

98,850 compounds (LeadQuest and Maybridge libraries)

filter for Zn^{2+}-binding anchor groups

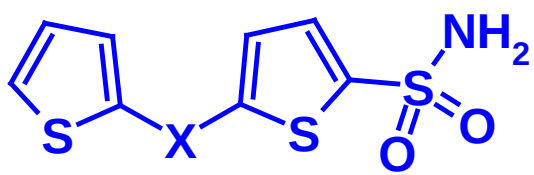
5,904 hits

2D and 3D pharmacophore searches (derived from binding site analysis)

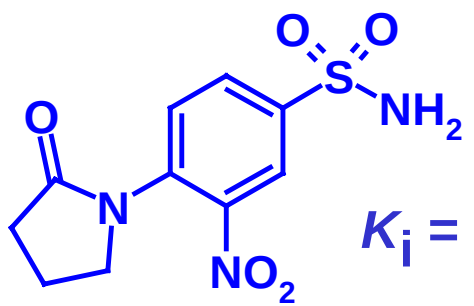
3,314 hits

FlexS superposition with dorzolamide, followed by FlexX docking of 100 hits into carbonic anhydrase binding site

13 hits



X = S
K_i = 0.9 nM
X = SO_2
K_i = 0.8 nM

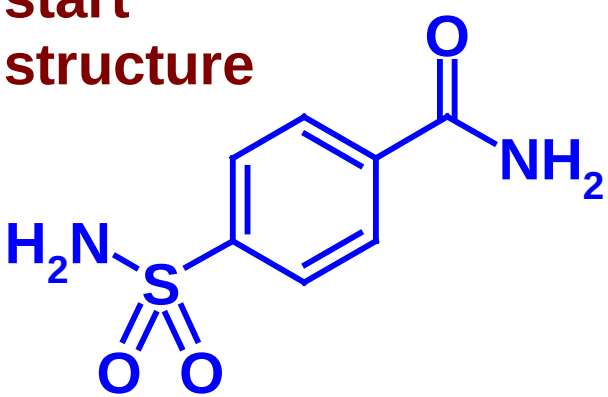


K_i = 0.6 nM

S. Grüneberg et al., Angew. Chem., Int. Ed. Engl. 40, 389-393 (2001);
S. Grüneberg et al., J. Med. Chem. 45, 3588-3602 (2002).

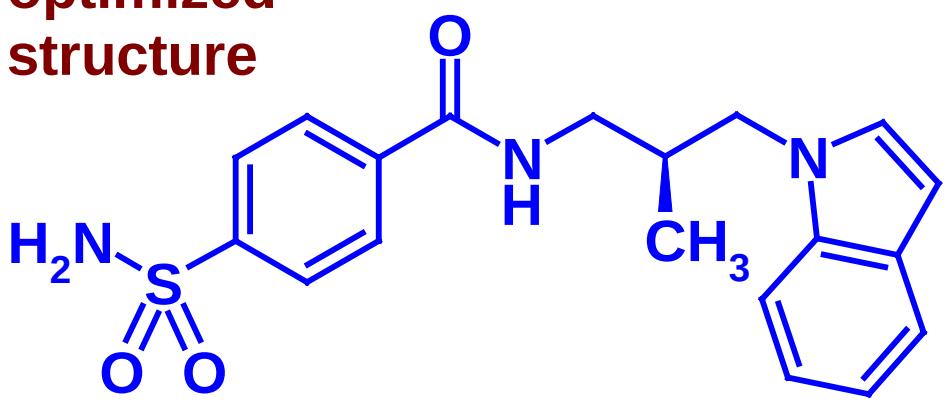
Combinatorial Design of Carbonic Anhydrase Inhibitors

start structure



K_d = 120 nM

optimized structure



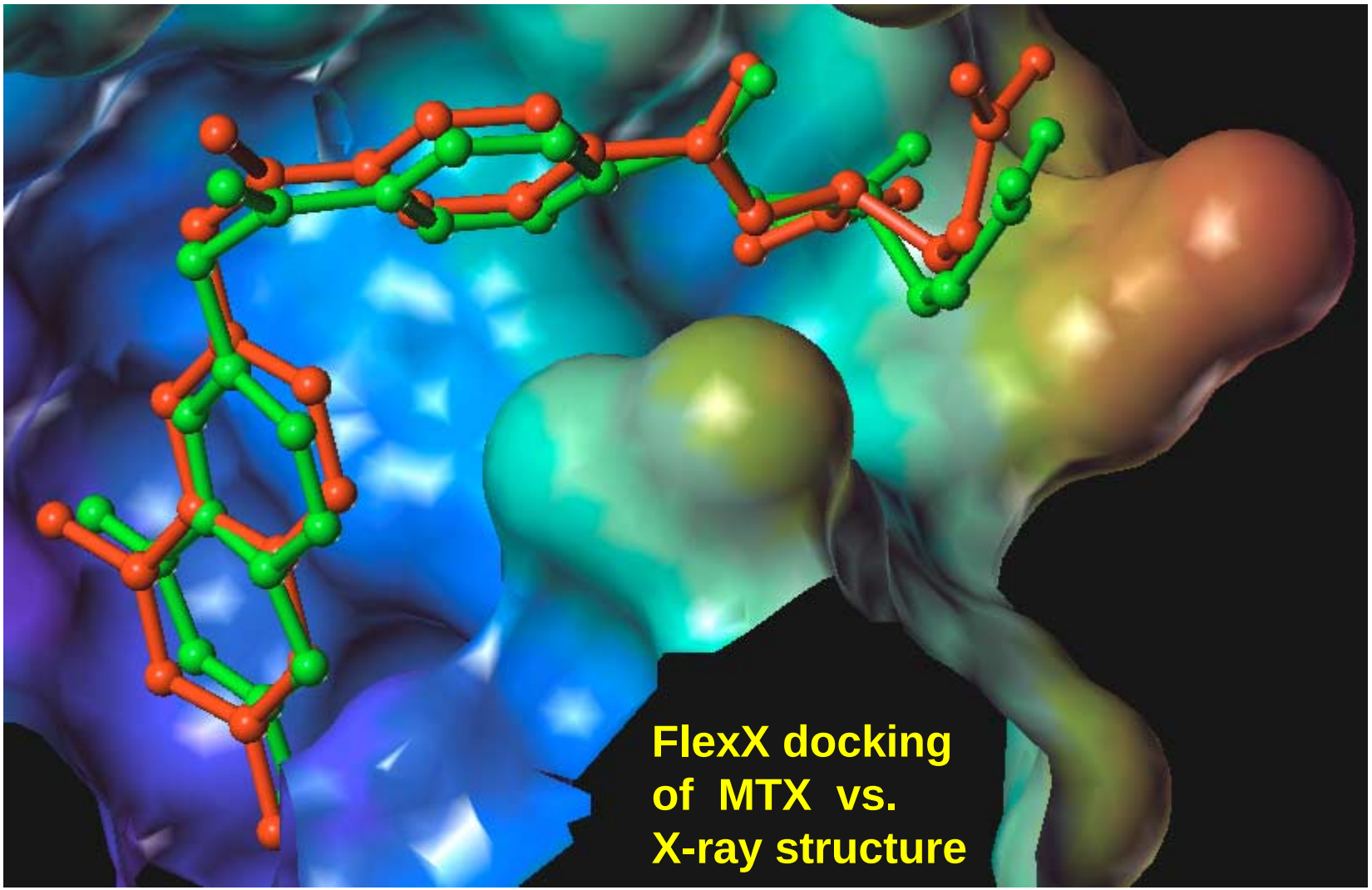
R enantiomer, K_d = 30 pM

(*S* enantiomer: K_d = 230 pM)

Program CombiSMoG, „best“ N-substituents from 100,000 candidates (20 scored by knowledge-based potentials)

B. A. Grzybowski et al., Acc. Chem. Res. 35, 261-269 (2002);
B. A. Grzybowski et al., Proc. Natl. Acad. Sci. USA 99, 1270-1273 (2002)

The Future: Combinatorial Drug Design



Summary and Conclusions

Virtual screening is a powerful tool to enrich libraries and compound collections

A proper preprocessing of the compound database is of utmost importance

Further experimental data and theoretical investigations are needed for better pK_a estimations and better scoring functions

Stepwise procedures (filters, pharmacophore searches, docking and scoring, visual inspection) are most efficient

Fragment-based approaches are a promising new strategy in lead structure search and optimization

Further progess needed in the understanding and scoring of ligand-receptor interactions

References

Böhm, H.-J., and Schneider, G., Eds., Virtual Screening for Bioactive Molecules (Volume 10 of Methods and Principles in Medicinal Chemistry, Mannhold, R., Kubinyi, H., and Timmerman, H., Eds.), Wiley-VCH, Weinheim, 2000.

Klebe, G., Ed., Virtual Screening: An Alternative or Complement to High Throughput Screening, Kluwer Academic Publ, Dordrecht, 2000; also published in Persp. Drug Discov. Design 20, 1-287 (2000).

Böhm, H.-J., and Schneider, G., Eds., Protein-Ligand Interactions. From Molecular Recognition to Drug Design, (Volume 19 of Methods and Principles in Medicinal Chemistry,, Mannhold, R., Kubinyi, H., and Folkers, G., Eds.), Wiley-VCH, Weinheim, 2003.

Alvarez, J., and Shoichet, B., Eds., Virtual Screening in Drug Discovery, CRC Press, Taylor & Francis Group, Boca Raton, FL, USA, 2005.

T. Langer and R. Hoffmann, Pharmacophores and Pharmacophore Searches (Volume 32 of Methods and principles in medicinal chemistry, R. Mannhold, H. Kubinyi and G. Folkers, Eds), Wiley-VCH, Weinheim, 2006.

H. Kubinyi, Success Stories of Computer-Aided Design, in: Computer Applications in Pharmaceutical Research and Development, S. Ekins, Ed. (Wiley Series in Drug Discovery and Development, B. Wang, Ed.), Wiley-Interscience, New York, 2006, pp. 377-424.

G. Klebe, Virtual ligand screening: strategies, perspectives and limitations, Drug Discov. Today 11, 580-594 (2006).