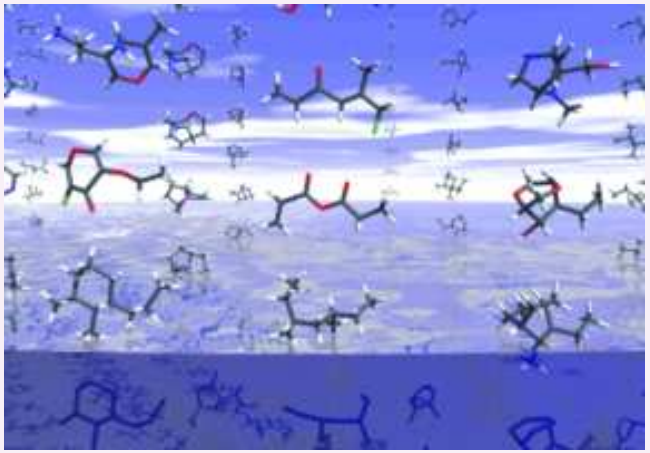
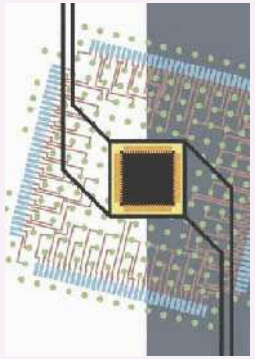
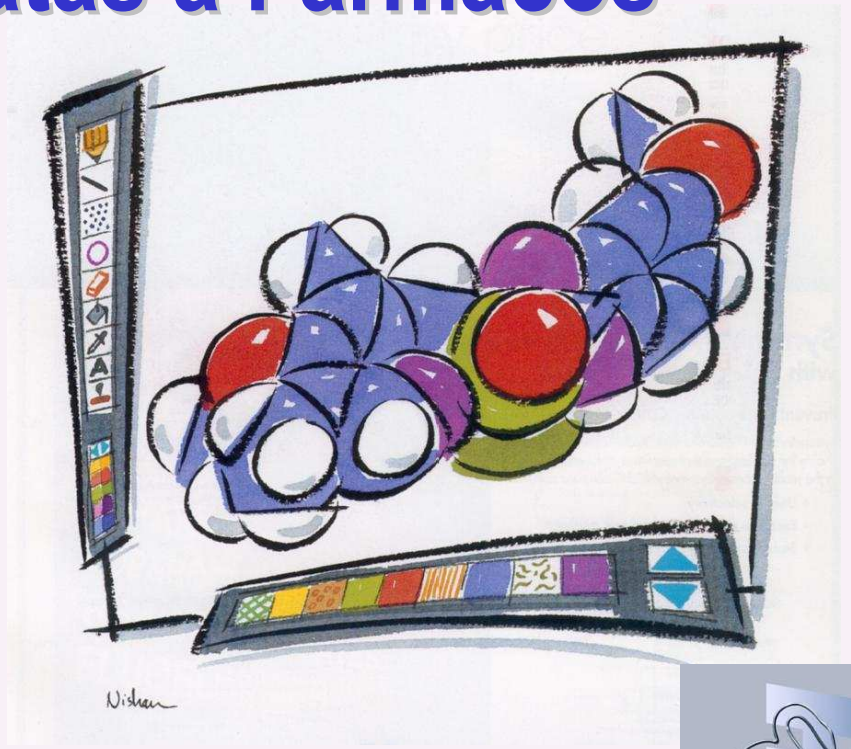
El Cribado Virtual de Quimiotecas como Herramienta útil para la Identificación de Moléculas Candidatas a Fármacos





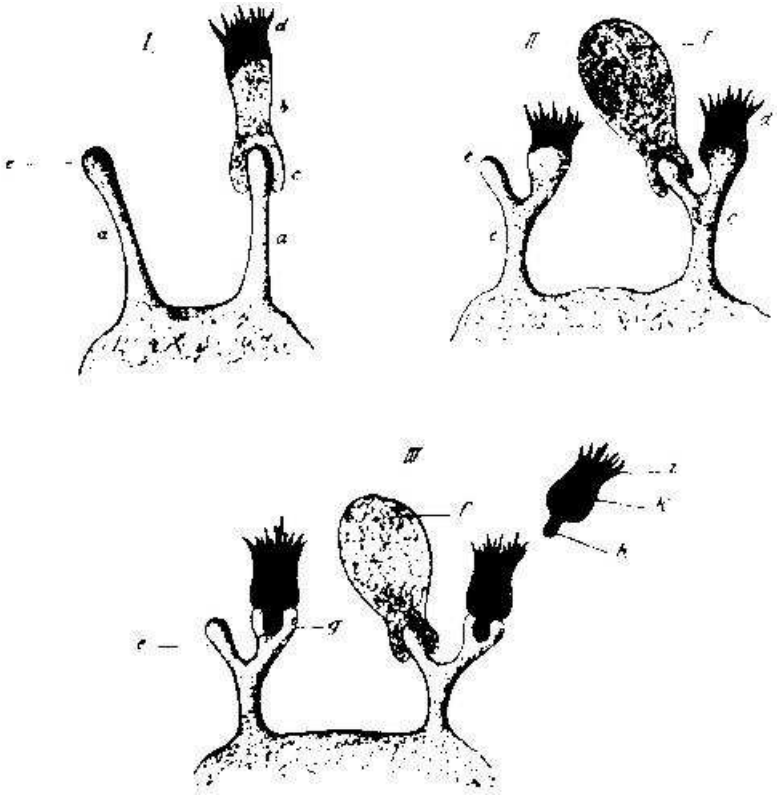


Antonio Morreale

CBM Severo Ochoa
CENTRO DE BIOLOGÍA MOLECULAR
UAM CSIC

Federico Gago
Departamento de Farmacología
Universidad de Alcalá, Madrid

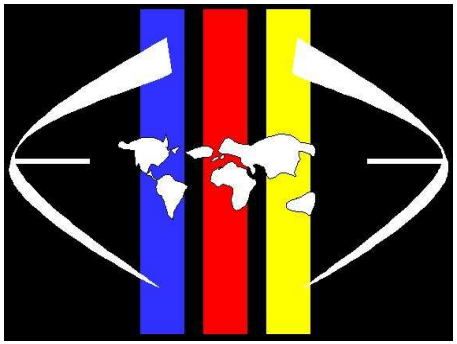
“Corpora non agunt nisi fixata”



Paul Ehrlich

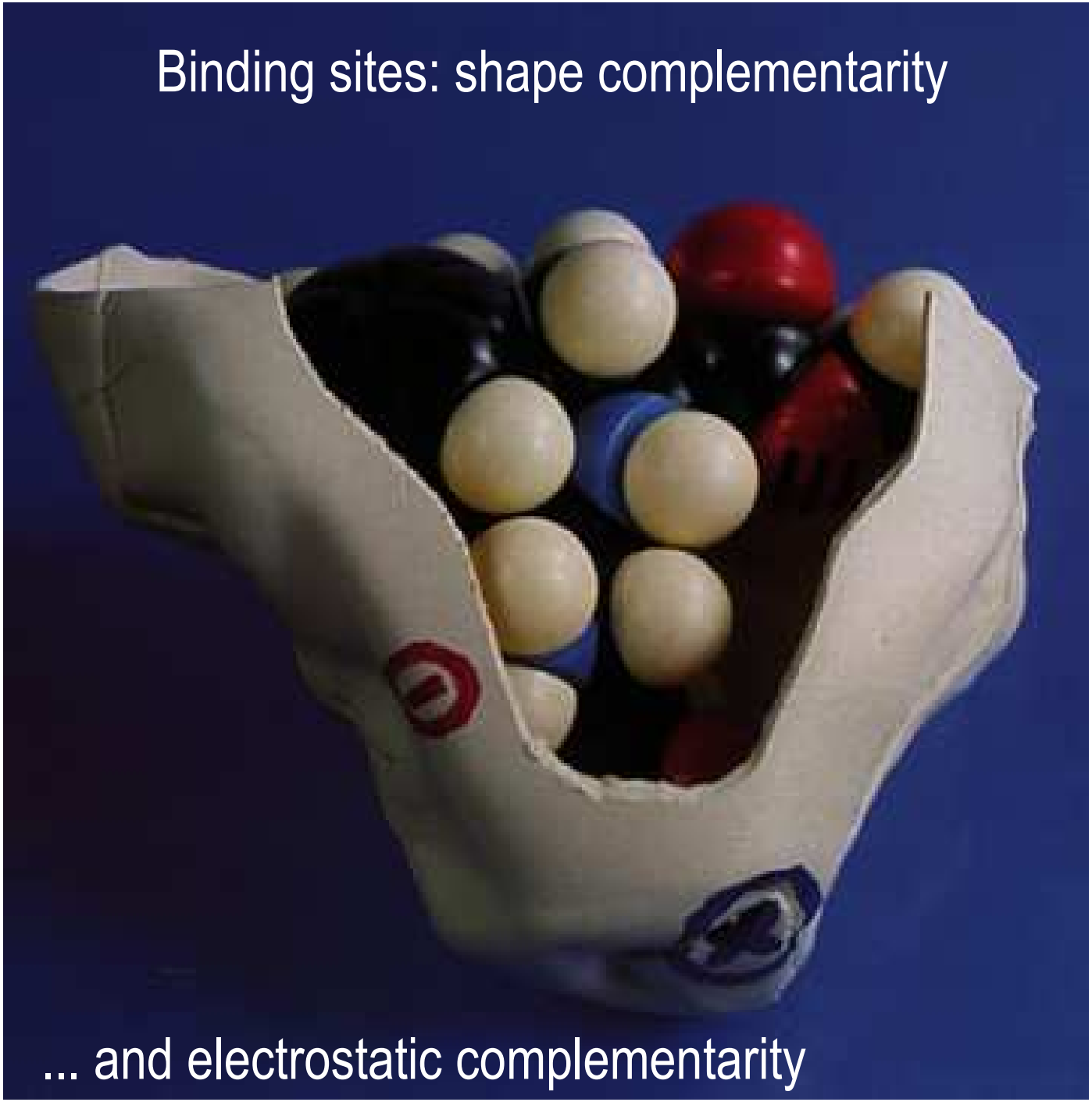
“Address in Pathology on Chemotherapeutics: Scientific Principles, Methods, and Results”

Lancet II, 445 (1913)

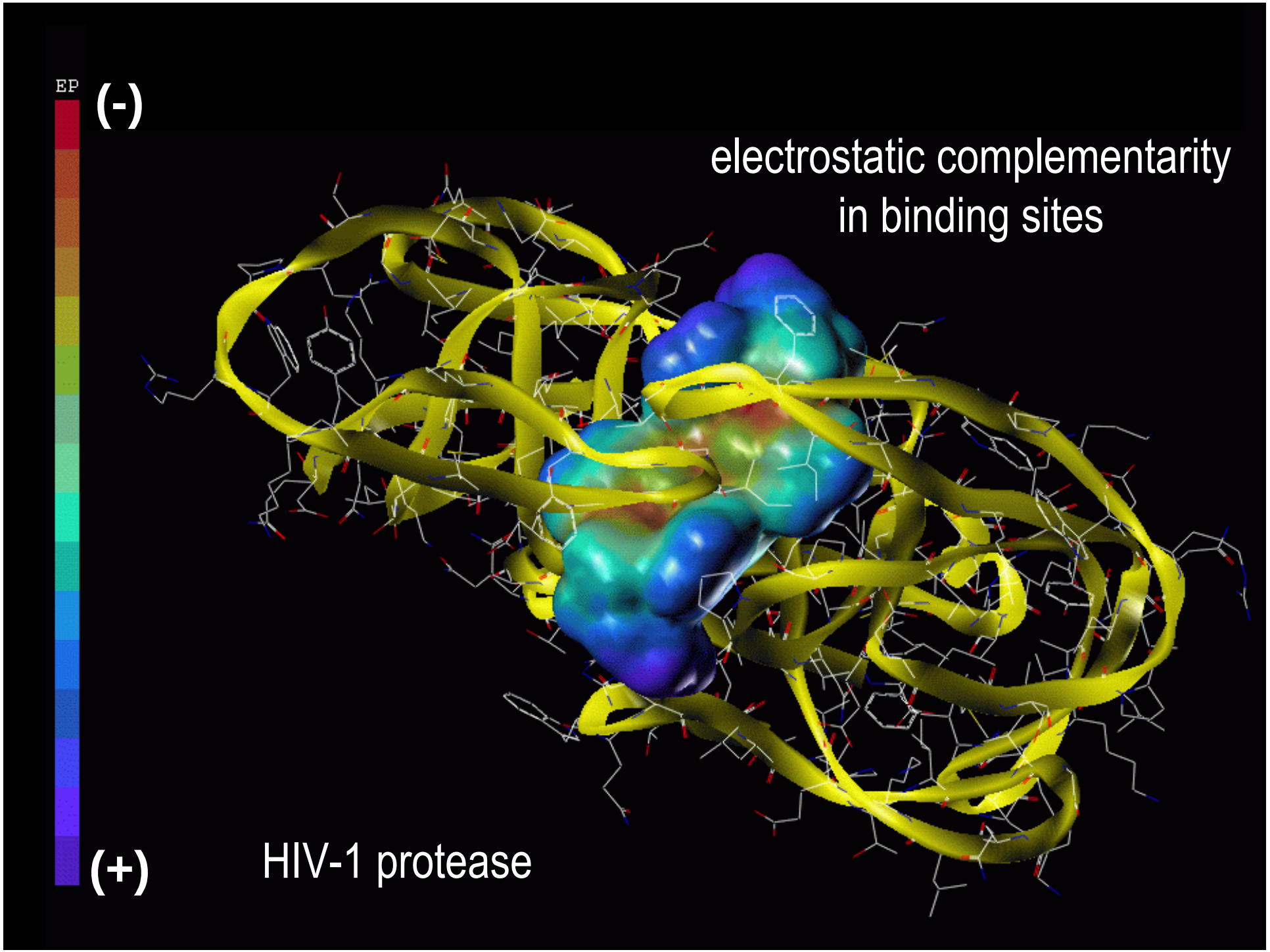


MAGIC BULLETS

Binding sites: shape complementarity



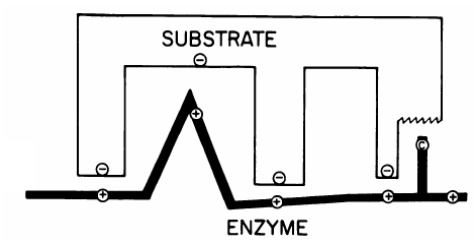
... and electrostatic complementarity

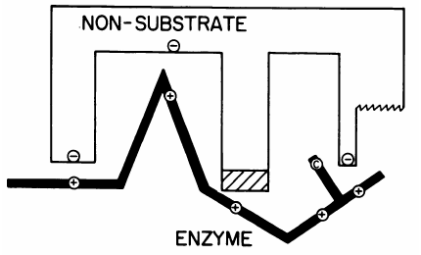

EP
(-)
(+)
electrostatic complementarity
in binding sites
HIV-1 protease

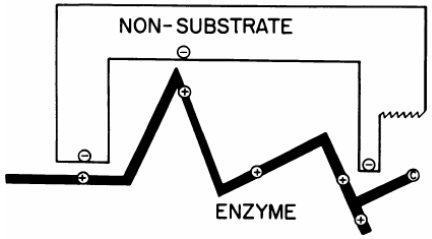
Historical perspective of ligand binding

1894: Emil Fischer's LOCK and KEY rationale

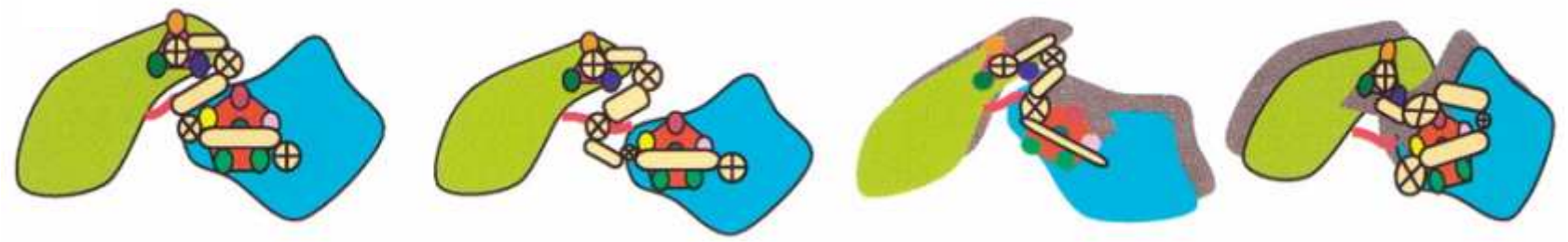
1958: Daniel Koshland's INDUCED FIT hypothesis

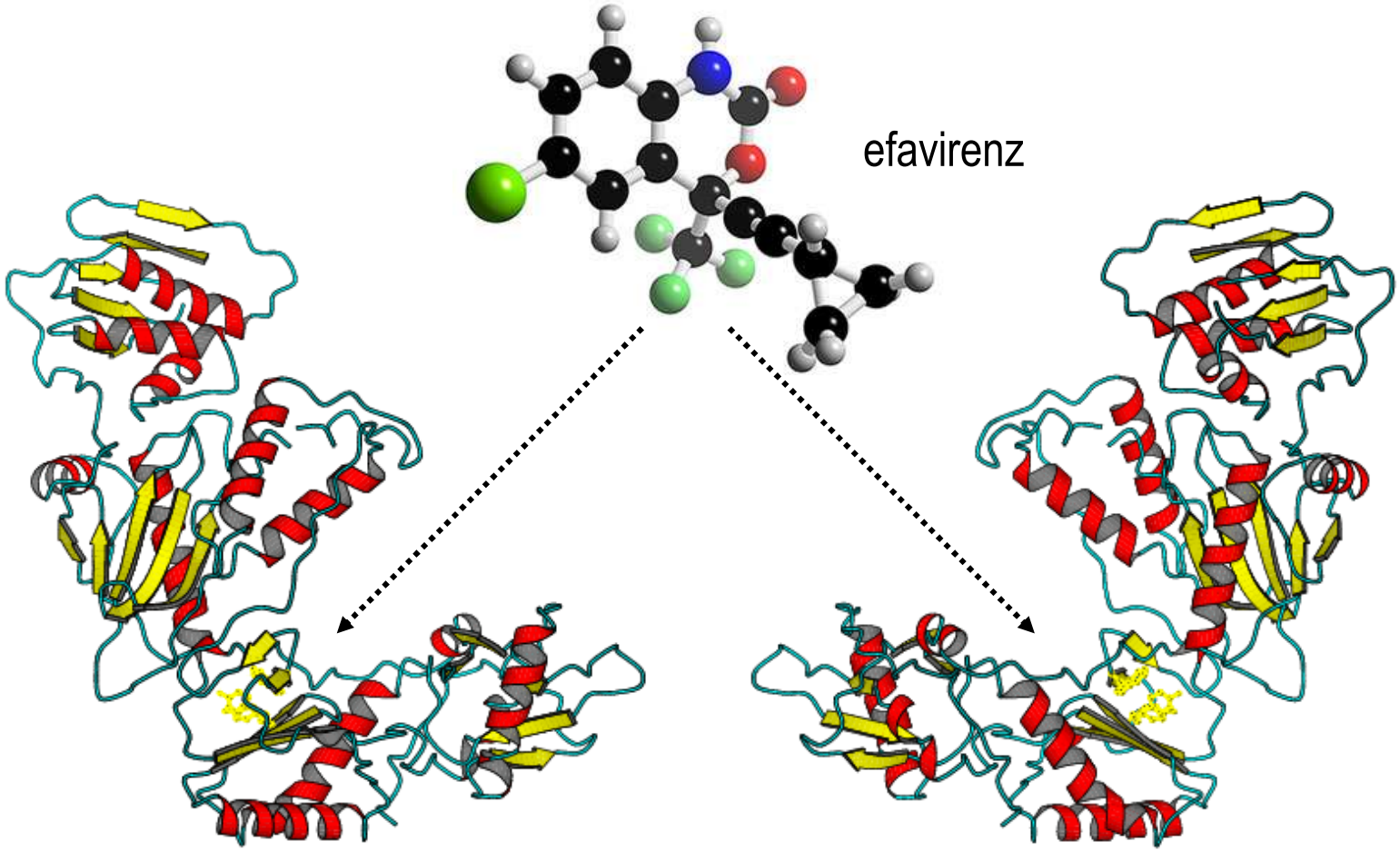






1999: STABILIZATION OF RECEPTOR CONFORMATIONAL ENSEMBLES

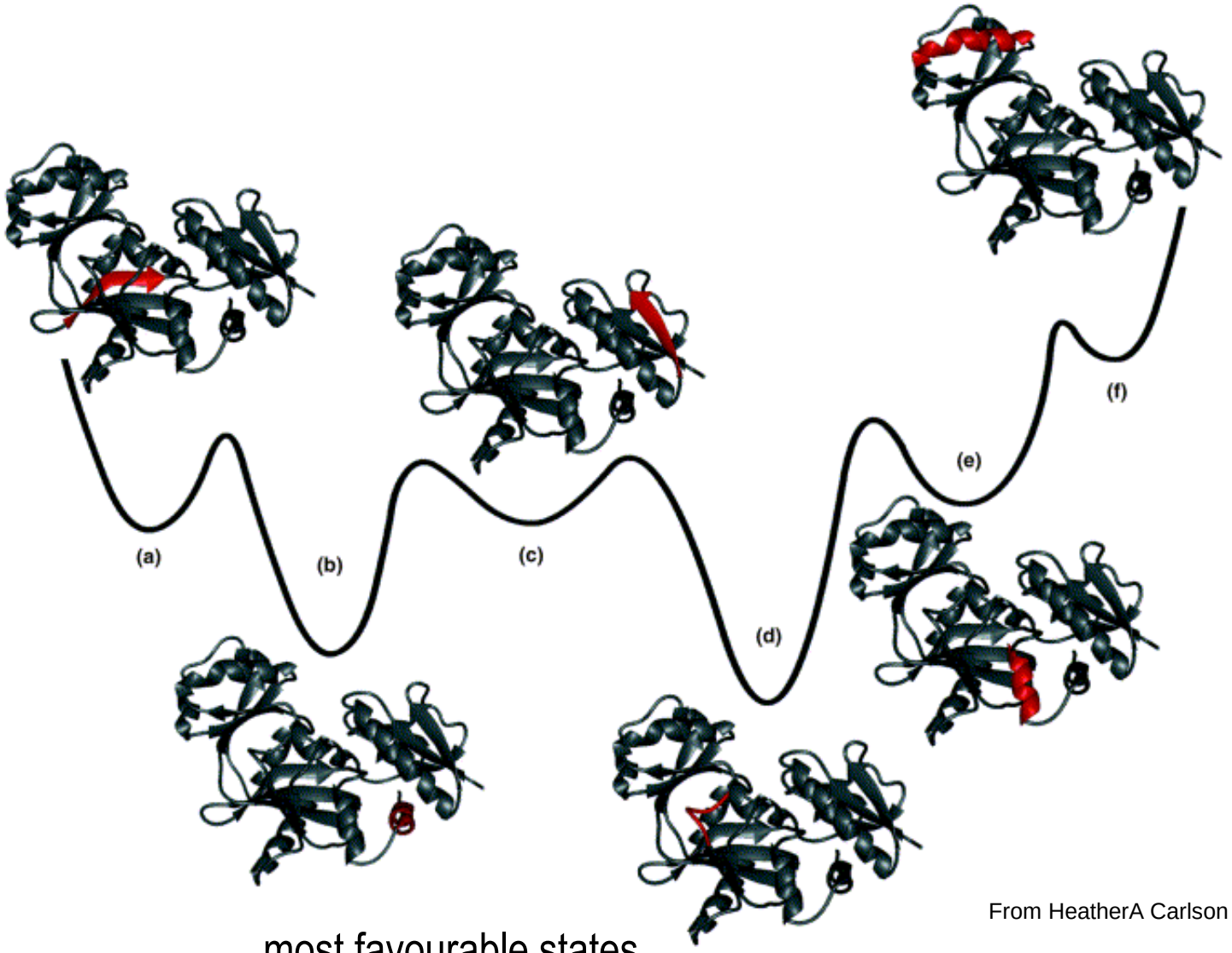




HIV-1 reverse transcriptase

1rtd → 1fk9

Energy profile of conformational space

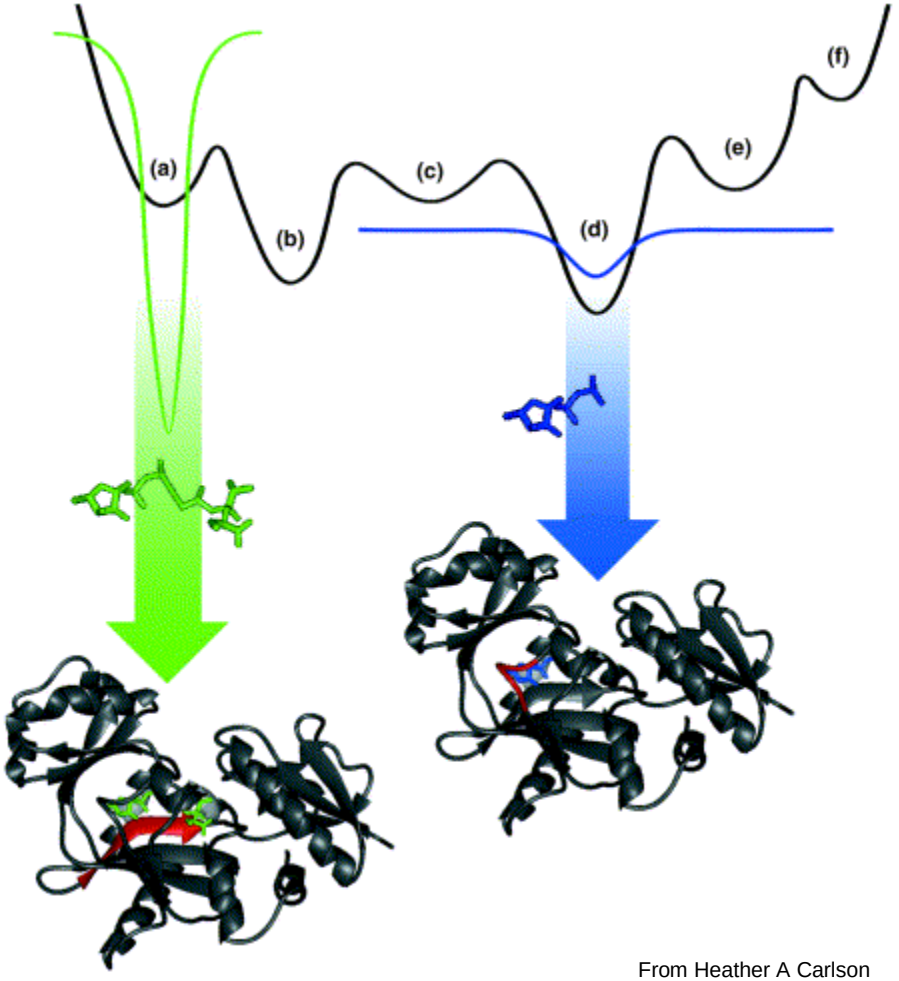


most favourable states

From HeatherA Carlson

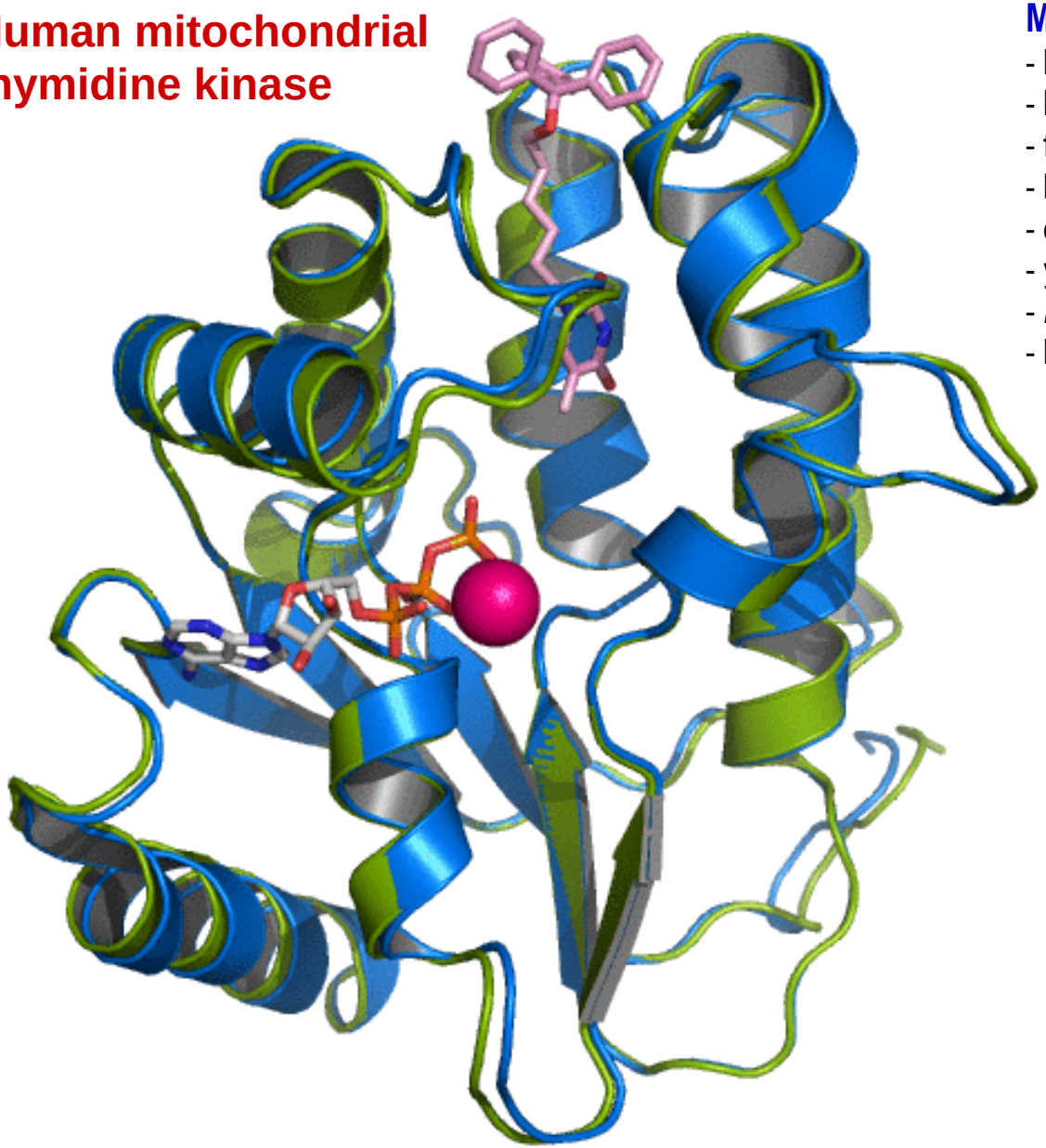
Current Opinion in Chemical Biology

The interaction with the ligand shifts the populations within the ensemble = ‘induced-fit’



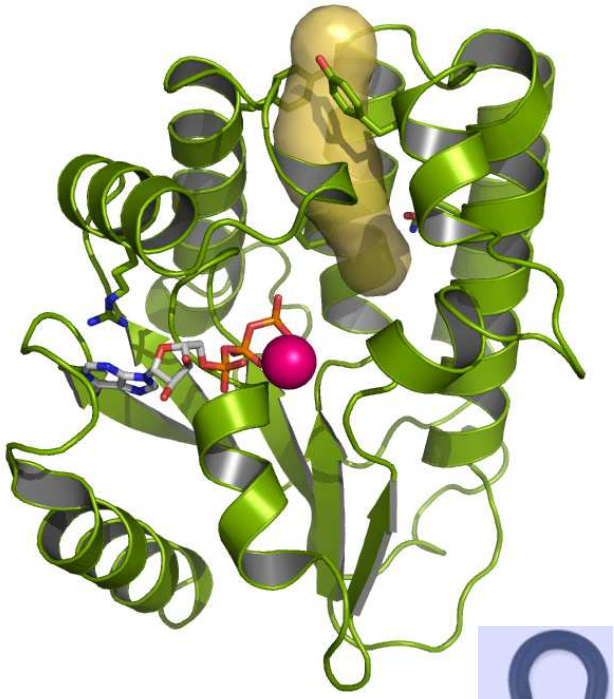
From Heather A Carlson
Current Opinion in Chemical Biology

Human mitochondrial thymidine kinase



Model derived from multiple alignment of:

- human deoxycytidine kinase [1p60]
- human deoxyguanosine kinase [1jag]
- fruitfly deoxyribonucleotide kinase [1oe0]
- herpes simplex virus 1 thymidine kinase [1p7c]
- equine herpes virus thymidine kinase [1p6x]
- yeast thymidylate kinase [3tmk]
- *E. coli* thymidylate kinase [4tmk]
- human thymidylate kinase [1e2f])



NOMAD-Ref Normal Mode Analysis, Deformation, and Refinement

http://lorentz.immstr.pasteur.fr/nma/index.php

http://loschmidt.chemi.muni.cz/caver/

"*The general aim ... has been to determine the extent to which the effects produced by drugs on cells can be interpreted as processes following known laws of physical chemistry*"

$$\text{Drug} + \text{Receptor} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \text{Drug-Receptor}$$

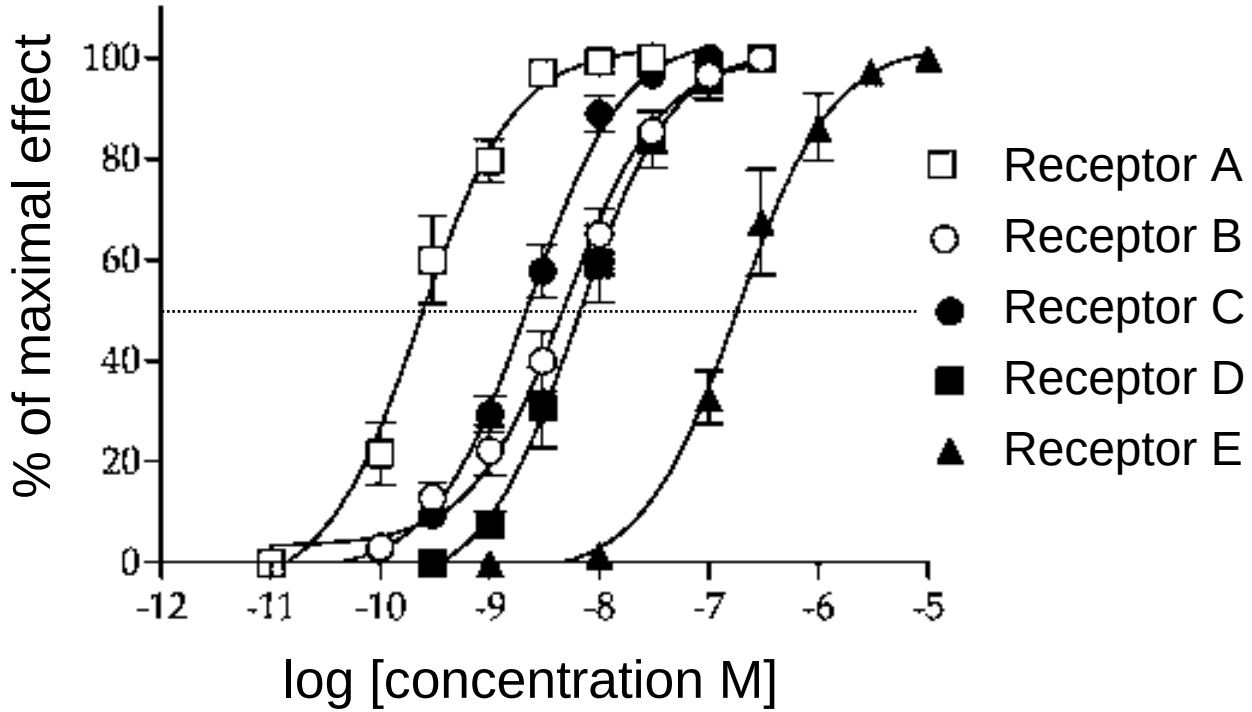
$$K_d = \frac{k_{off}}{k_{on}} = \frac{[\text{Drug}]\,[\text{Receptor}]}{[\text{Drug-Receptor}]}$$

Clark, A.J. (1937)

General Pharmacology: Heffter's Handbuch d. exp. Pharmacologie (Ergband 4), Springer.

Affinity *vs.* Specificity

$$\text{Ligand} + \text{Receptor} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \text{Ligand-Receptor}$$

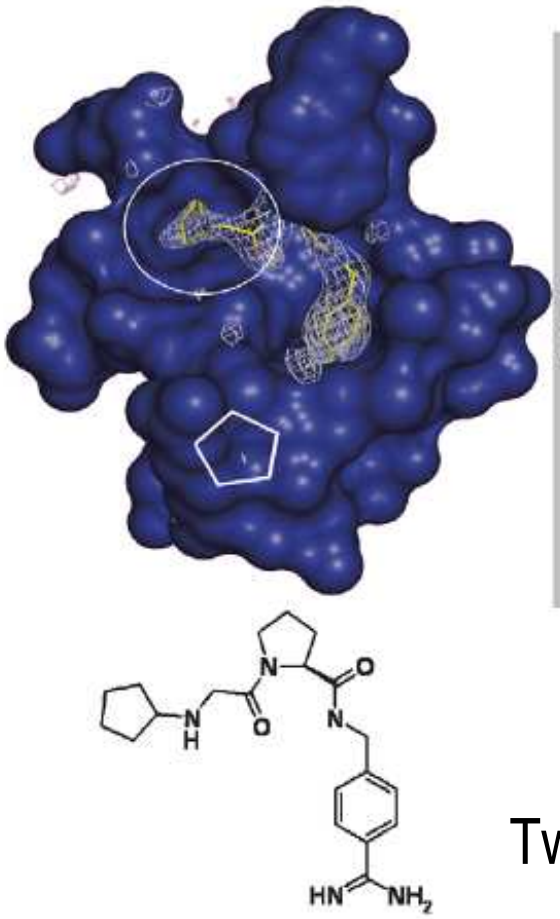


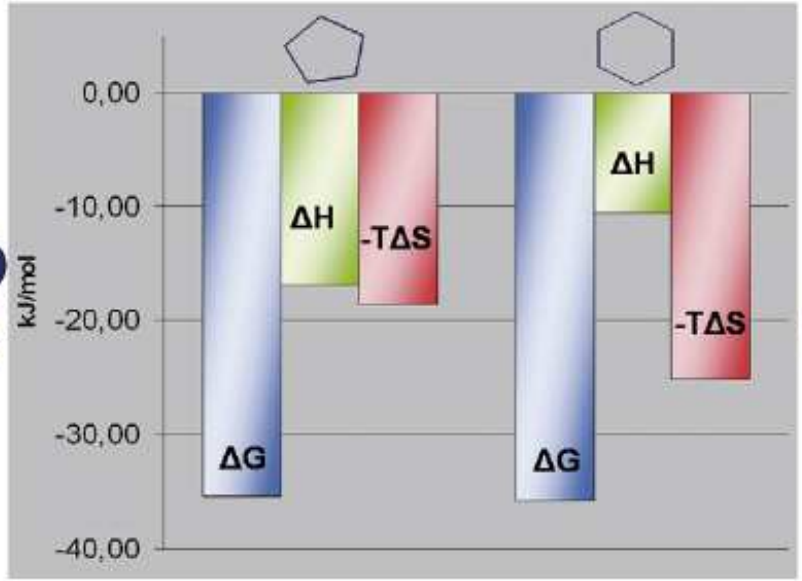
Binding Energy	Binding Constant
ΔG (kcal/mol)	ΔK_d
0.5	2x
1.0	5x
1.5	13x
2.0	29x
2.5	68x
3.0	158x

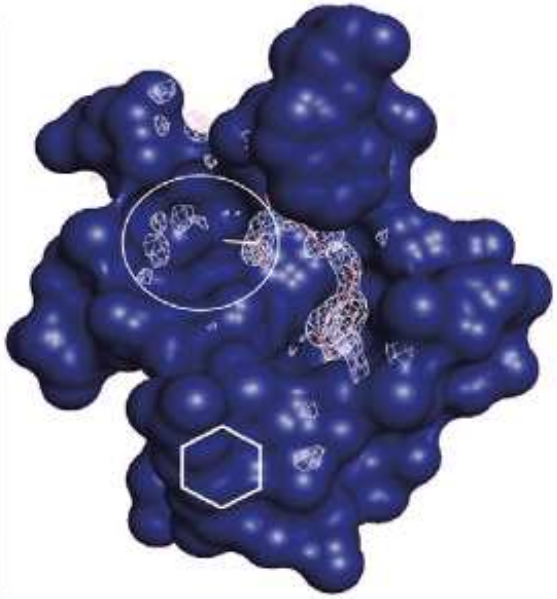
$$\Delta G = \Delta H - T\Delta S = 2.303\ RT \log K_d$$

Ajay, Murcko MA. Computational methods to predict binding free energy in ligand-receptor complexes. *J. Med. Chem*. **1995**, *38*, 4953-4967

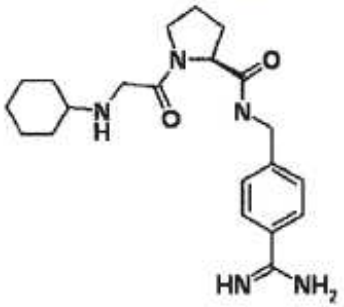
Similar ligands decompose differently into enthalpic and entropic binding contributions.







Two closely related thrombin inhibitors

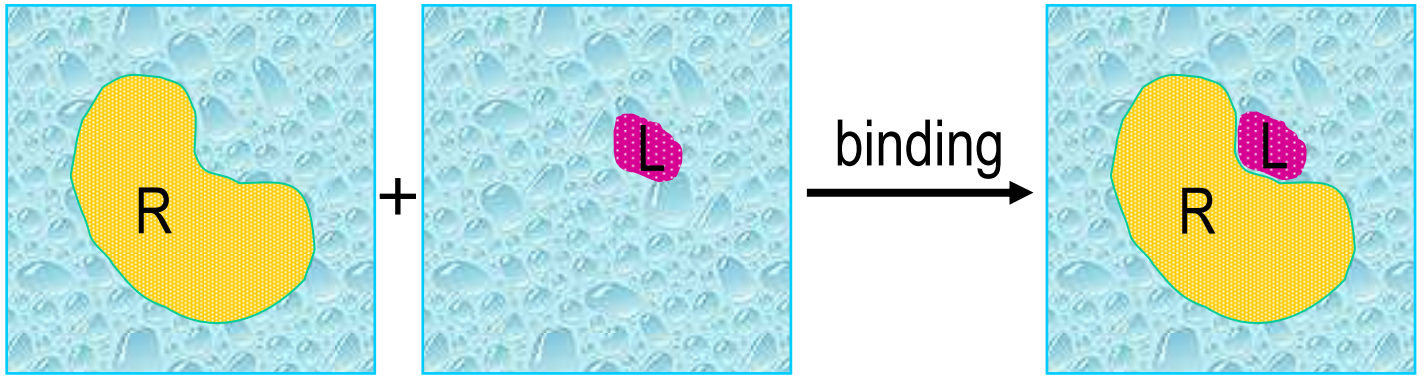


Drug Discovery Today • Volume 11, Numbers 13/14 • July 2006

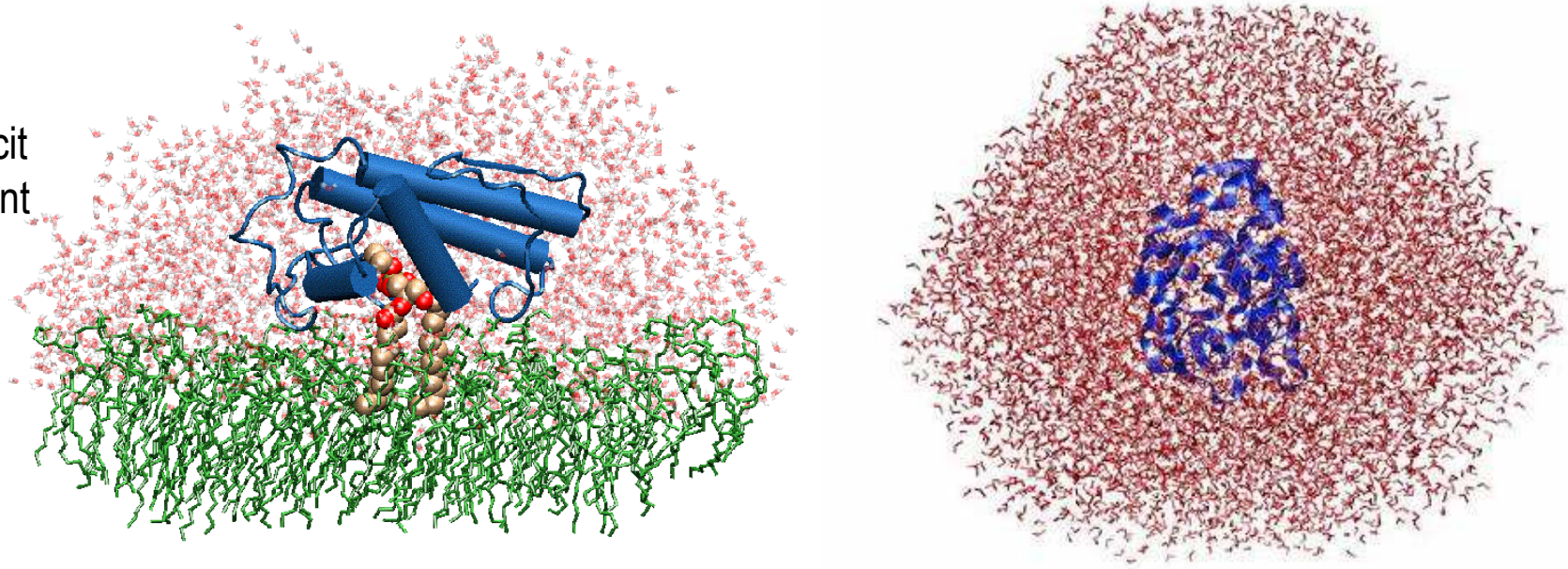
ENERGETICS OF COMPLEX FORMATION

$$\Delta G_{binding} = G_{LR} - (G_R + G_L) \Rightarrow \Delta E_{binding} = E_{LR} - (E_R + E_L)$$

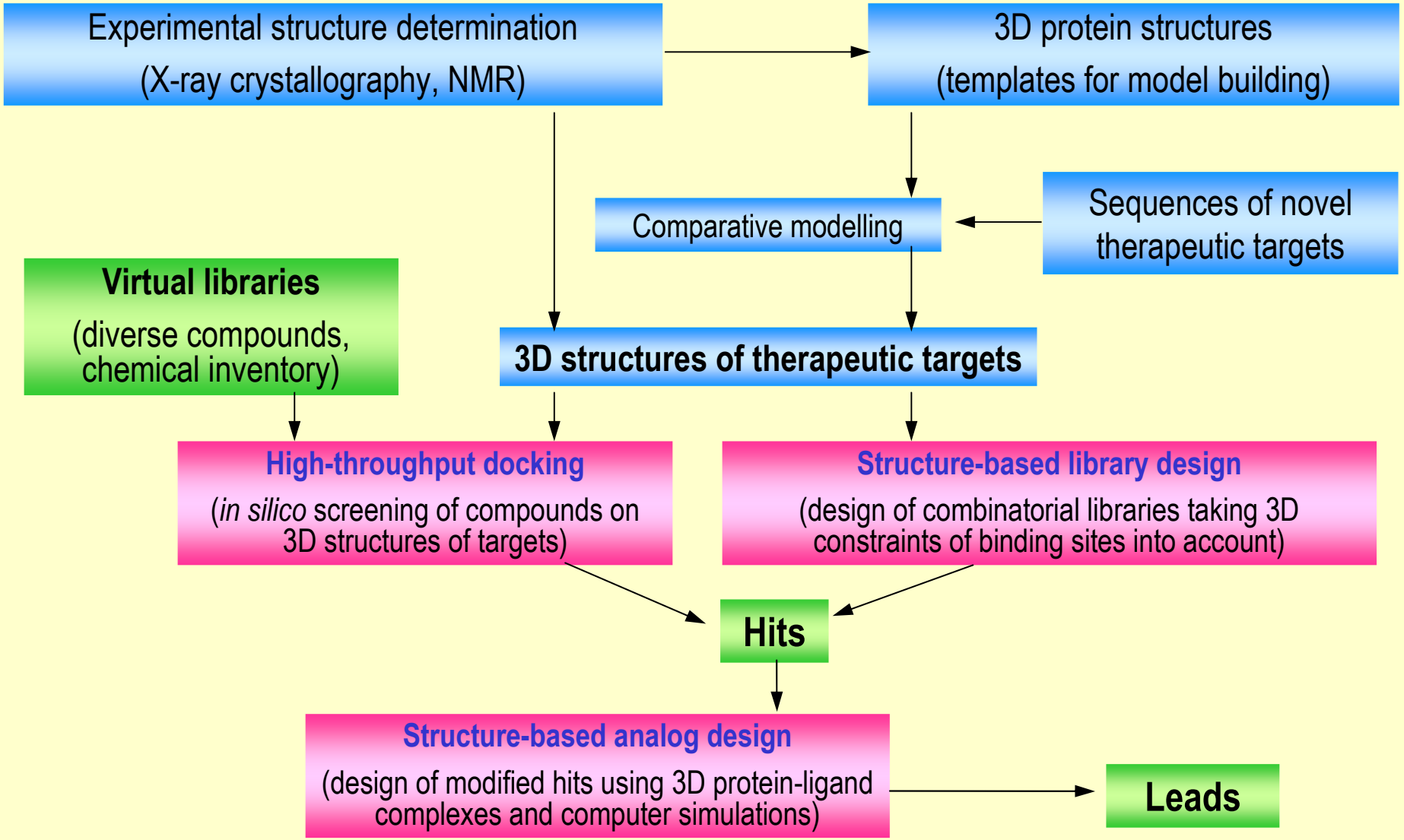
implicit solvent

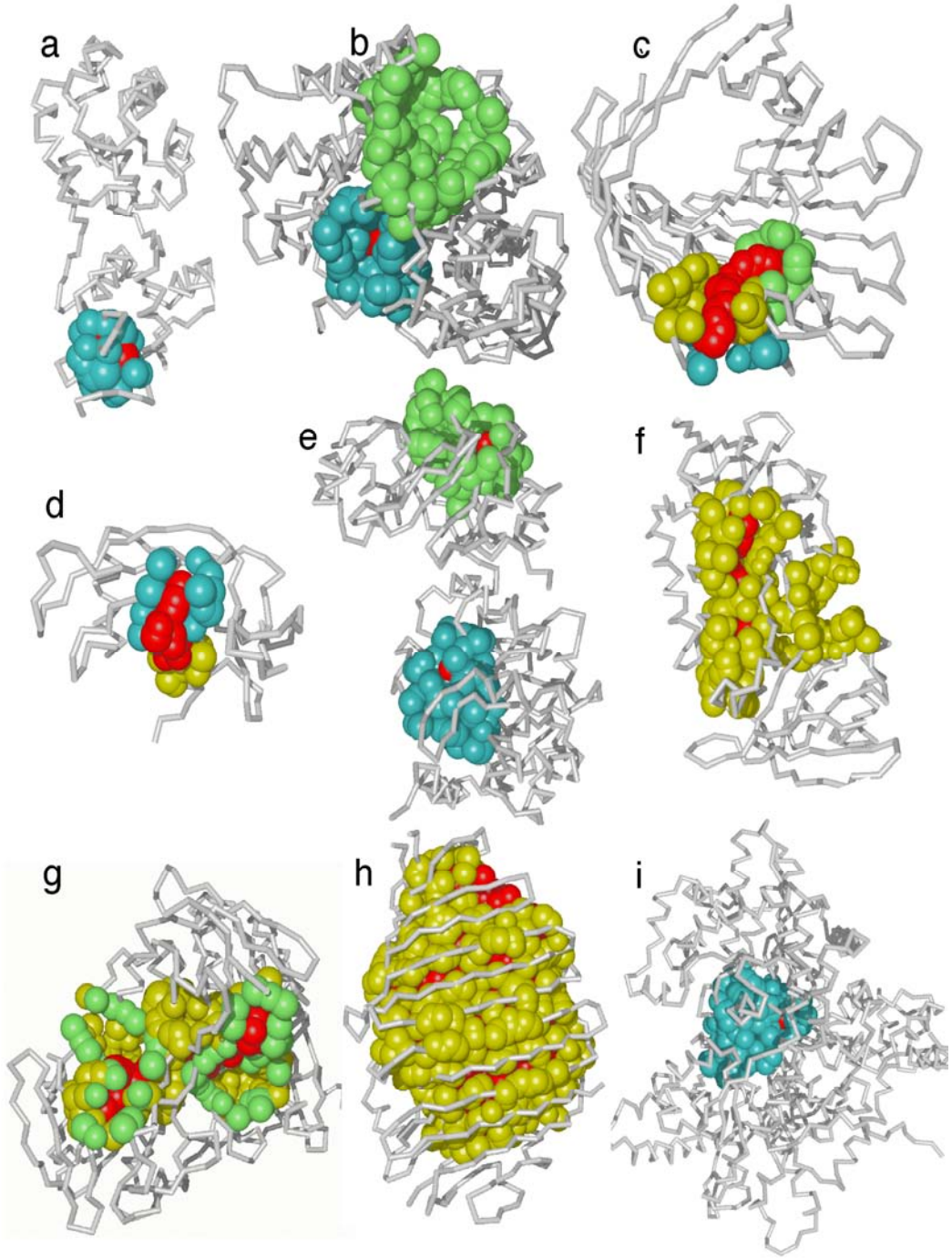


explicit solvent



Integration of Structure-Based and Informatics Methodologies for Drug Discovery



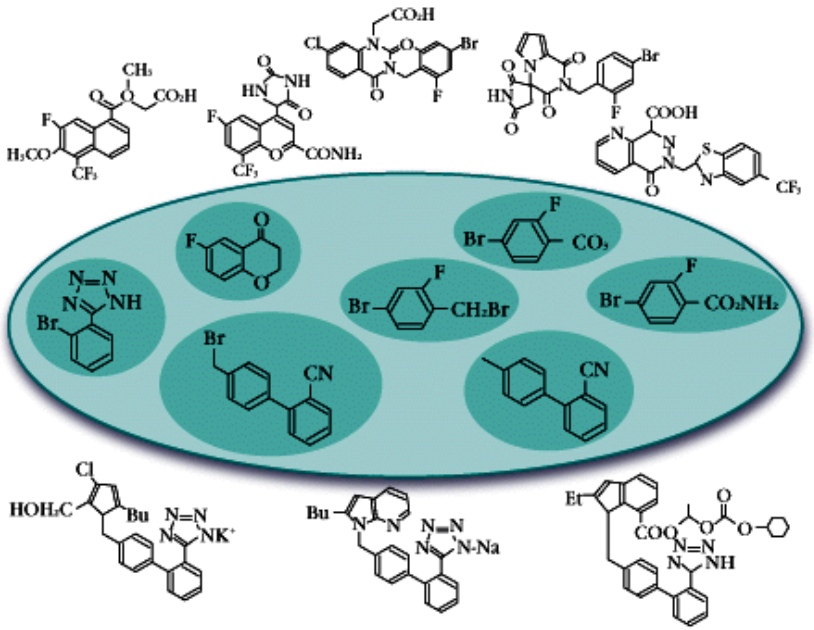


← Validated Targets

+

Combinatorial Chemistry /
Natural Products Libraries

↓

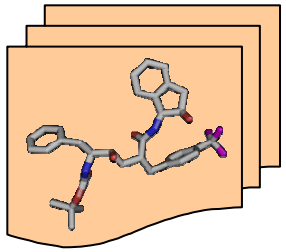


=

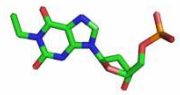
THE "*DOCKING*" PROBLEM

Ligand-Receptor Complexes

Virtual Screening

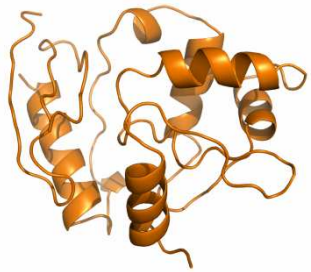


Chemical Library



Ligand

Protein (rigid)

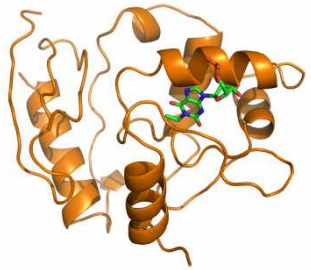


Docking

Search engine
Systematic Sampling
Stochastic Sampling
Deterministic Sampling

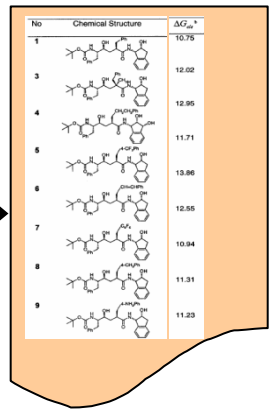
Evaluation
MM Force Fields
Empirical Data
Potentials of Mean Force

Pose + Score



Objective of VS:
to discriminate active from inactive ligands

candidate prioritization



Virtual (“in silico”) screening

Docking/scoring programs

Docking engines: search the conformational space in the binding site

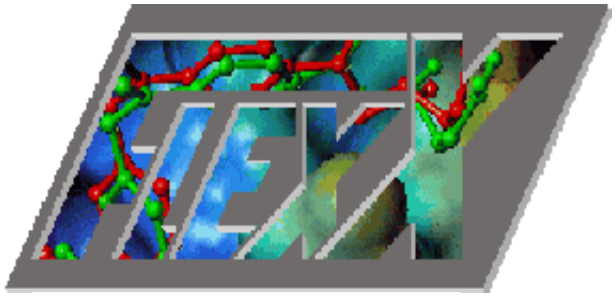
Scoring functions: discrimination of correctly docked from misdocked conformations

Examples of docking algorithms

Rigid ligand:

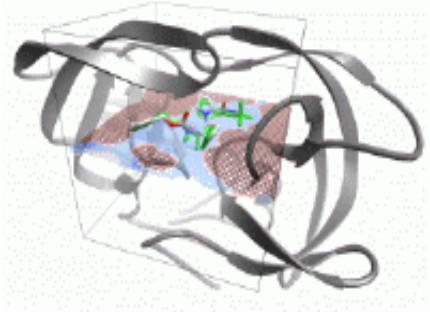
Fast shape matching (DOCK)

Flexible ligand:



Fast shape matching (DOCK 4.0)

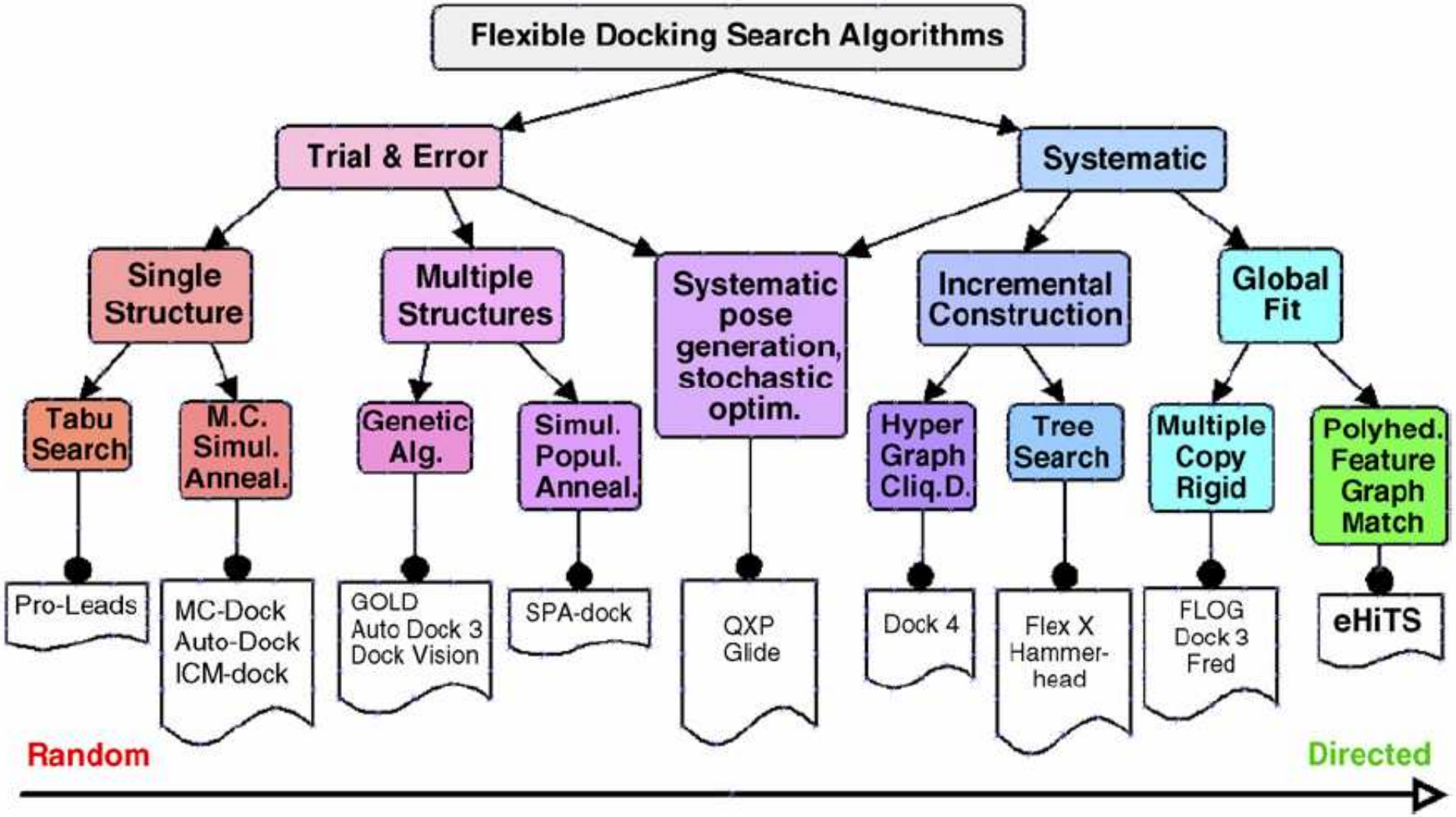
Incremental construction (FlexX)



Simulated annealing (AutoDock 2.4)

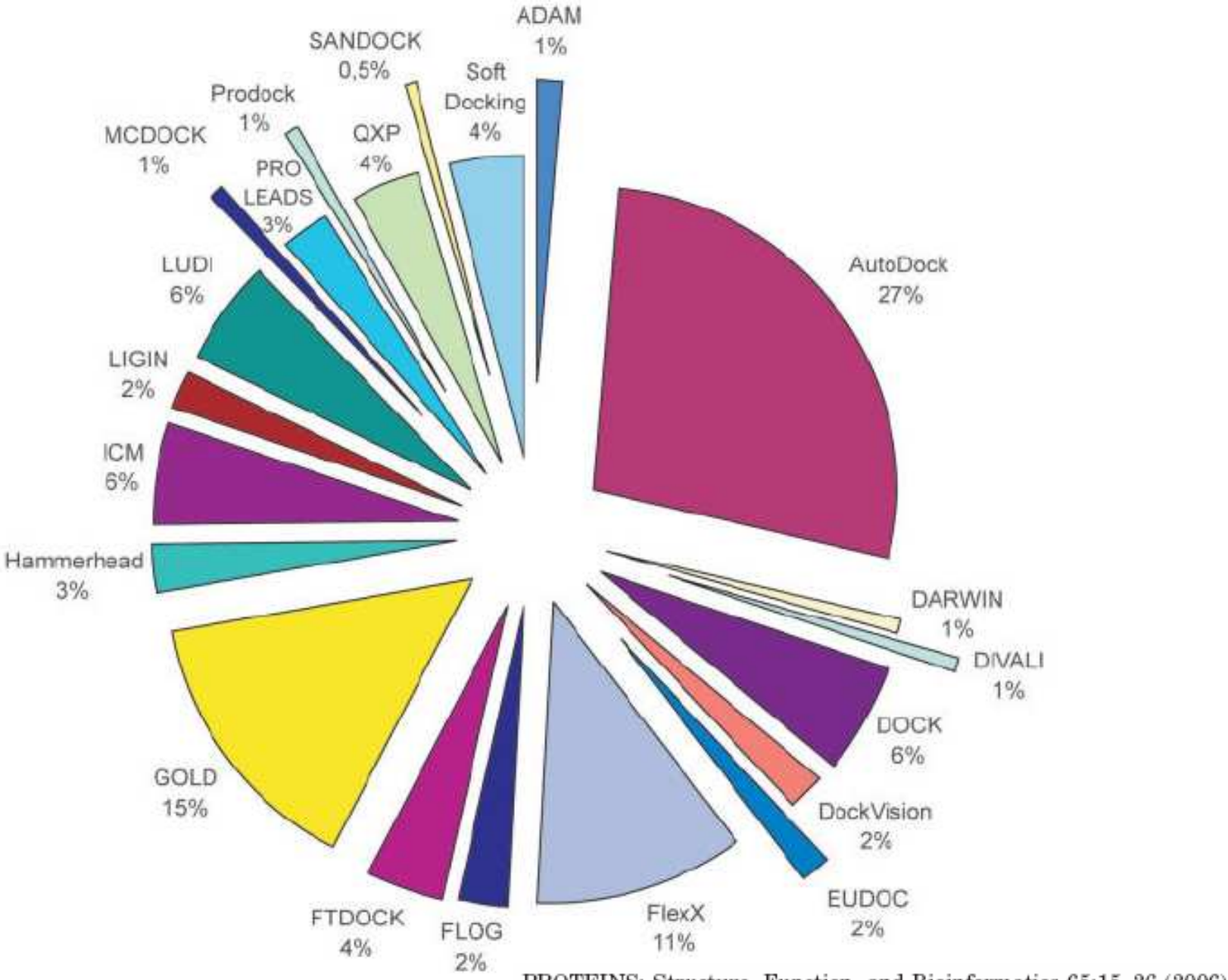
Monte Carlo simulations (MCDOCK)

Genetic algorithm (AutoDock 3.0, GOLD, GAMBLER)



Z. Zsoldos et al. / Journal of Molecular Graphics and Modelling 26 (2007) 198–212

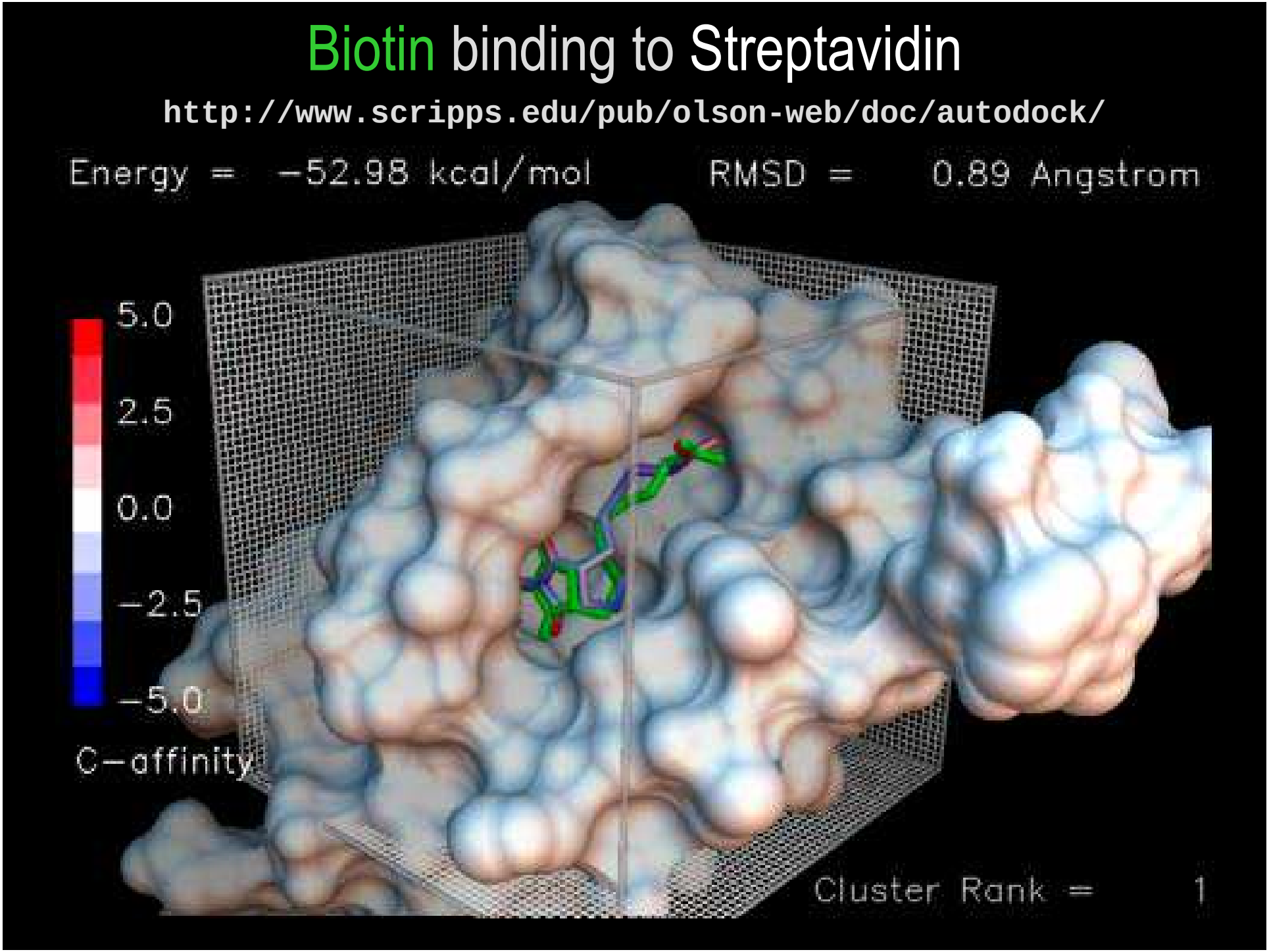
Docking software use according to number of citations in ISI Web of Science 2005



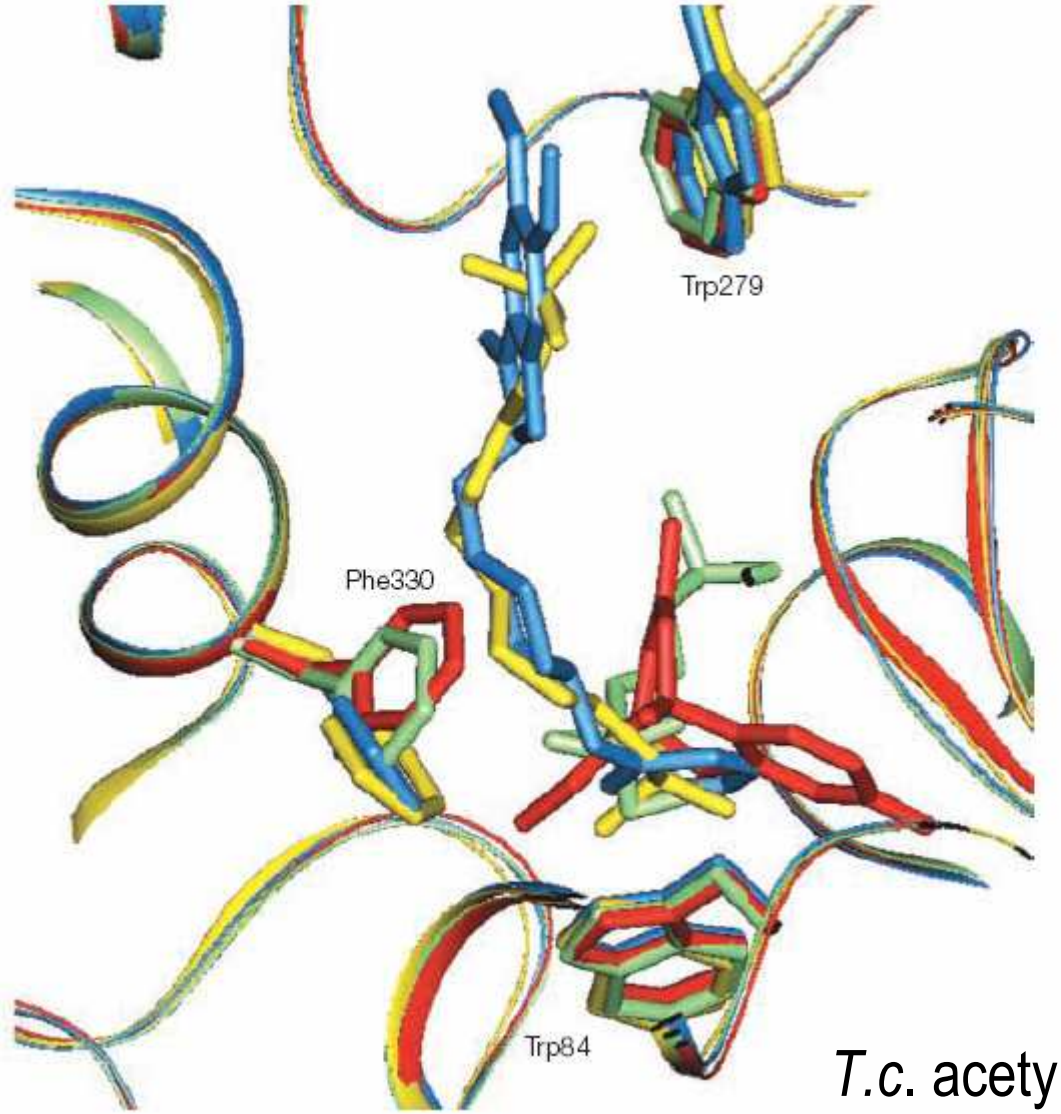
PROTEINS: Structure, Function, and Bioinformatics 65:15–26 (2006)

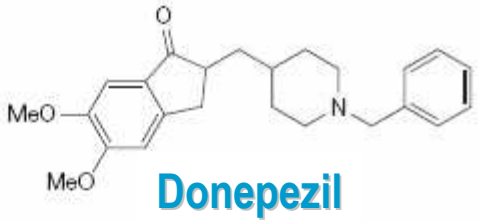
Biotin binding to Streptavidin

http://www.scripps.edu/pub/olson-web/doc/autodock/



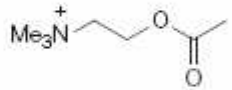
Implications of protein flexibility: Multiple conformations of a single residue



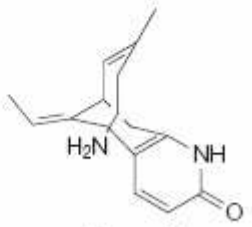


Donepezil

Decamethonium



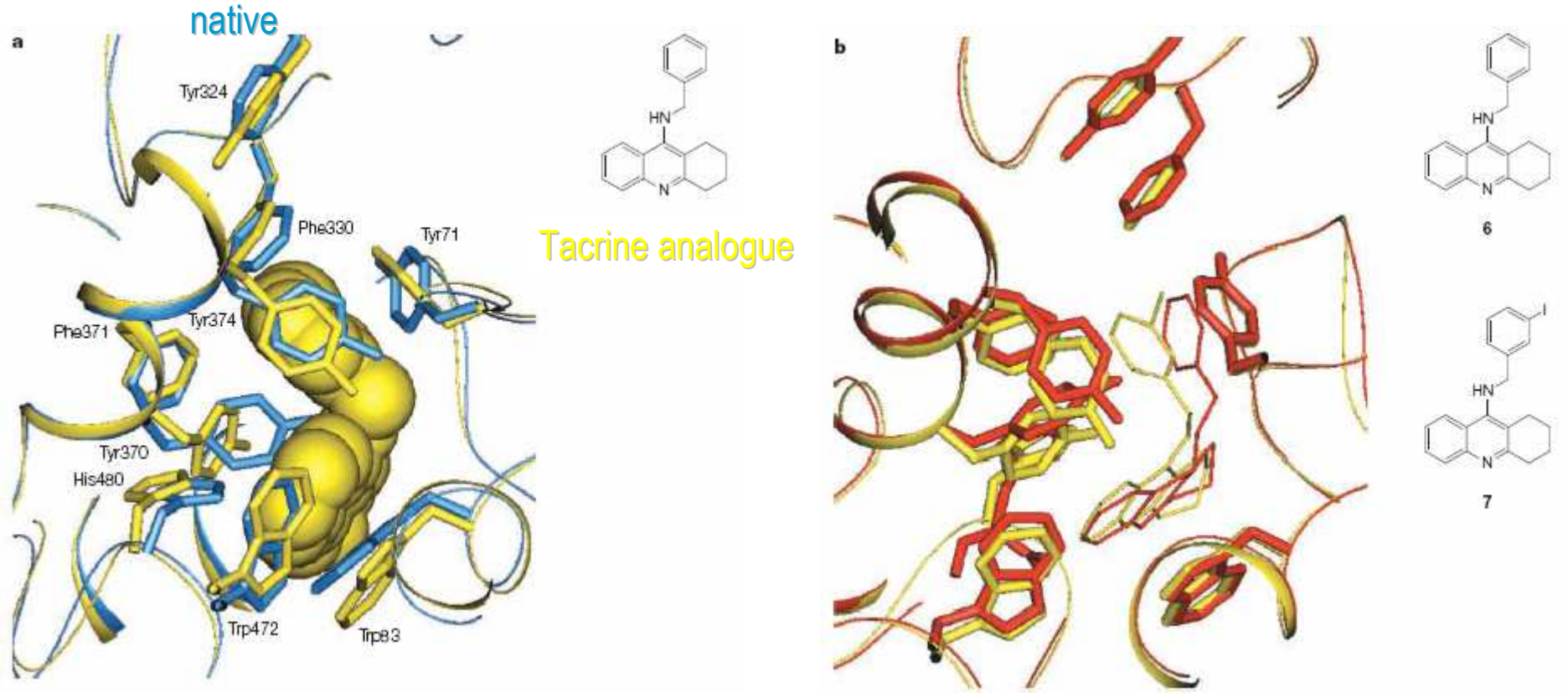
Acetylcholine



Huperzine

T.c. acetylcholinesterase binding site
(PDB entries 2ACE, 1EVE, 1VOT, 1ACL)

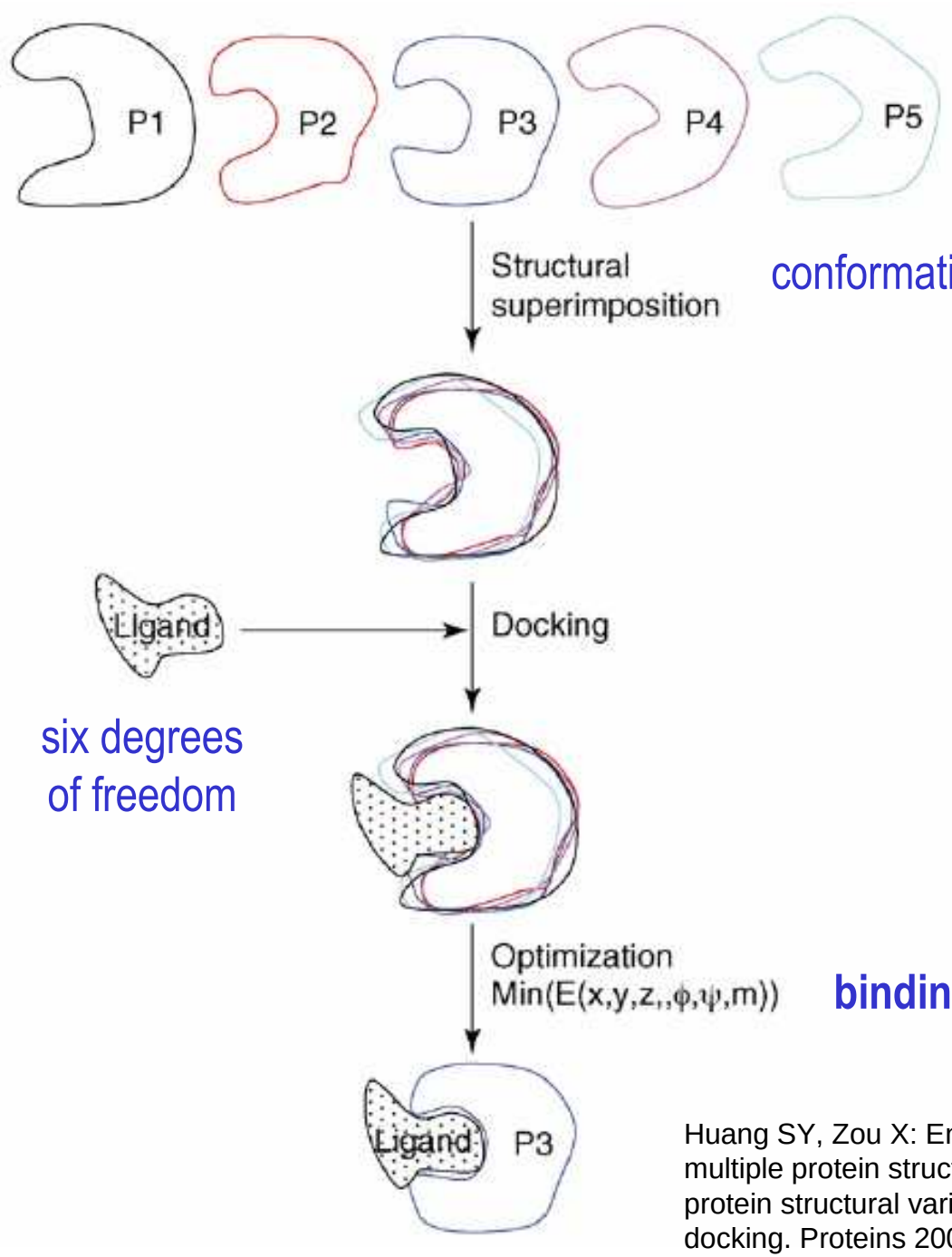
Implications of protein flexibility: Movement of a large number of residues



D.m. acetylcholinesterase binding site

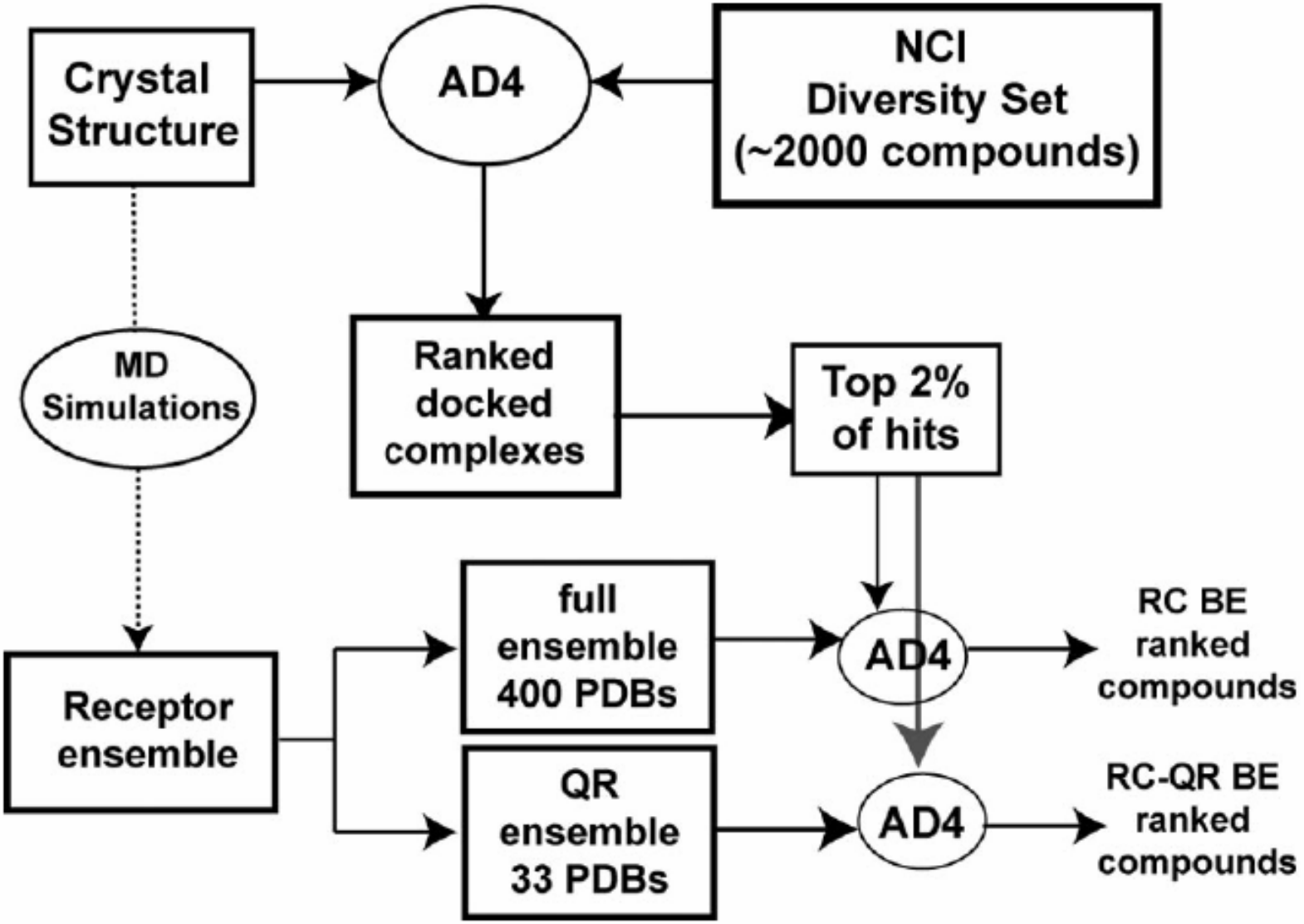
(PDB entries 1QO9, 1DX4) (PDB entries 1DX4, 1QON)

Ensemble docking



Huang SY, Zou X: Ensemble docking of multiple protein structures: considering protein structural variations in molecular docking. Proteins 2007, 66:399-421.

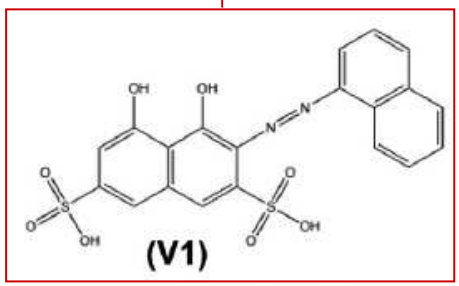
The “Relaxed Complex Scheme” for Virtual Screening



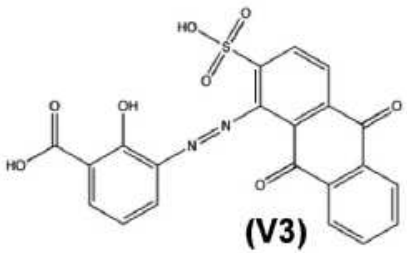
17278–17283 | PNAS | November 11, 2008 | vol. 105 | no. 45

Discovery of drug-like inhibitors of an essential RNA-editing ligase in *Trypanosoma brucei*

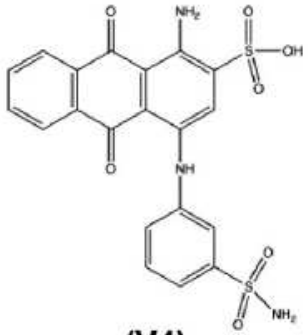
Rommie E. Amaro[a,1,2], Achim Schnaufer[b,1,2], Heidrun Interthal[c], Wim Hol[d,e,f], Kenneth D. Stuart[b,g], and J. Andrew McCammon[a,h,i]



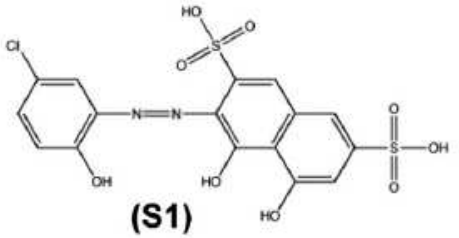
(V1)



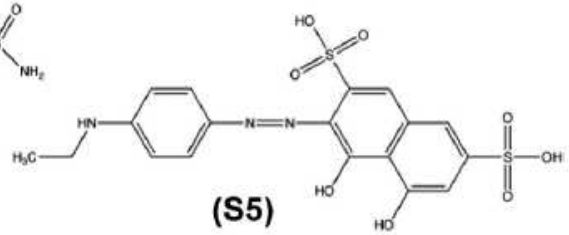
(V3)



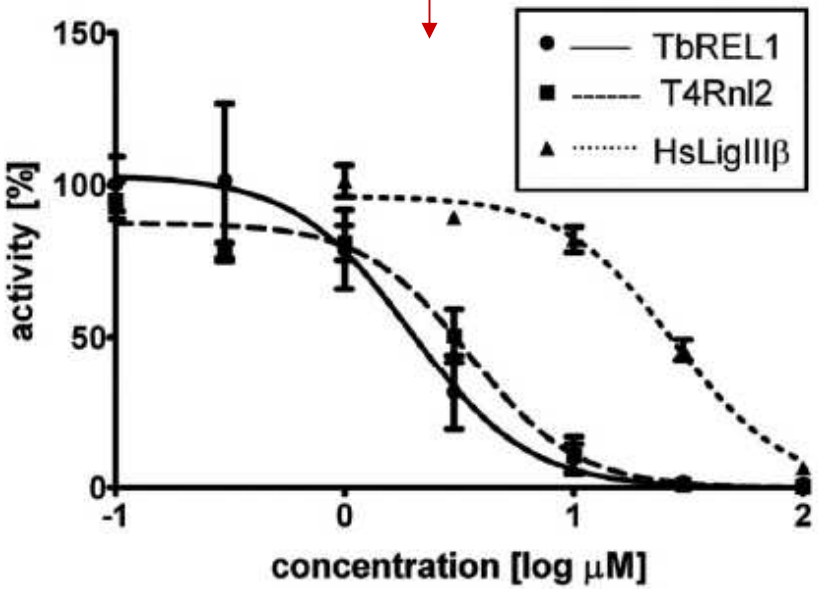
(V4)



(S1)

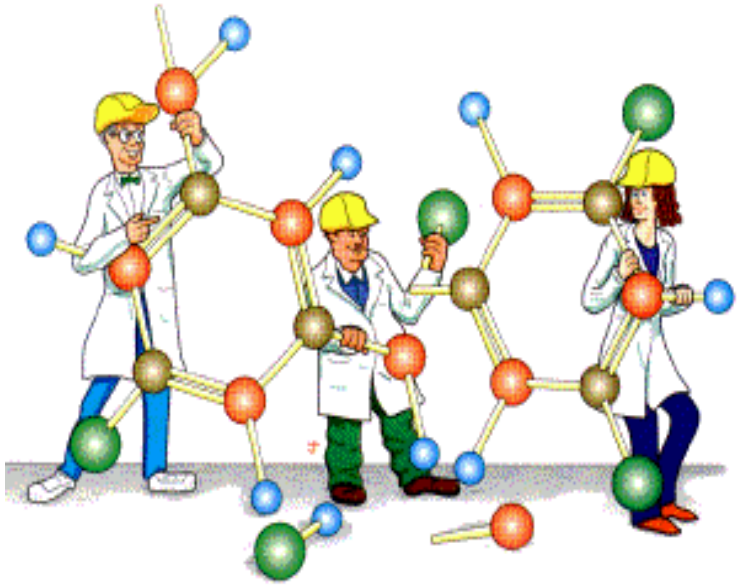


(S5)



	*Tb*REL1	T4Rnl2	*Hs*LigIIIβ
IC_{50} (μM)	1.95±0.33	3.53±1.17	27.49±6.40
R^2	0.9976	0.9868	0.9948

Molecular model building, geometry optimization, and energy calculations:



Molecular Mechanics

Also a scoring function for docking

$$E_{pot} = E_{bonded} + E_{non\text{-}bonded}$$

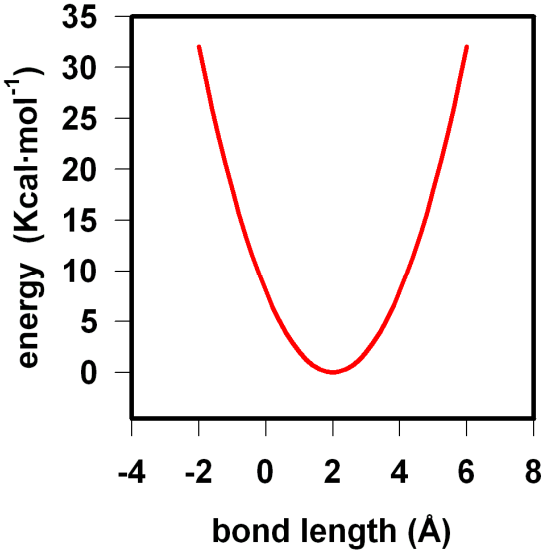
$$E_{bonded} = \Sigma_i \; E_{bond} + \Sigma_i \, E_{angle} + \Sigma_i \, E_{dihedral}$$

$$E_{non\text{-}bonded} = \Sigma_i \, E_{electrostatic} + \Sigma_i \, E_{van\ der\ Waals}$$

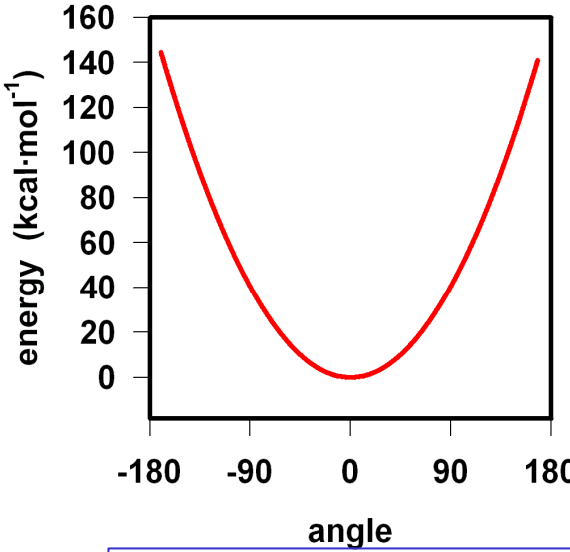
TINKER's *"Molecular Mechanics"* Logo Illustration by Jay Nelson.
Courtesy of Prof. Robert Paine, Chemistry Dept., Univ. of New Mexico.

Molecular Dynamics Simulations
CHARMM

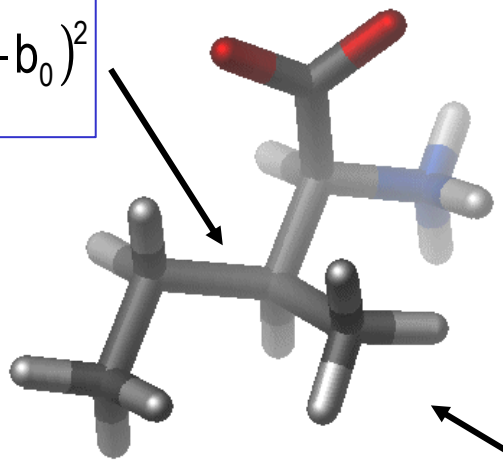
BONDING TERMS

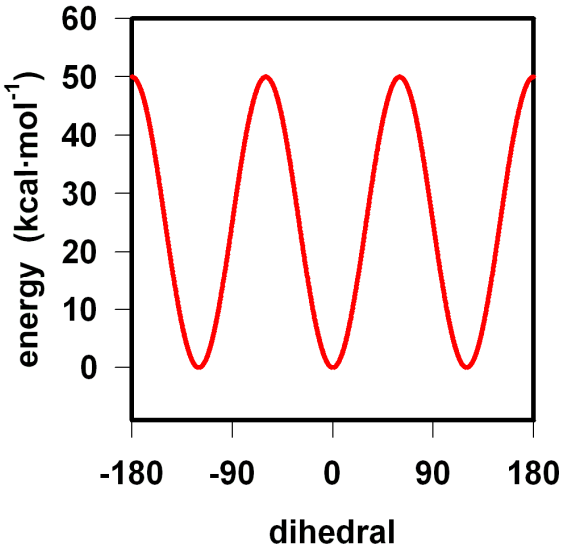


$$E_{bonds} = \sum_{bonds} \frac{1}{2} k_b (b - b_0)^2$$



$$E_{angle} = \sum_{angles} \frac{1}{2} k_\theta (\theta - \theta_0)^2$$



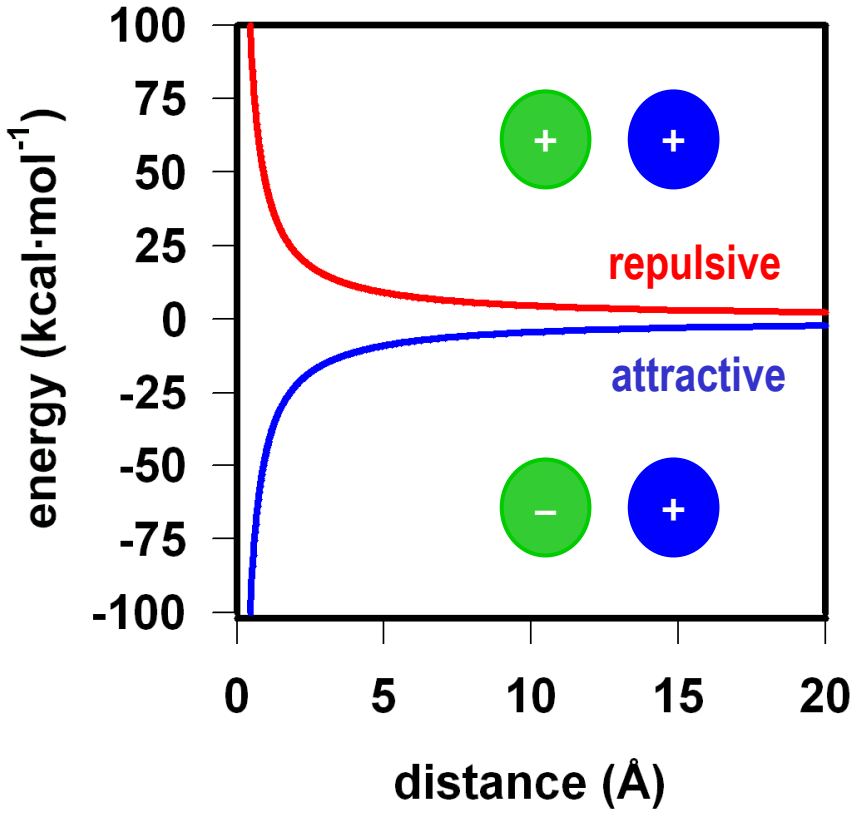


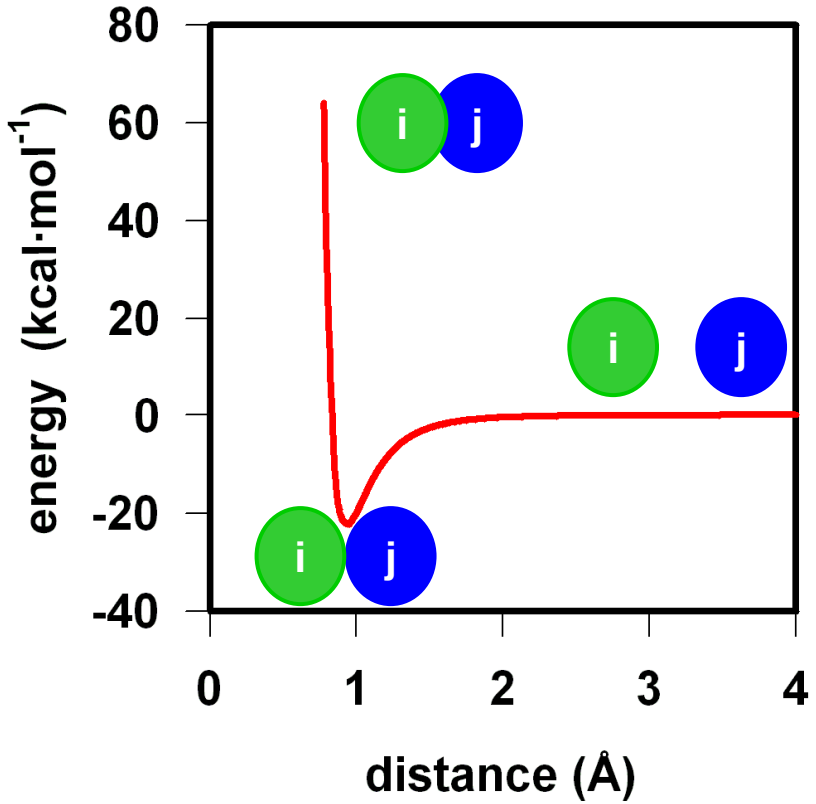
$$E_{dihedral} = \sum_{dihedrals} \frac{1}{2} k_d [1 + \cos(\varphi - \varphi_0)]$$

NON-BONDING TERMS

$$E_{electrostatic} = \frac{1}{4\pi\varepsilon_0\varepsilon}\sum_{ij}\frac{q_i q_j}{r_{ij}}$$

$$E_{Lennard-Jones} = \sum_{ij}\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}$$

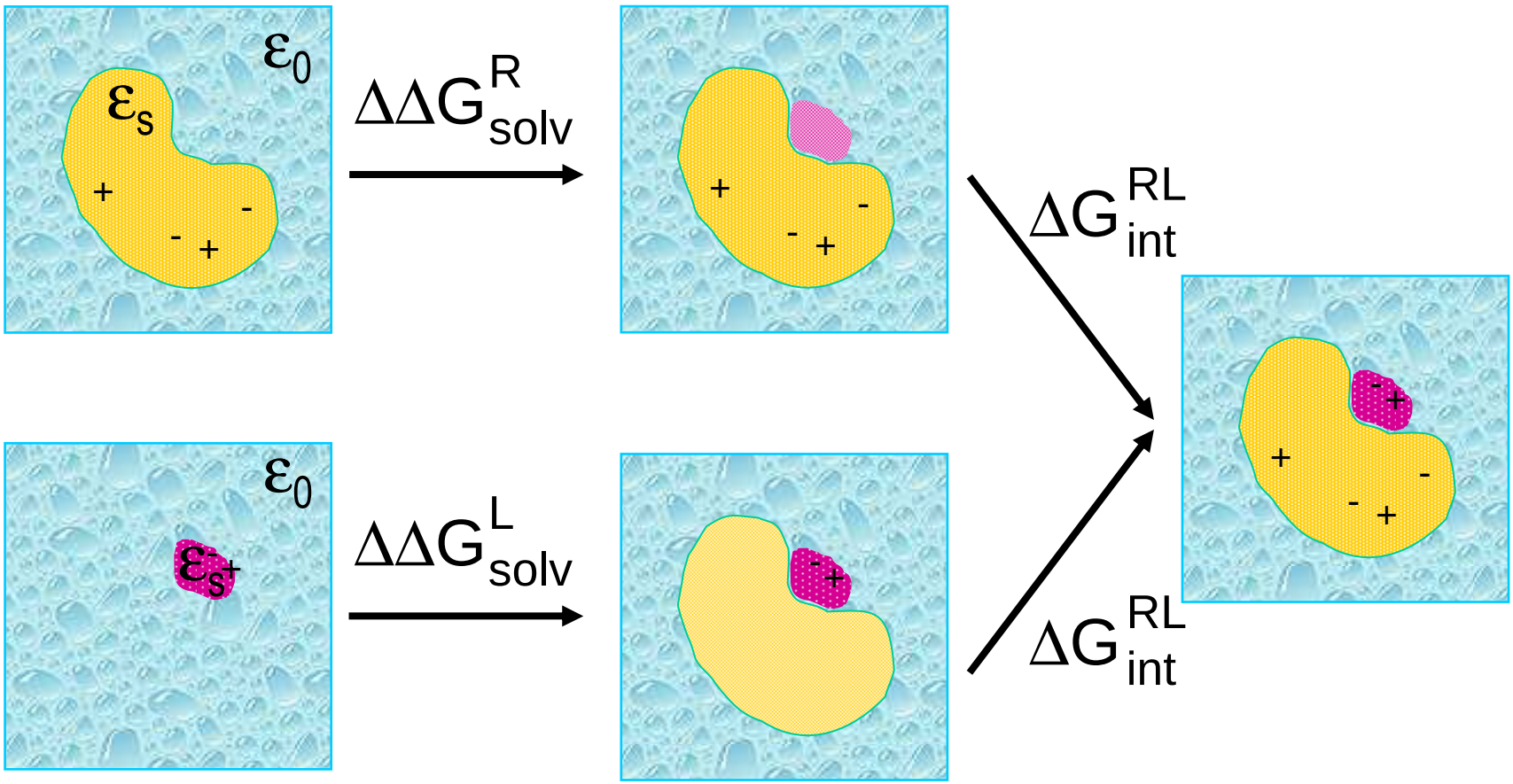




	Protein-Ligand	Internal Ligand
SYBYL™ G-Score	$E_{vdW} + E_{H-bond} =$ $\sum_{prot}\sum_{lig}\left[\left(\frac{A_{ij}}{d_{ij}^{8}} - \frac{B_{ij}}{d_{ij}^{4}}\right) + \left(E_{da} + E_{ww}\right) - \left(E_{dw} + E_{aw}\right)\right]$	$E_{vdW} + E_{torsion} =$ $\sum_{lig}\left(\frac{C_{ij}}{d_{ij}^{12}} - \frac{D_{ij}}{d_{ij}^{6}}\right) + \sum_{lig}\frac{1}{2}V\left[1 + \frac{n}{\lvert n\rvert}\cos\left(\lvert n\rvert\omega\right)\right]$
SYBYL™ D-Score	$E_{vdW} + E_{electrostatic} =$ $\sum_{prot}\sum_{lig}\left[\left(\frac{A_{ij}}{d_{ij}^{12}} + \frac{B_{ij}}{d_{ij}^{6}}\right) + 332.0\frac{q_i q_j}{\varepsilon\left(d_{ij}\right)d_{ij}}\right]$	
GoldScore	$E_{vdW} + E_{electrostatic} =$ $\sum_{prot}\sum_{lig}\left[\left(\frac{A_{ij}}{d_{ij}^{a}} + \frac{B_{ij}}{d_{ij}^{b}}\right) + 332.0\frac{q_i q_j}{\varepsilon\left(d_{ij}\right)d_{ij}}\right]$	$E_{vdW} + E_{electrostatic} =$ $\sum_{lig}\left[\left(\frac{A_{ij}}{d_{ij}^{a}} + \frac{B_{ij}}{d_{ij}^{b}}\right) + 332.0\frac{q_i q_j}{\varepsilon\left(d_{ij}\right)d_{ij}}\right]$ $+ optional\ E_{H-bond}$
AutoDock v3.05	$E_{vdW} + E_{H-bond} + E_{electrostatic} =$ $\sum_{prot}\sum_{lig}\left[\left(\frac{A_{ij}}{d_{ij}^{12}} - \frac{B_{ij}}{d_{ij}^{6}}\right) + E(t)\times\left(\frac{C_{ij}}{d_{ij}^{12}} - \frac{D_{ij}}{d_{ij}^{10}}\right) + 332.0\frac{q_i q_j}{\varepsilon\left(d_{ij}\right)d_{ij}}\right]$ where $E(t)$=angular weight factor	$E_{vdW} + E_{H-bond} + E_{electrostatic} =$ $\sum_{lig}\left[\left(\frac{A_{ij}}{d_{ij}^{12}} - \frac{B_{ij}}{d_{ij}^{6}}\right) + E(t)\times\left(\frac{C_{ij}}{d_{ij}^{12}} - \frac{D_{ij}}{d_{ij}^{10}}\right) + 332.0\frac{q_i q_j}{4\left(d_{ij}\right)d_{ij}}\right]$ where $E(t)$=angular weight factor

Poisson equation: $$\nabla^2\phi(r) = -\frac{4\pi\rho(r)}{\varepsilon}$$

Poisson-Boltzmann equation: $$\nabla\cdot[\varepsilon(r)\nabla\phi(r)] - k'\phi(r) = -4\pi\rho(r)$$



Empirical Scoring Functions

The binding free energy is broken down into a number of different *weighted* contributions (supposed to be additive: number of hydrogen bonds, ionic interactions, apolar contacts, entropy penalties...)

	Functional Form
LUDI	$\Delta G_{bind} = \Delta G_{H-bond} \sum_{H-bond} f(\Delta R, \Delta\alpha) + \Delta G_{ionic} \sum_{ionic} f(\Delta R, \Delta\alpha) + \Delta G_{hydrophobic} \sum_{hydrophobic} \lvert A_{hydrophobic} \rvert + \Delta G_{rotor} N_{rotor} + \Delta G_0$
SYBYL™ F-Score	$\Delta G_{bind} = \Delta G_{H-bond} \sum_{H-bond} f(\Delta R, \Delta\alpha) + \Delta G_{ionic} \sum_{ionic} f(\Delta R, \Delta\alpha) + \Delta G_{aromatic} \sum_{aromatic} f(\Delta R, \Delta\alpha) + \Delta G_{contact} \sum_{contact} f(\Delta R) + \Delta G_{rotor} N_{rotor} + \Delta G_0$
SYBYL™ ChemScore	$\Delta G_{bind} = \Delta G_{H-bond} \sum_{H-bond} f(\Delta R, \Delta\alpha) + \Delta G_{metal} \sum_{metal} f(\Delta R, \Delta\alpha) + \Delta G_{lipo} \sum_{lipo} f(\Delta R) + \Delta G_{rotor} \sum_{rotor} f(P_{nl}, P_{nl}') + \Delta G_0$

Knowledge-based Scoring Functions

A sum of *potentials of mean force* between receptor and ligand atoms is derived from statistical analysis of 3D complex structures

	Functional Form
PMF	Parametrized pairwise potential PMF score: $PMF = \sum_{prot}\sum_{lig} A_{ij}(d_{ij}) A_{ij}(d_{ij}) = -k_B T \ln\left[f^{j}_{Vol\_corr}(r) \frac{\rho^{ij}_{seg}(r)}{\rho^{ij}_{bulk}} \right]$ where: (a) k_B is the Boltzmann constant; (b) $f^{j}_{Vol\_corr}(r)$ is a ligand volume correction factor; (c) and $\frac{\rho^{ij}_{seg}(r)}{\rho^{ij}_{bulk}}$ indicates a radial distribution function for a protein atom *i* and a ligan atom *j*
DrugScore v1.2	$\Delta W = \gamma \sum_{prot}\sum_{lig} \Delta W_{ij}(r) + (1-\gamma)\left[\sum_{lig} \Delta W_i(SAS, SAS_0) + \sum_{prot} \Delta W_j(SAS, SAS_0) \right]$ where: (a) SAS correspond to the surface accessible area terms; (b) W_{ij} is a distance dependent pairwise potential; and (c) γ is an adjustable weight factor, normally set to 0.5
SMoG	$G = \sum_{ij} - kT \log\left[\frac{p_{ij}}{\overline{p}} \right] \Delta_{ij}$ where: (a) Δ_{ij} is 1 if the distance between atoms i and j is within 5.0 Å (and 0 if it is higher than 5.0 Å); and (b) p_{ij} and $\overline{p}$ are inter atomic and averaged inter atomic interactions

Some important questions....

- Is there any relationship between docking and ranking accuracies?

- Will docking/scoring combinations provide better results in terms of hit rates? If so, which ones?

- Does “consensus scoring” from two or three independent scoring lists outperform single scoring?

- Will it be possible to find a universal scoring function?

17 pairs of complexes of the same protein bound to 2 related ligands /

Molecular mechanics (AMBER) and statistical potentials (PMF)

Exhaustive enumeration of all possible docking solutions

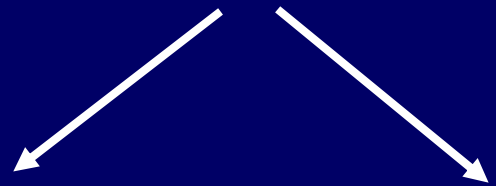
Reconstruction of the shape of the energy landscape (*coverage-error plots*)

Calculation of physico-chemical descriptors

Quantitative evaluation of success

Linear discriminant analysis

Physical origin of failures/successes



Desolvation effects

Directional effects of hydrogen bonds

Dispersive interactions

C. Pérez & A. R. Ortiz

J. Med. Chem. **44**, 3768-3785 (2001)

Virtual (“in silico”) screening

Chemical Libraries

Usual notations: SMILES, SMARTS...

2D $\rightarrow$ 3D conversion: CORINA...

Fiel format interconversion: BABEL...

Popular databases: ACD, ZINC, DUD...

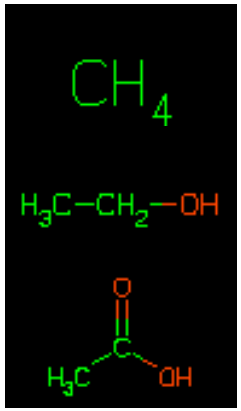
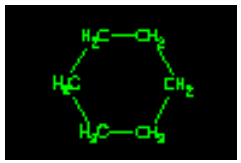
SMILES

Simplified Molecular Input Line Entry Specification

Rules

1. Atoms are represented by atomic symbols: B, C, N, O, F, P, S, Cl, Br, and I.
2. Double bonds are `=', triple bonds are `#'.
3. Branching is indicated by parentheses.
4. Ring closures are indicated by pairs of matching digits.

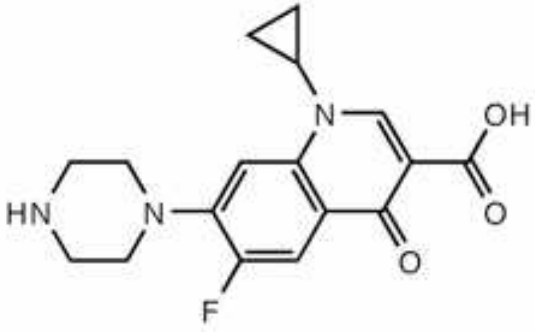
Examples

Depiction	SSMILES	Name	Remark
	C	methane	hydrogens fill normal valence
	CCO	ethanol	a single bond is assumed to join adjacent atoms
	CC(=O)O	acetic acid	parentheses are used to indicate branching
	C1CCCCC1	cyclohexane	bonds can also be represented by pairs of matching digits

DAYLIGHT
Chemical Information Systems, Inc.

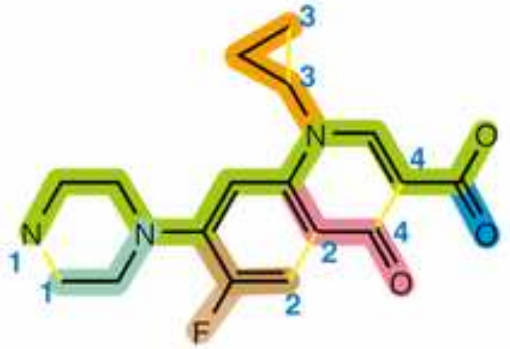
WIKIPEDIA
The Free Encyclopedia

A



B

C



D

N1CCN(CC1)C(C(F)=C2)=CC(=C2C4=O)N(C3CC3)C=C4C(=O)O

http://www.daylight.com/dayhtml/doc/theory/theory.smiles.html

SMARTS

SMiles ARbitrary Target Specification

A language for specifying substructural patterns in molecules

The SMARTS line notation is expressive and allows extremely precise and transparent substructural specification and atom typing:

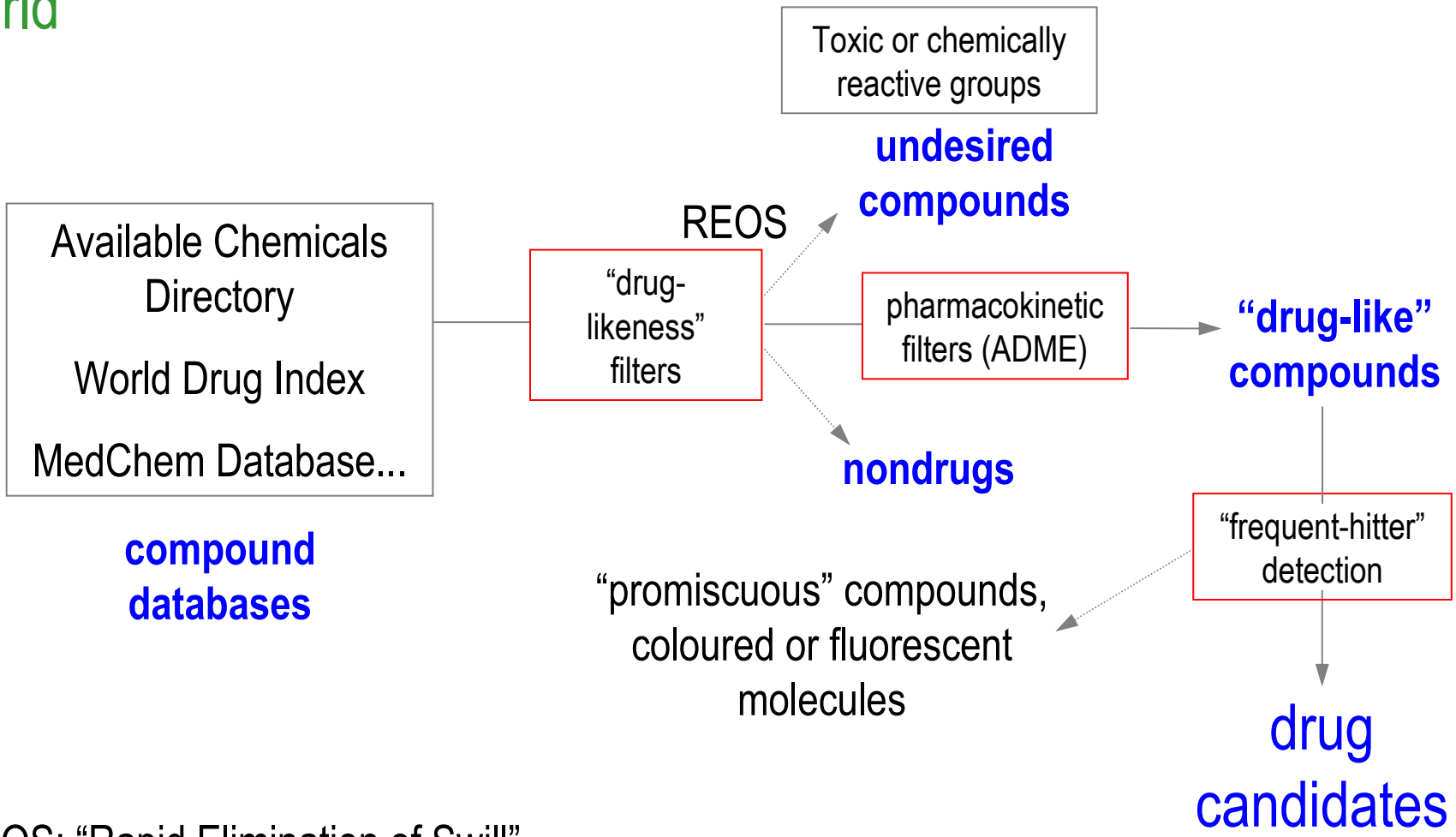
1. Atomic properties: *Aliphatic carbon* [C], *aromatic carbon* [c], *any carbon* [#6] or [C,c]...
2. Bonds: – (*single*), = (*double*), # (*triple*), : (*aromatic*), and ~ (*any*).
3. Connectivity: [CX4] matches a carbon atom with bonds to 4 other atoms.
4. Cyclicity: smallest set of smallest rings (SSSR).
5. Logical operators: and or , ; & ! (not)

Applications:

1. Definition of substructural filters e.g. to identify undesirable compounds
2. Definition of bond types in RECAP (Retrosynthetic Combinatorial Analysis Procedure)
3. Use in several programs: Leatherface (molecular editor), ALADDIN (pharmacophore matching), etc

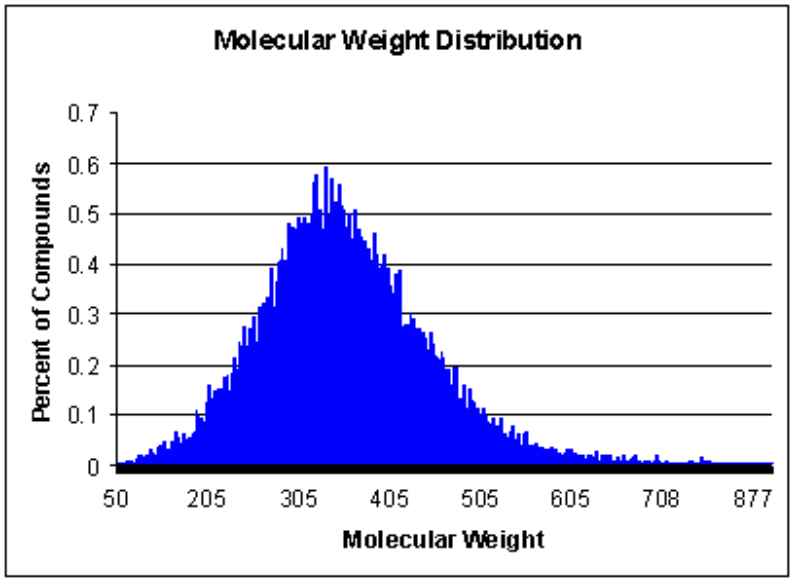
In silico **VIRTUAL SCREENING** and FOCUSED LIBRARY DESIGN

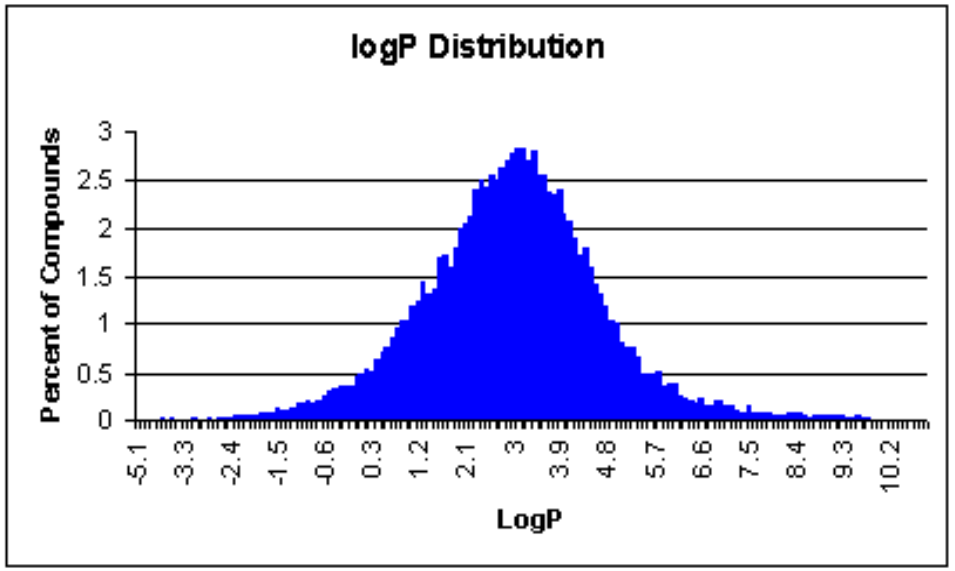
Near-perfect structures in an imperfect world

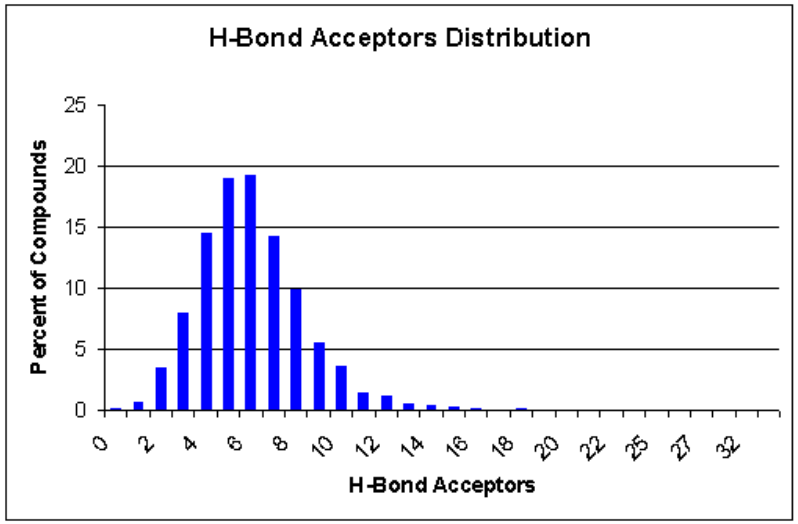


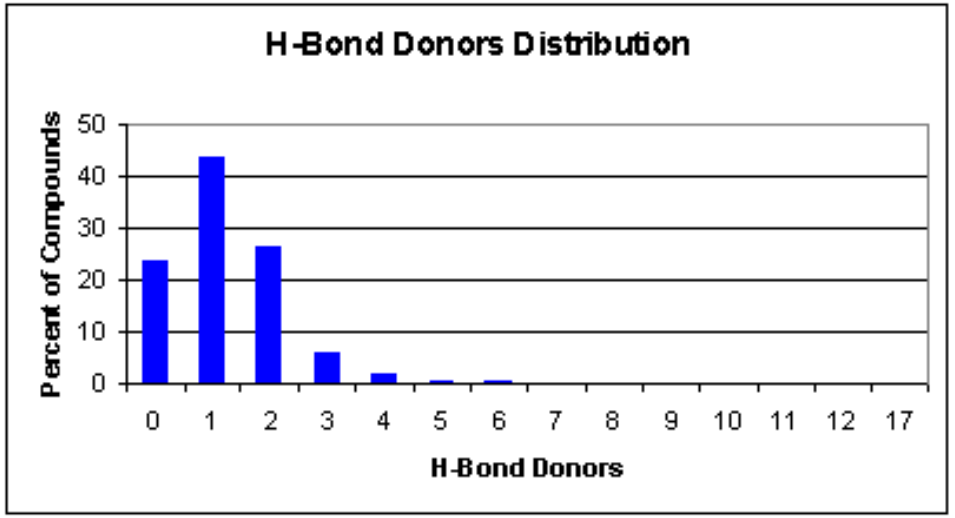
REOS: "Rapid Elimination of Swill"

Drug-like characteristics



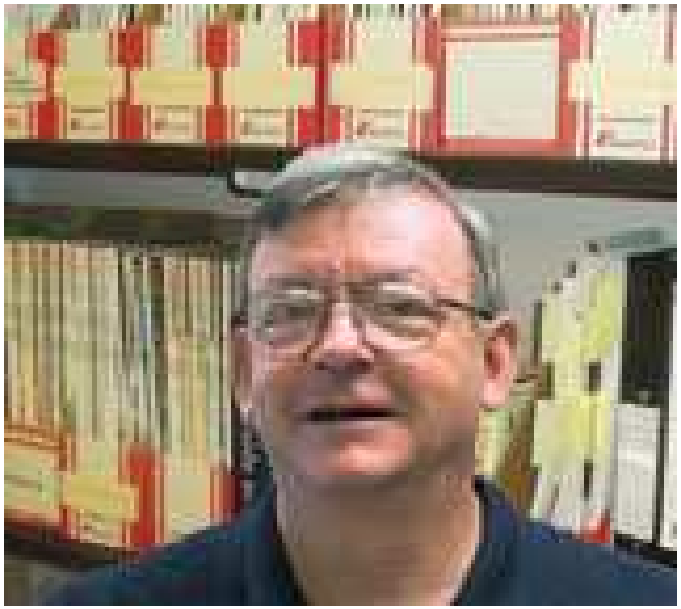






Lipinski, C. A. *et al.* Experimental and computational approaches to estimate solubility and permeability in drug discovery and developmental settings. *Advanced Drug Deliv. Rev*. **1997**, *23*, 3-29.

LIPINSKI's "rule of five" (Ro5)



H-bond donors <5

H-bond acceptors (N, O) <10

cLog P <5

Molecular Weight <500 Da

Lipinski, C. A. *et al.* Experimental and computational approaches to estimate solubility and permeability in drug discovery and developmental settings. *Advanced Drug Deliv. Rev*. **1997**, *23*, 3-29.

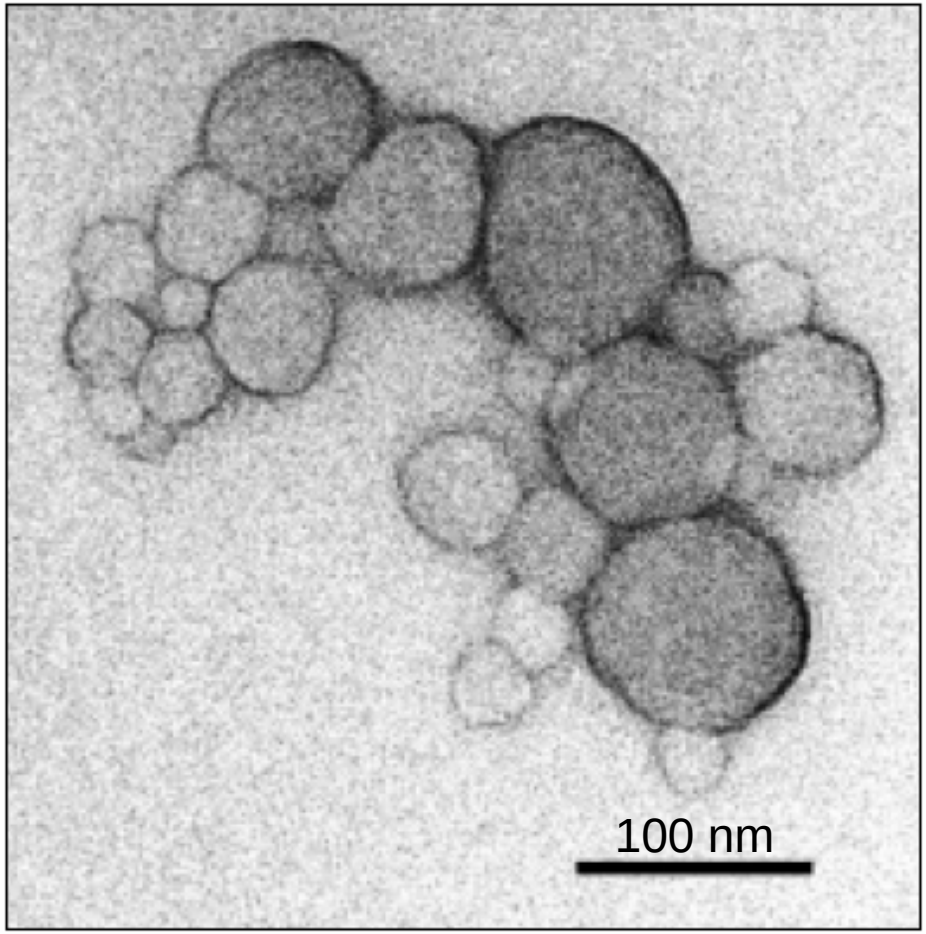
1 - compounds which satisfy 1 requirement - 1% of all compounds

2 - compounds which satisfy 2 requirements - 2% of all compounds

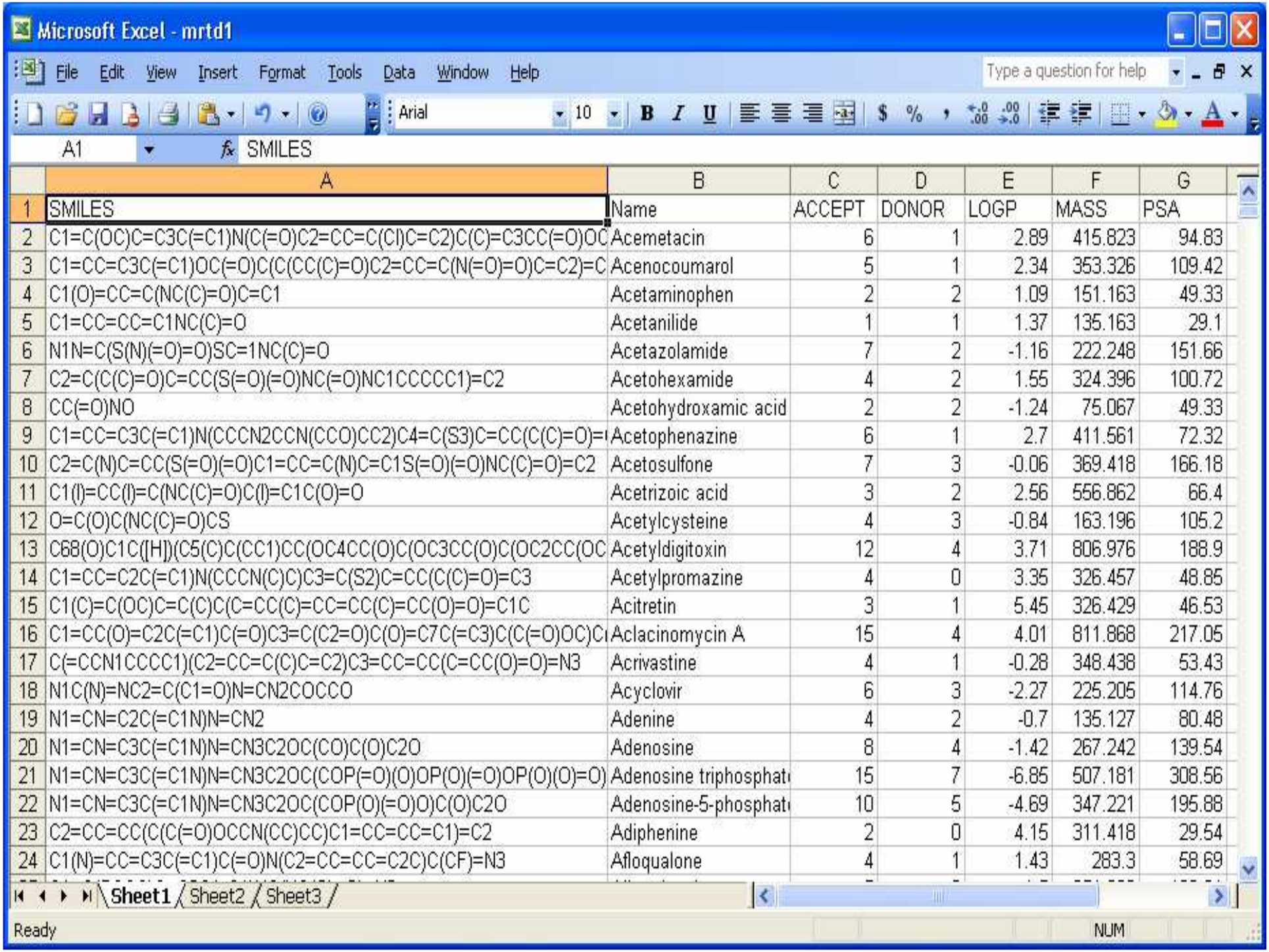
3 - compounds which satisfy 3 requirements - 6% of all compounds

4 - compounds which satisfy 4 requirements - 91% of all compounds

Aggregates formed by tetraiodophenolphthalein, a promiscuous inhibitor, as visualized by transmission electron microscopy.



McGovern SL, Caselli E, Grigorieff N, Shoichet BK: A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J Med Chem* 2002, 45:1712-1722.



	A	B	C	D	E	F	G
1	SMILES	Name	ACCEPT	DONOR	LOGP	MASS	PSA
2	C1=C(OC)C=C3C(=C1)N(C(=O)C2=CC=C(Cl)C=C2)C(C)=C3CC(=O)OC	Acemetacin	6	1	2.89	415.823	94.83
3	C1=CC=C3C(=C1)OC(=O)C(C(CC(C)=O)C2=CC=C(N(=O)=O)C=C2)=C	Acenocoumarol	5	1	2.34	353.326	109.42
4	C1(O)=CC=C(NC(C)=O)C=C1	Acetaminophen	2	2	1.09	151.163	49.33
5	C1=CC=CC=C1NC(C)=O	Acetanilide	1	1	1.37	135.163	29.1
6	N1N=C(S(N)(=O)=O)SC=1NC(C)=O	Acetazolamide	7	2	-1.16	222.248	151.66
7	C2=C(C(C)=O)C=CC(S(=O)(=O)NC(=O)NC1CCCCC1)=C2	Acetohexamide	4	2	1.55	324.396	100.72
8	CC(=O)NO	Acetohydroxamic acid	2	2	-1.24	75.067	49.33
9	C1=CC=C3C(=C1)N(CCCN2CCN(CCO)CC2)C4=C(S3)C=CC(C(C)=O)=C	Acetophenazine	6	1	2.7	411.561	72.32
10	C2=C(N)C=CC(S(=O)(=O)C1=CC=C(N)C=C1S(=O)(=O)NC(C)=O)=C2	Acetosulfone	7	3	-0.06	369.418	166.18
11	C1(I)=CC(I)=C(NC(C)=O)C(I)=C1C(O)=O	Acetrizoic acid	3	2	2.56	556.862	66.4
12	O=C(O)C(NC(C)=O)CS	Acetylcysteine	4	3	-0.84	163.196	105.2
13	C68(O)C1C([H])(C5(C)C(CC1)CC(OC4CC(O)C(OC3CC(O)C(OC2CC(OC	Acetyldigitoxin	12	4	3.71	806.976	188.9
14	C1=CC=C2C(=C1)N(CCCN(C)C)C3=C(S2)C=CC(C(C)=O)=C3	Acetylpromazine	4	0	3.35	326.457	48.85
15	C1(C)=C(OC)C=C(C)C(C=CC(C)=CC=CC(C)=CC(O)=O)=C1C	Acitretin	3	1	5.45	326.429	46.53
16	C1=CC(O)=C2C(=C1)C(=O)C3=C(C2=O)C(O)=C7C(=C3)C(C(=O)OC)C(	Aclacinomycin A	15	4	4.01	811.868	217.05
17	C(=CCN1CCCC1)(C2=CC=C(C)C=C2)C3=CC=CC(C=CC(O)=O)=N3	Acrivastine	4	1	-0.28	348.438	53.43
18	N1C(N)=NC2=C(C1=O)N=CN2COCCO	Acyclovir	6	3	-2.27	225.205	114.76
19	N1=CN=C2C(=C1N)N=CN2	Adenine	4	2	-0.7	135.127	80.48
20	N1=CN=C3C(=C1N)N=CN3C2OC(CO)C(O)C2O	Adenosine	8	4	-1.42	267.242	139.54
21	N1=CN=C3C(=C1N)N=CN3C2OC(COP(=O)(O)OP(O)(=O)OP(O)(O)=O)	Adenosine triphosphat	15	7	-6.85	507.181	308.56
22	N1=CN=C3C(=C1N)N=CN3C2OC(COP(O)(=O)O)C(O)C2O	Adenosine-5-phosphat	10	5	-4.69	347.221	195.88
23	C2=CC=CC(C(C(=O)OCCN(CC)CC)C1=CC=CC=C1)=C2	Adiphenine	2	0	4.15	311.418	29.54
24	C1(N)=CC=C3C(=C1)C(=O)N(C2=CC=CC=C2C)C(CF)=N3	Afloqualone	4	1	1.43	283.3	58.69

CORINA

Automatic generation of three-dimensional atomic COoRdINAtes

Molecular Networks GmbH
Inspiring Chemical Discovery

Partners

Home

Online Demonstrations

Demo - CORINA Interactively

Please enter a structure as SMILES string and an identifier in the form below and press the *Submit* button (or just use "alanin" for demonstration). CORINA will generate 3D coordinates for the given structure. A new page will be generated showing the 3D molecular model if you have RASMOL, CHIME, or some similar program installed on your computer).

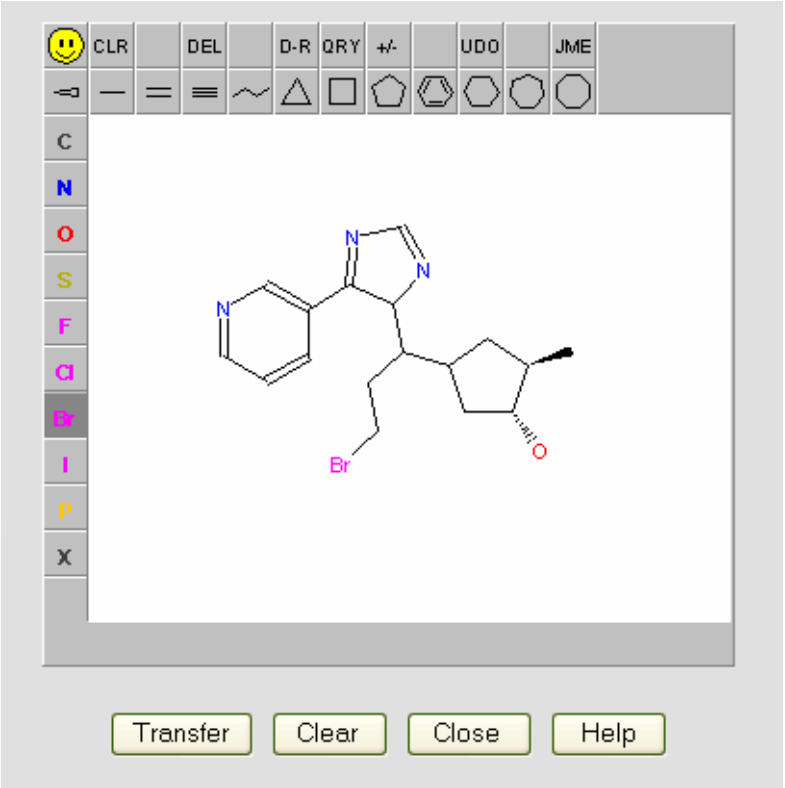
C1=CC=NC=C1C2=NC=NC2C(CCBr

Create Molecule

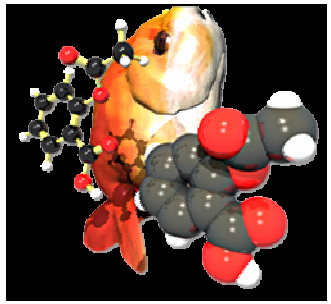
my-molecule

Reset to Demo

Submit



http://www.molecular-networks.com/online_demos/corina_demo.html



Open Babel is a community-driven scientific project including both cross-platform programs and a developer library designed to support molecular modeling, chemistry, and many related areas, including interconversion of file formats and data.

http://openbabel.sourceforge.net/wiki/Main_Page

OpenBabelGUI

File View Help

---- INPUT FORMAT ----

sdf -- MDL MOL format | Format Info

Use this format for all input files (ignore file extensions)

C:\Documents and Settings\Federico Gago\Mis documentos\

del.sdf | ...

Input below (ignore input file)

```
C11H12N2O
WAtclserve06100619443D 0   0.00000     0.00000

 26 27  0  0  0  0  0  0  0  0  0999 V2000
    4.1697    1.9351    0.0487 C   0  0  0  0  0  0  0  0  0  0  0  0
    3.3102    0.6981    0.0026 C   0  0  0  0  0  0  0  0  0  0  0  0
    3.8739   -0.5441    0.2278 C   0  0  0  0  0  0  0  0  0  0  0  0
    3.0857   -1.6786    0.1854 C   0  0  0  0  0  0  0  0  0  0  0  0
    1.7334   -1.5712   -0.0818 C   0  0  0  0  0  0  0  0  0  0  0  0
    1.1697   -0.3294   -0.3060 C   0  0  0  0  0  0  0  0  0  0  0  0
   -0.3041   -0.2124   -0.5977 C   0  0  0  0  0  0  0  0  0  0  0  0
   -1.0210   -0.0513    0.6279 O   0  0  0  0  0  0  0  0  0  0  0  0
   -2.3391    0.0484    0.3027 C   0  0  0  0  0  0  0  0  0  0  0  0
   -3.3612    0.2110    1.1702 C   0  0  0  0  0  0  0  0  0  0  0  0
   -4.5108    0.2633    0.4206 N   0  0  0  0  0  0  0  0  0  0  0  0
   -5.4134    0.3747    0.7580 H   0  0  0  0  0  0  0  0  0  0  0  0
   -4.1500    0.1311   -0.8730 C   0  0  0  0  0  0  0  0  0  0  0  0
   -2.8516    0.0082   -0.9377 N   0  0  0  0  0  0  0  0  0  0  0  0
    1.9574    0.8059   -0.2594 C   0  0  0  0  0  0  0  0  0  0  0  0
    4.5683    2.1379   -0.9453 H   0  0  0  0  0  0  0  0  0  0  0  0
    3.5692    2.7827    0.3793 H   0  0  0  0  0  0  0  0  0  0  0  0
    4.9933    1.7796    0.7456 H   0  0  0  0  0  0  0  0  0  0  0  0
    4.9302   -0.6279    0.4365 H   0  0  0  0  0  0  0  0  0  0  0  0
    3.5261   -2.6490    0.3607 H   0  0  0  0  0  0  0  0  0  0  0  0
    1.1176   -2.4578   -0.1156 H   0  0  0  0  0  0  0  0  0  0  0  0
   -0.4803    0.6520   -1.2380 H   0  0  0  0  0  0  0  0  0  0  0  0
   -0.6472   -1.1150   -1.1034 H   0  0  0  0  0  0  0  0  0  0  0  0
   -3.2886    0.2851    2.2452 H   0  0  0  0  0  0  0  0  0  0  0  0
```

CONVERT

min warning level displayed

Start import at molecule # specified

End import at molecule # specified

All input files describe a single molecule

Continue with next object after error, if possible

Compress the output with gzip

Delete Hydrogens

Add Hydrogens

Add Hydrogens appropriate for pH

Convert dative bonds e.g.[N+](=O)[O-] to N(=O)=O

Center Coordinates

Join all input molecules into a single output molecule

Convert only molecules matching SMARTS:

Convert only molecules NOT matching SMARTS:

---- OUTPUT FORMAT ----

mol2 -- Sybyl Mol2 format | Format Info

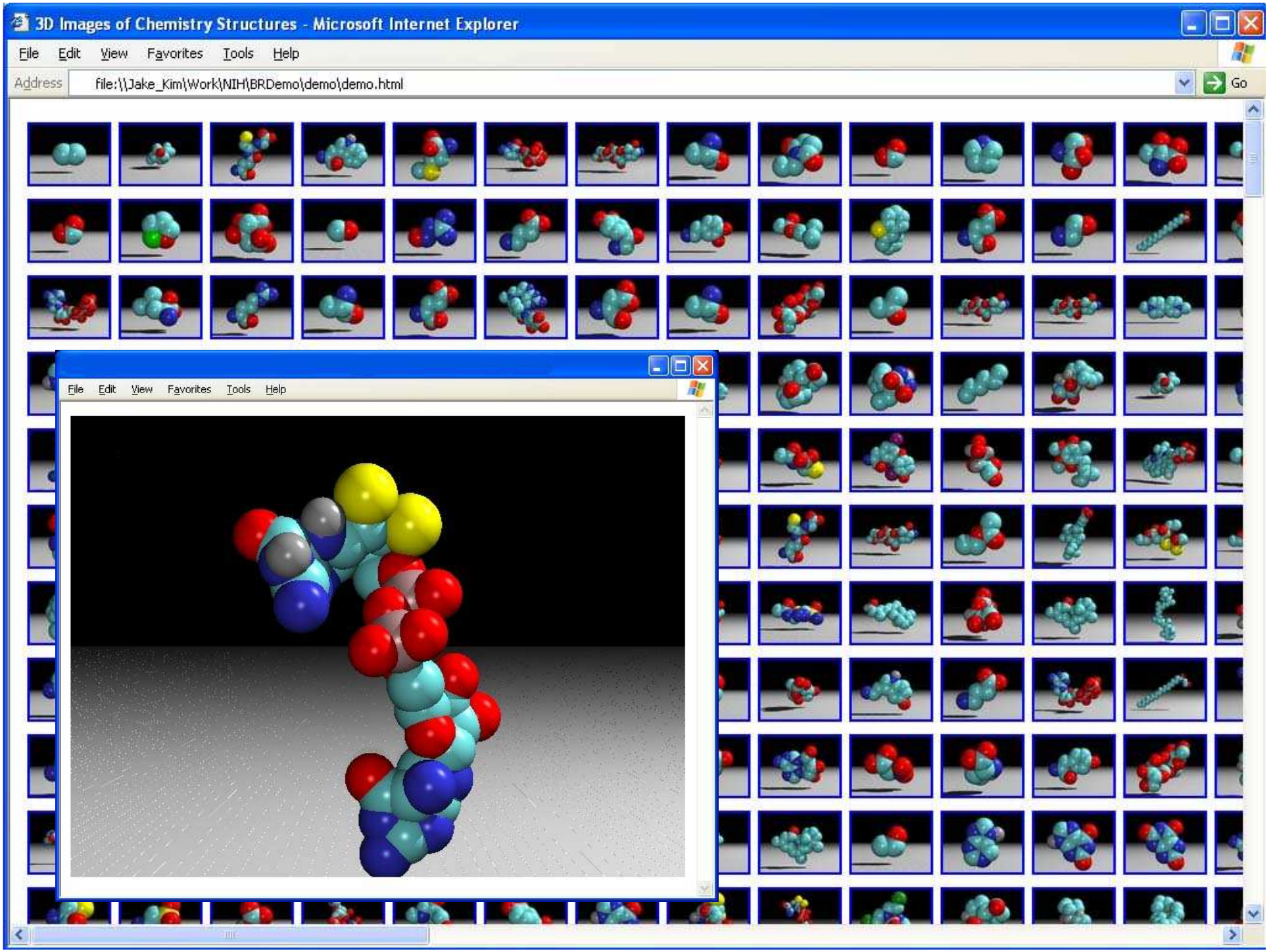
Output file | ...

Output below only (no output file)

```
1 molecule converted

@<TRIPOS>MOLECULE
C11H12N2O
 26 27 0 0 0
SMALL
GASTEIGER
Energy = 0

@<TRIPOS>ATOM
      1 C1          4.1697    1.9351    0.0487 C.3     1 <1>      -0.0397
      2 C2          3.3102    0.6981    0.0026 C.ar    1 <1>      -0.0501
      3 C3          3.8739   -0.5441    0.2278 C.ar    1 <1>      -0.0588
      4 C4          3.0857   -1.6786    0.1854 C.ar    1 <1>      -0.0611
      5 C5          1.7334   -1.5712   -0.0818 C.ar    1 <1>      -0.0553
      6 C6          1.1697   -0.3294   -0.3060 C.ar    1 <1>      -0.0101
      7 C7         -0.3041   -0.2124   -0.5977 C.3     1 <1>       0.1163
      8 O1         -1.0210   -0.0513    0.6279 O.3     1 <1>      -0.4706
      9 C8         -2.3391    0.0484    0.3027 C.ar    1 <1>       0.2329
     10 C9         -3.3612    0.2110    1.1702 C.ar    1 <1>       0.0652
     11 N1         -4.5108    0.2633    0.4206 N.ar    1 <1>      -0.3477
     12 H1         -5.4134    0.3747    0.7580 H       1 <1>       0.1663
     13 C10        -4.1500    0.1311   -0.8730 C.ar    1 <1>       0.0973
     14 N2         -2.8516    0.0082   -0.9377 N.ar    1 <1>      -0.2022
     15 C11         1.9574    0.8059   -0.2594 C.ar    1 <1>      -0.0524
     16 H2          4.5683    2.1379   -0.9453 H       1 <1>       0.0278
     17 H3          3.5692    2.7827    0.3793 H       1 <1>       0.0278
     18 H4          4.9933    1.7796    0.7456 H       1 <1>       0.0278
     19 H5          4.9302   -0.6279    0.4365 H       1 <1>       0.0620
```


3D Images of Chemistry Structures - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Address
file:\\Jake_Kim\Work\NIH\BRDemo\demo\demo.html
Go
File Edit View Favorites Tools Help

United States National Library of Medicine

ChemIDplus Advanced

http://chem.sis.nlm.nih.gov/chemidplus/

News | SIS Home | Site | About Us | Contact | Help

Env. Health & Toxicology | TOXNET | ChemIDplus Lite | Advanced

Search | Clear | History | Help

Display 5 results

Substance Identification

Name/Synonym | Equals

Data is available for 380,101 records.

Toxicity

Test: (any) | between

(mg/kg or ppm)

Species: (any)

Route: (any)

Effect: (any)

Toxicity data is available for 139,354 records.

Physical Properties

Melting Point

between

Either Measurement Type

Physical property data was provided by Syracuse Research Corporation and is available for 25,661 records.

Locator Codes

(any)

Structure

View | Help

Powered by ChemAxon Marvin

Structure Search Options

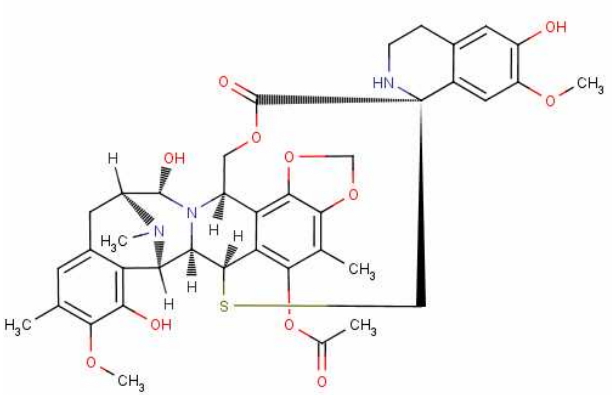
- Substructure Search
- Similarity Search 80 %
- Exact (parent only)
- Flex (parent, salts, mixture) *NEW*
- Flexplus (parent, all variations) *NEW*

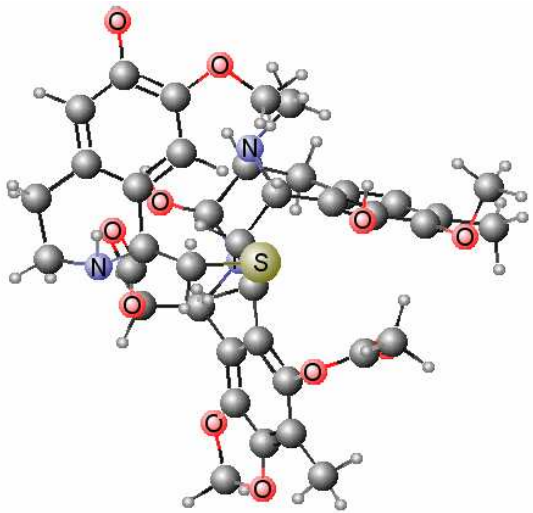
Display structures using

Marvin | Chime | Change

Structure data is available for 263,354 records.

Molecular Weight





DrugBank

http://drugbank.ca/

The DrugBank database is a unique bioinformatics and cheminformatics resource that combines detailed drug (i.e. chemical, pharmacological and pharmaceutical) data with comprehensive drug target (i.e. sequence, structure, and pathway) information.

The database contains nearly **4800 drug entries** including >1,350 FDA-approved small molecule drugs, 123 FDA-approved biotech (protein/peptide) drugs, 71 nutraceuticals and >3,243 experimental drugs. Additionally, more than 2,500 non-redundant protein (i.e. drug target) sequences are linked to these FDA approved drug entries. Each DrugCard entry contains more than 100 data fields with half of the information being devoted to drug/chemical data and the other half devoted to drug target or protein data.