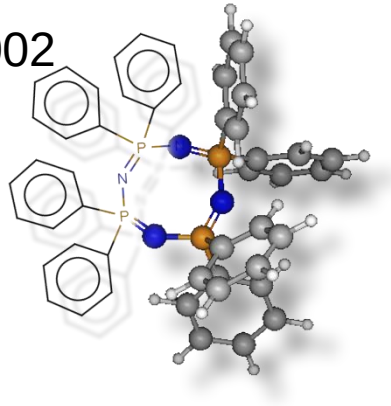
Flexible alignment in 3D & applications

Adrian Kalaszi, G. Imre, O. Farkas, M. Vargyas

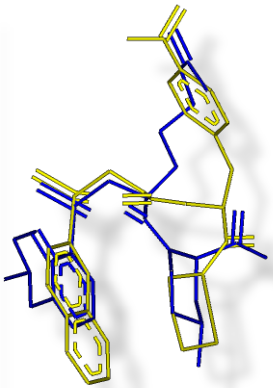
ChemAxon

ChemAxon in 3D

3D Structure generation – released in 2002



3D Alignment – released in 2009



3D Similarity – ligand based virtual screening 2010 Q3

fast & good quality ?

Flexible Alignment

Function

atomic coordinates $f(x)$

Gradient $\delta f(x) / \delta \Theta$

Kinematics: infinitesimal rotational tensor

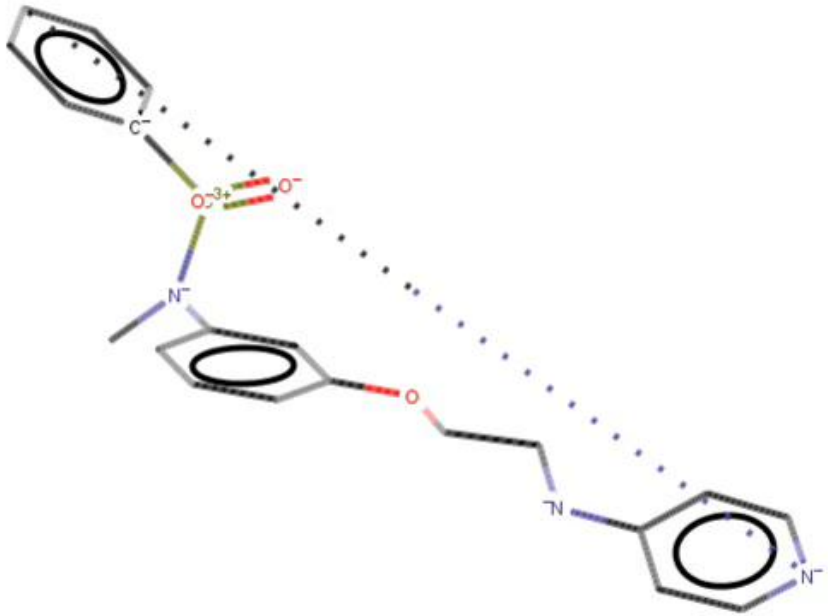
http://mathworld.wolfram.com/InfinitesimalRotation.html

Hurst, J. Chem. Inf Comput. Sci. 1994, 34, 190-196

MDS-CG (Multi-Dimensional Search in Conjugate Gradient)

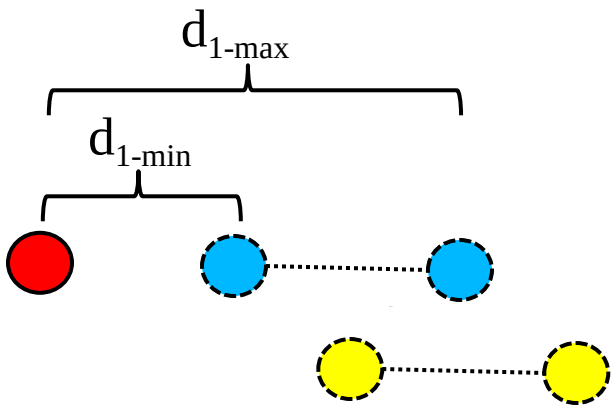
Farkas, Ö., Schlegel, H. B. J. Mol. Struct.-Theochem 2003, 666, 31-39.

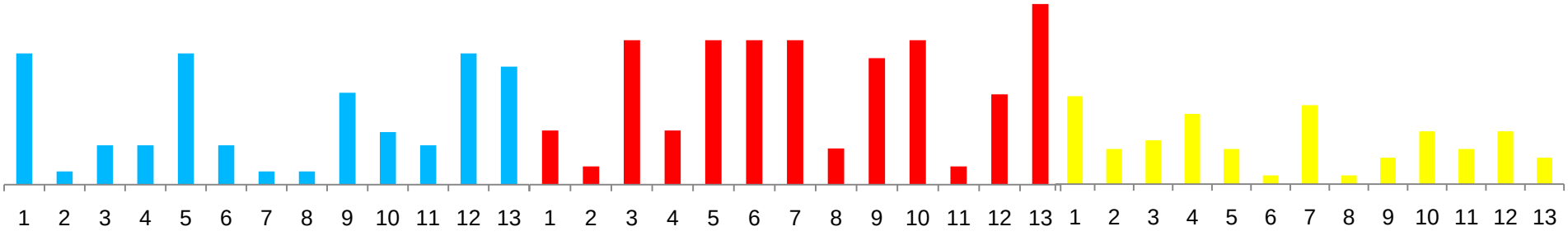
Minimum / maximum distance



Watch the video here: http://www.youtube.com/watch?v=FgMsvDmtGkA

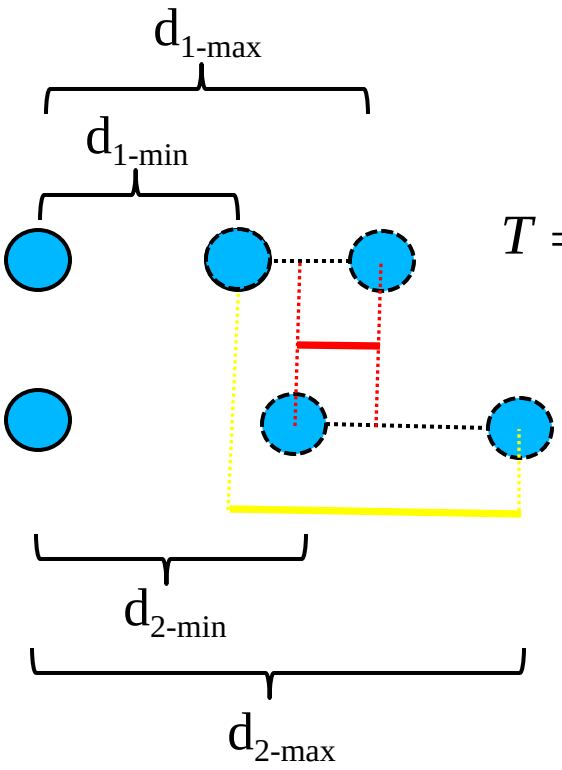
Atomic histograms





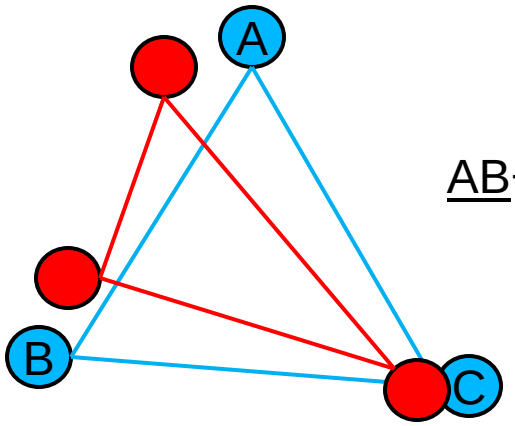
Distance range Tanimoto

For atom pair comparison



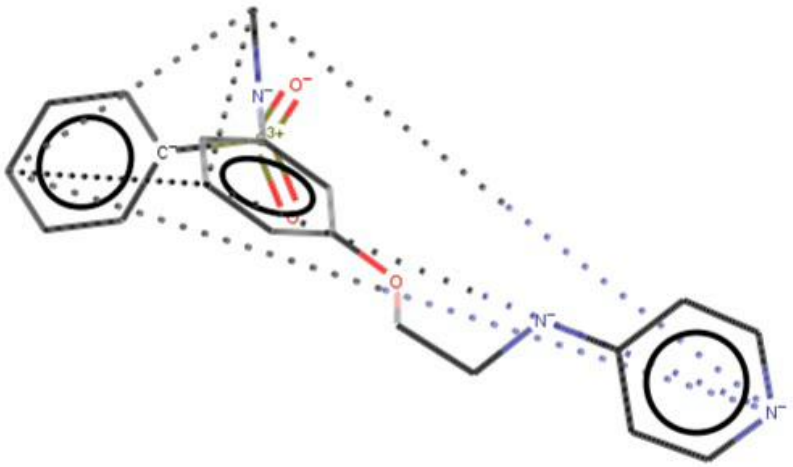
$$T = \frac{\min(d_{1-\max}, d_{2-\max}) - \max(d_{1-\min}, d_{2-\min})}{\max(d_{1-\max}, d_{2-\max}) - \min(d_{1-\min}, d_{2-\min})}$$

Triangle inequality



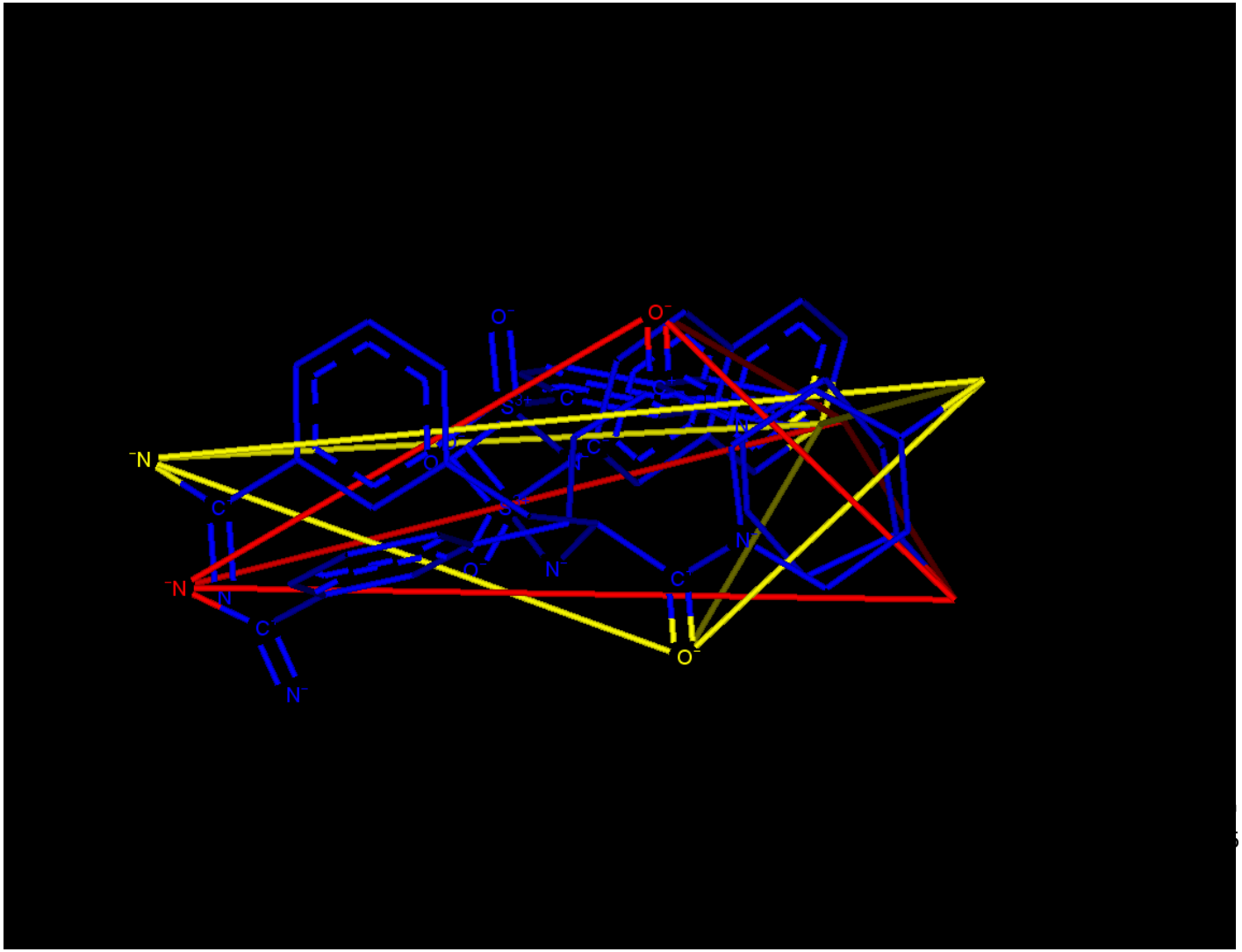
$\underline{AB}+\underline{AC} > \underline{BC}$

Experiment



Watch the video here: http://www.youtube.com/watch?v=22lfa5HeCts

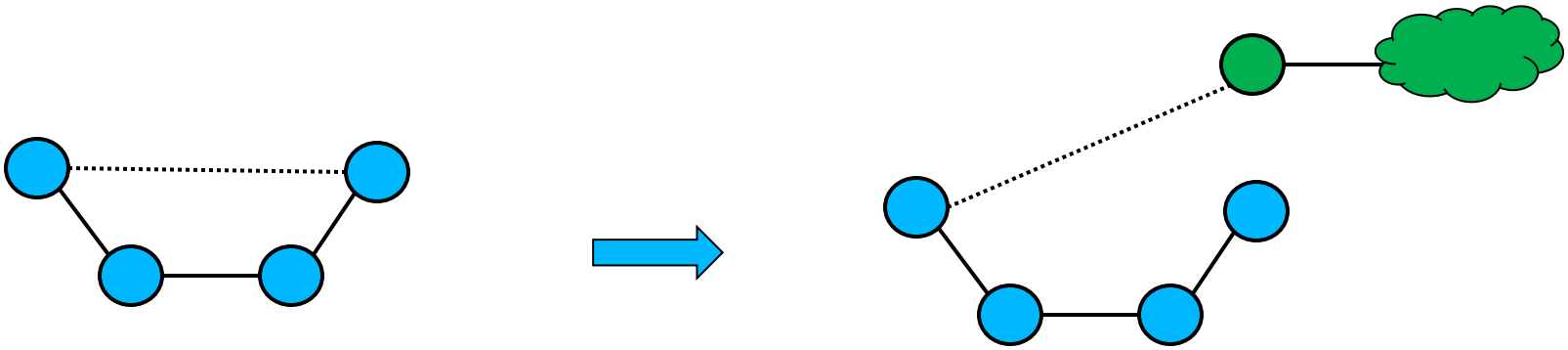
Internal distances



Distances between molecules

Function atomic coordinates or distances $f(x)$

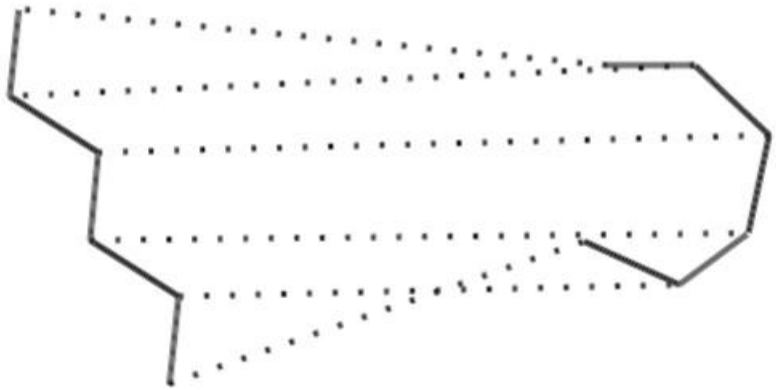
Gradient $\delta f(x) / \delta \Theta$



Add +6DF on molecules that we wish to translate & rotate

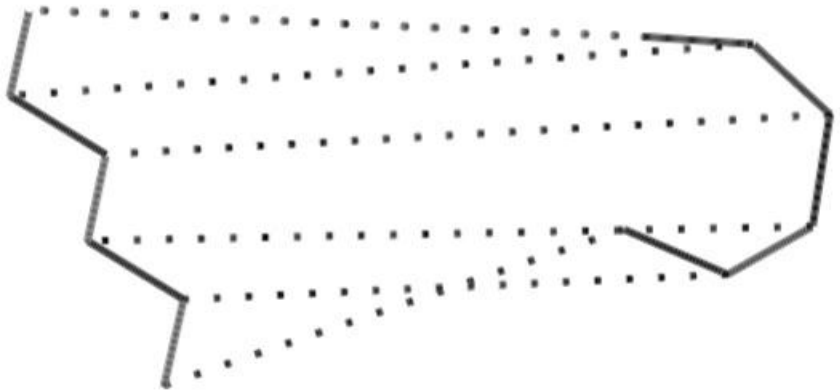
Optimization on mixed internal coordinate / Cartesian domain

Jump to external distances



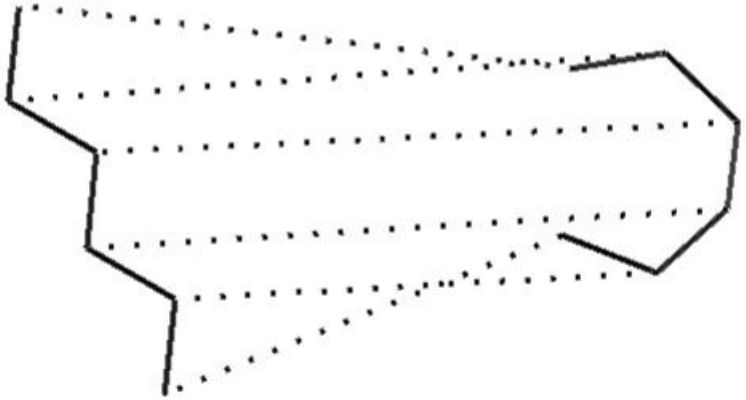
Watch the video here: http://www.youtube.com/watch?v=zNtxblx0-Dw

Jump to external distances



Watch the video here: http://www.youtube.com/watch?v=8KJqz5Emco0

Jump to external distances



Watch the video here: http://www.youtube.com/watch?v=kXfoukkAmQQ

Quaternion fit

2 molecules: translate & rotate one

Pros:

- **minimize RMSD on selected atom pairs (best alignment)**
- **fast (one step) – analytically solving the equation**

Cons:

- **need to know the atom mapping**
- **no flexibility**

Quaternion & Flexibility Hybrid

Delegate translation and rotation part to quaternion fit

Dihedrals are optimized

Pros:

- **very fast**
- **robust: usually finds the best alignment**

Cons:

- **only quadratic constraints are allowed**
- **two molecules per alignment**

Still need to solve the atom/atom mapping

Calculations before atom / atom match

Select atoms

Colors (extended atom types / pharmacophore types)

Topological features (e.g.: longest chain start/end/center)

Ring centers (aromatic, aliphatic)

Calculate

min/max internal distance ranges

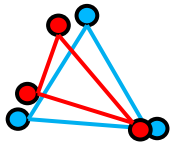
atomic histograms for selected atoms

+once in the life for every molecule

-slow (~ 10k mol/day)

Combination for atom pairs

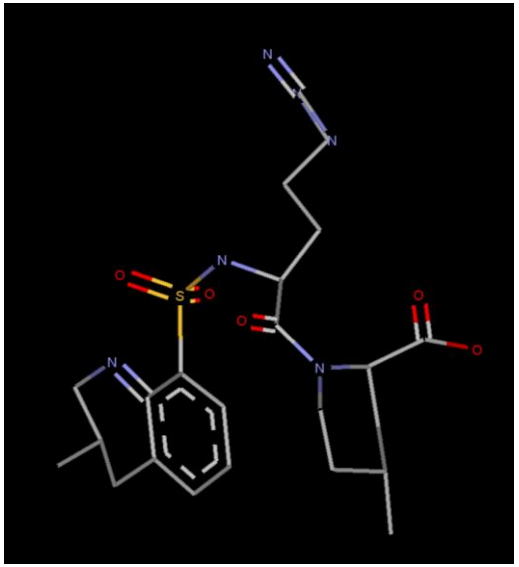
Backtracking with heuristics:

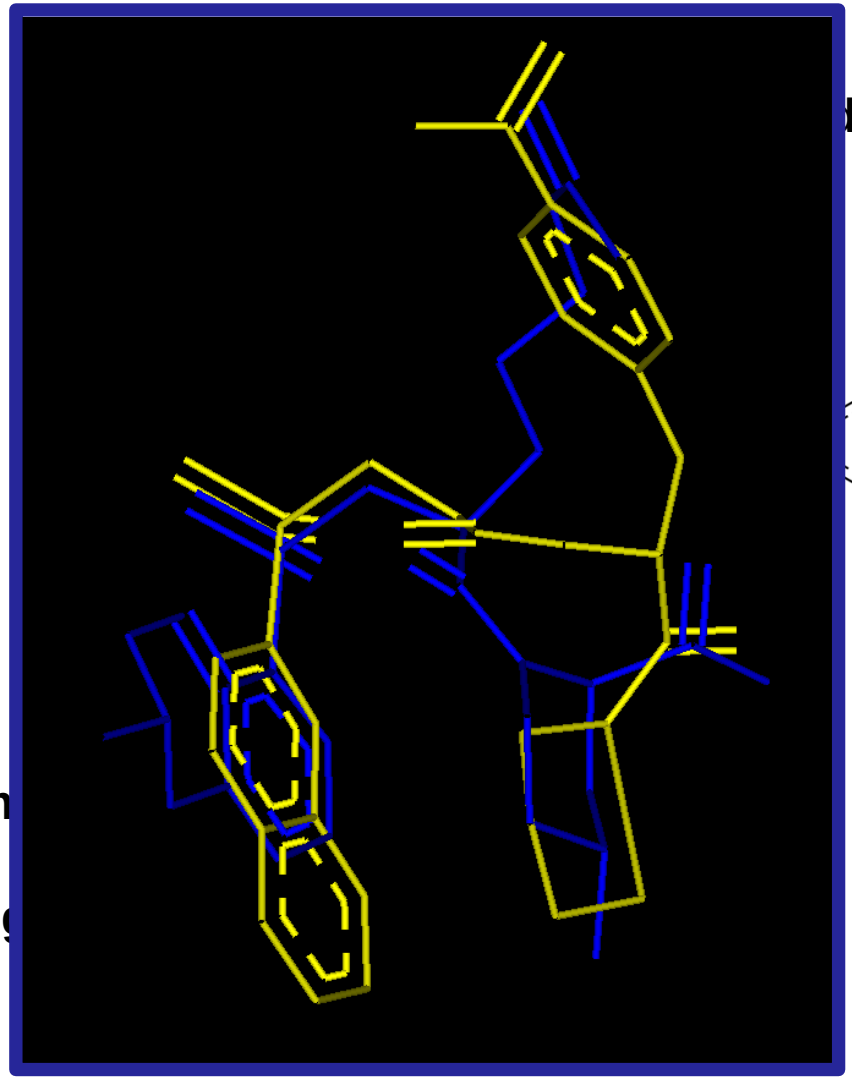
1. Same colors of atoms for the pairs
2. Distance ranges OK for any pairs of mapped atoms $T = \frac{\text{—}}{\text{———————}}$
3. Triangle inequality for any triplet of maps 
4. Quaternion Flexible Hybrid Alignment on map (filter for “chirality”)

Guaranteed: No good solution from the conformational space is lost!

Example: Thrombin inhibitors

1dwc





dwd

Sam

Distance rang

Combination for atom pairs

Backtracking with heuristics returns:

1. Atom/atom mapping
2. Molecules aligned
3. SCORE: Tanimoto of atomic histograms of the mapped atoms

Development version with default scoring weights (1, 0)

Test on DUD

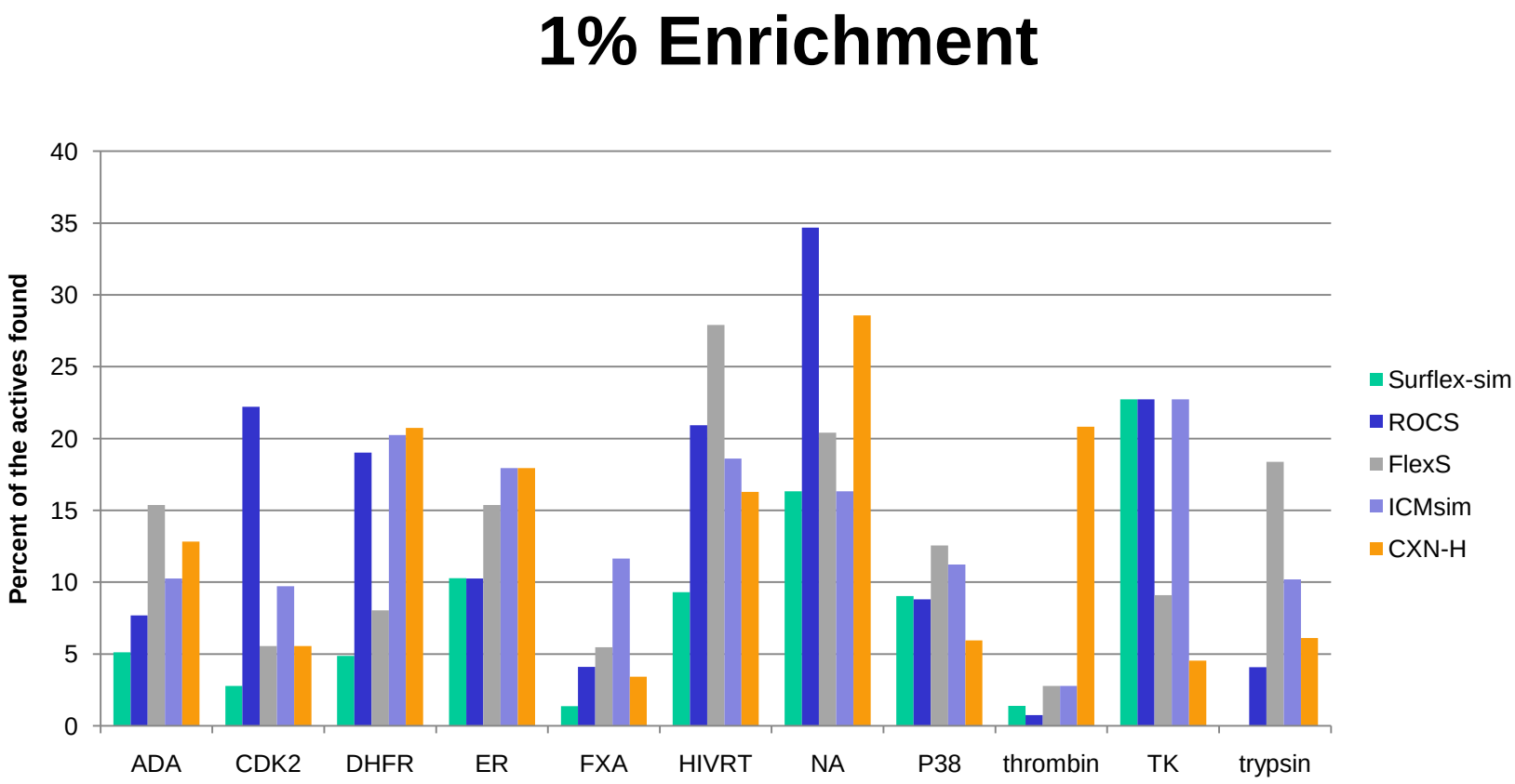
ADA	Adenosine deaminase	39/927
CDK2	Cyclin dependent kinase 2	72/2074
DHFR	Dihydrofolate reductase	410/8367
ER	Estrogen receptor antagonist	39/1448
FXA	Factor Xa	146/5745
HIVRT	HIV reverse transcriptase	43/1519
NA	Neuraminidase	49/1874
P38	P38 mitogen activated protein kinase	454/9141
THR	Thrombin	72/2456
TK	Thymidine kinase	22/891
TRY	Trypsin	49/1664

ROCS
Surflex-sim
FlexS
ICMsim

DOCK
Surflex-Dock
FRED
FlexX
ICM

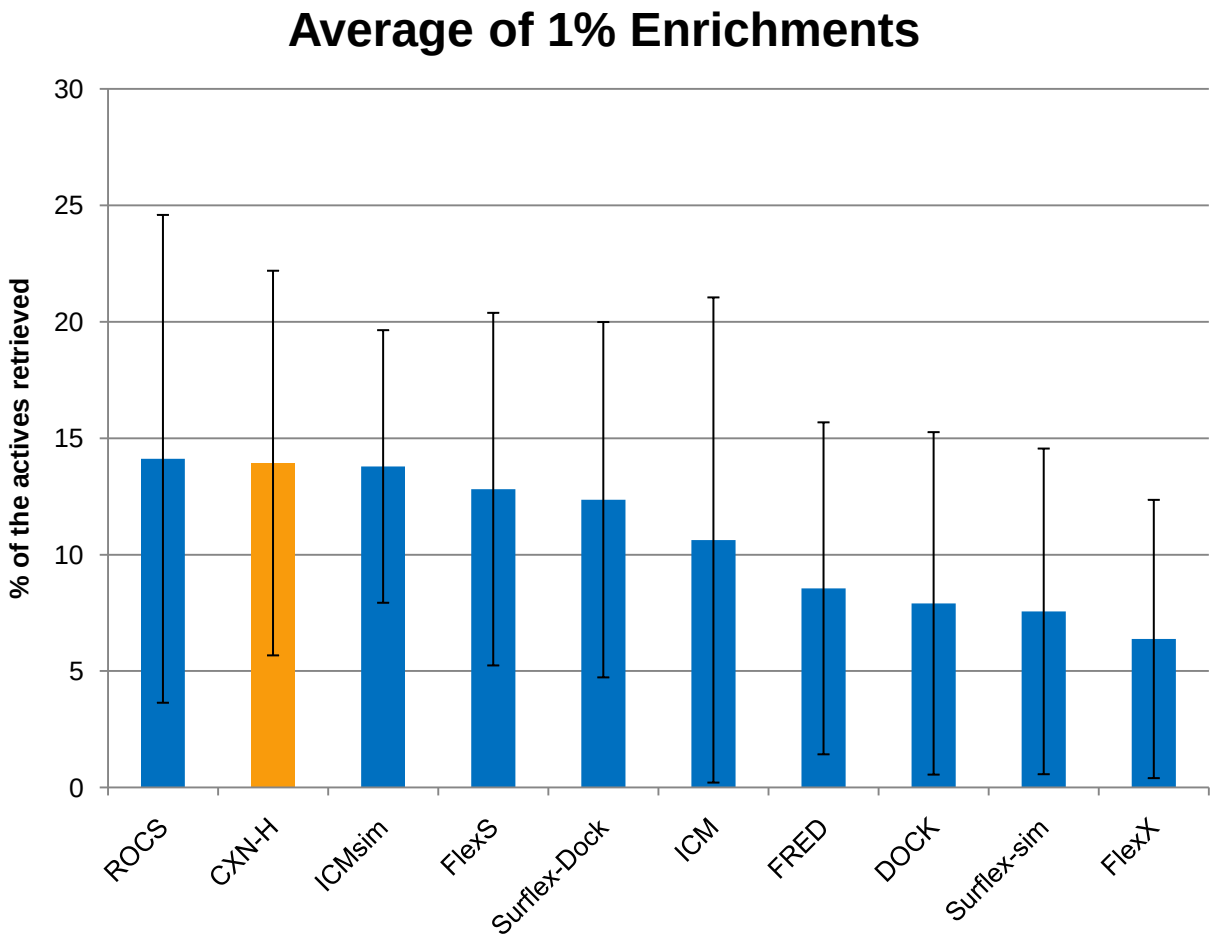
Giganti et al. J. Chem. Inf. Model. 2010, 50, 992

Test on DUD



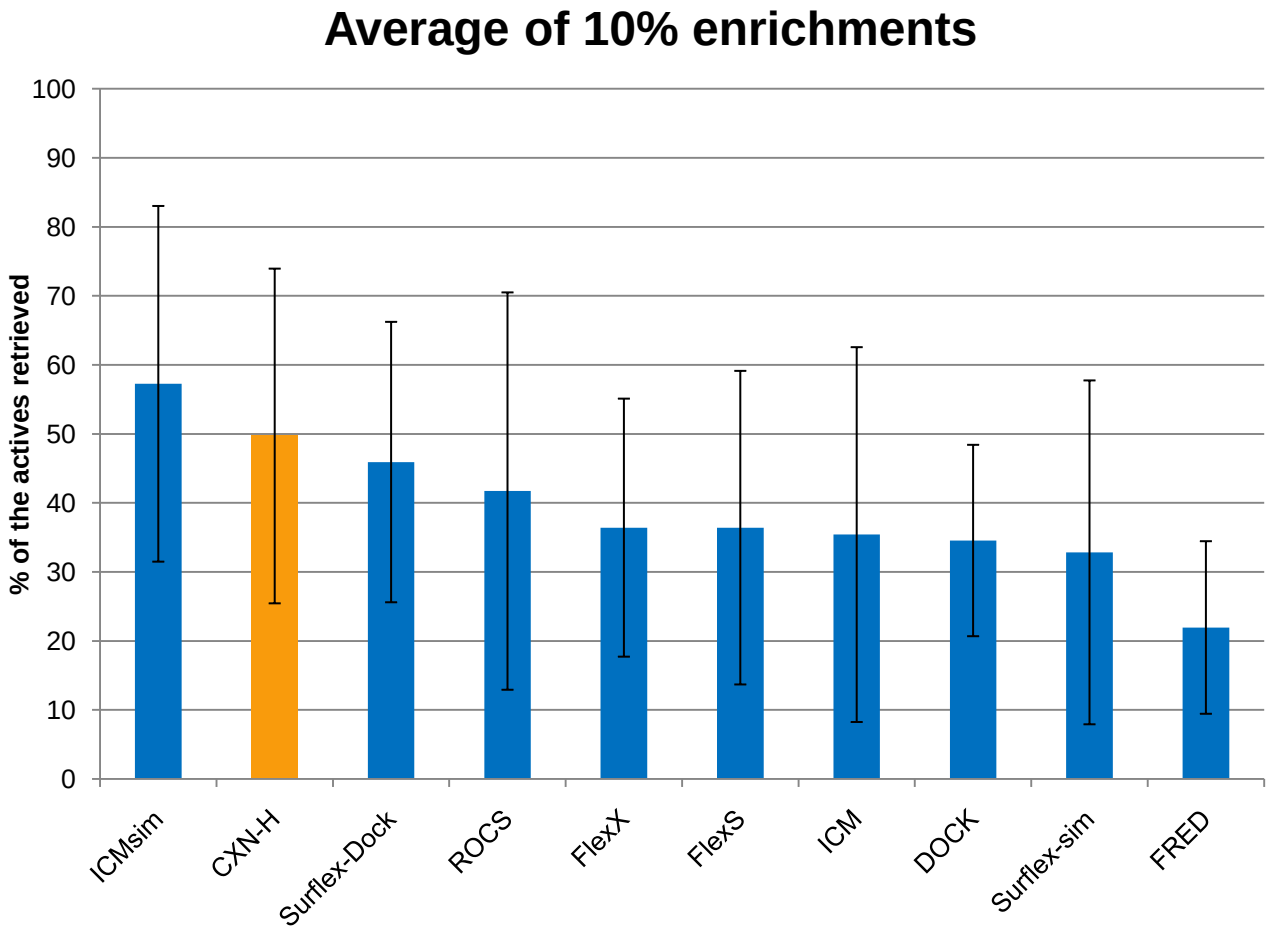
Giganti et al. J. Chem. Inf. Model. 2010, 50, 992

Test on DUD



Giganti et al. J. Chem. Inf. Model. 2010, 50, 992

Test on DUD



Giganti et al. J. Chem. Inf. Model. 2010, 50, 992

Test on DUD

Speed

average time per compound (s)		
CXN-H	0.07	Intel Q6600 2.4 GHz
ROCS	0.5	Intel Xeon 2.4 GHz
FRED	1.0	
ICMsim	2.4	
Surflex-sim	6.7	
FlexS	6.9	
Surflex-dock	14.6	
FLEXX	15.6	
ICM	17.7	

Giganti et al. J. Chem. Inf. Model. 2010, 50, 992

Conclusion

Screen3D:

Further developments: improve scoring

comparable to other methods

high speed

expected release in 2010

Evaluators are welcome!