

Drug Discovery (Drama?)

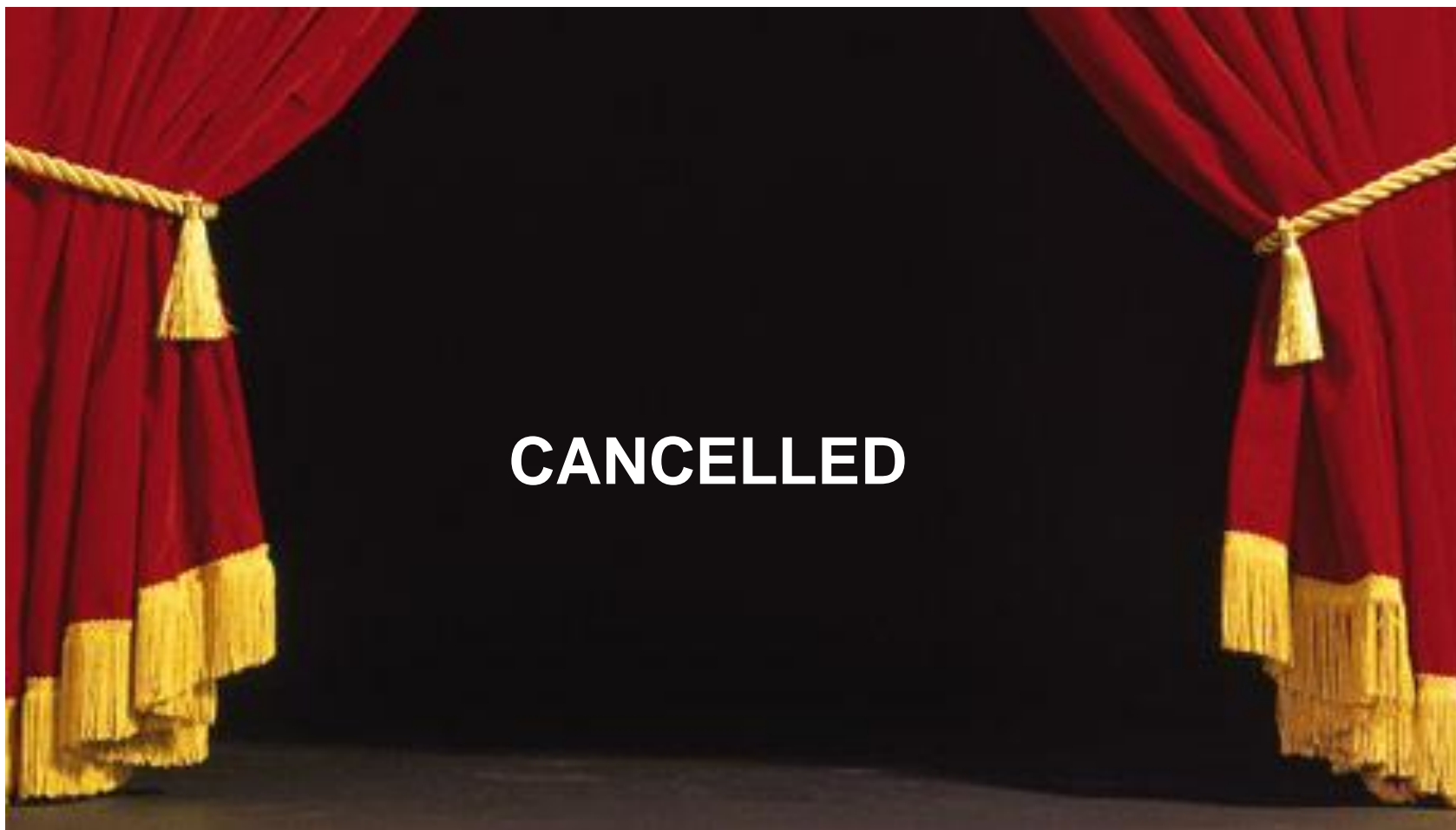
Tools for similarity calculations

Presented by Miklos Vargyas

EUGM 2010 Budapest



ChemAxon
Solutions for Cheminformatics



Is Drug Discovery a Drama?

No!

ChemAxon Rocks!

... not ROCS!

The Art of Componentry



- Cheminformatics specialized
- Toolkit / components provider
- Customer decides based on needs, strategies, politics
- Ease or replacement / upgrade

“Shimano of ChemInformatics”



Screen performs high throughput ligand based virtual screening of compound libraries using similarity comparisons by various molecular descriptors.

Available descriptors

- ChemAxon chemical fingerprint (CF)
- ChemAxon pharmacophore fingerprint (PF)
- BCUT
- Scalars (logP, logD, Szeged index ...)
- custom descriptors, in-house fingerprints

Availability

- JChem for Excel
- Instant Jchem
- JChemBase
- JChem Oracle cartridge
- JChem Web Services
- KNIME
- PipelinePilot
- InforSense workflow platform
- standalone command line application programs

Live demo – IJC

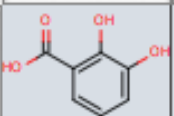
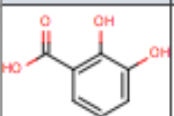
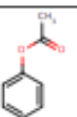
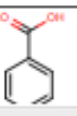
Instant JChem 5.3.1

1 / 7

localdb [as admin]

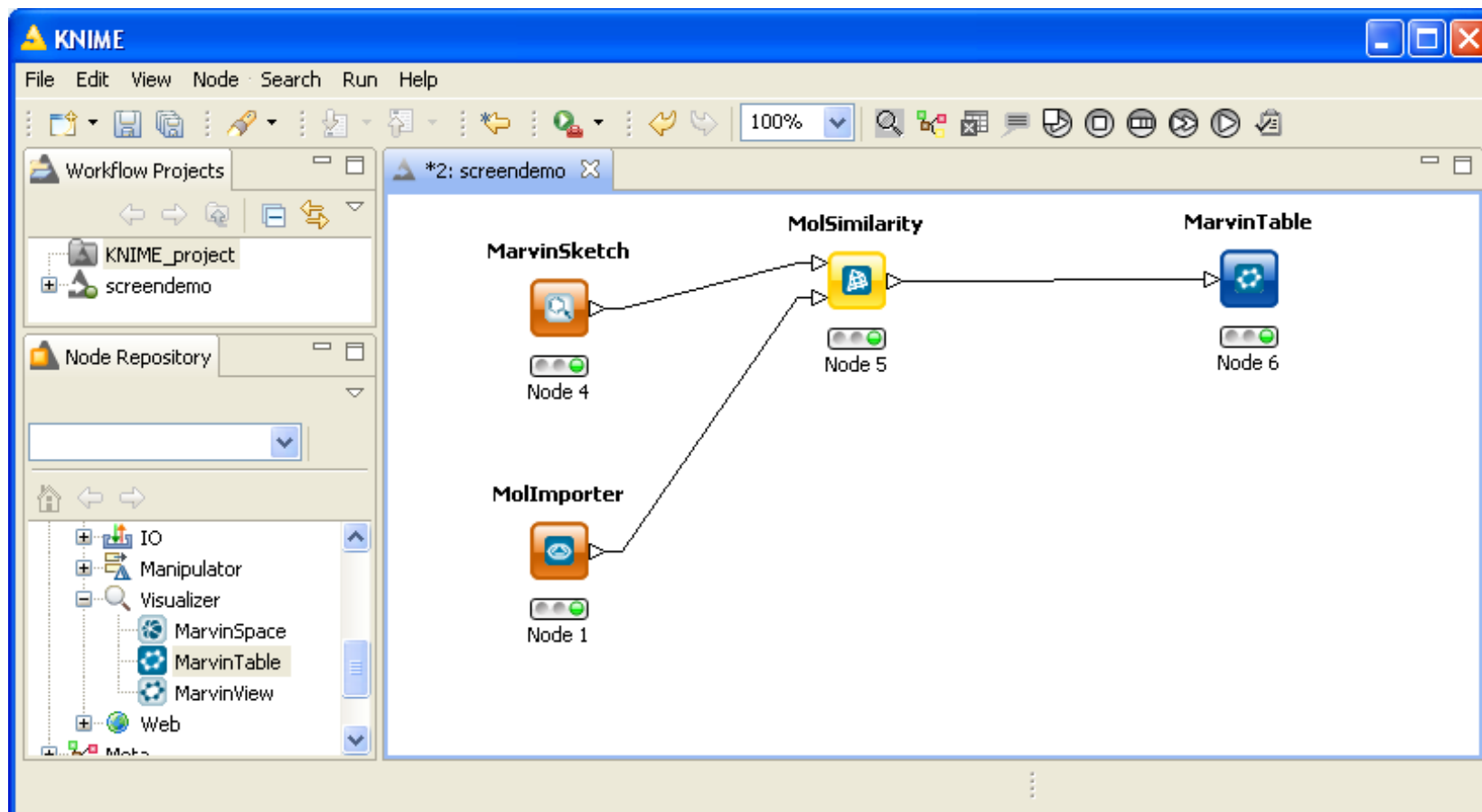
- Pubchem demo
 - Pubchem tabular f
 - Pubchem grid view
 - Pubchem single re
- Wombat (compound v
- Wombat (activities vie
- README.txt

Query Browse

	Structure	Cdid	Mol Weight	Formula	IUPAC name	Donors	Acceptors	Rot bonds	DB regid	DB name	XLogP
1		1,413	154.12	C7H6...	2,3-dihydroxybenzoic acid	3	4	1	C00196	KEGG	1.39
2		1,015	154.12	C7H6...	2,3-dihydroxybenzoic acid	3	4	1	2-3-D...	BioCyc	1.39
3		1,100	152.15	C8H8...	4-hydroxyacetophenone	1	3	2	4-HY...	BioCyc	1.48
4		1,106	138.12	C7H6...	4-hydroxybenzoic acid	2	3	1	4-hyd...	BioCyc	1.16

Pubchem demo: 7 out of 1,000 rows.

Live demo – KNIME



Live demo – Ajax

Ajax interface to JChem web services

http://www.chemaxon.com/ajax/

Anti-pattern...ncyclopedia Working ho...emAxonWiki JKlustor ChemAxon Apple SZTAKI Szótár Magyar Helyesírás

EUGM 19-20 May 2010, register ... Ajax interface to JChem web servi...

ChemAxon

Navigation

- Query
- Insert
- Export
- Print
- About

Hit alignment Results: 8 editexample

No.	Structure	Similarity	ID	Formula	Molweight
1		0.706	6100	C ₇ H ₆ O ₃	138.1207
2		0.64	5787	C ₇ H ₆ O ₄	154.1201
3		0.64	6103	C ₁₄ H ₁₀ O ₉	322.2238

Applet chemaxon/marvin/applet/JMSketch notstarted

Live demo – command line

screenmd caffeine.mol -a demo -k farmakofo -o SDF results.sdf
mview results.sdf

#	structure	CD_ID	q1_PF_Tan	q1_PF_Euc
1		36	0.35593218	10.954452
2		448	0.375	10.770329
3		520	0.4285714	12.4499
4		588	0.4705882	12.247449

Optimization of screening

Standard metrics

$$T(x,y) = \frac{\sum_i |x_i - y_i|}{\sum_i \max(x_i, y_i)} = 1 - \frac{\sum_i \min(x_i, y_i)}{\sum_i x_i + \sum_i y_i - \sum_i \min(x_i, y_i)}$$

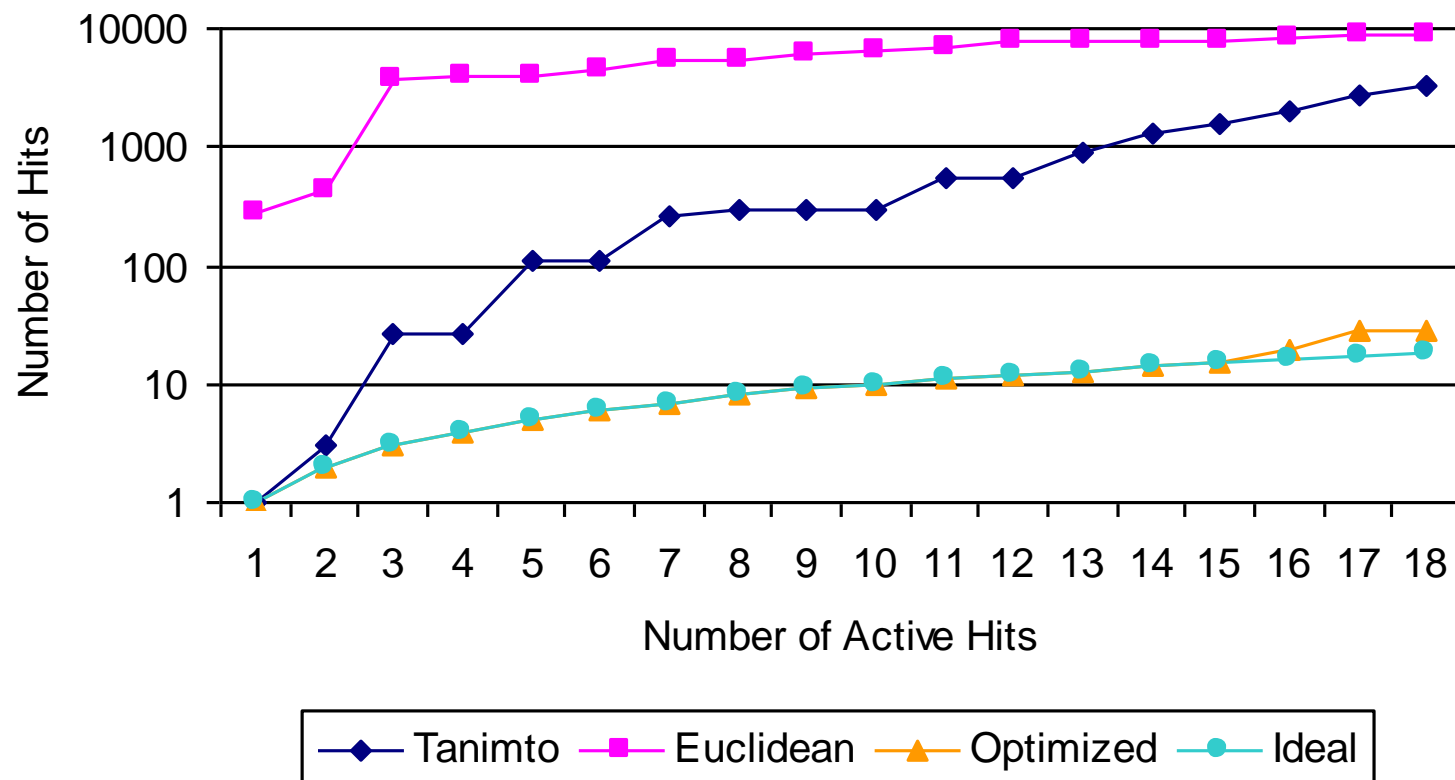
$$E(x,y) = \sqrt{\sum_i (x_i - y_i)^2}$$

Tunable parametrised metrics

$$T(x,y) = 1 - \frac{\sum_i s_i \min(x_i, y_i)}{\alpha \left(\sum_i x_i - \sum_i s_i \min(x_i, y_i) \right) + (1-\alpha) \left(\sum_i y_i - \sum_i s_i \min(x_i, y_i) \right) + \sum_i s_i \min(x_i, y_i)}$$

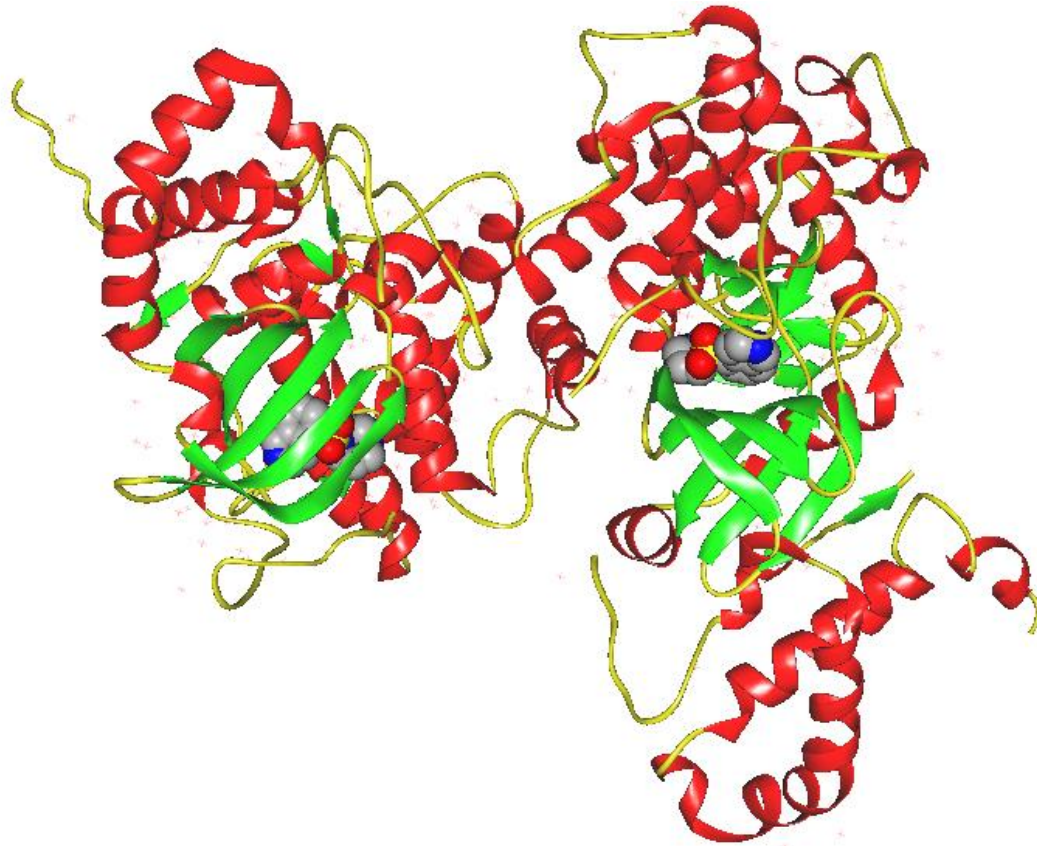
$$E(x,y) = \sqrt{\sum_{x_i < y_i} w_i \alpha (x_i - y_i)^2 + \sum_{x_i \geq y_i} w_i (1-\alpha) (x_i - y_i)^2}$$

Validation of screening configuration



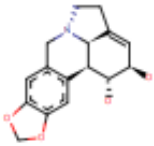
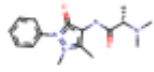
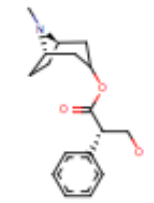
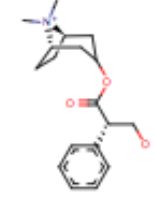
Validation of virtual hits

- Rho-kinase

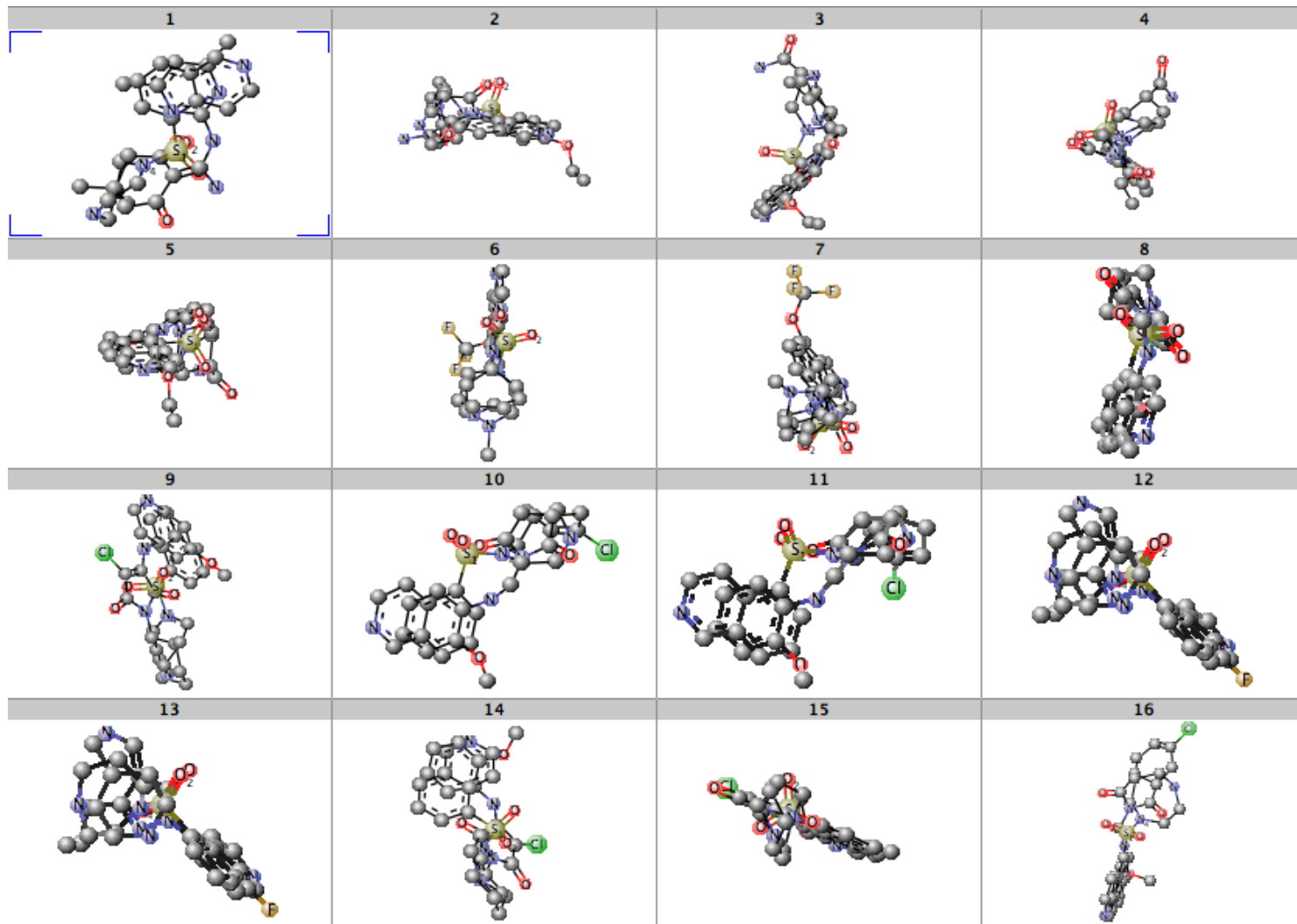


Virtual hits

pftansa.hits.sdf - MarvinView 5.3.0branch

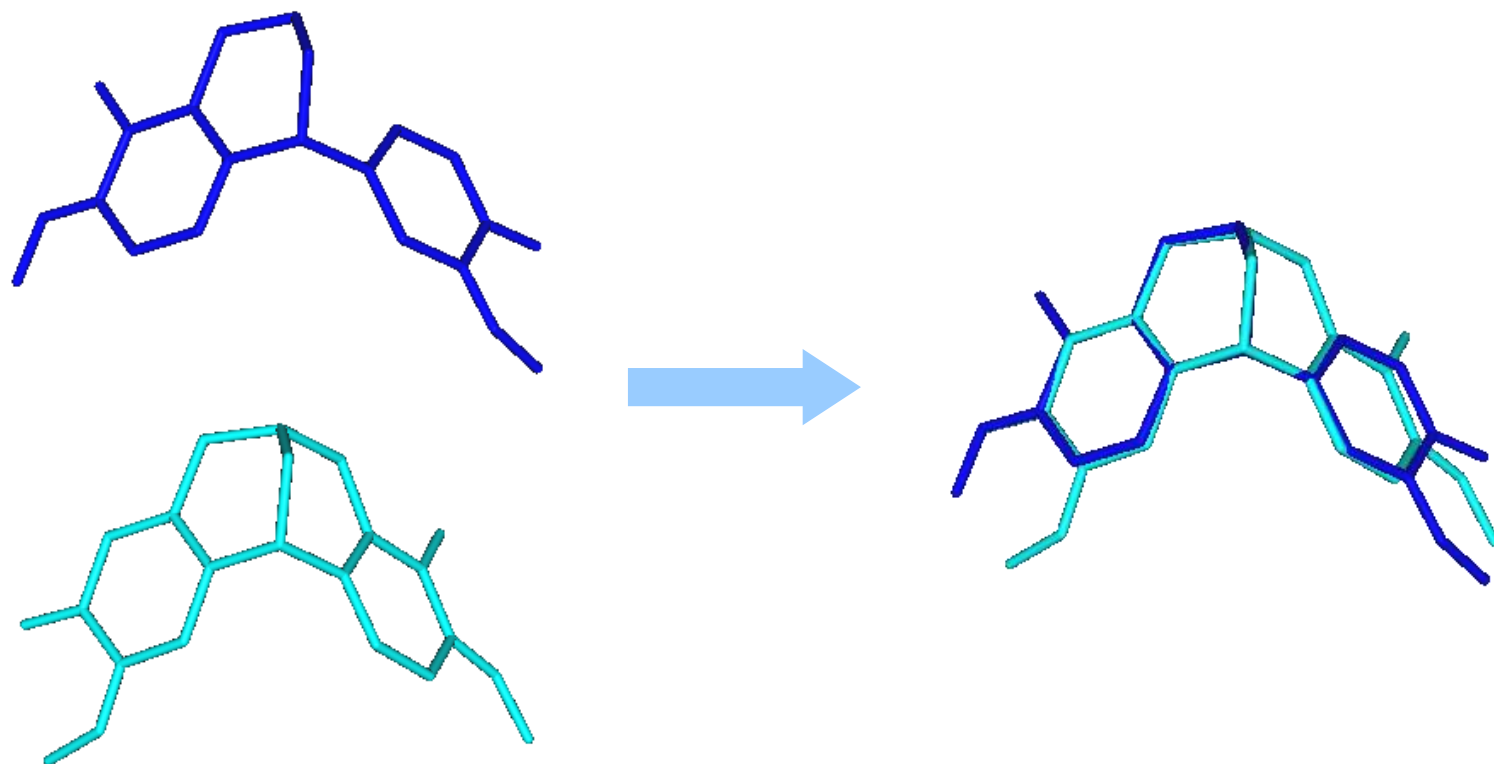
#	structure	SMolName	q1_PF_Tan	PMAP
1		ZINC00000024	0.2714178	r;r;r;r;r;h;d;h;h;h;h;a;d;a/d;h;a;h;a
2		ZINC00000031	0.2742359	h;r;r;r;r;d;h;a;h;h;d;h;a;r;r;r;r;h
3		ZINC00000056	0.15298718	h;d;h;h;h;h;h;a;h;a;h;h;a/d;r;r;r;r
4		ZINC00000057	0.1592223	h;+ /d;h;h;h;h;h;a;h;a;h;h;a/d;r;r;r;r;h

Aligned hits

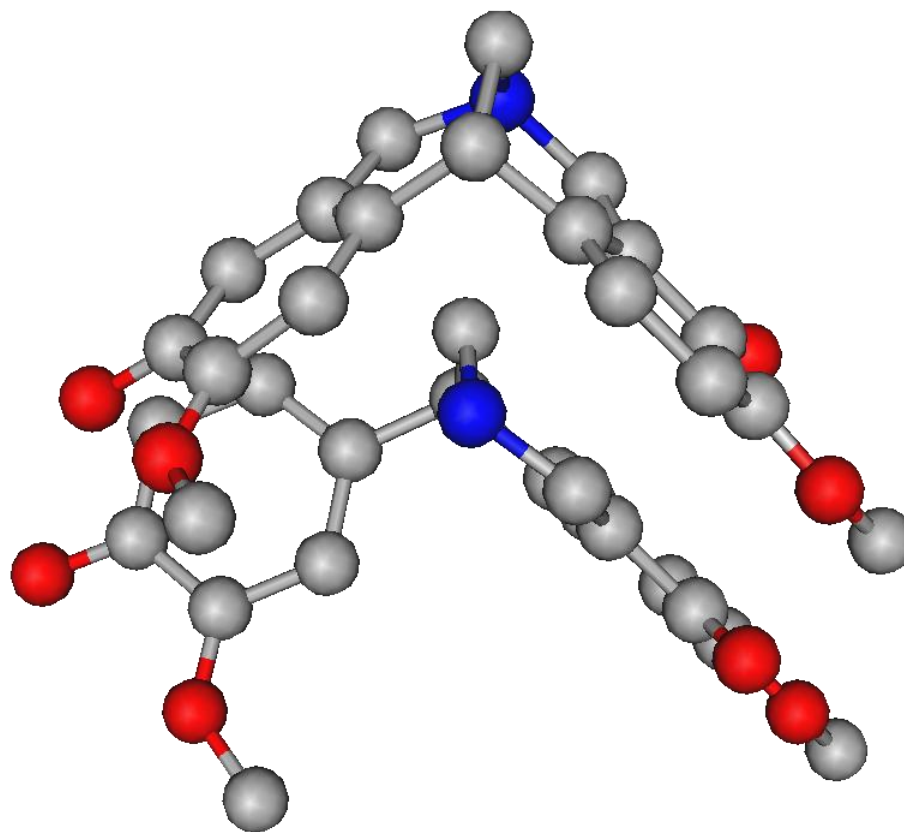


Under the hood

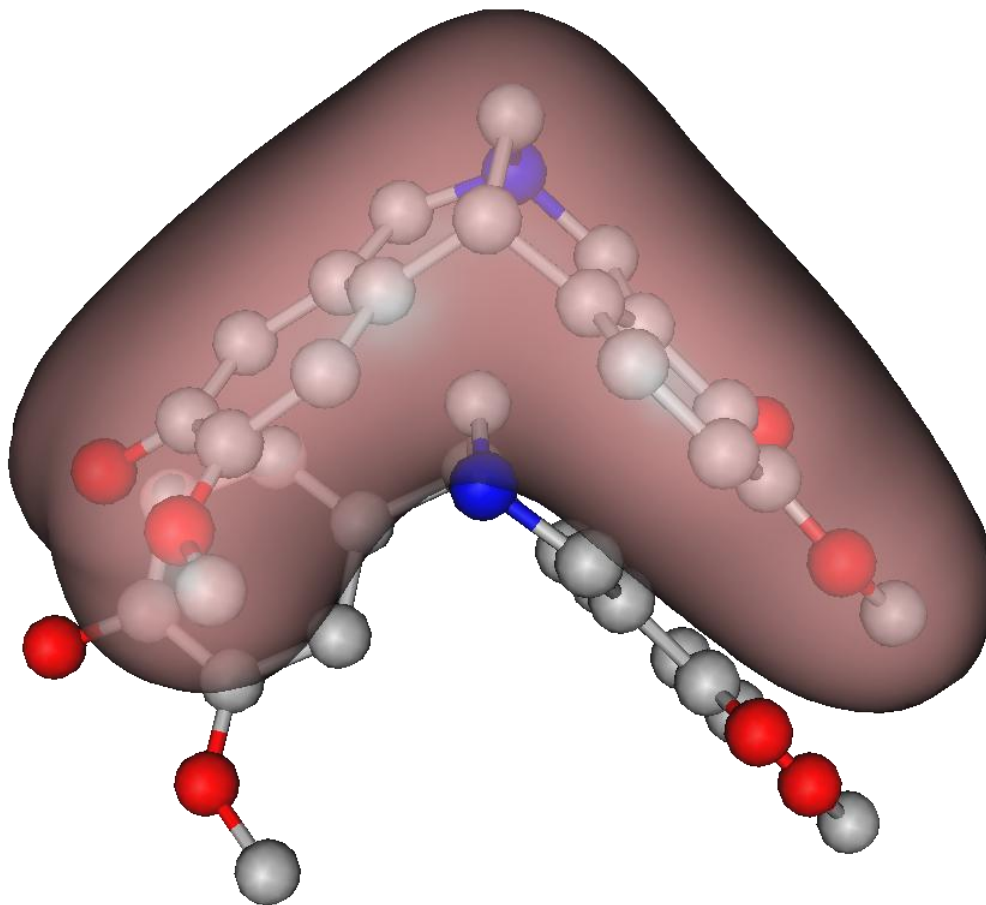
... but without the nuts and bolts



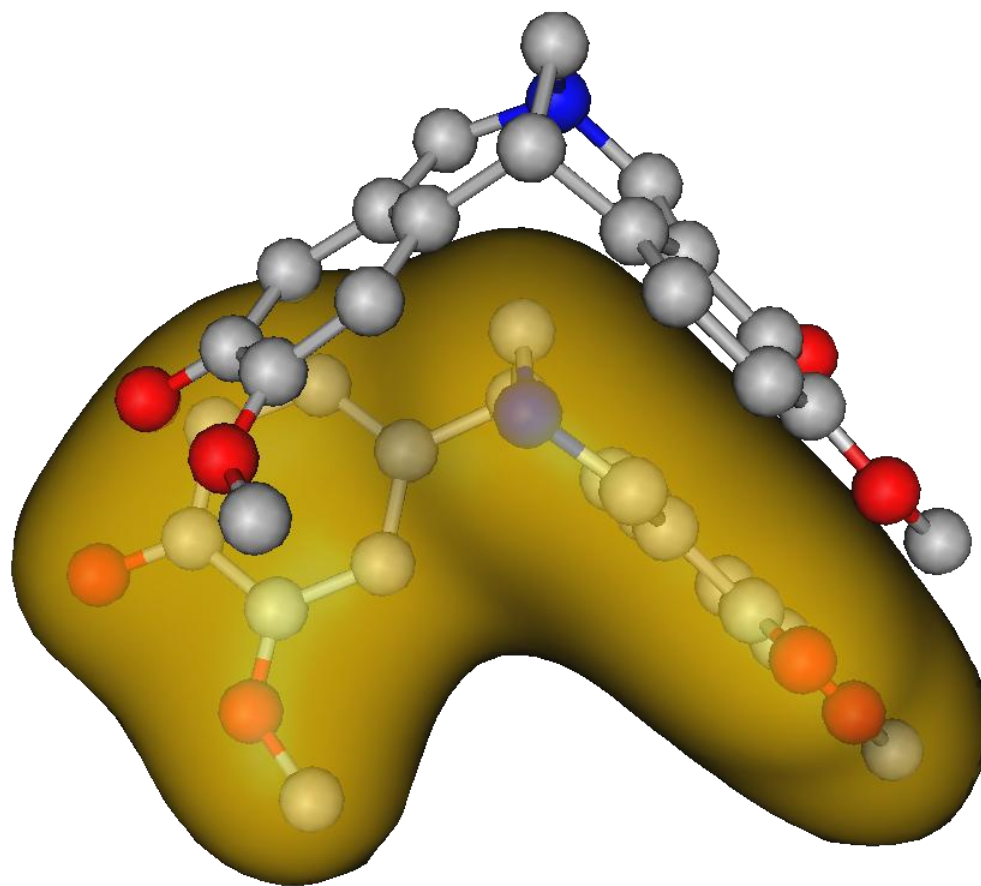
Similarity by volume alignment



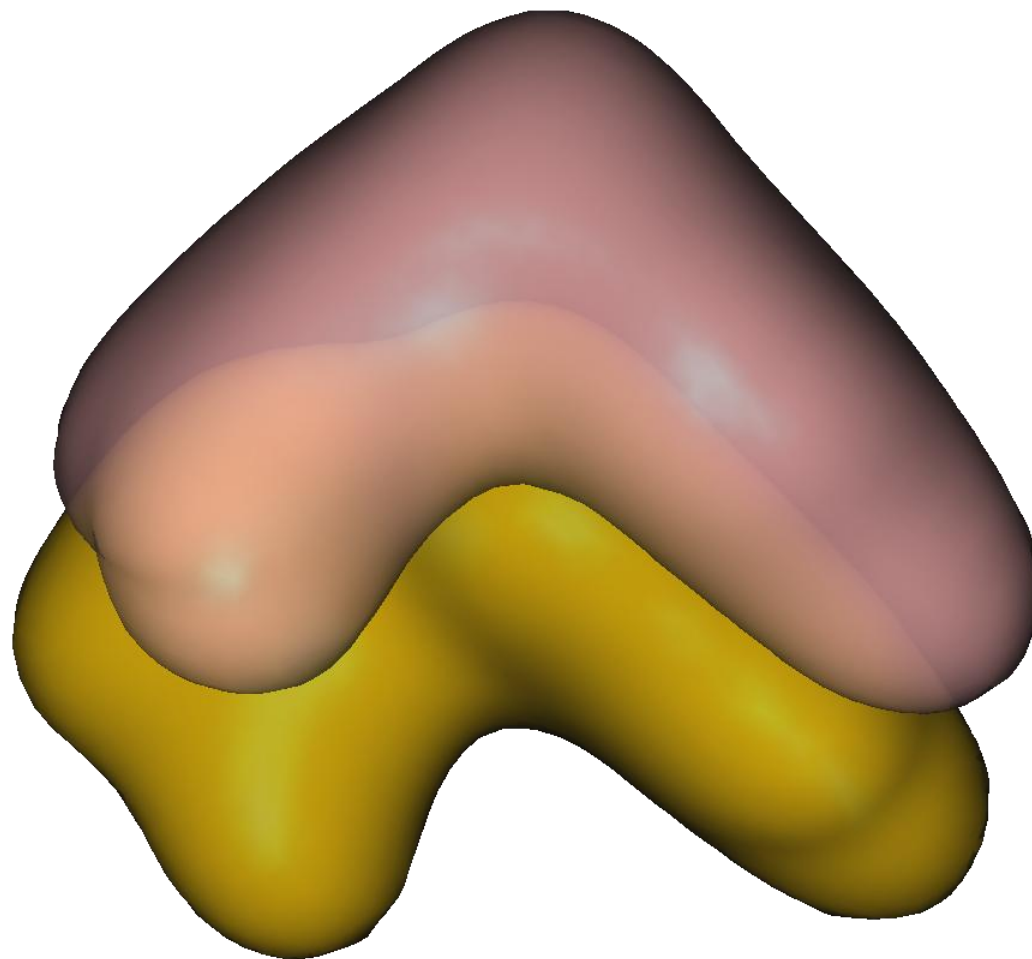
Similarity by volume alignment



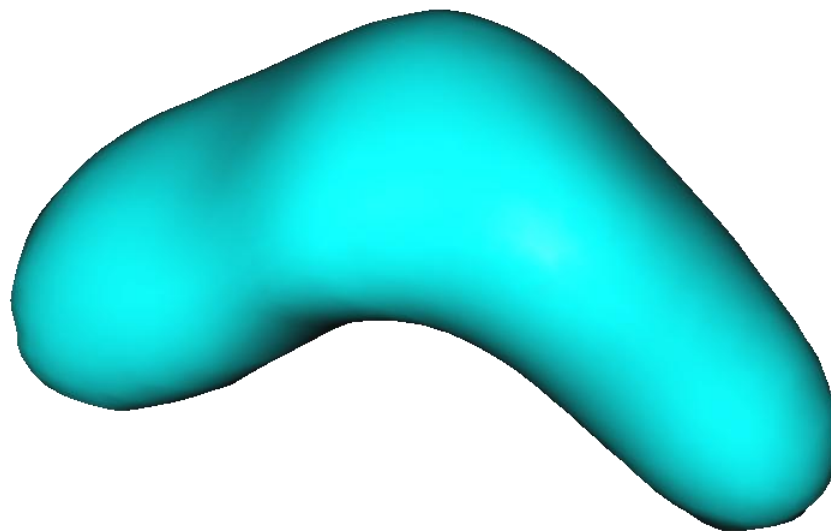
Similarity by volume alignment



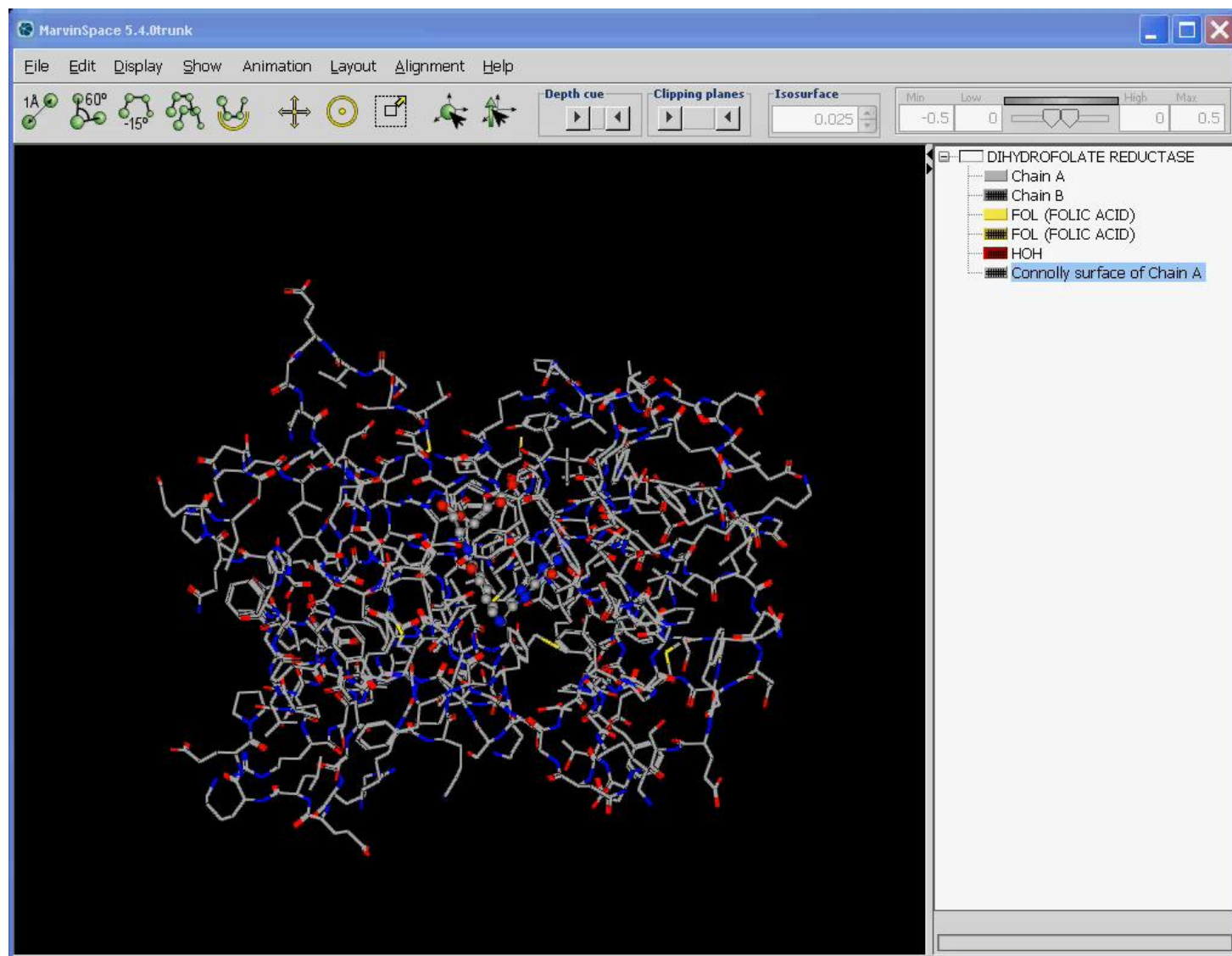
Similarity by volume alignment



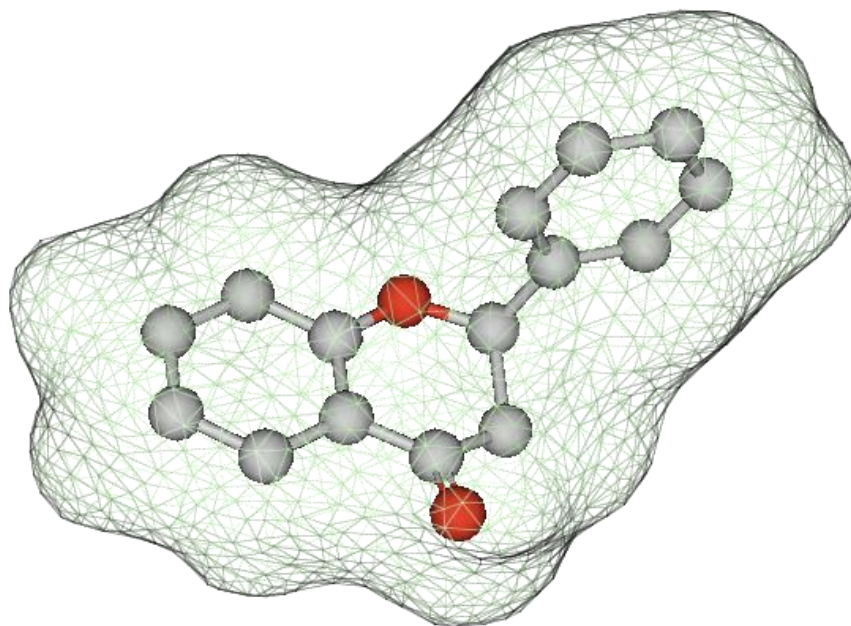
Similarity by volume alignment



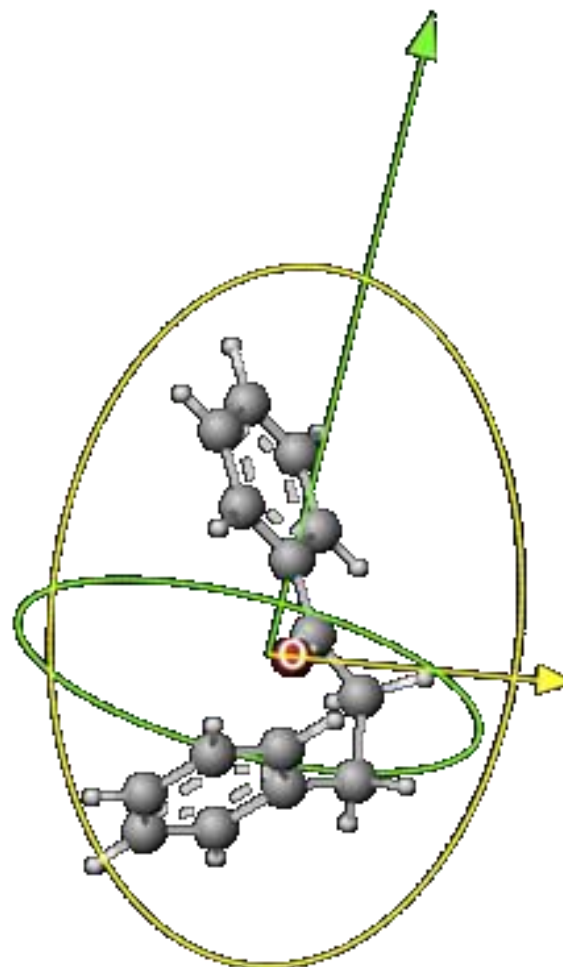
Align to ligand



Further applications of alignment



van der Waals volume = 204.55 \AA^3



Minimal projection area = 37.33 \AA^2

Length perpendicular to the max area = 7.01 \AA

Future directions

- Alignment based 3D similarity searching in 5.4
 - 3D shape descriptors
 - 3D vhts (ROCS)
- 3D pharmacophore searching
- Computing clusters, multi-core machines
- New descriptors (ECFP/FCFP, 3D shape) in 5.4

Acknowledgements

Adrián Kalászi



Gábor Imre



Ödön Farkas



Tímea Polgár



Zsuzsa Szabó



Judit Papp

Questions

Lessons we learned so far

- Customers need (and thus buy) descriptors, fingerprinting technology
- 2D would not satisfy them anymore
 - But they have 2D data

3D structure generation

