Structure-Based 3D Pharmacophores: An Alternative to Docking?

Gerhard Wolber*,
Johannes Kirchmair and Thierry Langer

*wolber@inteligand.com

inte:ligand

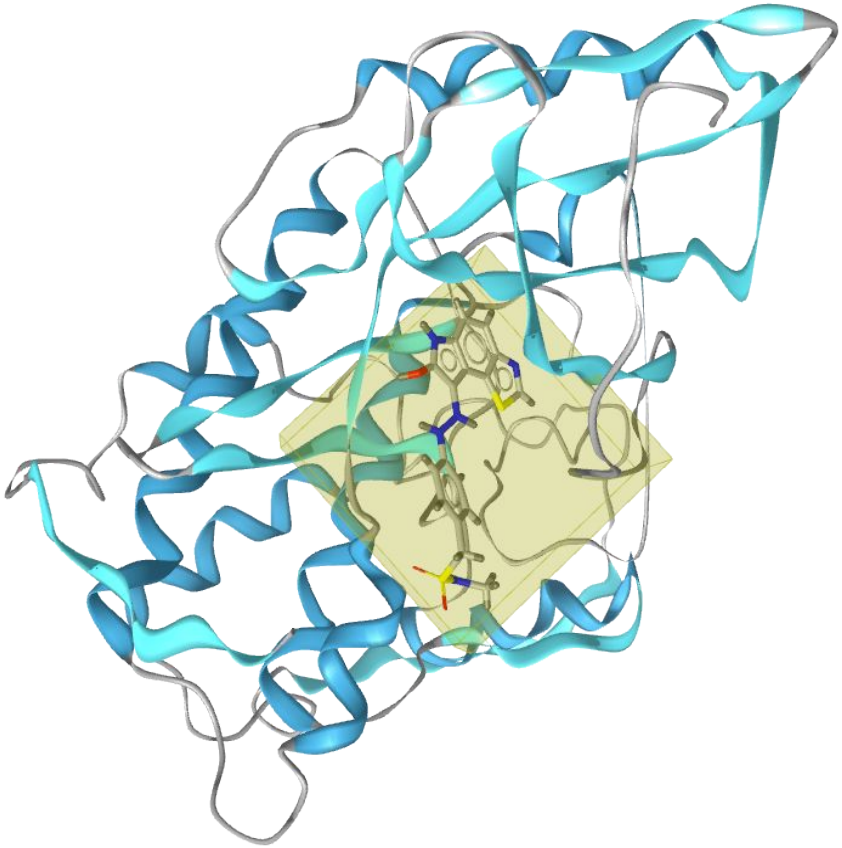
Abstract & Outline

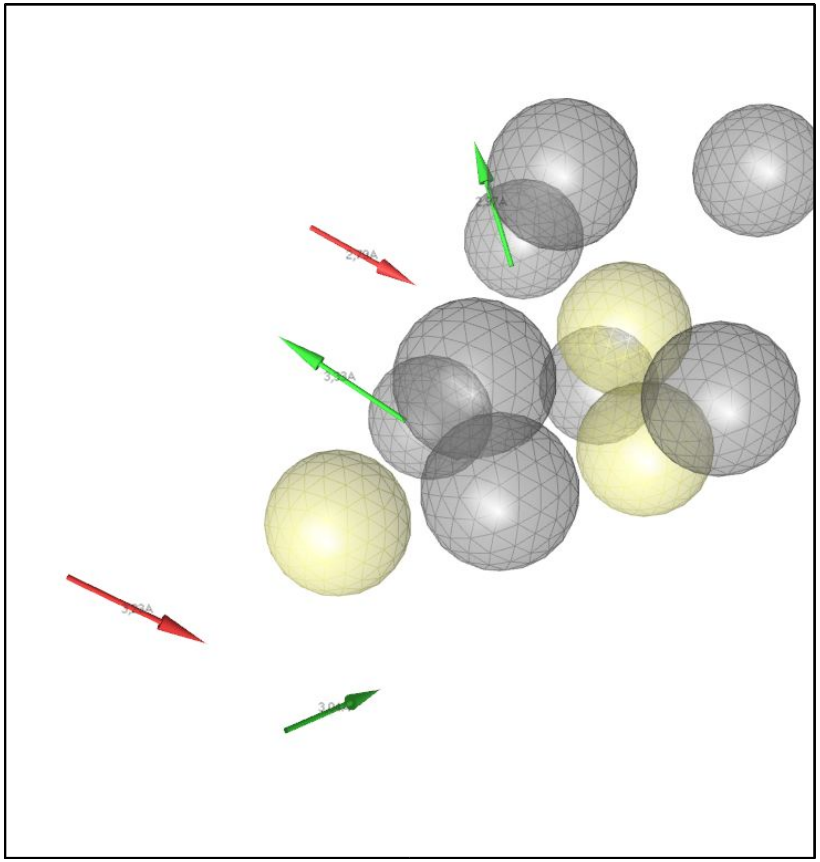
- **Pharmacophores & the Protein Data Bank**
  - 3D pharmacophore methodology
  - Primary data source: The Protein Data Bank
  - Motivation: Structure-based pharmacophore creation tool

- **LigandScout**
  - Ligand perception
  - 3D pharmacophore generation
  - Shared feature pharmacophores
  - Application example

- **Docking Comparison**
  - Compared active pose prediction
  - 58 relevant protein-ligand complexes

Structure-based pharmacophores



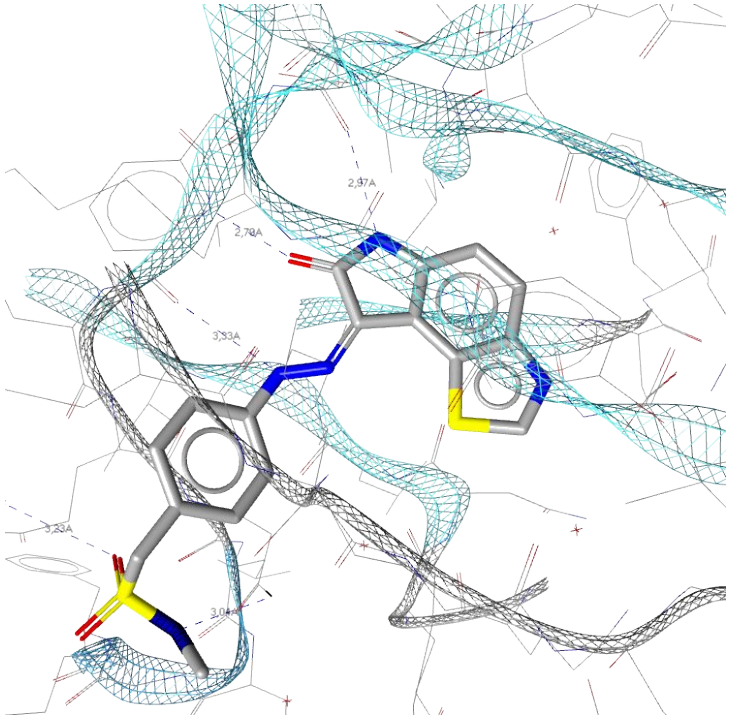


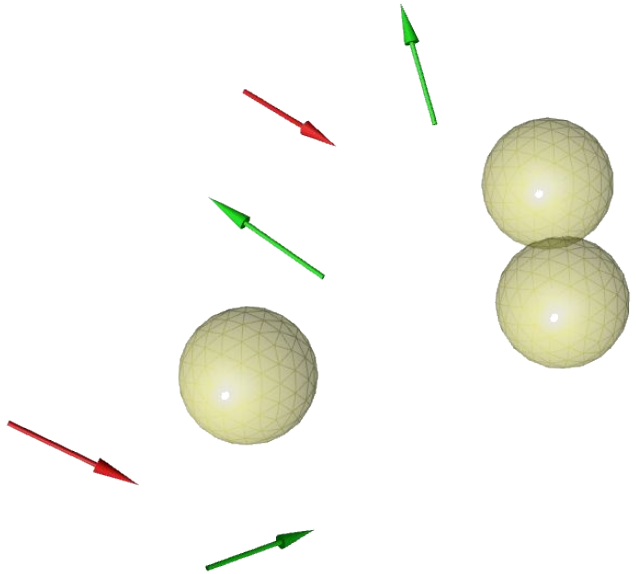
CDK2 Complexed With N-Methyl-{4- [2-(7-Oxo-6,7-Dihydro-8H-[1,3]Thiazolo[5,4-E]Indol-8-Ylidene)Hydrazino]Phenyl}Methanesulfonamide

Pharmacophore models

Pharmacophore = Ensemble of universal chemical features that represent a specific mode of action in 3D

Chemical Features: Hydrogen bonds, charge interactions, hydrophobic areas





inte:ligand

Why use structure-based pharmacophores?

- **Universal**

  Pharmacophores represent chemical functions, valid not only for the currently bound, but also unknown molecules

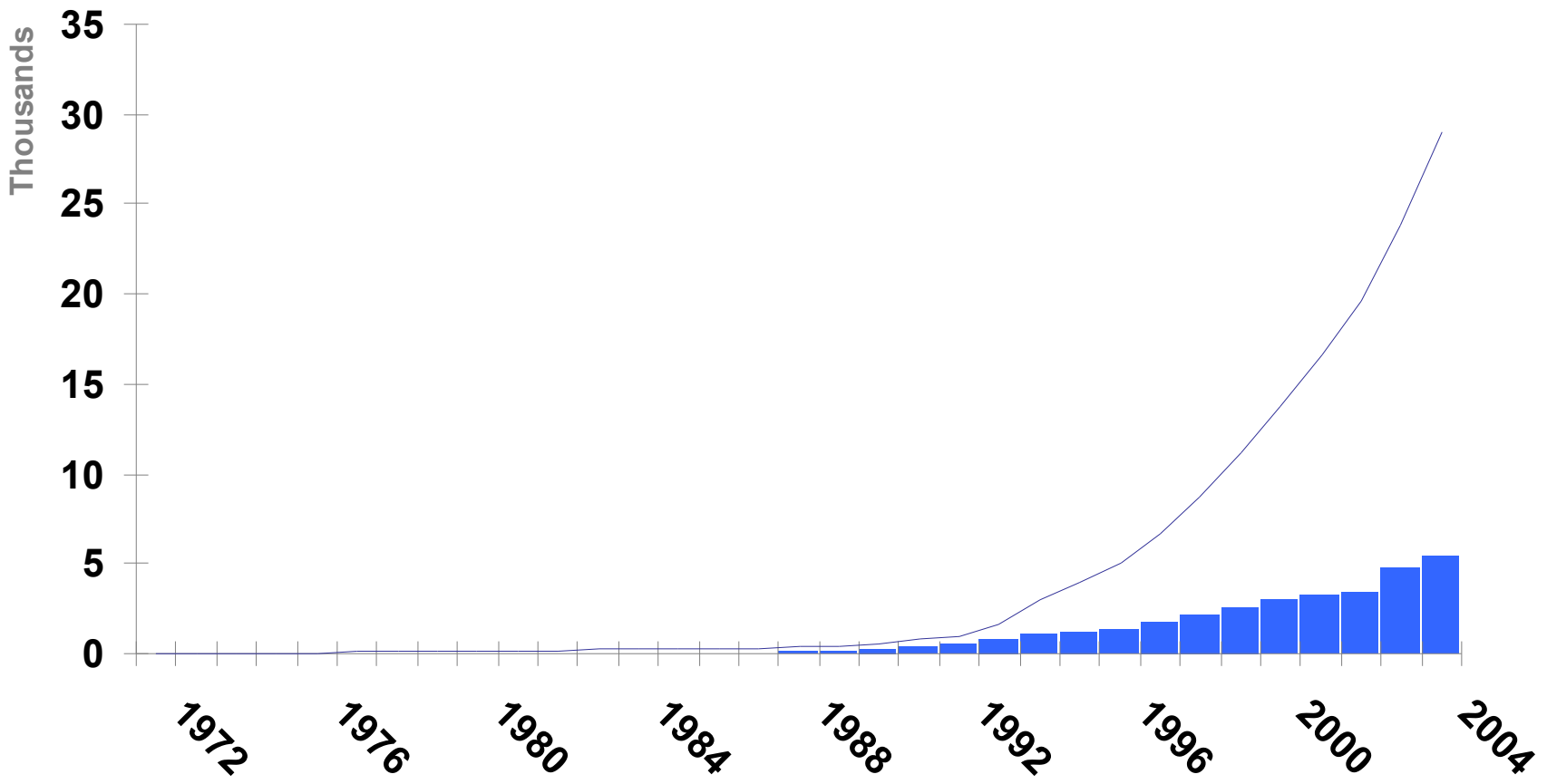
- **Computationally efficient**

  Due to simplicity (Suitable for virtual screening)

- **Comprehensive & Editable**

  Selectivity-tuning by adding or omitting feature constraints

PDB age !



inte:ligand

LigandScout: A structure-based pharmacophore creation tool

Structure-based pharmacophore creation from all PDB complexes:

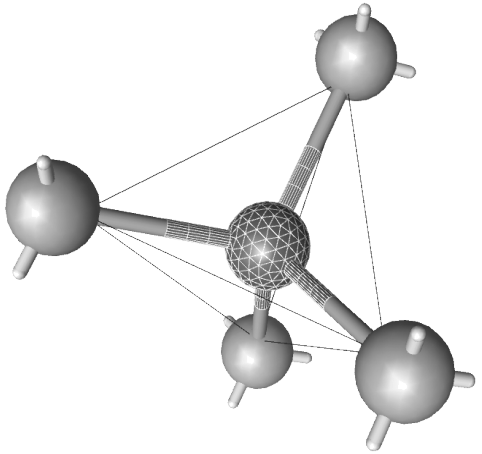
1. Extract, identify and interpret ligands (hybridization states, bonds)
2. Create pharmacophores
3. Visualize, allow user interaction and export for virtual screening

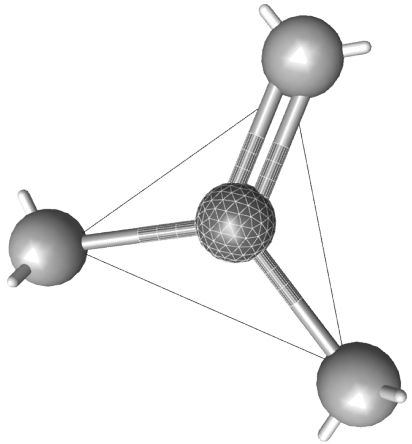
Hybridization state determination

Quantitative Geometry Templates
for all geometry types:

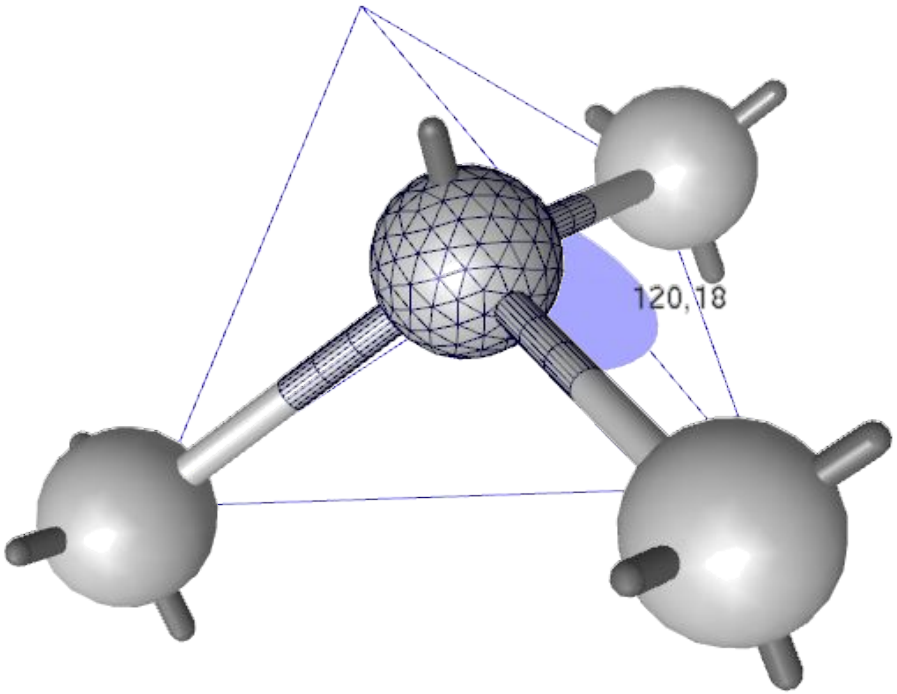
- **sp^3:** tetrahedral
- **sp^2:** trigonal planar
- **sp:** linear

Align along the first two points, numerically turn to match the third point





Geometry templates: Better than bond angles?



?

inte:ligand

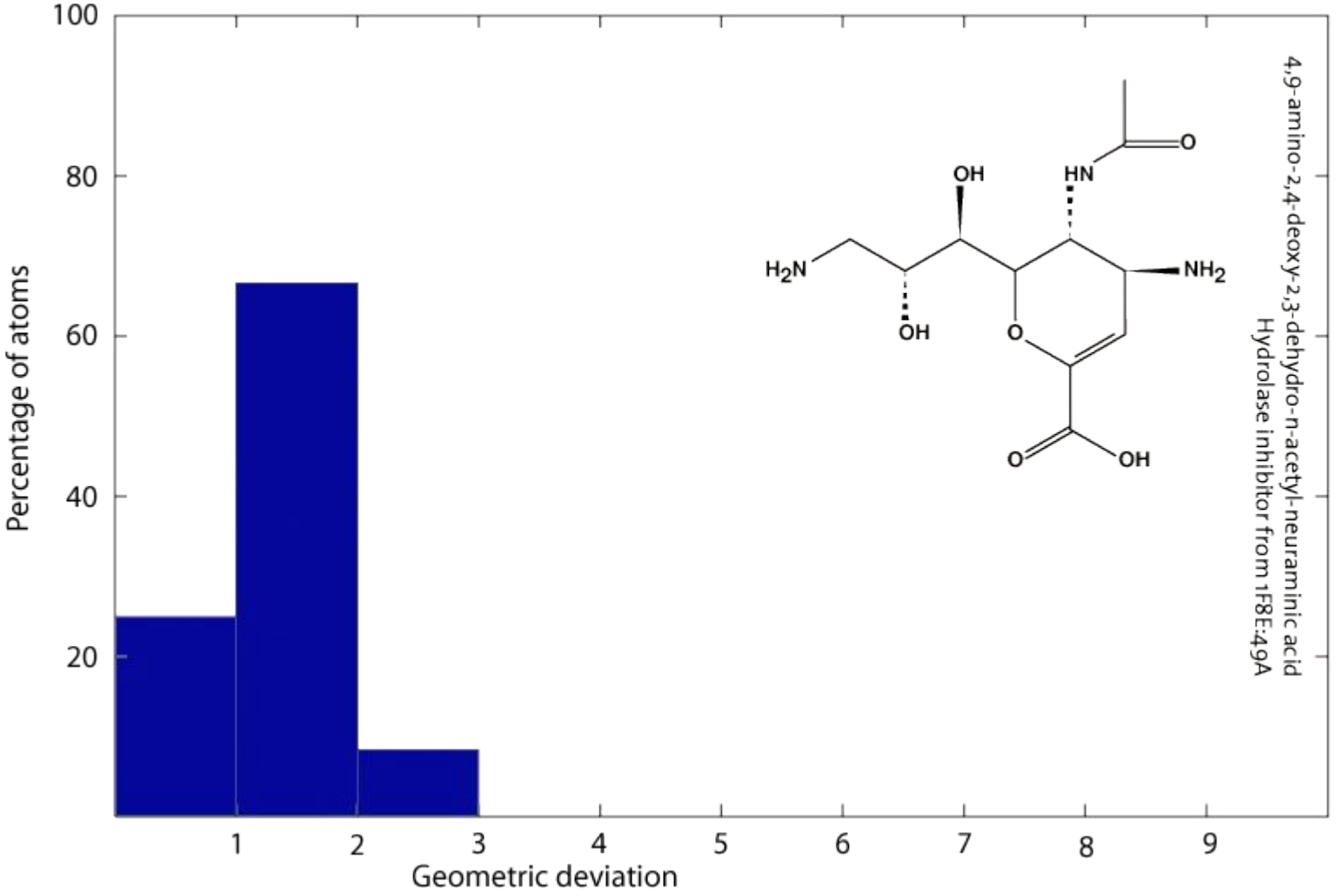
Hybridization state: Error determination

$$d_a = \sum_{i=0}^{n} \sqrt{(I_i - O_i)^2}$$

$$d_r = \frac{d_a}{n}$$

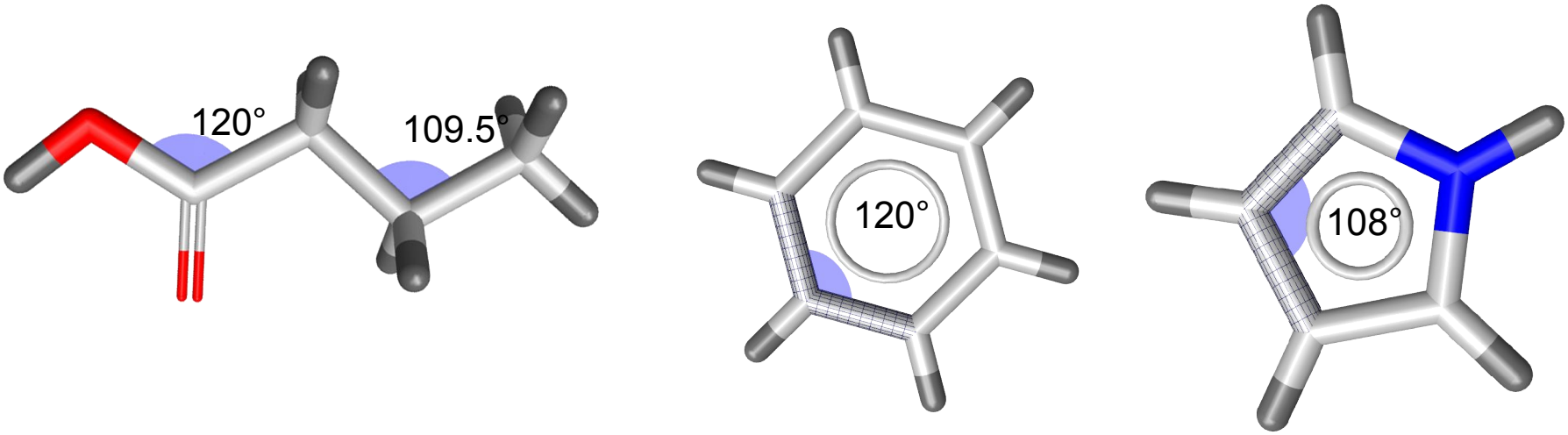
d_a, d_r	relative/absolute geometric deviation
I_i	ideal template positions
O_i	neighbor atom positions
n	number of atoms

Hybridization state: Error determination



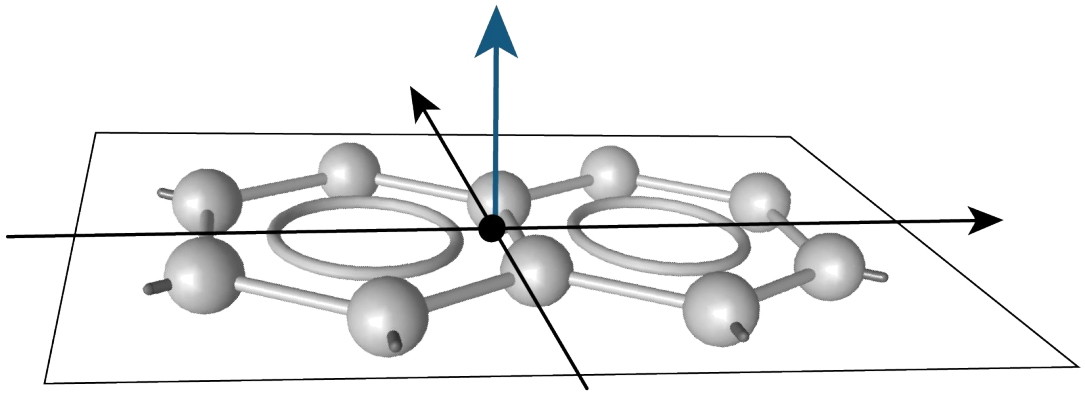
inte:ligand

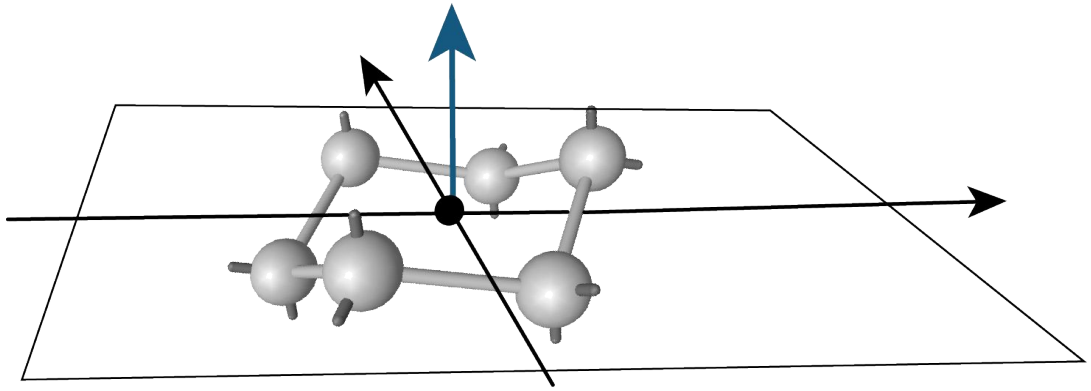
Ring geometry is different



Planar rings show different bond angles than non-ring sp^2 atoms: all planar ring atoms are to be sp^2 hybridized

Using PCA for planarity detection

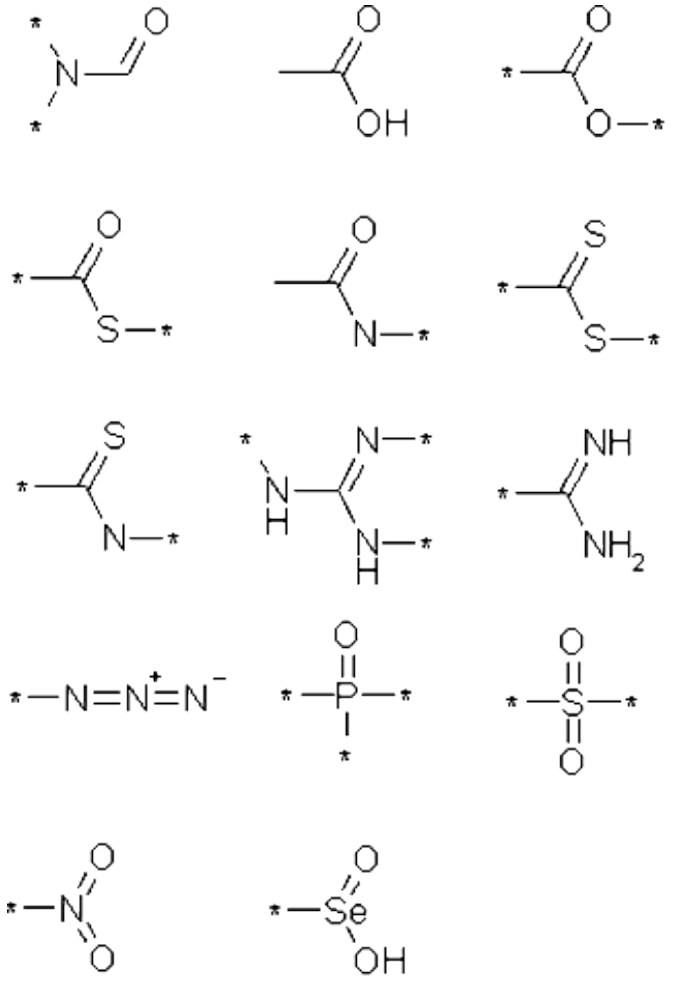




Distance from PCA plane < 0.4 A

Double bond distribution among sp^2 atoms

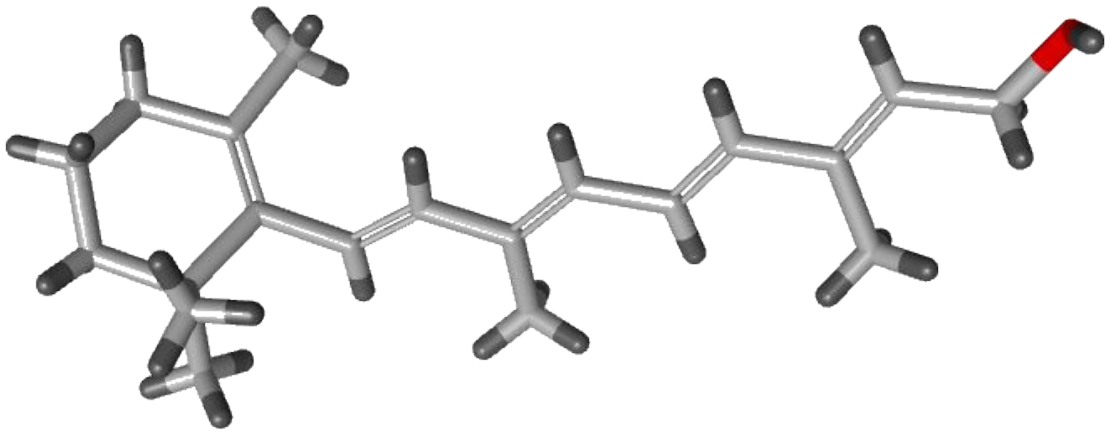
- No exact solution in many cases (e.g. Keto-enol tautomere)
- Use of patterns explicitely covering all known cases from the view of a central atom
- Weighted distribution of the maximum number of double bonds for the rest of the cases (nonbipartite maximum matching)



Patterns by Roger Sayle: Bioinformatics Group, Metaphorics LLC, Santa Fe, see http://www.daylight.com/meetings/mug01/Sayle/m4xbondage.html

Nonbipartite weighed matching

- **Double Bond Distribution along adjacent sp^2 paths**
  - Create bond classes: Identify longest and shortest bonds with non-linear geometry
  - Shortest bonds: high weights
  - Apply maximum number of double bonds using weighed nonbipartite complete matching



Pharmacophore creation

Chemical Features that are likely to occur in the complex:

- Hydrogen Bond Donors
- Hydrogen Bond Acceptors

Vectors: Direction and Distance constraint

- Negative Ionizable Areas
- Positive Ionizable Areas

- Hydrophobic Interactions

Location Spheres: Distance constraint only

Chemical features always refer to the ligand side.

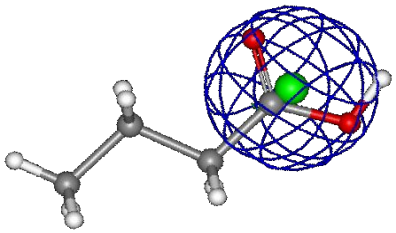
Chemical feature constraints

Distance Constraints

Relation between two points, one located on ligand side, one on macromolecular side.

Feature Type	Distance
H-Bond	2.5-3.8 A
Charge Transfer	1.5-5.6 A
Hydrophobic	1.0-5.8 A

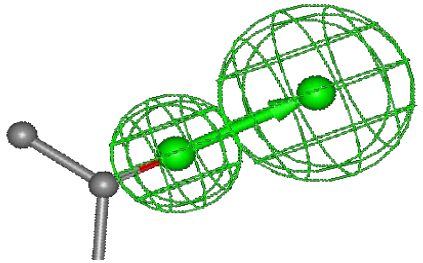
Result: one tolerance sphere on ligand side



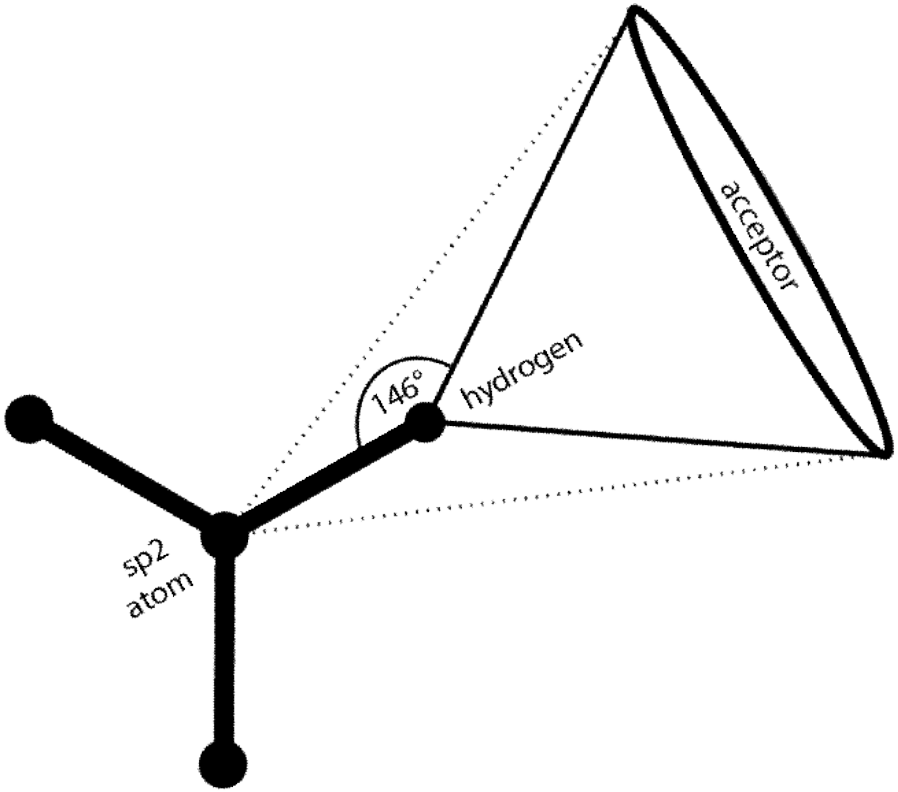
Direction Constraints

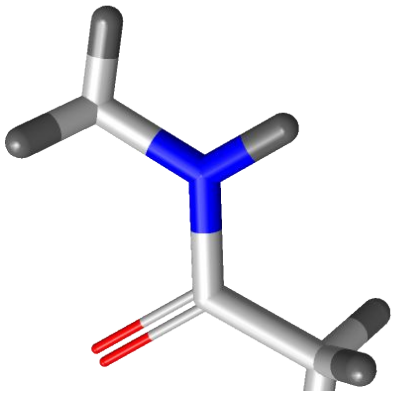
Relation between two atom groups, one located on ligand side, one on macromolecular side.

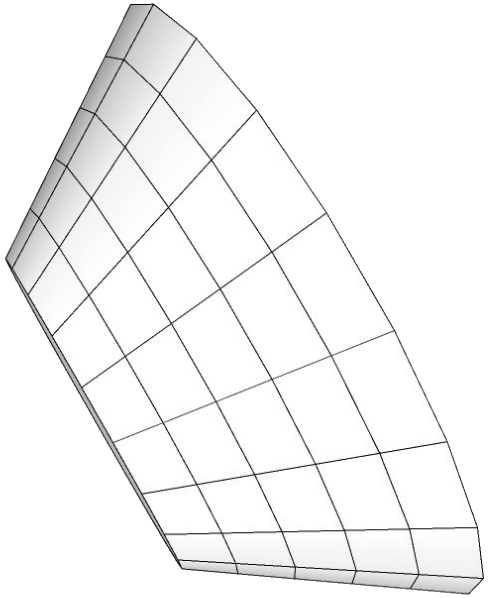
Groups form a rigid reference geometry, which are the basis for a directed vector.



Chemical feature constraints: Rigid H-bonds

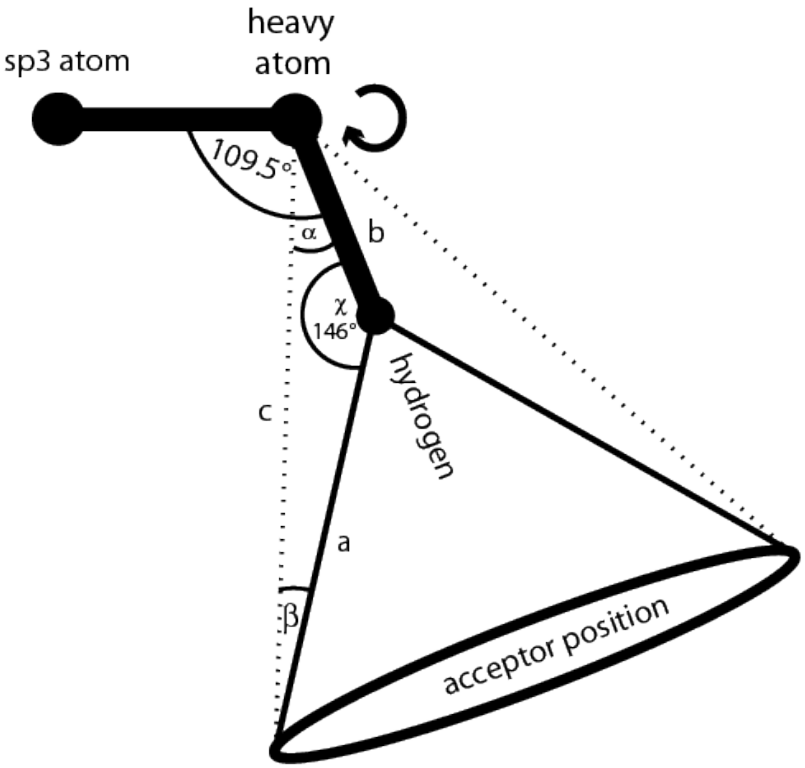


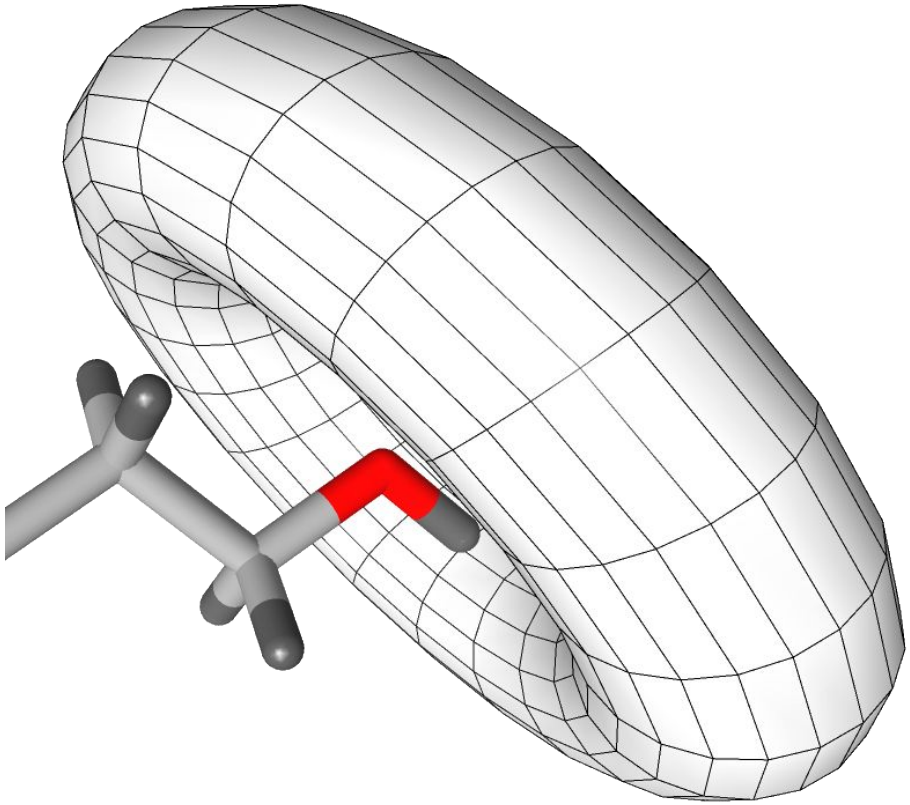




inte:ligand

Chemical feature constraints: Flexible H-bonds





$$\alpha = \delta - a\sin\left(\frac{b.\sin\chi}{c}\right)$$

inte:ligand

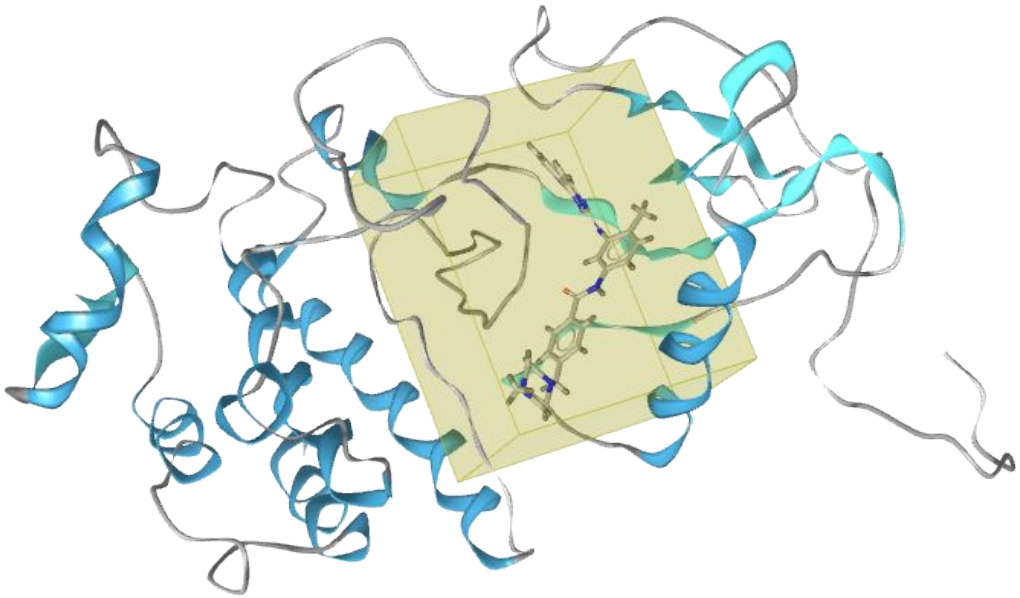
Chemical feature universality layers

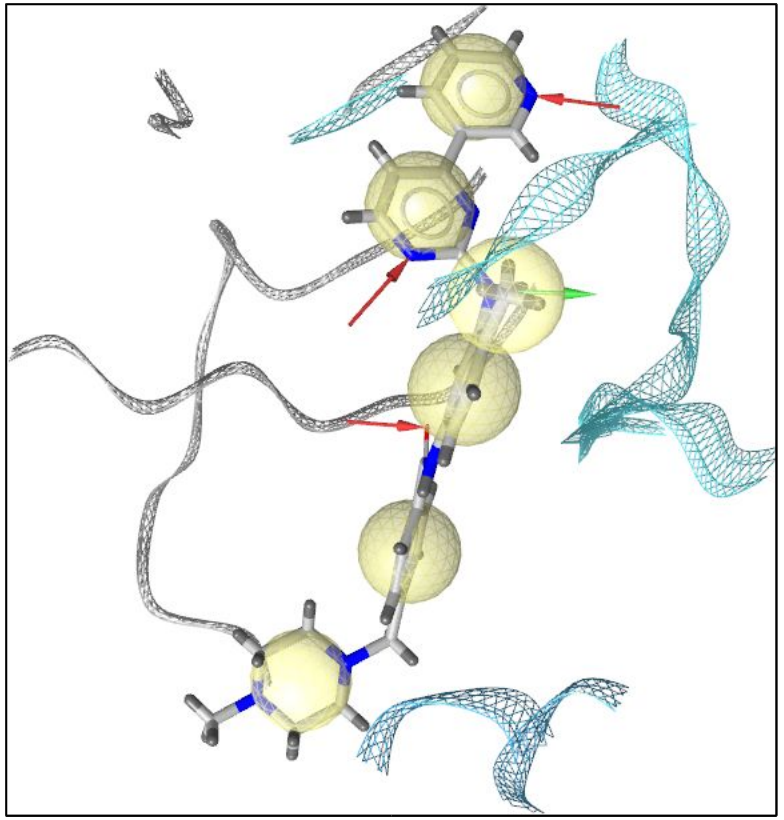
Layer 4	**Chemical Function**	Without geometry constraint	Lipophilic area, positive ionizable area
Layer 3		Including geometry constraint	Hydrogen bond Donor/Acceptor
Layer 2	**Subgraph**	Without geometry constraint	Hydroxylic group, Phenol Group
Layer 1		Including geometry constraint	Phenol group facing a parallel benzene

⇨ LigandScout creates pharmacophores using the universal Layer 3 and Layer 4 features

inte:ligand

Application example: Gleevec



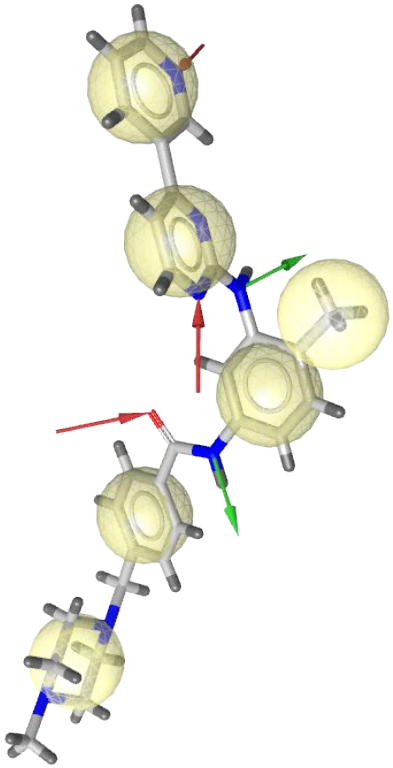


Gleevec in PDB complex
1IEP, 1OPJ; variant 1FPU

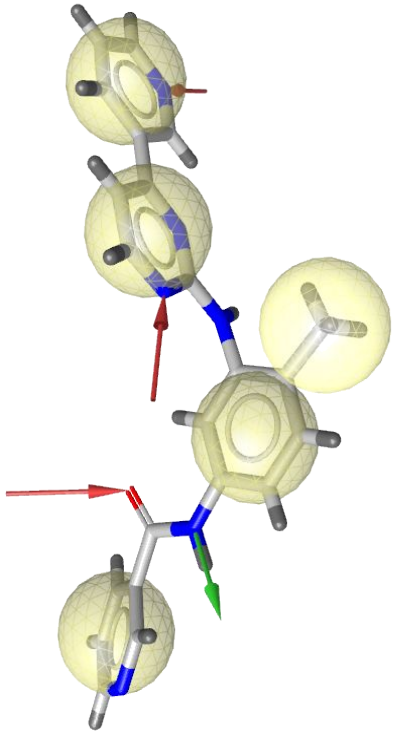
inte:ligand

Shared feature pharmacophore

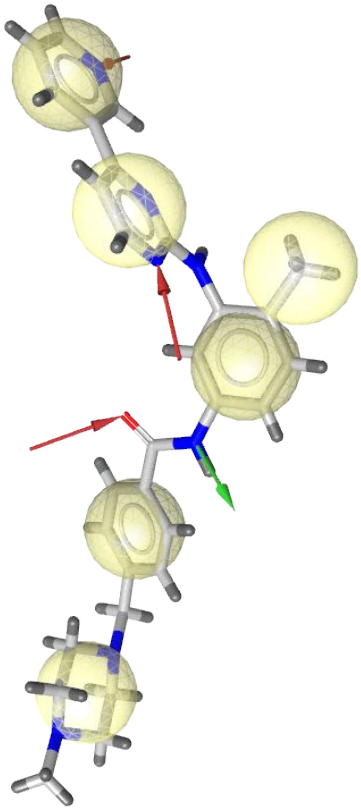
1iep



1fpu



1opj



inte:ligand

Pharmacophore overlaying

Pharmacophore model derived from one single bound ligand may not be able to retrieve other related compounds …

Starting set: Several ligand-protein complex pharmacophores

Creation of compatibility graphs

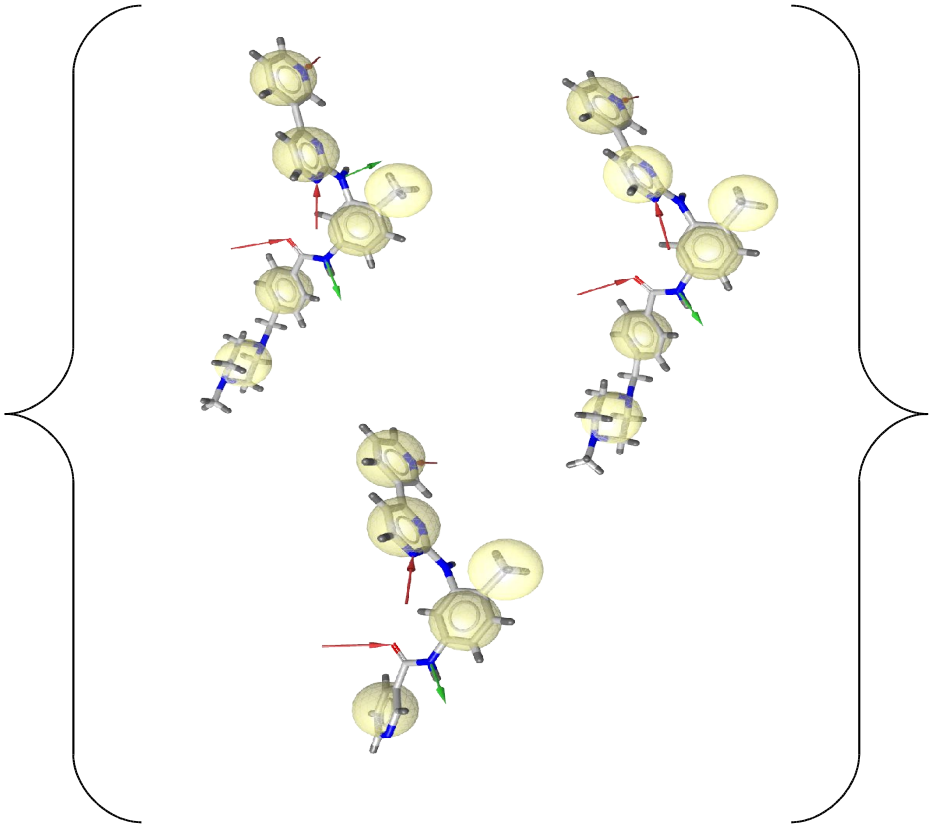
Maximum clique detection

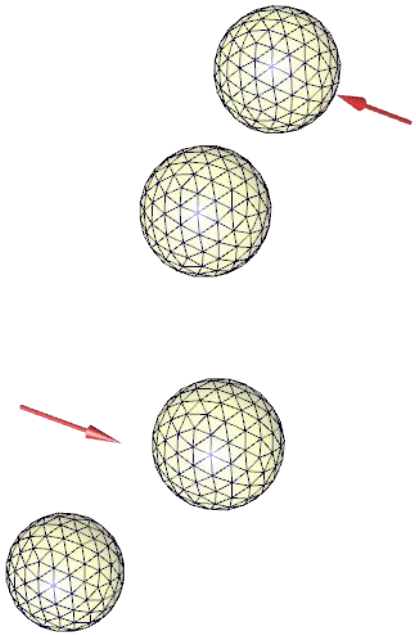
Feature alignment

Calculation of combined features

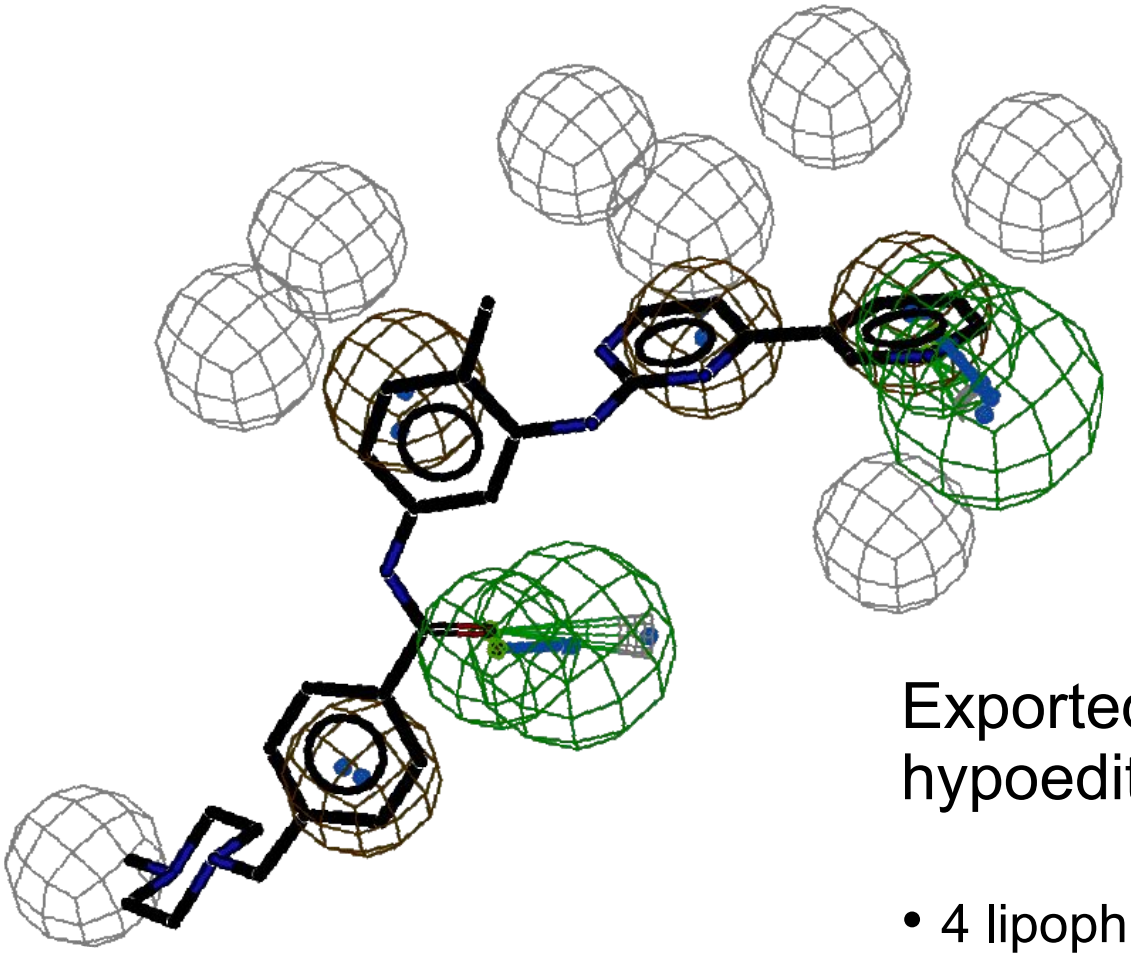
… new shared feature pharmacophore

Shared feature pharmacophore





Shared feature pharmacophore



Exported to **Catalyst** using hypoedit tool:

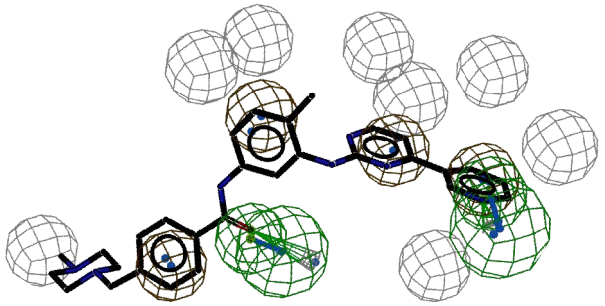
- 4 lipophilic aromatic areas
- 2 hydrogen bonding interactions

Virtual screening setup

- **Virtual screening using Catalyst**
- 3 Databases:
  - **PDB singleConf:** all PDB ligands with one single entry per conformation [67k]
  - **PDB multiConf:** all PDB ligands with one single entry per unique molecule and 50 conformers each (multiConf; 50 FAST) [7k]
  - **Maybridge 2003** (multiConf) [55k]

Virtual screening results

Gleevec
shared feature pharmacophore



Database	Hits	Drug-like hits
PDB singleConf (~67k)	7	7
PDB multiConf (~7k)	2	2
Maybridge (~55k)	19	7

LigandScout summary

LigandScout

- Extracts and interprets ligands and their protein environment from PDB files
- Automatically creates and visualizes 3D pharmacophore models
- Creates overlaid “shared feature” pharmacophores to broaden the scope of a single model

Docking comparison

Is it possible to predict the active pose of a ligand using a 3D pharmacophore?

Is fitting ligands to structure-based 3D pharmacophores as accurate as docking?

Method comparison: Discussion

Pharmacophores

- Pharmacophore biased to specific binding mode (multi-feature binders less)
- Editable
- Fully automated
- **Suitable for virtual screening** (60,000 compounds in minutes)
- Conformer generation might become a limit

Docking

- Not biased to bound ligand
- Generic – might detect different binding locations and modes
- Black Box
- Pre-processing necessary
- **Suitability for VS questionable** (30 sec to minutes per compound)

Docking and Pharmacophore Fitting

1. Docked bio-active ligands into 58 pharmacologically relevant complexes [1] using **FlexX** and **Gold**

3. Generated unbiased conformers and fitted into **LigandScout** hypotheses using Catalyst (maxConfs=50, FAST) [2]

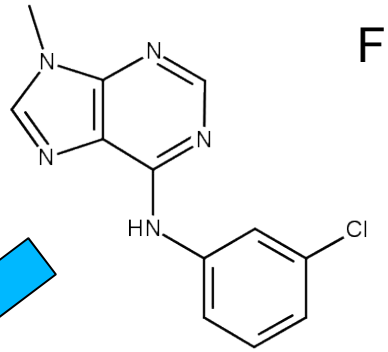
5. Compared best fitting conformation to best scored docking pose (CScore, GoldScore)

[1] M. Kontoyianni, L.M. McClellan, G.S. Sokol. Evaluation of Docking Performance: Comparative Data on Docking Algorithms. J. Med. Chem.; **2004**; 47(3); 558-565.

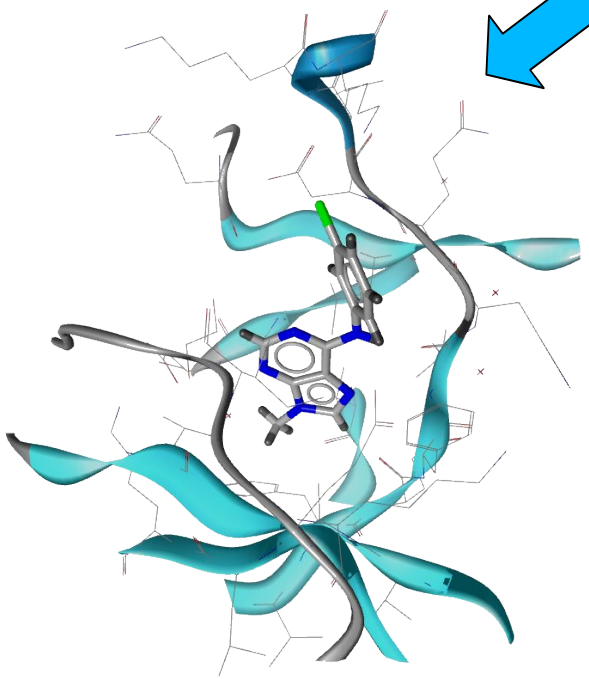
[2] J. Kirchmair, C. Laggner, G. Wolber, T. Langer. Comparative Analysis of Protein-Bound Ligand Conformations with Respect to Catalyst's Conformational Space Subsampling Algorithms. J. Chem. Inf. Model.; **2005;** 45(2) pp 422 - 430;

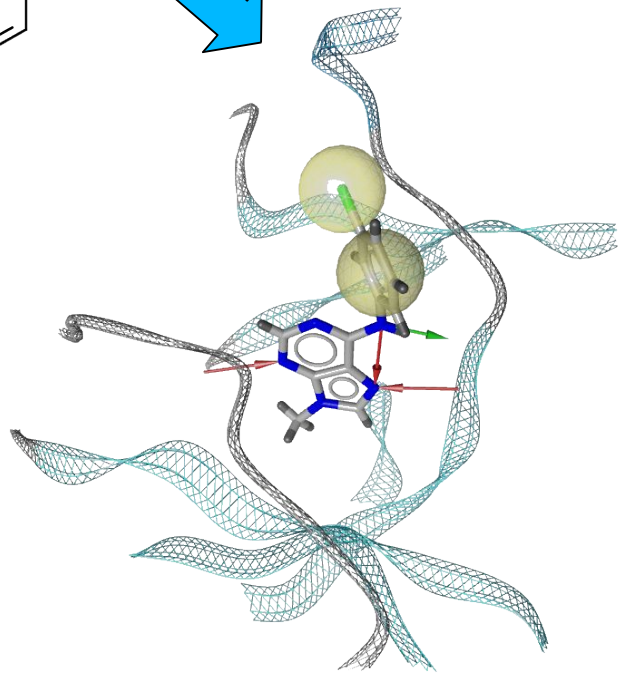
Docking and fitting

Docking into active site with FlexX and GOLD

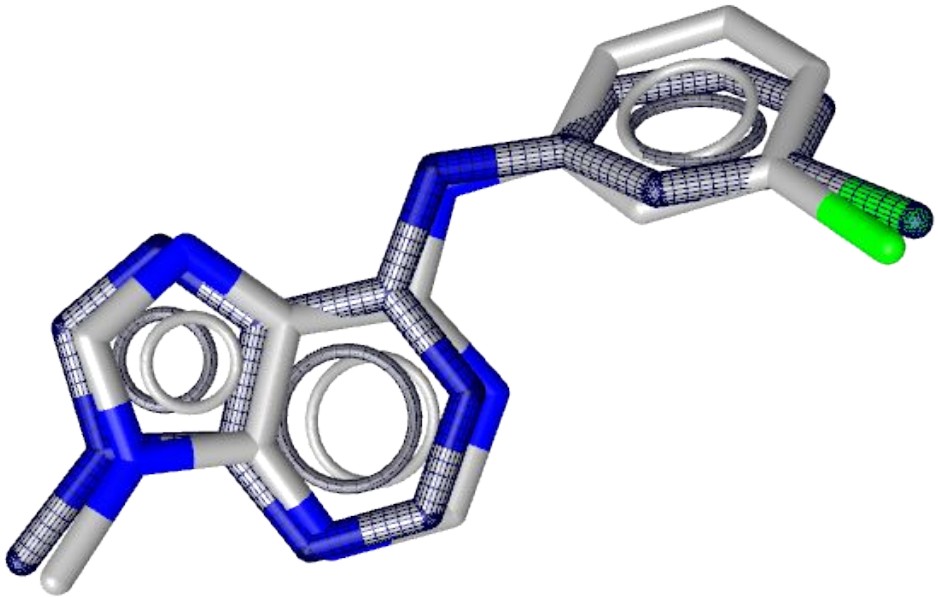


Fitting to pharmacophore in original coordinate frame



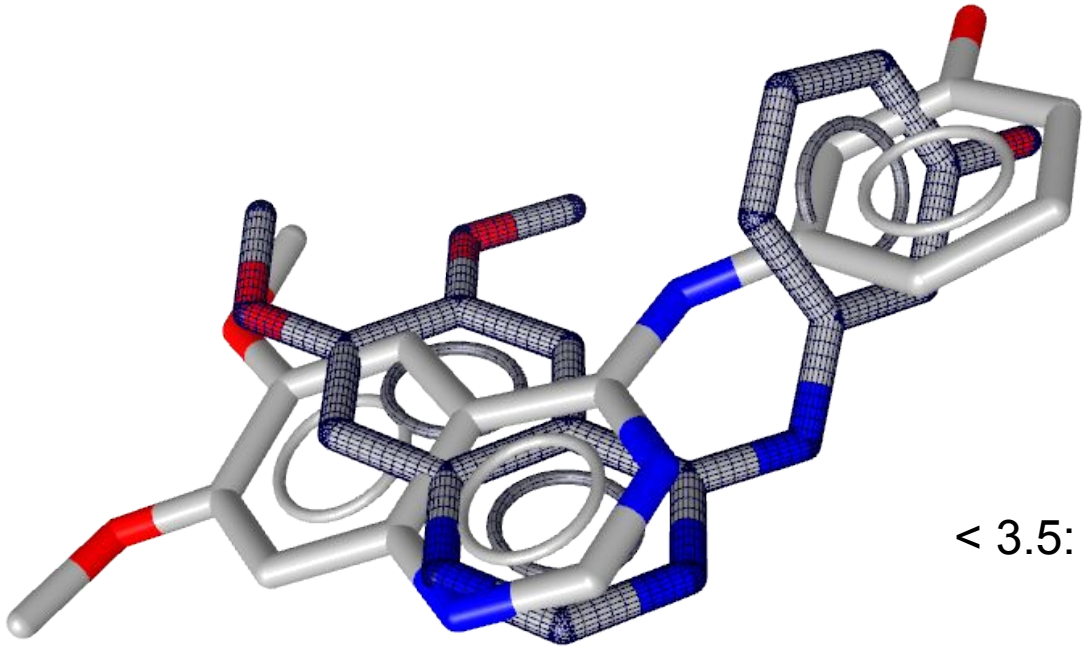


inte:ligand



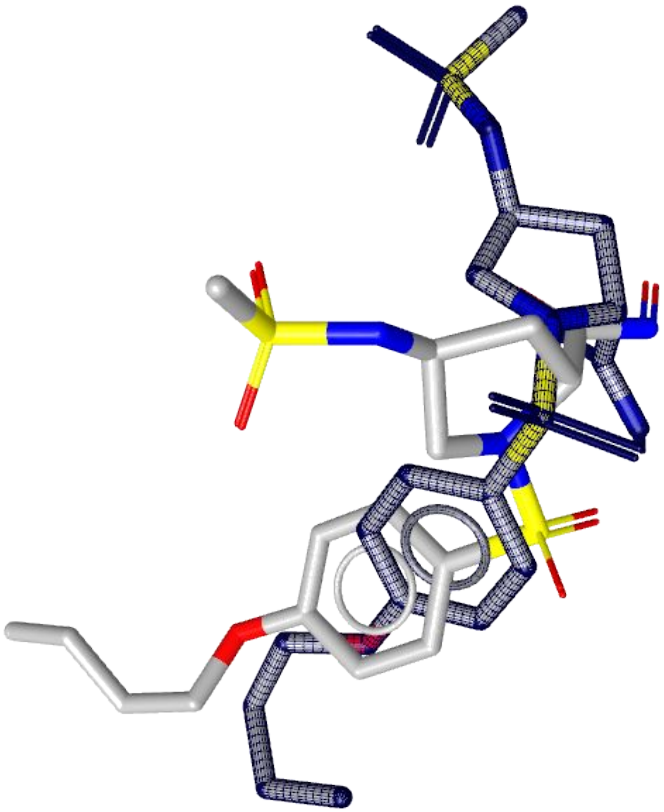
1CKP

RMS = 0.63
< 1.5: **„perfect fit“**



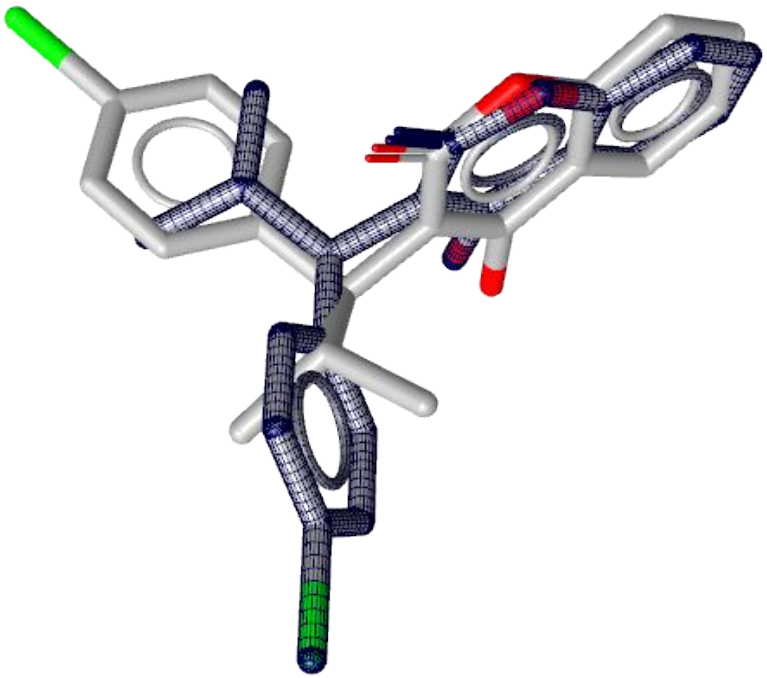
1DI8
RMS = 2.18
< 3.5: **„acceptable fit“**

inte:ligand



1G49
RMS = 3.69

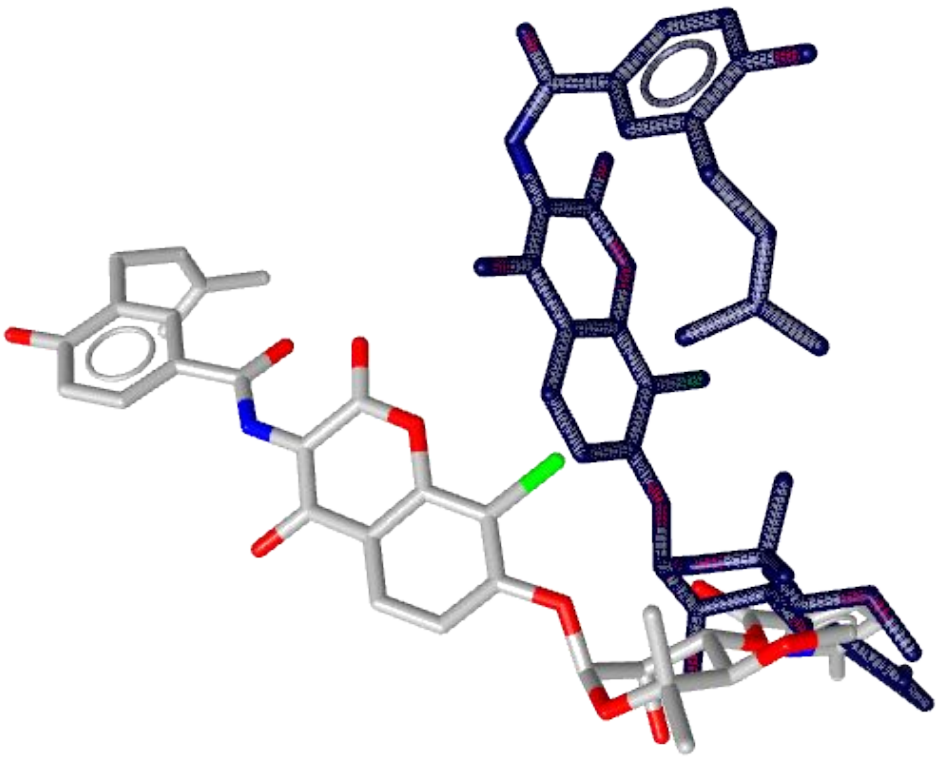
3.5 < RMS < 6:
„inadequate fit“



1UPJ
RMS = 3.78

LigandScout case
„inadequate fit“

But bioisosteric!



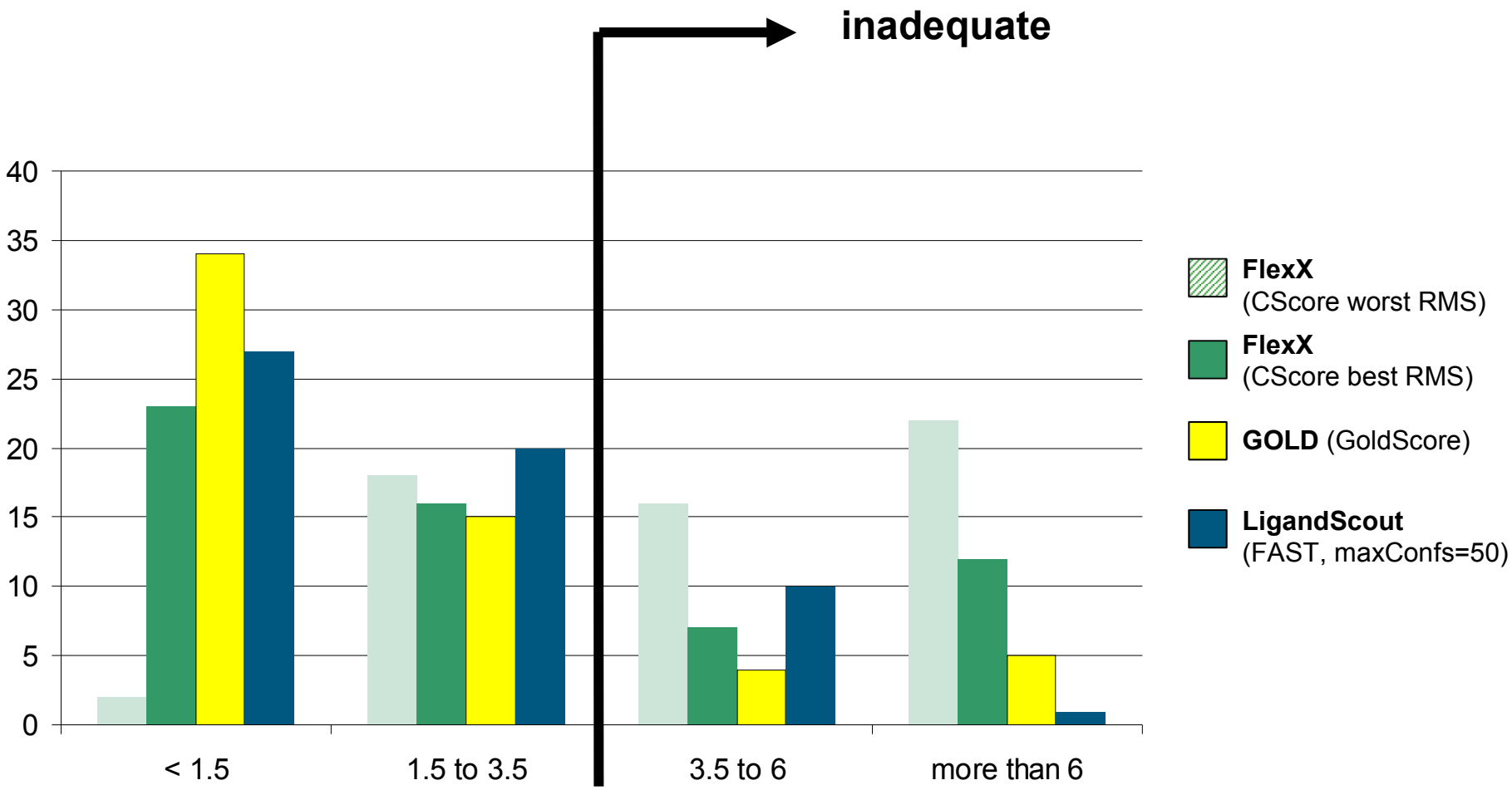
1KZN

RMS = 8.76
„inacceptable fit“

> 6

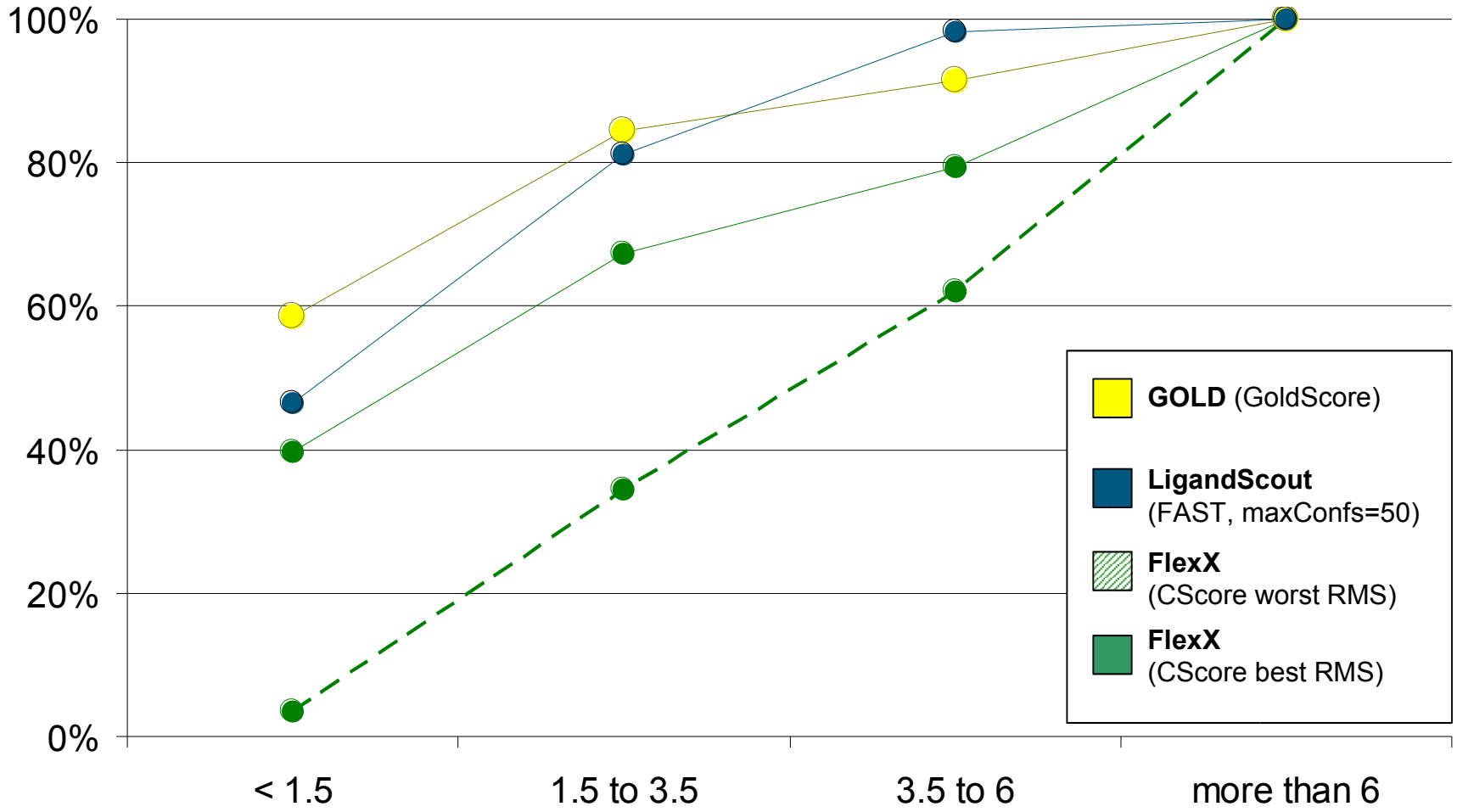
inte:ligand

Results



inte:ligand

Cumulative percentage



inte:ligand

Results summary

- More than 80% of the LigandScout complex fits are below an RMS of 3.5!
- „Binding site bias“ can be seen as an advantage
- Better conformer generation might further improve results

Conclusions

- 3D pharmacophores perform considerably well in predicting poses
- Accuracy is comparable to docking (with fewer complete failures)
- Virtual screening using 3D pharmacophores is much faster (pre-sampled multi-conformer databases)

>> Structure-based 3D pharmacophores are a viable alternative to docking!

Literature

- G. Wolber and T. Langer. LigandScout: **3-D Pharmacophores Derived from Protein-Bound Ligands and Their Use as Virtual Screening Filters** *J . Chem. Inf. Model.*; 2005; 45; 160-169

- T. Steindl and T. Langer. **Docking Versus Pharmacophore Model Generation: A Comparison of High-Throughput Virtual Screening Strategies for the Search of Human Rhinovirus Coat Protein Inhibitors.** *QSAR and Combinatorial Science*; 2005; in press

Acknowledgements

inte:ligand

- Alois Dornhofer
- Robert Kosara
- Martin Biely
- Fabian Bendix
- Eva Maria Krovat
- Theodora Steindl

universität innsbruck

- Thierry Langer
- Johannes Kirchmair
- Christian Laggner
- Daniela Schuster
- Oliver Funk