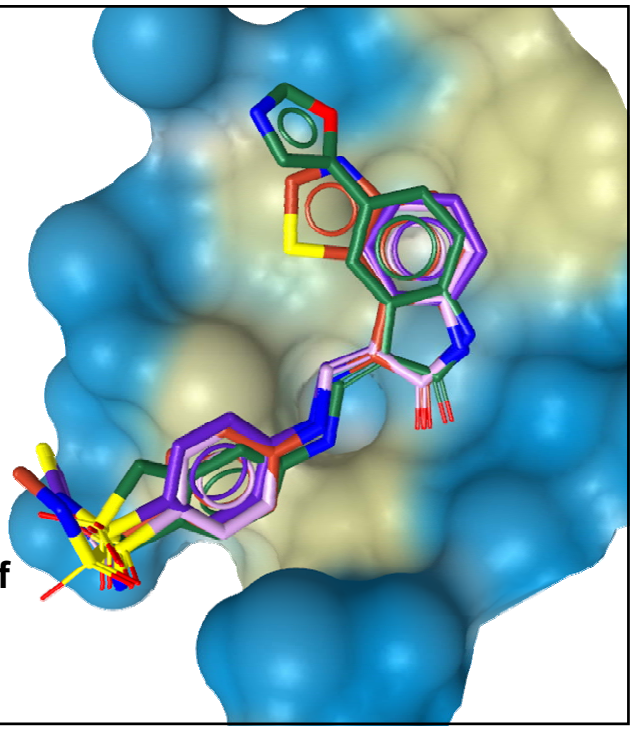
# Efficient overlay of molecular 3-D pharmacophores

**Gerhard Wolber***,
Alois A. Dornhofer & Thierry Langer

* E-Mail: wolber@inteligand.com



inte:ligand



**Superposition of molecules …**

# Alignment: Outline

**Scope, design goals & existing methods**

- o Scope = virtual screening & pharmacophore model building

**Our algorithm, validation & examples**

- o Matching and analytic alignment
- o Validation

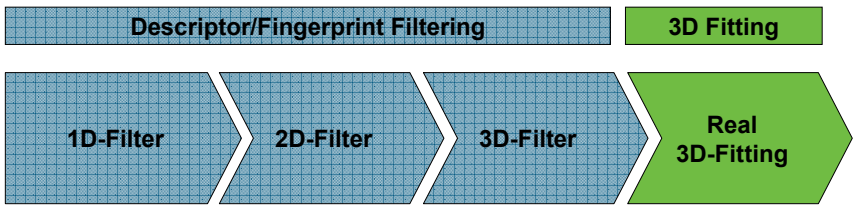
**Current applications & outlook**

- o Shared and merged pharmacophores



inte:ligand

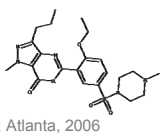
# Virtual Screening Methods



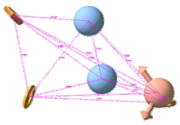

**1D Filtering**
- o Property Ranges
- o Fingerprints

**e.g. MW 200-500 Lipinsky**

**2D Filtering**
- o Topology, Molecular Graph
- o (Red. Graphs, FTrees, …)



**3D Filtering**
- o 3-point pharmacophores
- o Distance hashing



**3D Fitting**
- o Flexible
- o From pre-computed conformers

**Computationally expensive**



inte:ligand

# Design Goals

**1. Fast**
- Scalable and usable also for „large“ small molecules or pharmacophores
- Interactively usable
- Applicable for virtual screening

**2. Rigid Method**
- Save time by pre-computing conformers

**3. Correct geometry & correct chemistry**
- Produce real geometric alignments, not only hashing
- Represent molecule in a pharmacophoric way

**4. Pharmacophore applicability**
- Be able to align molecules to 3D pharmacophores and vice versa
- Align different molecules/pharmacophores to each other



# Existing methods, different scopes

**Distance-based combinatorial approach (brute-force)**
- Computationally very expensive

**Distance-based clique-detection:**
- Feasible for small molecules, but still growing exponentially with the number of features (np-complete)

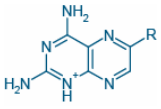
**Fingerprinting:**
- No real 3D alignment, only a pre-filter step

**Reduced Graphs:**
- Bound to the chemical graph, problems with aggregation/fragmentation

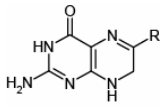
**Flexible Pharmacophore Elucidation:**
- Expensive for virtual screening

**[Lemmen C, Lengauer T. 2000]** Computational methods for the structural alignment of molecules. J Comput Aided Mol Des. 2000 Mar;14(3):215-32 **(27 methods!)**

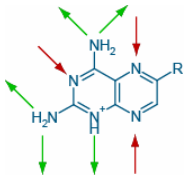
# Alignment by pharmacophore points

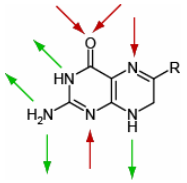


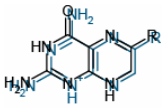
Methotrexate



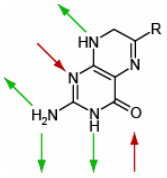
Dihydrofolate

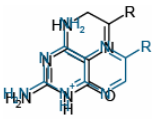






Wrong





Correct

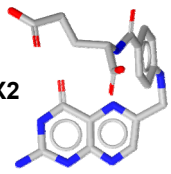
Böhm, Klebe, Kubinyi: Wirkstoffdesign (1999) p. 320f

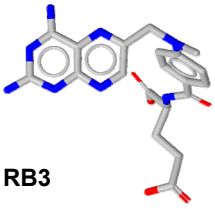


inte:ligand

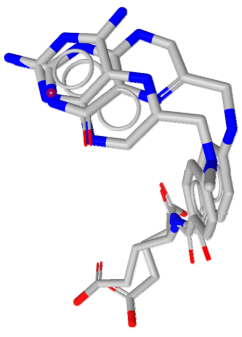
# Alignment by pharmacophore points

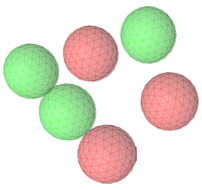
1RX2





1RB3



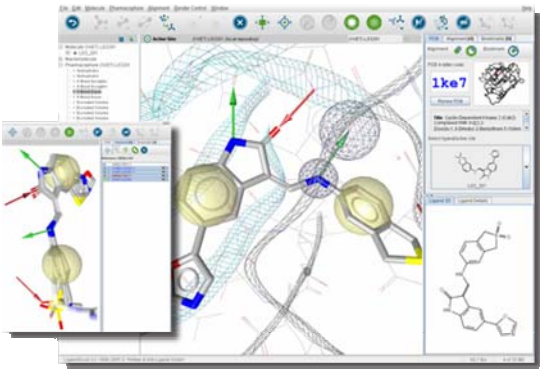




inte:ligand

# Pharmacophore Representation

## LigandScout

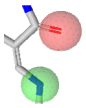


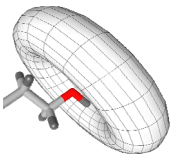

**[LigandScout 2005]** Wolber, G.; Langer, T. 3D Pharmacophores Derived from Protein-Bound Ligands and Their Use as Virtual Screening Filters J. Chem. Inf. Comput. Sci.; (Article); 2005; 45(1); 160-169. DOI: 10.1021/ci049885e

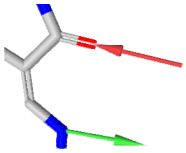


# 3D Pharmacophore: Chemical Features

### Hydrogen Bond Donors/Acceptors

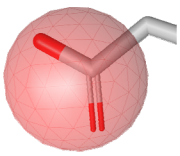


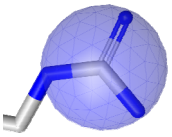




**Vectors:** Direction and Distance constraint

### Negative/Positive Ionizable Spheres



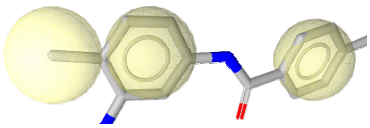


**Location Spheres:** Distance constraint only

# 3D Pharmacophore: Chemical Features

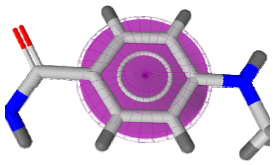
**Hydrophobic Interactions**

- Location spheres by aggregation



**π Interactions**

- Center & normal



Chemical features always refer to the ligand side!



# 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 |


⇨ Layer 3 and Layer 4 features provide good abstraction, and still sufficient characterization

# Existing methods: Kabsch

**Geometric Alignment**

**KABSCH**: Given two **pairwise defined** sets of points of equal size, there is an optimal rotation to minimize RMSD, related to distances.

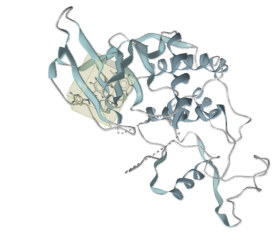
- uses matrix algebra to find the optimal rotation of two sets of points in N-dimensional space to minimize the RMSD between them
- Eigenvector decomposition to derive the orthogonal matrix that describes the best rotation


**[Kabsch, 1976]** Kabsch, W. (1976). A solution for the best rotation to relate two sets of vectors. Acta. Crystal, 32A:922-923.

**[Kabsch, 1978]** Kabsch, W. (1978). A discussion of the solution for the best rotation to related two sets of vectors. Acta. Crystal, 34A:827-828.

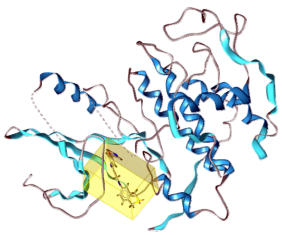



# Kabsch Alignment

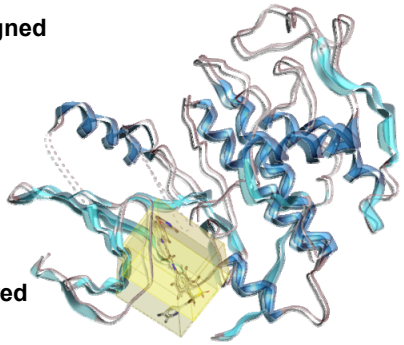


**For sequence-aligned proteins**

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



**One single centered rotation**

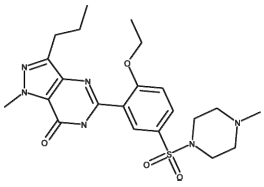


**Computationally efficient**

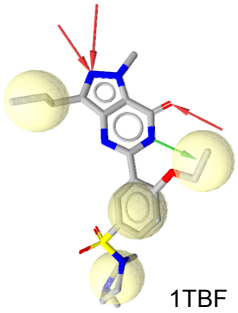
# Kabsch Superposition

## For molecules: atoms instead of sequence

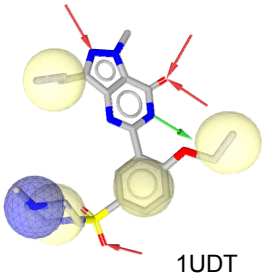
- Canonicalize atom order to create pairs



**Sildenafil**



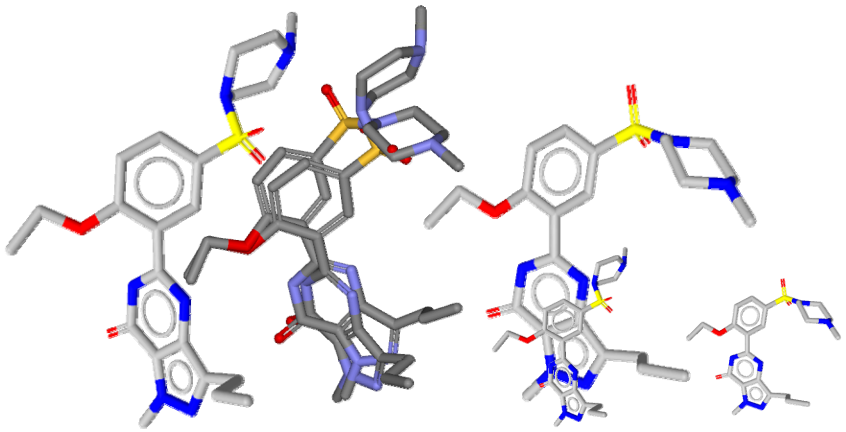
1TBF



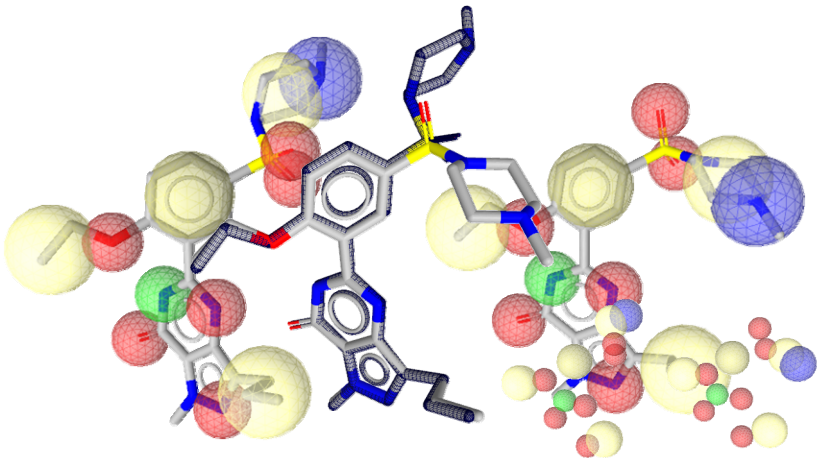
1UDT



# Atom by Atom



# Pharmacophore Points

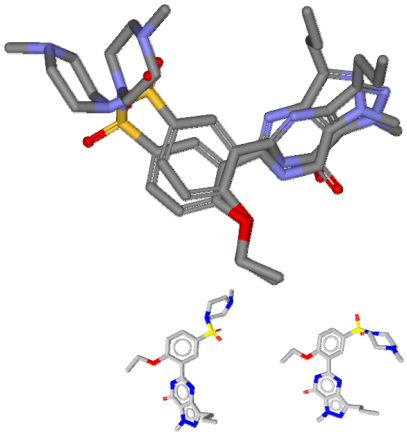


G. Wolber. ACS Spring Meeting, Atlanta, 2006

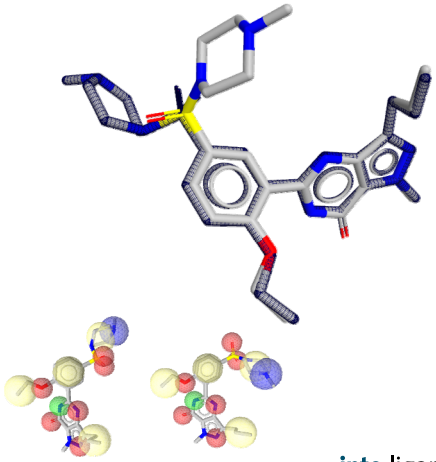
inte:ligand

# Pairwise Superposition

**Ideal rotation using all atom pairs**



**Ideal rotation using pharmacophore points**



G. Wolber. ACS Spring Meeting, Atlanta, 2006

inte:ligand

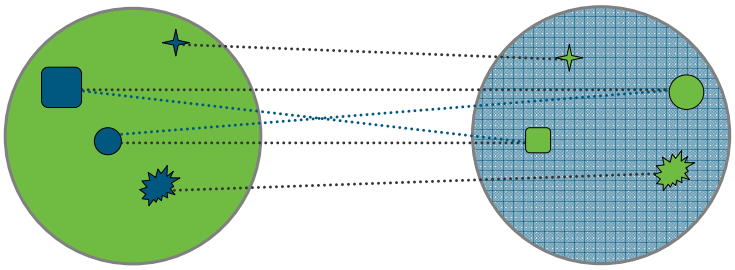
# What Is Missing?


## Best Pairing:

- Algorithm to find a **maximum # of pairs** with **minimum matching cost**
- **Similarity measure** that describes matching cost between pharmacophore points



# How to get optimal pairs?



## Hungarian Matcher (Marrying Problem)

**[Edmods 1965]** Matching and a Polyhedron with 0-1 Vertices. J. Res. NBS 69B (1965), 125-30 [nonbipartite application]

**[Kuhn 1995]** The Hungarian method for the Assignment Problem. Noval Research Quarterly, 2 (1995) [bipartite variant]

**[Richmond 2004]** Application to chemistry: N. Richmond et al. Alignment of 3D molecules using an image recognition algorithm. J. Mol. Graph. Model." 23, 2004, pp 199-209.

# Hungarian Matching

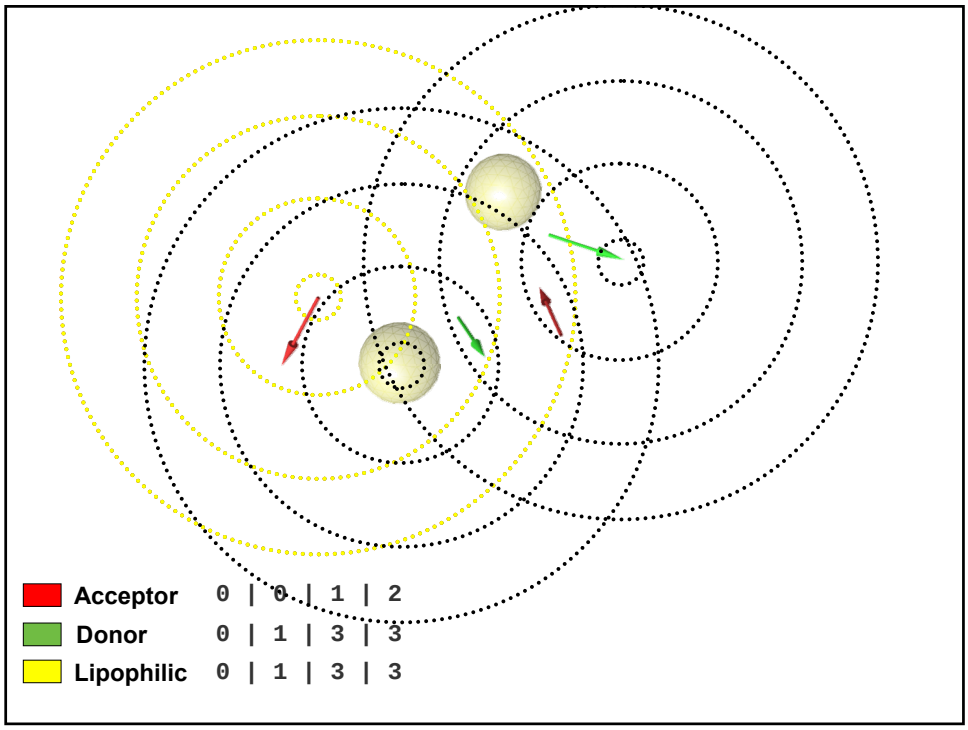
## How to define pharmacophore feature matching cost (similarity)?

- Few feature types
- Selectivity by geometric relations

=> Encode geometry in each feature!



inte:ligand



# Similarity measure

**Similarity =**

Create a "typed shared shell" list for each type


- Take the minimum from each typed shell.
- The higher the sum of the counts in each shell, the higher the potential correspondence for the respective point type.
- Subtract value from max. cost (normalize to minimize cost)



inte:ligand

# Cost function

$$cost(x,y) = \sum_{shell}\left[ weight(shell) * \sum_{type}\left(\min(n_x(shell,type), n_y(shell,type))\right)\right]$$

$n_x(shell,type)$ …. # elements in *shell* of *type* for element *x*
$n_y(shell,type)$ …. # elements in *shell* of *type* for element *y*

$$weight(i_{shell}) = \left(1-\left(\frac{i_{shell}}{\mathrm{n}}\right)^3\right) * 2\left(1-\frac{i_{shell}}{\mathrm{n}}\right) * m$$

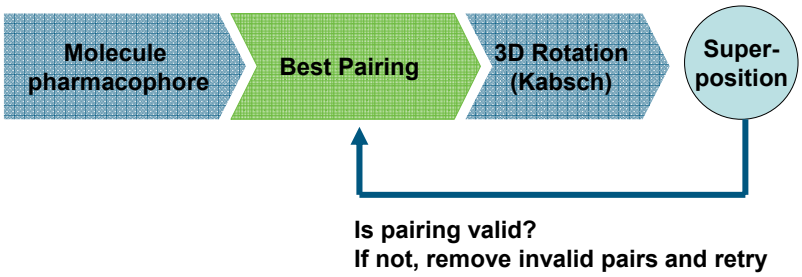
$m$ …. maximum element count for all shells and all types
$n$ …. maximum shell index
$i_{shell}$ … shell index



inte:ligand

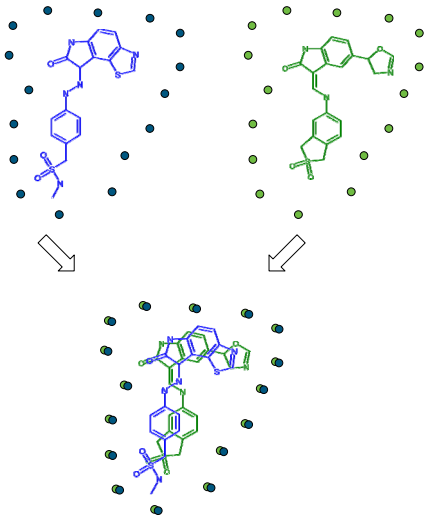
# The Symmetry Problem

- Reduced information may cause false ambiguity for some cases





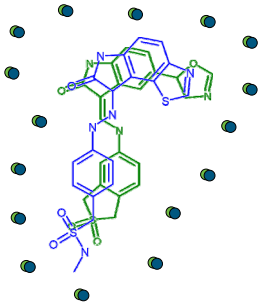

# Validation



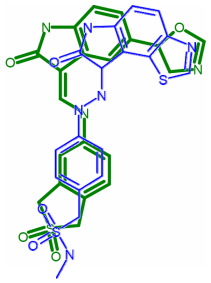
- Bound ligands from the PDB in the same target
- $C_\alpha$ (7A) as reference points
- Alignments of small molecules compared to alinged binding sites

# Validation

Alignment within protein



**compared to**



Pure Ligand-based alignment

**Success measure:**
**RMSD between protein-bound and predicted position**



# Test Set

## 4 Targets

- o COX-2
- o ABL-Tyrosin Kinase
- o PDE5
- o CDK2

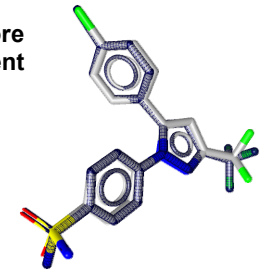
## Prerequisites

- o PDB quality (resolution, reasonable)
- o Binding site similarity (>= 40 similar $C_{\alpha;}$ RMS <1)

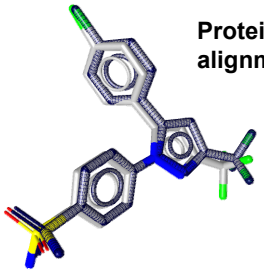
# COX-2

**Pharmacophore alignment**

**RMSD=0.4**



**Protein alignment**



| | 1CX2 | 4COX | 6COX |
|---|---|---|---|
| **1CX2** | - | 2.6 | 0.4 |
| **4COX** | 2.4 | - | 2.7 |
| **6COX** | 0.4 | 2.8 | - |

**1CX2, 6COX** – selective inhibitors (SC558)
**4COX** – non-selective inhibitor (indomethacine)

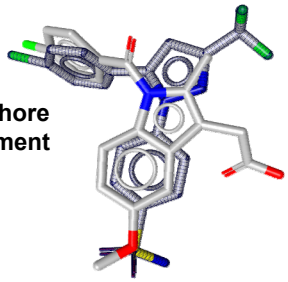
G. Wolber. ACS Spring Meeting, Atlanta, 2006

inte:ligand

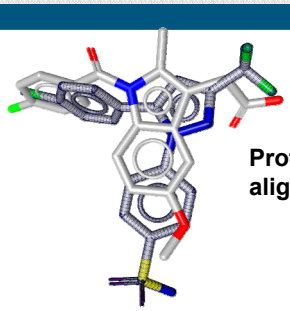
# COX-2

**Pharmacophore alignment**

**RMSD=2.4**



**Protein alignment**



| | 1CX2 | 4COX | 6COX |
|---|---|---|---|
| **1CX2** | - | 2.6 | 0.4 |
| **4COX** | 2.4 | - | 2.7 |
| **6COX** | 0.4 | 2.8 | - |

**1CX2, 6COX** – selective inhibitors (SC558)
**4COX** – non-selective inhibitor (indomethacine)

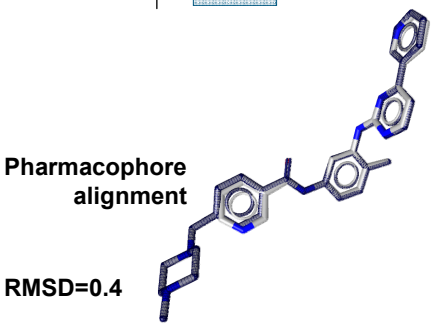
G. Wolber. ACS Spring Meeting, Atlanta, 2006

inte:ligand

# Abl Tyrosin Kinase/Gleevec

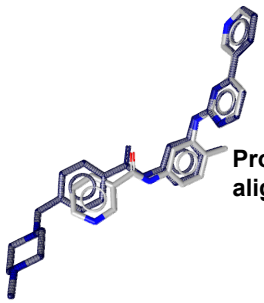
| | 1FPU | 1IEP |
|---|---|---|
| 1FPU | - | 0.4 |
| 1IEP | 0.6 | - |

Gleevec and a variant



Pharmacophore alignment

RMSD=0.4



Protein alignment



inte:ligand

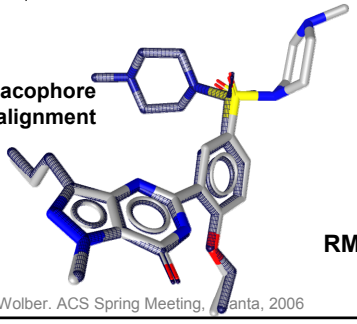
# Phosphodiesterase 5

| | 1UDT | 1TBF | 1XP0 | 1UHO | 1UDU | 1XOZ |
|---|---|---|---|---|---|---|
| 1UDT | - | 0.7 | 0.9 | 0.3 | 2.3 | 2.5 |
| 1TBF | 0.7 | - | 0.4 | 0.7 | 2.7 | 3.1 |
| 1XP0 | 0.8 | 0.4 | - | 0.8 | 2.6 | -- |
| 1UHO | 0.3 | 0.7 | 0.7 | - | 2.3 | 4.5 |
| 1UDU | 2.2 | 2.6 | 2.5 | 1.7 | - | 1.0 |
| 1XOZ | 2.6 | 2.9 | -- | 4.5 | 1.0 | - |

Sildenafil

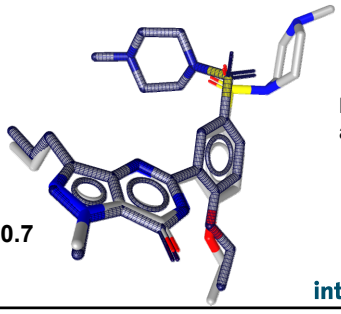
Tadalafil

Vardenafil



Pharmacophore alignment

RMSD=0.7



Protein alignment



inte:ligand

# Phosphodiesterase 5

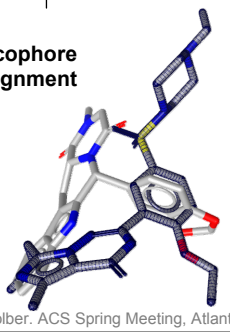
| | 1UDT | 1TBF | 1XP0 | 1UHO | 1UDU | 1XOZ |
|---|---|---|---|---|---|---|
| **1UDT** | - | 0.7 | 0.9 | 0.3 | 2.3 | 2.5 |
| **1TBF** | 0.7 | - | 0.4 | 0.7 | 2.7 | 3.1 |
| **1XP0** | 0.8 | 0.4 | - | 0.8 | 2.6 | -- |
| **1UHO** | 0.3 | 0.7 | 0.7 | - | 2.3 | 4.5 |
| **1UDU** | 2.2 | 2.6 | 2.5 | 1.7 | - | 1.0 |
| **1XOZ** | 2.6 | 2.9 | -- | 4.5 | 1.0 | - |

Sildenafil

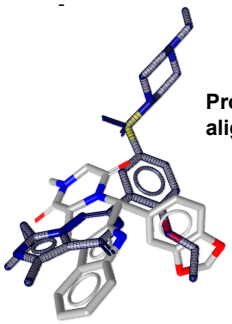
Tadalafil

Vardenafil

**Pharmacophore alignment**



**RMSD=4.5**

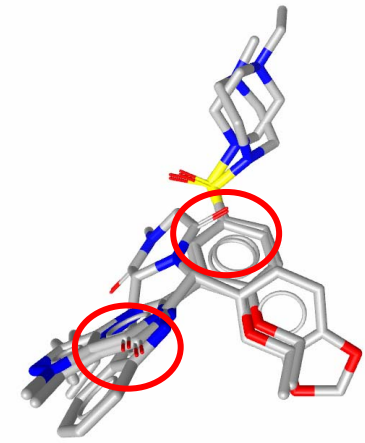


**Protein alignment**



inte:ligand

# Phosphodiesterase 5



**Protein alignment**



inte:ligand

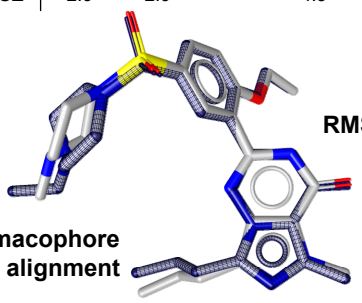
# Phosphodiesterase 5

| | 1UDT | 1TBF | 1XP0 | 1UHO | 1UDU | 1XOZ |
|---|---|---|---|---|---|---|
| 1UDT | - | 0.7 | 0.9 | 0.3 | 2.3 | 2.5 |
| 1TBF | 0.7 | - | 0.4 | 0.7 | 2.7 | 3.1 |
| 1XP0 | 0.8 | 0.4 | - | 0.8 | 2.6 | -- |
| 1UHO | 0.3 | 0.7 | 0.7 | - | 2.3 | 4.5 |
| 1UDU | 2.2 | 2.6 | 2.5 | 1.7 | - | 1.0 |
| 1XOZ | 2.6 | 2.9 | -- | 4.5 | 1.0 | - |

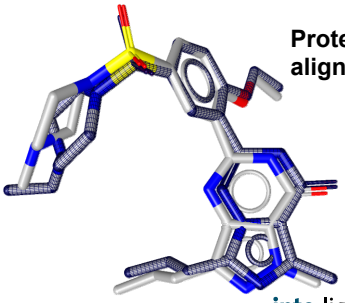
Sildenafil

Tadalafil

Vardenafil



RMSD=0.4



**Protein alignment**

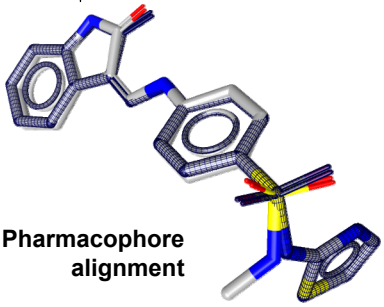
**Pharmacophore alignment**

G. Wolber. ACS Spring Meeting, Atlanta, 2006

inte:ligand

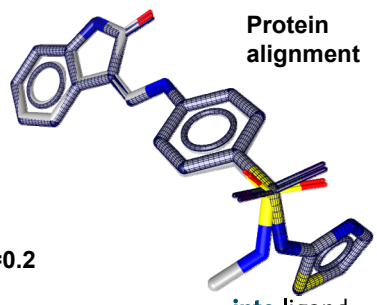
# CDK2

| | 1KE5 | 1KE6 | 1KE7 | 1KE8 | 1KE9 |
|---|---|---|---|---|---|
| 1KE5 | - | 0.4 | 0.6 | 0.2 | 0.3 |
| 1KE6 | 0.5 | - | 0.5 | 0.4 | 0.3 |
| 1KE7 | 0.6 | 0.6 | - | 0.7 | 0.3 |
| 1KE8 | 0.3 | 0.4 | 0.7 | - | 0.4 |
| 1KE9 | 0.3 | 0.3 | 0.3 | 0.3 | - |



**Pharmacophore alignment**

RMSD=0.2

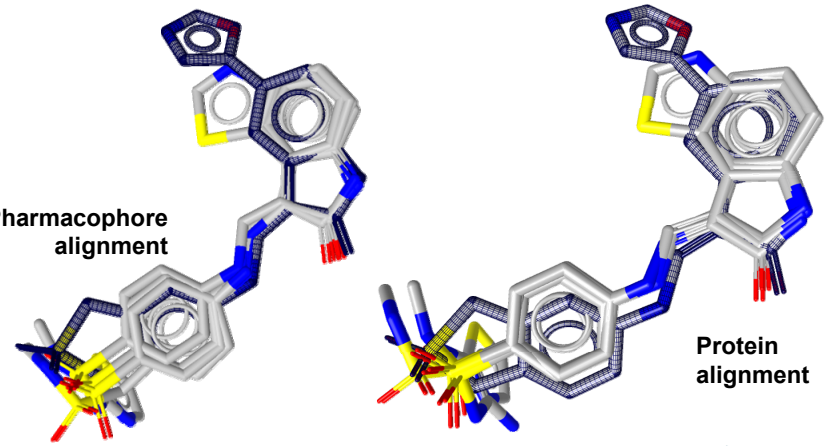


**Protein alignment**

G. Wolber. ACS Spring Meeting, Atlanta, 2006

inte:ligand

## CDK2 (all 5)

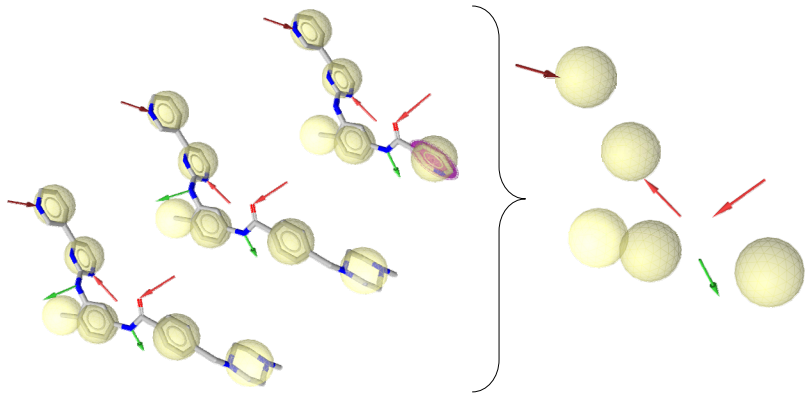





## Validation Summary

- o The alignment worked perfectly in nearly all of the cases
- o Very Fast (up to max. 50 ms per alignment)
- o Scales polynomially with the number of features
- o Provides pharmacophoric view on molecule (scaffold-hopping)

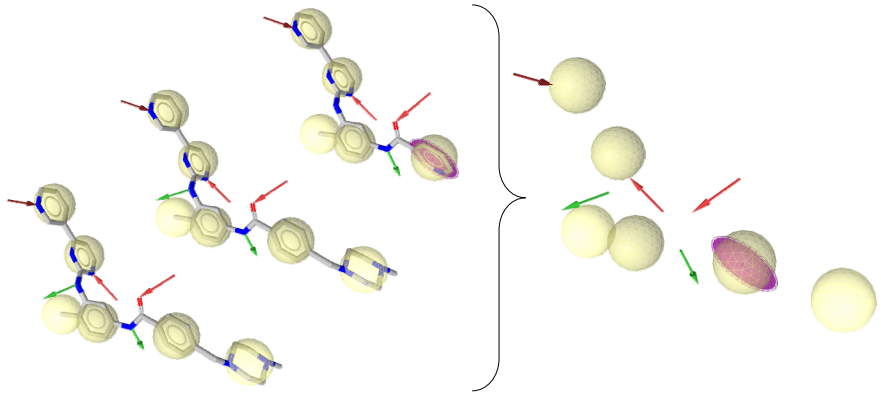
## Shared Feature Pharmacophore



G. Wolber. ACS Spring Meeting, Atlanta, 2006

inte:ligand

## Merged Feature Pharmacophore



G. Wolber. ACS Spring Meeting, Atlanta, 2006

inte:ligand

# Summary

**Universal, effective and fast (20-50ms) pharmacophore alignment method**

**Can be used for:**

- o Comparison of molecules by their pharmacophore features
- o Model & compare pharmacophores (share/merge)
- o Fast and accurate 3D pharmacophore screening



# Thanks to …

inte:ligand

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

universität innsbruck

- Thierry Langer
- Johannes Kirchmair
- Christian Laggner
- Daniela Schuster
- Markus Böhler