



syngenta

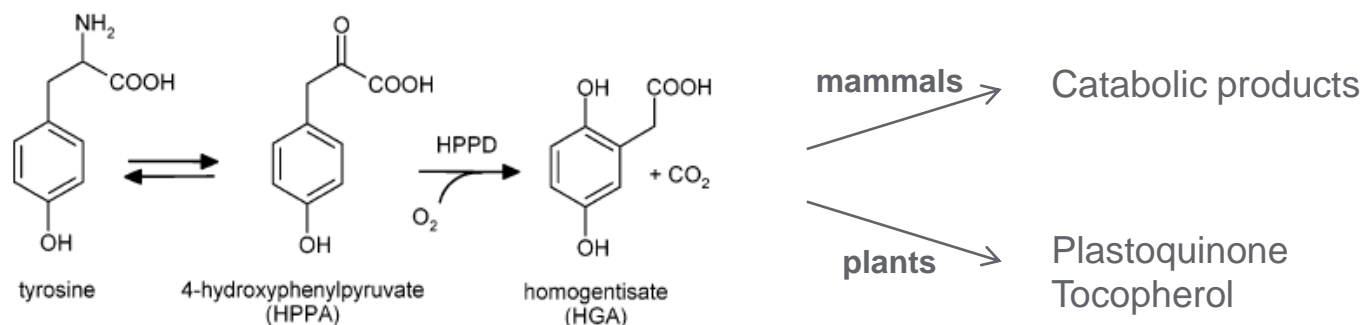
HPPD: ligand- and target-based virtual screening on a herbicide target

Miriam López Ramos

UK QSAR meeting, Jealott's Hill, 22-10-2009

HPPD: a herbicide target

- HPPD: 4-hydroxyphenylpyruvate dioxygenase



- A pharmaceutical target: treatment of hypertyrosinemia (unregulated degradation of tyrosine into toxic metabolites)
- A herbicide target: HPPD inhibition
 - prevents production of plastoquinones → no energy from photosynthesis
 - restricts vitamin E production → sensitivity to oxidative stress
 - prevents carotenoid biosynthesis → chlorophyll no longer protected against UV → bleaching

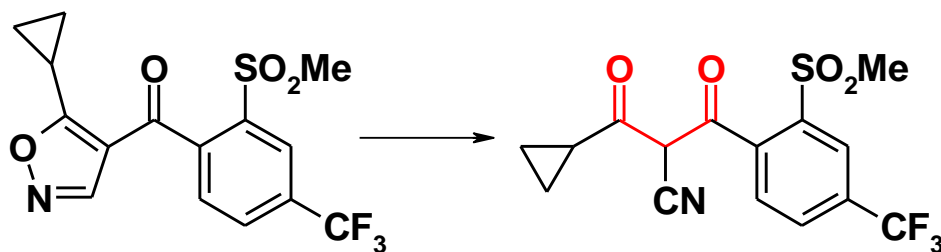
Biochemistry, 43, 10414 (2004)

2 *Phytochemistry*, 68, 2004-2014 (2007)

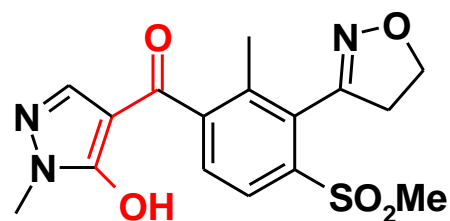
Outlooks on Pest Management, 20, 1, 27-30 (2009)

HPPD: a herbicide target

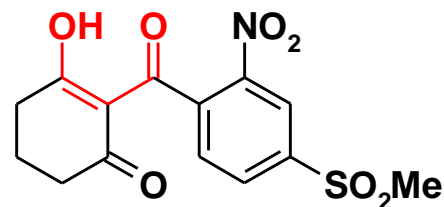
- HPPD is a Fe(II) containing, non-heme protein
- Typical inhibitors:



Diketonitriles: Isoxaflutole (IFT)
(Balance[®] and Merlin[®], Bayer 1996)



Pyrazolones: Topramezone
(Clio[®] and Impact[®], BASF 2006)



Triketones: Mesotrione
(Callisto[®] and Lumax[®], Syngenta 2002)

HPPD at Syngenta

- First inhibitor discovered in 1977 by serendipity (natural product)
- More than 25 years of research
- Well-known mode of action and clear SAR
- 6000 compounds synthesised
- Several co-crystal structures solved in house
- Blockbusters on the market: mesotrione reached \$270 million sales by 2004



Good case study to:

- implement a typical pharma computational approach in agro
- evaluate the performance of virtual screening methods available in house

HPPD: ligand- and target-based virtual screening

- Aim of the study: a retrospective work: let's imagine that...
 - the HPPD project is just starting
 - we only have a few active compounds and a crystal structure
 - we want to look for other compounds/scaffolds in our corporate collection and/or commercial libraries
- Which virtual screening tool would give the best chances to retrieve the highest number of active compounds?



Assess the relative performance of the virtual screening methods

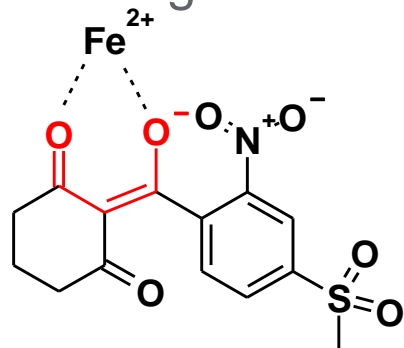
- using a set of true actives and a decoy
- evaluating active compounds retrieval and scaffold hopping possibilities

HPPD: ligand- and target-based virtual screening

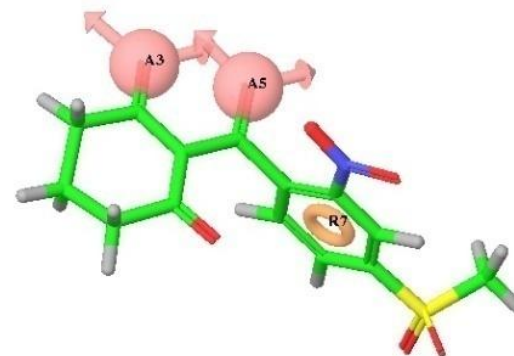
- Choice of reference ligand and target crystal structure
- Building of the database used
- Virtual screening methods used & results
- Conclusions

Reference ligand & target protein

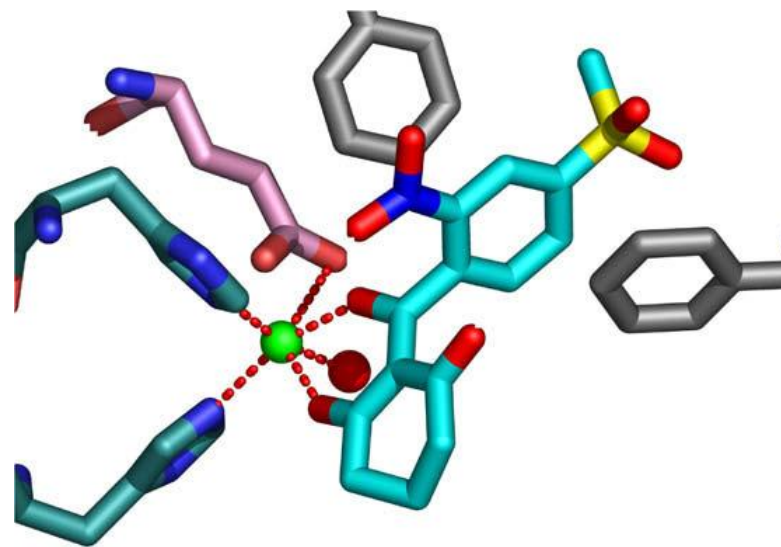
- Reference ligand: mesotrione



pharmacophore →



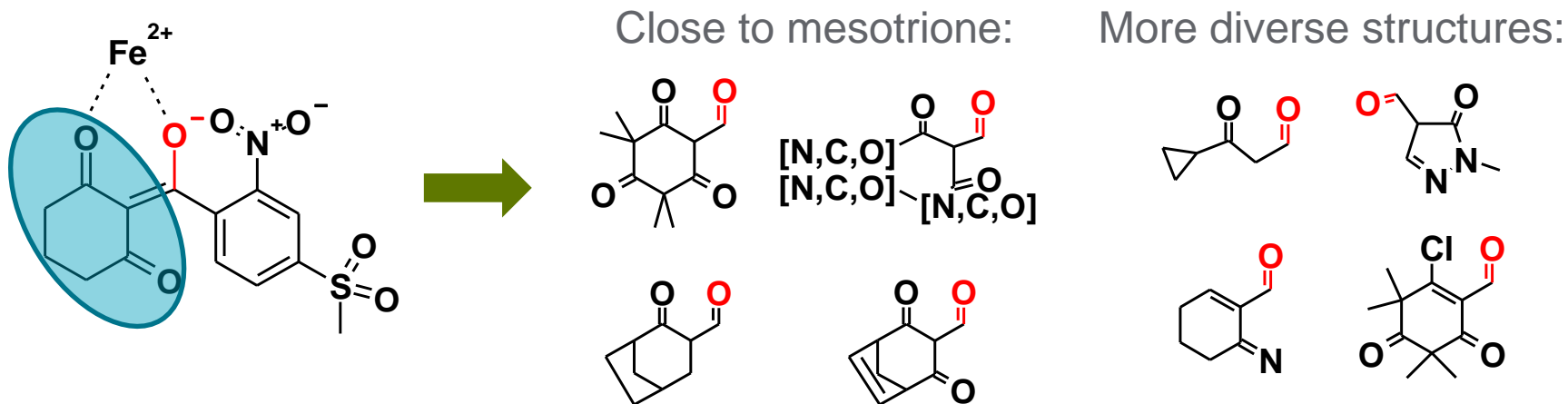
- Target protein: one of the most representative crystals



The immediate binding site around the Fe(II) from a crystal structure of mesotrione bound to Arabidopsis HPPD

Database building: active compounds

- In vitro IC_{50} values had been determined for the most active and representative compounds
 - 216 compounds with $IC_{50} < 1 \mu\text{M}$ as active set
- Chemical diversity: 15 scaffolds present
 - analyse minimal fraction of the database to select to find first compound from each group



Database building: decoy compounds

- No agro equivalent of the MDDR database
- 1st idea: use the corporate database
 - too diverse, difficult to filter and cluster to retain only agro-like compounds
 - the results were too good!!
- 2nd trial: use the Pesticide Manual compounds as starting point
 - 1,373 products commercialised for agricultural use
 - Select in the „cleaned“ corporate database the most similar compounds to each of them
 - Remove all compounds associated with the HPPD project
 - 49,549 compounds as decoy set

HPPD: ligand- and target-based virtual screening

- Choice of reference ligand and target crystal structure
- Building of the database used
- Virtual screening methods used & results
 - 2D ligand-based
 - 3D ligand-based
 - Target-based
- Conclusions

Virtual screening methods

- Ligand-based 2D:
 - 12 Pipeline Pilot's circular fingerprints, using

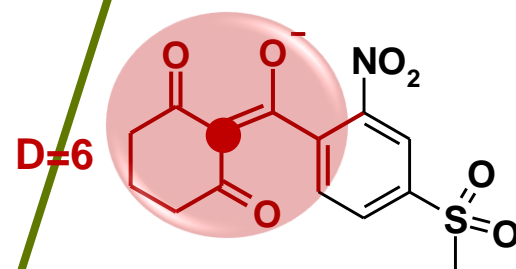
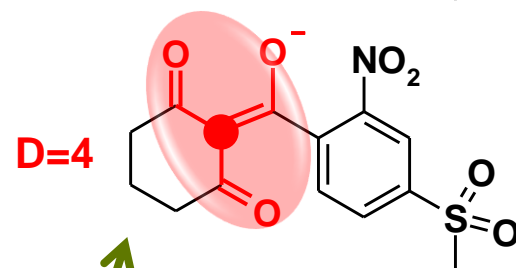
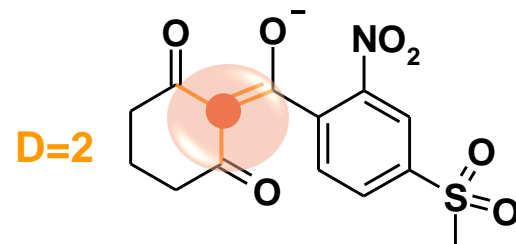
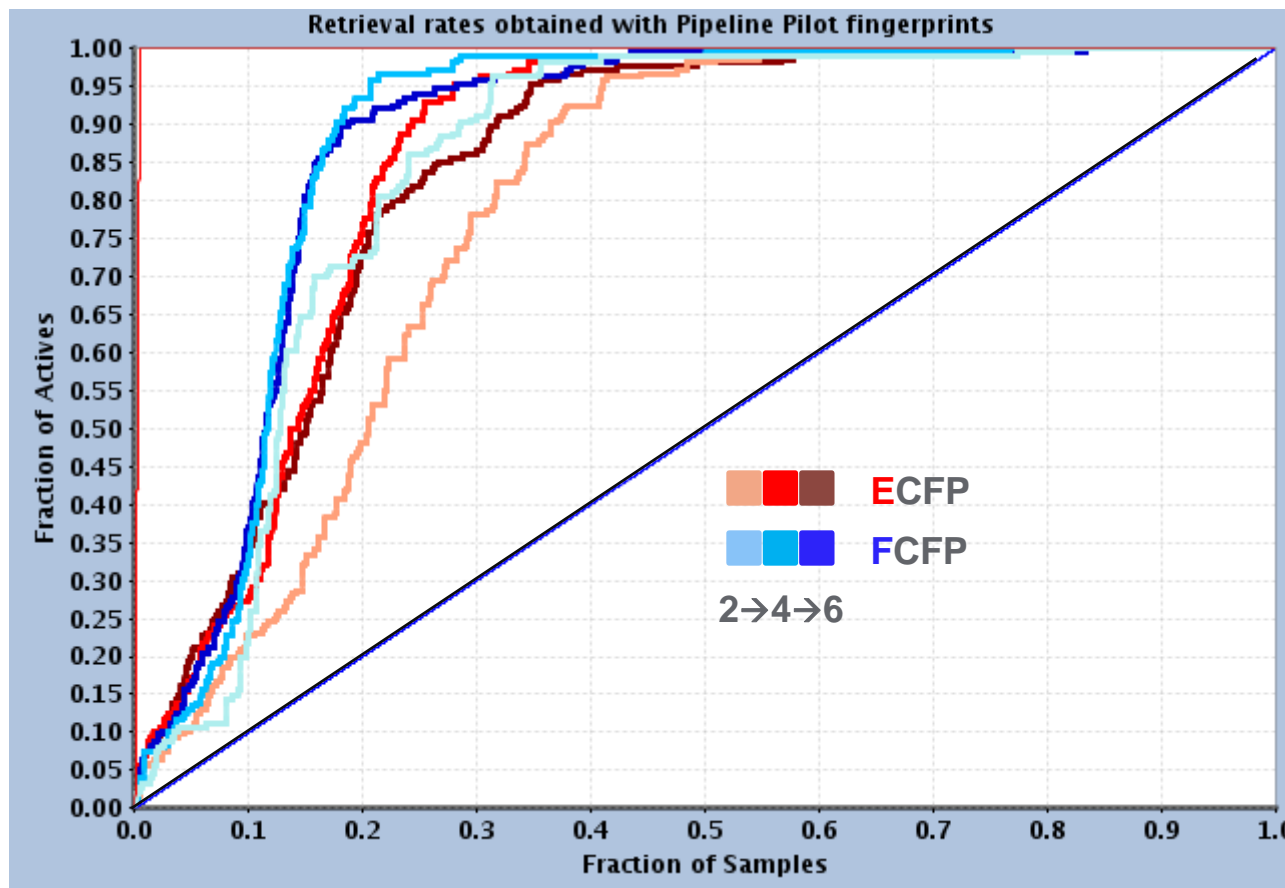
		Atom typing		
		Functional (F)	Atom types (E)	
Feature evaluation	Presence/ absence (P)	FCFP_2	ECFP_2	2
		FCFP_4	ECFP_4	4
		FCFP_6	ECFP_6	6
	Counts (C)	FCFC_2	ECFC_2	2
		FCFC_4	ECFC_4	4
		FCFC_6	ECFC_6	6

- Pipeline Pilot's MDL keys (not shown)

2D: Pipeline Pilot's fingerprints: diameter

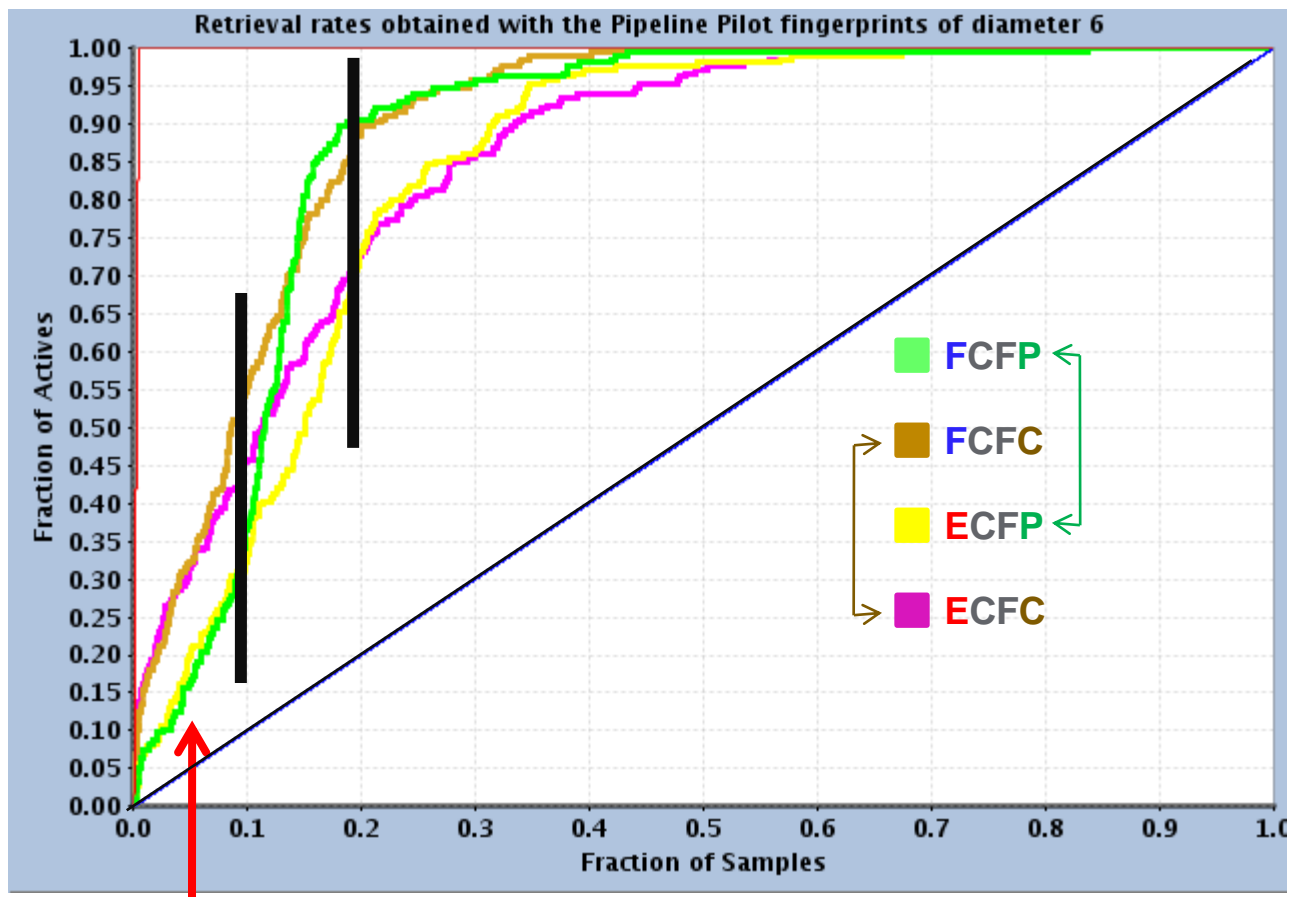
Influence of diameter:

$$2 < 4 \sim 6$$



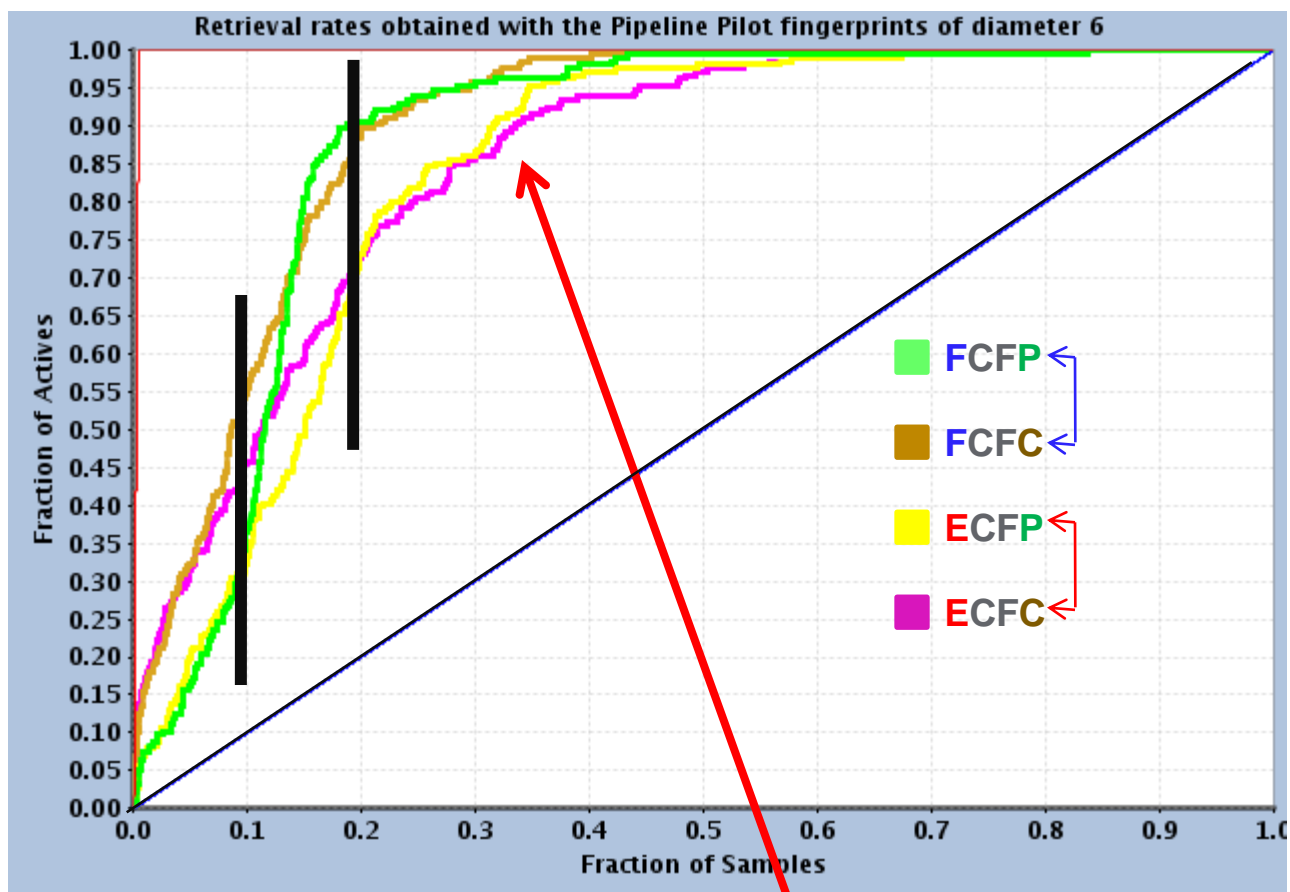
Has all the important features

2D: Pipeline Pilot's fingerprints: atom type and feature evaluation



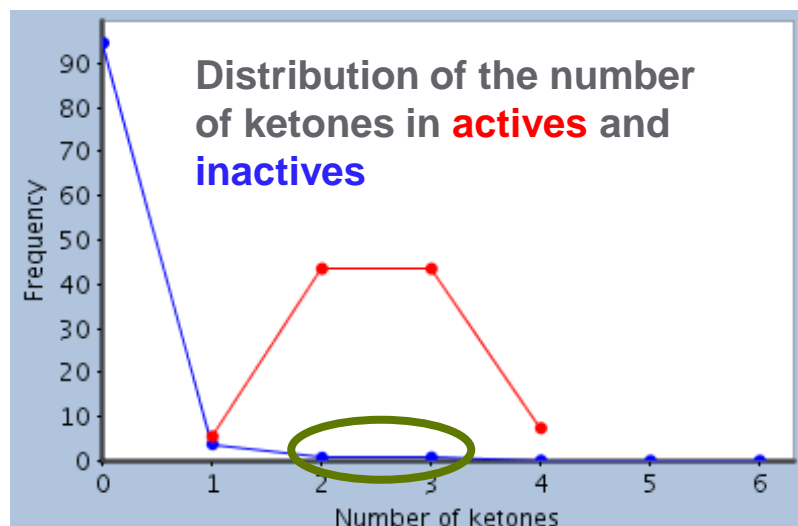
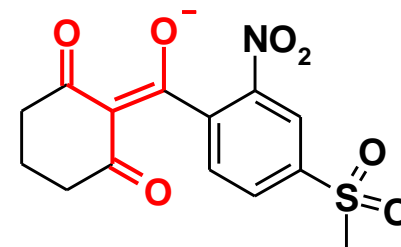
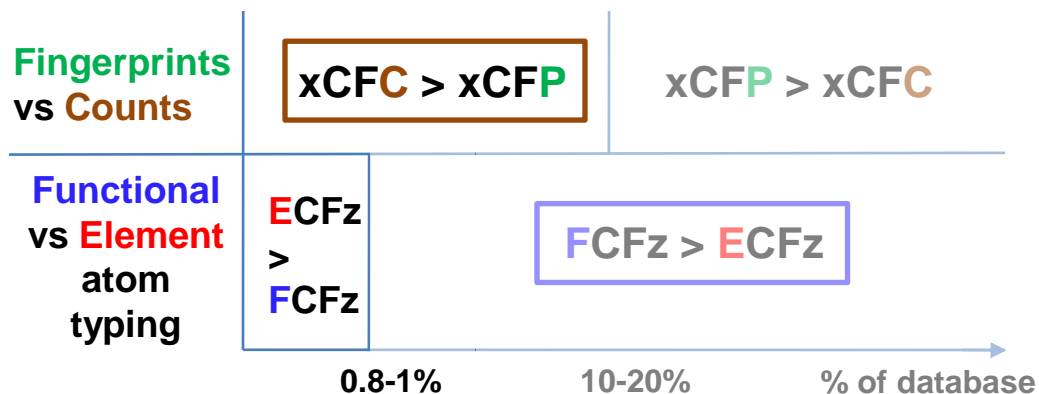
Counts > fingerprints
Atom types no influence

2D: Pipeline Pilot's fingerprints: atom type and feature evaluation



Functional > Element atom types
Feature evaluation no influence

2D: Pipeline Pilot's fingerprints: atom type and feature evaluation

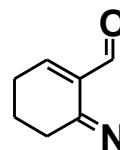
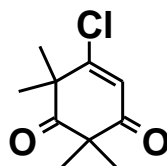
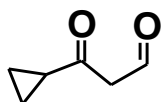
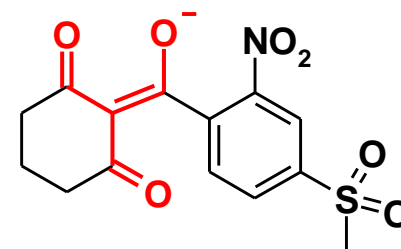
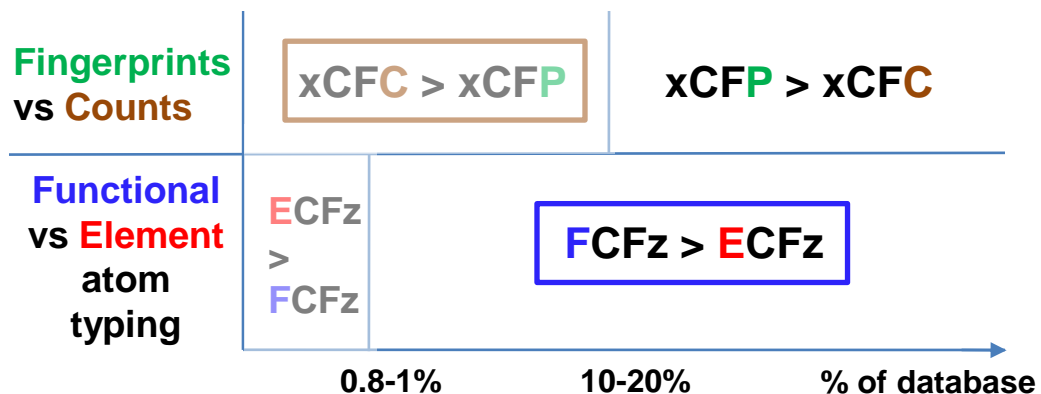


Retrieving compounds with 2 or 3 ketones and not just 1 will give more actives!



Counts > fingerprints as long as all diketones and triketones have not been retrieved

2D: Pipeline Pilot's fingerprints: atom type and feature evaluation



- Compounds with more diverse structures are the last to be retrieved
- **Functional atom-typing allows better scaffold hopping** and gives best results

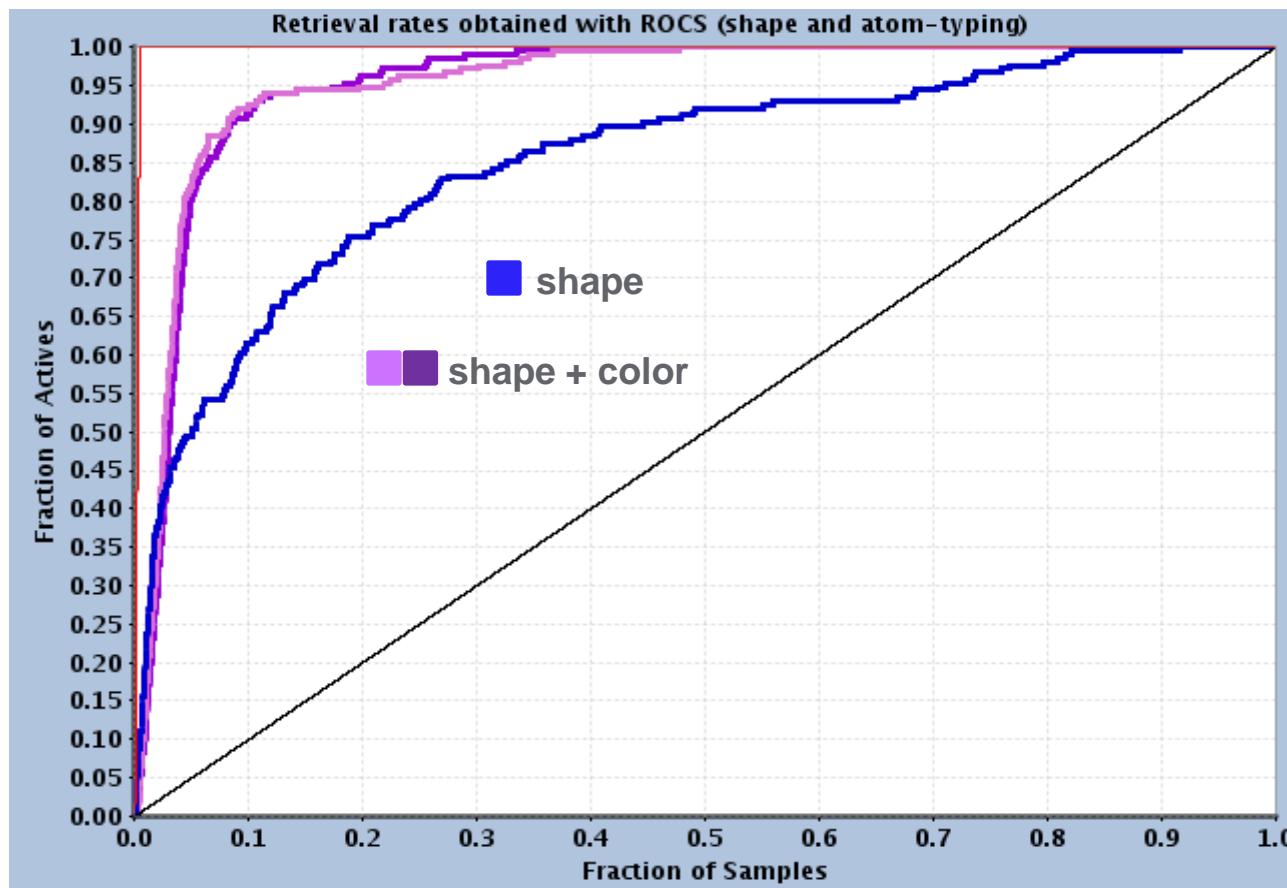
HPPD: ligand- and target-based virtual screening

- Choice of reference ligand and target crystal structure
- Building of the database used
- Virtual screening methods used & results
 - 2D ligand-based
 - 3D ligand-based
 - Target-based
- Conclusions

3D ligand-based virtual screening: methods

- Ligand-based 3D:
 - ROCS (Openeye):
 - Shape only
 - Using a color force field (ImplicitMillsDean)
 - Using a color force field and gradients (optchem)
 - ROCS + EON reranking of all ROCS runs
 - Phase (Schrödinger):
 - Shape only
 - Including atom-typing:
 - Elements
 - MacroModel types
 - Phase pharmacophore types
 - Phase pharmacophore search

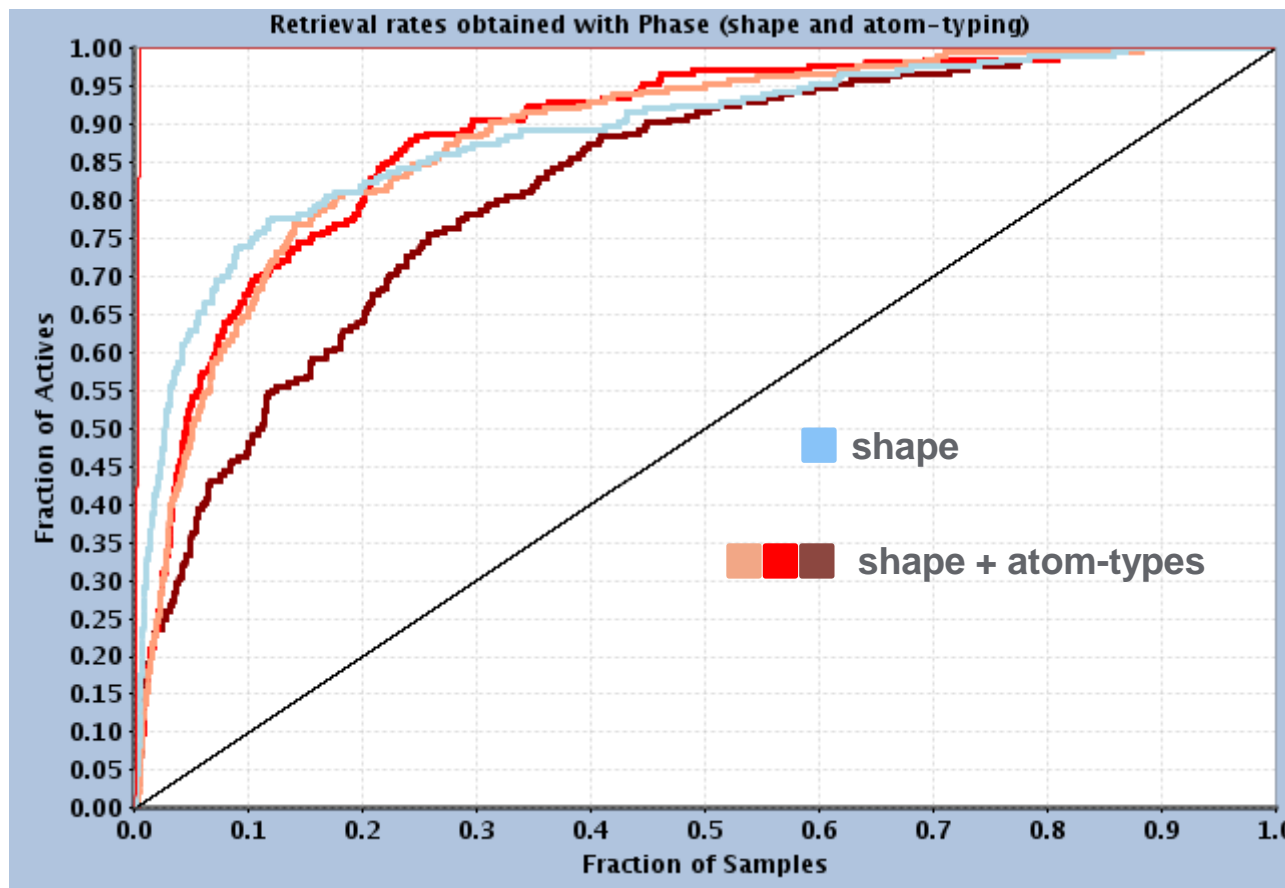
3D ligand-based: ROCS



shape + color
>>
shape only

- with atom types (color):
→ retrieval rate ↑ ↑
→ scaffold hopping ↓ ↓
- with gradients in color :
→ retrieval rate ~
→ scaffold hopping ↓

3D ligand-based: Phase



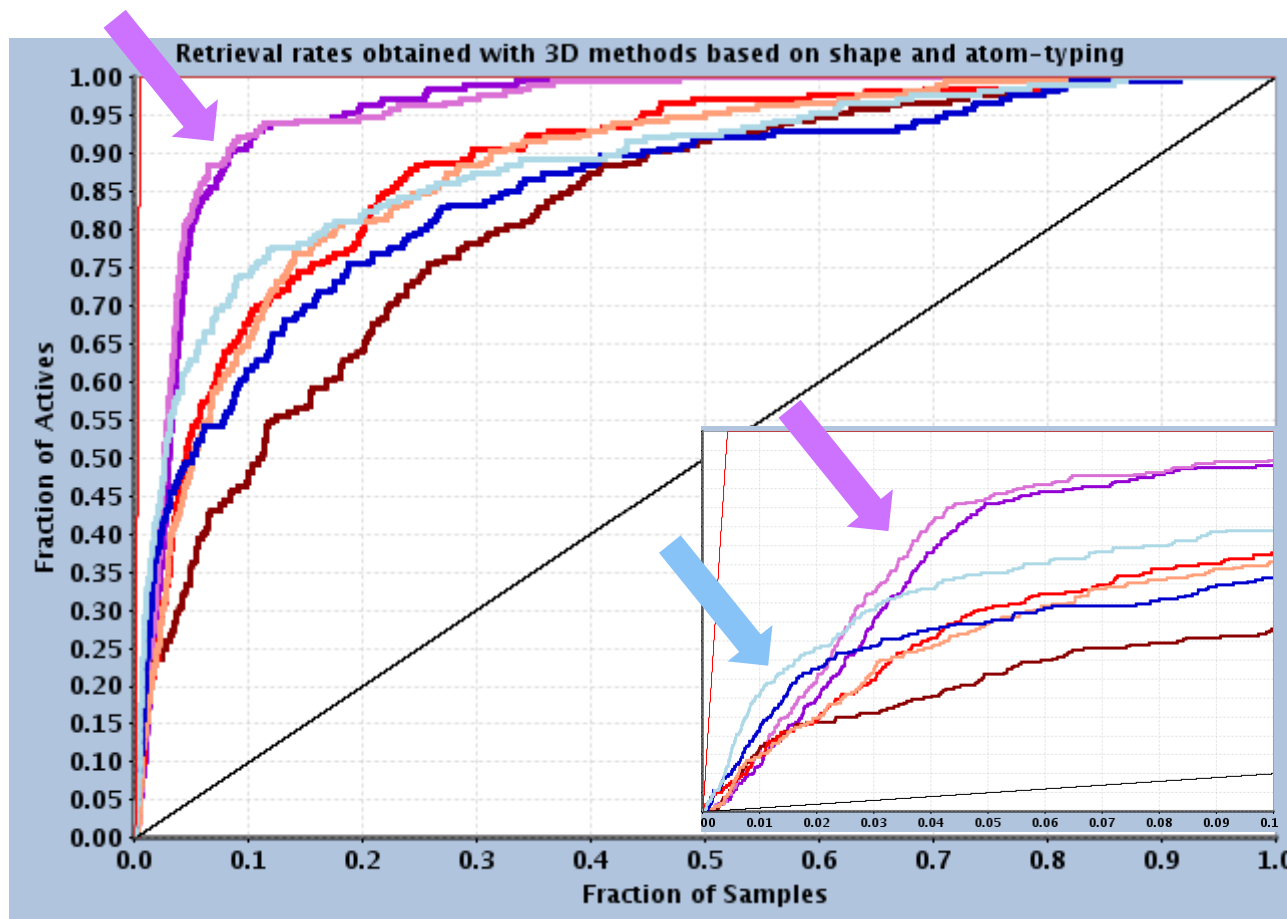
shape only
>
shape + atom-types

With atom types:

→ retrieval rate ↓
(**pharmacophore** ↓ ↓)

→ scaffold hopping ↓
(**macromodel** < **elements**)

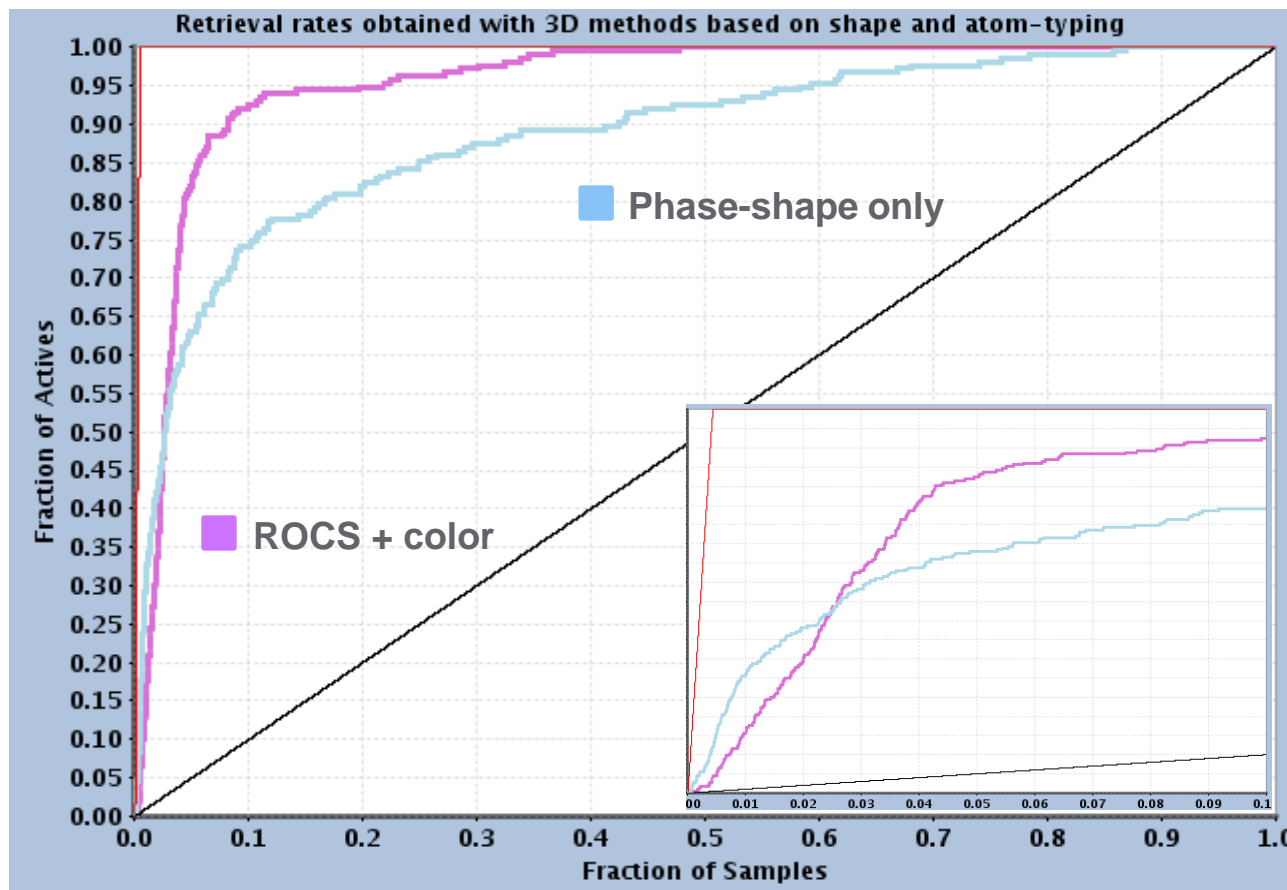
3D ligand-based: shape and atom-types



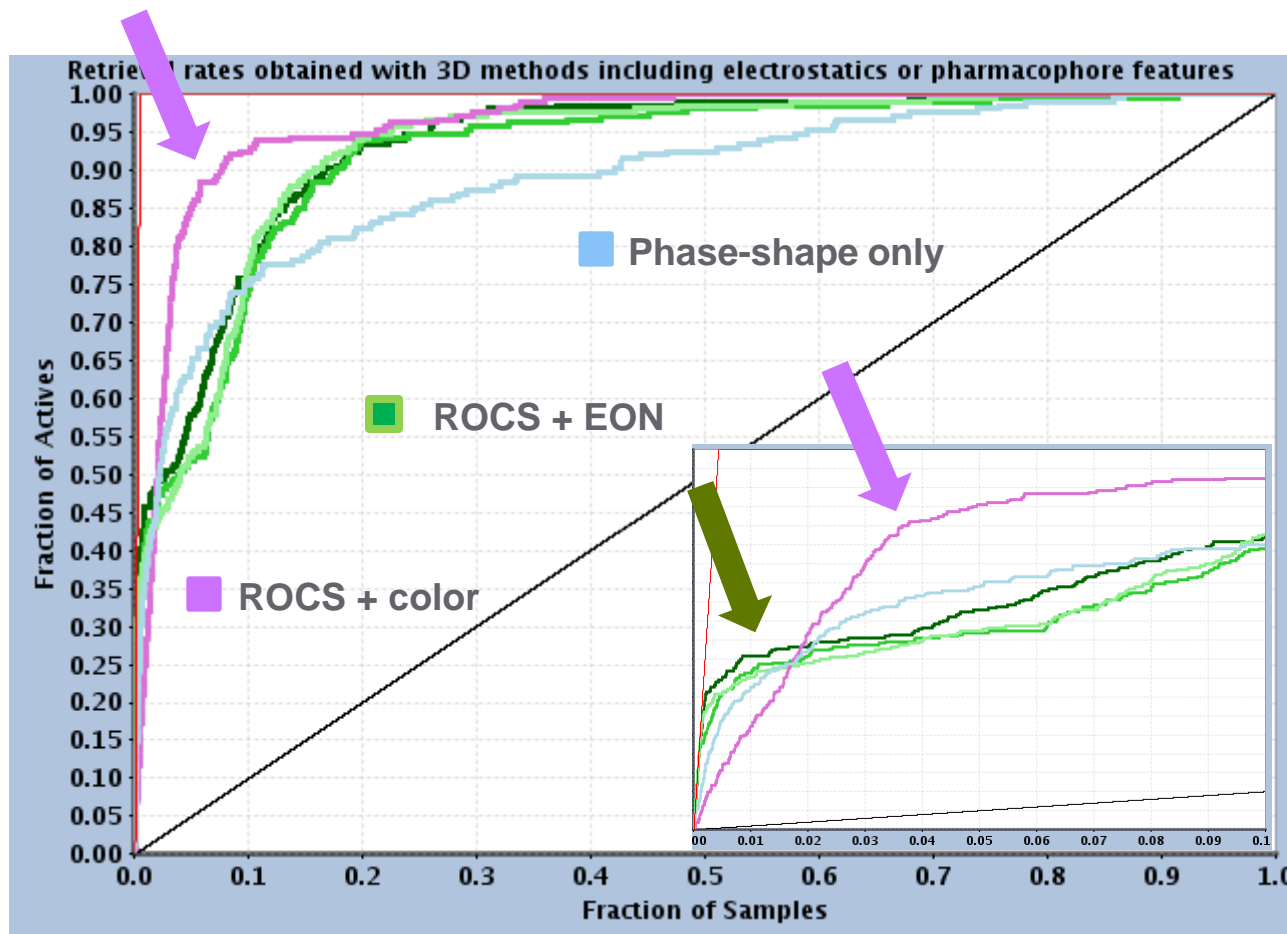
■ ROCS ■ + atom-types
■ Phase ■ + atom-types

Phase-shape and
ROCS + color perform
best

3D ligand-based: include electrostatics



3D ligand-based: include electrostatics



- All EON equivalent
- Plateau: difficulty to retrieve more diverse compounds

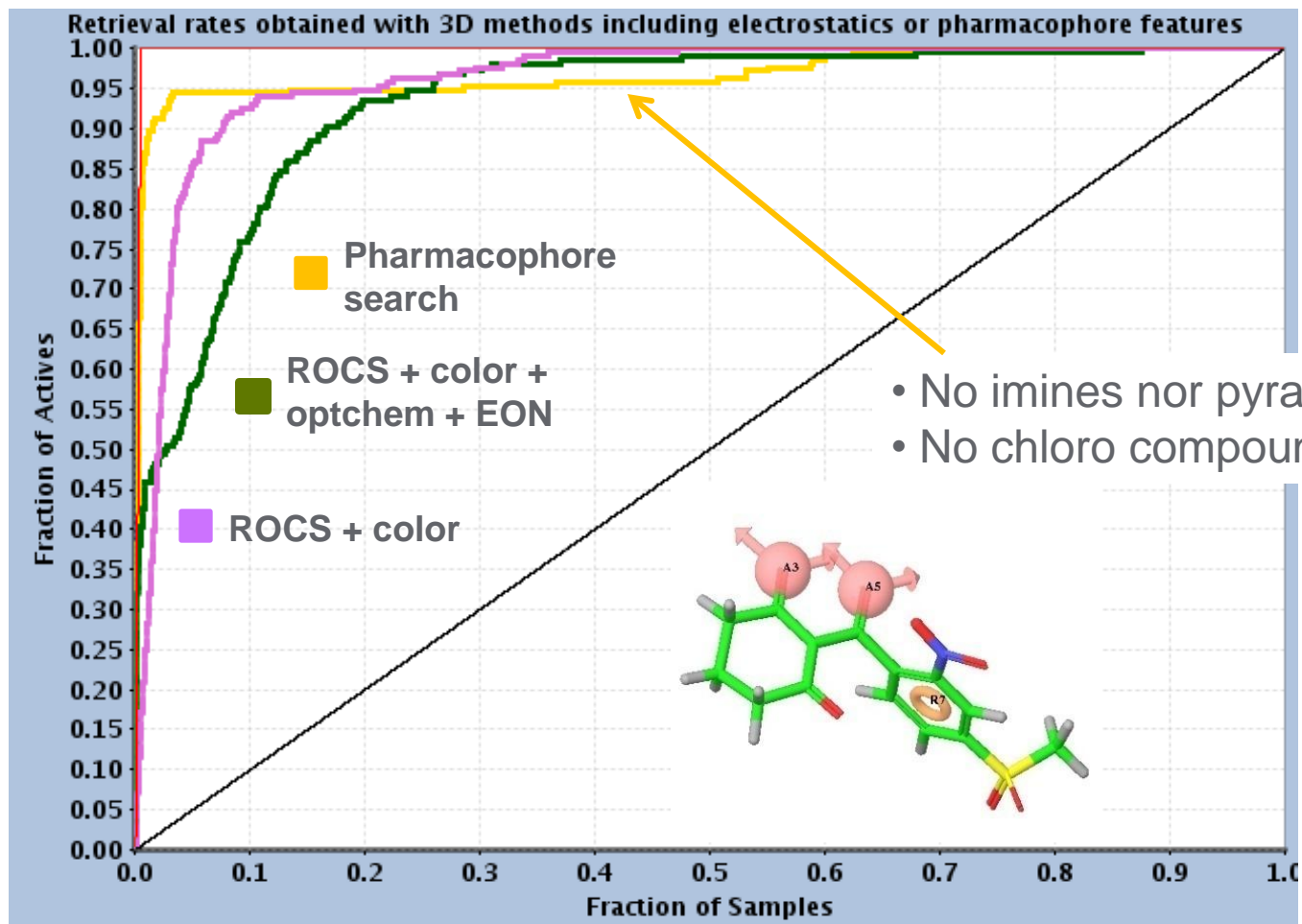


EON good to enhance retrieval rate but not for scaffold hopping

Best choices:

- ROCS + color + optchem + EON
- ROCS + color

3D ligand-based: search matches for pharmacophore (Phase)



3D ligand-based: conclusions

- Best 3D method is search for matches for pharmacophore hypothesis
But too easy test case, result might not be general
- ROCS & Phase methods:
 - Retrieval rates: best choice is ROCS+color (eventually with optchem and EON)
 - Scaffold hopping: ***the more strongly defined the atom types are, the less prone for scaffold hopping the method is***

For scaffold hopping:

- Phase-shape Macromodel < Phase-shape elements < Phase-shape only
- ROCS + EON < ROCS + color + optchem < ROCS + color < ROCS

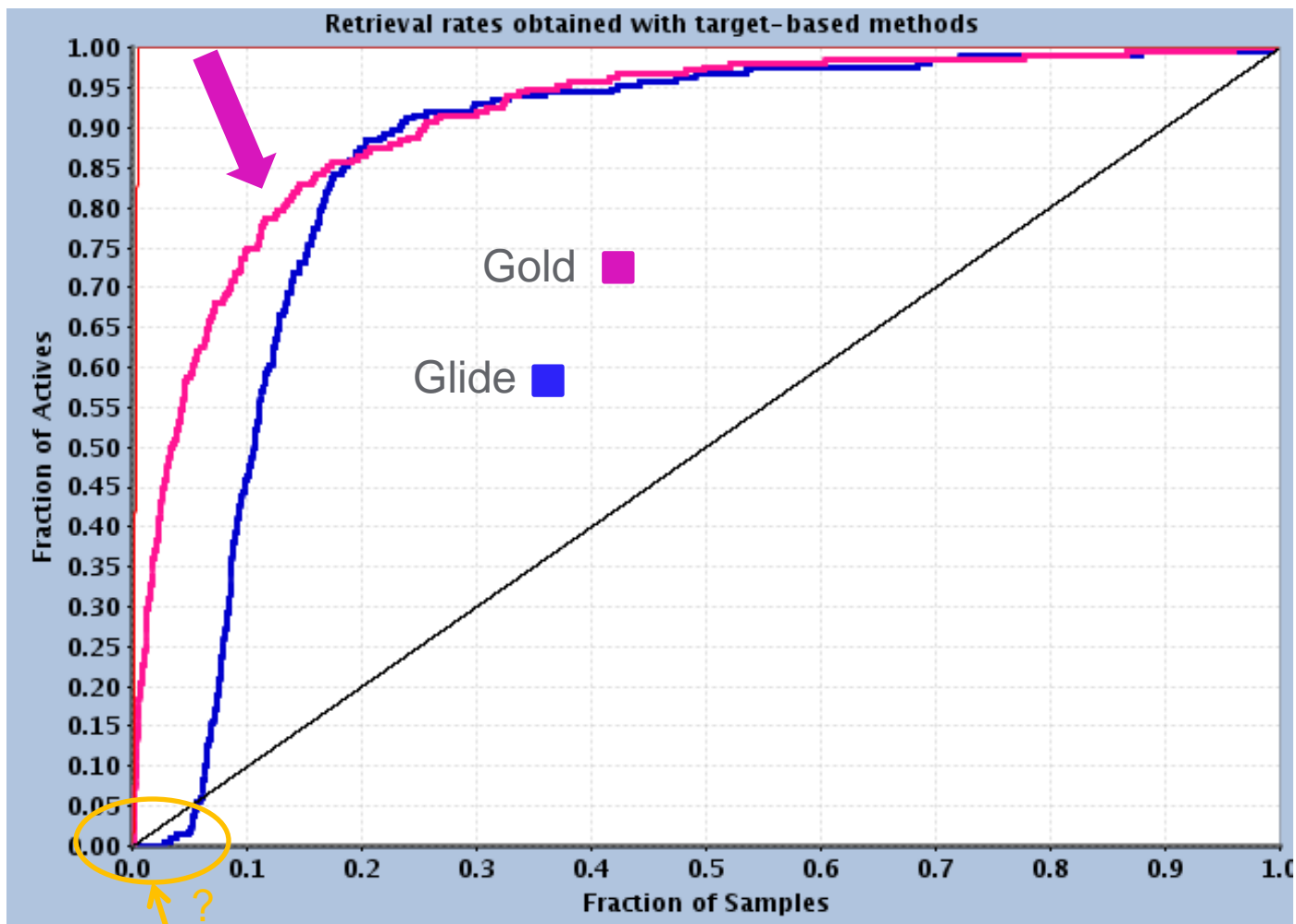
HPPD: ligand- and target-based virtual screening

- Choice of reference ligand and target crystal structure
- Building of the database used
- Virtual screening methods used & results
 - 2D ligand-based
 - 3D ligand-based
 - Target-based
- Conclusions

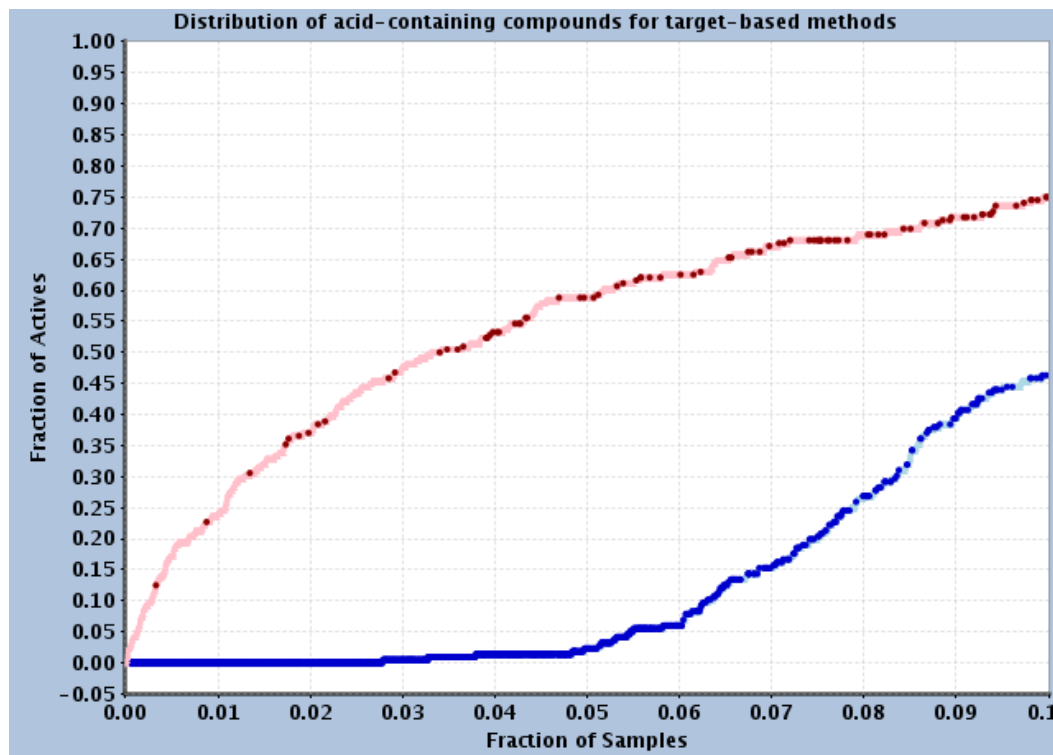
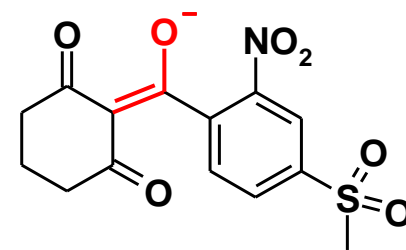
Target-based virtual screening: methods

- Target-based:
 - Glide (Schrödinger)
 - Gold (CCDC)

Target-based: results



Target-based: initial Glide plateau



	acid	no acid
Gold		
Glide		

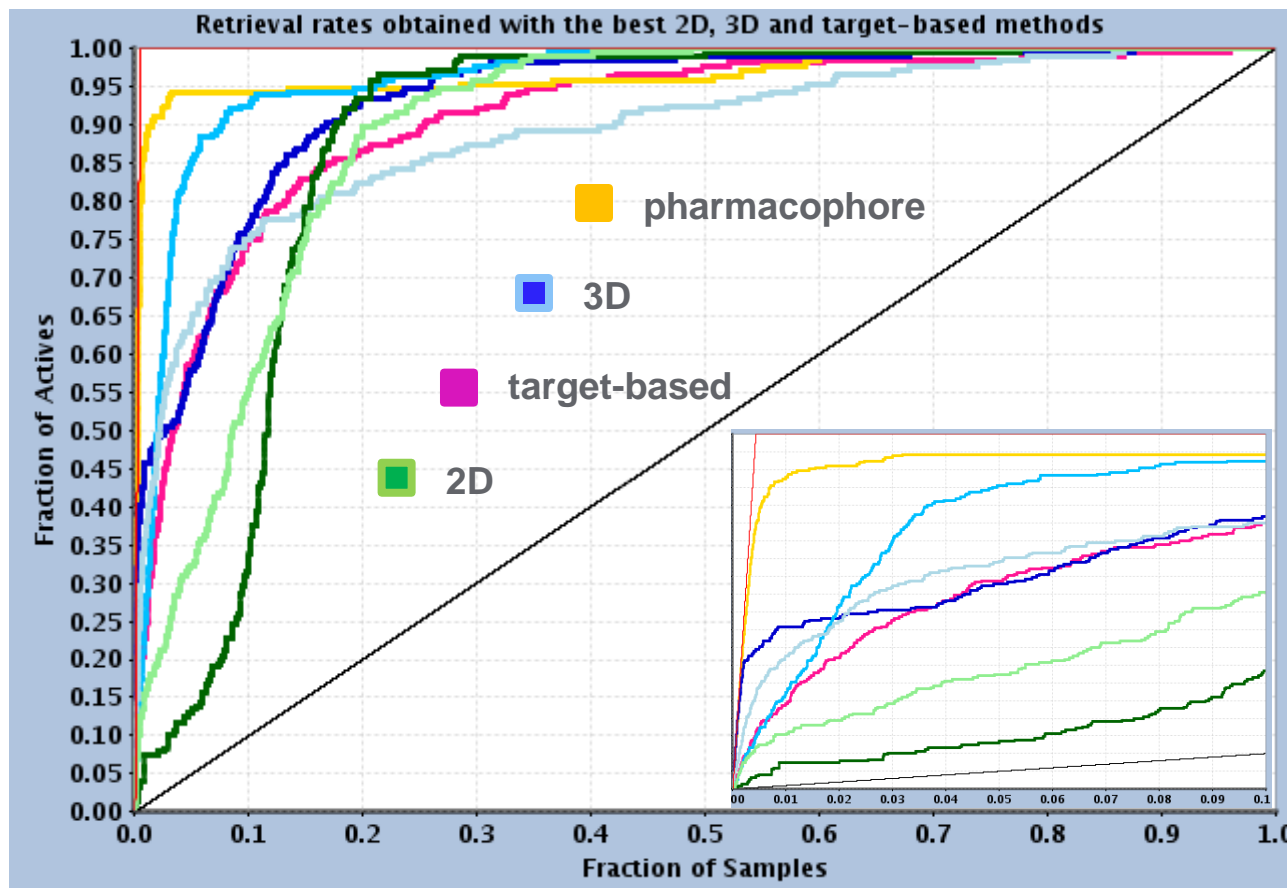
All the first picked compounds are acids!!
(wrong chelation geometry)

- 1) Glide scoring with metals: when metal positively charged, **only interactions with anionic ligands** are included
- 2) **No deprotonated tautomers** generated



No actives picked until 0.5%, then good performance

Conclusions



Conclusions

- Best methods for retrieval rates of active compounds:
 - Overall: pharmacophore search (not general result)
 - Otherwise: ROCS + color
- Scaffold hopping:
 - 2D, 3D and target-based methods are globally equivalent
 - In ligand-based methods, the more defined the atom type the less diversity retrieved
- Correlation between increased complexity and increased results:
 - Good within ligand-based methods
 - Weak 2D → 3D → target-based since target-based < 3D
- Pitfalls...
 - Database generation & preparation
 - Non optimal default parameters → adjust to target?

Acknowledgements

Thanks to...

- Francesca Perruccio (Computational Chemist)
- Jane Wibley (Crystallographer)
- Glynn Mitchell (HPPD Project Leader)