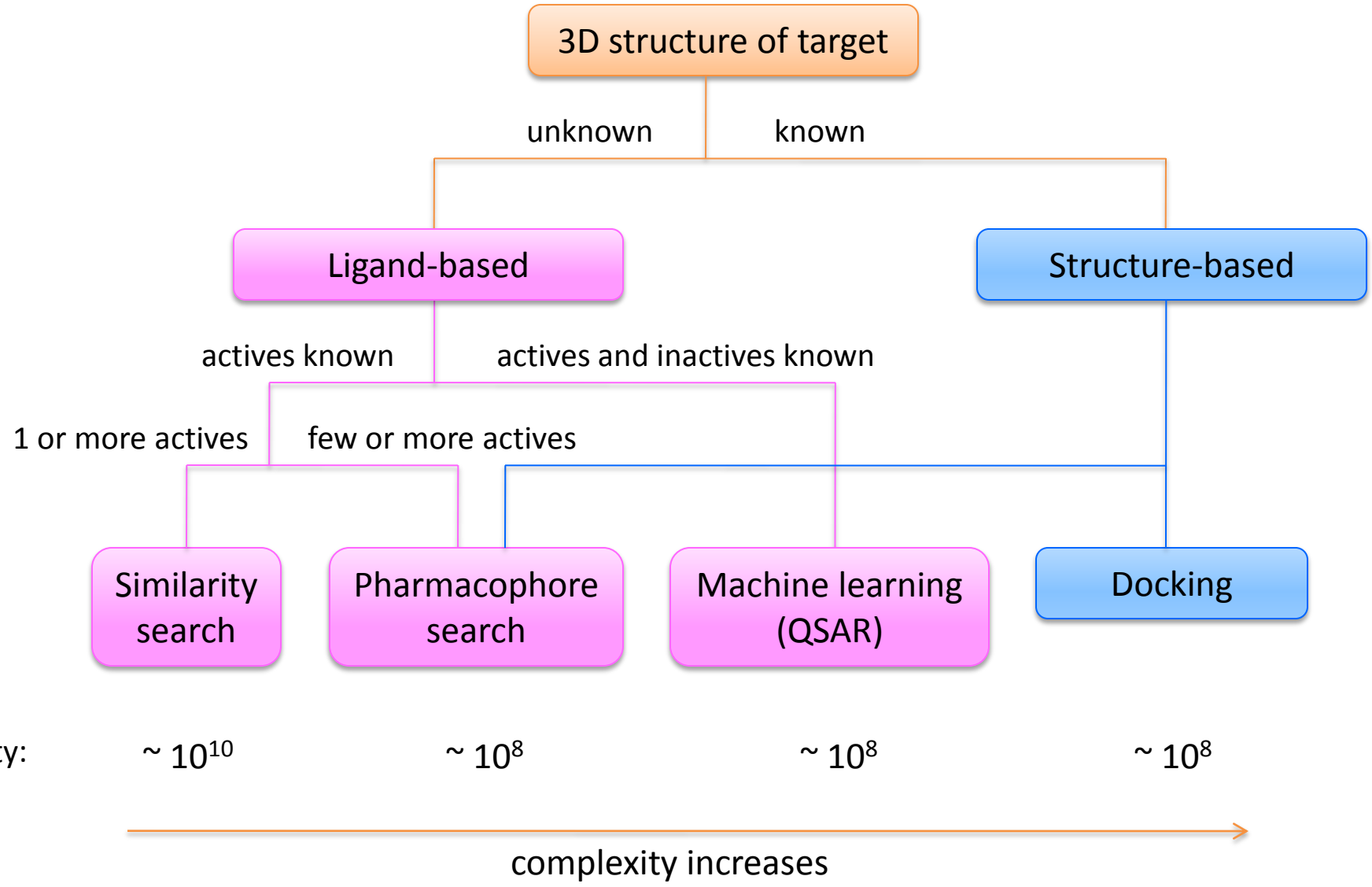
Pharmacophore modeling

Pavel Polishchuk

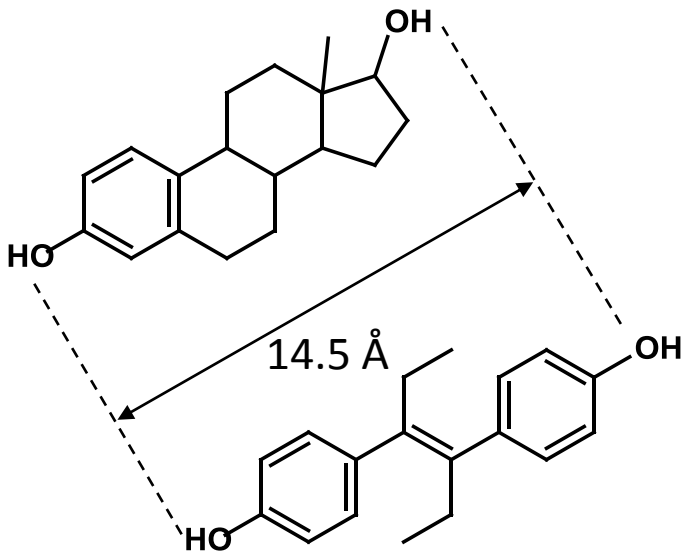
Institute of Molecular and Translational Medicine
Faculty of Medicine and Dentistry
Palacky University

pavlo.polishchuk@upol.cz

Virtual screening methods

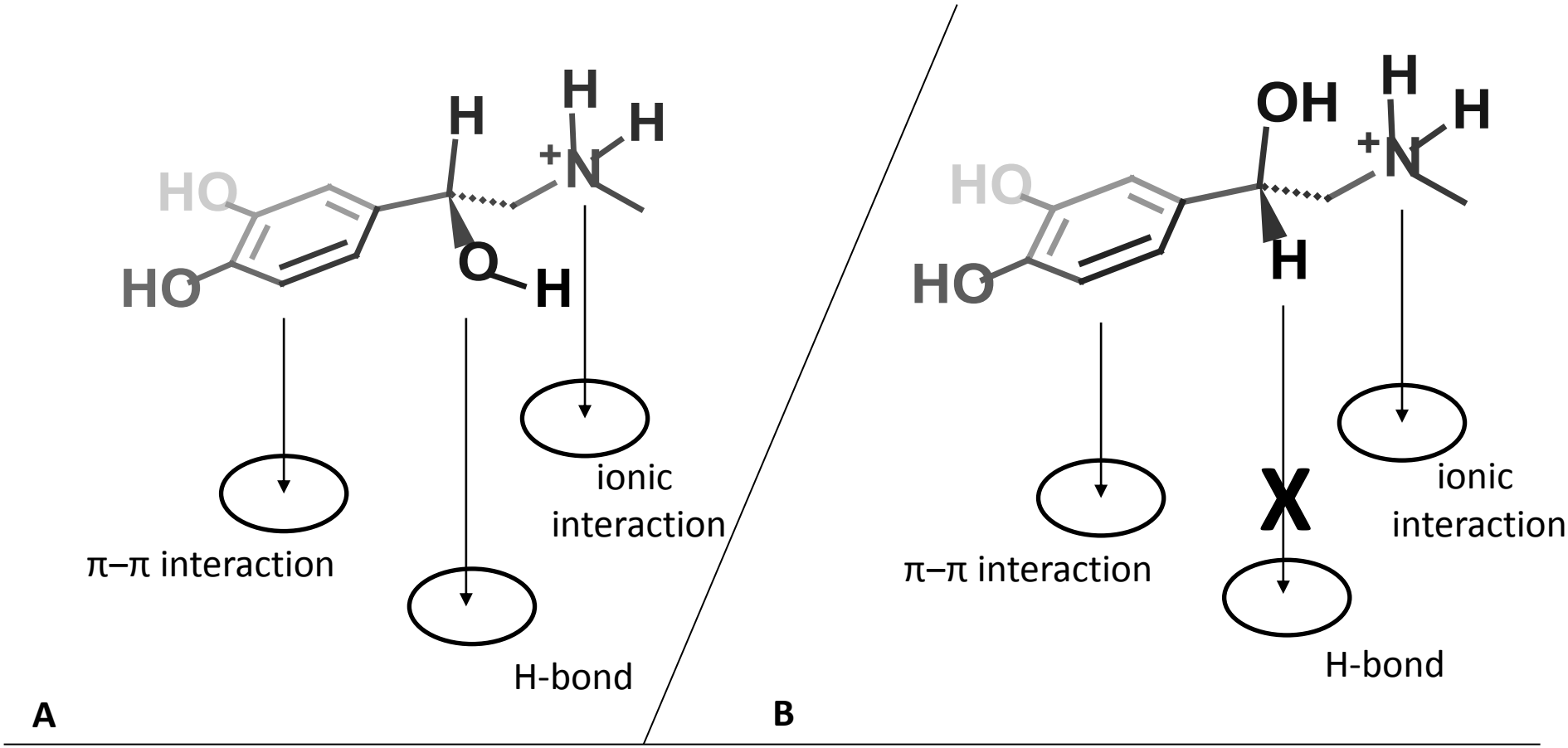


Early pharmacophore models



estradiol and *trans*-diethylsilbestrol.

Early pharmacophore models



(R)-(-) –Epinephrine
(Adrenalin)

(S)-(+) –Epinephrine

Pharmacophore definition

A **pharmacophore** is the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interaction with a specific biological target structure and to trigger (or block) its biological response.

Annu. Rep. Med. Chem. 1998, 33, 385–395

Advantages of pharmacophore models

Universal

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

Computationally Efficient

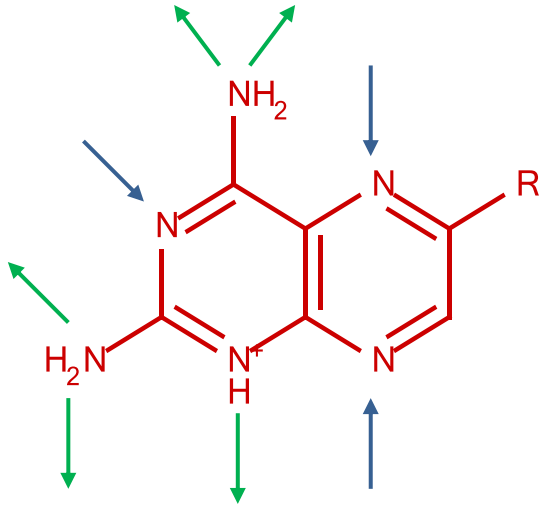
Due to their simplicity, they are suitable for large scale virtual screening (~10^8 compounds, also in parallel settings)

Comprehensive & Editable

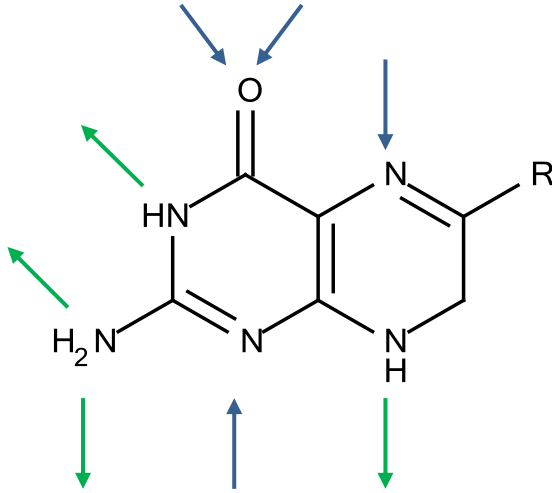
Selectivity-tuning by adding or omitting chemical feature constraints, information can be easily traced back

Atom- and pharmacophore-based alignment

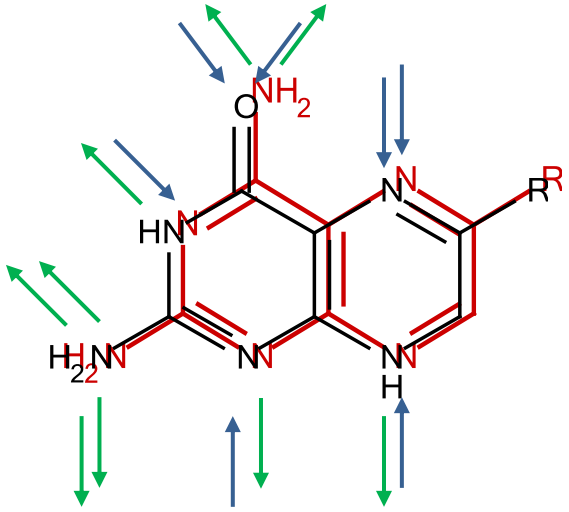
Methotrexate



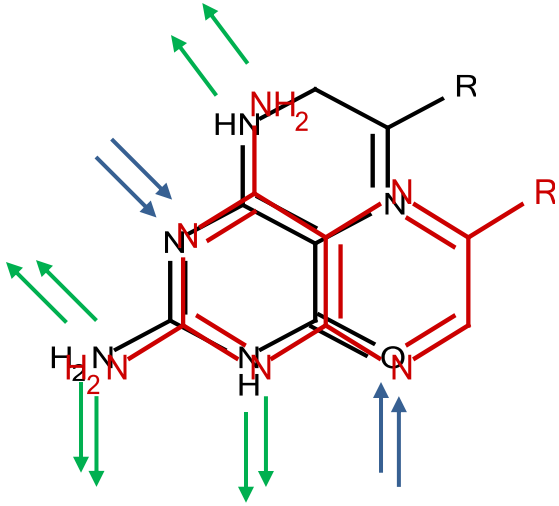
Dihydrofolate



Hydrogen bonding patterns



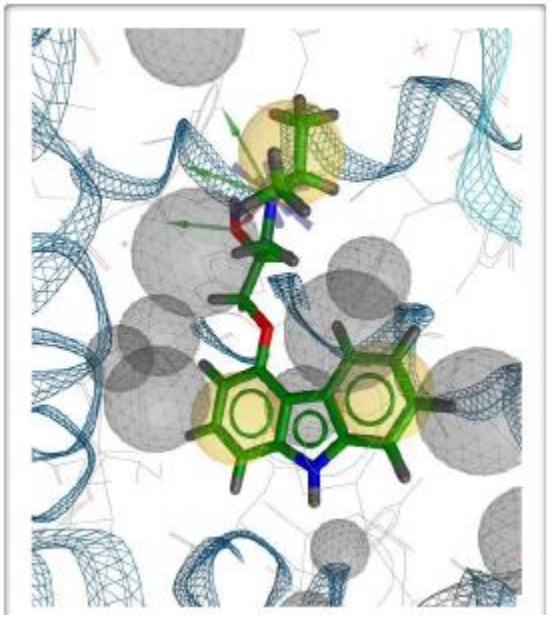
Atom-based alignment

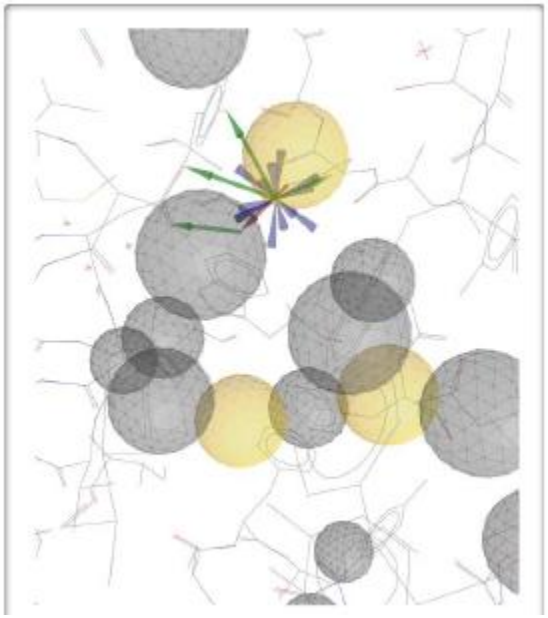


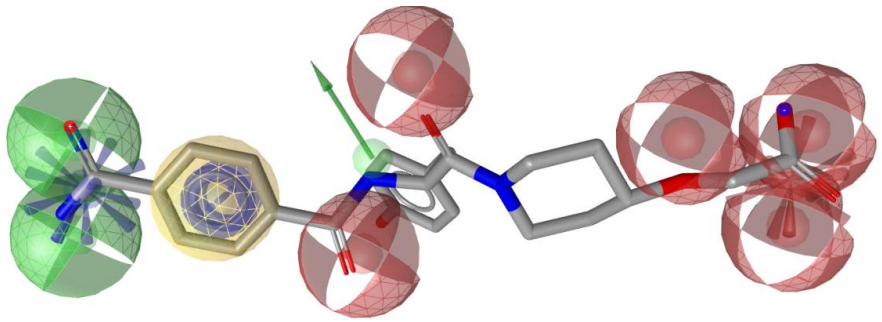
Pharmacophore alignment

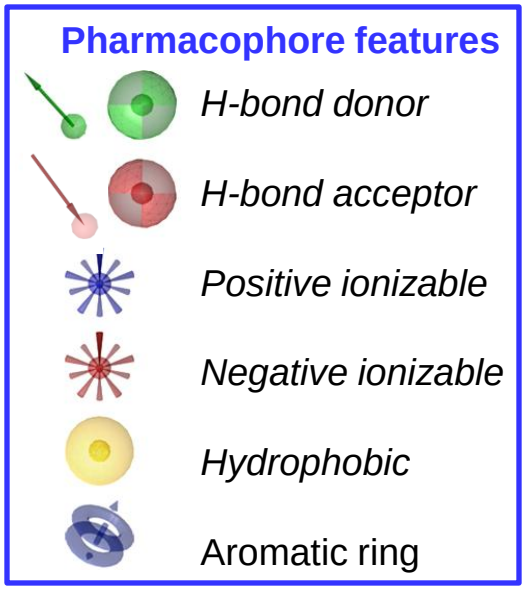
Feature-based pharmacophore models

Features: Electrostatic interactions, H-bonding, aromatic interactions, hydrophobic regions, coordination to metal ions ...









LS: Comparability vs Specificity of Chemical Features

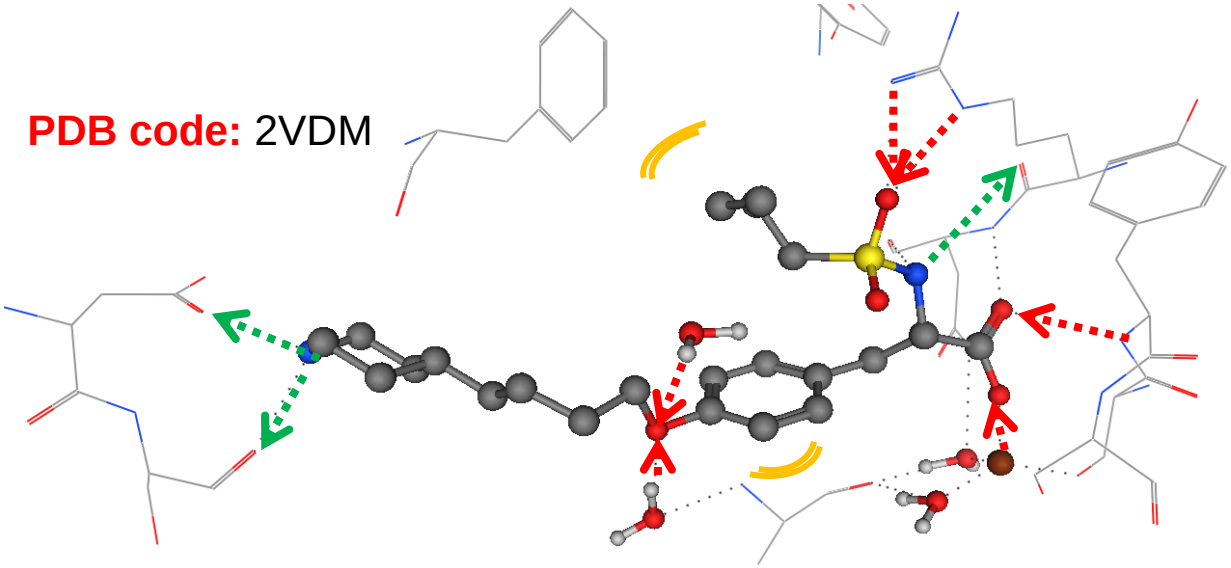
Levels	Universality	Specificity	Classification	Example
1	--	+++	Molecular graph descriptor (atom, bond) with geometric constraint	A phenol group facing a parallel benzenoid system within a distance of 2–4A
2	-	++	Molecular graph descriptor (atom, bond) without geometric constraint	A phenol group
3	++	+	Chemical functionality (hydrogen bond donor, acceptor) with geometric constraint	H-bond acceptor vector including an acceptor point as well as a projected donor point; aromatic ring including a ring plane
4	+++	-	Chemical functionality (positive ionizable area, lipophilic contact) without geometric constraint	H-bond acceptor without the projected point; lipophilic group

LigandScout SMARTS pharmacophore patterns

	Inclusion patterns	Exclusion patterns
HBA-F	{[O,S]}[#1]	c1nnnn1
	{N}[#1]	
	C{F}	
HBD	{[N,O,S;X1,X2]}	[–,–2,–3]
PI	{[NX3]}([CX4])([CX4,#1])[CX4,#1]	
	{N}=[CX3]({[N;H1,H2]})[! N]	
	N=[CX3]({[NH1]}){[NH1]}	
	{[+,+2,+3;! $(*[–,–2,–3])]}	
NI	[S,P](={O})(={O}){[OH]}	
	[S,C,P](={O}){[OH]}	
	{c}1{n}{n}{n}{n}1	
	{[–,–2,–3;! $(*[+,+2,+3])]}	

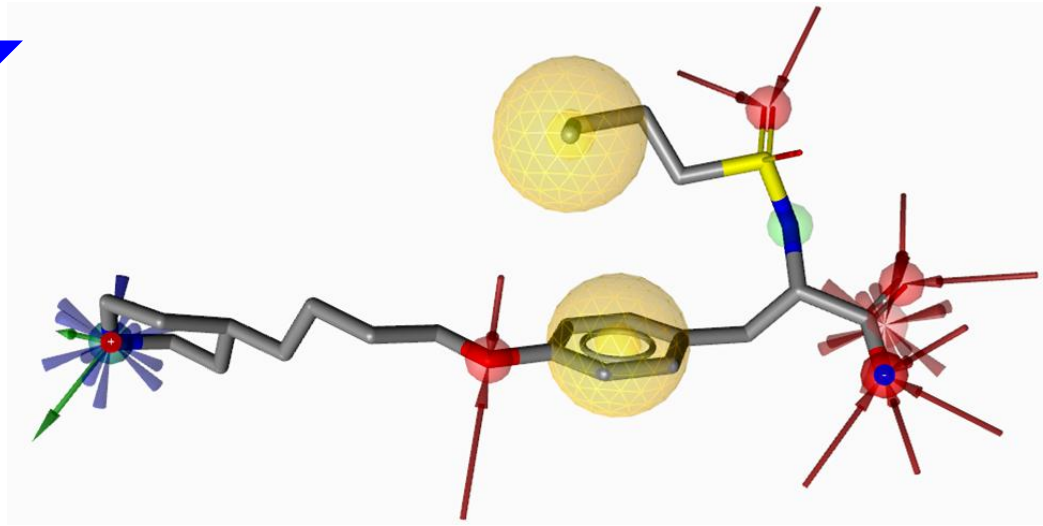
Structure-based pharmacophores

PDB code: 2VDM



H-bonds formed by the ligand

Hydrophobic interaction



Pharmacophore features

H-bond donor

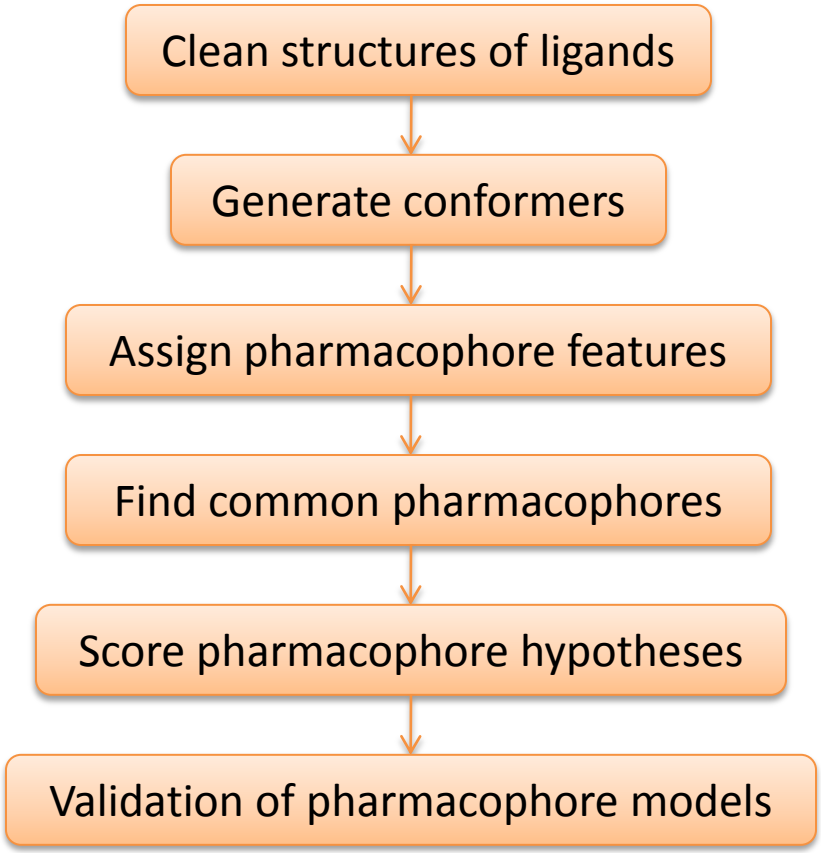
H-bond acceptor

Positive ionizable

Negative ionizable

Hydrophobic

Typical ligand-based pharmacophore modeling workflow



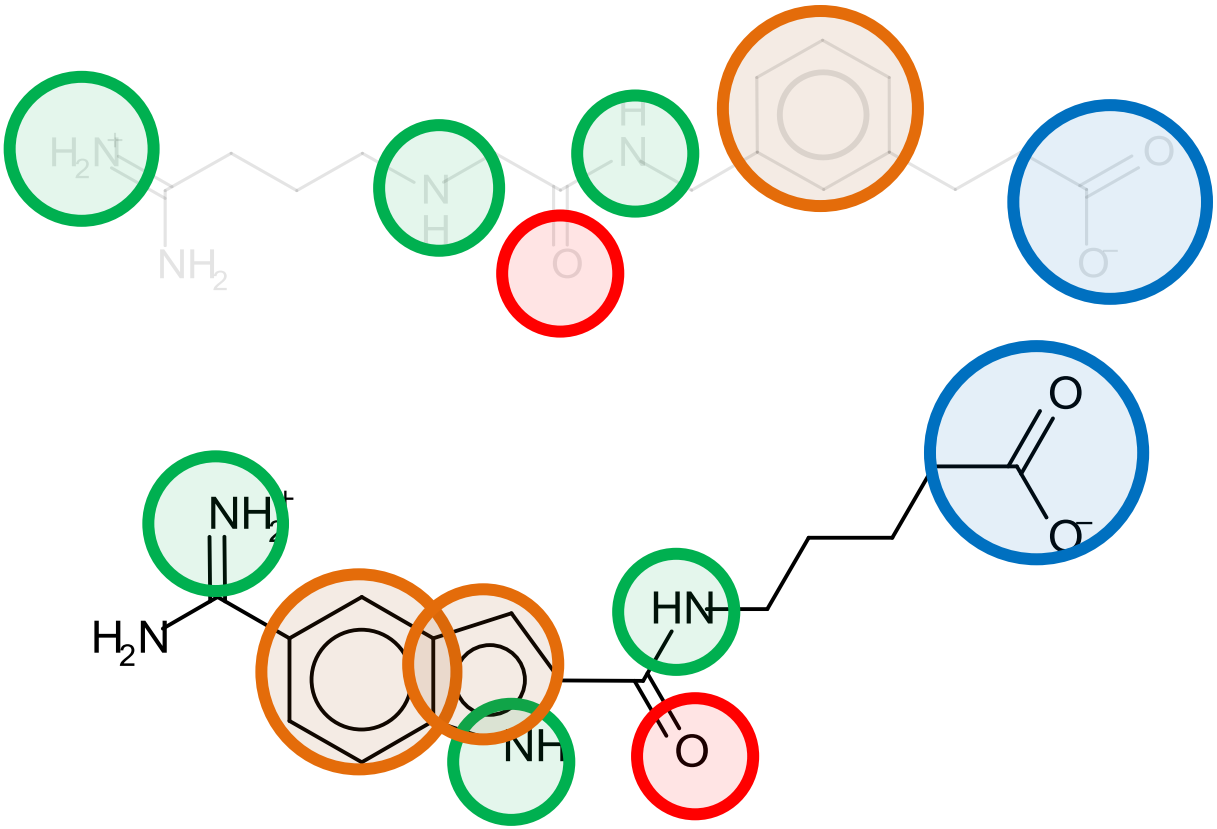
Pharmacophore software

Software	Input	Identification methods	Virtual screening capability	Free for academic use[a]
FLAP[9]	Ligand, complex, apo	Molecular field	Yes	No
Pharmer[10]	Ligand, complex	Substructure pattern, feature	Yes	Yes (GPLv2)
LigandScout[11]	Ligand, complex, apo	Substructure pattern, feature, molecular field	Yes	No
Catalyst[12]	Ligand, complex, apo	Substructure pattern, feature, molecular field	Yes	No
MOE[13]	Ligand, complex, apo	Substructure pattern, feature, molecular field	Yes	No
PHASE[14]	Ligand, complex, apo	Substructure pattern, feature, molecular field	Yes	No
Pharao[15]	Ligand	Substructure pattern	Yes	Yes (GPLv2)
UNITY[16]	Ligand, complex	Substructure pattern, feature	Yes	No
Forge[17]	Ligand	Molecular field	Yes	Free for PhD students

Program	Scoring method	Provider	Program/web-tool
DiscoveryStudio	Overlay	Biovia (formerly Accelrys)	Program
LigandScout	Overlay	Inte:Ligand	Program
MOE	RMSD	Chemical Computing Group	Program
PHASE	RMSD	Schrödinger	Program
GASP	Overlay	Tripos	Program
DISCOTech	RMSD	Tripos	Program
Pharmer	RMSD	Camacho Lab	Program code free
PharmaGist	Overlay	Tel Aviv University	Web tool
QUASI	Overlay	DeNovo Pharmaceuticals	Program
AnchorQuery[a]	RMSD	Cacho and Dömling laboratories	Web-tool
ROCS[b]	Overlay	OpenEye	Program
USR[b]	Overlay	Istar	Web-tool

1) Schaller, D.; et al. Next generation 3D pharmacophore modeling. *WIREs Computational Molecular Science* **2020**, 10 (4), e1468.
2(Vuorinen, A.; Schuster, D., Methods for generating and applying pharmacophore models as virtual screening filters and for bioactivity profiling. *Methods* **2015**, 71, 113-134.

Common features finding

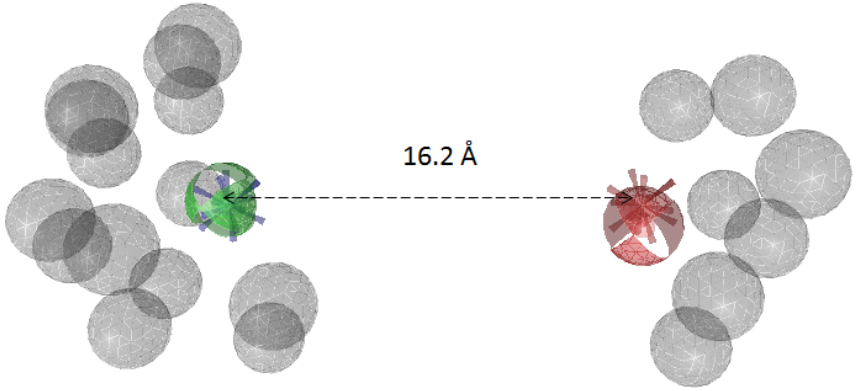


Pharmacophore features

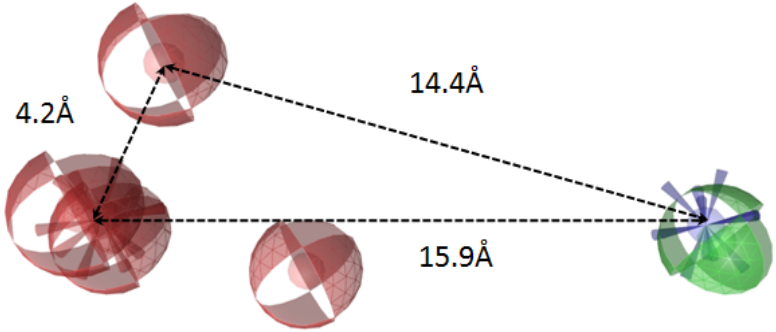
- *H-bond donor*
- *H-bond acceptor*
- *Negative ionizable*
- *Hydrophobic*

Ligand-based pharmacophore example

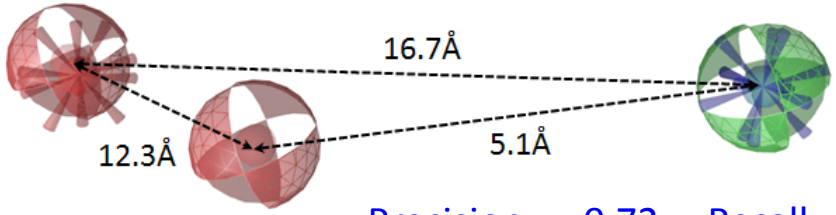
Shared model on 83 antagonists of fibrinogen receptor



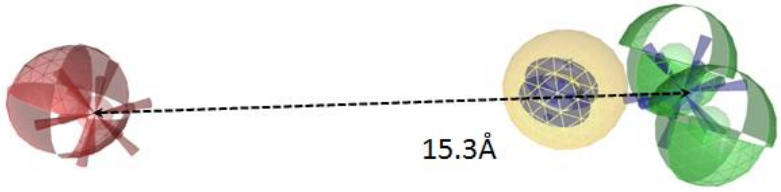
Pharmacophore models obtained for clusters of compounds



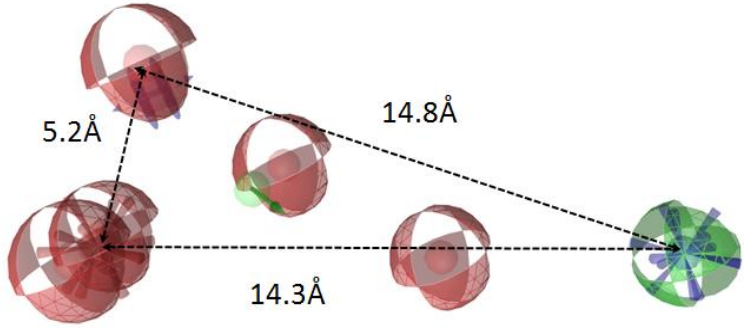
Precision: 0.77 Recall: 0.27



Precision: 0.72 Recall: 0.77



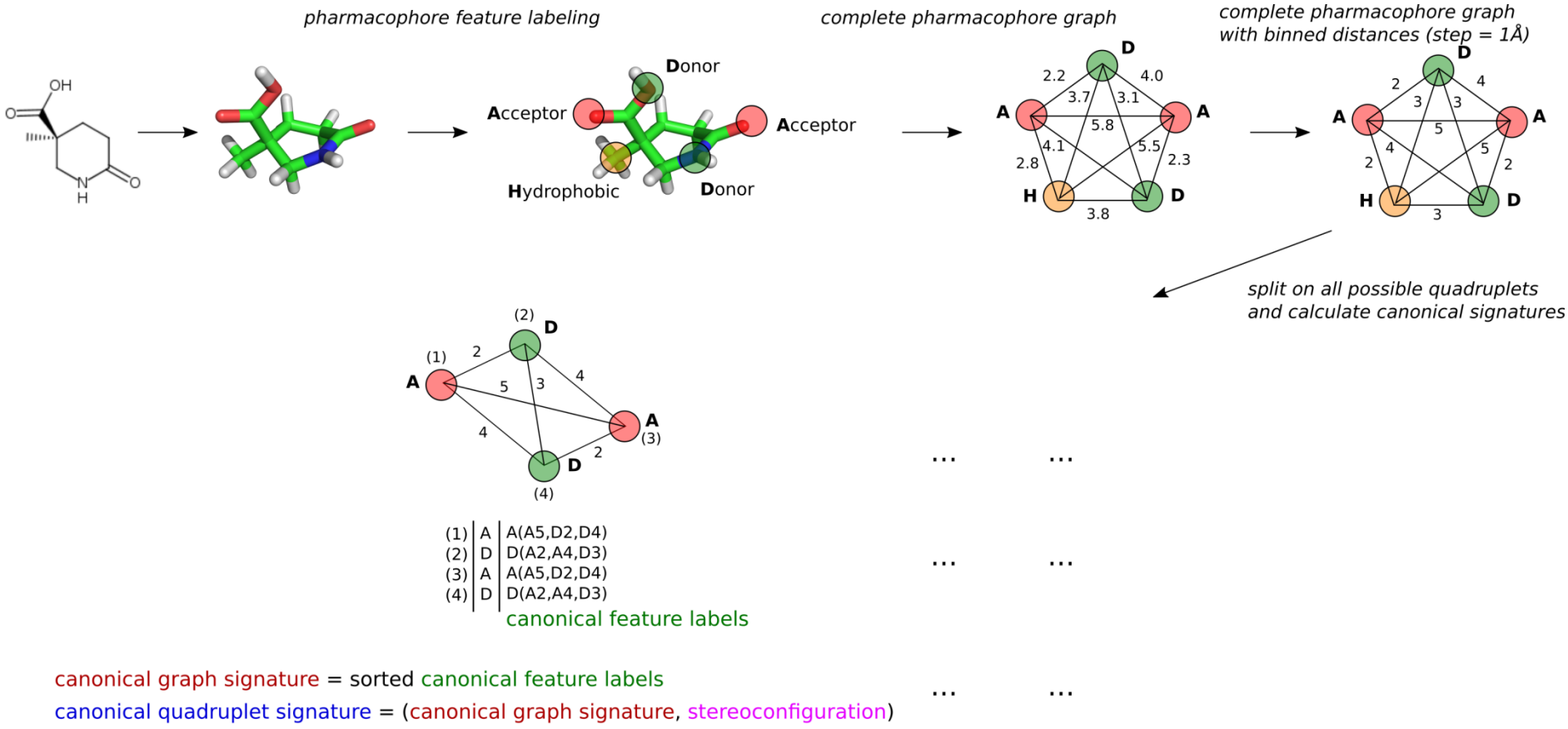
Precision: 0.67 Recall: 0.29



Precision: 0.90 Recall: 0.04

Polishchuk, P. G. et al., *Journal of Medicinal Chemistry* **2015**, 58, 7681-7694.

Pmapper: 3D pharmacophore descriptors

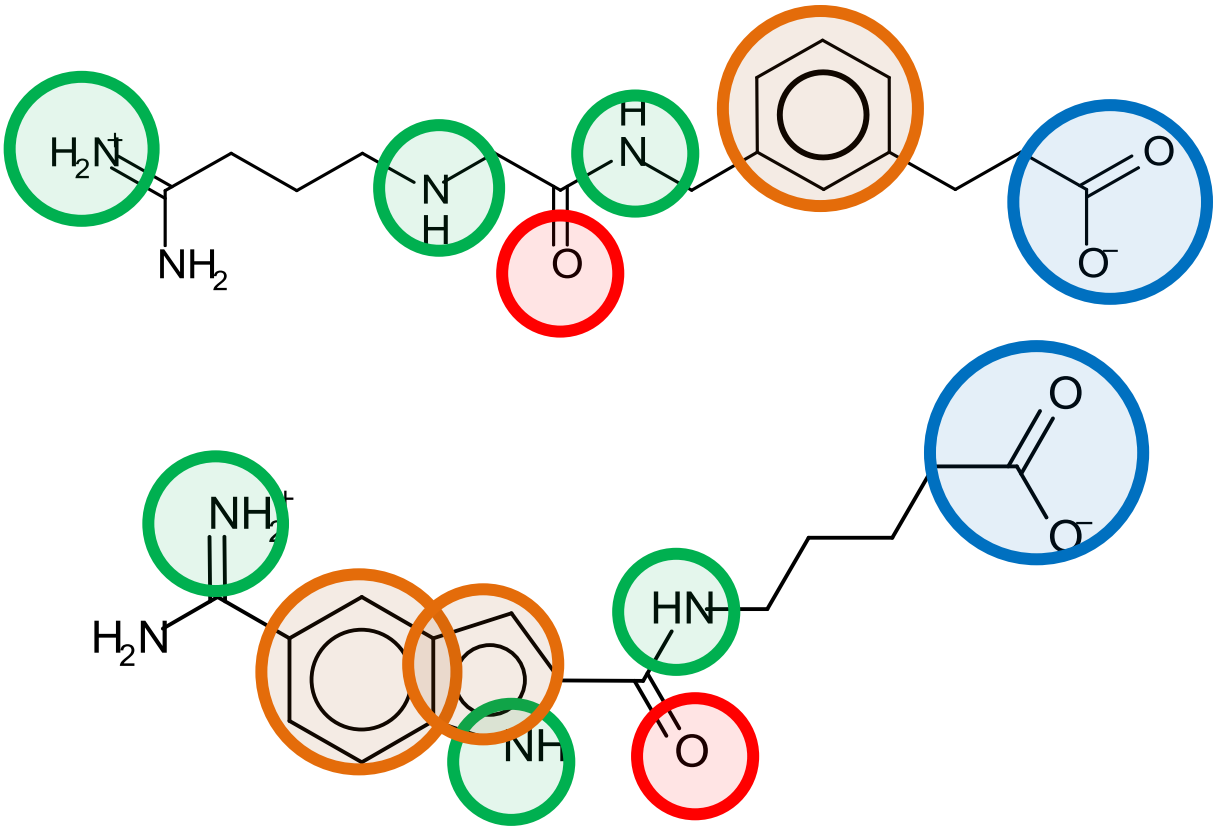


https://github.com/DrrDom/pmapper

Kutlushina, A. et al., Ligand-Based Pharmacophore Modeling Using Novel 3D Pharmacophore Signatures. *Molecules* **2018**, 23, 3094.

Common features finding

alignment-free

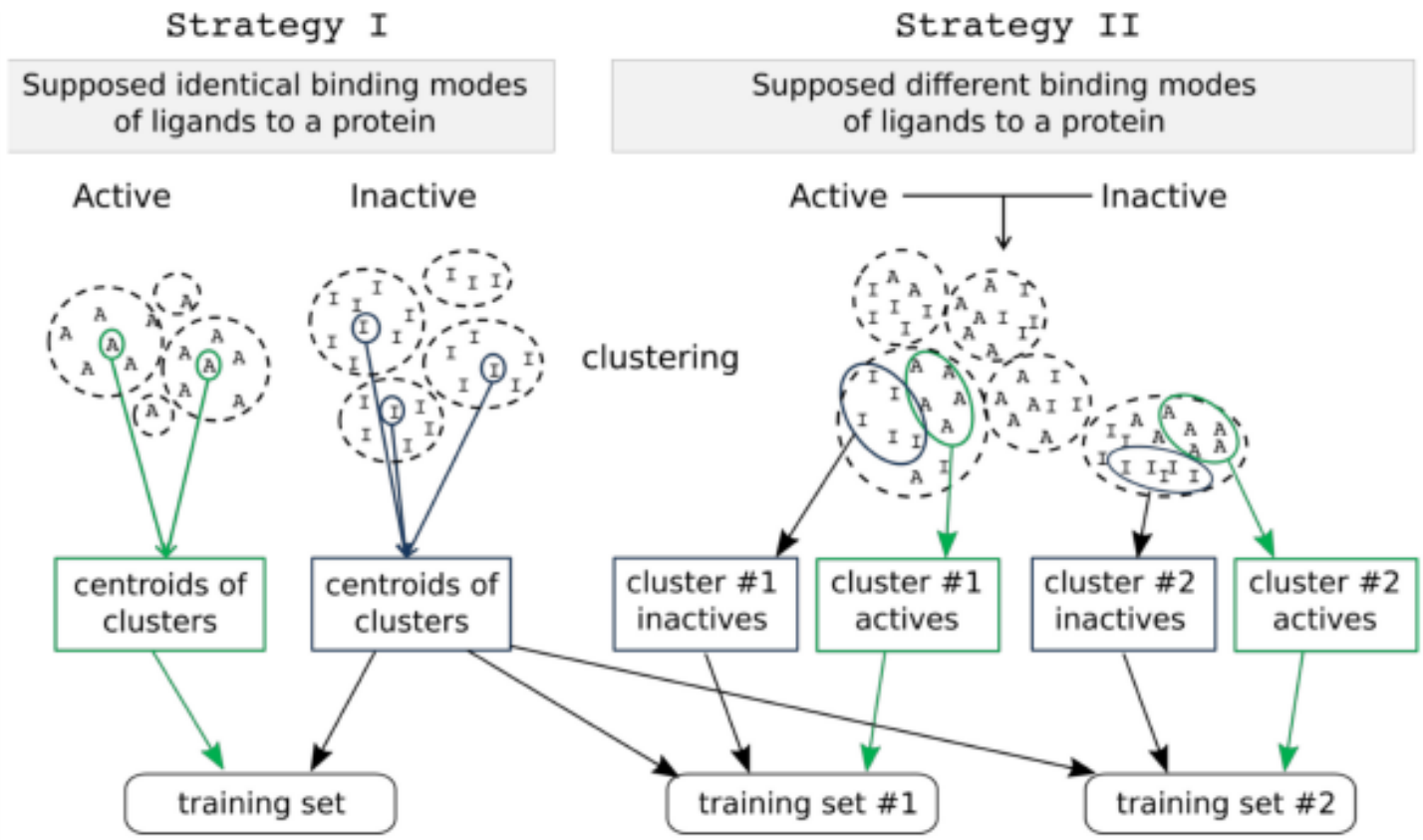


	3D pharmacophore hash
	gd6fdfh90a...
3-point	89kfnnmsao...
	...
4-point	kbxlpsh692...
	...
5-point	07dkfhksqz...
	...
N-point	...

	3D pharmacophore hash
	gd6fdfh90a...
3-point	jkspl789n1...
	...
4-point	9ondbnmq12...
	...
5-point	pmne91ml1q...
	...
N-point	...

Kutlushina, A. et al, Ligand-Based Pharmacophore Modeling Using Novel 3D Pharmacophore Signatures. *Molecules* **2018**, 23, 3094.

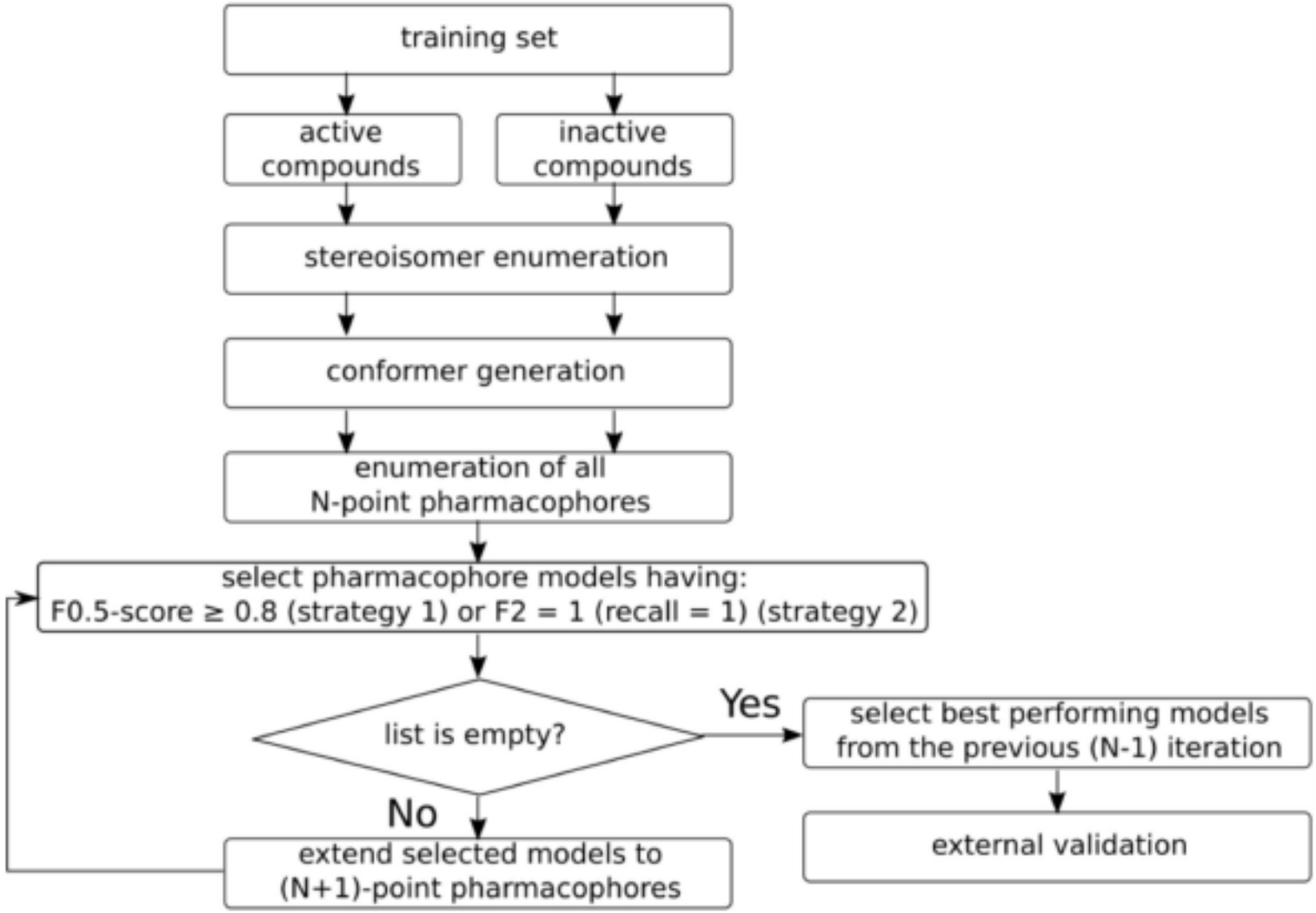
Alignment-free ligand-based pharmacophore modeling



https://github.com/meddwl/psearch

Kutlushina, A.; Khakimova, A.; Madzhidov, T.; Polishchuk, P., Ligand-Based Pharmacophore Modeling Using Novel 3D Pharmacophore Signatures. *Molecules* **2018**, 23 (12), 3094.

Alignment-free ligand-based pharmacophore modeling

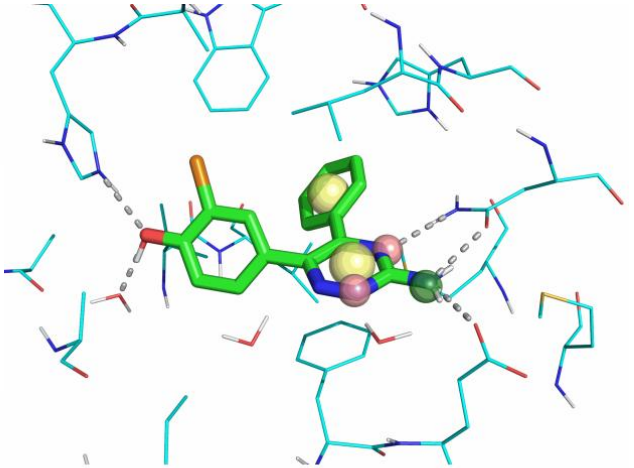


https://github.com/meddwl/psearch

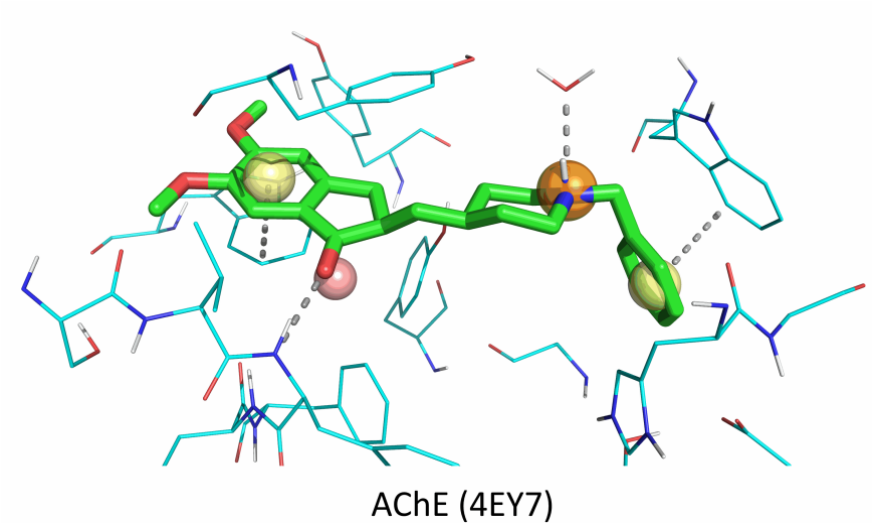
Kutlushina, A.; Khakimova, A.; Madzhidov, T.; Polishchuk, P., Ligand-Based Pharmacophore Modeling Using Novel 3D Pharmacophore Signatures. *Molecules* **2018**, 23 (12), 3094.

Alignment-free ligand-based pharmacophore modeling

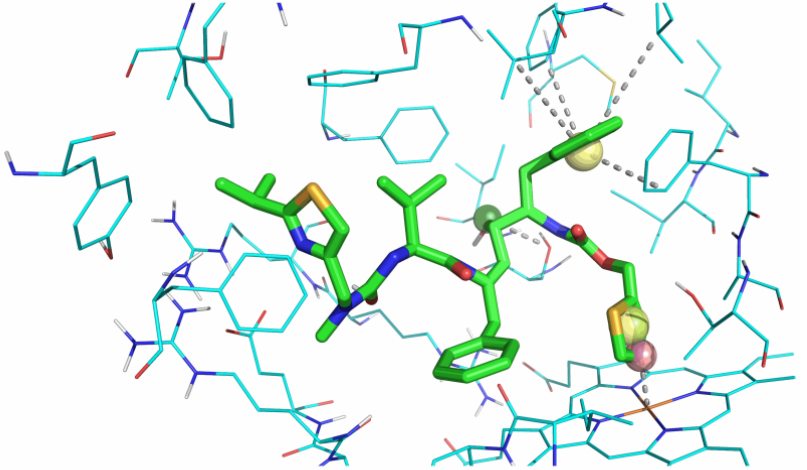
Data Set	Number of Actives	Number of Inactives	Total Number of Compounds
AChE	176 ($pIC_{50} \geq 8$)	1070 ($pIC_{50} \leq 6$)	1246
CYP450 3A4	138 ($pIC_{50} \geq 7$)	548 ($pIC_{50} \leq 5$)	686
A2a	293 ($pKi/pKd/pIC_{50} \geq 7$)	279 ($pKi/pKd/pIC_{50} \leq 5$)	574



A2a (5OLZ)



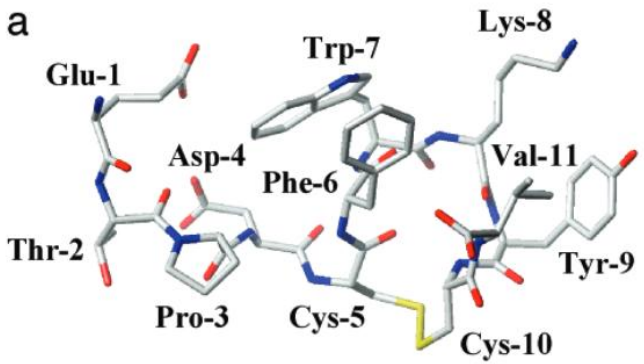
AChE (4EY7)



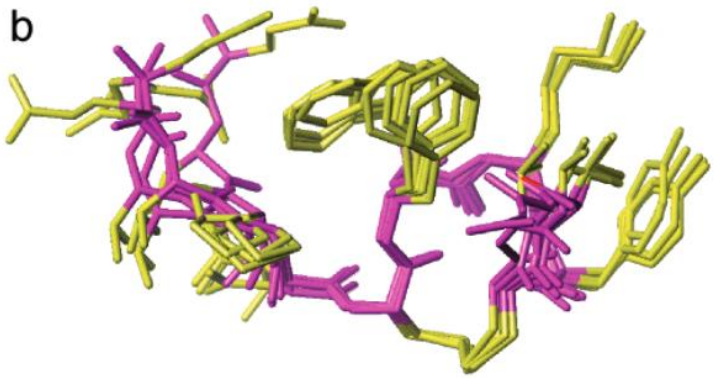
CYP450 3A4 (3NXU)

Kutlushina, A. et al, Ligand-Based Pharmacophore Modeling Using Novel 3D Pharmacophore Signatures. *Molecules* **2018**, 23, 3094.

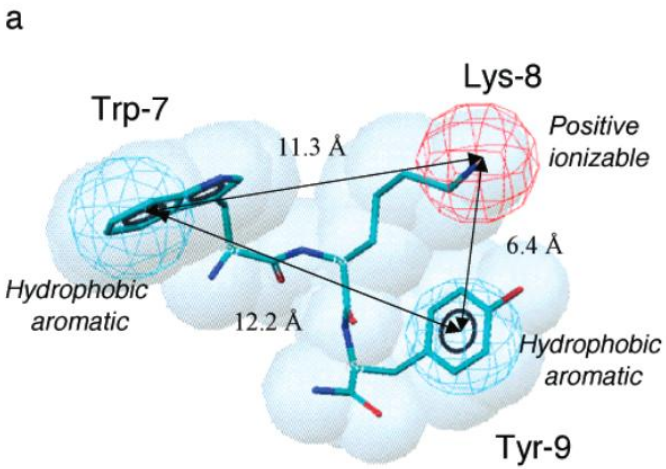
Structure-based & MD pharmacophore example



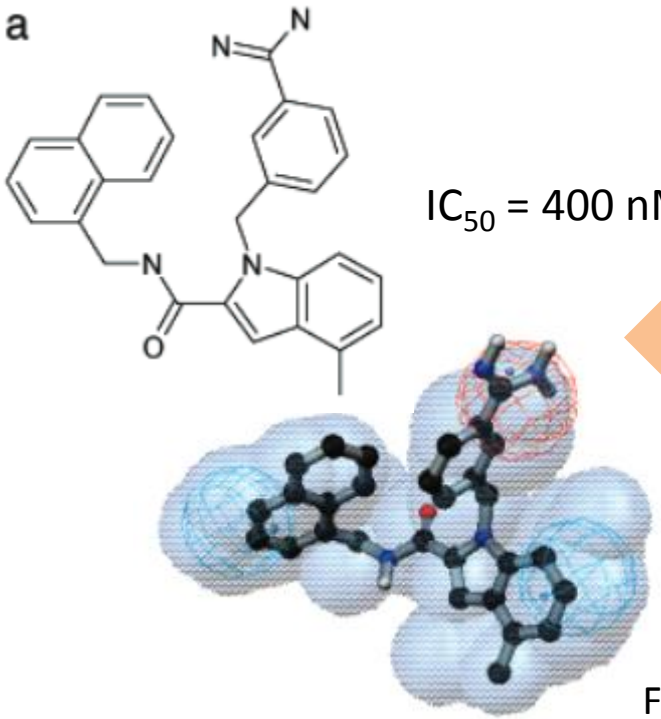
Ala scan
NMR
MD



Urotensin II - ETPDc[CFWKYCV]
potent vasoconstrictor



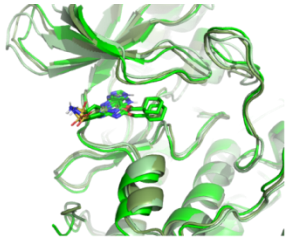
500 hits

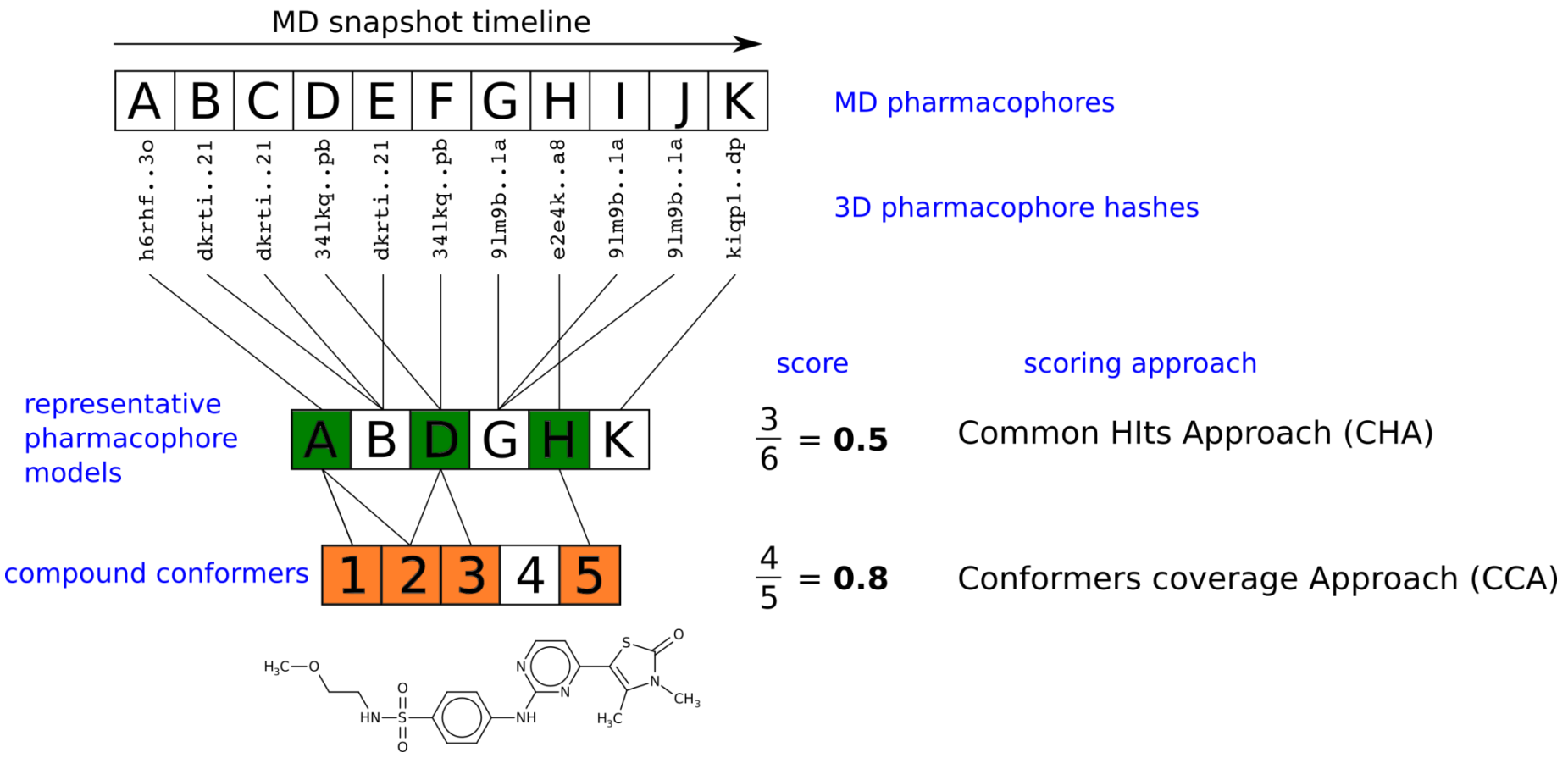


IC_{50} = 400 nM

Flohr S. et al., J. Med. Chem., **2002**, 45 (9), pp 1799–1805

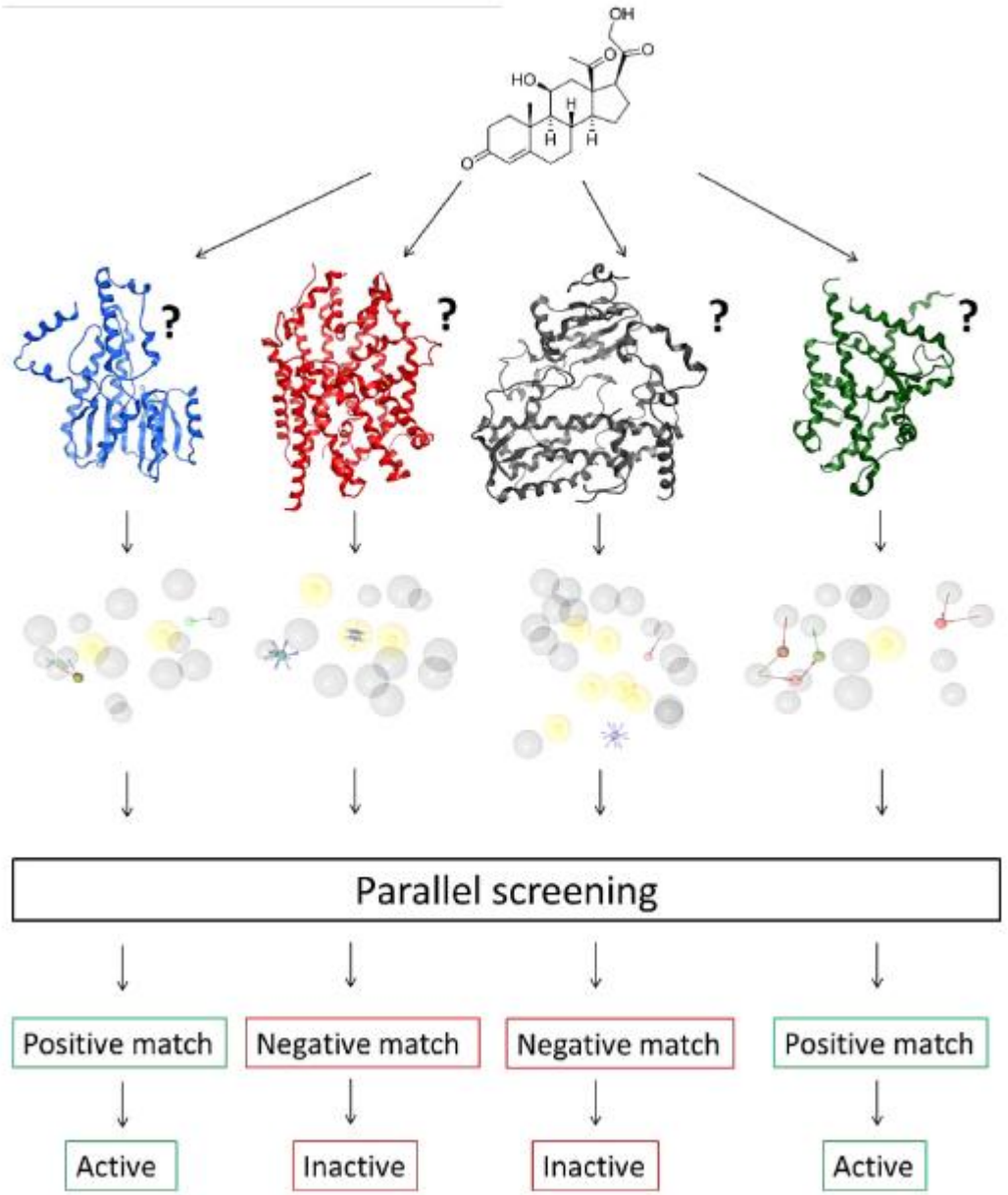
MD pharmacophores





Polishchuk, P. et al. Virtual Screening Using Pharmacophore Models Retrieved from Molecular Dynamic Simulations. *International Journal of Molecular Sciences* **2019**, 20, (23), 5834.

Pharmacophore: ligand profiling

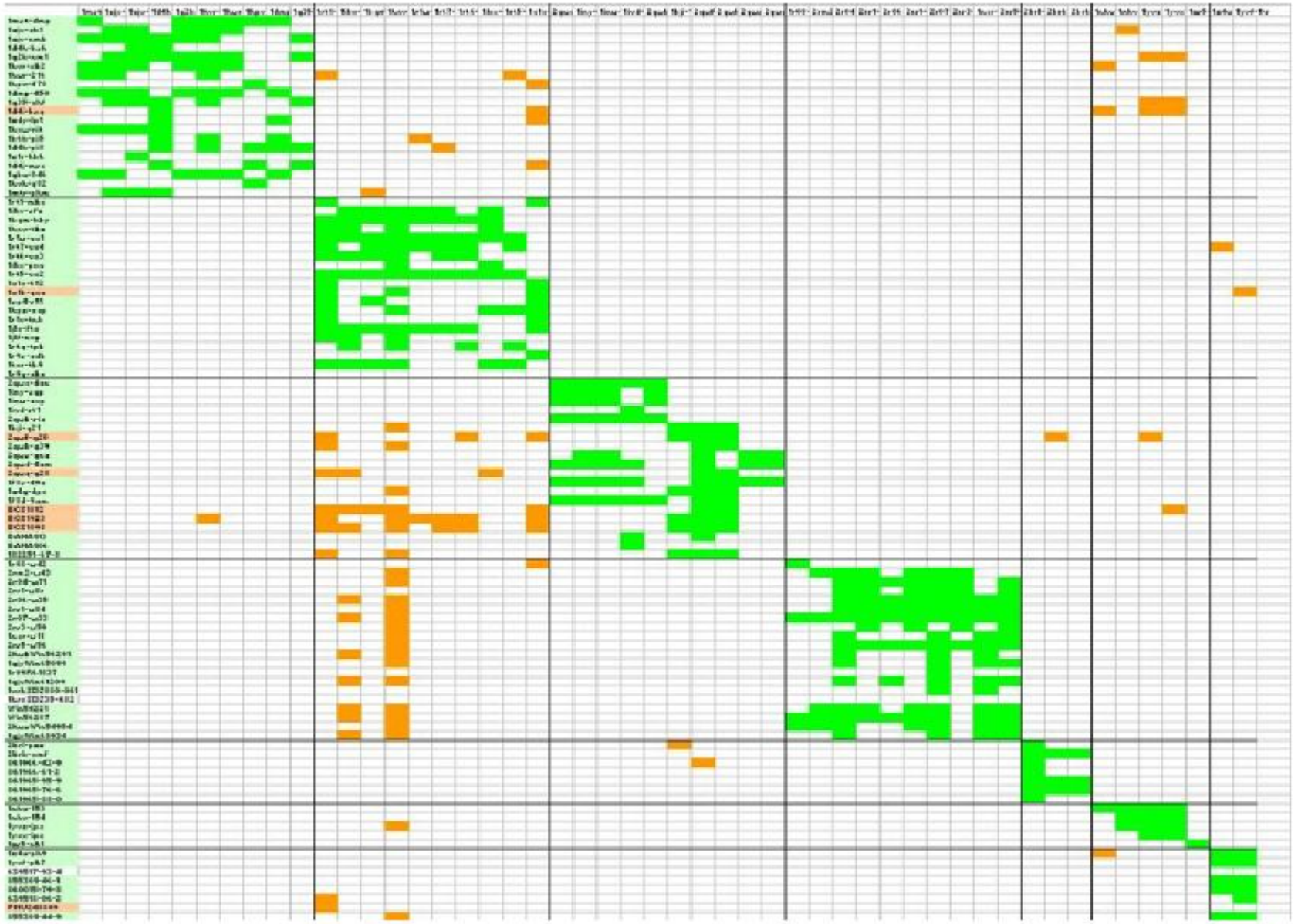


Vuorinen, A.; Schuster, D., Methods for generating and applying pharmacophore models as virtual screening filters and for bioactivity profiling. *Methods* **2015**, 71, 113-134.

Pharmacophore: ligand profiling

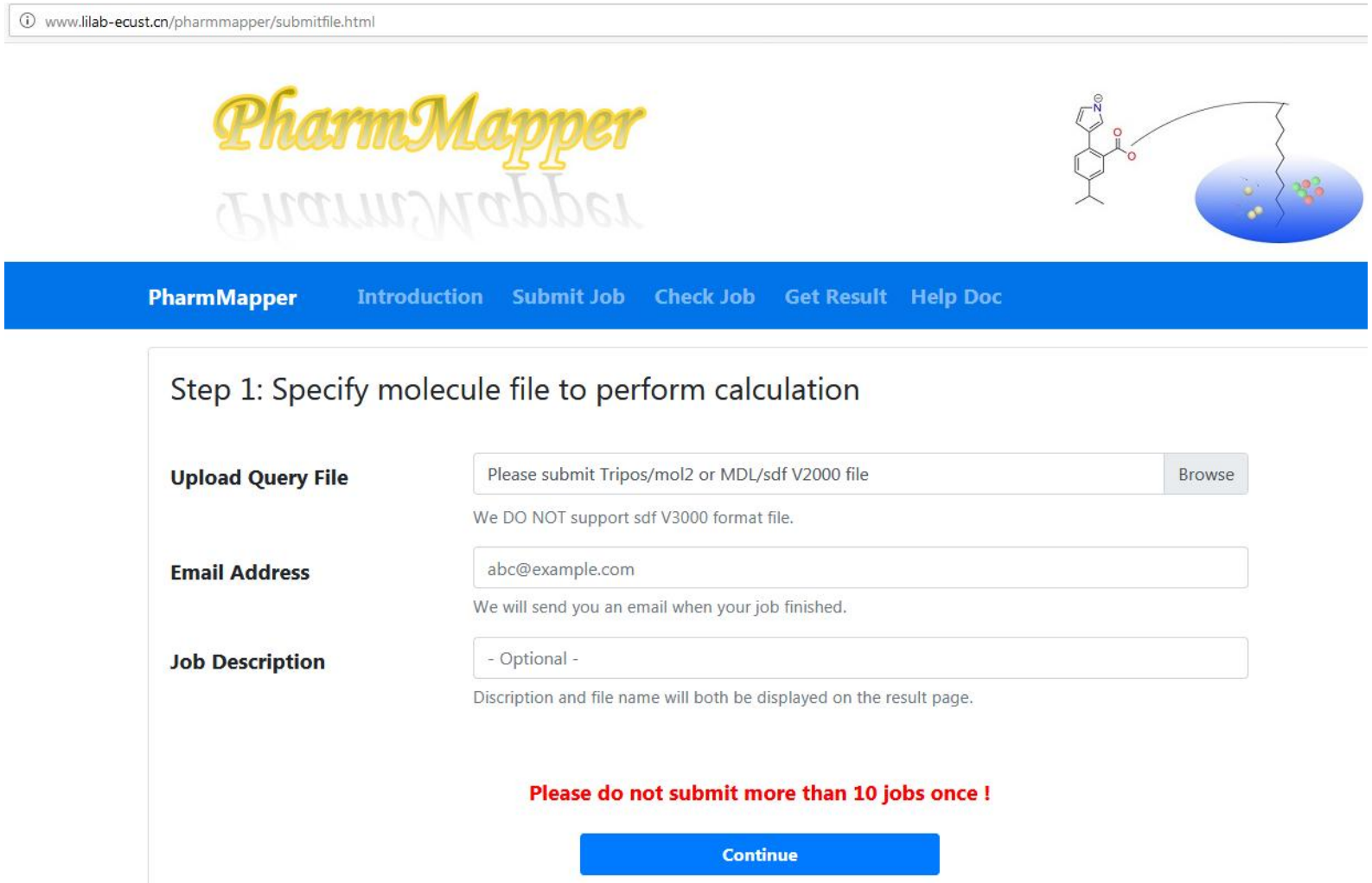
Pharmacophore models

Ligands



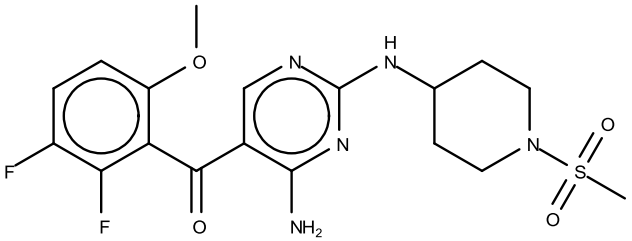
T. Steindl et al., *J. Chem. Inf. Model.*, 46, 2146-2157 (2006)

PharmMapper



Xia Wang et al, PharmMapper 2017 update: a web server for potential drug target identification with a comprehensive target pharmacophore database. *Nucleic Acids Res.*, **2017**, 45, W356-W360.

PharmMapper



2FVD

Result of 211114173050

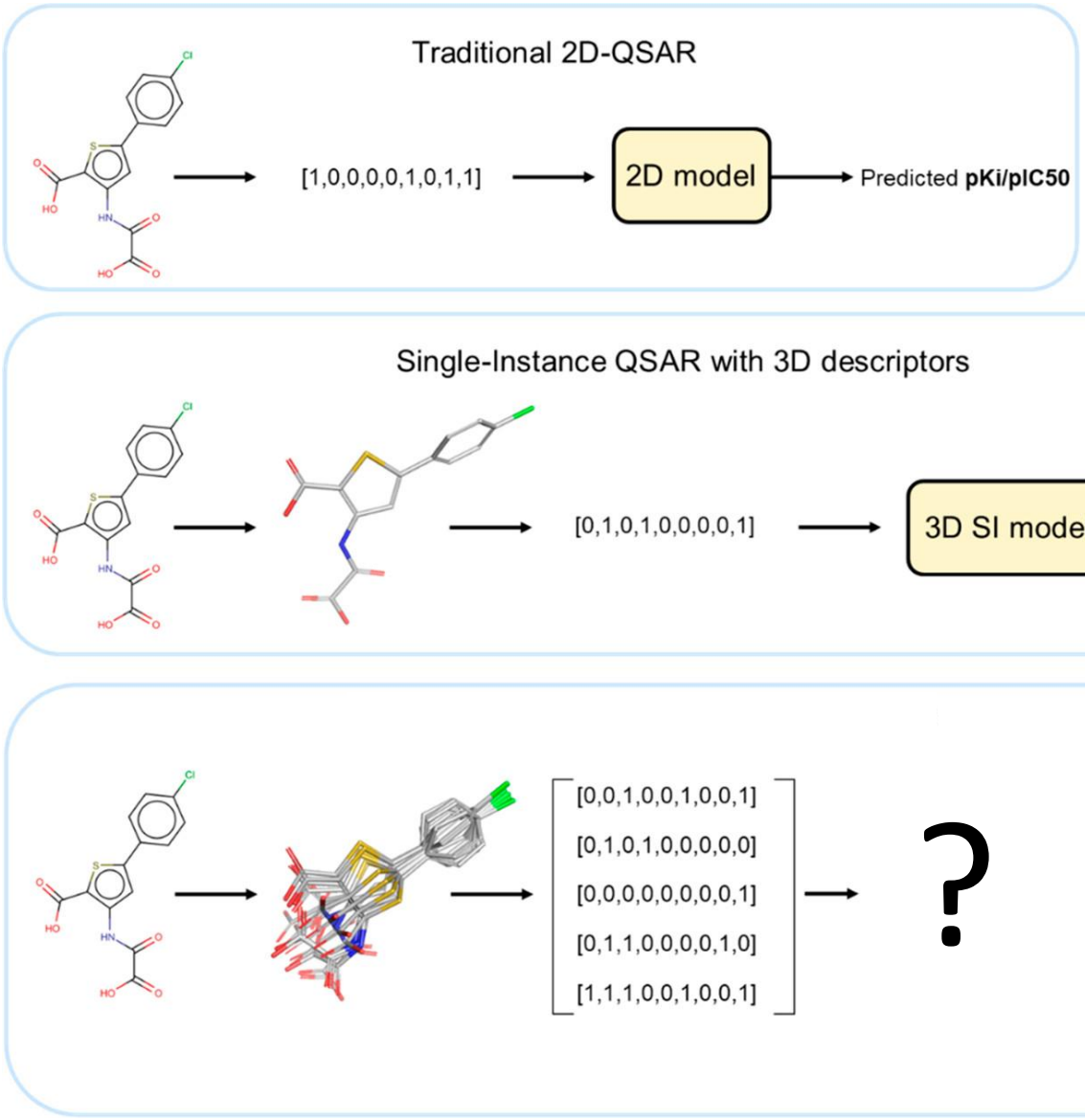
Top 300 targets ranked by normalized fit score in descending order

1.mol2 -

Ligand: LIA					
Rank	PDB ID	Target Name	Number of Features	Fit Score	Normalized Fit Score
+ 1	3MAH	NONE	3	2.998	0.9994
+ 2	2CT7	RING finger protein 31	3	2.998	0.9993
+ 3	1T0T	UPF0447 protein GK3416	3	2.998	0.9993
+ 4	1I4W	Mitochondrial replication protein MTF1	3	2.995	0.9984
+ 5	1EVY	Glycerol-3-phosphate dehydrogenase [NAD+], glycosomal	3	2.99	0.9967
+ 6	2KDD	Borealin	3	2.99	0.9967
+ 7	1XPP	DNA-directed RNA polymerase subunit L	3	2.981	0.9938

Xia Wang et al, PharmMapper 2017 update: a web server for potential drug target identification with a comprehensive target pharmacophore database. *Nucleic Acids Res.*, **2017**, 45, W356-W360.

4D QSAR



Zankov, D. V. et al, QSAR Modeling Based on Conformation Ensembles Using a Multi-Instance Learning Approach. *Journal of Chemical Information and Modeling* **2021**, 61, 4913-4923.

Multi-instance learning

ELSEVIER

Artificial Intelligence 89 (1997) 31–71

Artificial Intelligence

Solving the multiple instance problem with axis-parallel rectangles

Thomas G. Dietterich [a,*], Richard H. Lathrop [b], Tomás Lozano-Pérez [c,d]

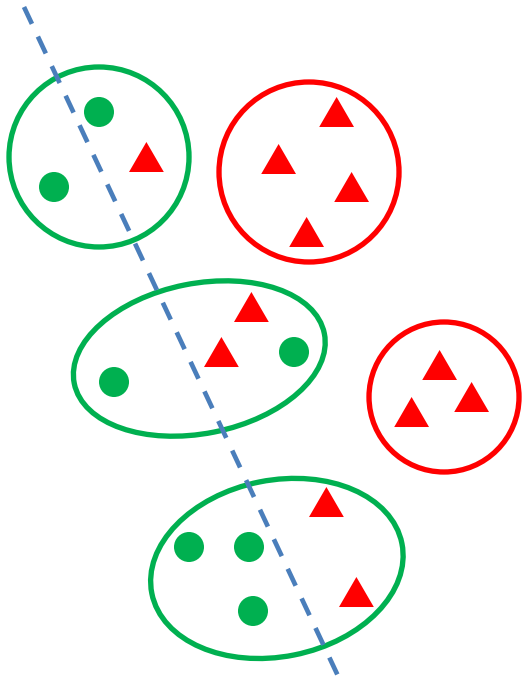
[a] *Department of Computer Science, Oregon State University, Dearborn Hall 303, Corvallis, OR 97331-3202, USA*

[b] *Department of Information and Computer Science, University of California, Irvine, CA 92697, USA*

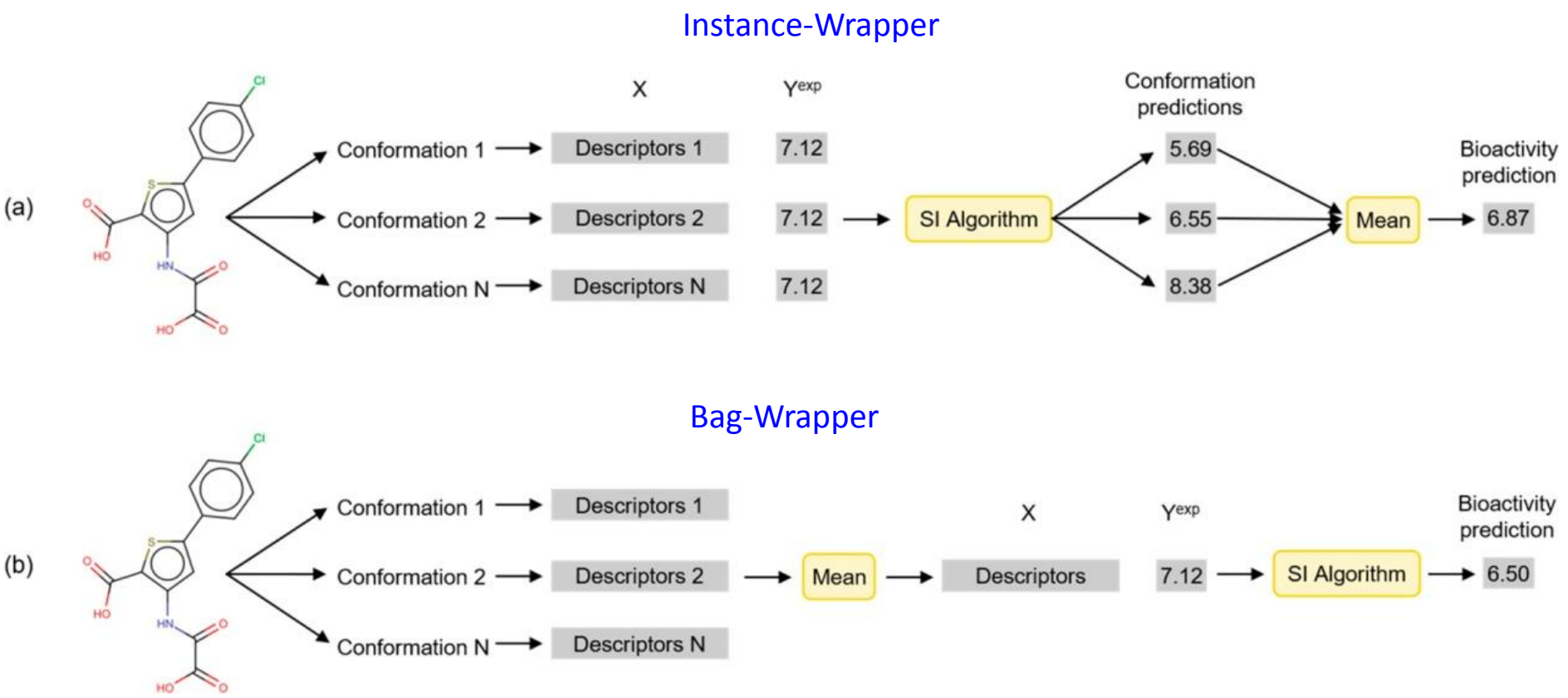
[c] *Arris Pharmaceutical Corporation, 385 Oyster Pt. Blvd., South San Francisco, CA 94080, USA*

[d] *MIT Artificial Intelligence Laboratory, 545 Technology Square, Cambridge, MA 02139, USA*

Received August 1994; revised July 1996



4D QSAR



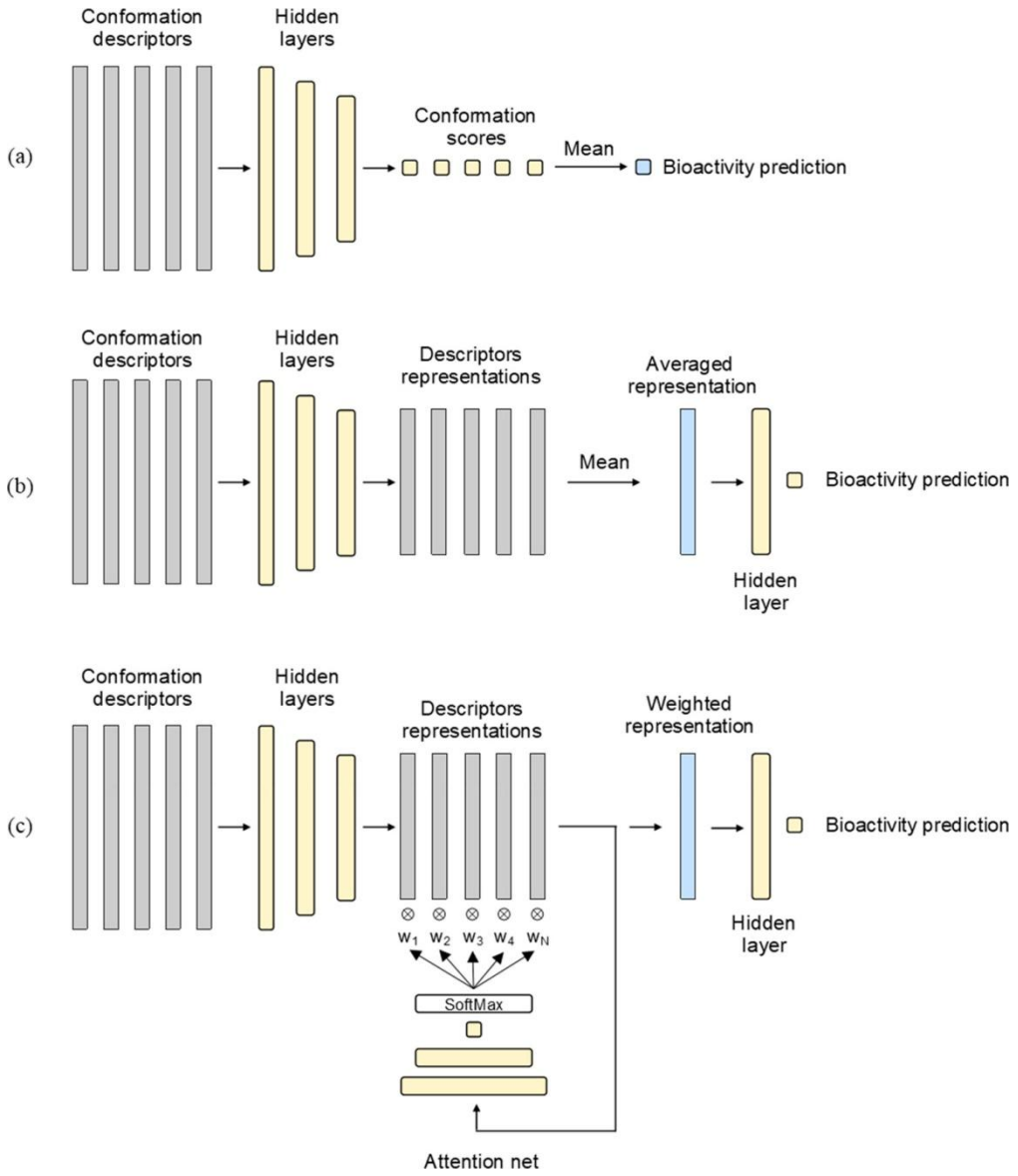
Zankov, D. V. et al, QSAR Modeling Based on Conformation Ensembles Using a Multi-Instance Learning Approach. *Journal of Chemical Information and Modeling* **2021**, 61, 4913-4923.

Multiple-instance QSAR

Instance-Net

Bag-Net

Bag-AttentionNet



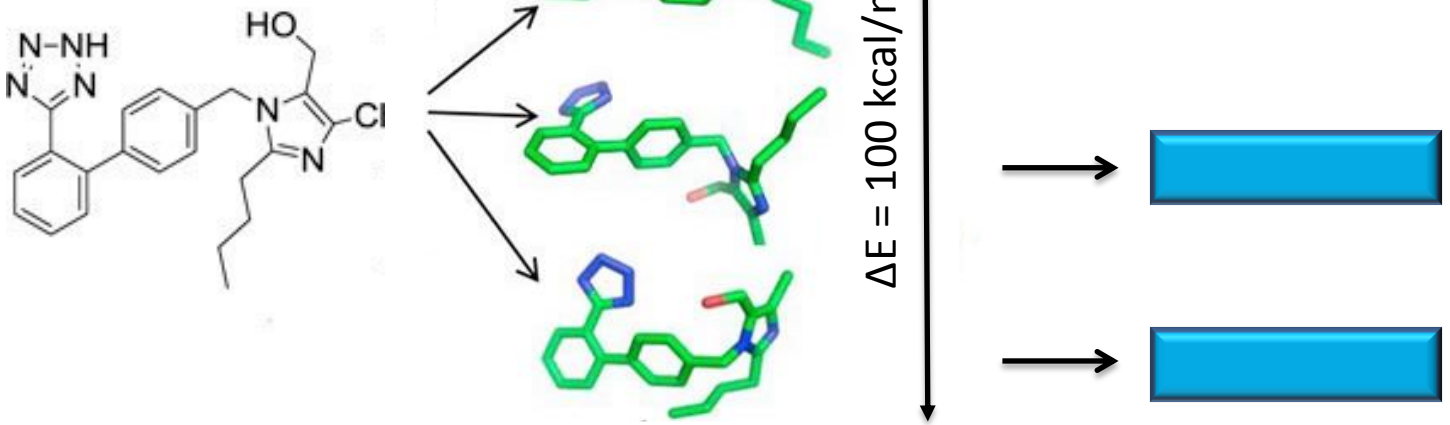
Zankov, D. V. et al, QSAR Modeling Based on Conformation Ensembles Using a Multi-Instance Learning Approach. *Journal of Chemical Information and Modeling* **2021**, 61, 4913-4923.

MIL study: conformer and descriptor generation

175 data sets from ChEMBL

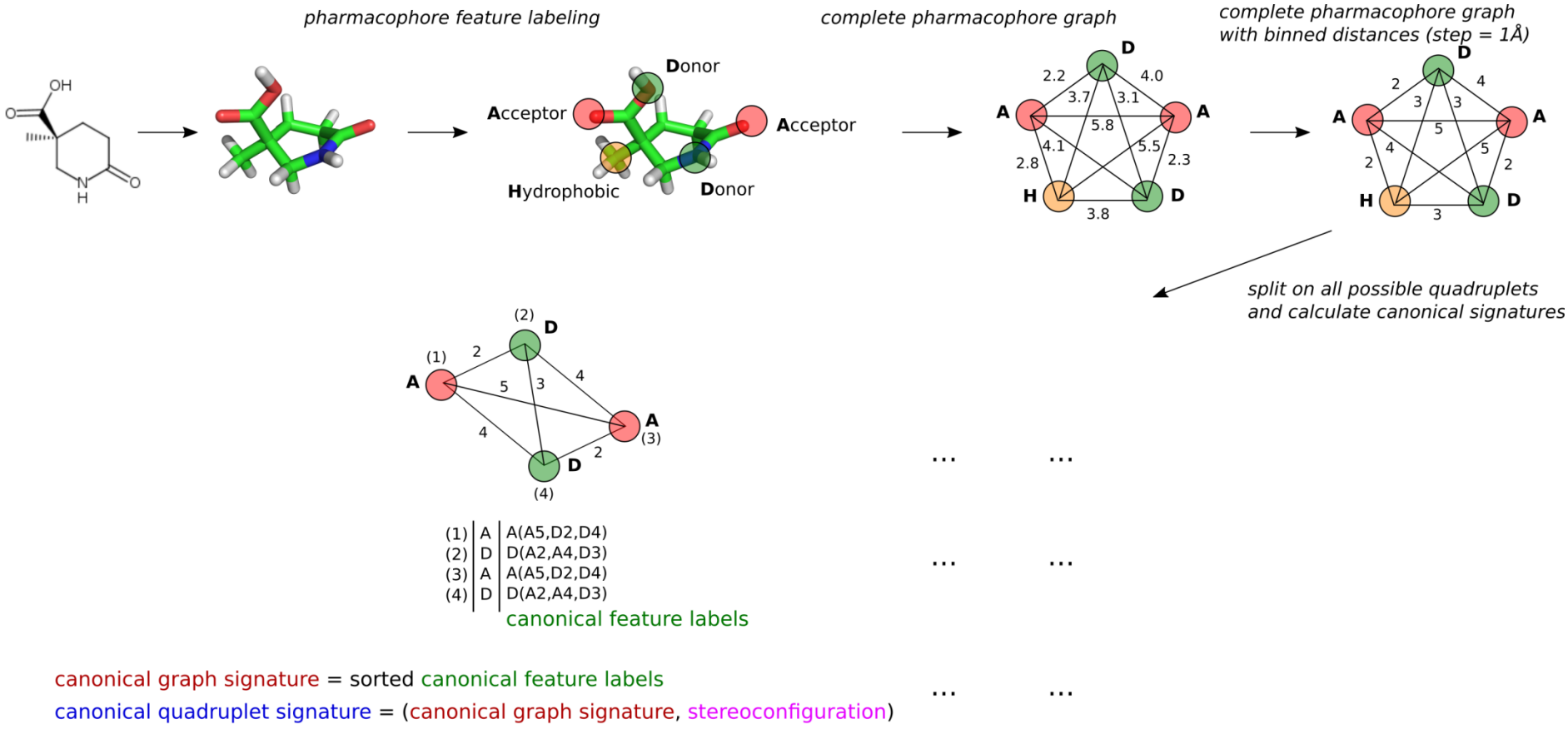
Generation of
up to **100**
conformations per
molecule
RDKit

Calculation of
Pmapper 3D descriptors



Kutlushina, A. et al. Ligand-Based Pharmacophore Modeling Using Novel 3D Pharmacophore Signatures. *Molecules* **2018**, 23, 3094.

Pmapper: 3D pharmacophore descriptors



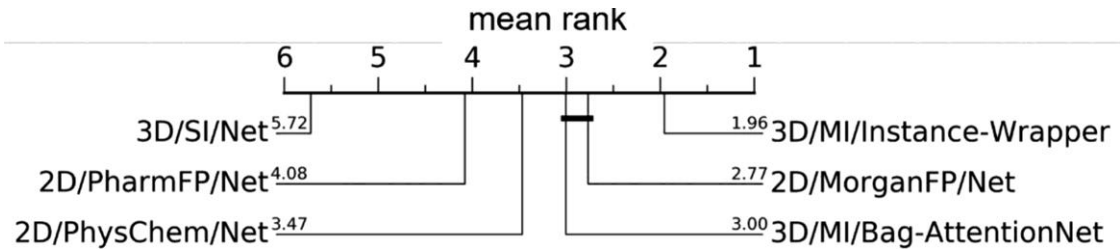
https://github.com/DrrDom/pmapper

Kutlushina, A. et al., Ligand-Based Pharmacophore Modeling Using Novel 3D Pharmacophore Signatures. *Molecules* **2018**, 23, 3094.

Multiple-instance QSAR

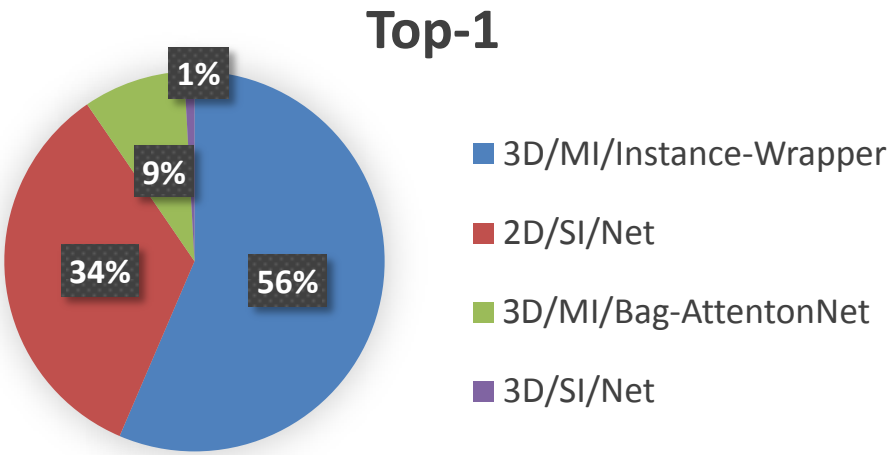
model	mean	median	top 1	top 2
3D/MI/Instance-Wrapper	0.524 ± 0.131	0.526	69	105
3D/MI/Bag-Attention	0.468 ± 0.161	0.474	12	57
2D/MorganFP/Net	0.464 ± 0.199	0.502	39	66
2D/PhysChem/Net	0.450 ± 0.144	0.443	17	37
2D/PharmFP/Net	0.382 ± 0.216	0.404	4	17
3D/SI/Net	0.024 ± 0.372	0.089	1	2

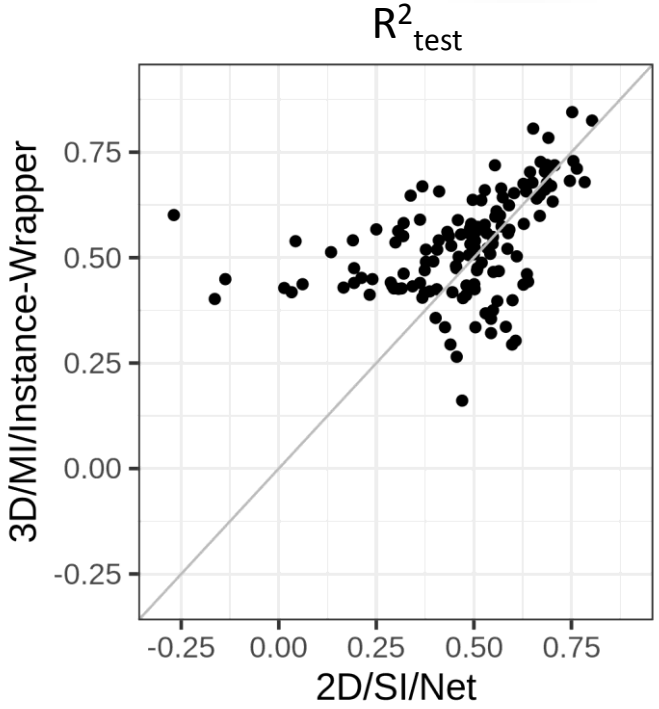
[a]Table reports mean, standard deviations, and median of R_{test}^2. Top 1 is the number of cases where the model was the best. Top 2 is the number of cases where the model was the first- or second-best one.

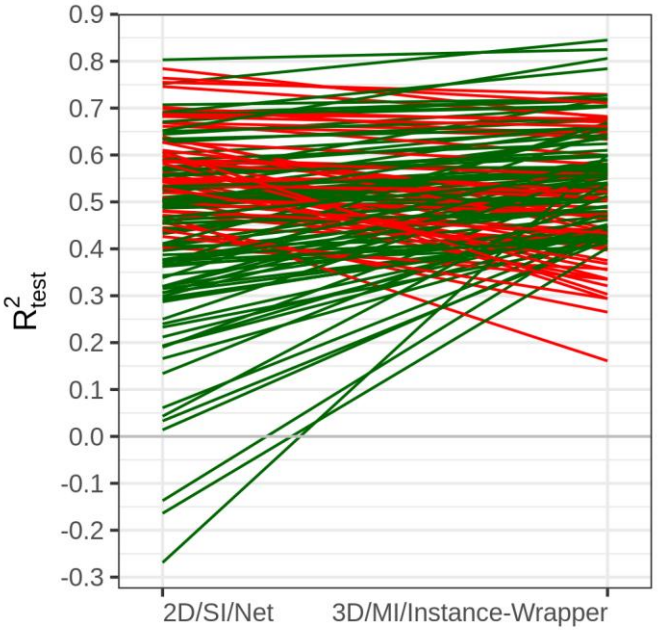


Zankov, D. V. et al, QSAR Modeling Based on Conformation Ensembles Using a Multi-Instance Learning Approach. *Journal of Chemical Information and Modeling* **2021**, 61, 4913-4923.

MIL study: comparison between 2D, 3D and MIL models



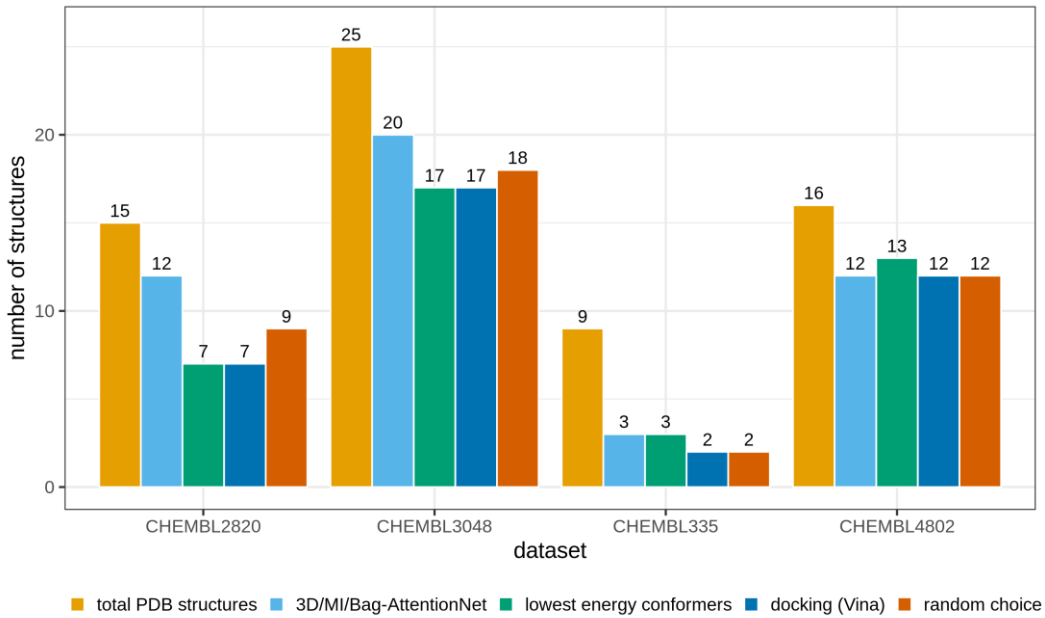


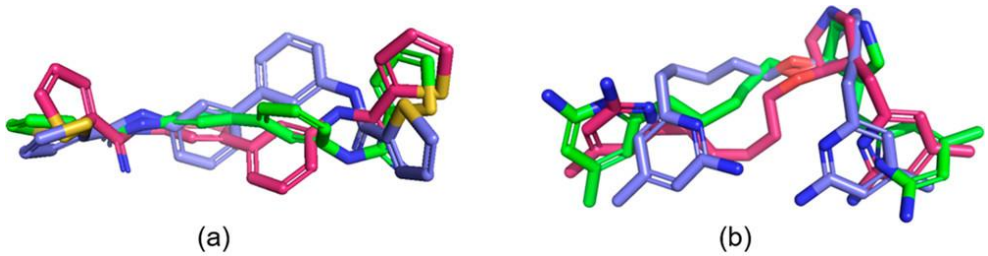


Zankov, D. V. et al, QSAR Modeling Based on Conformation Ensembles Using a Multi-Instance Learning Approach. *Journal of Chemical Information and Modeling* **2021**, 61, 4913-4923

MIL study: identification of “bioactive” conformers

Selected compounds had average RMSD of generated conformers > 2A relative to PDB structure





(a)

Experimental pKi: 6.10
3D/SI/Net pKi: 6.48, RMSD = 2.42 Å
3D/MI/Bag-AttentionNet pKi: 6.31
(attention weight: 0.83), RMSD = 1.70 Å

(b)

Experimental pKi: 7.42
3D/SI/Net pKi: 7.86, RMSD = 2.78 Å
3D/MI/Bag-AttentionNet pKi: 7.41
(attention weight: 0.59), RMSD = 1.55 Å

Zankov, D. V. et al, QSAR Modeling Based on Conformation Ensembles Using a Multi-Instance Learning Approach. *Journal of Chemical Information and Modeling* **2021**, 61, 4913-4923

Conclusions

+ Universal representation of binding pattern
+ Qualitative output
+ Very fast screening
+ Scaffold hopping

- Structure-based models can be very specific
- Ligand-based models depend on conformational sampling