Promiscuous Enzyme Activity

Predicting Substrates from 3-Dimensional Structures

Maximilian Faissner

Main Supervisor: Christoph Flamm

What is Enzyme Promiscuity?

Def. **Broad Substrate Specificity**

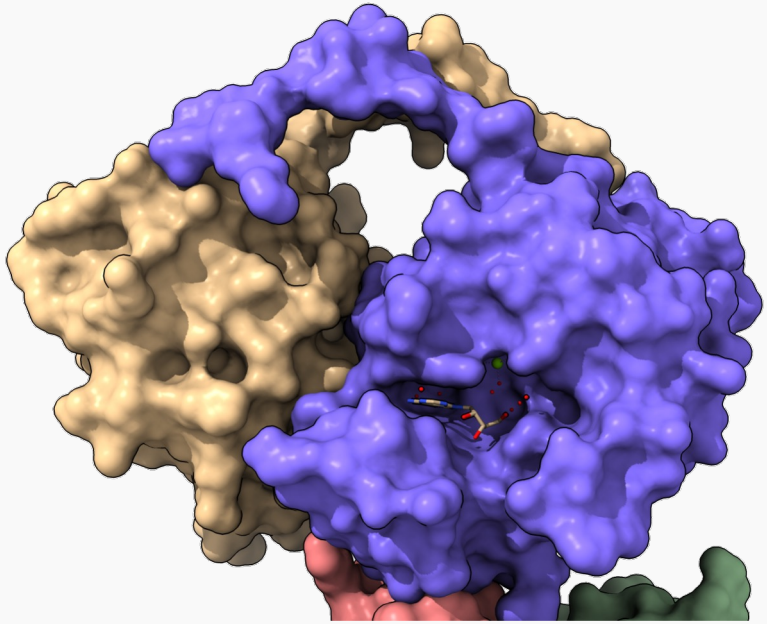
act on a range of structurally similar substrates with a relatively high efficiency.

Def. **Enzyme Promiscuity**

Secondary activities, usually less efficient, may involve structurally distinct substrates

Reactions are **physiologically irrelevant**

- Activities are too inefficient to affect fitness
- Enzyme never encounters substrate



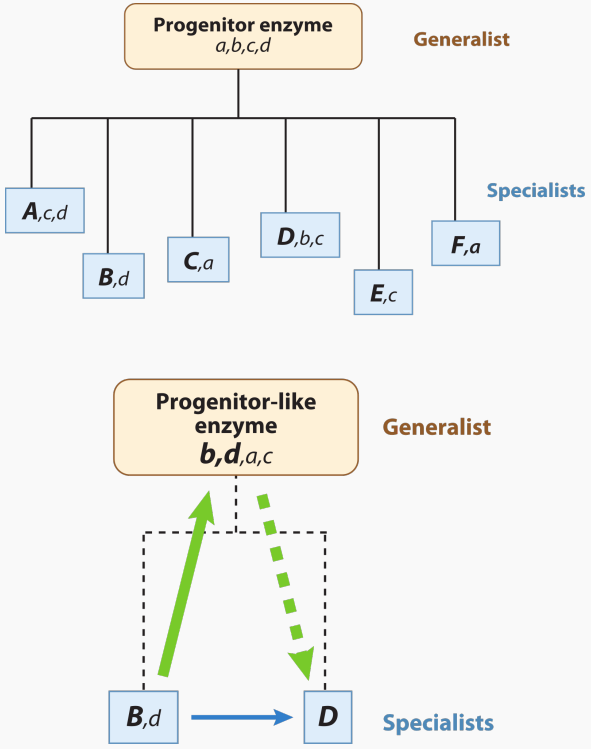
Crystal Structure of a phosphatase (RCSB ID: 1RMT)

Applications of Enzymatic Promiscuity

- **Evolutionary Studies**
  - Specificity is shaped by natural selection
  - Natural selection optimizes specificity, such that other reactions do not impair the fitness.
  - Promiscuous activities are evolutionary starting points of new enzymes

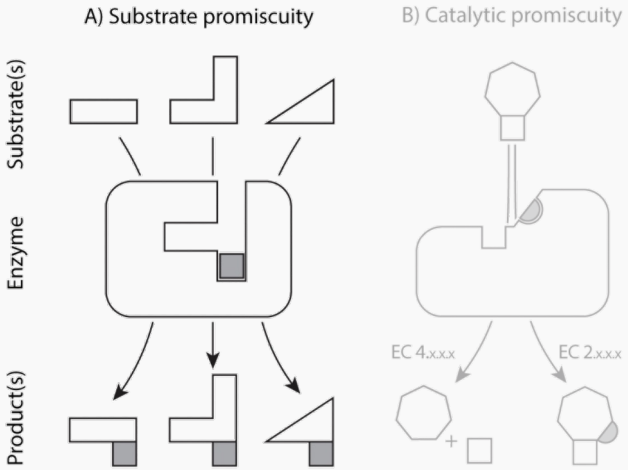
- **Metabolic networks**
  - flexibility and adaptability of metabolic networks
  - utilize diverse substrates and respond to changes in environmental conditions

- **Drug Discovery, Biotechnology**



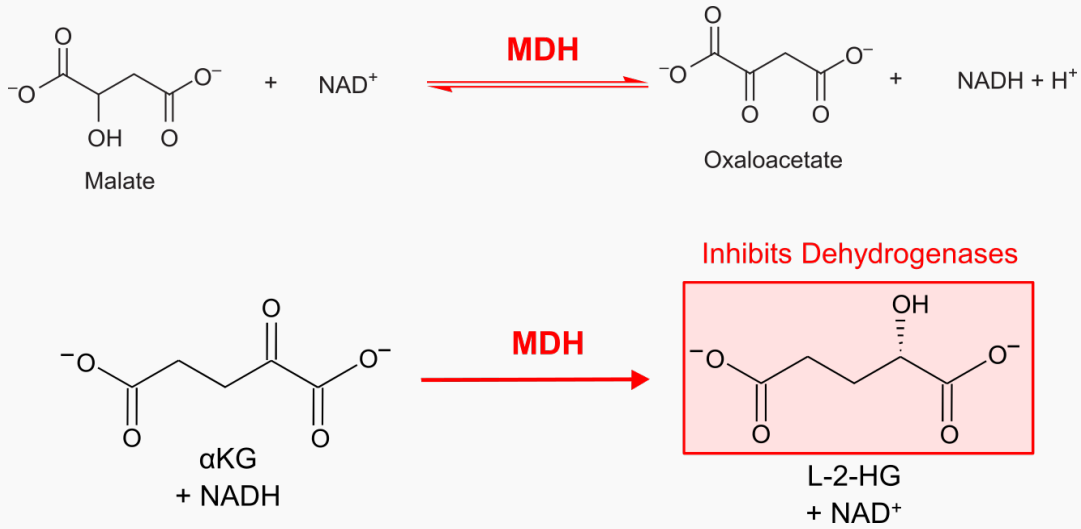
Khersonsky and Tawfik (2010)

Substrate Promiscuity



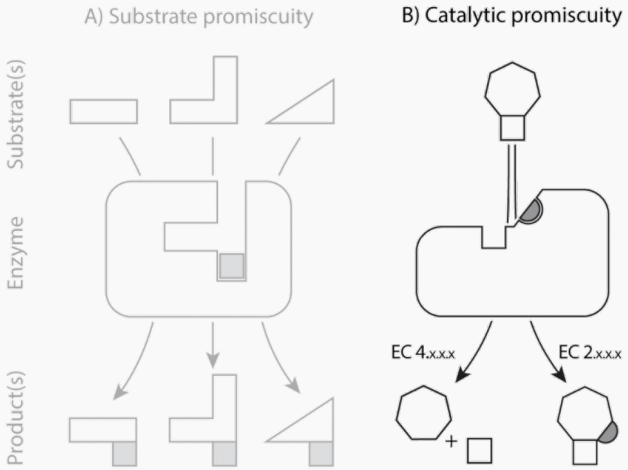
Adapted from: Piedrafita G et al. (2015)

Example: Malate dehydrogenase



Adapted from: Niehausand T and Hillmann K (2019)

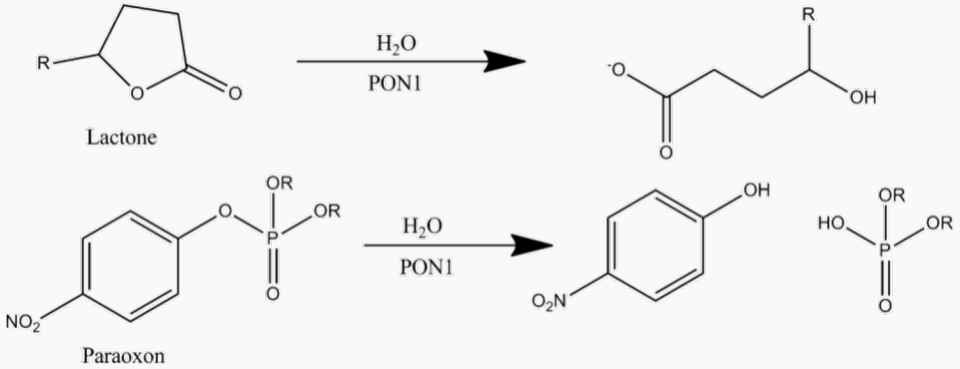
Catalytic Promiscuity



Adapted from: Piedrafita G et al. (2015)

Example: Serum Paraoxonase (PON1)

native lactonase activity

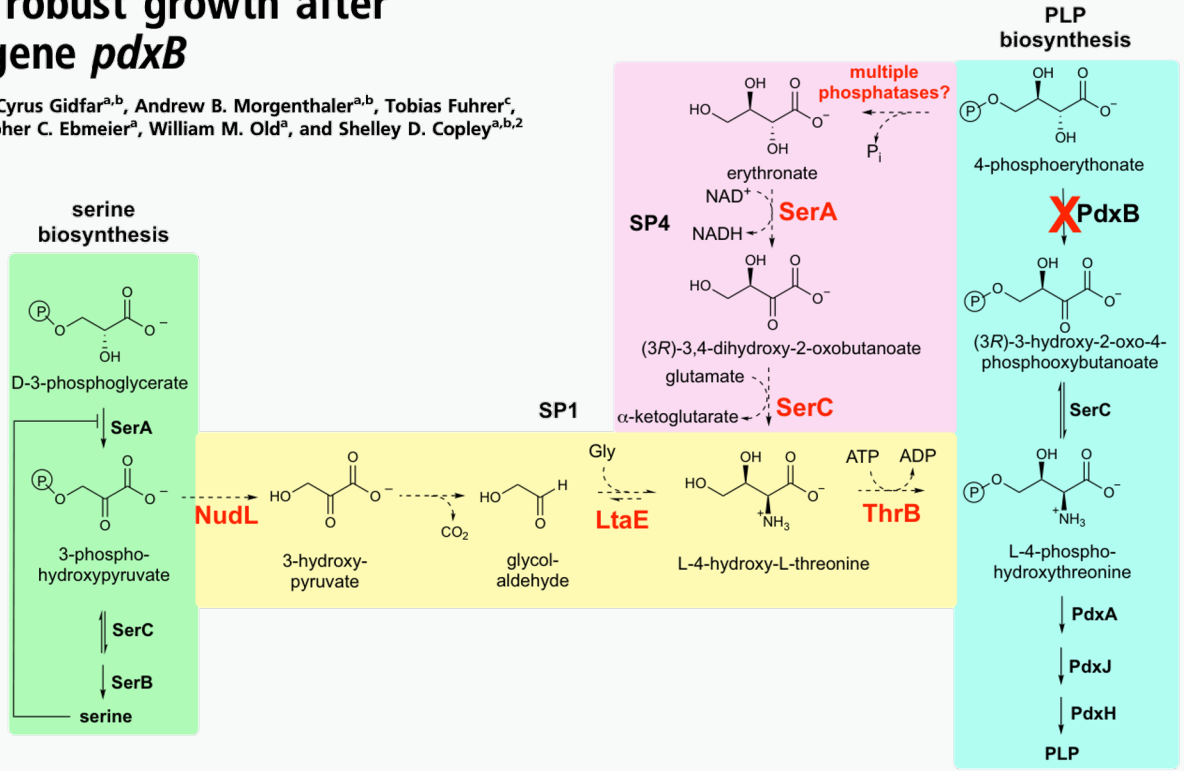


promiscuous phosphotriesterase activity

Adapted from: Gupta RD (2016)

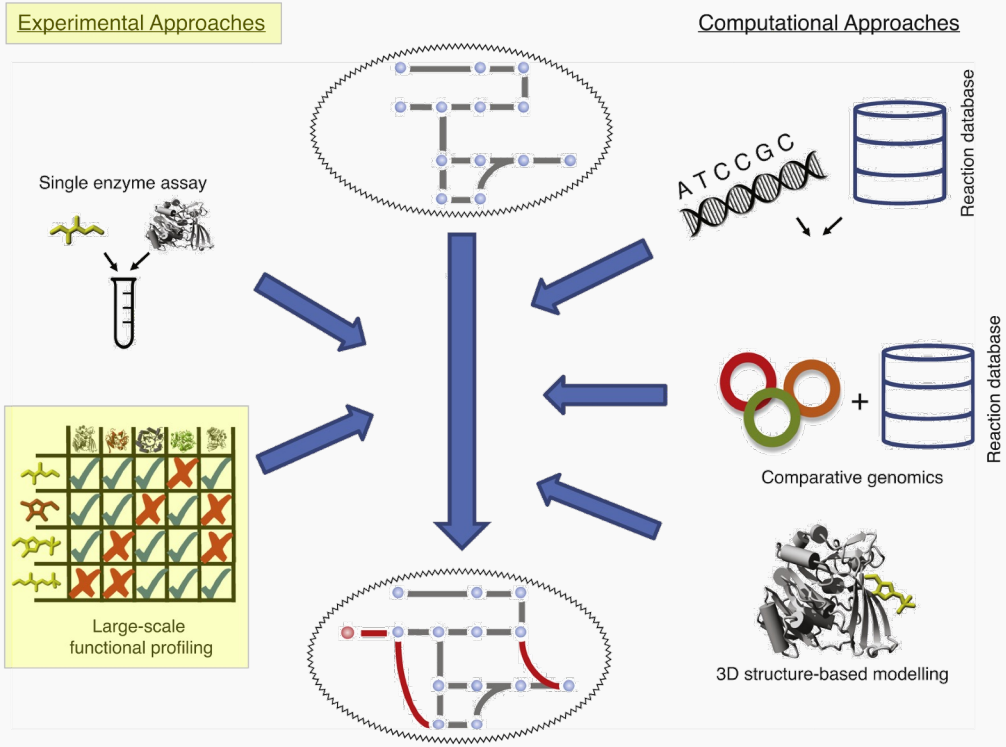
Hidden resources in the *Escherichia coli* genome restore PLP synthesis and robust growth after deletion of the essential gene *pdxB*

Juhan Kim[a,b,1], Jake J. Flood[a,b,1], Michael R. Kristofich[a,b], Cyrus Gidfar[a,b], Andrew B. Morgenthaler[a,b], Tobias Fuhrer[c], Uwe Sauer[c], Daniel Snyder[d], Vaughn S. Cooper[d], Christopher C. Ebmeier[a], William M. Old[a], and Shelley D. Copley[a,b,2]



Two different “Serendipitous pathways” (SPs) restore PLP synthesis (pyridoxal 5′-phosphate) in the absence of PdxB.

On the hunt for promiscuous activities



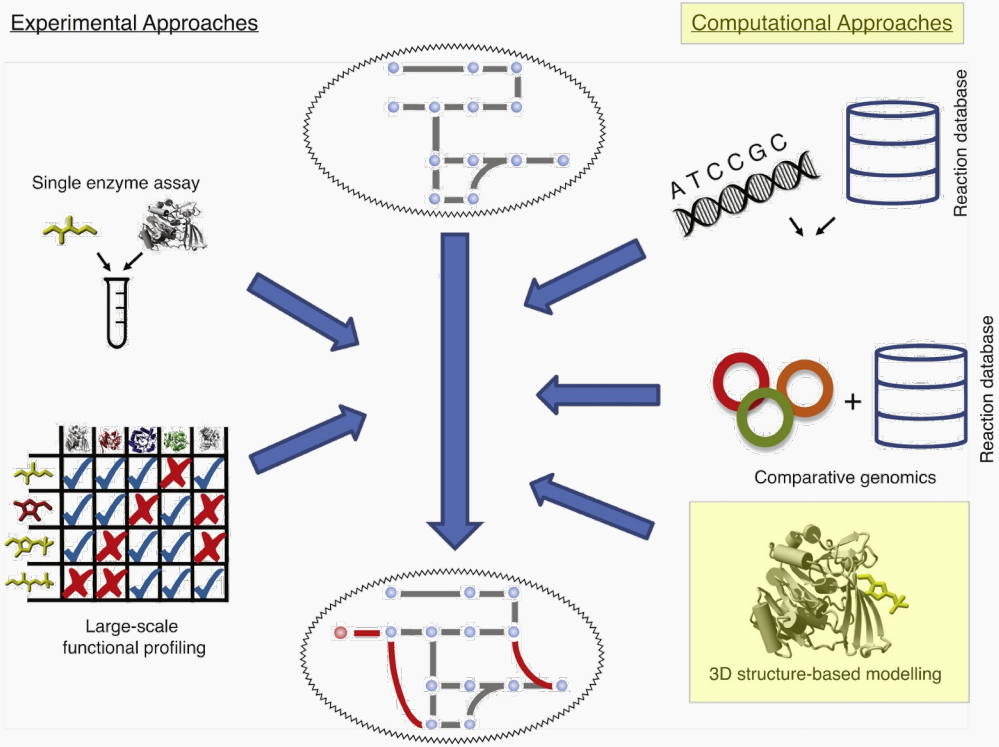
adopted from Notebaart R. et al. (2018)

High-Throughput challenges

- enzyme availability
- activity may be undetectable if inhibited by other metabolites
- substrate availability (fluorogenic substrates)
- activity screening: only small change in absorbance

Demand for robust in silico predictions

On the hunt for promiscuous activities

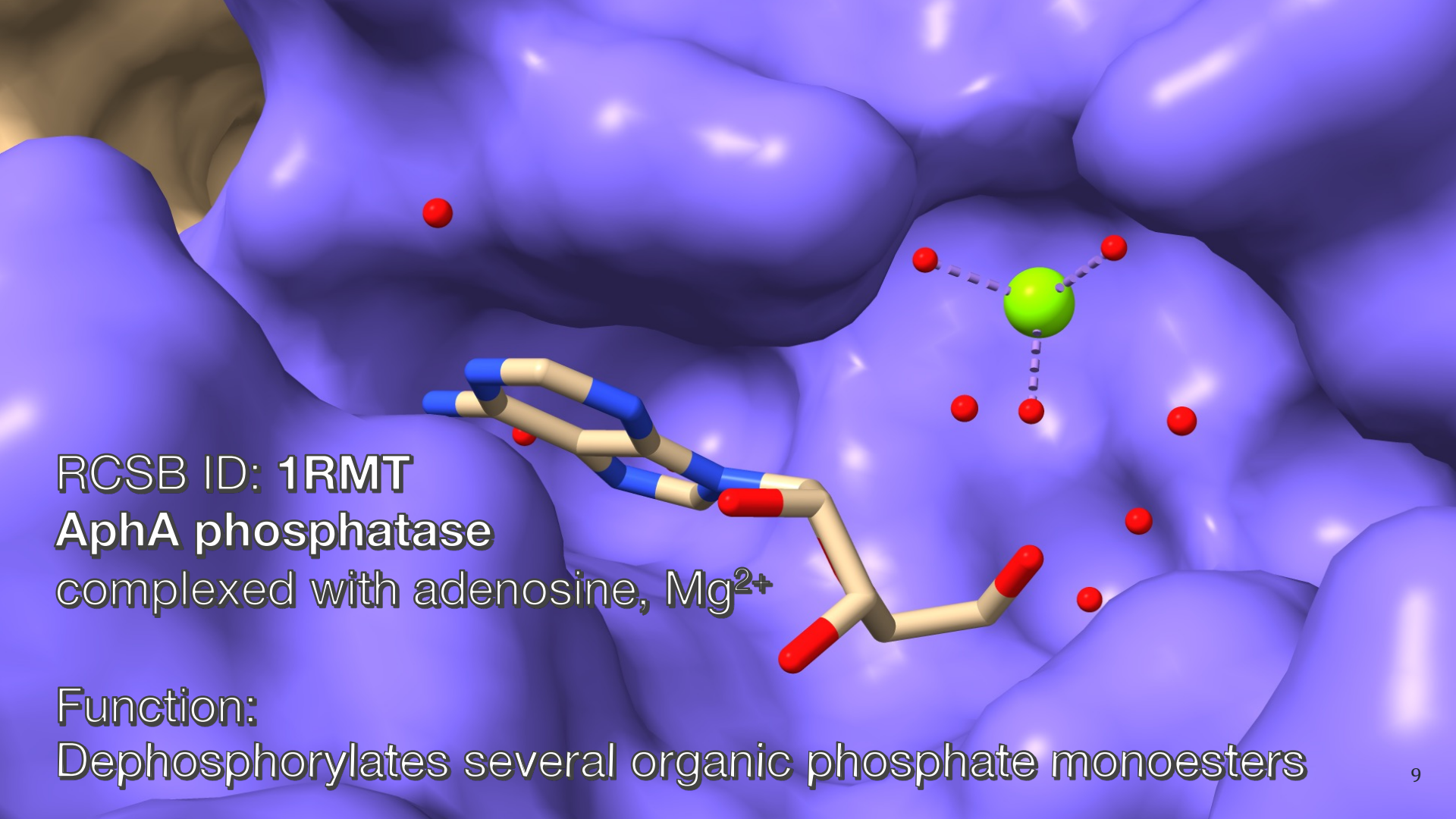


adopted from Notebaart R. et al. (2018)

Example: ML-based Enzyme-Substrate Pair Prediction, by Kroll A et al. 2023, https://esp.cs.hhu.de/

- Good predictions even for unseen enzymes
- Low model performance for unseen small molecules

Structure-based modelling as alternative approach



RCSB ID: **1RMT**
AphA phosphatase
complexed with adenosine, Mg^{2+}

Function:
Dephosphorylates several organic phosphate monoesters

Framework for structure-based drug design

Pharmacophore Models: (next slide)

- simplest and most abstract models to represent essential features of a ligand complex

Molecular Docking:

- Prediction of preferred binding orientation of a ligand
- Estimation of binding affinities
- Software: Autodock VINA, GOLD, …

Molecular Dynamics Simulations:

- Simulating the physical movement of atoms and molecules over time

Constraints

- M-CSA (Mechanism and Catalytic Site Atlas)
- EzMechanism (additional and more generic catalytic rules)

Catalytic Residues Roles

Residue	Roles
Glu147A(AA)	hydrogen bond acceptor
Asp7A	activator, hydrogen bond acceptor
His180A	hydrogen bond donor
Cys70A	activator, hydrogen bond donor
Ser8A	hydrogen bond donor, electrostatic stabiliser, increase basicity
Cys70A	**proton donor**
Asp7A	**proton acceptor**

Pharmacophore Modelling

Definition:

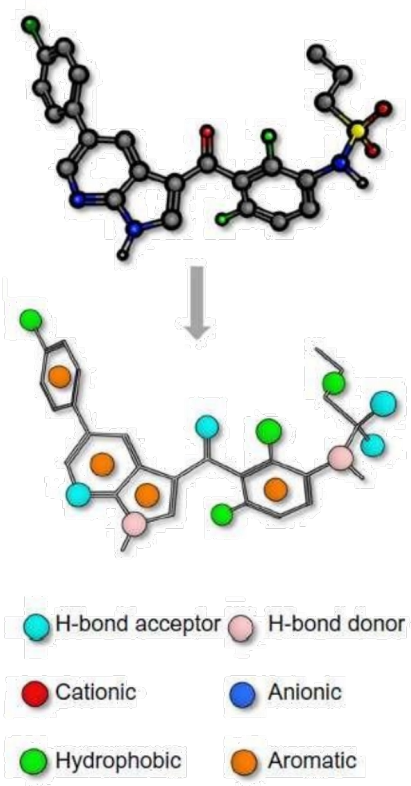
"largest common denominator"
(w.r.t. interactions)

Represent the essential features of a ligand necessary for its biological activity

Essential keystone in drug discovery workflows:
Screening large databases of compounds, identify potential ligands with matching properties

Collaboration with Thierry Langer, Thomas Seidl:

- LigandScout (closed source), cdpkit (open source)



Adapted from: Pandagale N.V. et al. (2022)

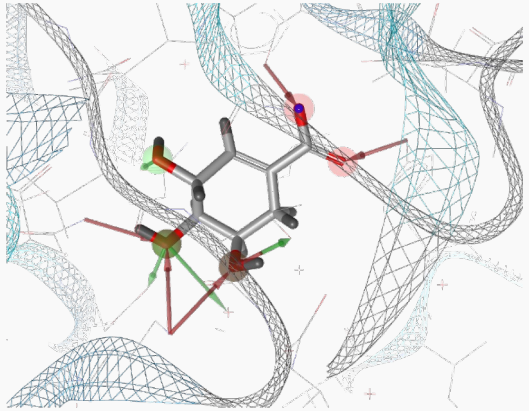
Generating Pharmacophores, Screening

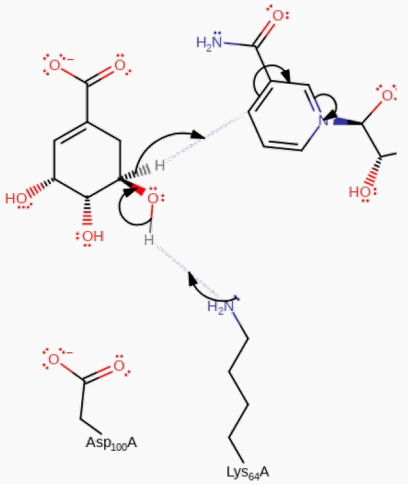
Requirement: ligand pose within active site

- Simple case: modelled ligand within crystal structure
- Molecular docking (with/without constraints)
- Fragment docking (apo-pharmacophores)

After screening process against library:

Pharmacophore-Fit Score is a starting point, postprocessing of results required (e.g. redocking of resulting conformers





Project Roadmap

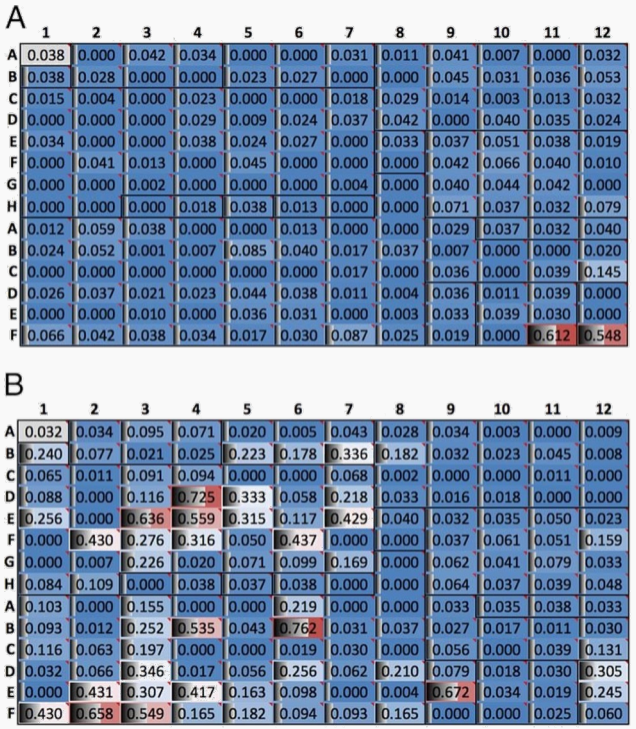
Starting Point

- Substrate Promiscuity Prediction
- Compare prediction with vitro enzyme assays (e.g. phosphatases)
- Crystal structures of enzymes with modelled ligands

Extensions

- Predicted structures (AlphaFold, homology modelling) / predicted ligands
- Extend framework to also include more sophisticated enzymes (additional cofactors like NADH), multi-step reactions

Promiscuity Database, Metabolic Networks



Huang H et al. (2015)

Acknowledgements

Christoph Flamm	University of Vienna, TBI
Shelley Copley	University of Colorado, Boulder
Walter Fontana	Harvard University
Thierry Langer	University of Vienna, Department of Pharmaceutical Chemistry
Thomas Seidel	University of Vienna, Department of Pharmaceutical Chemistry
MATOMIC Team	Syddansk Universitet

Thank You for Your Attention

Enzyme Focus – late 2023 PDB statistics

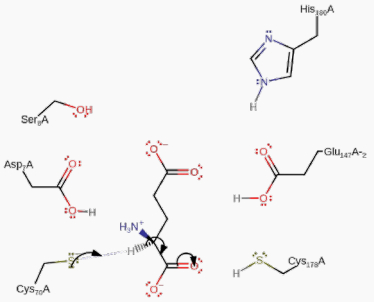
UniProtKB: query: E.coli strain K-12 MG165

- 5279 results of genes
- filter for catalytic activity: 1559
- filter for 3d structure: 873
- Information about ligand/receptor complex:
  - Uniprot filter for active/binding site annotations: 363
  - better resource: catalytic site atlas (M-CSA)
  - RCSB: many ligand/receptor complexes

- 873 structures: include serA, serC, some of the glutamine-dependent amidotransferases (carAB, pyrG, purF)
- Related work: Find structures from related species (example: ThrB: homoserine kinase), analyse differences (alignment, …)

Catalytic Residues Roles

Residue	Roles
Glu147A(AA)	hydrogen bond acceptor
Asp7A	activator, hydrogen bond acceptor
His180A	hydrogen bond donor
Cys70A	activator, hydrogen bond donor
Ser8A	hydrogen bond donor, electrostatic stabiliser, increase basicity
Cys70A	**proton donor**
Asp7A	**proton acceptor**



M-CSA (EC 5.1.1.3, Marvin Files)

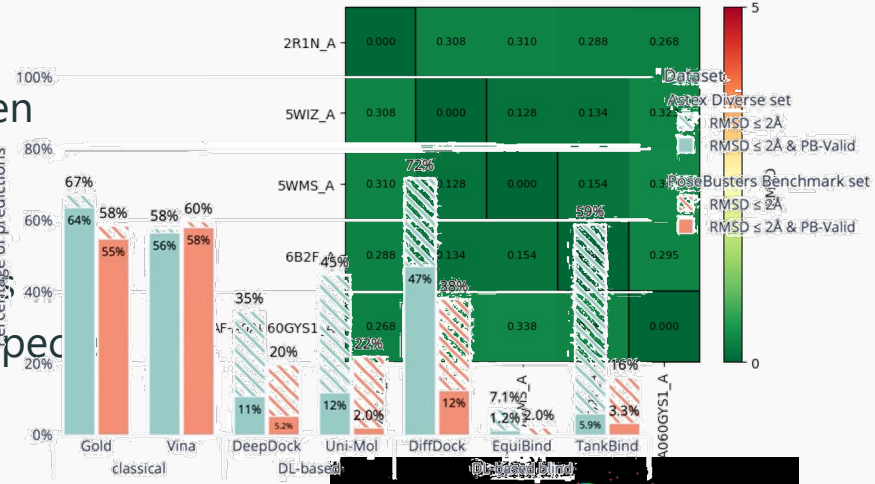
Current Project State

Starting point: Uniprot ID

- Downloads and analyses referenced / chosen structures
- sequence and structural alignment
- Ligand analysis, if present / missing residues
- TODO: structures / metadata from related spec

Define ligands / SMILES as starting point:

- Conformer Generation (.sdf files) with CONFORGE (Thomas Seidl, Thierry Langer group)
- Docking with SMINA (related to Autodock VINA) and GOLD (not yet automated)
- Best docked poses are the starting point for the pharmacophore generation (cdpkit)



Source: 2023-11-17 [link]
https://www.youtube.com/watch?v=nWDzUcoDM7w&t=3615s

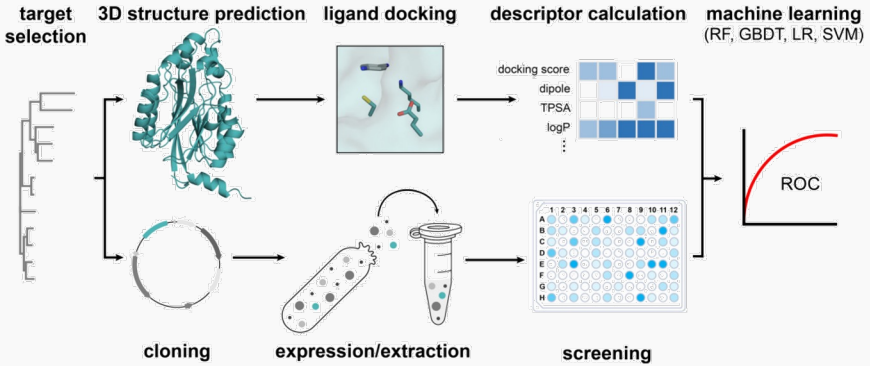
solid bars have additional plausibility requirements, besides RMSD

A few words about molecular docking (2)

Proteins. 2021 Mar;89(3):336-347. doi: 10.1002/prot.26019. Epub 2020 Nov 10.

Machine learning–based prediction of enzyme substrate scope: Application to bacterial nitrilases

Zhongyu Mou [1], Jason Eakes [1], Connor J Cooper [2], Carmen M Foster [1], Robert F Standaert [1], Mircea Podar [1], Mitchel J Doktycz [1 2], Jerry M Parks [1 2]



Substrate Prediction by combining structural modelling, docking, physiochemical properties, and machine learning models.

"Unsurprisingly, docking scores do not correlate with enzymatic activity"

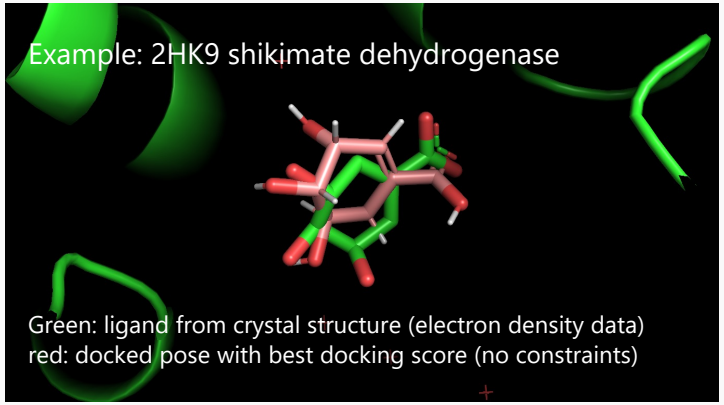
Improving the Docking Procedure

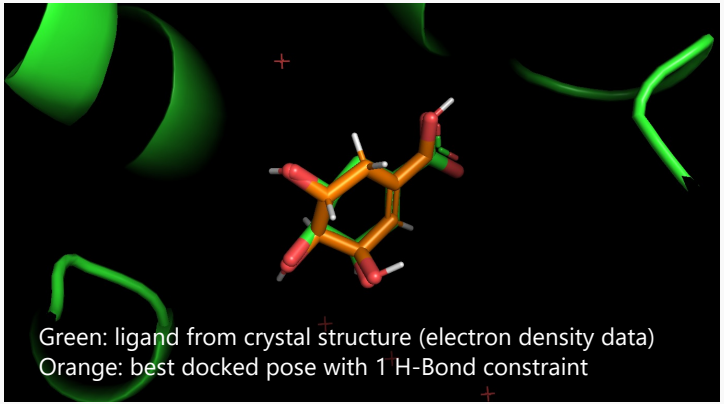
Improving the docking procedure by adding constraints (GOLD):

- similarity constraints
- scaffolding and regional constraints (hydrophobic regions)
- pharmacophore constraints

Start with “native” ligand, generate pharmacophore, dock other substrates with constraints

In addition, check all poses for chemical validity and consistency check (e.g. PoseBusters)





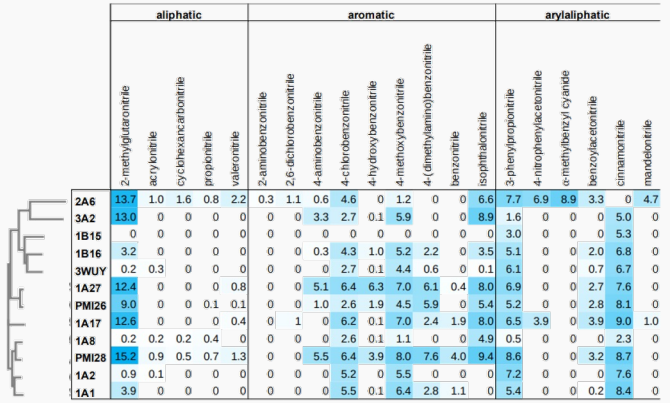
Future Work and Discussion

- create pharmacophore models / virtual screening for ligand/receptor complexes with known negative experimental activity data
- Evaluate and compare different virtual screening options (LigandScout, cdpkit, …)
- Verify Screening Results with docking (with and without constraints)

› Proteins. 2021 Mar;89(3):336-347. doi: 10.1002/prot.26019. Epub 2020 Nov 10.

Machine learning-based prediction of enzyme substrate scope: Application to bacterial nitrilases

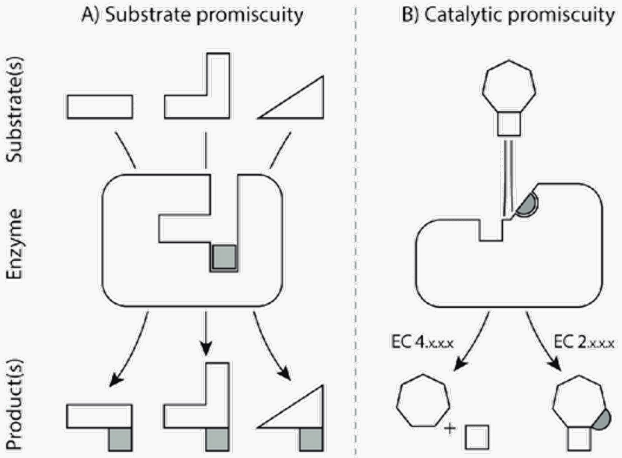
Zhongyu Mou [1], Jason Eakes [1], Connor J Cooper [2], Carmen M Foster [1], Robert F Standaert [1], Mircea Podar [1], Mitchel J Doktycz [1,2], Jerry M Parks [1,2]



	aliphatic					aromatic									arylaliphatic					
	2-methylglutaronitrile	acrylonitrile	cyclohexancarbonitrile	propionitrile	valeronitrile	2-aminobenzonitrile	2,6-dichlorobenzonitrile	4-aminobenzonitrile	4-chlorobenzonitrile	4-hydroxybenzonitrile	4-methoxybenzonitrile	4-(dimethylamino)benzonitrile	benzonitrile	isophthalonitrile	3-phenylpropionitrile	4-nitrophenylacetonitrile	α-methylbenzyl cyanide	benzoylacetonitrile	cinnamonitrile	mandelonitrile
2A6	13.7	1.0	1.6	0.8	2.2	0.3	1.1	0.6	4.6	0	1.2	0	0	6.6	7.7	6.9	8.9	3.3	0	4.7
3A2	13.0	0	0	0	0	0	0	3.3	2.7	0.1	5.9	0	0	8.9	1.6	0	0	0	5.0	0
1B15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3.0	0	0	0	5.3	0
1B16	3.2	0	0	0	0	0	0	0.3	4.3	1.0	5.2	2.2	0	3.5	5.1	0	0	2.0	6.8	0
3WUY	0.2	0.3	0	0	0	0	0	0	2.7	0.1	4.4	0.6	0	0.1	6.1	0	0	0.7	6.7	0
1A27	12.4	0	0	0	0.8	0	0	5.1	6.4	6.3	7.0	6.1	0.4	8.0	6.9	0	0	2.7	7.6	0
PMI26	9.0	0	0	0.1	0.1	0	0	1.0	2.6	1.9	4.5	5.9	0	5.4	5.2	0	0	2.8	8.1	0
1A17	12.6	0	0	0	0.4	0	1	0	6.2	0.1	7.0	2.4	1.9	8.0	6.5	3.9	0	3.9	9.0	1.0
1A8	0.2	0.2	0.2	0.4	0	0	0	0	2.6	0.1	1.1	0	0	4.9	0.5	0	0	0	2.3	0
PMI28	15.2	0.9	0.5	0.7	1.3	0	0	5.5	6.4	3.9	8.0	7.6	4.0	9.4	8.6	0	0	3.2	8.7	0
1A2	0.9	0.1	0	0	0	0	0	0	5.2	0	5.5	0	0	0	7.2	0	0	0	7.6	0
1A1	3.9	0	0	0	0	0	0	0	5.5	0.1	6.4	2.8	1.1	0	5.4	0	0	0.2	8.4	0

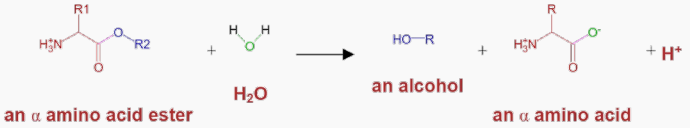
Figure S2. Activity data (ammonia concentration in mM) for putative nitrilases with 20 nitrile substrates at 25% dilution. A phylogenetic tree generated with Clustal Omega is shown on the left.

What is Enzyme Promiscuity? (dt.: Wahllosigkeit, Freizügigkeit)



Adapted from: Piedrafita G et al. (2015)

Example for broad substrate specificity:
EC 3.1.1.43 amino acid ester hydrolase



Definition: **The ability of an enzyme to catalyze more than one type of reaction.**

Examples

Active Site Plasticity

- different active-site conformations
- alternatively, different sub-sites within the same active site

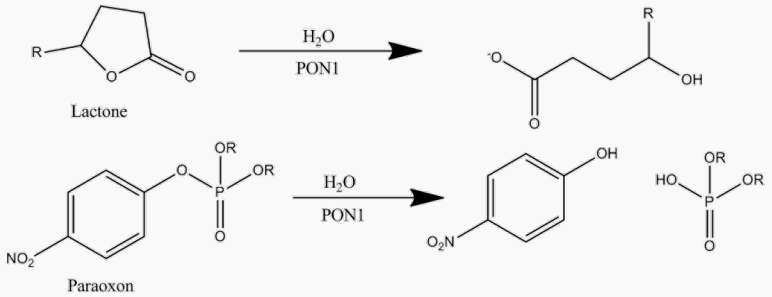
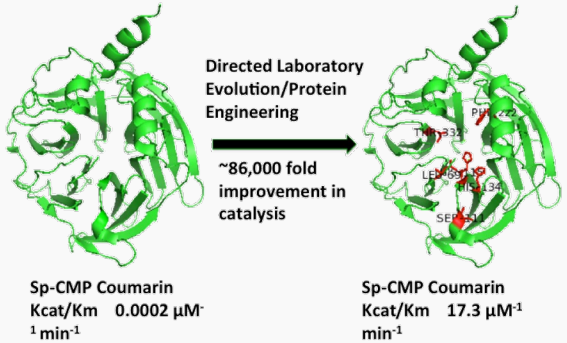


Fig. 2 The native lactonase activity (with lactone substrate) and promiscuous phosphotriesterase activity (with paraoxon) catalyzed by serum paraoxonase (PON1)



Rinkoo Devi Gupta (2016)

Ambiguous substrates

- Cytochromes P450 (diverse enzyme group with common fold, heme cofactor)

Introduction to Molecular Docking

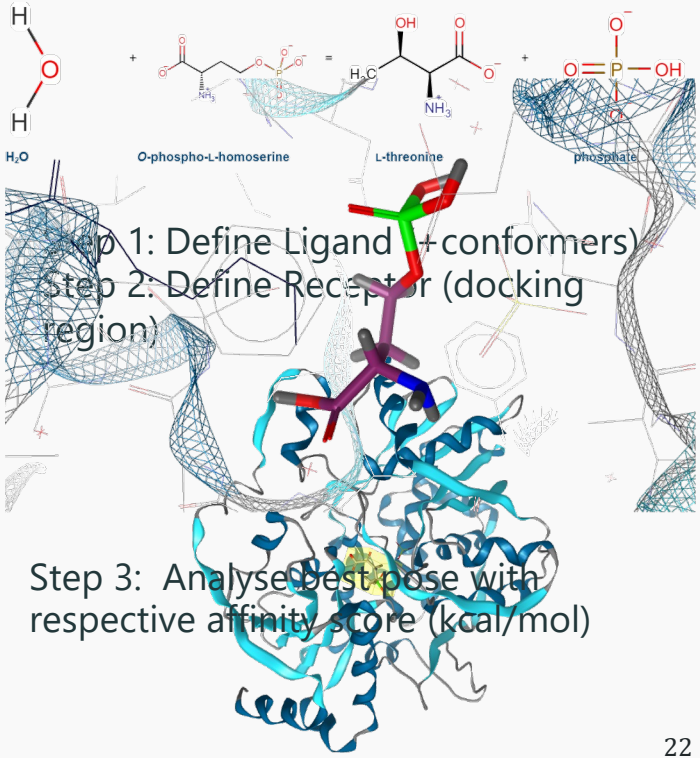
Searches for the preferred binding modes of ligand molecules to bound receptor proteins.

Optimization Problem: Finding the most energetically favourable conformations / orientations of ligands within a defined binding site.

Goal:
- understanding key interactions
- Identify drug candidates / virtual screening

Modern Algorithms: Flexible Docking (induced fit theory), open source examples: Autodock, Autodock VINA

Example: threonine synthase EC 4.2.3.1



From Docking to Pharmacophores

Definitions:

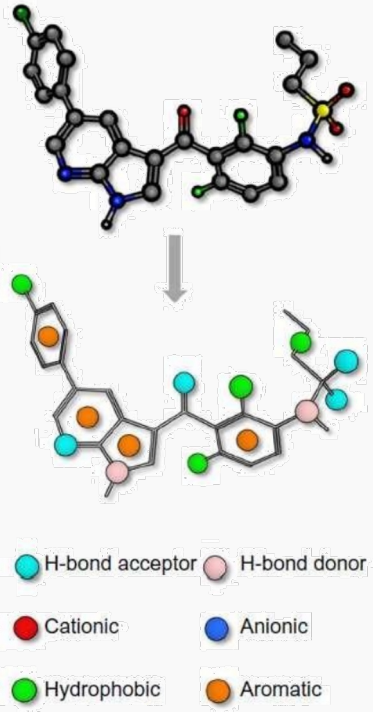
- "A pharmacophore is a pattern of features of a molecule that is responsible for a biological effect."
- "largest common denominator" (w.r.t. interactions)

It represents the geometric arrangement of a ligand's functional groups, crucial for its chemical functionality and biological impact.

Key Features (see figure)

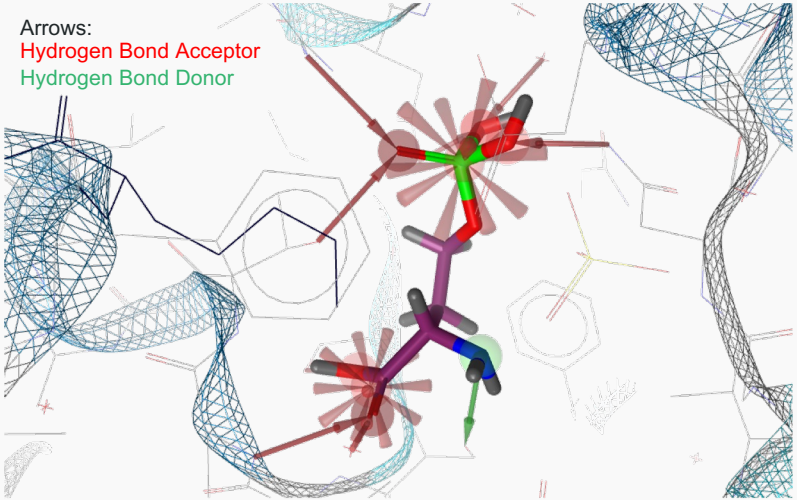
Pharmacophores are essential in drug discovery workflows, e.g. software packages like LigandScout

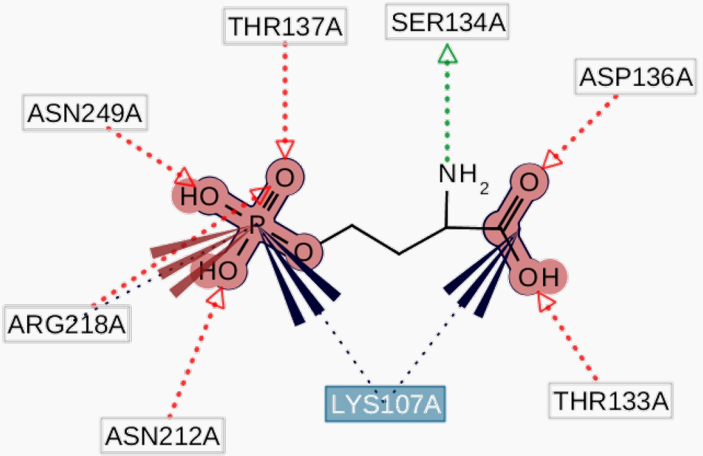
Example: Start with docked ligand/receptor complex



Adapted from: Pandagale N.V. et al. (2022)

Sample Pharmacophore (threonine synthase EC 4.2.3.1)





Various approaches for pharmacophore generation:

- Ligand/receptor-complex based (this example)
- Apo-site pharmacophores (only receptor, e.g. via fragment docking)

Next step: virtual screening against metabolite library

THE END