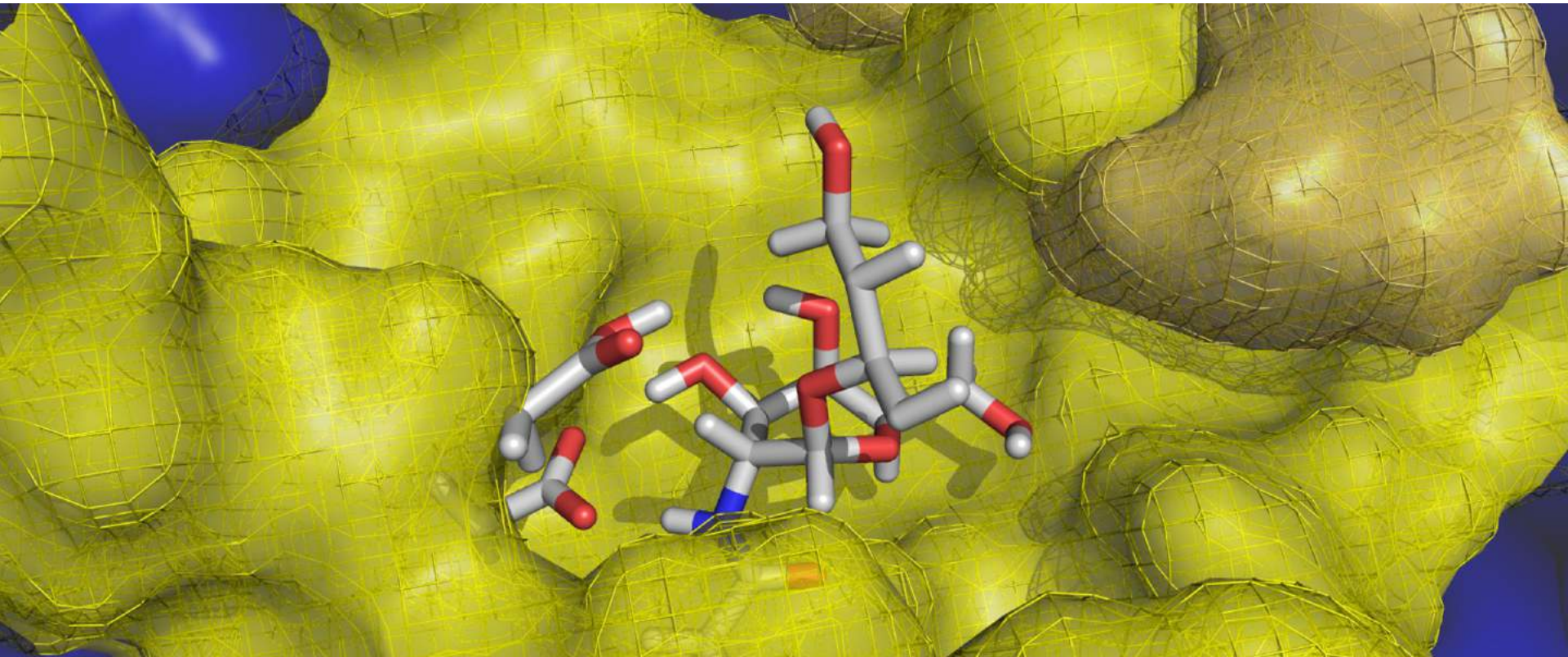
Virtual Screening of Compound Libraries

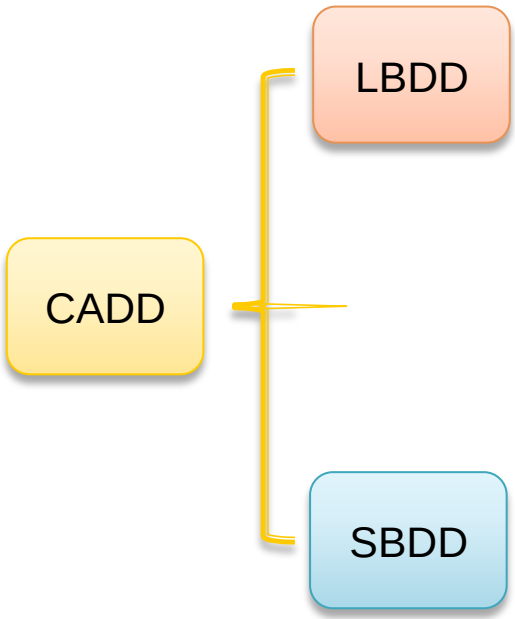
Pedro Alexandrino Fernandes
Department of Chemistry and Biochemistry
University of Porto
Portugal



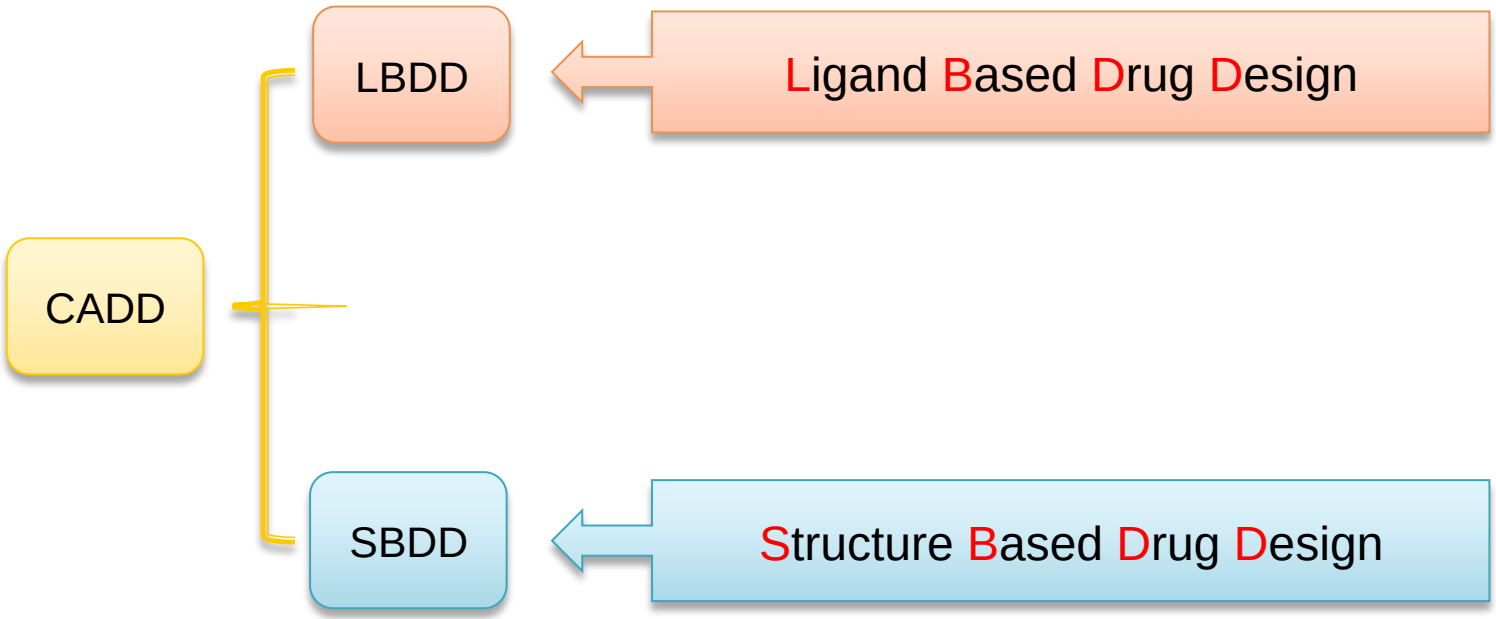
Computer Aided Drug Discovery

CADD

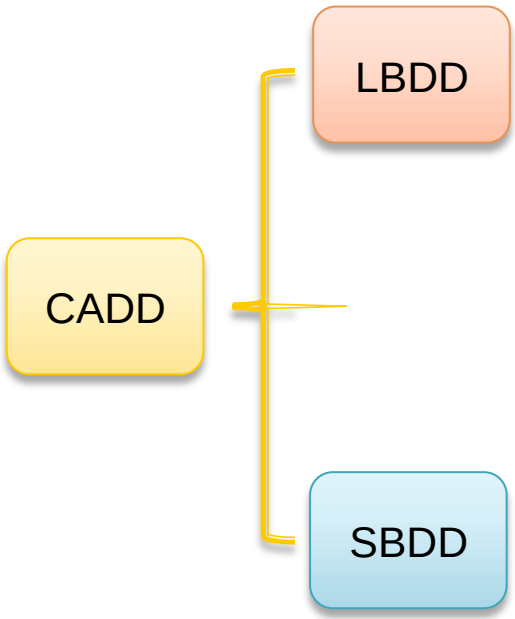
Computer Aided Drug Discovery



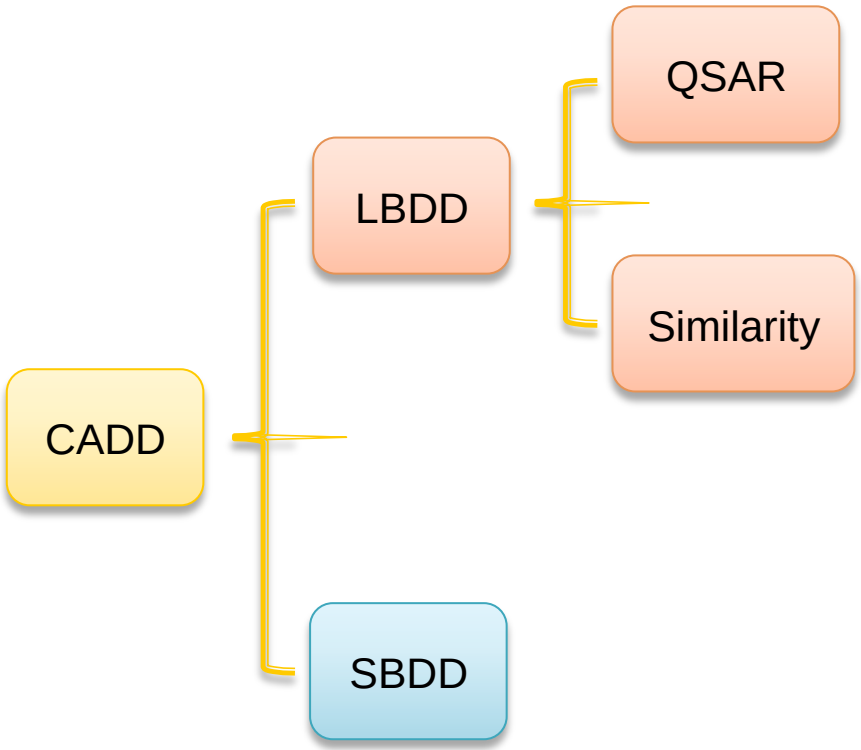
Computer Aided Drug Discovery



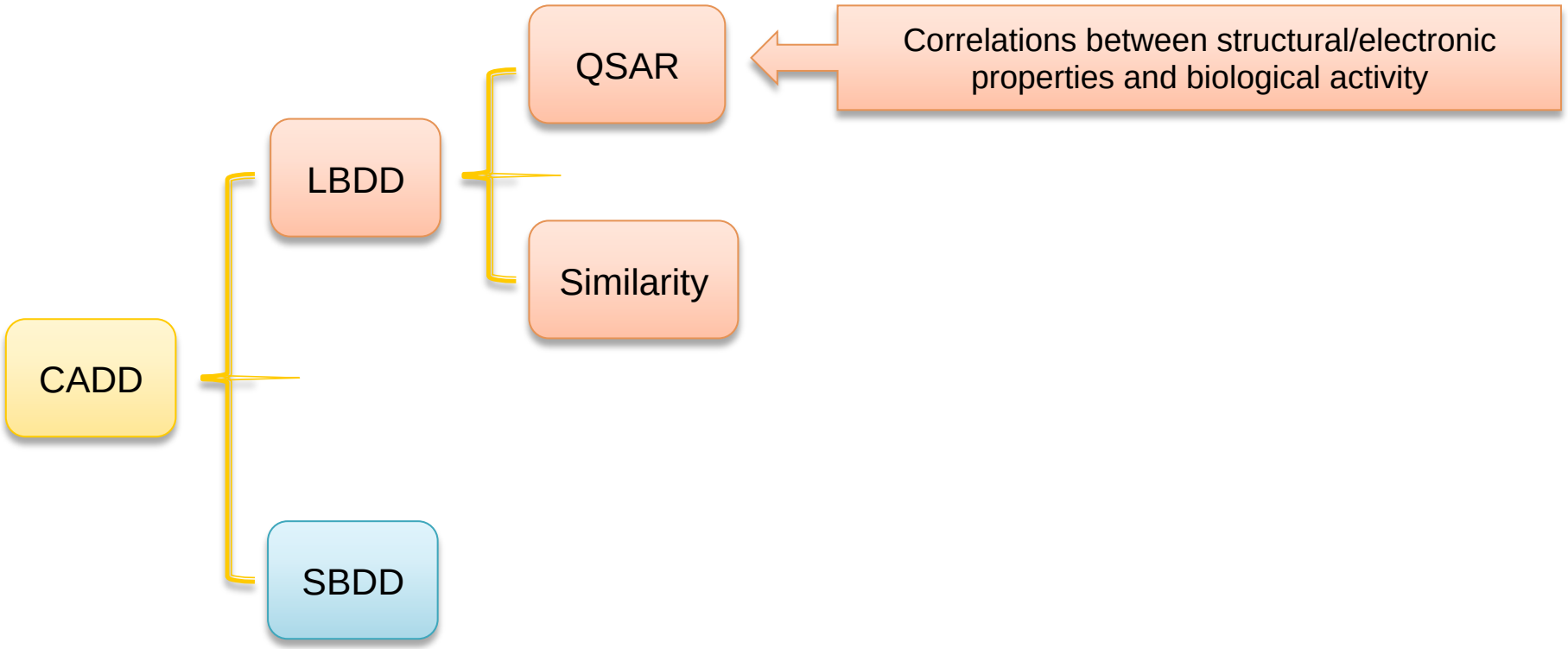
Computer Aided Drug Discovery



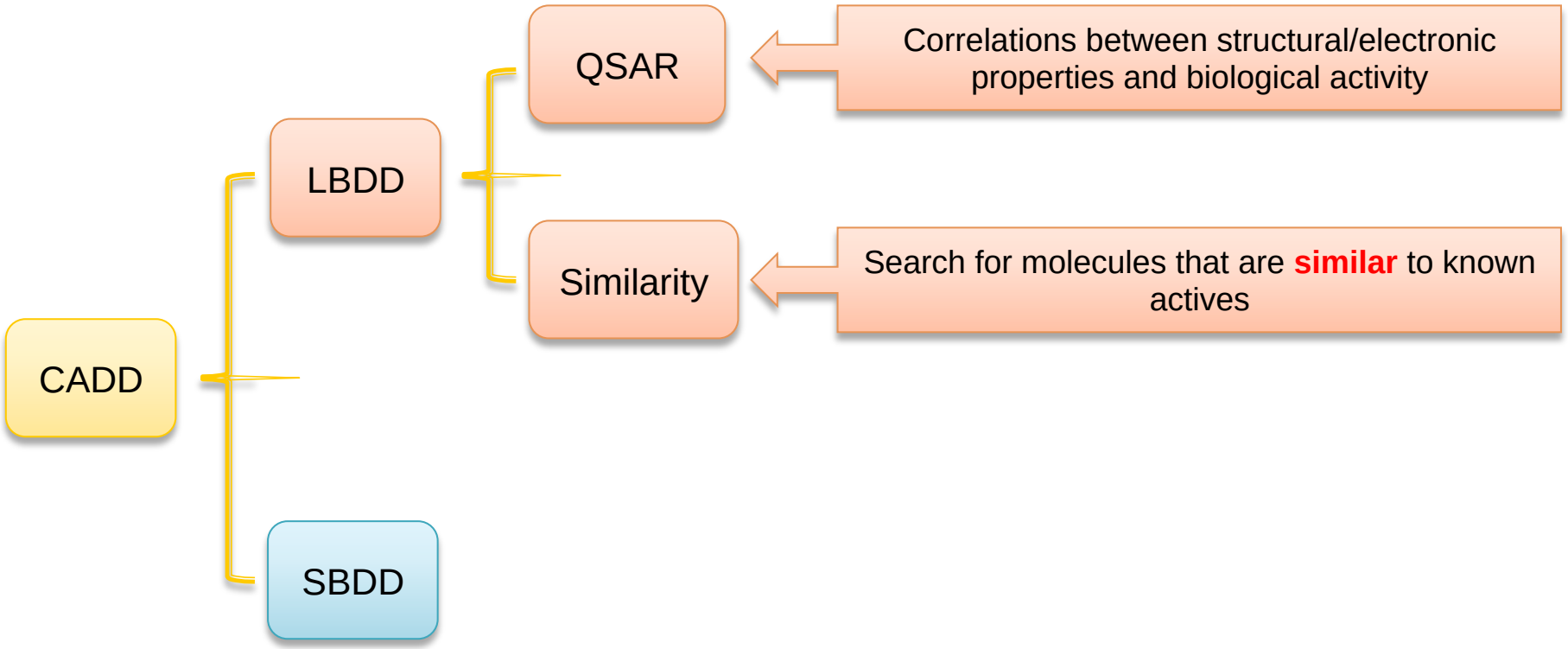
Computer Aided Drug Discovery



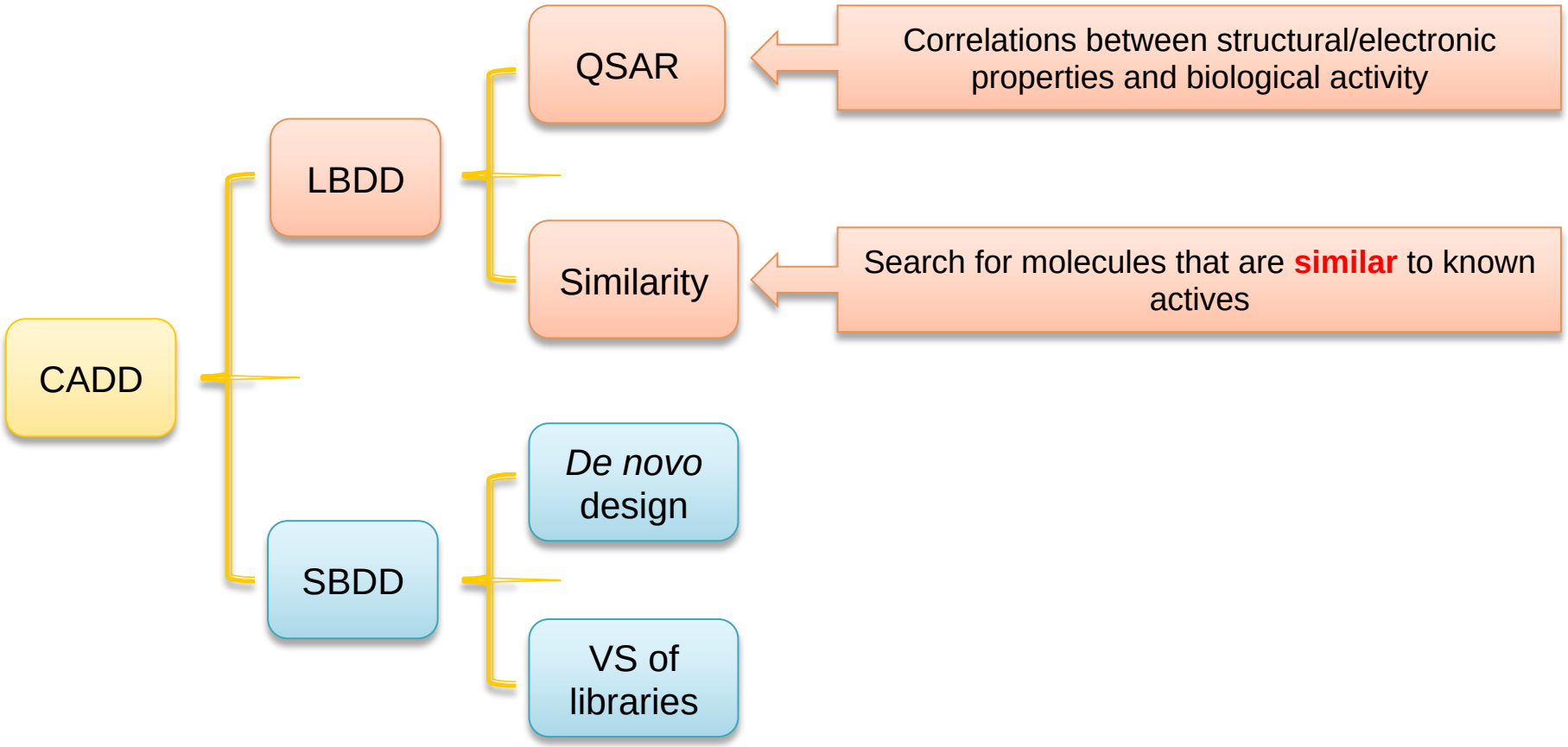
Computer Aided Drug Discovery



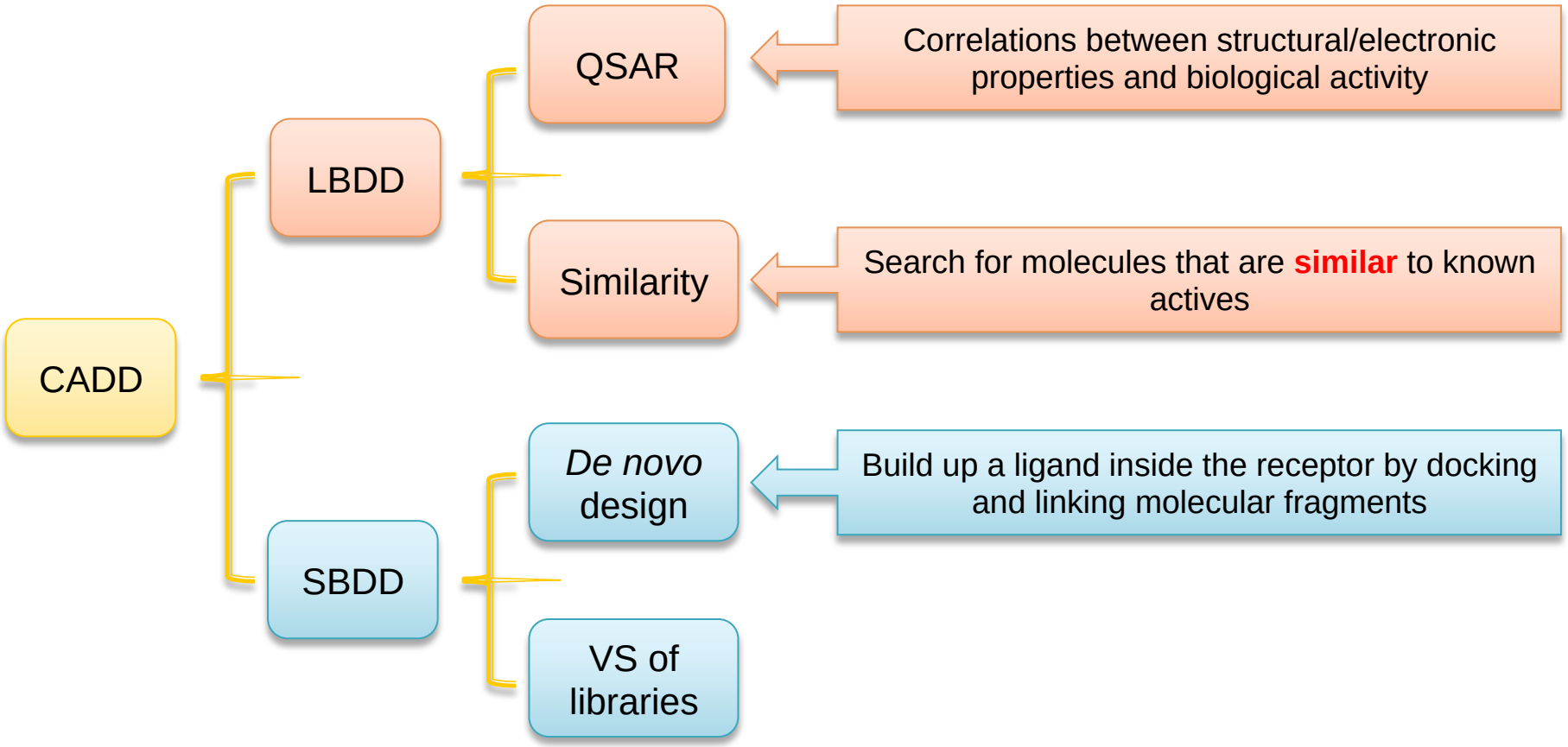
Computer Aided Drug Discovery



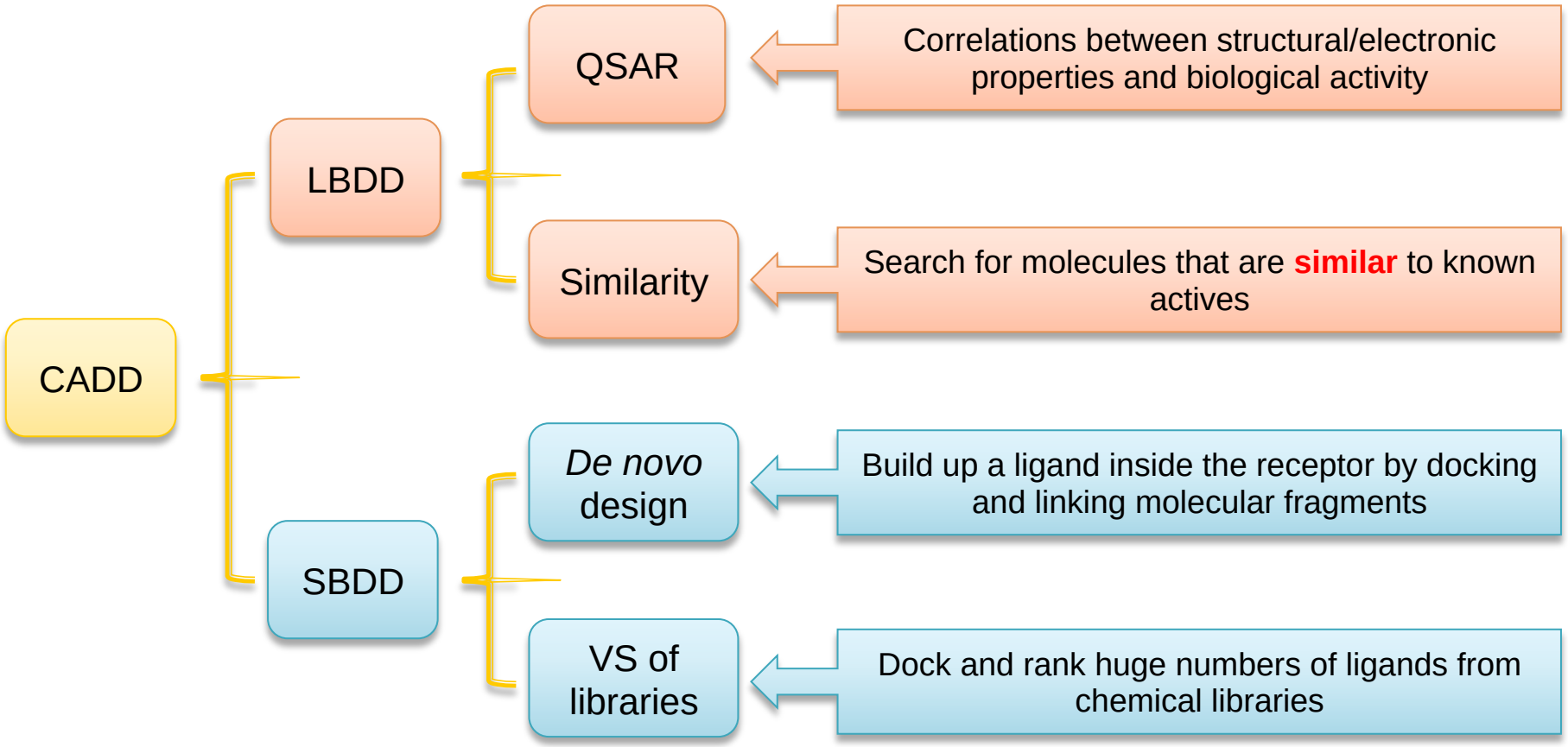
Computer Aided Drug Discovery



Computer Aided Drug Discovery



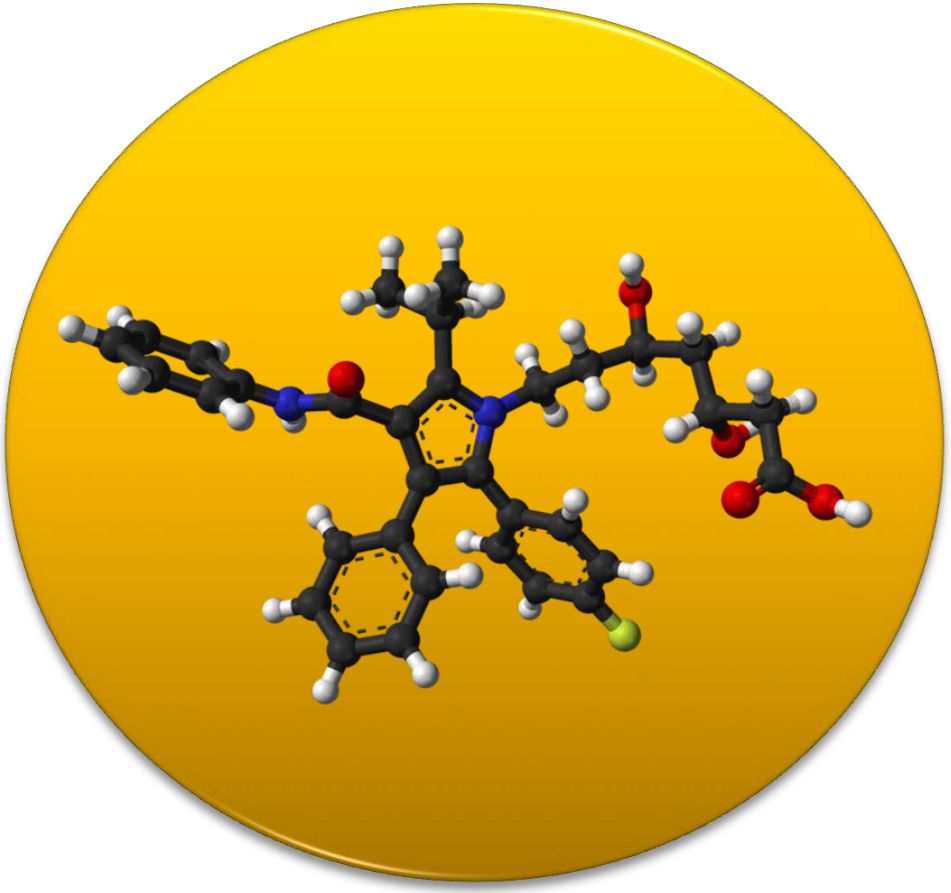
Computer Aided Drug Discovery



Structure-Based Drug Discovery

SBDD

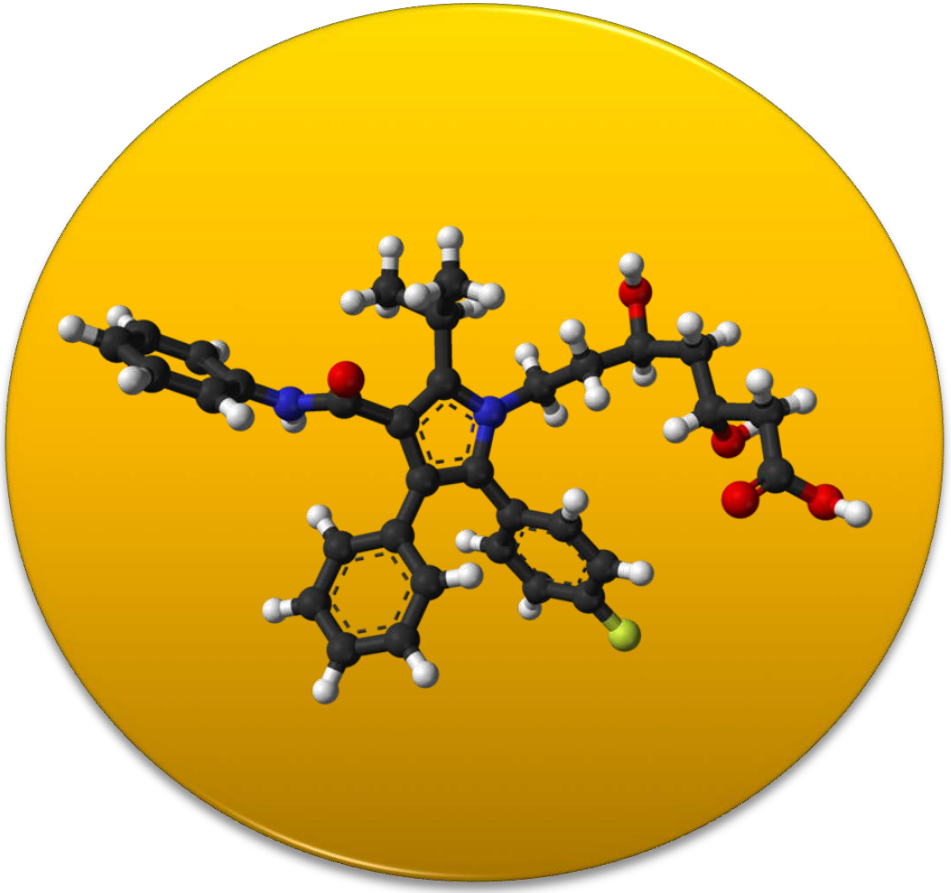
First step: finding a **HIT**



Structure-Based Drug Discovery

SBDD

Second step: improving the **HIT**



Find a hit through Virtual Screening

SBDD

First step: finding a **HIT**

Find a Hit Through Virtual Screening

What is a HIT molecule?

Hit Molecule

What is a HIT molecule?

VIRTUAL **HIT**: Molecule that is predicted to bind the receptor with good affinity

Hit Molecule

What is a HIT molecule?

VIRTUAL **HIT**: Molecule that is predicted to bind the receptor with good affinity

Computational **HIT** molecules may have never been synthesized

Hit Molecule

What is a HIT molecule?

VIRTUAL **HIT**: Molecule that is predicted to bind the receptor with good affinity

Computational **HIT** molecules may have never been synthesized

Experimental **HIT**: Molecule that binds the receptor in solution with μM K_I

Hit Molecule

What is a HIT molecule?

VIRTUAL **HIT**: Molecule that is predicted to bind the receptor with good affinity

Computational **HIT** molecules may have never been synthesized

Experimental **HIT**: Molecule that binds the receptor in solution with μM K_I

Experimental **HITS** do not need previous ADME-T studies

Lead Molecule

What is a LEAD molecule?

Lead Molecule

What is a LEAD molecule?

Molecule that binds the receptor in a cell culture with nM K_I

Lead Molecule

What is a LEAD molecule?

Molecule that binds the receptor in a cell culture with nM K_I

LEADS need to be synthesized

Lead Molecule

What is a LEAD molecule?

Molecule that binds the receptor in a cell culture with nM K_I

LEADS need to be synthesized

LEADS must show good ADME-T properties *in vitro* in cell cultures

Lead Molecule

What is a LEAD molecule?

Molecule that binds the receptor in a cell culture with nM K_I

LEADS need to be synthesized

LEADS must show good ADME-T properties *in vitro* in cell cultures

LEADS need to be patentable

Lead Molecule

What is a LEAD molecule?

Molecule that binds the receptor in a cell culture with nM K_I

LEADS need to be synthesized

LEADS must show good ADME-T properties *in vitro* in cell cultures

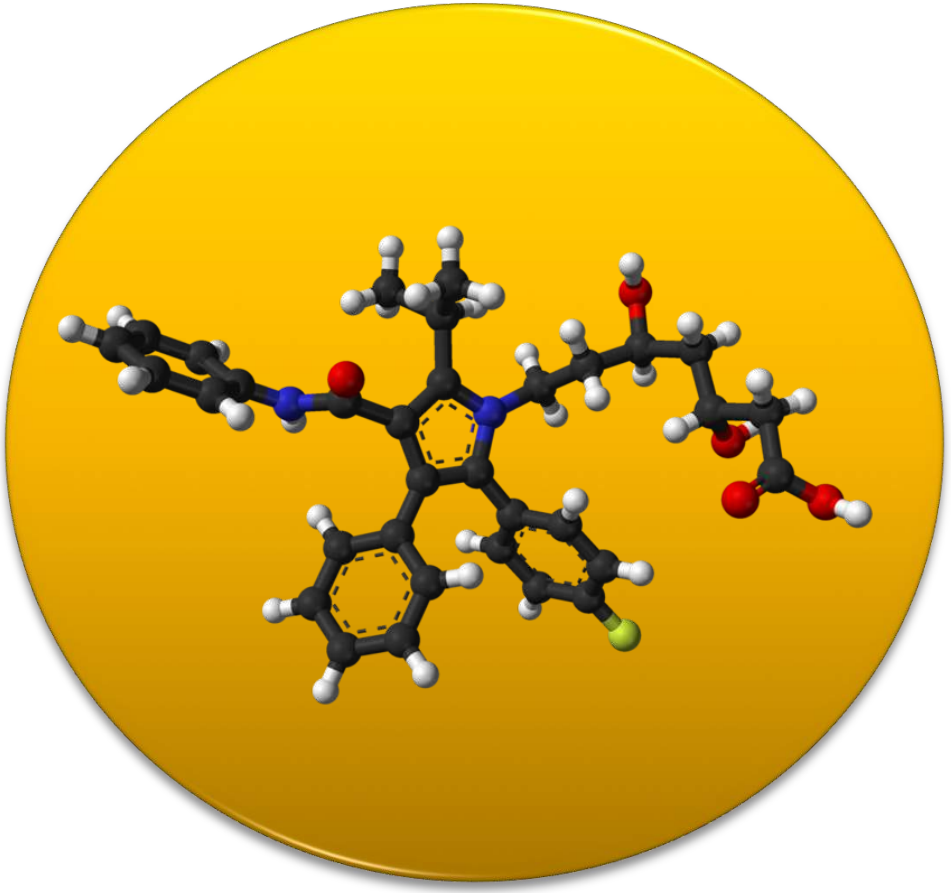
LEADS need to be patentable

Only experiments can validate **LEADS**

Find a Hit Through Virtual Screening

SBDD

First step: finding a **HIT**



SBDD

First step: finding a **HIT**

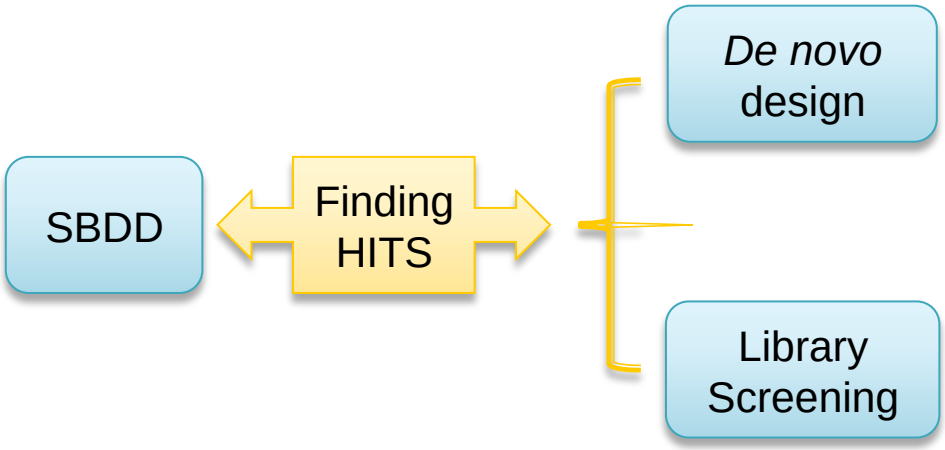
Virtual Screening of Chemical Libraries

The objective of a Virtual Screening Campaign is to find **NEW PHARMACOPHORES** and not **BETTER VARIANTS** of a known pharmacophore

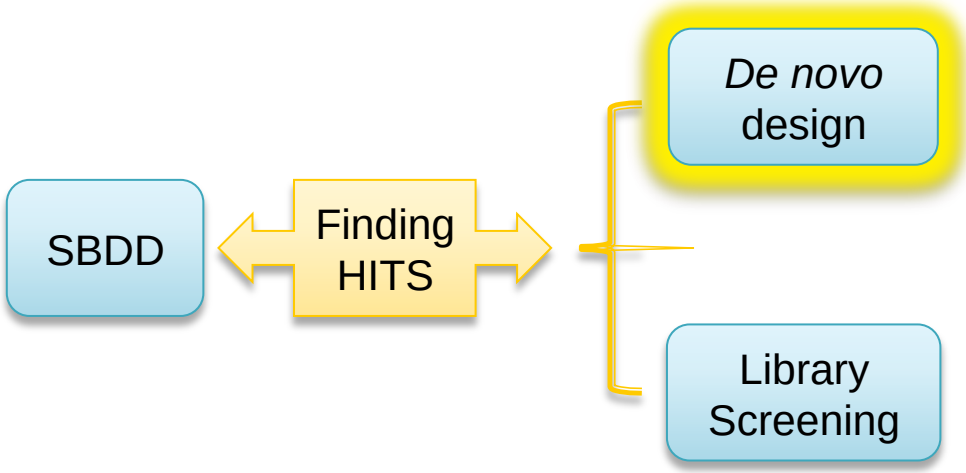
Virtual Screening of Chemical Libraries

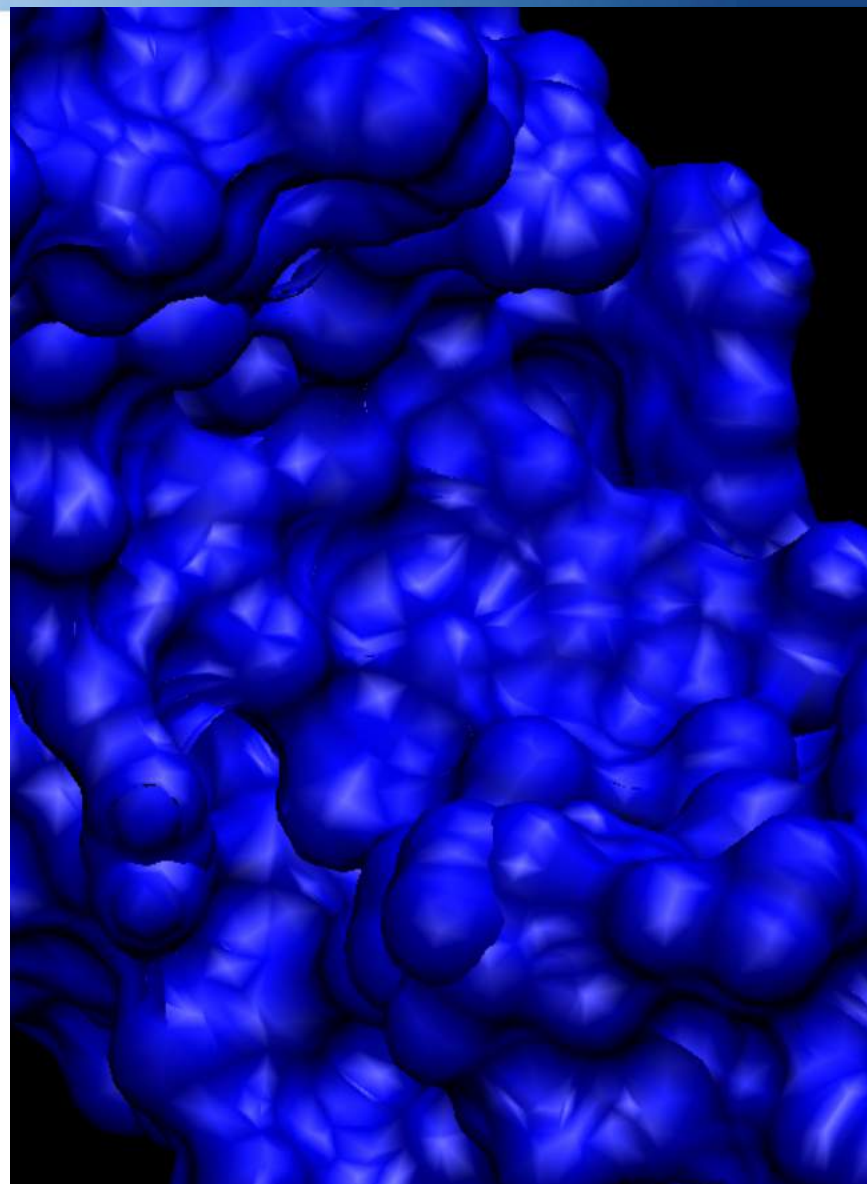
In a Virtual Screening Campaign we must be ready to tolerate FALSE POSITIVES and accept FALSE NEGATIVES in exchange of finding a new pharmacophore

Virtual Screening



Virtual Screening by *de novo* Design





Virtual Screening by *de novo* Design

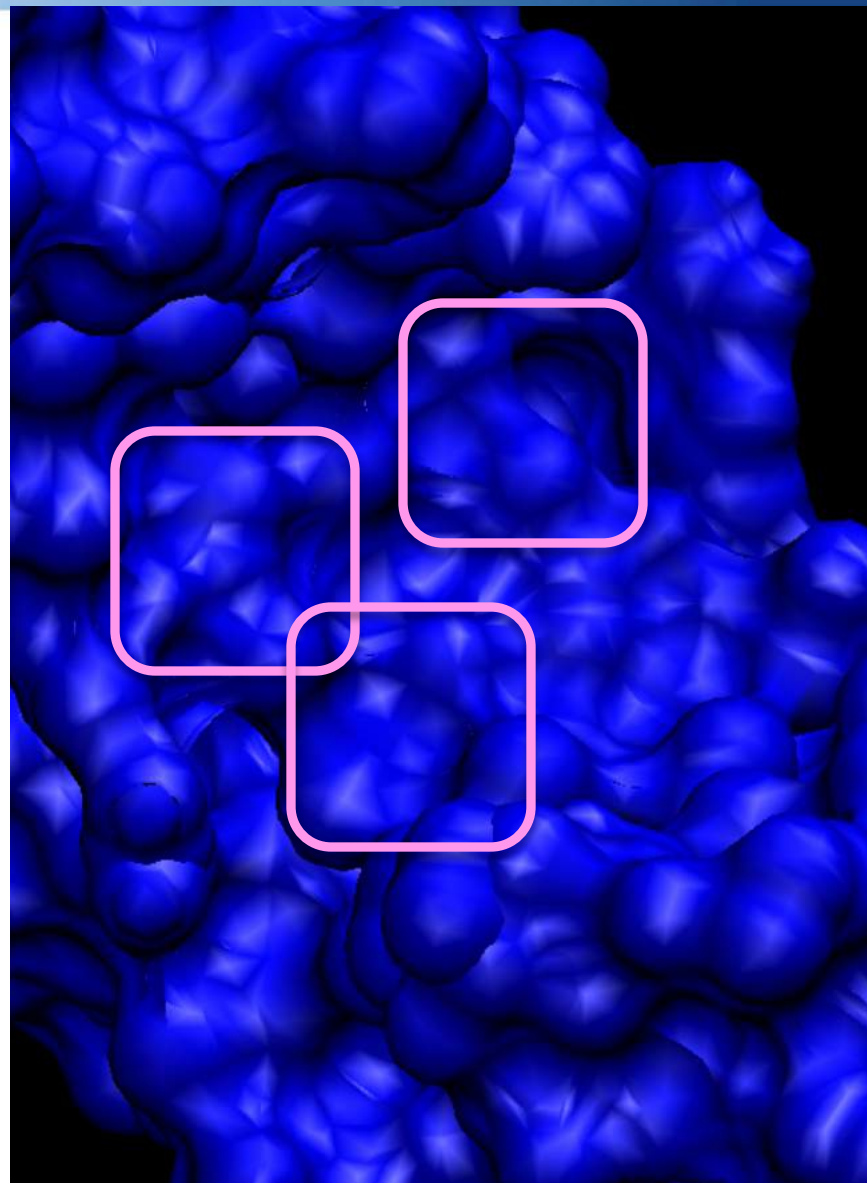
SBDD

Finding HITS

De novo design

Library Screening

Search for small, druggable pockets



Virtual Screening by *de novo* Design

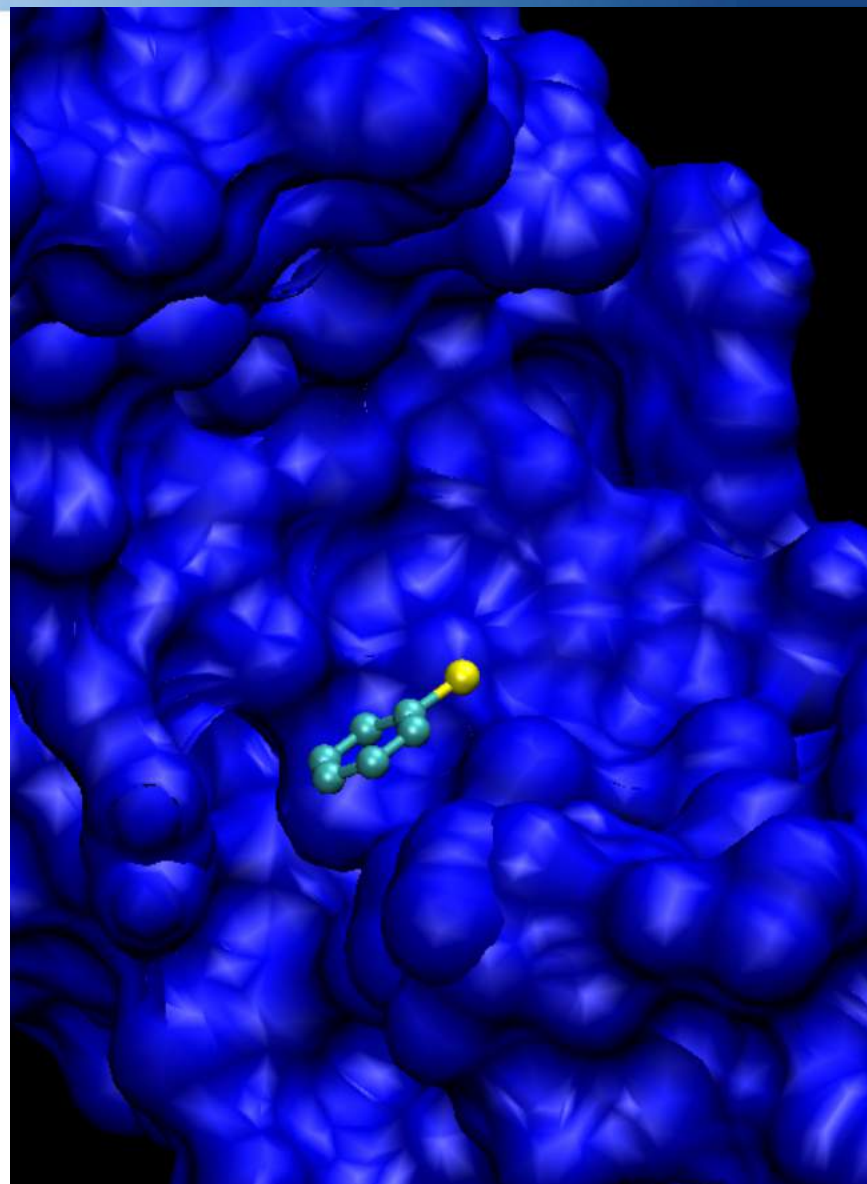
SBDD

Finding HITS

De novo design

Library Screening

Use a **large fragment library** to search for the **fragments** that best **bind the pocket**



Virtual Screening by *de novo* Design

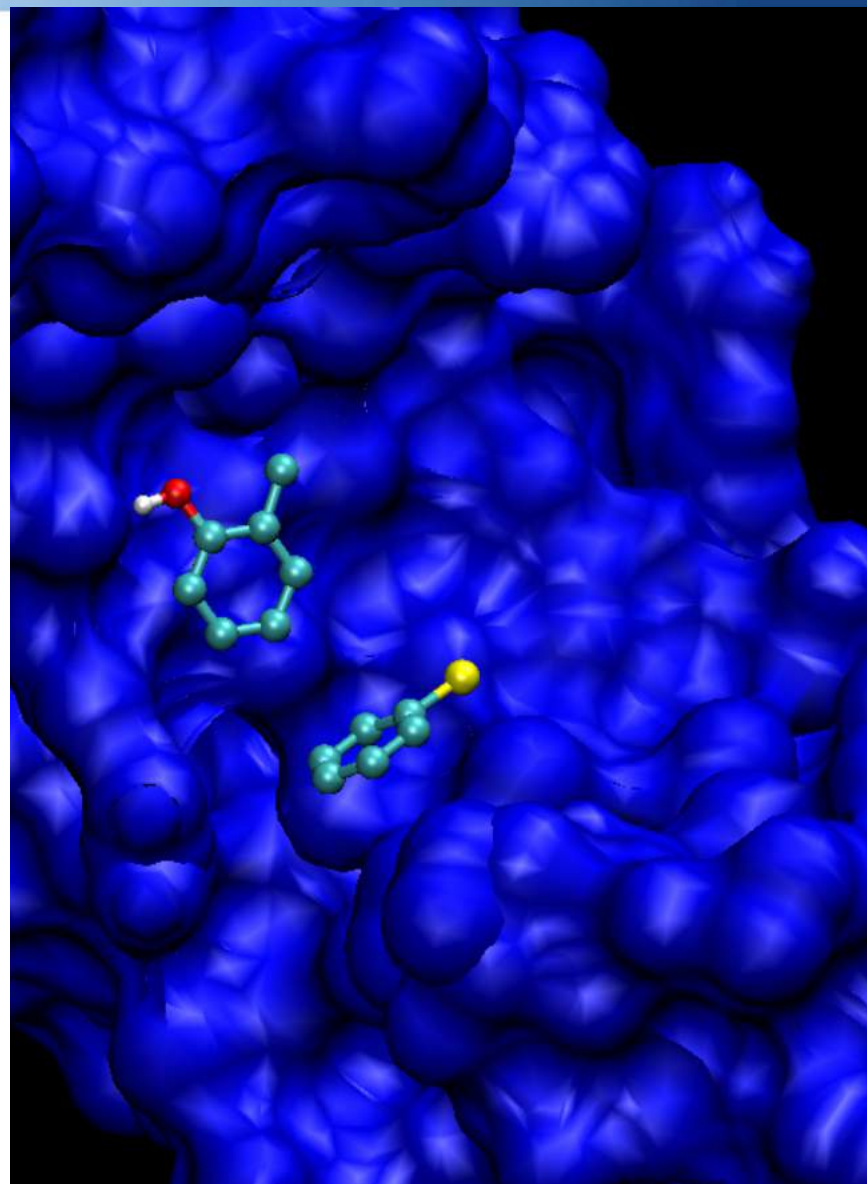
SBDD

Finding HITS

De novo design

Library Screening

Use a **large fragment library** to search for the **fragments** that best **bind the pocket**



Virtual Screening by *de novo* Design

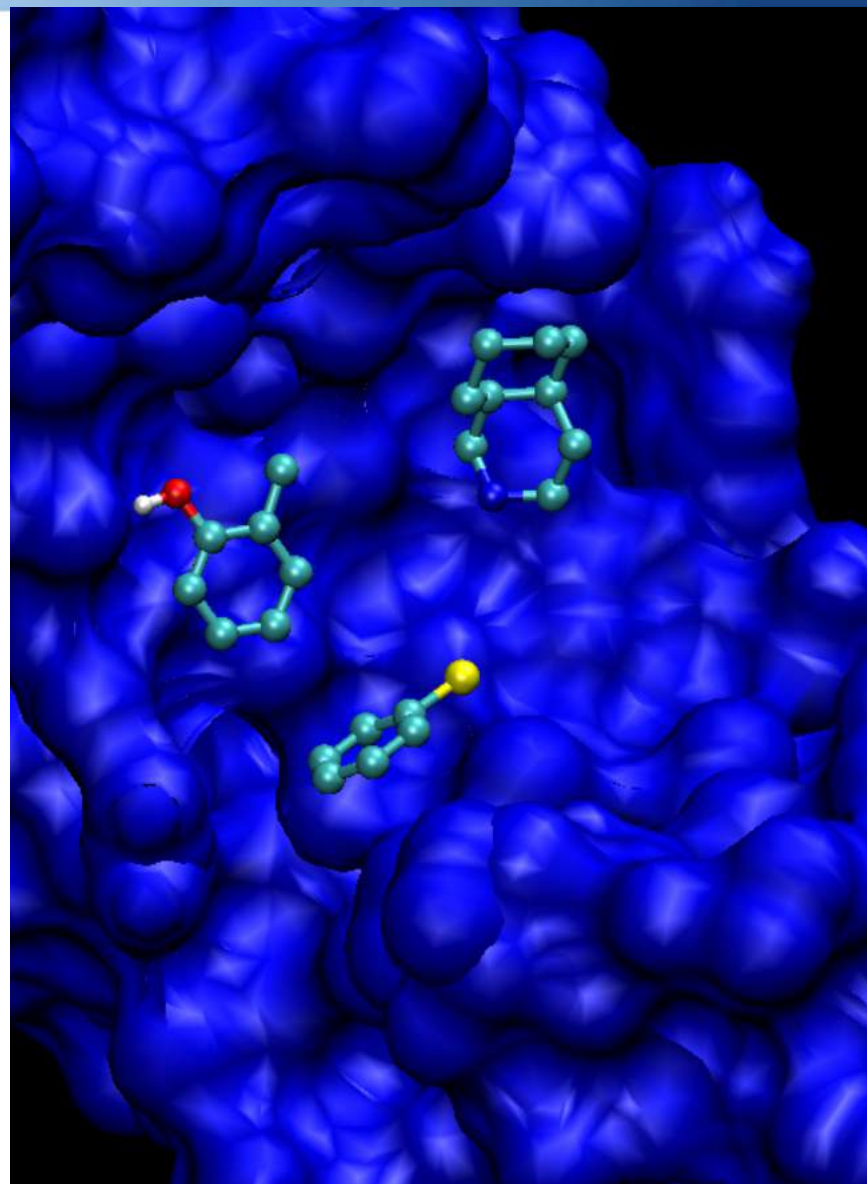
SBDD

Finding HITS

De novo design

Library Screening

Use a **large fragment library** to search for the **fragments** that best **bind the pocket**



Virtual Screening by *de novo* Design

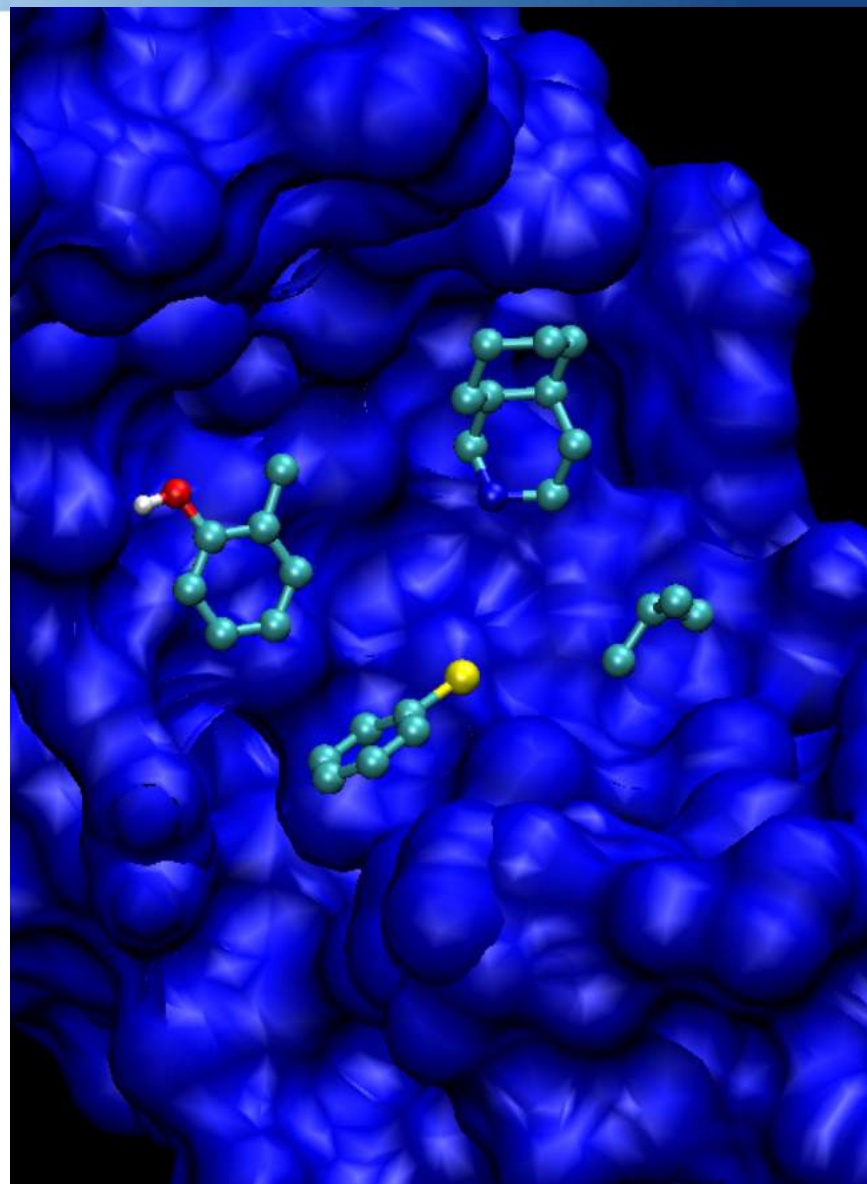
SBDD

Finding HITS

De novo design

Library Screening

Use a **large fragment library** to search for the **fragments** that best **bind the pocket**



Virtual Screening by *de novo* Design

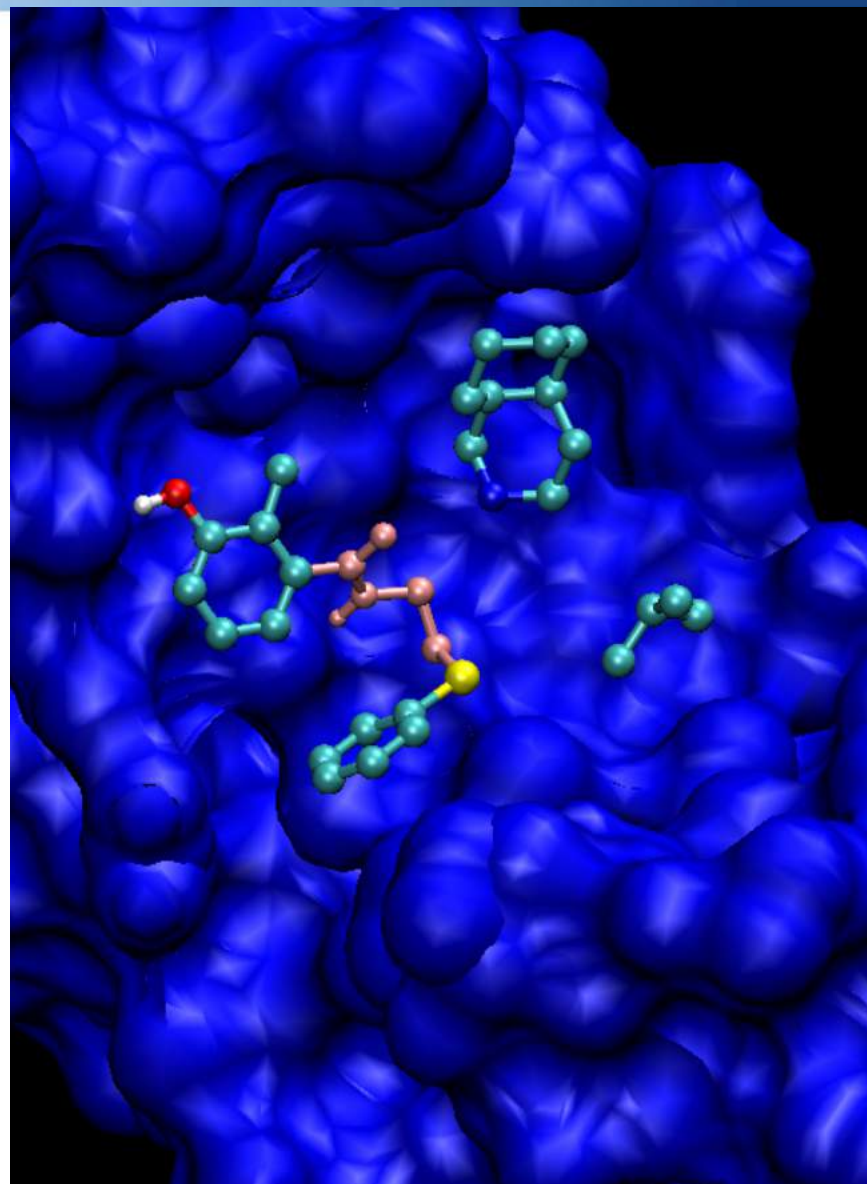
SBDD

Finding HITS

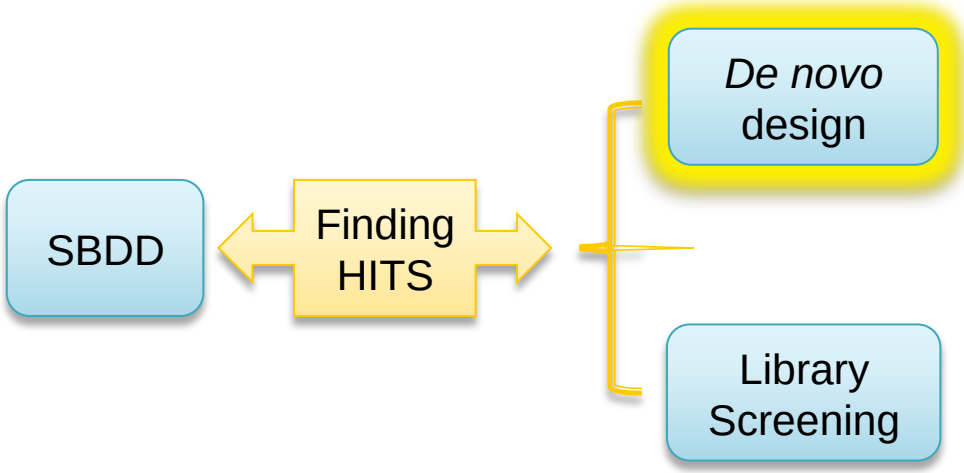
De novo design

Library Screening

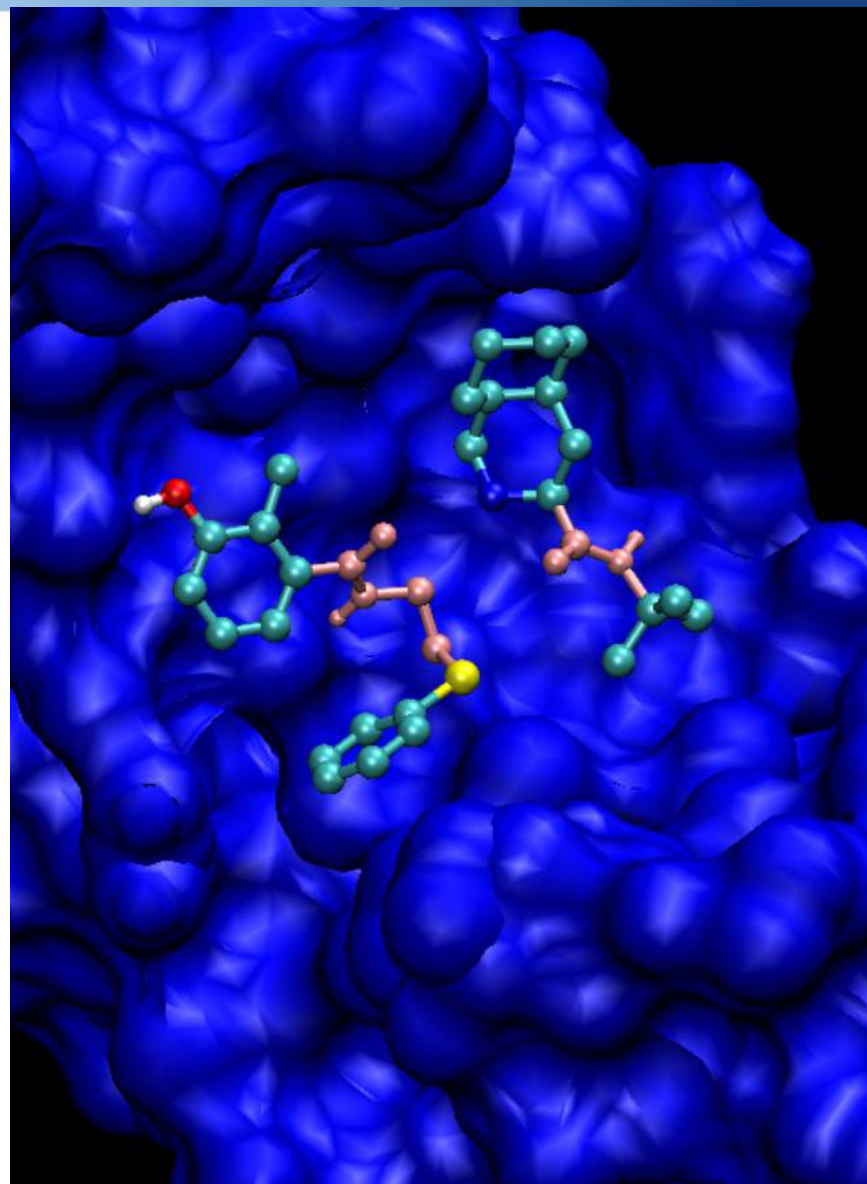
Use a **large link library** to search for the **links** that best **connect the fragments**



Virtual Screening by *de novo* Design



Use a **large link library** to search for the **links** that best **connect the fragments**



Virtual Screening by *de novo* Design

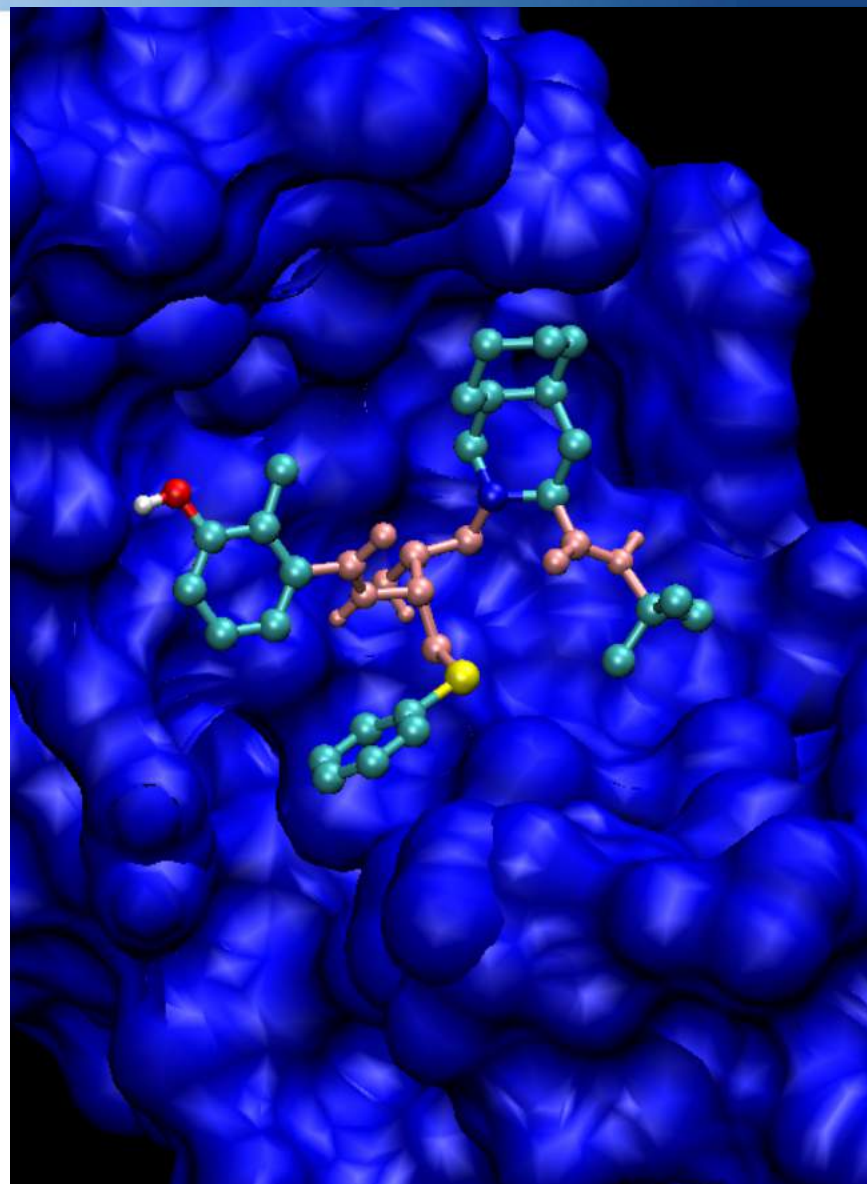
SBDD

Finding HITS

De novo design

Library Screening

Use a **large link library** to search for the **links** that best **connect the fragments**



Virtual Screening by *de novo* Design

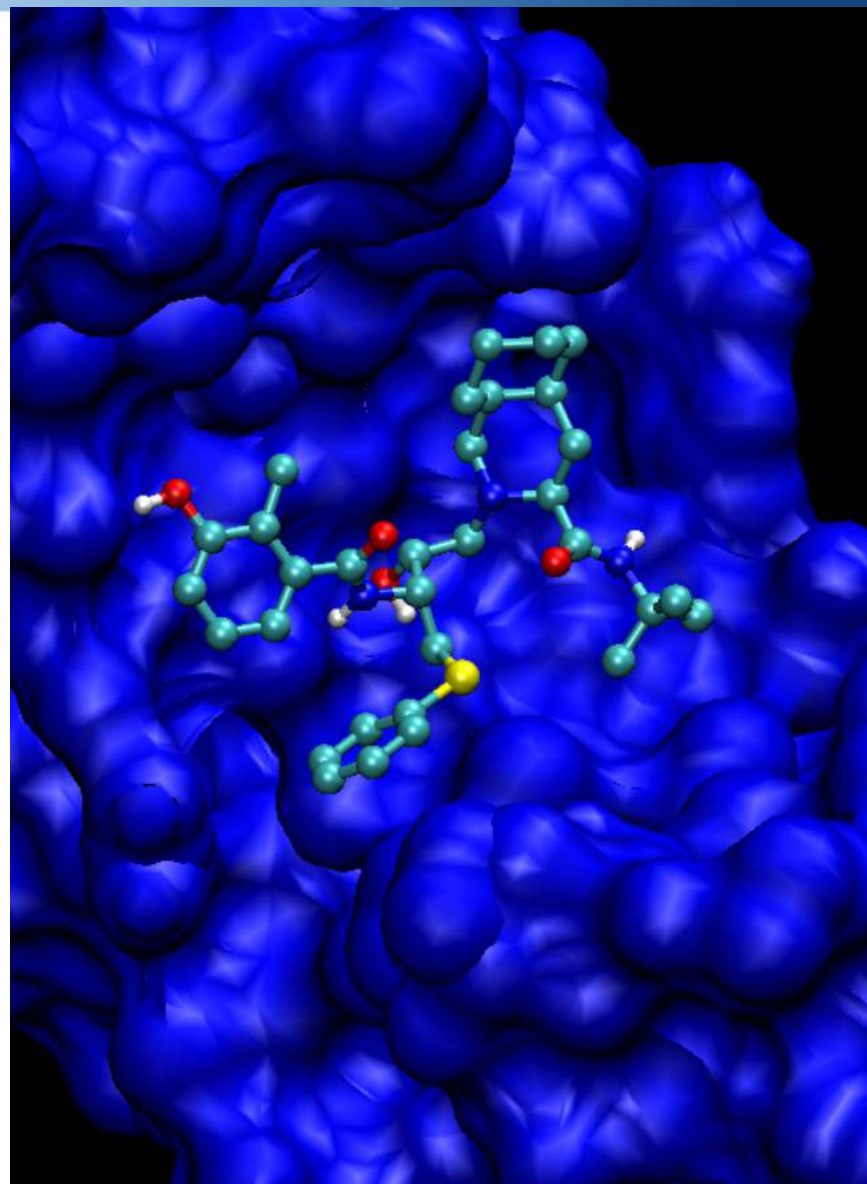
SBDD

Finding HITS

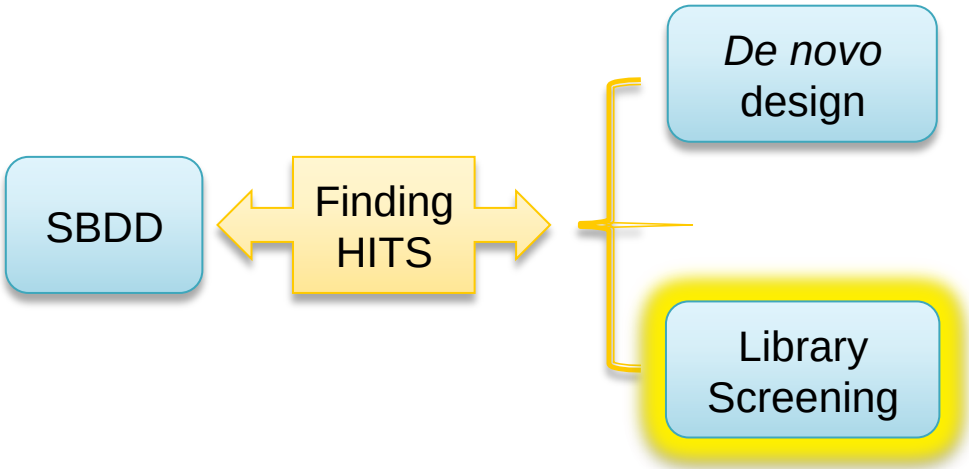
De novo design

Library Screening

De novo molecule!
Hit?



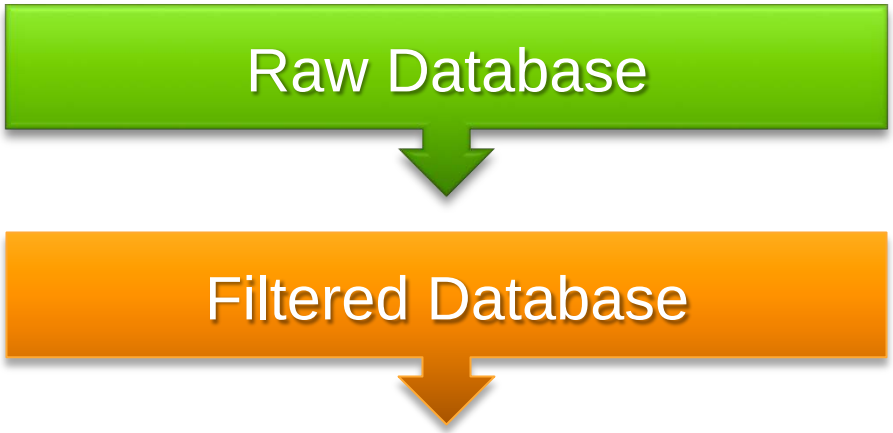
Strutcure Based Drug Discovery



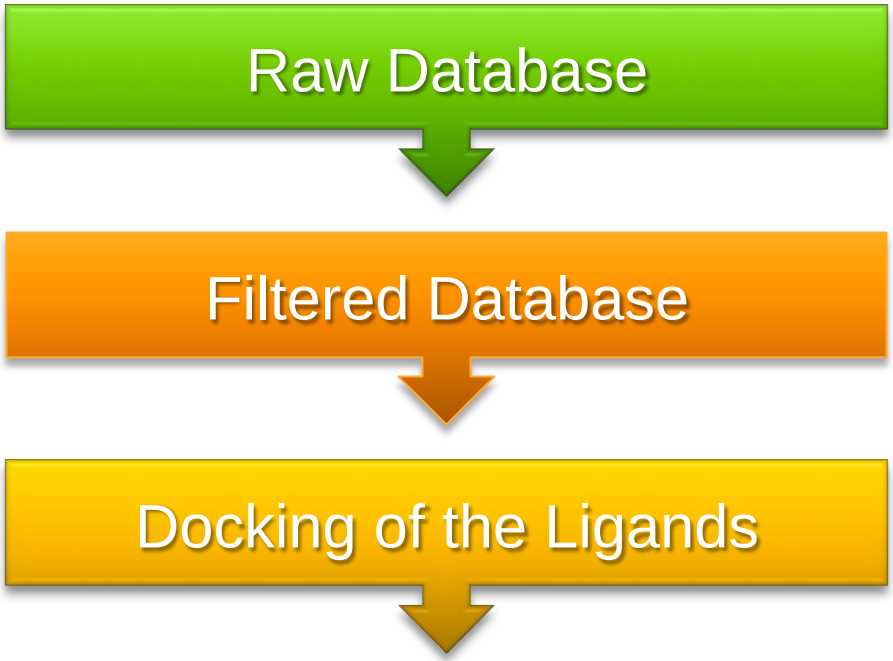
Virtual Screening of Chemical Libraries

Virtual Screening of Chemical Libraries

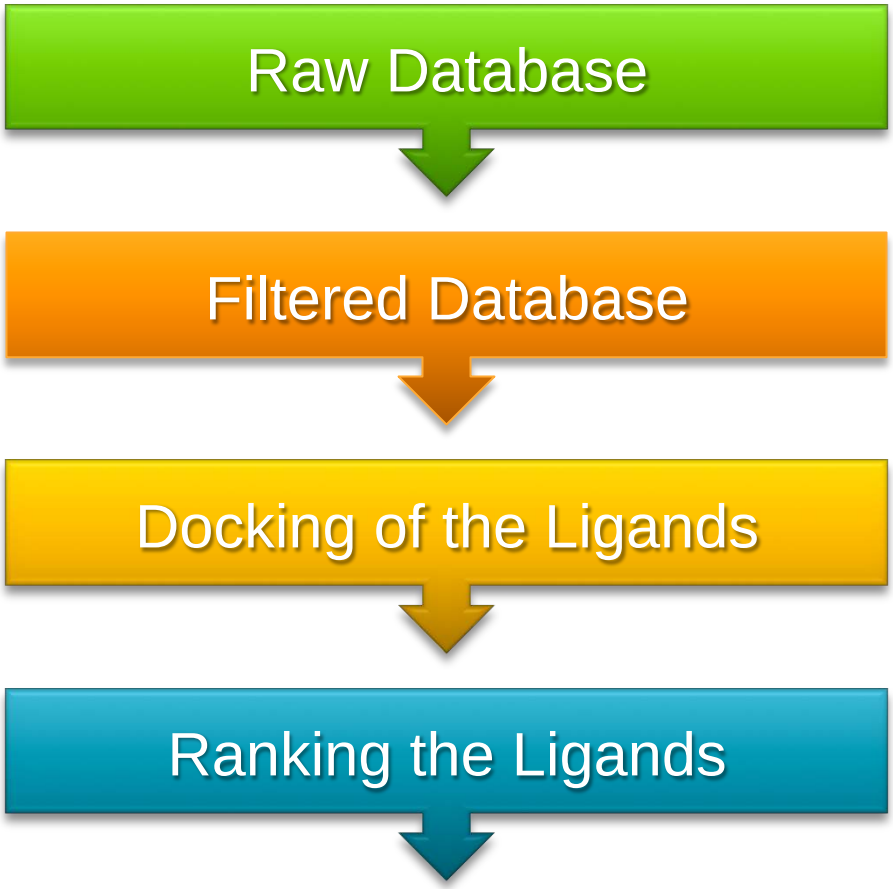
Virtual Screening of Chemical Libraries



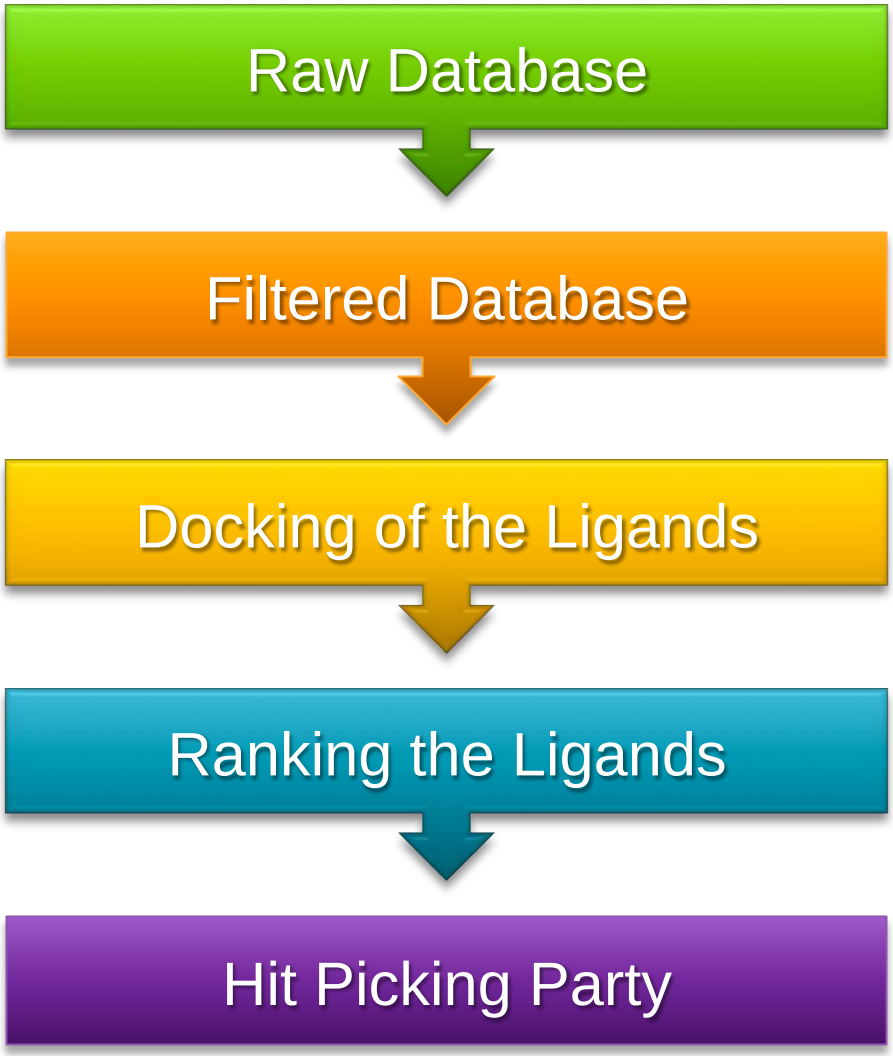
Virtual Screening of Chemical Libraries



Virtual Screening of Chemical Libraries



Virtual Screening of Chemical Libraries



Some of the Available Online Compound Libraries

databases and ref	web pages	number of molecules
Binding Database[61]	http://www.bindingdb.org	284 206 small ligands with 648 915 binding data, for 5662 protein targets
Chem ID[62]	http://chem.sis.nlm.nih.gov/chemidplus/	388 000
ChemBank[63]	http://chembank.broadinstitute.org	800 000
ChEMBL db[64]	https://www.ebi.ac.uk/chembldb/	658 075 differing bioactive compounds and 8091 targets
Chemical Entities of Biological Interest (ChEBI)[65]	http://www.ebi.ac.uk/chebi/init.do	584 456
ChemMine[66]	http://bioweb.ucr.edu/ChemMineV2/	6 200 000
Chimiotheque nationale[67]	http://chimiotheque-nationale.enscm.fr/index.php	44 817 compounds, 32 573 compounds in plate, 14 514 natural extracts
Commercial Compound Collection (CoCoCo)[68]	http://cococo.unimore.it/tiki-index.php	6 957 134 molecules, more than 144 millions conformations
Developmental Therapeutics Program (DTP)[69]	http://dtp.nci.nih.gov Downloads: http://cactus.nci.nih.gov/	473 965
DrugBank[70]	http://www.drugbank.ca	6827 drugs, 4477 nonredundant protein sequences
GVK BIO[71]	http://www.gvkbio.com/informatics.html	not specified (focused libraries with target inhibitor or toxicity collections applied in the field of bio- and chemo-informatics)
i:lib diverse[72]	http://www.inteligand.com/	drug-like fragment set for combinatorial library generation
Mother of All Databases (MOAD)[73]	http://www.bindingmoad.org	14 720 ligand-protein complexes, 4782 structures with binding data, 7064 ligands
PDB bind[74,75]	http://www.pdbbind.org/	3214 ligand−protein complexes
PubChem[76]	http://pubchem.ncbi.nlm.nih.gov/	49 875 000
Therapeutic Target Database[77,78]	http://bidd.nus.edu.sg/group/cjttd/TTD_HOME.asp	1906 targets, 5124 drugs
Traditional Chinese Medicine Database (TCM)[79]	http://tcm.cmu.edu.tw/	more than 20 000 pure compounds isolated from 453 TCM ingredients
WOMBAT[80]	http://www.sunsetmolecular.com/	305 727 molecules, 1966 unique targets
ZINC[81]	http://zinc.docking.org/	13 000 000

Zinc Database

UCSF University of California, San Francisco | About UCSF | Search UCSF | UCSF Medical Center

Shoichet Laboratory docking.org

ZINC 12

Not Authenticated – sign in

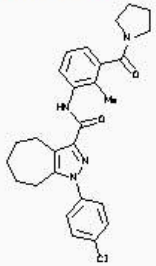
Active cart: Temporary Cart (0 items)

About Search Subsets Help Social g+1 77

Quick Search Bar... Go

Welcome to ZINC, a free database of commercially-available compounds for virtual screening. ZINC contains over 35 million purchasable compounds in ready-to-dock, 3D formats. ZINC is provided by the Shoichet Laboratory in the Department of Pharmaceutical Chemistry at the University of California, San Francisco (UCSF). To cite ZINC, please reference: Irwin, Sterling, Mysinger, Bolstad and Coleman, *J. Chem. Inf. Model. 2012* DOI: 10.1021/ci3001277. The original publication is Irwin and Shoichet, *J. Chem. Inf. Model. 2005;45(1):177-82* PDF, DOI. We thank NIGMS for financial support (GM71896).

Molecule of the Year 72052718



ZINC ID, Drug Name, SMILES, Catalog, Vendor Code, Go

Structure/Draw Physical Properties Catalogs & Vendors ZINC IDS Targets Rings Combination

What's NEW? Feedback Like us
@chem4biology Blog RSS
Video Walkthroughs

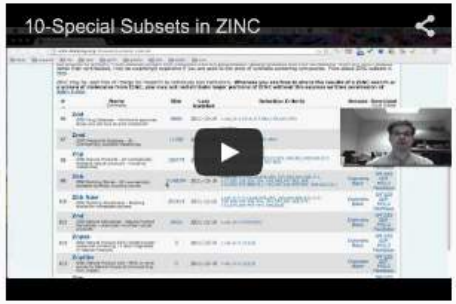
ZINC Database
Like 640

Quick Links

Download
Target focused
Natural Products
Search By Target
Rings
Search
Thanks
Special Subsets
PBCs
Carts

Your Carts

Create an account or login to have multiple carts.



Bioinformatics and Chemical Informatics Research Center (BCIRC) Terms of use Privacy policy Questions, Discussion, Bug reports, Feature requests Thank you NIGMS! GM71896 Go Secure

Zinc Database

ZINC 12

Physical Property Subsets

Ready-to-download subsets of ZINC filtered by physical properties are available below. All molecules in all of these subsets are purchasable. To go to the download page, click on the subset name. To browse sample compounds, click on the number of compounds. Three versions of subsets are available:

Standard: For delivery in 0-10 weeks, including in-stock, make-on-demand and agent compounds
Clean: Stricter filtering rules have been applied – e.g. aldehyes and thiols have been removed
In Stock: For immediate delivery – typically less than 2 weeks
Boutique: Boutique compounds – you should be willing to "pay any price". These are NOT included in standard, clean and in stock subsets, where the target price is $100 or less per sample

More about ZINC subsets is here. ZINC may be used free of charge for research by individuals and institutions. **Whereas you are free to share the results of a ZINC search or a screen of molecules from ZINC, you may not redistribute major portions of ZINC without the express written permission of John Irwin.**

	Lead-Like	Fragment-Like	Drug-Like	All	Shards
Standard Size Updated	**Lead-Like** 6,053,287 2014-09-29	**Fragment-Like** 847,909 2015-02-04	**Drug-Like** 17,900,742 2014-11-24	**All Purchasable** 22,724,825 2014-11-28	**Shards** 635,159 2014-05-16
Clean Size Updated	**Clean Leads** 4,591,276 2014-09-25	**Clean Fragments** 1,611,889 2014-09-24	**Clean Drug-Like** 13,195,609 2013-11-05	**All Clean** 16,403,865 2013-12-18	**Clean Shards** 325,950 2014-11-24
In Stock Size Updated	**Leads Now** 3,687,621 2014-06-25	**Frags Now** 704,041 2015-02-04	**Drugs Now** 10,639,555 2014-11-24	**All Now** 12,782,590 2014-05-01	**Shards Now** 424,775 2014-09-24
Boutique Size Updated	**Boutique Leads** 5,114,169 2012-12-24	**Boutique Frags** 2,755,555 2013-11-08	**Boutique Drugs** 10,292,210 2012-11-27	**All Boutique** 12,217,845 2012-11-27	**Boutique Shards** 80,698 2013-11-08
Comments/Citation	Teague, Davis, Leeson, Oprea, Angew Chem Int Ed Engl. 1999 Dec 16;38(24):3743-3748.	Carr RA, Congreve M, Murray CW, Rees DC, Drug Discov Today. 2005 Jul 15;10(14):987	Lipinski, J Pharmacol Toxicol Methods. 2000 Jul-Aug;44(1):235-49.	Purchasable chemical space	Type I binding sites

Filtering the Databases

Filtering is necessary to allow for a better evaluation of the candidates afterwards

Filtering the Databases

Filtering is necessary to allow for a better evaluation of the candidates afterwards

Large number of candidates are needed to find good hits

Filtering the Databases

Filtering is necessary to allow for a better evaluation of the candidates afterwards

Large number of candidates are needed to find good hits

The larger the number of candidates the worse the predictions will be

Filtering the Databases

Filtering is necessary to allow for a better evaluation of the candidates afterwards

Large number of candidates are needed to find good hits

The larger the number of candidates the worse the predictions will be

Filtering the Databases

How to filter?

Filtering the Databases

How to filter?

ADME descriptors?

Filtering the Databases

How to filter?

ADME descriptors?

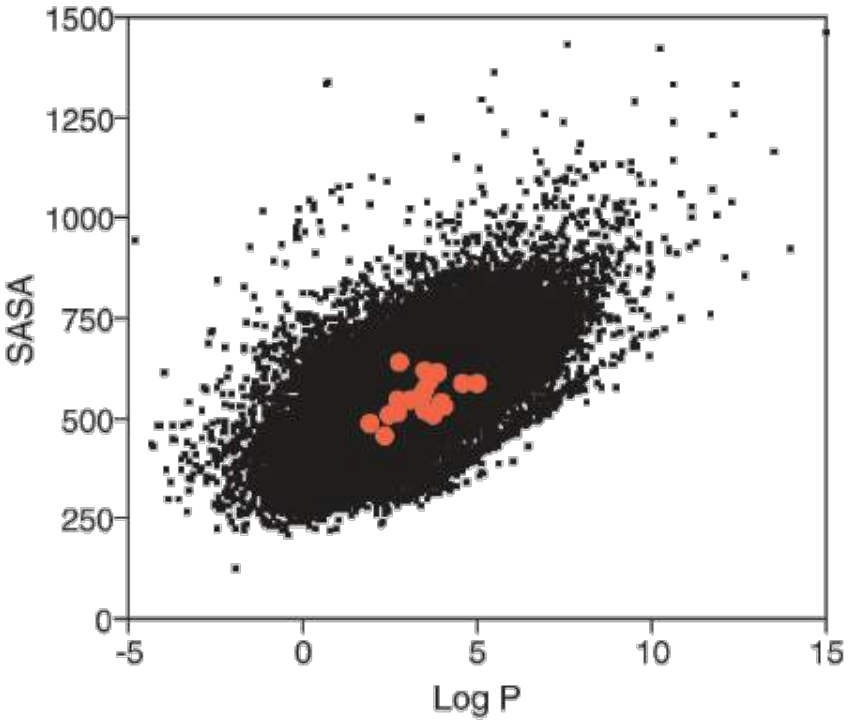


Fig. 1. Plot of solvent accessible surface area (Å^2) versus log P, as computed by QikProp (*27*), for the >70,000 compounds in the 2001 Maybridge catalog with the addition of 20 non-nucleoside inhibitors of HIV-1 reverse transcriptase (NNRTIs) in red.

Filtering the Databases

How to filter?

ADME descriptors?

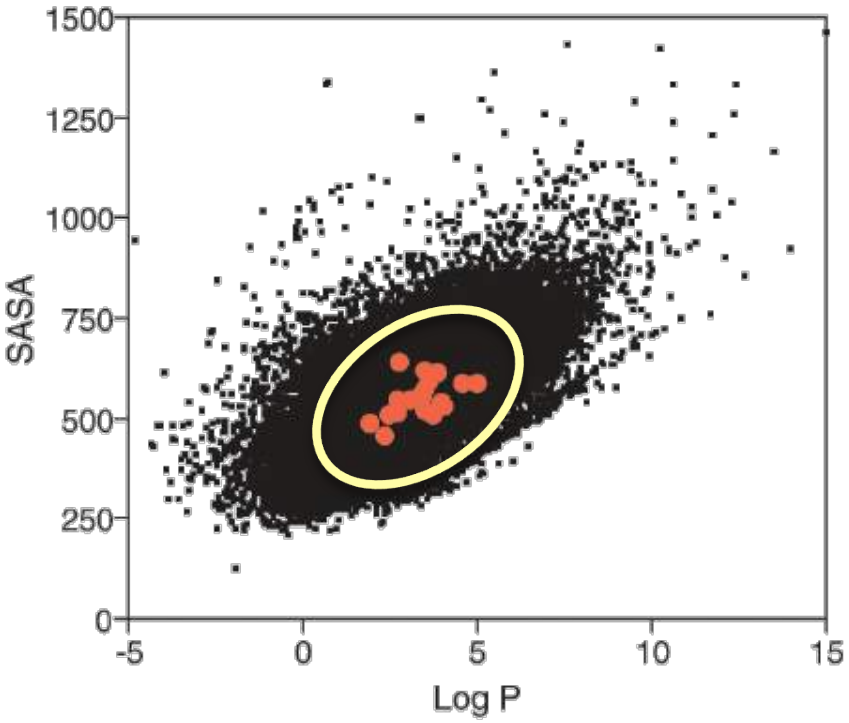


Fig. 1. Plot of solvent accessible surface area (Å^2) versus log P, as computed by QikProp (*27*), for the >70,000 compounds in the 2001 Maybridge catalog with the addition of 20 non-nucleoside inhibitors of HIV-1 reverse transcriptase (NNRTIs) in red.

Filtering the Databases

How to filter?

ADME descriptors?

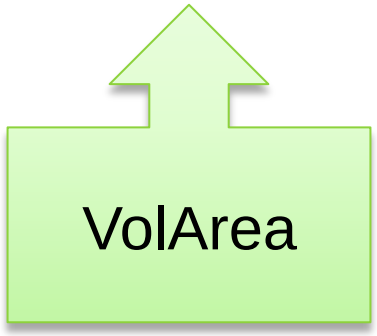
Size/Shape/Polarity?

Filtering the Databases

How to filter?

ADME descriptors?

Size/Shape/Polarity?



Filtering the Databases

How to filter?

ADME descriptors?

Size/Shape/Polarity?

Tanimoto Similarity?

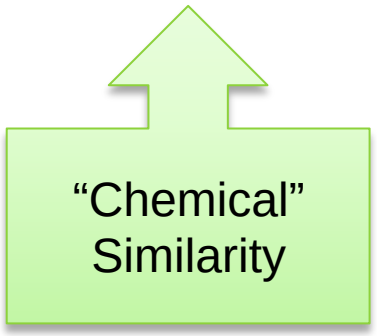
Filtering the Databases

How to filter?

ADME descriptors?

Size/Shape/Polarity?

Tanimoto Similarity?



Filtering the Databases

How to filter?

ADME descriptors?

Size/Shape/Polarity?

Tanimoto Similarity?

Filtering focus the search in known leads/pharmacophores.
It increases chance of hit detection but reduces the diversity of the ligands.

Filtering the Databases

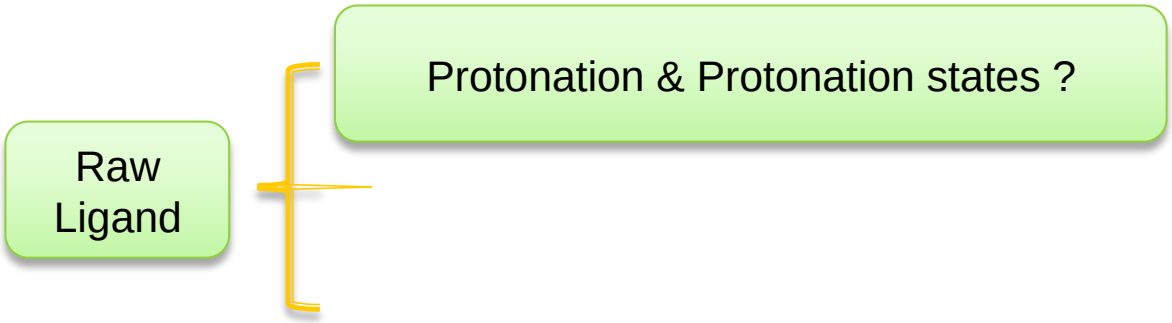
The objective of a Virtual Screening Campaign is to find **NEW PHARMACOPHORES** and not **BETTER VARIANTS** of a known pharmacophore

Preparing the Ligand and Receptor

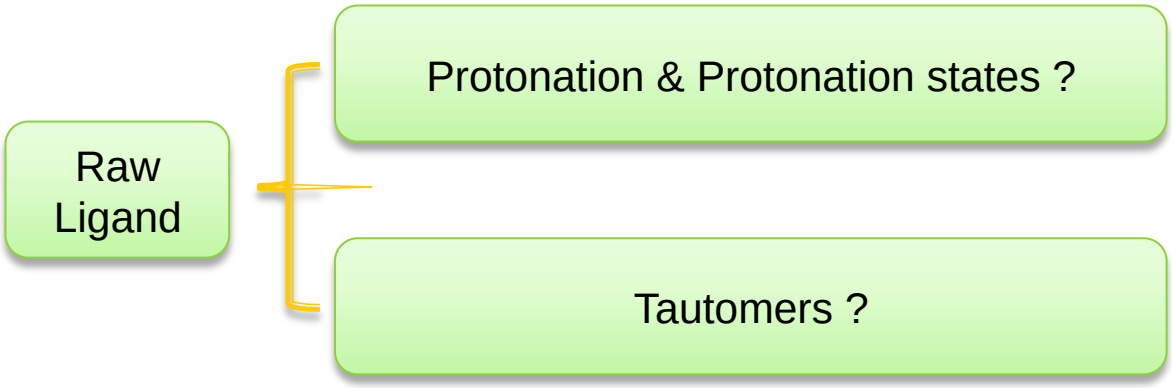
Preparing the Ligands

Raw Ligand

Preparing the Ligands



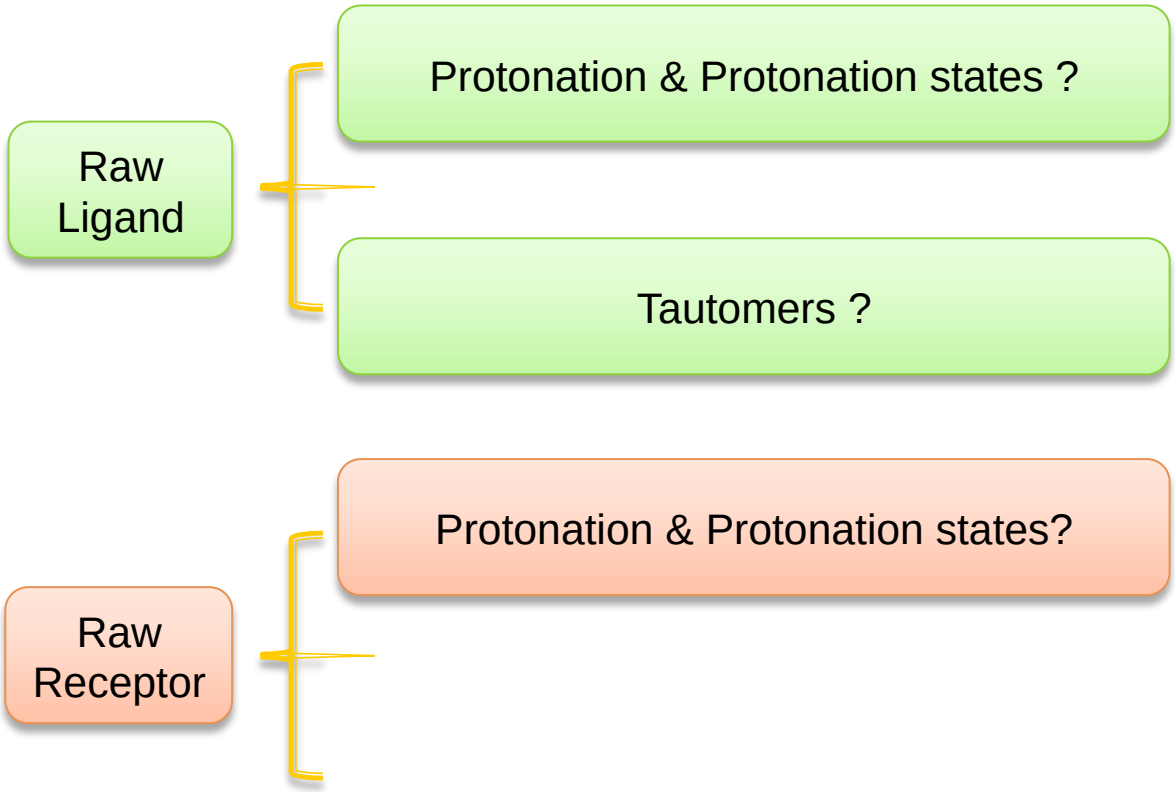
Preparing the Ligands



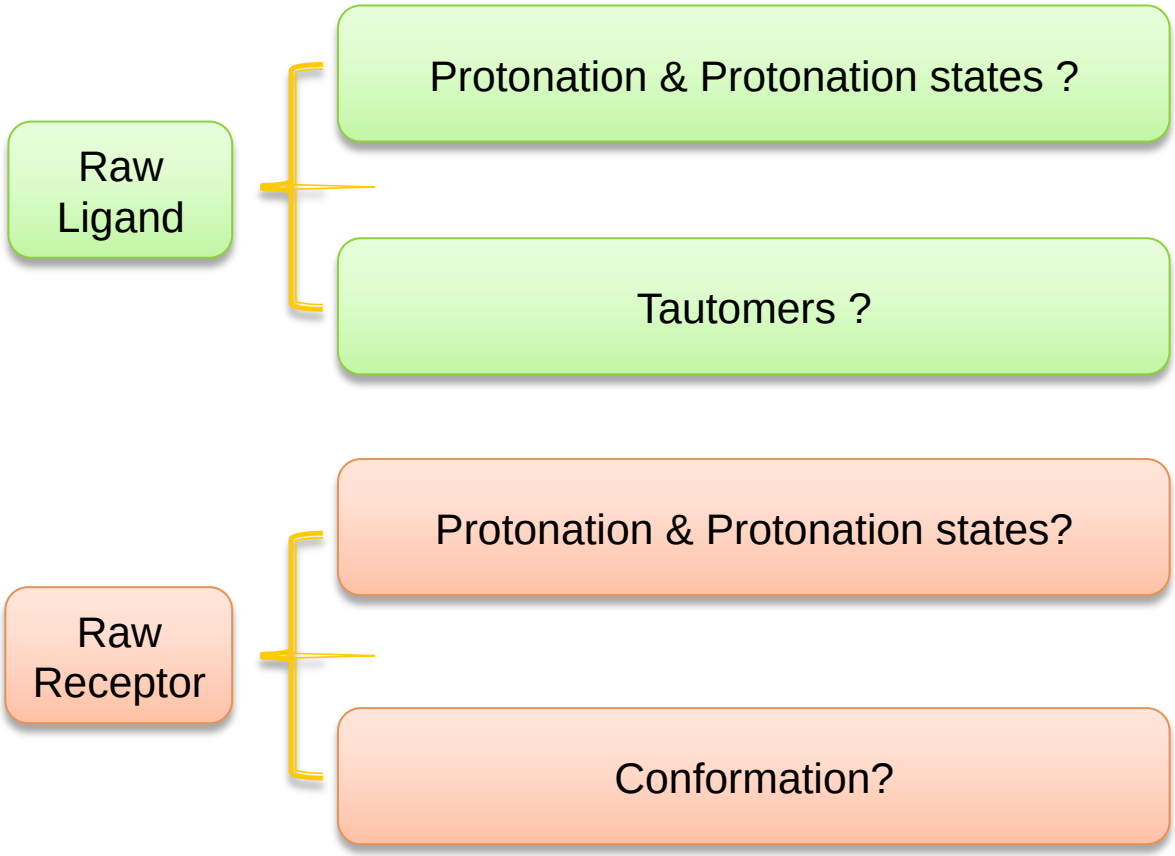
Preparing the Ligands

Raw Receptor

Preparing the Ligands



Preparing the Ligands



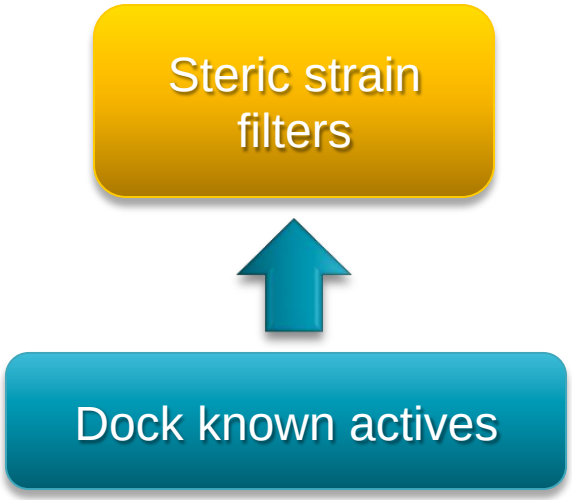
Docking the Ligands

Molecular Docking:
Finding the binding pose of a ligand inside a receptor, and estimating its affinity.
More of this on the next chapter

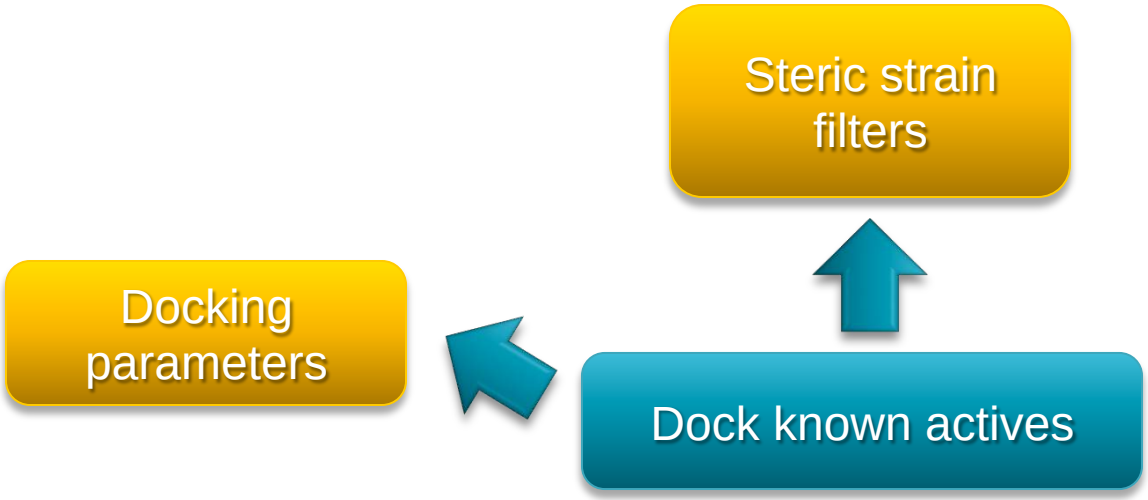
Tuning the docking method

Dock known actives

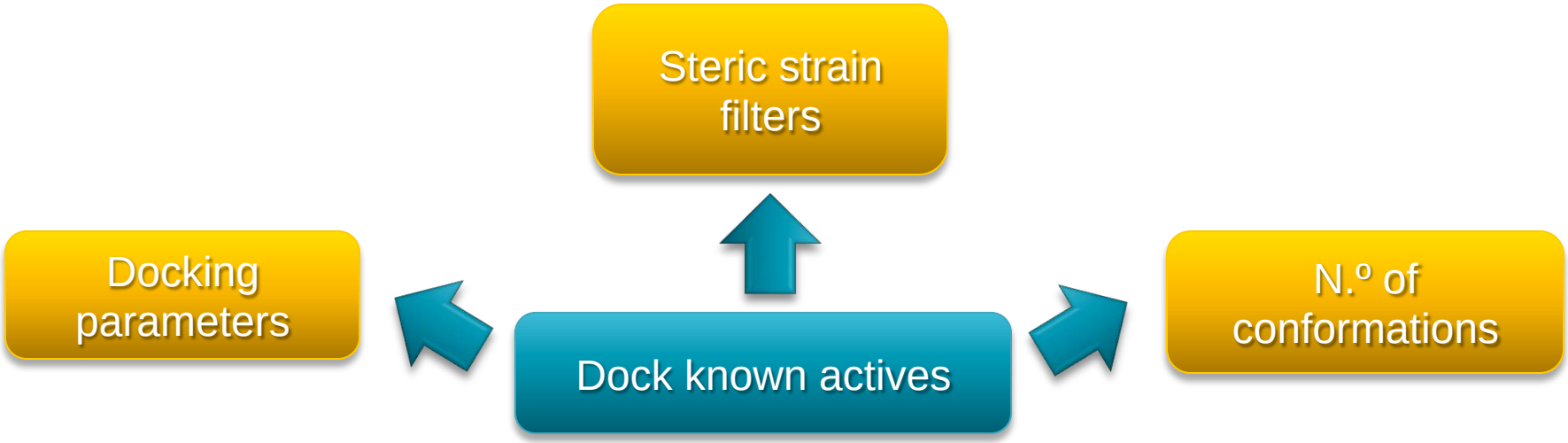
Tuning the docking method



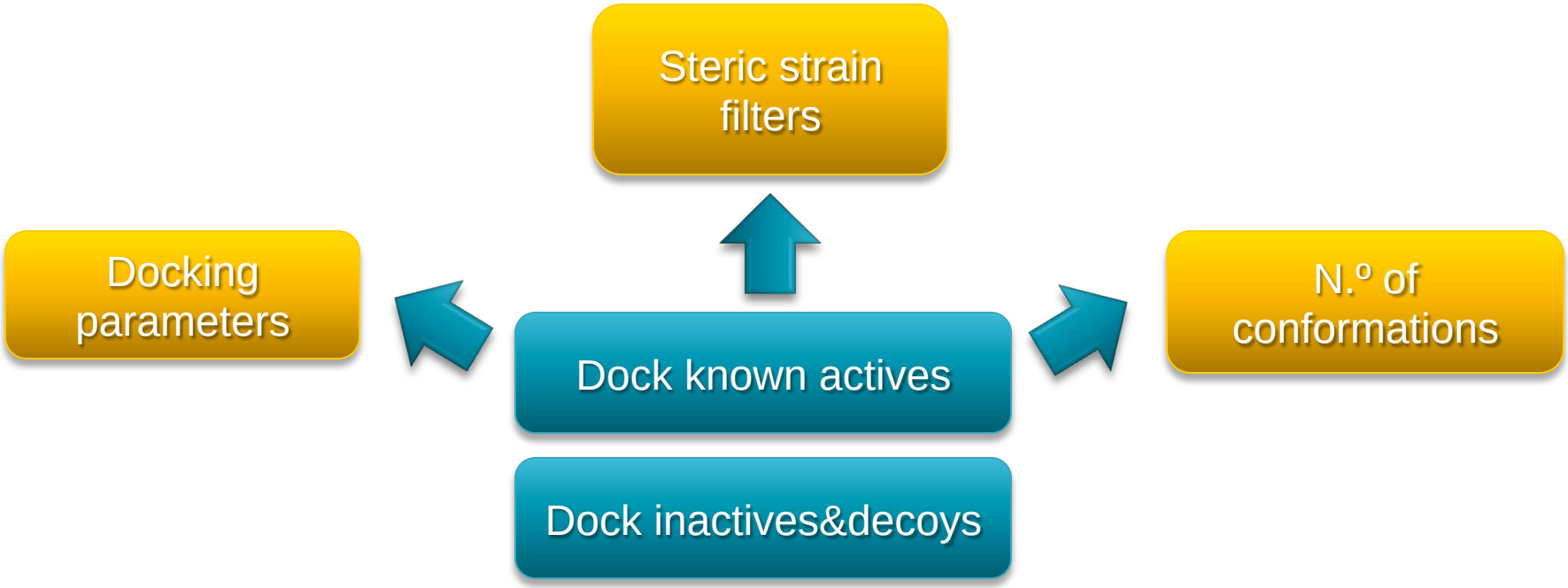
Tuning the docking method



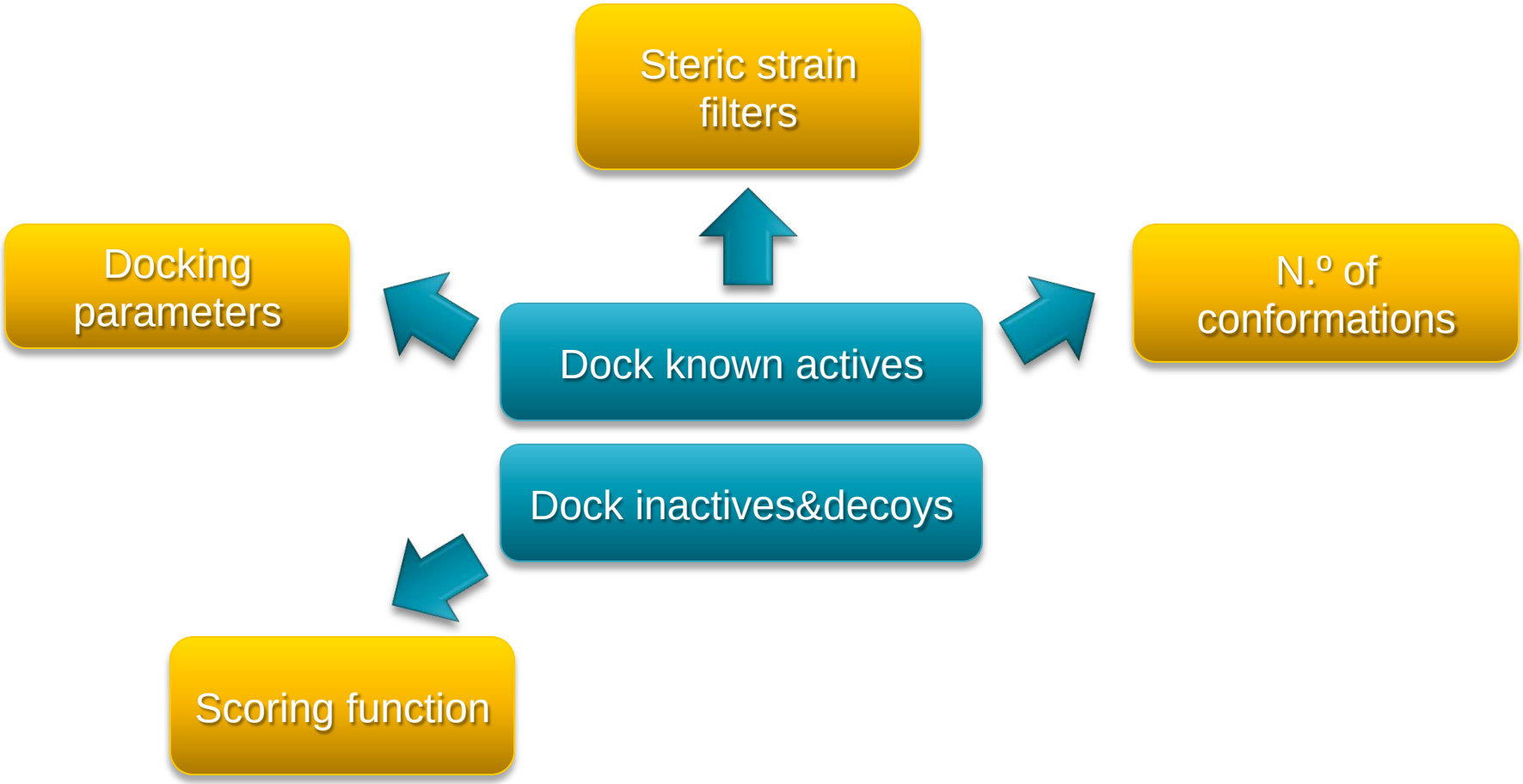
Tuning the docking method



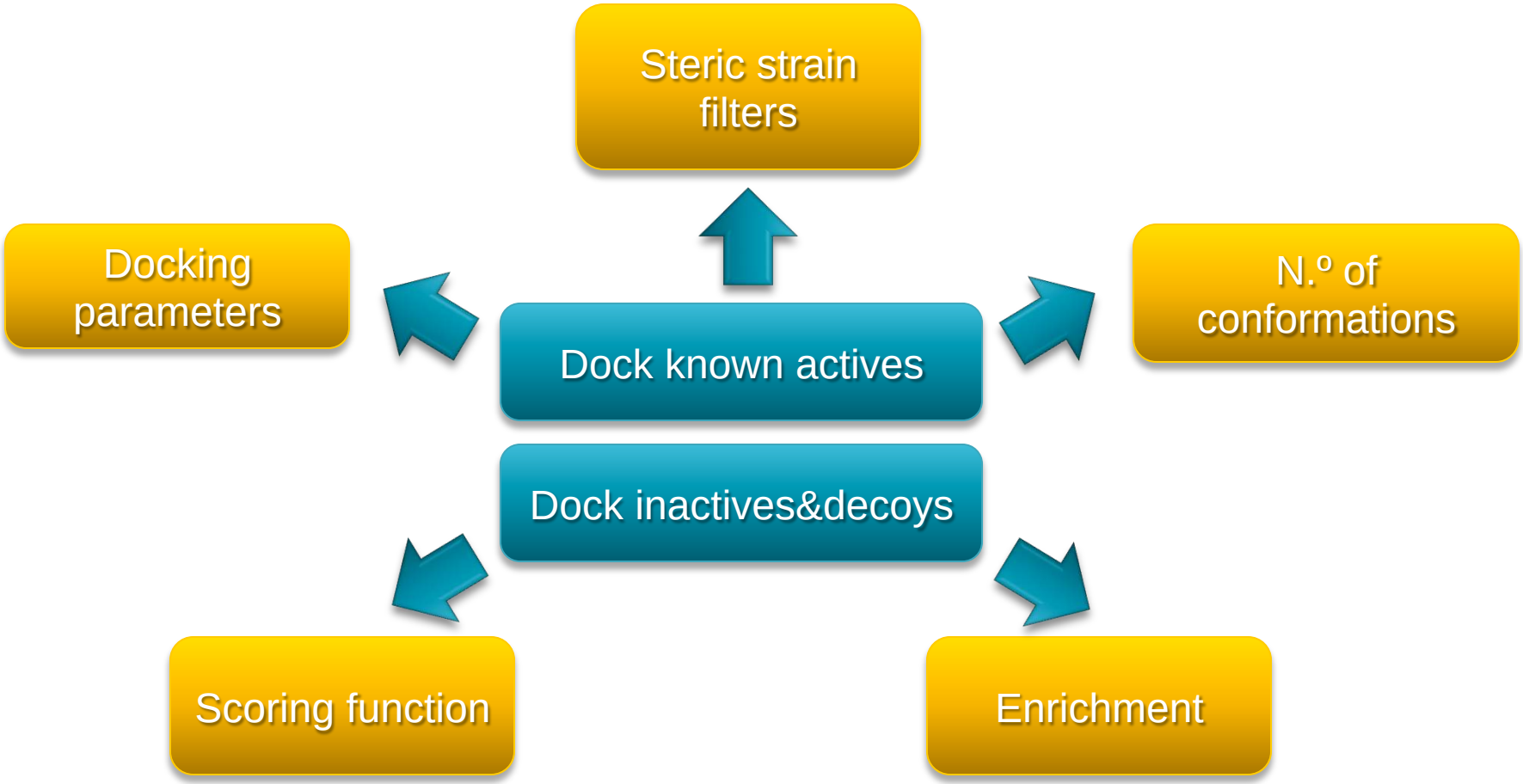
Tuning the docking method



Tuning the docking method

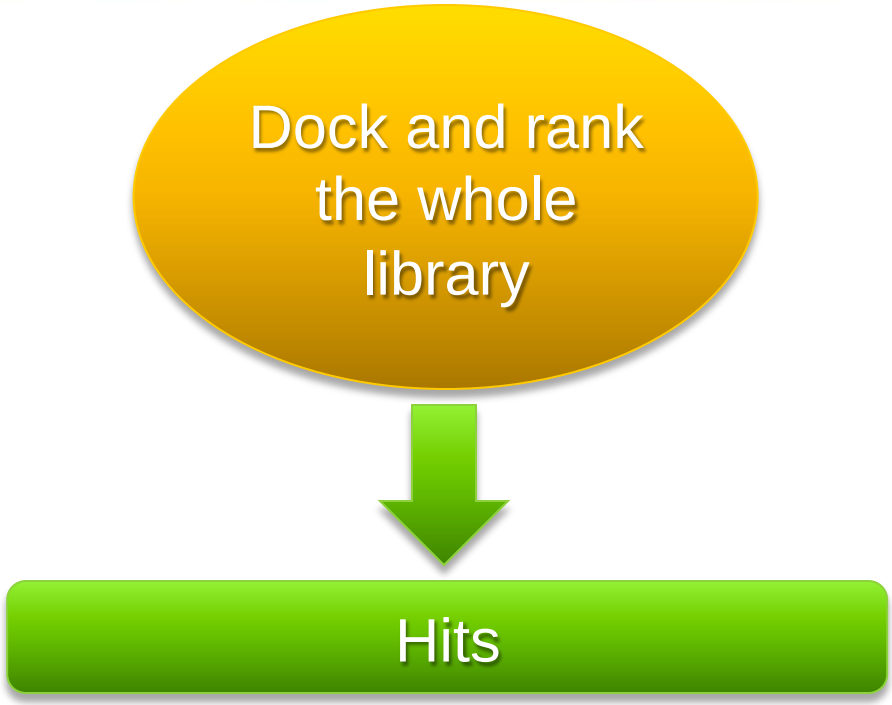


Tuning the docking method

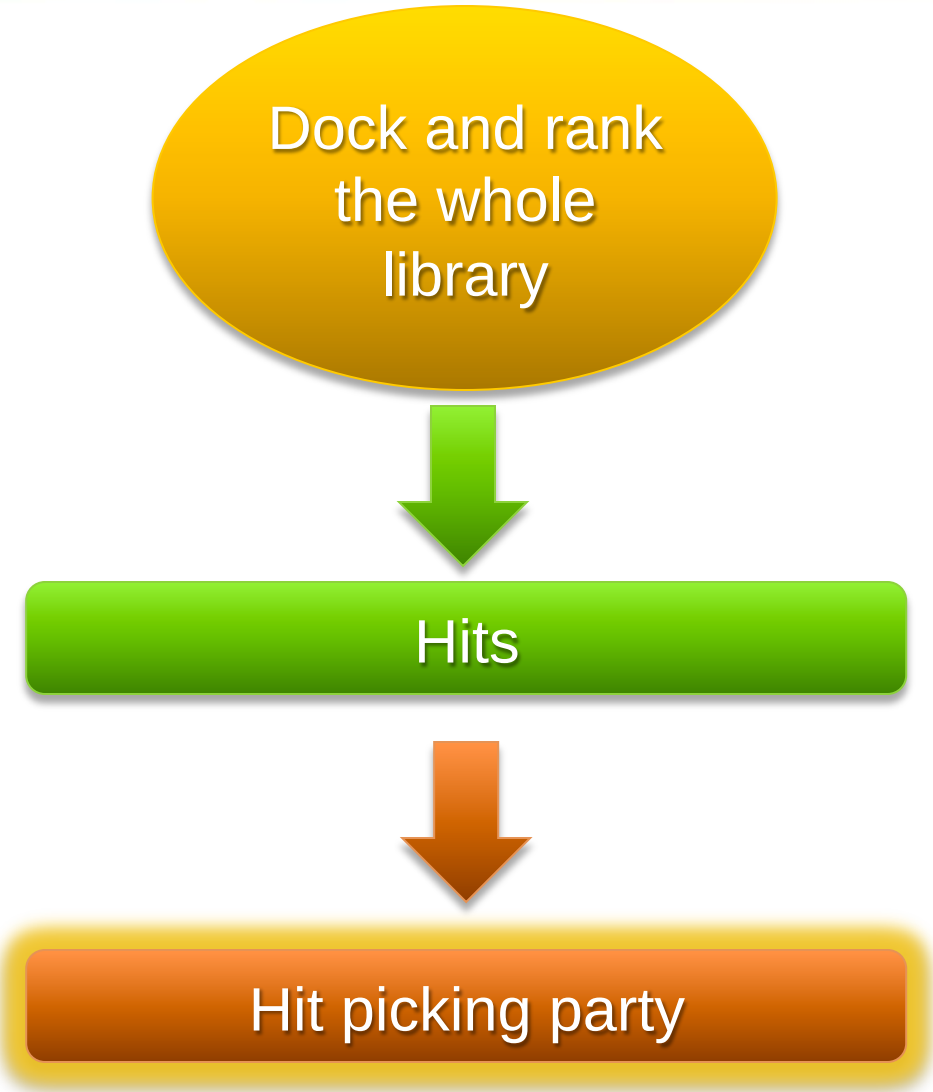


Docking the library

Docking the library



Docking the library



Virtual Screening of Compound Libraries

Pedro Alexandrino Fernandes
Department of Chemistry and Biochemistry
University of Porto
Portugal