BHARATHIDASAN UNIVERSITY

Tiruchirappalli- 620024, Tamil Nadu,
India

Programme: M.Sc., Biomedical Science

Course Title : Drug Discovery and Assay Development
Course Code : 18BMS48ES

Unit-V

Computer Aided Drug Designing

Dr. P.S.Dhivya
Guest Lecturer
Department of Biomedical Science

Computer-Aided Drug Discovery (CADD)

- Investigates molecular properties to develop novel therapeutic solutions by way of **computational tools and data resources.**
- Computational approaches for designing or selecting compounds as potential candidates **before they are synthesized and tested for their biological activity**
- **in silico** approach to reducing costs and time.

- To date, in vitro screening is expensive and time-consuming
- Virtual Screening (VS) is a CADD method
- It involves in- silico screening of a library of **chemical compounds**, to identify those **that are most likely to bind to a specific target.**

Pharmacophore Models

- Compound libraries can be searched to pull out molecules of interest with desired properties.
- Chemo-informatics
- Virtual screening
- Scaffold hopping,
- Lead optimization
- Ligand profiling
- Target identification
- Multi-target Drug or de novo drug design

- 19th century when Langley first suggested **that certain drug molecules might act on particular receptors**.
- Only later, with the discovery of Salvarsan by Paul Elrich, the selectivity of drug–target interactions was recognized.
- **Emil Fisher** who, "Lock & Key" in 1894,
- To interact with each other through a chemical bond

- “Pharmacophore” was used to indicate the **functional or structural capacity** of a compound with specific characteristics towards a biological target

- **Schueler (IUPAC) as**
- "the ensemble of steric and electronic features that is necessary to ensure the optimal supra-molecular interactions with a specific biological target structure and to trigger (or to block) its biological response"

- It is based on the theory that having **common chemical functionalities,** and maintaining a similar **spatial arrangement**, leads to **biological activity on the same target.**
- **Interactions with its ligand** are represented in the pharmacophoric model **as geometric entities such as spheres, planes and vectors.**

Figure 1. Pharmacophoric features. Main pharmacophoric feature types are represented by geometric entities and include: 1—hydrogen bond acceptor (HBA), 2—hydrogen bond donor (HBD), 3—negative ionizable (NI), 4—positive ionizable (PI), 5—hydrophobic (H), 6—aromatic (AR), 7—exclusion volume (XVOL).

- Models themselves do not focus on actual atoms, but on chemical functionalities, they are good tools in **recognizing similarities between molecules**.
- Pharmacophore activity is independent of the scaffold, and this **explains why similar biological events can be triggered by chemically divergent molecules**.

Two Different Approaches

STRUCTURE-BASED

- Uses the structural information of the **target** proteins like enzymes or receptors, **to identify compounds that can potentially be used as a drug.**

LIGAND-BASED

- Development of **3D pharmacophore models and modelling (QSAR) or (QSPR**), using only the physicochemical properties of known ligand molecules.

The choice of the best approach Depends on

- Data availability,
- Data quality,
- Computational resources
- The intended use of the generated pharmacophore models.

Structure-Based Pharmacophore Modelling

- 3D structure of a macromolecule target - protein provides **significant details at the-atomic level** - for the new drugs
- the **stereo-electronic features,** which make a ligand bioactive toward a specific target,
- Type of information represented in its holo or apo form.

Workflow

- Protein preparation
- Identification of ligand binding site
- Pharmacophore features generation
- Selection of relevant features for ligand
- 3D structure of the target or the ligand–target complex

- RCSB Protein Data Bank (PDB) [www.rcsb.org]
- Presence of a bound ligand, mainly solved by X-ray crystallography or
- Computational techniques
  - Homology modelling
  - Molecular docking

Protein Structure Preparation

- The residues' protonation states,
- The position of hydrogen atoms
- The presence of non-protein groups which may have or not some functional roles,
- Eventual missing residues or atoms,
- The Stereochemical and Energetic parameters accounting for the general quality **biological-chemical sense** of the investigated target

LIGAND-BINDING SITE

- Experimental data such as **site directed mutagenesis or X-ray structures of the protein co-crystallized with ligands,**
- inspect the protein surface to search for potential ligand-binding sites according to various properties **(evolutionary, geometric, energetic, statistical, or a combination of them),**
- Examples of computer programs developed for this purpose are **GRID and LUDI**

- Regular grid to identify grid points with which they make energetically favorable interactions, thus generating molecular interaction fields

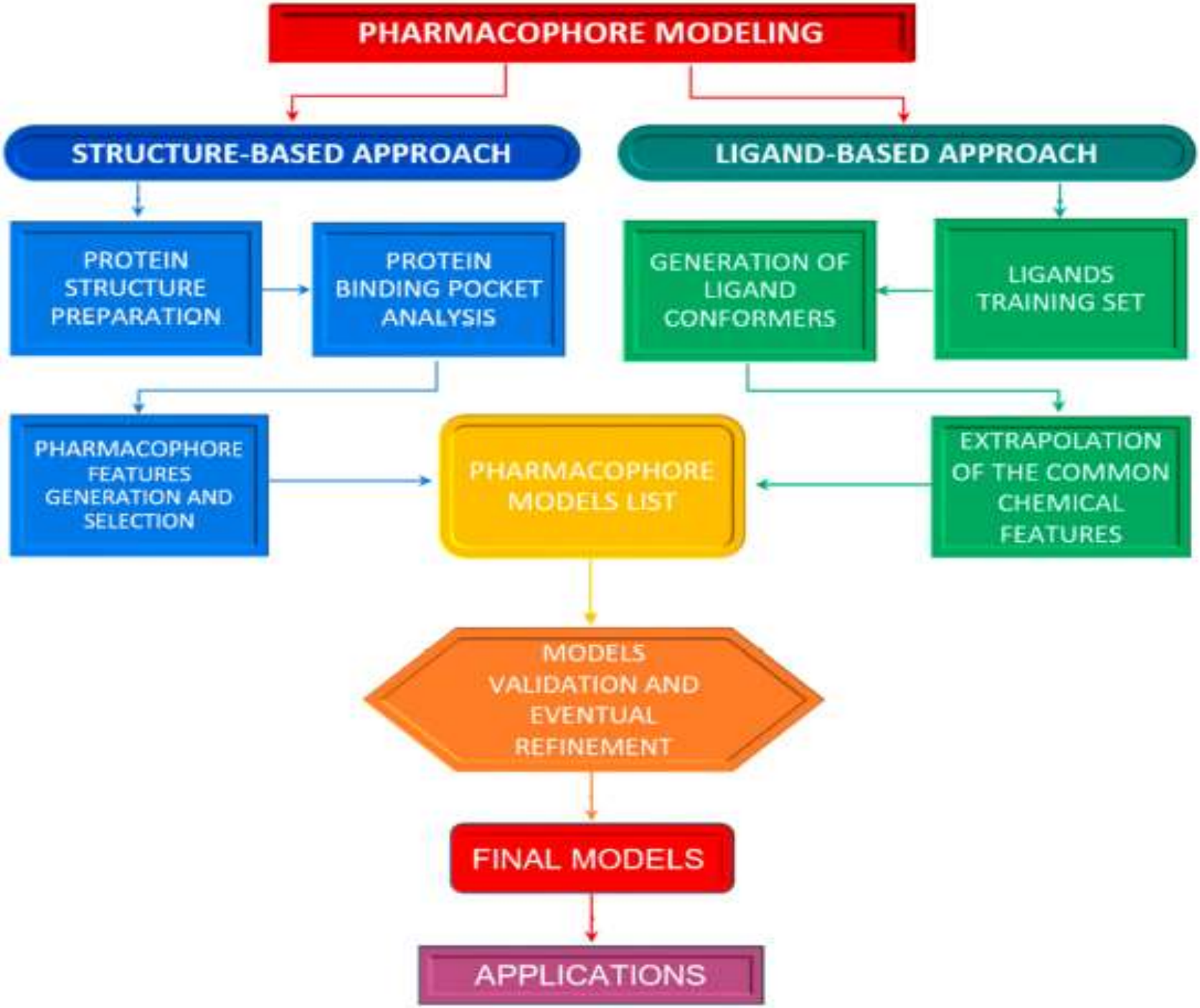
- LUDI- predicts potential interaction sites using the knowledge from distributions of non-bonded contacts in experimental structures or geometric rules.

Goal

- To generate a **pharmacophore model**
- From the **interactions** between an active molecule (the ligand) and its protein target (enzyme or receptor),
- **Characterization** of the ligand-binding site
- **Describing the type and the spatial arrangement** of the chemical features required for a ligand to interact with the residues of the binding region.

- Essential for ligand bioactivity should be **selected and incorporated** into the final model
- **Removing features** that do not strongly contribute to the energy binding,
- Identifying the **conserved interaction** if multiple protein–ligands structures exist,
- **Preserving residue with key functions** from sequence alignments or variation analysis,
- **Incorporating spatial constraints** from the receptor information

- In case the available structural data do not include a bound ligand, the pharmacophore modelling depends **only on the target structure** which can be analyzed **to detect all possible ligands interaction**
- However, in the absence of a ligand counterpart, **several pharmacophoric features are** calculated and approximately positioned resulting in less **accurate models that should be manually refined**


PHARMACOPHORE MODELING
STRUCTURE-BASED APPROACH
LIGAND-BASED APPROACH
PROTEIN STRUCTURE PREPARATION
PROTEIN BINDING POCKET ANALYSIS
GENERATION OF LIGAND CONFORMERS
LIGANDS TRAINING SET
PHARMACOPHORE FEATURES GENERATION AND SELECTION
PHARMACOPHORE MODELS LIST
EXTRAPOLATION OF THE COMMON CHEMICAL FEATURES
MODELS VALIDATION AND EVENTUAL REFINEMENT
FINAL MODELS
APPLICATIONS

Ligand-Based Modelling

- **Target** molecule is **not** available.
- Knowledge of active compounds, **which bind the same protein target** with a similar orientation
- Extraction of **chemical features** in common
- Pharmacophore construction.
- Extraction of the **shared features**, is the generation of ligand conformers,
- Each set of conformers, should correspond to the bioactive conformation of the ligand

- Two datasets i.e., a training set and a test set.
- The training set composition differs according to the algorithm employed and the data availability, varying from the simplest, composed of at least two active compounds, to the most complex, for which it is possible to set multiple compounds with a different activity.

- The test set - many active and inactive structurally different compounds as possible,
- experimentally confirmed inactive compounds and compounds judged as active by experimentally
- proven direct interaction and with suitable activity cutoffs (i.e., low EC50/IC50 and high binding affinity values).

- In the case of unknown inactive ligands,
- **decoy molecules (Directory of Useful Decoys)**
- OpenPHACTS
- ChEMBL
- Drugbank
- PubChem Bioassay
- Tox21

Common Feature Approach

- An algorithm of alignment is involved, to extrapolate and align specific & common features:
- Starting from a defined feature for each active element the algorithm combines it with the corresponding features of the other molecules,
- From this alignment a novel feature, combination of the previous, is generated;
- Reiterating this procedure for each set of overlapping features .

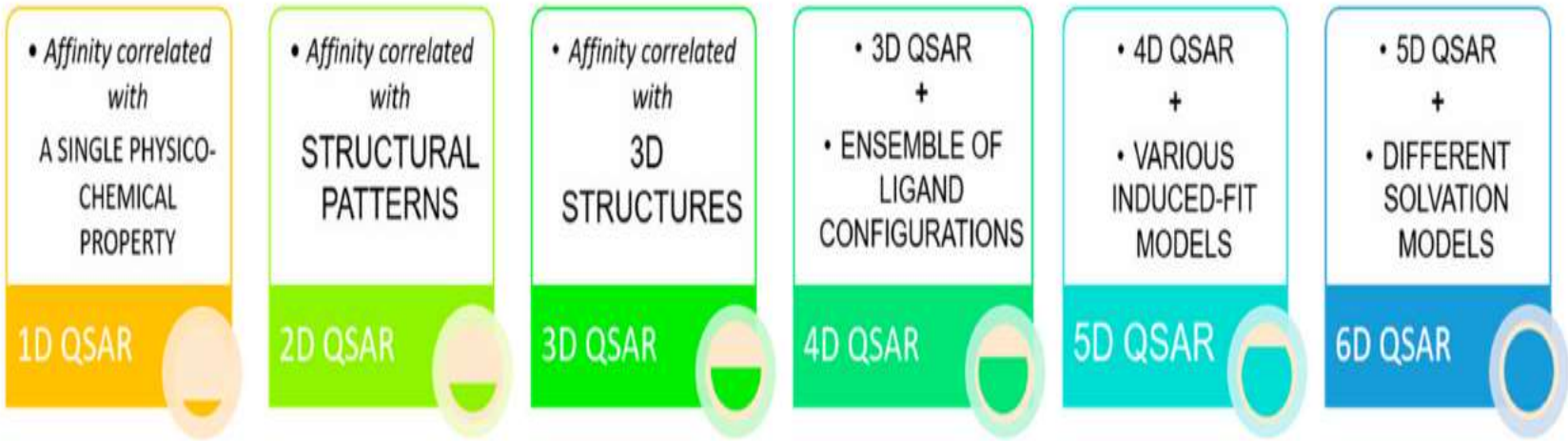
- Based on a combination of shared Pharmacophore feature
- Based on a merged pharmacophore feature.
- Based on common features present in all training compounds
- Based on all features of the aligned compounds are summated in a merged feature pharmacophore

- **Pharmacophore of shared feature is built**.
- Espresso algorithm implemented in LigandScout

- HipHop algorithm by Discovery Studio
- Identifying the common chemical features arrangement of the training set structures.
- Beginning with a few groups of features, which are extended until no more shared configuration features are found,
- Several models are obtained and ranked based on the fitness of the training compounds.

- Based on the molecules **with similar physico-chemical properties** show a similar binding affinity for a protein target
- **QSAR approach** is used in drug discovery to build statistical models derived from these correlations exploiting for the prediction of the **bioavailability, the ADMET profile (toxicity), adsorption, distribution, metabolism, binding affinity, biological activity and elimination of novel Compounds.**

- QSAR (Quantitative structure - activity relationships) based pharmacophore is a mathematical model that tries to find **statistical correlations** between structures and functions,
- Quantifying the impacts in a biological activity of specific structural modifications in existing or predicted molecules.


• Affinity correlated with A SINGLE PHYSICO-CHEMICAL PROPERTY
1D QSAR
• Affinity correlated with STRUCTURAL PATTERNS
2D QSAR
• Affinity correlated with 3D STRUCTURES
3D QSAR
• 3D QSAR + • ENSEMBLE OF LIGAND CONFIGURATIONS
4D QSAR
• 4D QSAR + • VARIOUS INDUCED-FIT MODELS
5D QSAR
• 5D QSAR + • DIFFERENT SOLVATION MODELS
6D QSAR

- 1D-QSAR - **single** physico-chemical property of the ligand, eg. pKa value.
- 2D-QSAR, **instead**, affinity is correlated with structural patterns,
- 3D-QSAR **3D structure** of the ligand and its interactions.
- 4D-QSAR incorporates an **ensemble of ligand configurations**
- 3D-QSAR, 5D-QSAR adds to 4D-QSAR various **induced-fit models**, 6DQSAR implements 5D-QSAR with different **solvation models**

Pharmacophore Models Validation

- Once one or more pharmacophore models have been computed, a validation step is crucial before their implementation for practical purposes.
- Pharmacophore validation - several methods,
- **Goodness of hit list (GH),**
- **Receiver operating characteristic (ROC) curves construction,**
- **Fischer's method,**
- **Satistical analysis, which relies on screening a test set**
- **Decoy set (if needed).**
- To evaluate the model ability to distinguish active and inactive molecules and provide an estimation of its quality.

Quality of a model can be described by four parameters

- the sensitivity (capacity to detect active compounds),
- the specificity (capacity to exclude the inactive molecules),
- the yield of actives (the ratio between true positives and the number of hits)
- the enrichment factor (which relates the yield of actives to the composition of the screening dataset)

Güner–Henry score

- Calculating the percentage of sensitivity,
- Evaluation of the efficiency of the screening dataset search and can vary from 0 to 1
- 1 is the value for the ideal model
- The ROC curve shows the enrichment power of a model plotting the sensitivity against specificity (the false positive rate).
- The (AUC) gives a measure of the pharmacophore's Performance and it is useful for multiple models evaluation.

- AUC can vary from 0 to 1,
- where 1 corresponds to an ideal case in which all the active compounds are detected at first, 0 to the collection of the inactive ones at first and 0.5 to random results.
- **Fisher's randomization test** validation method is instead used to analyze the **significance of the statistical correlation between structure and biological activity.**
- Regression analysis can also be applied to check for the correlation between the compounds predicted as active and those whose bioactivity is experimentally confirmed.

Pharmacophore-Based Virtual Screening

- They increase the speed discovery process through computational simulations by selecting from large libraries the compounds that are most likely to interact with the identified target.
- VS identifies compounds that may be toxic or have unfavorable pharmacodynamic and pharmacokinetic properties.

- To filter large collections of compounds to find the so-called hits,
- i.e., novel molecules matching the pharmacophoric features required to be potentially active against a specific target.
- A pharmacophore **does not represent exact chemical groups** but **chemical functionalities and their spatial relationships**,
- The retrieved hits usually include **structurally different compounds**, making pharmacophores useful tools for scaffold hopping

pharmacophore-based screening

- Pharmit (http://pharmit.csb.pitt.edu)
- ZINCPharmer (http://zincpharmer.csb.pitt.edu/)
- LigandScout, Schrodinger-Maestro, MOE, and Discovery Studio.

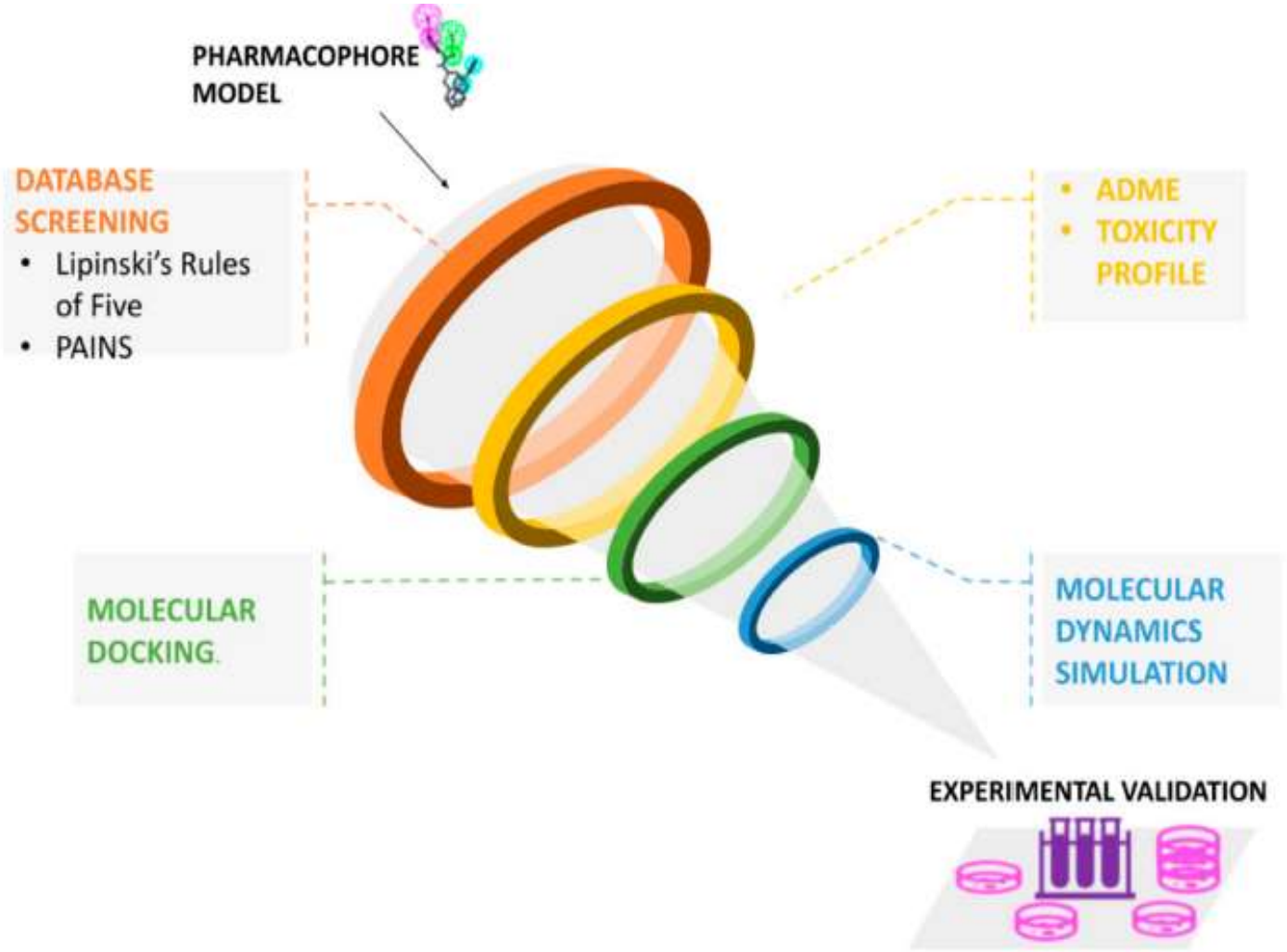


Figure 4. Virtual screening workflow. Basic filters and methods applicable for a virtual screening process. Scheme of the multistage vs. approach. The application of vs. follows a typical sequence of processes with a cascade of filters able to narrow down and choose a set of hits with potential biological activities.

Ligand- Scout

- Ligand topology interpretation
- Aromatic ring detection,
- Assignment of functional group patterns,
- Determination of hybridization state and bond types,
- Ligand and binding site analysis
- classify protein–ligand interactions into hydrogen bonding, ionic, aromatic and lipophilic contacts according to geometric constraints.

TABLE 1 3D pharmacophore modeling software, their components, and availability of free academic licenses

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

Note: Software name describes virtual screening component. Names of 3D pharmacophore generation components in respective software suite may vary. 3D pharmacophores are commonly derived from an overlay of small molecules (ligand), from a ligand bound to its macromolecular target (complex) or from the macromolecular target alone (apo). Pharmacophore features are identified by analyzing molecular fields describing the potential interaction energy toward molecular probes (molecular field), by searching for substructure patterns able to perform the respective interaction (substructure pattern) or by filtering for geometric descriptors fulfilling the criteria for molecular interactions (feature).

[a]The software website was searched, or holder of rights was asked for academic license conditions.

Acknowledgement

- The Presentation is being used for educational and non commercial Purposes
- Thanks are due to all the original contributors and entities whose pictures were used in the creation of this presentation.