# 2.2 3D pharmacophore models

BSR3101: Computer Aided Drug Design

The goal of this tutorial is to learn how to obtain pharmacophore models for a set of ligands.

The tutorial consists of the following steps:

1. Save a new project and open pharmacophore modeling interface
2. Adding Ligands
3. Choosing the active and inactive sets
4. Creating pharmacophore Sites
5. Find common pharmacophores
6. Score hypotheses

What will you learn:
- Identify chemical moieties with common spatial orientation in a set of active ligands
- Build 3D pharmacophore models

Review before you start:
- 3D ligand structures
- pharmacophore modeling

Approximate time required to complete this tutorial:
**15 minutes.**

One of the main ligand-based approaches consists in identifying specific 3D arrangement of the distinctive chemical features of active ligands (i.e. a pharmacophore), and eventually use such arrangements to identify new active ligands. The goal of this tutorial is to build a pharmacophore model of a set of active ligands.

## 1. Save a new project and open pharmacophore modeling interface

First, download the tutorial folder from the link https://goo.gl/h3djH6, move the downloaded archive to the Desktop and extract it. This folder contains example output files that can be used to compare the results of the exercises executed in this tutorial.

Open **Maestro**, set the working directory to the tutorial folder with File > Change Working directory …
Save your project using File > Save Project As…
In the **File name** text box, type **`pharmacophore_modeling`**. Click Save

## 2. Loading the Ligands

Before we begin modeling, we will open the set of ligands for which we have experimental information.

- Click on File > Import structures... to open the file selector panel
- Find the file **`ligands_small.maegz`** in the tutorial directory, and click Open. This file contains different conformations of 25 ligands. For this exercise, the ligands have been previously prepared and conformers generated for each of them.
  (The file **`ligands.maegz`** contains the full set of ligand conformations, while file **`ligands_small.maegz`** is a random subset of the complete set, which is used here to obtain results within the time limits of the tutorial execution).

In the Task panel, type “pharmacophore modeling”, and click on the “Develop pharmacophore hypothesis” entry. Make sure that the “Create pharmacophore model using:” field is set to “Multiple ligands (selected entries)”.
Make sure that the **Genrate Conformers** checkbox is NOT checked, since the conformers for this set of ligands have already been generated and are included in the imported ligands.

## 3. Choosing the active and inactive sets

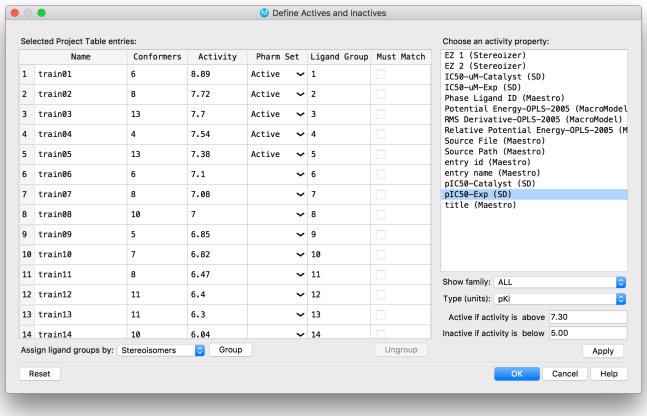
We will now define the actives and inactives: click on the “Define” button in the “Actives/Inactive split”. **Choose Activity Property** window will allow us to pick the variable that will be used to distinguish the active compounds from inactive ones.
Select **`pIC50-EXP`** as the experimental activity value, and make sure the **Units** field is set to “pKi”, then click OK.

In this section, you will set a threshold for actives of IC50 ≤ 50 nM, which translates to pIC50 ≥ 7.3, and a threshold for inactives of pIC50 ≤ 5.0 (IC50≥10 μM).

$-\text{Log}_{10}50\times10^{-9} = 7.3$
$10^{-\text{pIC50}}=10^{-5}=10\times10^{-6} = 10$ μM

- In the **Active if activity above** text box, type **7.3**
- In the **Inactive if activity below** text box, type **5.0**
- Click Apply.

- The **`Pharm Set`** column will indicate whether a molecule is in the set of “actives” used to identify common pharmacophore hypotheses, or in the set of “inactives” used to eliminate nondiscriminatory hypotheses. An empty value indicates that the ligand doesn’t belong to either set. If ligands of widely varying activity are present, you would normally want to use only the most active ones in the set of ‘actives’. The most active ligands are assumed to contain the strongest binding and greatest number of pharmacophore features that are involved in binding to the protein target. On the other hand, the set of actives should contain as much structural diversity as possible, so that the resulting pharmacophore models are applicable across different chemical families.
- The **`#Conformations`** column indicates how many conformations are present for each ligand.



Make sure ligands with activity between 5.0 and 7.3 have blank **`Pharm Set`** fields. You can do so by selecting certain ligand and click in the **`Pharm Set`** field to switch among active/inactive/blank. In the end, five ligands should be in the active set: **`train01`** through **`train05`**, and six ligands in the inactive set: **`train20`** through **`train25`**.

## 4. Inspect the Pharmacophore Sites and set Hypothesis settings

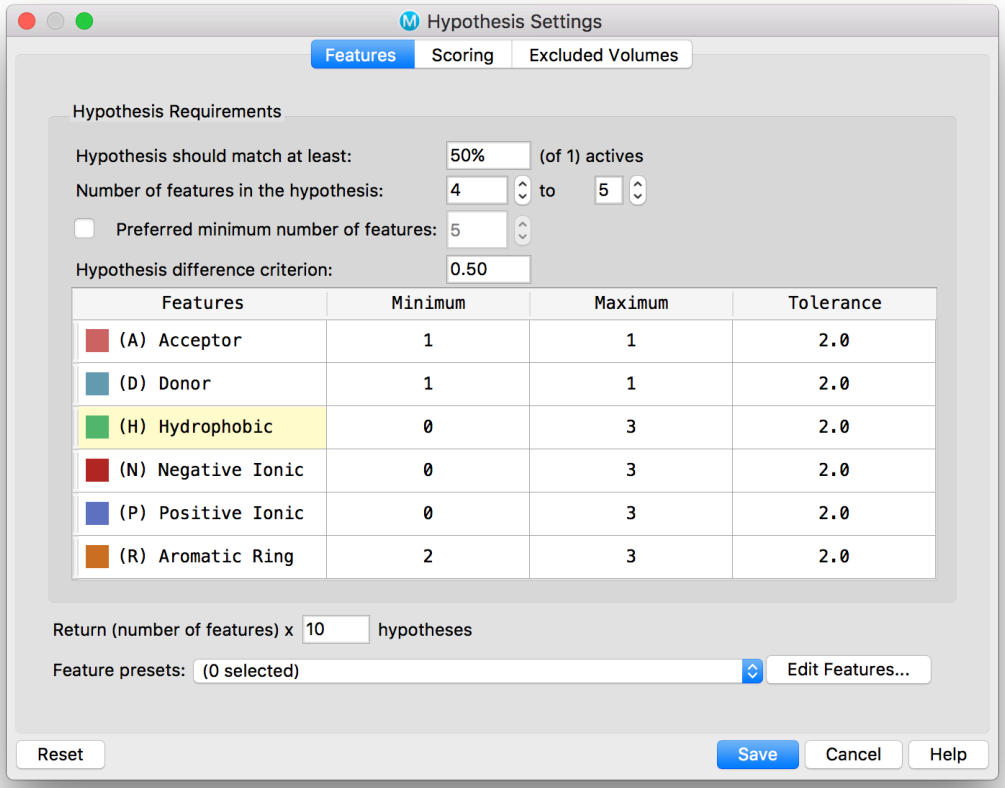
Since the ligands have already been aligned, select "**Use prealigned ligands (consensus model)**", in the **Pharmacophore method** section. Inspect the pharmacophore features for the ligands by displaying different ligands.

A pharmacophore variant is a set of chemical features. In this step, the users are asked to select which variants they want to explore by selecting how many sites the variant is composed of. In a real-like situation, it is recommended to repeat calculation with different to find the hypotheses that optimally distinguish the active ligands from inactive. The parameters that the users can change are:

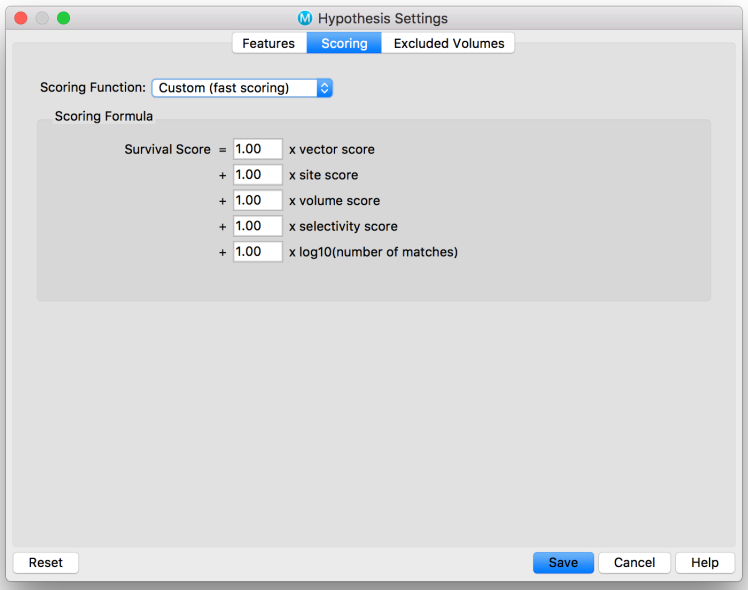
a. Maximum/minimum number of sites.
b. Must match at least…: how many of the active ligands need to match all the pharmacophore features. If you set this to less than the number of active ligands, then the retained hypotheses would explaining the activity of a subset of the active ligands.
c. Feature frequencies. Here the maximum number of each feature is given by counting each ligand we are analyzing, and the user can change the minimum and maximum frequencies of each feature in the pharmacophore models that would be generated.
d. Tolerance: the max tolerance (in Ångstrom) between the pharmacophore feature and the position of the chemical moiety in the actual ligand conformation.

For this tutorial (and just as an instance),

- set the feature frequency of H-bond acceptor (A) to 1 (min=1, max=1)
- set the frequency of H-bond donors (D) to 1 (min=1, max=1)
- set the frequency of aromatic rings (R) to be between 2 and 3



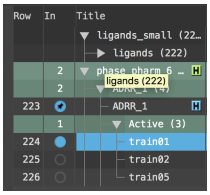

Switch to the **Scoring tab** of the **Hypothesis Settings** panel. The default scoring function for ranking hypothesis in common pharmacophore perception is the PhaseHypoScore. This metric is a linear combination of the Survival score and the BEDROC enrichment performance on a small-scale virtual screen automatically set up and executed by the common pharmacophore perception process. Each hypothesis is evaluated in its ability to rank actives used in common pharmacophore perception against the 1000 compound Glide decoy set. The Glide decoy set was created by selecting 1000 ligands from a one million compound library that were chosen to exhibit "drug-like" properties.



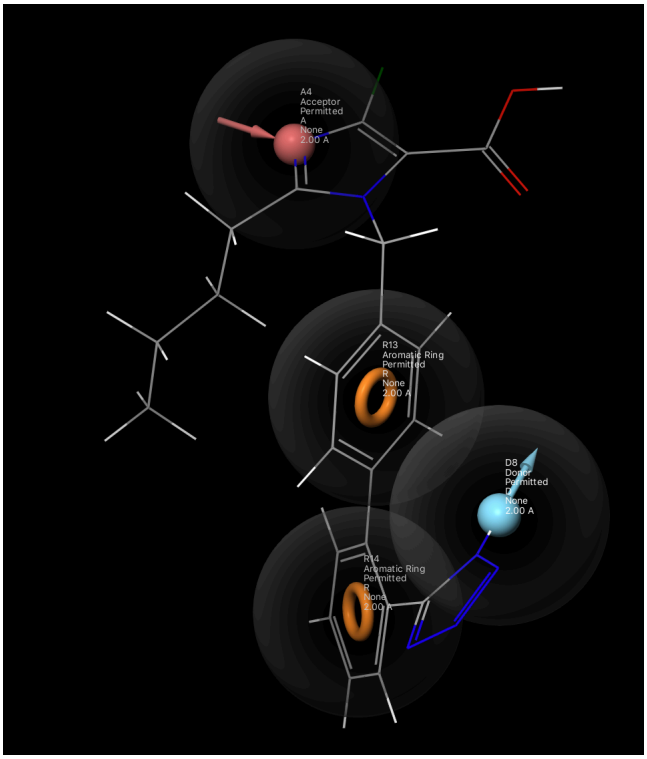

To speed the current calculation, change the **Scoring Function** to **Custom (fast scoring).** This option avoids execution of the small-scale virtual screen and the returned hypothesis are rank ordered by their survival scores. You can leave the coefficients for the Survival score at their default values.

## 5. Identify the Pharmacophore Hypotheses

Once the calculation is done, one single hypothesis (ADRR) should be returned given the settings described above.




Click twice on the blue circle to display and lock the pharmacophore hypothesis and expand the “Active” group to identify which ligands in the Actives set match the pharmacophore within the tolerance specified.


A4
Acceptor
Permitted
A
None
2.00 A
R13
Aromatic Ring
Permitted
R
None
2.00 A
D8
Donor
Permitted
D
None
2.00 A
R14
Aromatic Ring
Permitted
R
None
2.00 A