GPCRmd Documentation

Release 1.0

GPCR drug discovery group

Aug 05, 2020

Contents:

1 Workbench **1**
- 1.1 Viewer . . . 1
- 1.2 Toolkit . . . 3

2 Receptor meta-analysis **7**
- 2.1 About the heatmap: . . . 7
- 2.2 About the Top representative interactions in clusters: . . . 8
- 2.3 Customized heatmaps . . . 8

3 Simulation submission **9**
- 3.1 Protein information . . . 9
- 3.2 Small molecule information . . . 10
- 3.3 Complex information . . . 10
- 3.4 Dynamics information . . . 10

4 Search **11**
- 4.1 List of simulated systems . . . 11
- 4.2 Browser . . . 11

5 Data download **15**

6 Citing GPCRmd **17**

CHAPTER 1

Workbench

Contents

- *Workbench*
  - *Viewer*
  - *Toolkit*

The GPCRmd Workbench includes a set of online tools for the interactive visualization (*Viewer*) and analysis (*Toolkit*) of individual simulations.

1.1 Viewer

1.1.1 General features

Mouse controls:

- **Left button hold and move**: rotate camera around center.
- **Middle button hold and move**: zoom camera in and out.
- **Middle button click**: center camera on atom.
- **Right button hold and move**: translate camera in screen plane.
- **Left button click: pick atom or distance.**
  - *On click show distance* mode:
    * When an atom is clicked, a label with information about it appears. Click at the background to deselect it, the label will disappear. To maintain a label, double-click on an atom. Double-click again on the atom to remove the label.

    * To draw a distance line between two atoms just single-click one atom after the other. Distances can be removed by double-clicking on one of the atoms at the edges.
    * It is also possible to remove all the atom labels and distances at once, with the **Clear dists. button**.
  - ***On click show variants* mode:** Click on the blue dots to obtain information on known natural variants of a residue. Data obtained from the gnomAD database.
  - ***On click show mutations* mode:** Click on the blue dots to obtain information on mutational experiments done on a residue. Data obtained from GPCRdb.

If you wish to use more complex visualization options, you can open the structure and trajectories in MDsrv just by clicking at the gear button.

1.1.2 Selection tools

Quick selection

Quick-selection buttons allow to rapidly display the molecules present at the dynamics. Hover the buttons with your mouse to see the abbreviated name of these molecules, which can be used to create your own selections.

In the case of Structure selections, It is also possible to select the residues or molecules that are found **within a certain distance of a ligand**. It is only necessary to:

1. Indicate what you want to visualize (residues or molecules found at the simulation).
2. Input the wanted threshold distance (in angstroms).
3. Indicate the molecule type around which the selection is made. Apart from predefined molecules, it is also possible to show the residues/molecules that are close to a personalized selection which, again, can include generic GPCR residue numbering.

If the selection is correct, a green checkmark will appear on the left. More than one distance selection can be displayed at the same time. Selections made with this tool will appear in coral red. The distance selection will be updated for each trajectory frame, as the disposition of the atoms may change.

Custom selection

Use the text input field to specify your personalized representations. You can choose a representation type (licorice, cartoon, etc.) and a coloring scheme (color by element, by chain, etc.).

Selections must be expressed using the NGL selection language. Moreover, to indicate protein residues it is also possible to use **generic GPCR residue numbering**: Ballesteros-Weinstein (ex. 1.50), GPCRdb structure-based numbering (ex. 1x50) or a combination of both (ex. 1.50x50).

For example, if you input `40-70:P or CLZ`, residues numbered from 40 to 70 at the PDB belonging to chain P and Clozapine will be displayed. Another example, this time using a combination of different generic GPCR numbering styles, could be `1.50 - 2x48 or 3.35x35 or SOD`.

If your selection includes water and/or membrane molecules that you wish to display, check 'Include membrane and water'. This may slow down the playback speed of the simulation.

Sequence selection

The GPCR Workbench also provides the option to select a protein segment from its sequence. Set your selection by clicking at the desired range or ranges of residues. Selected segments will appear at the sequence in green. To deselect a residue range from the sequence, just click on it. Finally, click at **Confirm selection**: the residue range(s) will be

added to a text input field, which you can further modify to adjust the selection. If you want to add new sequence selections, click at the plus button.

GPCR conserved positions

This section provides the possibility to rapidly select positions or domains conserved in the different GPCR family classes. The GPCR class of the protein being represented will be selected by default, and therefore the conserved positions/domains corresponding to that GPCR class will be available to visualize.

It is also possible to visualize the positions that correspond to conserved positions from other GPCR classes. For example, if your protein belongs to class A, you can represent the residue that corresponds to class B *2.50 (2.50b)*. Hover the buttons with your mouse for more information about the conserved positions and motifs, if available.

Experimental density maps

Display X-ray or electron microscopy density maps by defining any selection within this section. By default, this will also modify the representation of the molecules of the system. To change this behavior, un-select the option "Apply default representations" within *Selection settings*.

1.2 Toolkit

1.2.1 Interaction network (Flare plots)

Flare Plots are a tool for the study and representation of intra-protein interactions developed at Stanford University by Dr. Fonseca and Dr. Venkatakishnan. This approach makes it possible to obtain a highly visual depiction of complex data, such as the set of interactions formed between protein residues throughout MD simulations, in the form of circular interactive networks named Flare plots. Residue-residue interactions are represented as lines connecting residue pairs. Hover or click a residue to highlight the lines representing the interactions in which it participates.

There are several options available

- **Interaction type: Select the type of interaction to display on the plot.**
  - **Hydrogen bonds:**
    - **Wernet Nilsson criteria (MDTraj):** Any combination of donor atoms (NH or OH) and acceptor atoms (N or O) that holds the condition:
      |AD| < 3.3 Å 0.00044 * HDA * HDA
      Where |AD| is the distance in Angstroms between donor and acceptor heavy atoms, and HDA is the angle formed by the hydrogen atom, donor, and acceptor atoms in degrees. Defined by the MDTraj module function wernet_nilson.
    - **GetContacts criteria:**
      |AD| < 3.5Å
      AHD < 70°
      Where A (acceptor) and D (donor) are any atom except hydrogen, carbon or sulphur.
      Based on GetContacts.
  - **Salt bridges:**
    |AC| < 4.0Å
    Where:
    A (anion): ASP/OD1+OD2, GLU/OE1+OE2
    C (cation): LYS/NZ, ARG/NH1+NH2, HIS/ND1+NE2

Based on GetContacts.

- **Pi-cation:**
  |AC| < 6.0Å
  CAn < 60°
  Where:
  A (aromatic): center(PHE/CG+CE1+CE2), center(TRP/CD2+CZ2+CZ3), center(TYR/CG+CE1+CE2), center(HIS/CG+CD2+CE1)
  C (cation): LYS/NZ, ARG/NH1+NH2, HIS/ND1+NE2
  Based on GetContacts.
- **Pi-stacking:**
  |A1A2| < 7.0Å
  (n1, n2) < 30°
  (n1, A1A2) < 45°
  (n2, A1A2) < 45°
  Where:
  A1, A2 (aromatic rings): center(PHE/CG+CE1+CE2), center(TRP/CD2+CZ2+CZ3), center(TYR/CG+CE1+CE2), center(HIS/CG+CD2+CE1)
  Based on GetContacts.
- **T-stacking:**
  |A1A2| < 5.0Å
  60° < (n1, n2) < 90°
  (n1, A1A2) < 45°
  (n2, A1A2) < 45°
  Where:
  A1, A2 (aromatic rings): center(PHE/CG+CE1+CE2), center(TRP/CD2+CZ2+CZ3), center(TYR/CG+CE1+CE2), center(HIS/CG+CD2+CE1)
  Based on GetContacts.
- **Van der Waals:**
  |AB| < Rvdw(A) + Rvdw(B) + 0.5
  Where A and B are any non-hydrogen atoms.
  Based on GetContacts.
- **Water bridges:** Two different residues forming a Hydrogen bond with the same water molecule. Based on GetContacts.
- **Extended water bridges:** Two different residues forming a Hydrogen bond with two different water molecules which also form a hydrogen bond between them. Based on GetContacts.
- **Hydrophobic:**
  |AB| < Rvdw(A) + Rvdw(B) + 0.5
  Where:
  A, B: ALA+CYS+PHE+GLY+ILE+LEU+MET+PRO+VAL+TRP and element C or S
  Based on GetContacts.

- **Display:**
  - **Interacting pairs**: Show only a subset of interactions (intra- or inter-helix) or all of them.
  - **Simulation**: It is possible to summarize the interactions formed through all the trajectory frames. The frequency of each interaction is represented by the thickness of the lines connecting residues.

- **Show in structure**: Click to display structural representations of the residues selected (clicked) at the flare plot. Unclick to hyde them. If there are no residues selected at the flare plot, nothing will happen.
- **Clear plot**: Click to delete all selections made on the plot.
- **Download data**: Click to download the plot data.

1.2.2 Interaction frequencies

Hydrogen bonds

This tool identifies Hydrogen Bonds formed in a simulation, splitting the results between protein-protein hydrogen bonds and protein-not protein bonds. We use the MDTraj module function wernet_nilson, which establishes a threshold distance of 3.3 Angstroms between the donor and acceptor atoms; this threshold becomes progressively stricter as the angle formed by H-D-A increases (a perfect straight bond is 0 degrees, as the donor atom is central). It's possible to choose between a few options:

1. **Do not include hydrogen bonds between neighbors**: If selected, excludes hydrogen bonds among residues which are less than 5 residues apart. These are usually the hydrogen bonds stabilizing alpha helices.
2. **All hydrogen bonds**: If selected, includes hydrogen bonds formed between backbone (BB) atoms or side chains (SC) atoms, in any combination (SC-SC, BB-BB, SC-BB).
3. **Only side-chain hydrogen bonds**: If selected, only includes hydrogen bonds formed between side-chain atoms.

Finally, you can set a frequency threshold so only those hydrogen bonds which hold the cited condition in a proportion of the frames greater than the value you have set will appear in the results. You can also define an interval of frames into which perform the analysis.

Results have a "Show Hbond" button next to them which displays the bond in the viewer. At the end of the results table, you can find a "Show All" button, which displays all the bonds in that table at once.

Ligand receptor contacts

This analysis tool calculates the frequency of interaction between the protein residues and a given ligand across a trajectory. When the distance between any of their atoms and the ligand is smaller than the threshold, it is considered to be an interaction. It is possible to chose which residue atoms will be considered (heavy atoms only or all atoms). The result is presented as a table and a plot, which can be downloaded as an image. The residues that are found to interact can be displayed at the viewer screen (shown in purple), which can be deactivated using the "Display interacting residues" checkbox. It is also possible to download the interaction data obtained.

Salt bridges

This tool allows you to identify the salt bridges formed through a simulation. Salt bridges are defined as any combination between these two sets: {Arg-NH1, Arg-NH2, Lys-NZ, His-NE2, His-ND1} and {Glu-OE1, Glu-OE2, Asp-OD1, Asp-OD2} in which the participating atoms are closer than 4 Angstroms. Histidine atoms are only considered if the residue is protonated. As with hydrogen bond analysis, you can select a percentage threshold, and the results include a "Show Salt Bridge" button and a "Show All" button. Furthermore, you can select an interval of frames, instead of the whole trajectory.

1.2.3 Distance

This tool is used to calculate the distance between atom pairs across the different frames of a trajectory, and therefore across time. To calculate a distance, you need to indicate the pair or pairs of atoms you are interested in. This can be

done in different ways:

- Select a pair of atoms at the viewer screen by clicking on them and, afterward, **importing the created distances** with the blue arrow button.
- Indicate the desired atom pairs manually, by selecting "Compute distance between" **atoms** and inputting a pair of atom indices at the text input fields.
- Indicate the desired atom pairs manually, by selecting "Compute distance between" **residues** and indicating the residue, chain and atom name you are interested in. The residue number and chain name must be indicated according to the NGL selection language (ex. 50:P), and the atom name selected from the droplist.

It is also necessary to select the trajectory that will be used for the calculation. Finally, just click at **Compute**. Only atom pairs that are marked with a green checkmark will be considered, since the absence of a checkmark indicates an error in the input (only numbers are allowed). The result will appear as a plot of distance by time or by frame, which can be downloaded as an image. It is also possible to download the data obtained as a CSV file. Moreover, the distances calculated can be displayed at the viewer screen, in the colors indicated at the plot legend. Such distance representations can be deactivated by deselecting the "Display distance" checkbox.

1.2.4 RMSD

This tool computes the RMSD of all the conformations in a target trajectory to a reference conformation. It is necessary to indicate the trajectory to be used and the frames to be considered. Also, a reference frame of a given trajectory. It is possible to chose which atoms are going to be considered in the calculation: only alpha carbons, non-hydrogen protein atoms, protein C-alpha, etc. As in the case of distance analysis, the result will be shown in a plot (RMSD by time or by frame). It is possible to download the plot as an image or all the obtained data as a CSV file.

1.2.5 Water volume distribution

Displays an averaged water density map of the MD trajectory under study. Maps are precomputed VMD VolMap Tool. They are generated only for oxygen atoms of a water molecule in a cutoff distance of 10 Å to the protein using a resolution of 1 Å. Atoms are treated as spheres using the atomic radii.

1.2.6 Tunnels and channels

Displays the tunnels and channels formed in the receptor during the simulation. Tunnels are defined as void pathways leading from a cavity buried in a protein core to the surrounding solvent, while in channels both endings are opened to the surrounding solvent.

Tunnels/channels are precalculated using the software Caver 3.0. The starting point coordinates for apo forms and receptor-ligand structures are set to the center of mass of ligand-interacting residues in the respective PDB structure. The following input parameters are used: probe_radius=1.4, shell_radius=3, shell_depth=4. Note that we focus our analysis on the ligand-binding pocket, so tunnels/channels unrelated to the ligand-binding pocket may not be detected.

All the tunnels/channels identified in the simulation are clustered by similarity. Such clusters of identified tunnels can be displayed by selecting them in the "Clusters" column.

It is also possible to display the tunnel with the highest throughput of each cluster. As defined by Caver, the throughput of a tunnel or channel corresponds to the importance of the pathway, which is the probability that the pathway is used as a route for transportation. Tunnel throughput is calculated based on the radius and length of the tunnel. The frame at which the highest-throughput tunnel of each cluster is found can be displayed by clicking at the "Display frame *x*" button.

CHAPTER 2

Receptor meta-analysis

Contents

- *Receptor meta-analysis*
  - *About the heatmap:*
  - *About the Top representative interactions in clusters:*
  - *Customized heatmaps*

2.1 About the heatmap:

This plot compares the interaction pattern among several MD simulations, based on the contact frequency of each residue pair of the simulated systems (including protein-protein and protein-ligand interactions).

Interaction frequencies have been calculated with GetContacts scripts, created by Rasmus Fonseca (fonseca.rasmus@gmail.com) and Anthony Ma (anthonyma27@gmail.com). Interaction frequencies are considered as the proportion of frames in which a residue pair is interacting, according to GetContacts interaction criteria. For simulations with more than one trajectory available, the average frequencies are obtained.

Based on the frequency values, a clustering analysis is performed to hierarchically classify each simulation. The obtained dendrogram is generated and displayed using plotly python library.

Protein residues are labeled using generic GPCR residue numbers (GPCRdb structure-based numbering). This numbering system differs for each GPCR class, and so a multi-class index has been used on the heatmap's axis.

Only residue pairs with a minimum of 50% interaction frequency in at least one simulation are displayed. Same-helix interactions are not displayed.

How to use this heatmap:

1. Select the simulation dataset to display in the heatmap. You can filter out simulations provided by individual contributions (not generated by the GPCRmd community) by unselecting this dataset.

2. Select the type of interaction you want to display. It is also possible to display the "total interaction frequency", which refers to the percentage of frames on which the residue pair is interacting in any of the considered interaction types.
3. Select the partners of the interactions to be displayed (ligand-protein or protein-protein).
4. Select the number of clusters to differentiate on the dendrogram.
5. Check "Show reversed residue pairs" to display interactions for the reversed residue pair (Eg: 5x43-7x32 and 7x32-5x43).
6. Click "Apply" to load the interactions selected.

More information from a selected interaction can be displayed by clicking upon its cell.

2.2 About the Top representative interactions in clusters:

These plots show the *n* most frequent interactions (on average) in each of the clusters. Interactions are colored based on the average frequency of the interaction within the cluster, according to the same color scale as the heatmap.

Such plots were created using FlarePlots, by Rasmus Fonseca (fonseca.rasmus@gmail.com) and Anthony Ma (anthonyma27@gmail.com).

How to use them:

1. To select a cluster to display, use the dropup on top of each flareplot. Cluster 1 and Cluster 2 are shown by default.
2. To select a position on the flareplot, click on it. If available, this position will be displayed on the structure viewer below.
3. Press "show in structure" to disable or re-enable the visualization of selected positions.

The viewers at the bottom of the page (based on NGL viewer) display each of the simulations present in the cluster selected in the flareplot.

How to use them:

1. Use the top-right dropdown to select the simulation to display. Notice only simulations of the selected cluster are available.
2. Use the top-left dropdown to select the trajectory file of this simulation to display.
3. Press play to reproduce the simulation on the viewer.
4. Click on any position in the flareplot to display it on the corresponding viewer.
5. For a more detailed visualization of this simulation, click on "open with GPCRmd Workbench".

2.3 Customized heatmaps

It is possible to analyze only a specified set of simulations using the Customized heatmaps. For that, use the "Select simulations" dropdown to specify the simulations of interest. Simulations can also be selected by clicking on their dendrogram labels. Once the selection is done, click on "Apply".

New plots showing the interactions of the simulations selected are generated. This allows for a more detailed one-to-one comparison of the interaction pattern of such simulations.

Please take into account that a large simulation selection may require a long time to load.

CHAPTER 3

Simulation submission

The simulation submission process consists in 4 steps:

- *Simulation submission*
  - *Protein information*
  - *Small molecule information*
  - *Complex information*
  - *Dynamics information*

3.1 Protein information

Please provide relevant data for GPCRs as well as other proteins (e.g. G protein, arrestin, nanobodies, peptide ligands, etc.) that are included in the simulation.

(A) Protein information is automatically retrieved by its UniProtKB accession number (AC). In case no AC exist, fill in manually the requested information in **(B)**.

(B) Automatically retrieved protein details. In case the protein is not available in UniProtKB the user has to manually provide this information.

By clicking the "**+ Add protein**" button, information about another protein in the system can be provided. Several proteins can be added in a single submission but at least one must be a GPCR whose sequence must have been deposited to the GPCRdb.

3.2 Small molecule information

Please, provide a detailed chemical description of all non-protein molecules in your system. This extensive information will help to provide a platform of well-characterized molecules for screening purposes to medicinal chemists and chemoinformaticians.

(A) Upload for each small molecule (non-protein) a '.sdf' or '.mol' file. A 2D chemical structure will be automatically displayed.

(B) Indicate if the uploaded structure is a co-crystallized molecule or if it belongs to bulk.

(C) Revise chemoinformatics data retrieved from the upload structure.

(D) Obtain molecule information from PubChem and ChEMBL databases.

Click the "**+ Add molecule**" button on the bottom of the panel corresponding to the last uploaded molecule to add information about another molecule in your system.

Molecule data submission can be made individually by clicking the "**Submit molecule**" button on the bottom of every molecule panel or simultaneously by clicking the "**Submit all molecules**" button on the bottom of the form.

3.3 Complex information

(A) Fill in the form with general information about the crystallized part of the simulated system.

(B) Fill the "Curated protein data". As the coordinates of a protein may have different sources (X-ray, homology modeling. . .), the table must be filled in a protein segment-wise fashion. A segment stands for a continuum fragment of the protein whose coordinates come from the same source. In cases where the coordinates of the whole protein come from the same source, only one row in the table is needed for that protein. Importantly, information about the segments of all the proteins submitted in step 1 must be provided. Every segment must be labeled with the "**Prot #**" number of the protein to which it belongs (the "**Submitted proteins summary**" table shows the "**Prot #**" number for each submitted protein). In order to assign correctly the canonical GPCR numbering to your simulation system, please indicate the starting ("**From res**" field) and ending ("**To res**") residue numbers for each protein segment. After clicking the "Autocomplete resid" button, the Uniprot numbering will be later automatically assigned. In addition, indicate the source for coordinates of each crystallized or modeled protein segment, e.g. 'threading' for a modeled loop. Note that the "**PDB ID**" field in the table must be filled when the segment source is "X-ray" or "NMR".

(C) Indicate the "**Resname**" in the retrieved list of co-crystallized small molecules. Extra rows can be added to the table after clicking the "**+ Add molecule**" if molecules with more than one "Resname" exist in the set of crystallized molecules. Then, click the "**Validate**" button and the "**Num of mol**" will be automatically filled.

3.4 Dynamics information

Please, fill in the following form concerning specific details on the MD simulation protocol.

(A) Upload simulation files including coordinate file, topology file, trajectory files, simulation parameters, other files. In the case of replicates, you can upload all files simultaneously (e.g. 20 to 100 replicates). After the file upload, please click the validate button to ensure the correct upload.

(B) Please complete the table of all simulation components by adding non-crystallized bulk molecules (e.g. waters, ions and lipids). For this, you only need to assign the corresponding residue name (resname) to the bulk molecules which are part of your simulation system. Then, click the "**Validate**" button to automatically fill the number of molecules (Num of mol) based on your uploaded simulation files

(C) Specify the simulation setup including information of Method, software, etc.

CHAPTER 4

Search

Contents

- *Search*
  - *List of simulated systems*
  - *Browser*

4.1 List of simulated systems

This table provides an easy and quick tool to search for any simulation in GPCRmd. Use the *Search* input to filter out the simulations shown at the table. It is possible to search by any of the terms included in the table (receptor name, Uniprot entry name, Uniprot ID, etc.).

The table rows contain a summary of the most important features of the simulations. This includes links to further details of the receptor and molecules present in the simulated system. Click on "Go to the Workbench" to visualize and analyze the simulation. By clicking on "Go to the Report", you will access the details of the system setup and simulation protocol, as well as links to download the simulation data.

4.2 Browser

For more complex searches, we provide our Browser, which includes more advanced browsing tools. Moreover, it has the advantage of allowing to search by synonyms.

Overall, the Browser is used to find simulations based on the elements included in it. First, use panel **1** to search the desired elements (GPCRs, other proteins, molecules, etc). Add the elements to the search by clicking "Add to Search". Then, use the panel **2** to search for simulations containing the elements added.

4.2.1 Usage

Let's say you are looking for a simulation containing serotonin and the serotonin receptor. You should use the panel **1** to search for these elements (e.g. search by the word "serotonin"). You will obtain a list of elements related to your search (serotonin, serotonin receptors, . . .). Select the elements you are interested in by clicking "Add to Search". They will appear in panel **2**. If you do not want to apply any more filters or use the advanced search, you can just click at the search button of panel **2**. A list of simulations corresponding to your search will appear.

Panel 1, searching the elements:

Use panel **1** to search for molecules or proteins (the "elements" which constitute simulations). You can use SMILES, InChI, InChIKey, PubChem CIDs, ChEMBL id's and common names to search for molecules. In order to find proteins, Uniprot accession codes and common names can be used. You can indicate if your search is a GPCR or another protein using the dropdown. You can also use it to search by dynamics ID. If you are not looking for a protein, just use the option "All". Then press the search button to obtain the resulting elements. Next to every result, you will find an "Add to search" button, which will add the element to the panel **2**.

Panel 2, joining the elements

Panel **2** holds all the elements you have added. You can directly click on the search button to find all simulations matching that combination of elements, or you can tune your results by filtering by some parameters, which you will find at "See simulation filters". Such parameters include time step, software used, etc. For molecules, you can select whether your molecule is an "Orthosteric ligand", "Allosteric Ligand", "Other" or "All" (meaning any of the previous). For proteins, you can select if it is a GPCR or not, using the *"is GPCR" checkbox*. If it is not selected, we assume that the protein is NOT a GPCR. If you need to use boolean operators like OR, AND, NOT, switch to the Advanced search.

4.2.2 Difference between standard form and specific state

The difference between standard form and specific states is that a standard form represents a **set** of related molecules that only differ in the protonation state, tautomerization state, or some other subtle features. Each of these different states is a single specific state. As the concept of specific state is more precise, when searching with it, you will only find those simulations containing exactly that specific state. However, if you search by a standard form, you will find all the simulations which have any of the related molecules belonging to that standard form. Therefore, if, for example, you search systems using Clozapine (standard form) you will always get more results (or the same) than if you search by Clozapine (specific state X).

4.2.3 Simple and Advanced Search

In panel **2**, you can choose between *Simple and Advanced search*, the difference being that the second allows you to use "AND", "OR" and "NOT" boolean operators, as well as brackets. Simple search only allows "AND" operator (which is enough in most cases). The following is an example of how to perform a search with brackets:

P28222	P28222
AND (Serotonin	AND Serotonin
OR Clozapine	OR Clozapine
OR POPC)	OR POPC

The example search in the left will return the systems matching any of the following:

- P28222 + Serotonin

- P28222 + Clozapine
- P28222 + POPC
- P28222 + Serotonin + Clozapine
- P28222 + Serotonin + POPC
- P28222 + Serotonin + Clozapine + POPC
- P28222 + POPC + Clozapine

However, the example search in the right will return the systems matching any of the following:

- P28222 + Serotonin
- Every system with Clozapine
- Every system with POPC

Please, keep in mind that the results may not be composed only by the elements you have included in the search. So, when we say that the system will match "P28222 + Serotonin", it means that at least, it has these two components - it may also have something else, like ions, water, etc. Read the *Exact match* section to know how to avoid this greedy behavior.

4.2.4 Exact Match

If you select "Exact Match", the search will return only those simulations whose composition matches exactly the one you have built in the right panel. If there is any other element in the simulation that you have not included in the search, it will not appear as a result. So, if you look for P28222 AND Serotonin, and select exact match, only those systems composed by P28222 and serotonin and nothing else will be retrieved; a system with P28222, water and serotonin will not be a result.

4.2.5 Empty search

You can search simulations without adding any elements. This will return every simulation we have in our database. You can still filter by any of the fields: Force Field, Software, etc.

CHAPTER 5

Data download

Contents

- *Data download*

All files related to the simulations stored in GPCRmd are available for download at the *Simulation report* page. This includes simulation outputs (trajectories, topologies, and coordinates) as well as simulation protocol and starting files.

The *Simulation report* can be accessed from the *List of simulated systems* tab of our Search page or directly from the *GPCRmd Workbench* page of the simulation, through the *Simulation report and files* link.

CHAPTER 6

Citing GPCRmd

If you use GPCRmd in your work, please cite:

- Rodríguez-Espigares, I., Torrens-Fontanals, M., Tiemann, J.K.S. et al. GPCRmd uncovers the dynamics of the 3D-GPCRome. *Nat Methods*. 2020;17(8):777-787. doi:10.1038/s41592-020-0884-y