The GPCRmd project

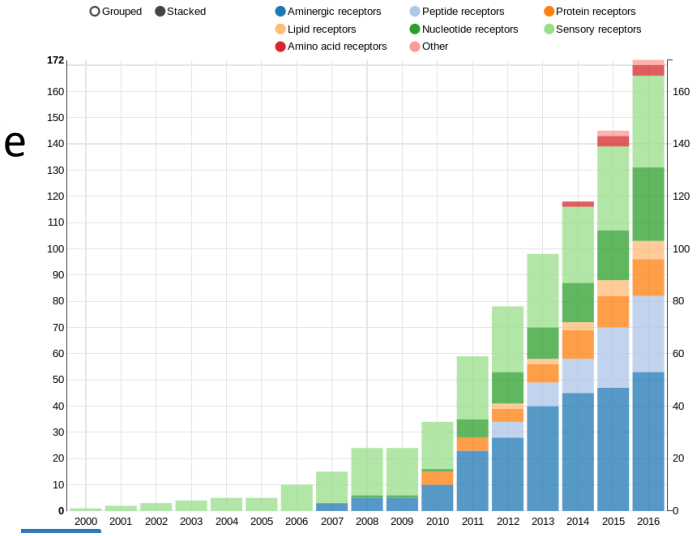
Jana Selent
GPCR drug discovery lab
IMIM & Pompeu Fabra University
Barcelona

1. Pompeu Fabra University: Jana Selent group
2. University of Copenhagen: David Gloriam
3. GPUGRID: Gianni de Fabritiis
4. Universitat Autonoma de Barcelona (Spain): Arnau Cordomi, Eduardo Mayol and Leonardo Pardo
5. Philipps-Universität Marburg: Maria Martí Solano and Peter Kolb
6. Paul Scherrer Institut (PSI): Xavier Deupi
7. Stockholm University: Jens Carlsson
8. Uppsala University: Hugo Gutiérrez de Terán
9. Institut für Medizinische Physik und Biophysik: Johanna Tiemann, Ramon Guixá and Peter Hildebrand
10. University of Warsaw: Sławomir Filipek
11. Stanford University: Rasmus Fonseca and A. J. Venkatakrishnan

Objective

1. **Creation of a GPCRmd database** (domain www.gpcrmd.org) accessible via GPCRdb

2. **Standard protocol for GPCR MDs**: an initial set of MDs for the GPCRmd database comprising 91 unique GPCR-ligand complexes



- Data transparency and reproducibility
- Exploitation of the data by other GPCR researcher
- Acceleration of GPCR research

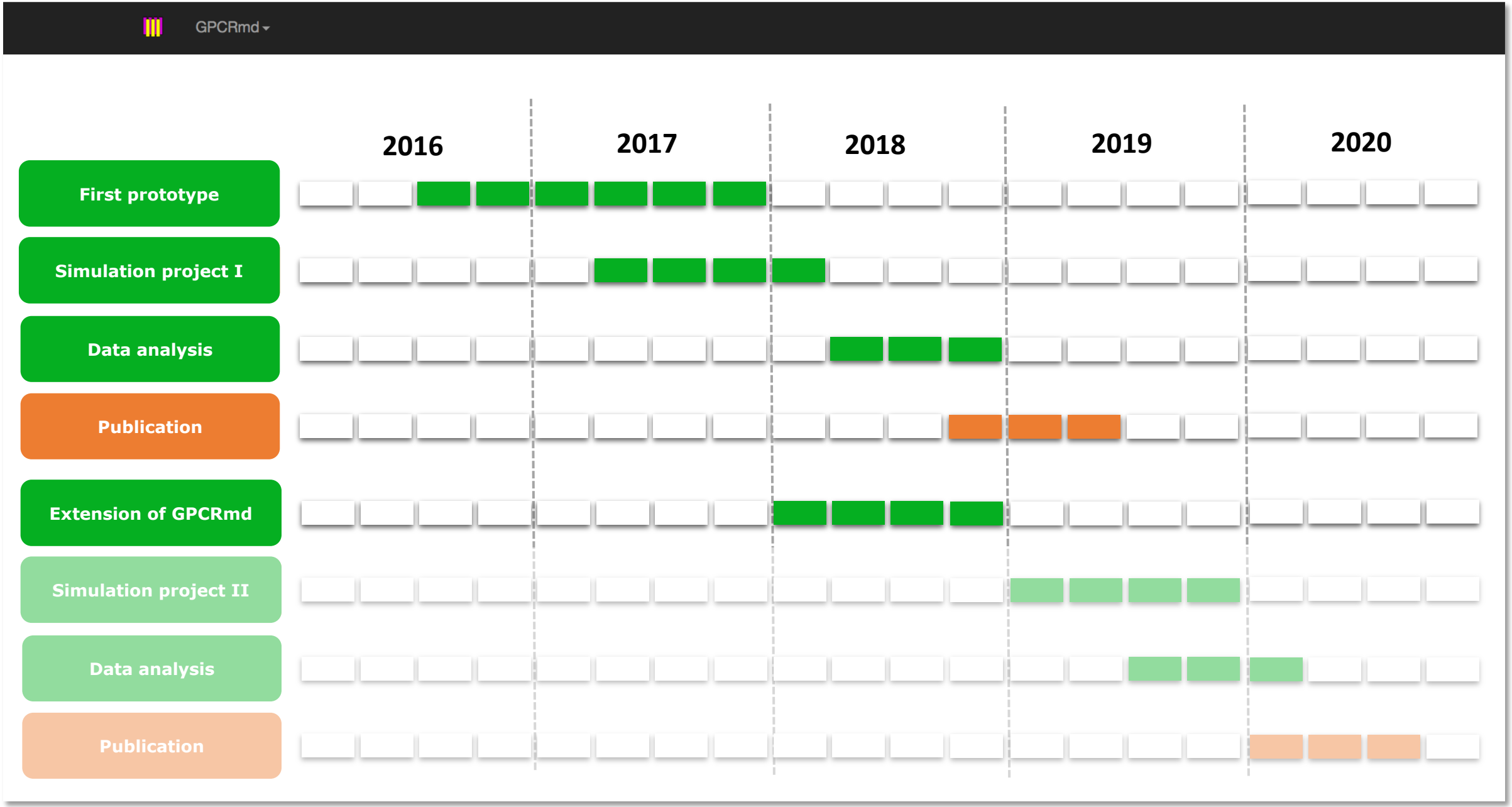
GPCRmd database

Main features

- Deposition /storage of MD data
- Query tools
- Visualization of MDs for relevant features using the NGL viewer
- Basic Analysis Tools
  - RMSD progression
  - Statistics about ligand-receptor interaction (contact frequencies) and receptor dynamics (helix bending, rotation, micro switches)
  - Interaction network dynamics - flare plot tool by R. Fonseca and AJ Venkatakrishnan

Extras

- Upon submission, a DOI and MD-ID is provided for **citations in papers**
- Non-MD experts can easily browse and visualize MDs of the database


GPCRmd
2016
2017
2018
2019
2020
First prototype
Simulation project I
Data analysis
Publication
Extension of GPCRmd
Simulation project II
Data analysis
Publication

Database design

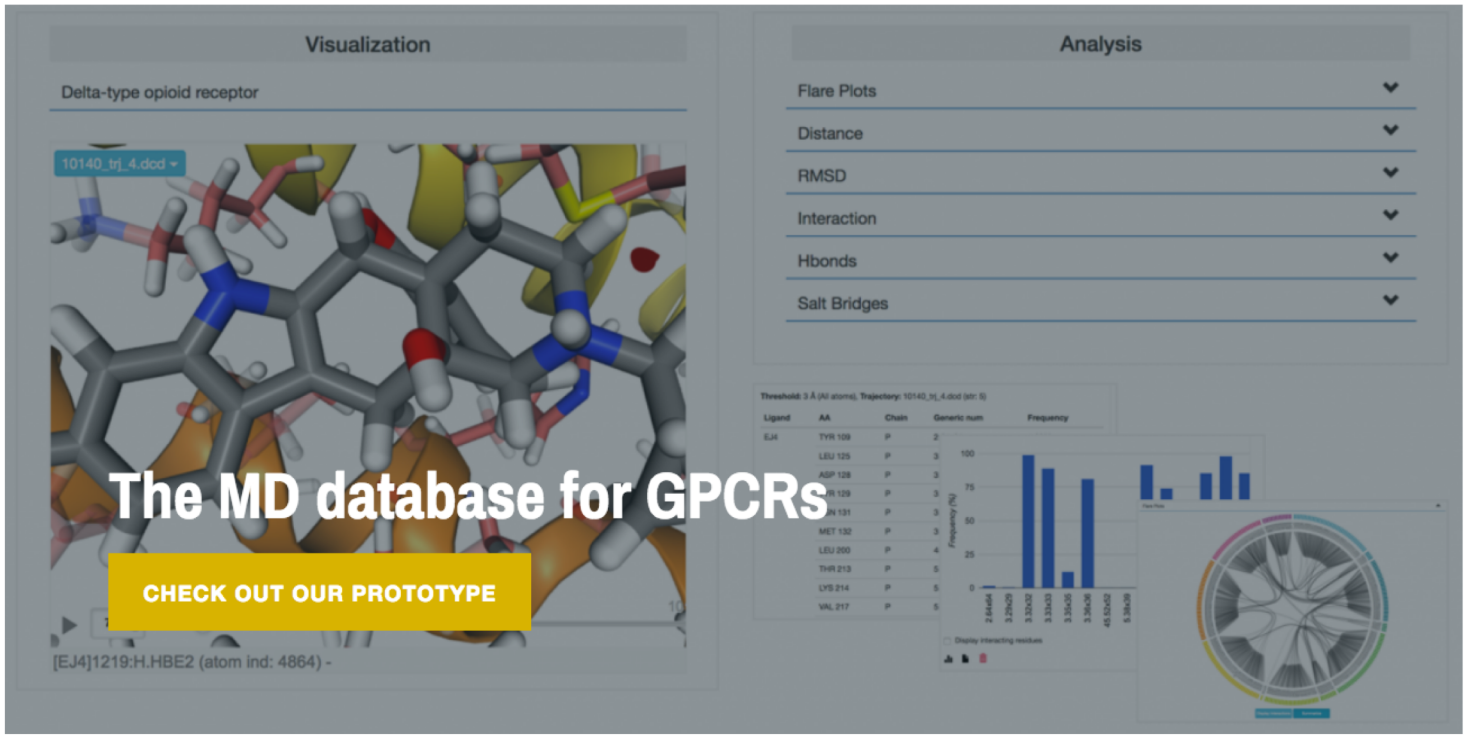
Database is created using the philosophy of the PDB and GPCRdb database:

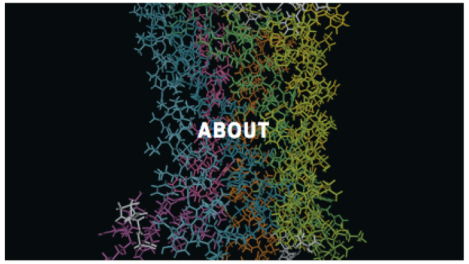
- Detailed information about proteins (GPCRs, signaling proteins and their mutants) and small molecules (ligand, lipids, etc.)
- Information about system preparation and dynamics protocol
- No special software needs to be installed locally to browse data

A first prototype GPCRmd database

GPCRMD

HOME ABOUT FEATURES COMMUNITY & MEETINGS CONTACT BROWSE SIMULATIONS

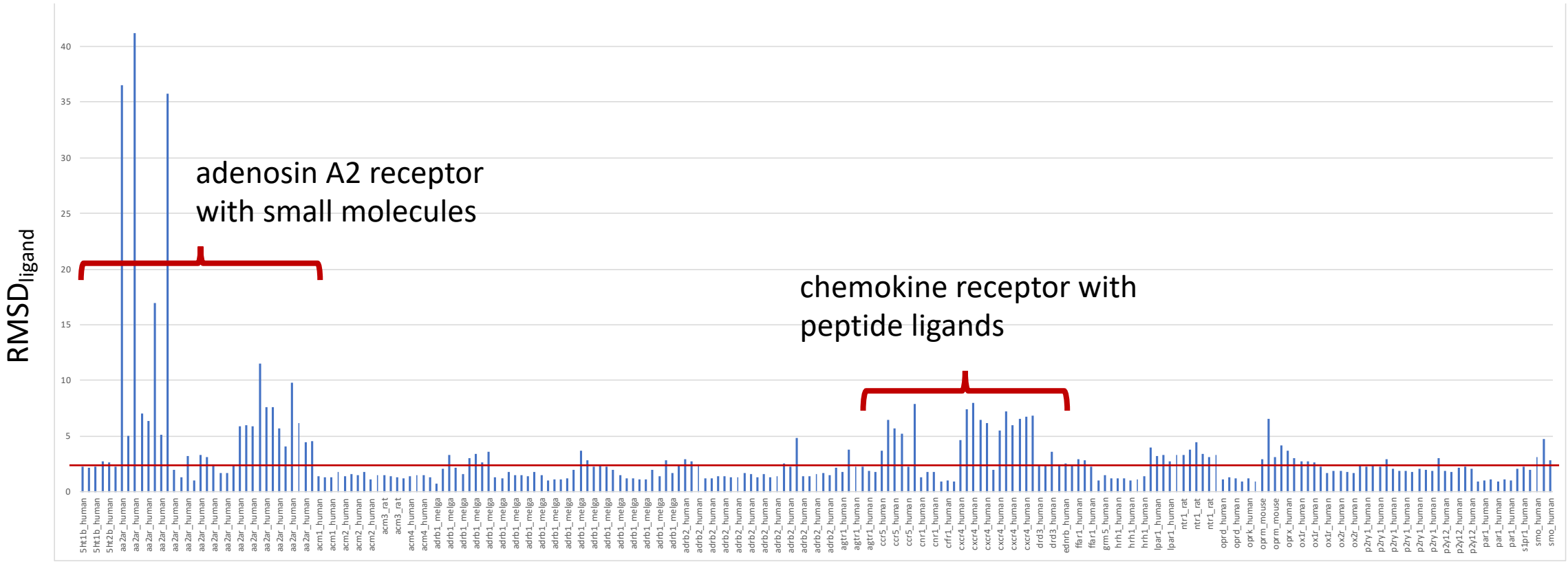




www.gpcrmd.org

GPCR simulation project

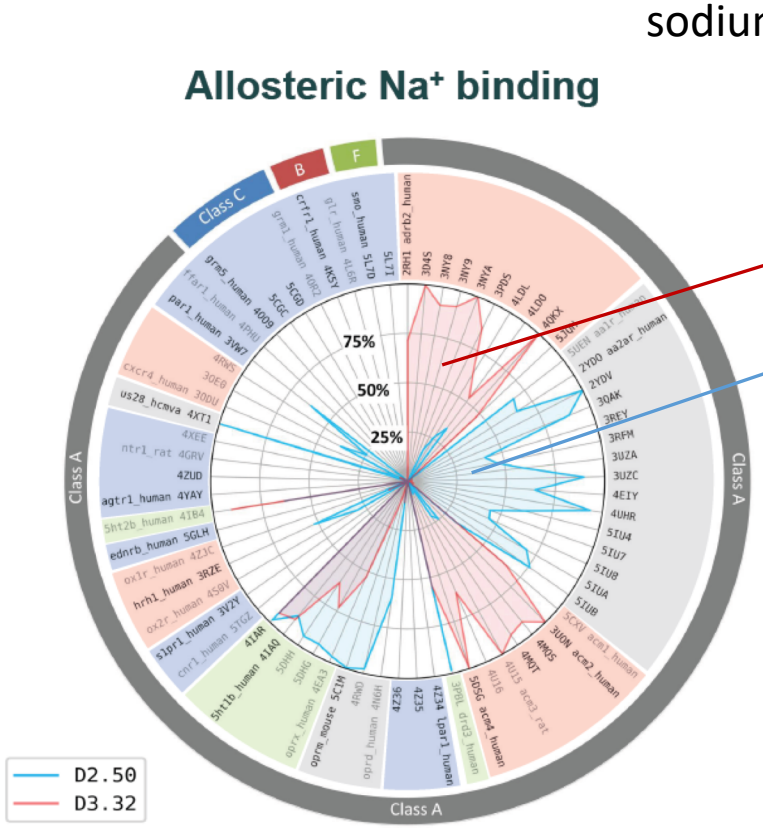
Quality of the simulation protocol for ligand-receptor complexes: ligand stability

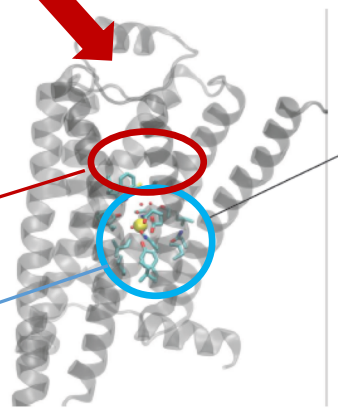


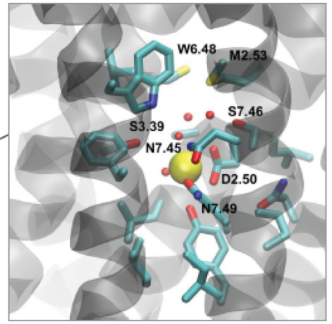
GPCR simulation project

Sodium binding to receptor apo-forms

sodium ions







Groups	Na^+ D2.50 freq	Na^+ D3.32 freq	HB/ ERB	Extracellular region net charge	D3.32	D2.50
Group 1	+++/++	–	–	-4 to -2, +0 (δOR), +2 (μOR)	+	+
Group 2a	+/–	+++/++	HB	-4 to -1	+	+
Group 2b	+/–, ++ (H1)	–	HB, ERB (H1)	-5 to -3, +1 (H1)	–	+
Group 3	+++/++	+++/++	–	-5 to -3	+	+
Group 4	–	–	CER (4b), ERB (4c)	NA	–	+ (4b, 4c), – (4a)

– < 10% freq.
\+ ≥ 10% freq.
++ ≥ 50% freq.
+++ ≥ 75% freq.

HB: Hydrophobic barrier
ERB: Extracellular region barrier
CER: Closed extracellular region

Extracellular region net charge = $N_{+\text{ charged residues}} - N_{-\text{ charged residues}}$

Submission system

GPCRmd

Simulation Submission Index

MOLECULAR DYNAMICS DATABASE FOR CRYSTALLIZED GPCR COMPLEXES

Follow the 4 steps for uploading simulations.

For more information on this form, see the docs.

Step 1:
Protein Information

Step 2:
Small Molecule Information

Step 3:
Complex Information

Step 4:
Dynamics Information

Simulation Submission Step 1: Protein Information

FILL IN THE PROTEIN FORM

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.

For more information on this form, see the docs.

Protein #1 General Information

(A) UniProtKB Accesion Code (AC) PROT #1

UniProtKB AC: P28222 Not available in UniprotKB?

Isoform: 1

It is not a GPCR.

Retrieve data

(B) Protein details PROT #1

Name: 5-hydroxytryptamine receptor 1B

Species: Homo sapiens (HUMAN)

You can search for scientific names and UniprotKB mnemonics here. Also, you can download UniprotKB species list and species not included there.

Aliases: Serotonin receptor 1B; 5-HT-1B; 5-HT1B; S12; Serotonin 1D beta receptor; 5-HT-1D-beta

Wild type sequence: MEEPGAQCAPPPPAGSETWVPQANLSSAPS QNCSAKDYIYQDSISLPWKVLLVMLLALITLAT

Is it a mutant?

Simulation Submission Step 2 Small Molecule Information

FILL IN THE SMALL MOLECULE FORM

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-crystalised molecule or if it belongs to bulk.

For more information on this form, see the docs.

Small Molecule #5 General Information

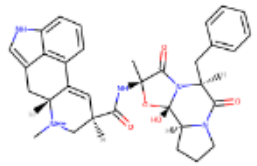
(A) Upload molecule — SMOL #5

Upload .mol/.sdf

Choose File | No file chosen

Upload

Uploaded Molecule



(B) Indicate if the uploaded structure is a co-crystalized molecule or if it belongs to bulk — SMOL #5

Co-crystalized molecule: Orthosteric ligand

Bulk (not co-crystalized):

Simulation Submission Step 4: Dynamics Information

FILL IN THE DYNAMICS FORM

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 case of replicated, you can upload all files simultaneously (e.g. 20 to 100 replicates). After the file upload, please click the validate button in to order to ensure correct upload.

(**B**) Please complete the table of all simulation components by adding non-crystallized bulk molecules (e.g. waters, ions and lipids). For

For more information on this form, see the docs.

Simulation Details and Files

(A) Upload the simulation files

Coordinate file — Choose File — No file chosen — Upload

Topology file — Choose File — No file chosen — Upload

Trajectory files — Choose Files — No file chosen — Upload

Simulation parameters — Choose File — No file chosen — Upload

Other files — Choose File — No file chosen — Upload

Validate

(B) Simulation components

Prot #	Protein name	UniProtKB AC
1	5-hydroxytryptamine receptor 1B	P28222

GPCRmd

Simulation Submission Step 4: Dynamics Information

click the validate button in to order to ensure 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.

For more information on this form, see the docs.

(B) Simulation components

Prot #	Protein name	UniProtKB AC
1	5-hydroxytryptamine receptor 1B	P28222

Resname	Molecule	Mol name	Num of mol	Type	Cryst
Cl-	0	None	1	Ions	☐
Na+	1	None	1	Ions	☐
POP	2	None	1	Lipid	☐
WAT	3	None	1	Water	☐
LIG	4	None	1	Ligand	☑

☑ Check the box if a simulation molecule has several resnames

+ Add Molecule | - Remove Molecule

Update the "**Molecule**" field in the added row to the value corresponding to the molecule with more than one "**Resname**".

Validate

Simulation Submission Step 4: Dynamics Information

click the validate button in to order to ensure 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.

For more information on this form, see the docs.

LIG	4	None	1	Ligand

Check the box if a simulation molecule has several resnames

Validate

(C) Simulation specifications

Method:	Molecular Mechanics		
Software:	ACEMD		
Software version:	2.11		
Force Field:	CHARMM		
FF version:	36		
Assay type:	Allosteric modulation		
Membrane type:	Homogeneous		
Solvent type:	TIP3P		
Solvent num:	45000	Num. Atoms:	45010
Time step:	4.0	Delta:	0.1
Additional Info:			

Submit

Back to step 3: Crystalized Components Information

Main challenges

- Data reproducibility
  - Starting files of simulation: coordinates and parameter files
  - Detailed information about the simulation (simulation software, simulation protocol, etc.)
- Data quality
  - We trust peer-reviewed data
- Data check upon upload to GPCRmd server
  - Small molecules: coordinates and topologies
  - Proteins: coordinates and sequence for pdb files

Main challenges

- Visualization tools for MD
  - Specialized software to display simulation data in Browser
  - Streaming simulation data (bandwidth, specialized software)
  - Solution: NGL viewer powered by WebGL and the MDsrv
- File formats accepted by GPCRmd
  - Visualization depends on NGL viewer
  - Analysis depends on MDtraj
- Simulation techniques
  - Classical unbiased MD
  - Limitation: Metadynamics, umbrella sampling, replica exchange, etc. ?
- Disk space and bandwidth: simulation size $\leq$2 GB

Acknowledgement

Creator of GPCRmd

Ismael Rodriguez-Espigares
Mariona Torrens Fontanals
Alejandro Varela
Juan Manuel Ramirez

GPCR MD community

1. Pompeu Fabra University: Jana Selent group
2. University of Copenhagen: David Gloriam
3. GPUGRID: Gianni de Fabritiis
4. Universitat Autonoma de Barcelona (Spain): Arnau Cordomi, Eduardo Mayol and Leonardo Pardo
5. Philipps-Universität Marburg: Maria Martí Solano and Peter Kolb
6. Paul Scherrer Institut (PSI): Xavier Deupi
7. Stockholm University: Jens Carlsson
8. Uppsala University: Hugo Gutiérrez de Terán
9. Institut für Medizinische Physik und Biophysik: Johanna Tiemann, Ramon Guixá and Peter Hildebrand
10. University of Warsaw: Sławomir Filipek
11. Stanford University: Rasmus Fonseca and A. J. Venkatakrishnan