



The Hitchhiker's Guide to Assay Central





Section I: Navigating the Bundle

Fundamentals are the building blocks of fun.

- If using OSX or Linux, locate the AssayCentral.app and double-click it to launch Assay Central.
 - If using Windows, locate the AssayCentral.jar file double-click it to launch Assay Central. A window will pop-up with instructions to access Assay Central that consist of clicking the Launch button. The same pop-up window will return information and your default browser (Chrome is preferred) will open a new window to a list of your models.

Assay Central Predictions

draw, paste or drag molecule

Prescriptions

- [Prescription Example](#)

Available Assay Datasets

Tagged Categories: [ADME](#) [CYP](#) [NTD](#)

	Title ↑	Target	Organism	Actives	Size	ROC	F1	Kappa	MCC	Domain	Invalid
Data Model Log	Chagas Disease: Broad (Cytotox)	Tulahuen-C4, CA-1/72 strains	T. cruzi	1692	4157	0.8023	0.6875	0.4891	0.4905	0.4146	
Data Model Log	CytochromeP450: 1A2	CytochromeP450 - 1A2	Human	310	810	0.9064	0.7945	0.6551	0.6580	0.2985	
Data Model Log	CytochromeP450: 2B6	CytochromeP450 - 2B6	Human	78	198	0.8387	0.7375	0.5597	0.5602	0.2287	
Data Model Log	CytochromeP450: 2C19	CytochromeP450 - 2C19	Human	246	789	0.8835	0.7410	0.6203	0.6206	0.2972	
Data Model Log	CytochromeP450: 2C9	CytochromeP450 - 2C9	Human	458	1203	0.8594	0.7394	0.5589	0.5634	0.3056	
Data Model Log	CytochromeP450: 2D6	CytochromeP450 - 2D6	Human	732	1668	0.8965	0.8049	0.6379	0.6408	0.3024	
Data Model Log	CytochromeP450: 3A4	CytochromeP450 - 3A4	Human	1106	2135	0.8919	0.8113	0.6129	0.6131	0.3224	
Data Model Log	ER alpha	Estrogen Receptor alpha	Human	890	1629	0.9607	0.9266	0.8347	0.8353	0.2821	
Data Model Log	ER beta	Estrogen Receptor beta	Human	770	1407	0.9563	0.9285	0.8353	0.8378	0.2692	
Data Model Log	HERG (HOU)	Kv11.1	Human	426	795	0.8511	0.7858	0.5775	0.5848	0.3207	
Data Model Log	Hepatitis B	Hepatitis B Virus	Hepatitis B Virus	139	895	0.9093	0.6089	0.5128	0.5446	0.2299	
Data Model Log	L. donovani: DNDi (Intracellular)	MHOM-ET-67/L82 strain	L. donovani	51	98	0.7664	0.7200	0.4286	0.4289	0.2511	
Data Model Log	Permeability: Caco-2	Caco-2	Human	175	424	0.8363	0.7500	0.5396	0.5505	0.2635	
Data Model Log	Toxoplasma gondii	multiple strains	Toxoplasma gondii	216	1080	0.9597	0.8242	0.7731	0.7837	0.3260	

- This page of Assay Central displays three headers:
 - Assay Central Predictions - includes a gray box for drawing molecules and a “Predict” button linking to another page.
 - Prescriptions - one example link is provided and is discussed in Section IV.
 - Available Assay Datasets - a list of models in this bundle and pertinent metrics.
- Each row in this list is a single Bayesian model. Columns include information pertaining to the data or model statistics:
 - Title - usually named according to the target or disease
 - Target - a specific protein or organism stage/strain
 - Organism - the host of the target
 - Actives - how many compounds are designated as active in the model
 - Size - how many total compounds comprise the model
 - Model metrics (all range from 0-1 ; the closer to 1.0, the more predictive the model is)
 - ROC - a percent probability of correctly predicting a compound
 - Additional Metrics - F1 Score, Cohen’s Kappa, and Matthew’s Correlation Constant
 - Domain - how much chemical space is covered by the model in relation to ChEMBL, ranging from 0-1

- The vertical arrow next to a column header indicates how the list is sorted, and its direction indicates ascending or descending order.
 - By default, the list is sorted in ascending order of the model titles.
- Above your list of models are a series of keywords following the word “Tagged Categories”.

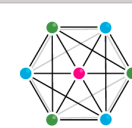
Tagged Categories: ADME CYP NTD

- Tags are a way to organize Assay Central with keywords.
- By clicking a keyword, only the subset of models associated with that tag are visible.
- Selecting multiple tags will visualize multiple categories, but click again to deselect any tags and regain the original list.

Assay Central

Assay Central Predictions

draw, paste or drag molecule



Prescriptions

- [Prescription Example](#)

Available Assay Datasets

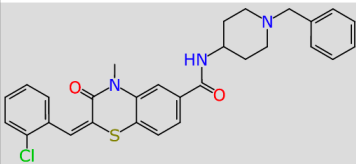
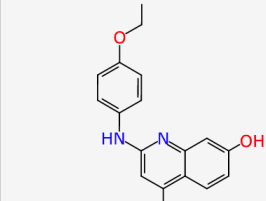
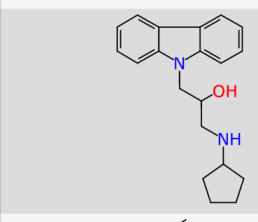
Tagged Categories: ADME CYP NTD

			Title ↑	Target	Organism	Actives	Size	ROC	F1	Kappa	MCC	Domain	Invalid
Data	Model	Log	Chagas Disease: Broad (Cytotox)	Tulahuen-C4, CA-1/72 strains	T. cruzi	1692	4157	0.8023	0.6875	0.4891	0.4905	0.4146	
Data	Model	Log	Hepatitis B	Hepatitis B Virus	Hepatitis B Virus	139	895	0.9093	0.6089	0.5128	0.5446	0.2299	
Data	Model	Log	L. donovani: DNDi (Intracellular)	MHOM-ET-67/L82 strain	L. donovani	51	98	0.7664	0.7200	0.4286	0.4289	0.2511	
Data	Model	Log	Toxoplasma gondii	multiple strains	Toxoplasma gondii	216	1080	0.9597	0.8242	0.7731	0.7837	0.3260	

- Each dataset has two buttons to the far left of the row, “Data” and “Model”.
 - Click **Data** and a new window similar to the figure on the left will appear.
 - Model input details are listed here. Details include thresholds, target organisms and proteins, relevant assay information (i.e. cell lines and incubation times), and data sources (i.e. ChEMBL/PubChem assay IDs).
 - The first 100 compounds that are present in the model of interest are displayed beneath input details, as well as molecule identifiers, activity, and links to external sources.
 - Activity values listed here are in units of $-\log[M]$ (i.e. 7 = 100 nM and 6 = 1 μ M) or are binary (i.e. 1 = active and 0 = inactive). Thresholds are always listed under “Description”.
 - If CAS numbers are available for molecules, links to the NIST Chemistry WebBook are provided here.
 - Click **Model** and a new window similar to the figure on the right will appear.
 - This area includes all of the metrics from the Introduction page with additional common metrics (precision, recall, and specificity) as well as an image of the ROC curve and truth table outline true/false positives, true/false negatives.

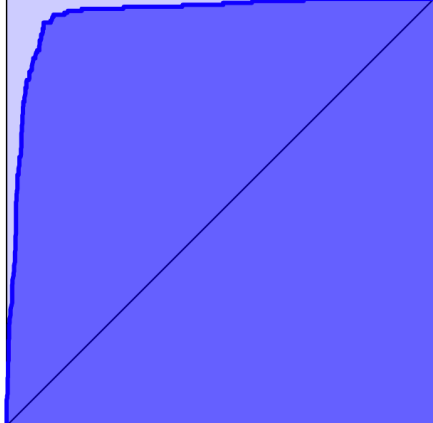
Toxoplasma gondii

Title: Toxoplasma gondii
Description: Collated by KMZ from ChEMBL Target ID 613497, specifically viral DNA replication ; IC50/EC50 measurements ; Threshold set by KMZ at 1 μ M
Target: multiple strains
Organism: Toxoplasma gondii
Target Type: Organism
Assay Type: Functional
Source: (ChEMBL assay IDs 730643, 813899, 833904, 927191, 939263, 1006440, 1112924, 1228266, 1775524, 2072979, 2185137, 3301450, 3372741, 3773071, 3832822, 1217005, 1264574, 2330479, 3223458, 3736777)
Units: $-\log_{10}[M]$

Molecule	Name	Value	Source
		< 4.6327	act:3048876
		< 4.5480	act:3048877
		< 4.5229	act:3048878 act:14773218

Assay Central

Toxoplasma gondii

Origin: Assay Central
Field: Toxo_gondii
Comments: Domain Compatibility: 0.32596973
Training Actives: 216 / 1080
ROC: 0.9597 (five-fold)
Curve: 

Truth Table:

		Predicted	
		Yes	No
Actual	Yes	204	12
	No	75	789

Precision: 0.7312
Recall: 0.9444
Specificity: 0.9132
F1 score: 0.8242
Kappa: 0.7731
MCC: 0.7837





Section II: Generating & Interpreting Predictions

Read the tarot cards.

- From the Introduction page, click **Predict** to open the Prediction page.
- The page is quite blank, except for a list of models at the right, separated into “Targets” and “Off-targets” categories.
 - In general, “Target” implies that the model is based off of a disease endpoint (i.e. Ebola, Schistosomiasis) while “Off-target” implies that the model is based off of an undesirable target (i.e. Cytochrome P450s) or a molecular property (i.e. Permeability, Solubility).
- Hover (do not click) over a model title, and a small pop-up bubble will appear with the model tags, input details, and total compounds (size) in addition to ROC, F1 Score, and Cohen’s Kappa metrics from the Introduction page.

The screenshot shows the Assay Central web interface. The main heading is "Predictions". Below it, there is a text prompt: "Provide molecules by sketching, pasting or dragging, and use the existing models to predict activities." There are two buttons: "Draw Molecule" (highlighted in blue) and "Example". Below this is the "Molecules" section, which currently says "No molecules defined yet." To the right, there are two panels: "Targets" and "Off-targets". The "Targets" panel has "All" and "None" buttons and a list of targets with checkboxes and honeycomb icons: Chagas Disease: Broad (Cytotox), Hepatitis B, L. donovani: DNDi (Intracellular), and Toxoplasma gondii. The "Off-targets" panel also has "All" and "None" buttons and a list of off-targets with checkboxes and honeycomb icons: CytochromeP450: 1A2, 2B6, 2C19, 2C9, 2D6, 3A4, ER alpha, ER beta, HERG (HOU), and Permeability: Caco-2. A pop-up bubble is shown over the "Permeability: Caco-2" target, displaying the following information: Tags: ADME; Target: Caco-2; Organism: Human; Target Type: Membrane; Cell: Heterogeneous human epithelial colorectal adenocarcinoma; Compounds: 424; ROC: 0.8363; F1: 0.7500; Kappa: 0.5396. At the bottom left, there is a "Honeycomb" section with the text: "Provide a molecule, select a target and click the honeycomb (honeycomb icon) icon."

- Generate a list of molecules by either clicking **Draw Molecule** to manually draw a structure, or **Example** up to five times.
 - There are delete and edit buttons to the left of each structure represented by  and , respectively.
- Additional methods of generating a list of molecules:
 - A structure file (.sdf or .ds) containing up to 100 compounds can also be dragged onto this page to generate a list of structures and names! Files over 100 compounds can be dragged in, but 100 compounds at a time can be predicted.
 - A molecule can also be drawn from clicking the gray box on the Introduction page.

- By default, the Prediction Display is set to “Selected” (highlighted blue).
- Select models by either click individual white squares to the left of each model at the right of the page, or click **All** above each model category.
 - If tags were selected on the Introduction Page, only that subset will show up on the Prediction page.
 - A colored square for each model will appear to the right of each compound, and a color scale is shown beneath.
- Hover (do not click) your cursor on one of the colored squares.
 - A pop-up bubble will appear with the model title, as well as Prediction and Applicability scores for the compound.
 - These squares are in the same order as the model list at the right, separated by categories.

Assay Central

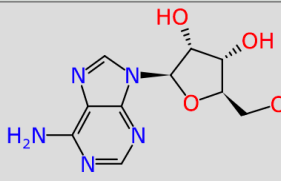
Predictions

Provide molecules by sketching, pasting or dragging, and use the existing models to predict activities.

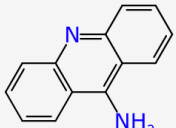
[Draw Molecule](#) [Example](#)

Molecules

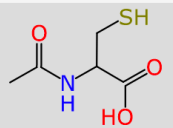
Prediction Display: **Selected** Ranked Flower



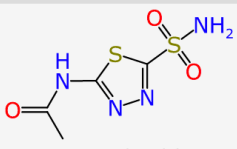
adenosine



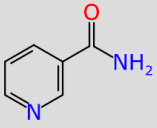
9-acridinamine



N-acetylcysteine





acetazolamide



nicotinamide

0 0.5 1





Target  (green is good)

Off-Target  (red is bad)

[Download DataSheet](#)
[Download SDfile](#)
[Download Text](#)











Targets

[All](#) [None](#)

- Chagas Disease: Broad (Cytotox) 
- Hepatitis B 
- L. donovani: DNDi (Intracellular) 
- Toxoplasma gondii 

Off-targets

[All](#) [None](#)

- CytochromeP450: 1A2 
- CytochromeP450: 2B6 
- CytochromeP450: 2C19 
- CytochromeP450: 2C9 
- CytochromeP450: 2D6 
- CytochromeP450: 3A4 
- ER alpha 
- ER beta 
- HERG (HOU) 
- Permeability: Caco-2 

- Interpretation of a Prediction score:
 - This generally ranges from 0-1 (zero being inactive at the target of interest) and colored according to the scale below the predicted compound.

It is possible to have larger or negative scores, due to the calculation method. See Appendix A for more information about how prediction scores are generated

Prediction Scores remain the same between Target and Off-Target models, the only difference is how you interpret the score.
 - Despite category, a higher score is always predicted to be more active at the model target.
 - For example, a higher score in caco-2 permeability means the compound is predicted to cross the cell membrane, and a higher score for a CYP450 means it is more likely to be metabolized by that enzyme.
- Interpretation of an Applicability score (also known as Model Overlap):
 - This is the percent of fragments from the predicted molecule that are present in the model.
 - An Applicability of 1.000 can indicate that your compounds is present in the model, but this is not always true, especially for small molecules. See if this is the case with hexplots, described on page 13.
 - The Applicability score is also represented by the length of the black bar beneath each square.
- Once you generate information you would like to save, you can download a single file containing molecular formulas, compound names, as well as the Prediction and Applicability scores for each compound, in .ds, .sdf, or .tsv formats.
 - Click the corresponding button beneath the predicted molecules, above the Honeycomb viewing area of the page.
 - This information for multiple models will be downloaded if multiple models are selected.

[Download DataSheet](#)[Download SDfile](#)[Download Text](#)



Section III: Visualizing Predictions

A picture is worth a thousand words.

- Generate a list of compounds to predict on the Predictions page and select the models of your choice to view scores.
- Click on one of the colored squares that appear to generate a visualization of the chemical fragment activity.
 - Atoms are colored according to “active” and “inactive” features.
 - Note that atom-coloring can only be implemented in this view.
 - A color scheme for Target and Off-Target models is presented below the predicted compounds list, applicable for both prediction scores and atom-coloring of features.

Assay Central

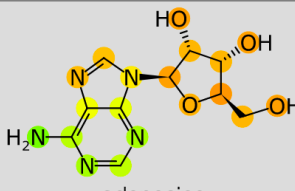

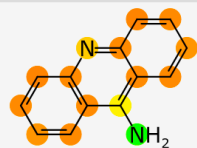

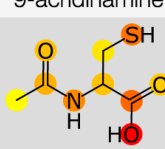

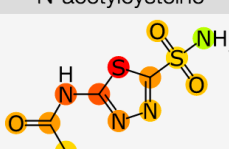

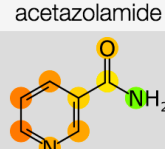

Predictions

Provide molecules by sketching, pasting or dragging, and use the existing models to predict activities.


[Draw Molecule](#) [Example](#)


Molecules

Prediction Display: [Selected](#) [Ranked](#) [Flower](#)

		Toxoplasma gondii Prediction: 0.4480 Applicability: 0.7170
		
		
		
		

0 0.5 1





Target  (green is good)

Off-Target  (red is bad)

[Download DataSheet](#) [Download SDfile](#) [Download Text](#)











Targets

All None

- Chagas Disease: Broad (Cytotox) 
- Hepatitis B 
- L. donovani: DNDi (Intracellular) 
- Toxoplasma gondii 

Off-targets

All None

- CytochromeP450: 1A2 
- CytochromeP450: 2B6 
- CytochromeP450: 2C19 
- CytochromeP450: 2C9 
- CytochromeP450: 2D6 
- CytochromeP450: 3A4 
- ER alpha 
- ER beta 
- HERG (HOU) 
- Permeability: Caco-2 

- Change your Prediction Display from “Selected” to “Flower”.
 - All of the squares become slices of a prediction pie, with a thick black line separating on/off targets.
 - The length of the slice of pie is the visual representation of Applicability, like the black bar beneath the squares.
 - Color schemes of the slices are identical to the squares and are shown at below the compound list. Atom-coloring can be viewed in this Prediction Display, but cannot be changed.

Assay Central

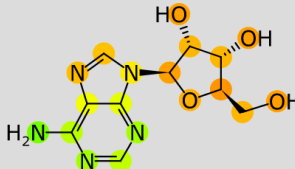
Predictions

Provide molecules by sketching, pasting or dragging, and use the existing models to predict activities.

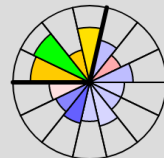
[Draw Molecule](#) [Example](#)

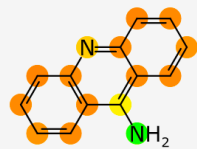
Molecules

Prediction Display: [Selected](#) [Ranked](#) [Flower](#)

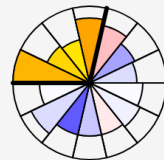



adenosine



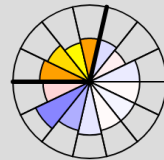


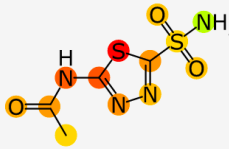
9-acridinamine



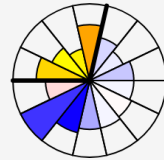


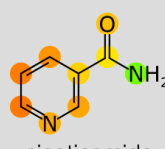
N-acetylcysteine



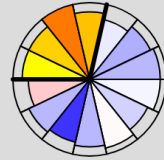


acetazolamide







nicotinamide



0 0.5 1

Target  (green is good)

Off-Target  (red is bad)

[Download DataSheet](#)
[Download SDfile](#)
[Download Text](#)

Targets

[All](#) [None](#)

- Chagas Disease: Broad (Cytotox) 
- Hepatitis B 
- L. donovani: DNDi (Intracellular) 
- Toxoplasma gondii 

Off-targets

[All](#) [None](#)

- CytochromeP450: 1A2 
- CytochromeP450: 2B6 
- CytochromeP450: 2C19 
- CytochromeP450: 2C9 
- CytochromeP450: 2D6 
- CytochromeP450: 3A4 
- ER alpha 
- ER beta 
- HERG (HOU) 
- Permeability: Caco-2 

- Deselect all models by clicking **None** above each category, and change the Prediction Display to “Ranked”.
- Colored squares will appear representing the five models that the compound scores highest in, from all of the models provided on the right within both model categories.
- If you select individual models prior to ranking, only five models of those selected are displayed.
- Color schemes of the slices are identical to the squares and are shown at below the compound list. Atom-coloring can be viewed in this Prediction Display, but cannot be changed.

Assay Central

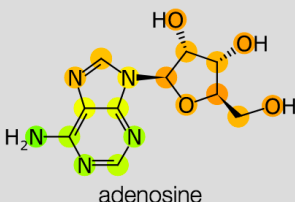
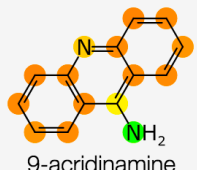
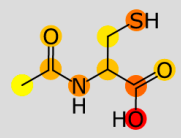
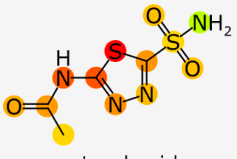
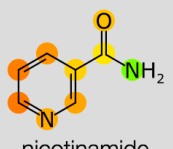
Predictions

Provide molecules by sketching, pasting or dragging, and use the existing models to predict activities.

Draw Molecule Example

Molecules

Prediction Display: Selected **Ranked** Flower

 <p>adenosine</p>	<input checked="" type="checkbox"/> Hepatitis B <input checked="" type="checkbox"/> Toxoplasma gondii <input checked="" type="checkbox"/> L. donovani: DNDi (Intracellular) <input checked="" type="checkbox"/> Chagas Disease: Broad (Cytotox)	<input type="checkbox"/> CytochromeP450: 2B6 <input type="checkbox"/> Permeability: Caco-2 <input type="checkbox"/> CytochromeP450: 2C9 <input type="checkbox"/> CytochromeP450: 3A4 <input type="checkbox"/> ER alpha
 <p>9-acridinamine</p>	<input checked="" type="checkbox"/> L. donovani: DNDi (Intracellular) <input checked="" type="checkbox"/> Hepatitis B <input checked="" type="checkbox"/> Chagas Disease: Broad (Cytotox) <input checked="" type="checkbox"/> Toxoplasma gondii	<input type="checkbox"/> CytochromeP450: 1A2 <input type="checkbox"/> CytochromeP450: 3A4 <input type="checkbox"/> Permeability: Caco-2 <input type="checkbox"/> CytochromeP450: 2C9 <input type="checkbox"/> CytochromeP450: 2D6
 <p>N-acetylcysteine</p>	<input checked="" type="checkbox"/> L. donovani: DNDi (Intracellular) <input checked="" type="checkbox"/> Hepatitis B <input checked="" type="checkbox"/> Toxoplasma gondii <input checked="" type="checkbox"/> Chagas Disease: Broad (Cytotox)	<input type="checkbox"/> Permeability: Caco-2 <input type="checkbox"/> CytochromeP450: 2B6 <input type="checkbox"/> CytochromeP450: 3A4 <input type="checkbox"/> CytochromeP450: 2C9 <input type="checkbox"/> CytochromeP450: 2D6
 <p>acetazolamide</p>	<input checked="" type="checkbox"/> Hepatitis B <input checked="" type="checkbox"/> L. donovani: DNDi (Intracellular) <input checked="" type="checkbox"/> Chagas Disease: Broad (Cytotox) <input checked="" type="checkbox"/> Toxoplasma gondii	<input type="checkbox"/> Permeability: Caco-2 <input type="checkbox"/> CytochromeP450: 2D6 <input type="checkbox"/> CytochromeP450: 3A4 <input type="checkbox"/> CytochromeP450: 2C9 <input type="checkbox"/> CytochromeP450: 2B6
 <p>nicotinamide</p>	<input checked="" type="checkbox"/> Chagas Disease: Broad (Cytotox) <input checked="" type="checkbox"/> Hepatitis B <input checked="" type="checkbox"/> Toxoplasma gondii <input checked="" type="checkbox"/> L. donovani: DNDi (Intracellular)	<input type="checkbox"/> Permeability: Caco-2 <input type="checkbox"/> CytochromeP450: 3A4 <input type="checkbox"/> CytochromeP450: 2C9 <input type="checkbox"/> CytochromeP450: 1A2 <input type="checkbox"/> CytochromeP450: 2D6

Targets

All None


- Chagas Disease: Broad (Cytotox)
- Hepatitis B
- L. donovani: DNDi (Intracellular)
- Toxoplasma gondii


Off-targets




All None

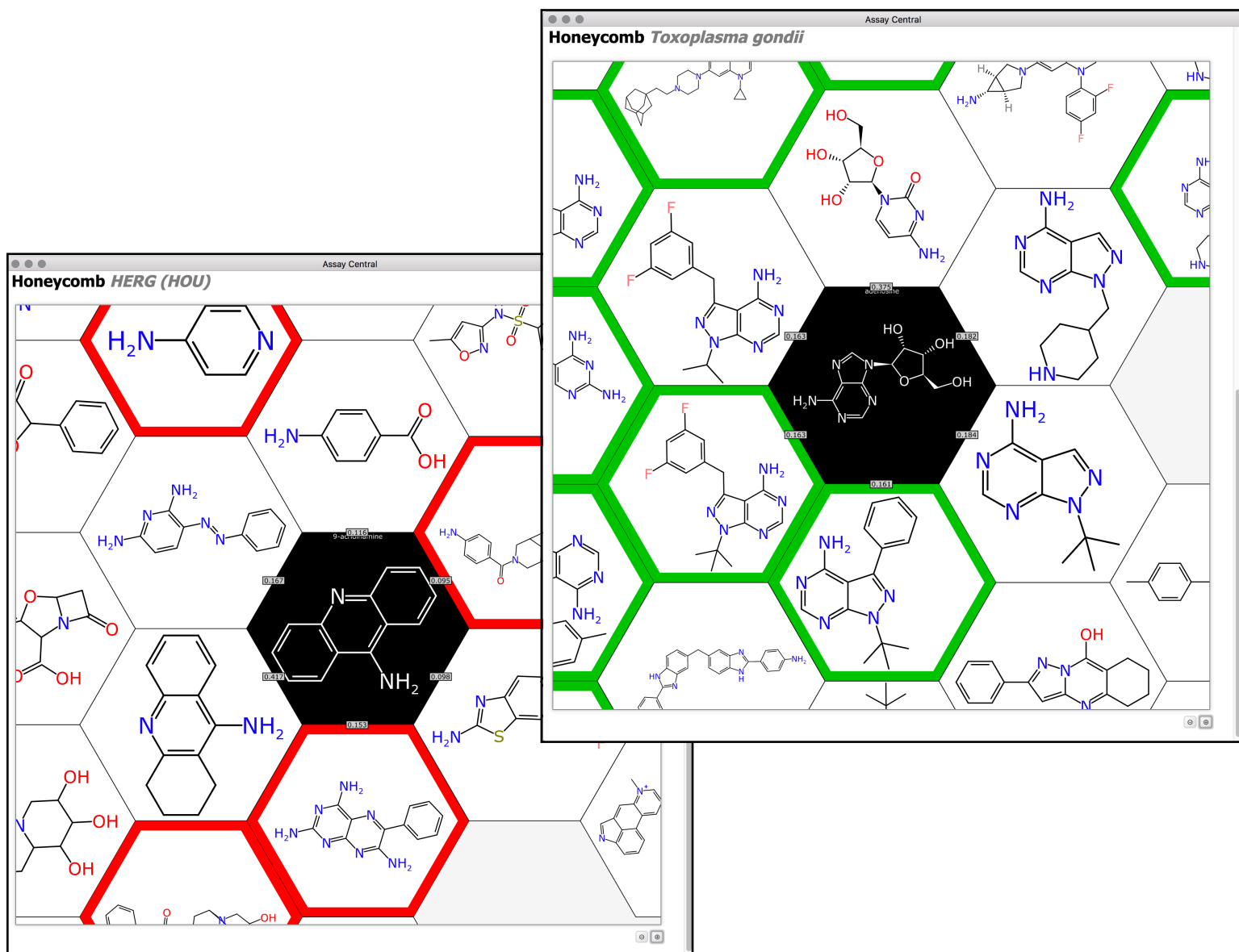
- CytochromeP450: 1A2
- CytochromeP450: 2B6
- CytochromeP450: 2C19
- CytochromeP450: 2C9
- CytochromeP450: 2D6
- CytochromeP450: 3A4
- ER alpha
- ER beta
- HERG (HOU)
- Permeability: Caco-2

0 0.5 1

Target  (green is good)

Off-Target  (red is bad)

- Click  to the right of a model to generate a honeycomb, or hexplot.
 - Predicted compounds (black hexagons) are presented in relation to compounds within the model input (white hexagons).
 - Depending on if the selected model is from the Target or Off-Target category, hexagons of active compounds are outlined in either green or red, respectively.
 - Left-click anywhere in this space and move the mouse to pan around.
 - Zoom out/in using   at the bottom right of the image.
 - In each hexagon, you will see name/CAS/identifier of each compound if they were provided.
- The hexplot is organized by compound similarity, evaluated with ECFP6 descriptors.
 - Honeycomb layout depends on the starting molecule chosen by the user (by default the first molecule in the list).
 - As the hexagons get farther from the starting molecule, they are more dissimilar.
 - Double-clicking on a different black hexagon will regenerate the entire hexplot.
 - If you hover (do not click) your cursor a compounds within a hexagon, you will see the Tanimoto similarity coefficient of the compounds to its six nearest neighbors, ranging from 0-1.
 - The user can use human intuition and knowledge to further deduce how reliable the prediction score is. For example, compare the atom-coloring of the predicted molecules from page 11 to the honeycomb on the right.





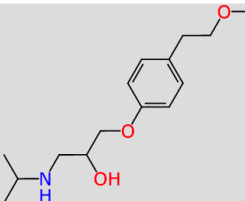


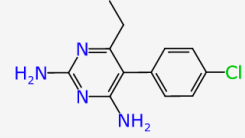


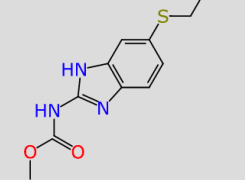


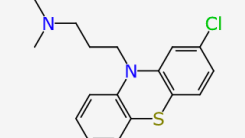


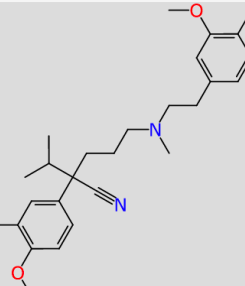


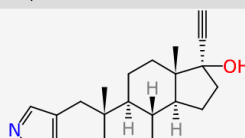


Section IV: Prescriptions

The Good, The Bad, & The Ideal.

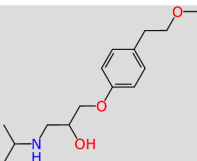


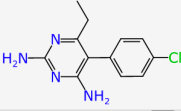


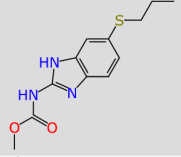


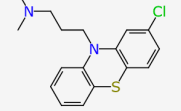


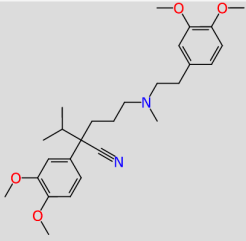


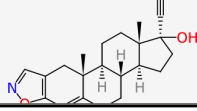


- Return to the Introductory page and click on the link at the top of the page titled “Prescription_Example”.
 - Prescriptions are a method of selecting compounds that hit your desired targets, have good molecular properties, and are inactive at your undesired targets.
 - At this time, only a CPI employee can generate a prescription. Please contact CPI if you would like to receive this service.
- This page is set up similarly to the Predictions page.
 - There is a compound with a name, a score, and several boxes representing a model colored according to the same on and off-target schemes seen previously.
 - Hovering over a square gives you the full path to the model’s directory on CPI machines, as well as the individual prediction score and applicability seen previously.
 - When a square has a black triangle in half of the area, it means that the molecule was present in the model and returns a statement of whether it was considered active or inactive, instead of a prediction score.

Prescription_Example

Edit Download DataSheet

	Molecule	Compound	Score	Targets	Off-targets
1		metoprolol	0.3507		
				ADME/Permeability/caco2 Known: active	
2		pyrimethamine	0.3319		
3		albendazole	0.2820		
4		chlorpromazine	0.2820		
5		verapamil	0.2817		
6		danazol	0.2802		

- Here, “Targets” and “Off-targets” do not necessarily correspond to what is presented on the Prediction page.
 - When a prescription is generated, each model is redefined by ‘desirable’ or ‘undesirable’ activity.
 - For example, Permeability classified as an Off-target on the Prediction page, but it is a desirable trait for potential drugs and is present in the “Targets” column below. If requested, Permeability can be defined as an undesirable trait and would be in the “Off-targets” column.
- The “Score” provided on this page is calculated by taking the difference of the average, calibrated Prediction score of the desired models and the the undesired ones.
 - A positive score indicates higher prediction scores from desired models over undesired, and vice versa.
- Click [Edit](#) at the top left corner of the page.
 - A new column “Measurement” appears between the Compound and Score.
 - After testing a compound, a measurement value can be added here as shown in row 1.
 - This page can be downloaded as a file in DS format (more formats to come) by clicking at [DataSheet](#) the top left corner of the page.

	Molecule	Compound	Measurement	Score	Targets	Off-targets
1		metoprolol	50 ± 2.4 nM	0.3507		
2		pyrimethamine	± Units	0.3319		
3		albendazole	± Units	0.2820		
4		chlorpromazine	± Units	0.2820		
5		verapamil	± Units	0.2817		
6		danazol	± Units	0.2802		



Section V: Important Notes

Gentle reminders to make life easier.

- Take care with how you depict compound structures, as the slightest change can alter the prediction score.
 - Show stereochemistry when possible, and please remove any salt/adducts prior to predicting!
 - When models are built the salts are removed, so to be consistent salts should be removed from prediction datasets beforehand. Scores will not *drastically* change, but will be a difference and one should be mindful of this.
- Larger models mean larger data to process. These require more time to load, so please be patient!
- Models are only as good as the data that populates it and our models are constantly being updated!
- At your request, CPI able to predict compounds from a supplied dataset or one of our available compound libraries, while solely exporting the molecules NOT present within models of interest.
 - Ranking of a compound library is available by request. Rather than what is shown in Assay Central, this prediction method outputs the most diverse structures of the top scoring molecules.
 - Prescriptions cannot be generated in Assay Central. If requested, a CPI employee will generate it and recreate and deliver a new bundled app that will include the new prescription.
- Do not hesitate to contact someone at CPI if...
 - ...you would like private models created.
 - ...you would like to receive updated bundles with new data, or request modifications to existing model details.
 - ...you would like to request a model of a specific target/measurement that would aid research efforts.
 - ...you are experiencing difficulties or have questions not addressed in this manual.
 - ...you have any feedback that would optimize the user experience.



Appendix A

The jargon translator.

General Jargon

- Prediction score
 - Score generated on the prediction page
 - Translates to: A probability that a compound will be active against the target modeled.
 - Current threshold for active/inactive is a score of 0.5
 - For more details, see DOI: 10.1021/acs.jcim.5b00143
- Prescription score
 - A consensus score between desirable and undesirable activity.
 - A positive score means the compound scores collectively higher (more probable to be active) in the desired models (“Targets” on the Prescription page) than in the undesired models (“Off-Targets” on the prescription page).
- Domain
 - Translates to: How much chemical space is covered by the model in relation to ChEMBL, ranging from 0-1.
- Applicability or Model Overlap
 - Translates to: The percent of fragments from the predicted molecule that are present in the model.
 - Lesser score indicates that the prediction is not as useful, because the molecules comprising the model do not contain a large portion of the predicted compound.
- Honeycomb
 - A method of visualizing the predicted molecules in relation to those comprising the model itself.
 - Molecules are ordered by decreasing similarity with distance from the initial molecule designated by the user.
 - For more details, see DOI: 10.1021/acs.jcim.6b00004



Appendix B

Aid for the treacherous journey ahead.

- Useful tools available in the Apple App Store, also created by our CTO (Dr. Alex M. Clark)
 - Polypharma
 - TB Mobile
 - MMDS (mobile) and XMDS (desktop)
- Public web-version of Assay Central available at <http://www.assaycentral.org>
- Contacting Collaborations Pharmaceuticals, Inc.
 - ★ Company website: <http://www.collaborationspharma.com>
 - Dr. Sean Ekins, CEO: sean@collaborations.com
 - Dr. Maggie Hupcey, COO: maggie@collaborations.com
 - Dr. Thomas R. Lane, PostDoctoral Fellow: tom@collaborationspharma.com
 - Kimberley M. Zorn, Staff Scientist: kim@collaborationspharma.com