Druggability

A protein is 'druggable' if its activity can be modulated *in vivo* by its binding to a drug-like compound. A compound is considered drug-like if it has properties that would make it suitable for use as a drug, e.g it must have chemical properties that enable it to cross cell membranes.

In contrast, the term 'tractable' is used to describe proteins whose activity can be modulated *in vitro* by a chemical probe compound, for example in a biochemical assay. Chemical probes are generally not drug-like, and are used to facilitate drug development.

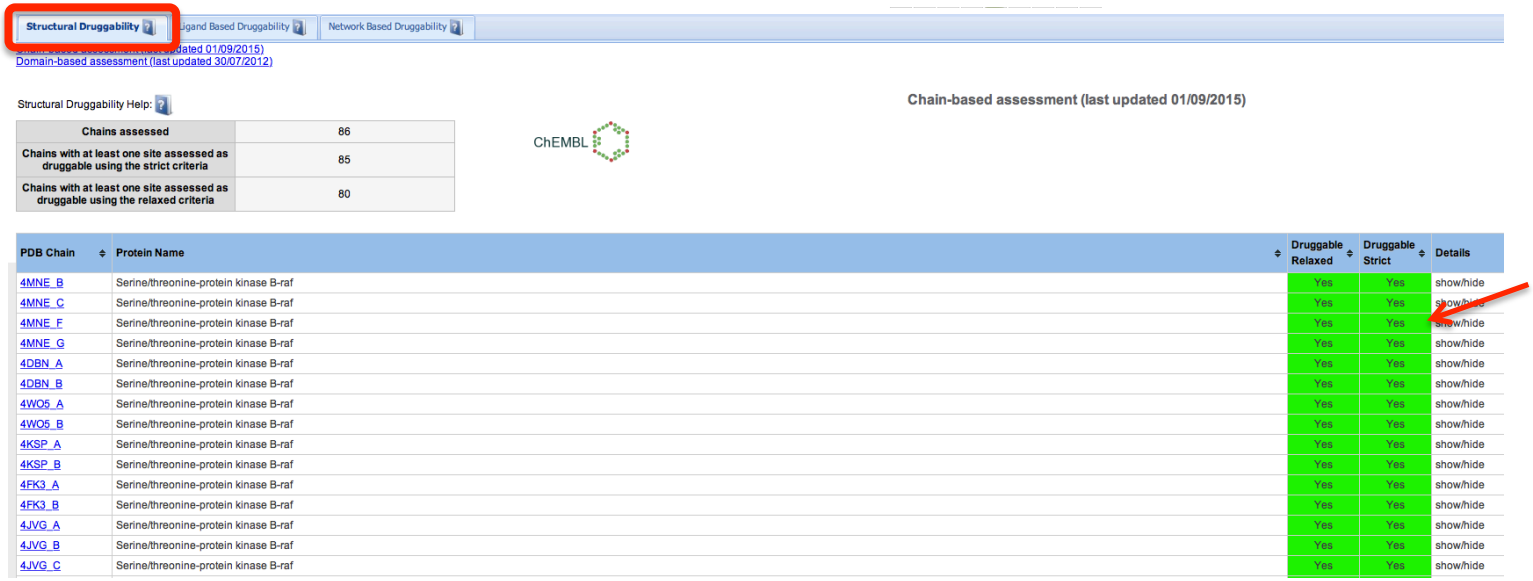
canSAR scores proteins on their druggability using three approaches:

- Structural druggability
- Ligand-based druggability
- Network-based druggability

Druggability scores can be found by clicking the 'Druggability' link on the Molecular Target Synopsis side menu.

Structural Druggability

The first tab to be displayed is for structural druggability. This method for assessing druggability examines the structure of the protein and identifies any cavities on the protein surface where a drug-like compound could bind.



Each PDB chain in the protein is assessed as druggable/non-druggable

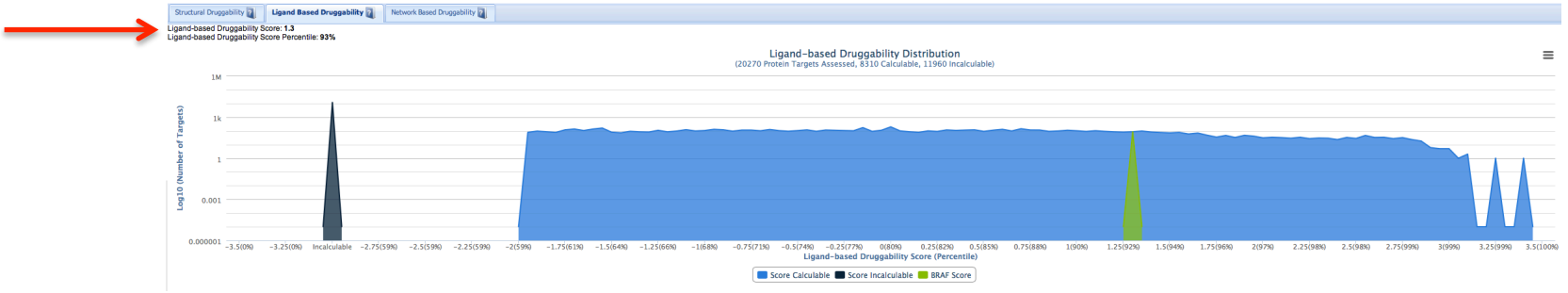
This analysis is based on the ChEMBL Strudel (https://www.ebi.ac.uk/chembl/drugebility/) (DrugEBIlity) methodology. All structures are processed through the pipeline and up to 10 possible cavities are identified in every structure, regardless of the presence or absence of a ligand.

The geometric and physicochemical properties of each cavity, such as volume, buried surface area, number of hydrogen-bond donors and acceptors are fed into an ensemble of decision trees and data mining algorithms that predict the likely druggability of the cavity based on these properties. More details can be found on the DrugEBIlity website and Patel et al, Nat Rev Drug Discov. 2013 Jan; 12(1):35-50.

We apply the same algorithm weekly on all PDB-chains and present the results in the chain-based table. In addition, we provide the domain-based assessment as obtained from the DrugEBIlity resource and is calculated for independent structural domains, as defined by SCOP (http://scop.mrc-lmb.cam.ac.uk/scop/) and Asteral Asteroids (http://astral.berkeley.edu)

Ligand-based druggability

The second tab displays the ligand-based druggability score of the target. Ligand-based druggability looks at the properties of compounds that have been tested against the target. If the target binds drug-like compounds, it is more likely to be druggable than a target that only binds compounds with very un-drug like properties.
The percentile for the druggability score is an easy way to see how the target's druggability compares to all other targets in the proteome.



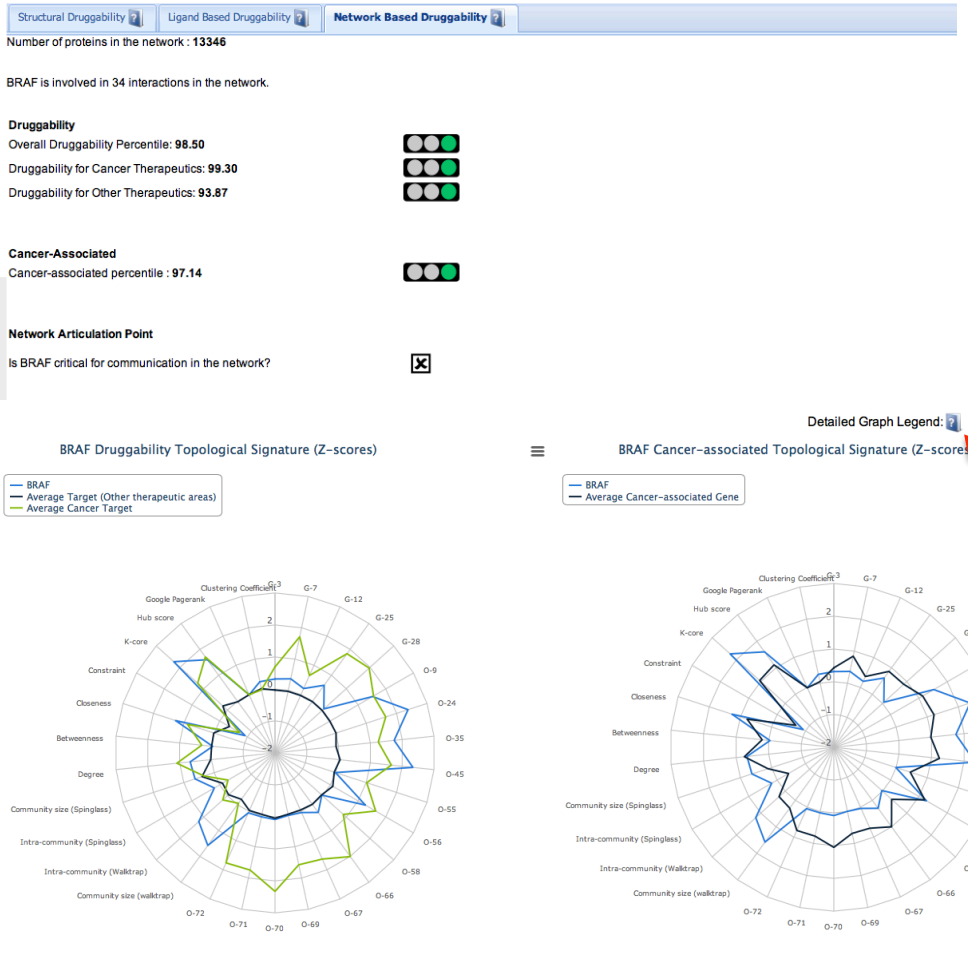
This methodology aims to estimate the likely druggability of a target based on the chemical properties and bioactivity parameters of small molecule compounds that have been tested against the protein itself and/or its homologues.

In summary, compounds are selected from canSAR if they have been tested against the target in question and its homologues. The compounds are weighted based on their bioactivity levels and the homology of their target to the initial protein query. The properties of the compounds (including molecular weight, med-chem friendliness and ligand-efficiency) are used to score each protein and rank it against the human proteome. Over half the proteins in the proteome cannot be assessed as neither they nor their homologues have known chemical matter.

For >8000 proteins that can be calculated, we display the scores distribution and show where the query protein falls within the distribution as a comparator.

Network-based druggability

The third tab shows the druggability score for the target using a network-based approach. This examines the structure of the protein-protein interaction around the target to decide how good a drug target the target is. For example, a protein involved in many hundreds of signalling and protein-protein interaction events will be a good drug target as disrupting its activity will affect many processes within the cell. A protein with few interactions will make a poor target as this would affect very few processes in the cell.



More detailed descriptions of these radar plots are displayed here