

## FAQ

### StarDrop: ADME QSAR Models

StarDrop's optional ADME QSAR module incorporates a comprehensive suite of high quality quantitative structure activity relationship (QSAR) models for ADME properties. These models exhibit a combination of high accuracy, computational efficiency and ease of interpretation through the use of chemically meaningful descriptors.

The models currently included in the ADME QSAR module comprise:

- logP (Octanol/Water)
- logD<sub>7.4</sub> (Octanol/buffer at pH 7.4)
- Aqueous Solubility
- Human Intestinal Absorption
- CNS (Blood-Brain Barrier) Penetration
- Cytochrome P450 Affinities
  - CYP2C9
  - CYP2D6
- P-gp Transport
- hERG pIC<sub>50</sub>
- Plasma Protein Binding

Features of the ADME QSAR models include:

- All models require only the chemical structure as input. Compound properties can be predicted to identify optimal chemistries before committing to expensive synthesis and testing.
- Every model is rigorously validated using an independent test set to ensure high predictive capability.
- Each prediction is accompanied by an estimate of the confidence in that prediction, enabling uncertainties to be explicitly considered.
- Every model is accompanied by a measure of its applicable chemical space. Predictions for molecules outside of the chemical space, where confidence is lower, are explicitly identified.

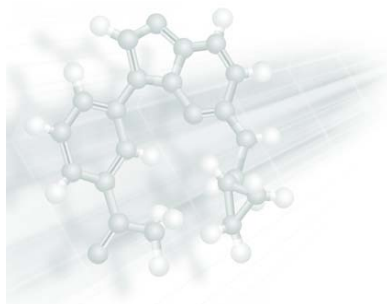
## Frequently Asked Questions

### Where do the data used to build the QSAR models come from?

The QSAR models are based on either in-house measurements or literature data. Larger data sets are often available from the published literature, particularly for properties requiring *in vivo* measurements. Regardless of the source of data, **strict quality control** by both experimental and computational scientists is always applied to any data sets used to develop StarDrop models.

### What kind of molecular descriptors are used?

The QSAR models are based on **chemically meaningful descriptors** to provide guidance regarding the effects of chemical modifications on the predicted property. 1D and 2D descriptors such as molecular weight, volume, lipophilicity terms, count of atoms and count of fragments are used in all models. 3D descriptors are not used in the current model suite. Due to the additional computational cost they would incur, they would only be included when they offered a significant improvement in the accuracy of the model.



## FAQ

### StarDrop: ADME QSAR Models

#### How do you validate your models?

The QSAR models are validated with **independent test sets**, i.e. sets of molecules not used to train the models. Such validation results are **statistically rigorous tests of the predictive ability of the model**, as opposed to cross correlation or training set results, which are often reported for other models. These latter values measure the ability of the model to fit the data in the training set, which is typically better than the true predictive ability of the model.

#### How do StarDrop models compare to other models?

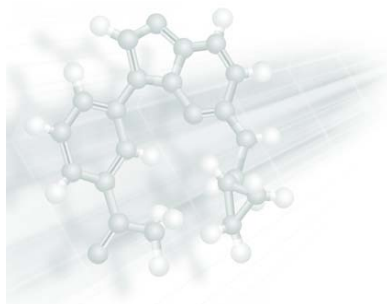
In independent tests, the StarDrop models have matched or exceeded the performance of other commercially available or published models.

The results for the StarDrop models are shown in the tables below. If comparing the StarDrop models with other *in silico* models, it is important to note that the reported validation results here are for **independent test sets** (see "How do you validate your models?" above). It is also important, when comparing StarDrop predictions with experimental data, to ensure that this is a **like-for-like comparison**. For example, HIA predictions cannot be compared directly with oral bioavailability data on which both absorption and first-pass clearance will have an impact. Similarly, the StarDrop predictions for aqueous solubility are unlikely to correlate well with data obtained in a high-throughput kinetic assay based on dilution of DMSO compound stocks and will not predict the solubilities of different salt forms.

#### Summary of Statistical Results for Continuous QSAR models

Model	Definition	N <sub>Test</sub> <sup>a</sup>	R <sup>2</sup> <sup>b</sup>	RMSE <sub>IN</sub> <sup>c</sup>
logP	Predicts the logarithm of the octanol/water partition coefficient for neutral compounds.	2950	0.92	0.44
logD <sub>7.4</sub>	Predicts the logarithm of the octanol/buffer at pH 7.4 distribution coefficient.	427	0.81	0.68
logs	Predicts the logarithm of the intrinsic aqueous solubility, S in $\mu$ M, for neutral compounds.	663	0.82	0.70
log([brain]:[blood])	Predicts the logarithm of the Brain/Blood ratio.	43	0.73	0.36
hERG pIC <sub>50</sub>	Predicts the pIC <sub>50</sub> values for inhibition of hERG K <sup>+</sup> channels expressed in mammalian cells.	29	0.68	0.70
2C9 pKi	Predicts the pKi values for CYP2C9 affinity.	25	0.62	0.63

<sup>a)</sup> N = number of compounds in **independent test set**. <sup>b)</sup> R<sup>2</sup> gives the correlation between calculated and experimental values for the compounds in the independent test set. <sup>c)</sup> The root mean squared error (RMSE) statistic gives the error for the corresponding correlation coefficient. When possible, the RMSE values are calculated for compounds within (IN) or outside (OUT) the chemical space of the model. "unknown" is reported if there are not enough training and test compounds outside the chemical space to calculate an RMSE value.



## FAQ

### StarDrop: ADME QSAR Models

#### Summary of Statistical Results for Classification QSAR models

Model	Definition	N <sub>Test</sub>	Accuracy	Specificity		
HIA category	Returns a binary prediction for human intestinal absorption, based on a threshold of 30% absorbed.	245	\-\'	66%	\-\'	91%
			\+\'	99%	\+\'	95%
BBB category	Returns a binary prediction for Blood/Brain barrier penetration.	52	\-\'	93%	\-\'	91%
			\+\'	83%	\+\'	83%
P-gp category	Returns a binary prediction for P-gp transport.	51	\yes\'	86%	\yes\'	78%
			\no\'	68%	\no\'	79%
PPB category	Returns a binary prediction for human plasma protein binding, based on a threshold of 80% absorbed	159	\-\'	78%	\-\'	78%
			\+\'	77%	\+\'	77%
2D6 affinity category	Returns a 4-class prediction for 2D6 affinity.	45	Root mean square error = 0.87 classes			

<sup>a)</sup> N = number of compounds in independent test set. <sup>b)</sup> The accuracy for each class is reported as the percentage of compounds correctly classified. <sup>c)</sup> The specificity refers to the percentage of correct classifications within the overall set of compounds predicted to be in that class.

#### Global vs local models?

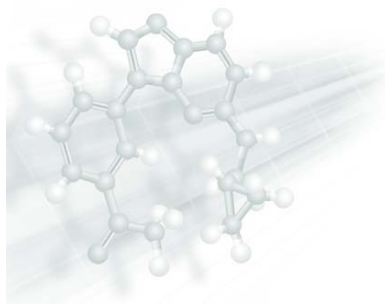
The QSAR models are 'global' models of ADME properties, i.e. compounds used to build the model (training compounds) cover as wide a range of chemical diversity as possible. In contrast, local models are specific to one region of the chemical space.

As a drug discovery project progresses, the chemistry under consideration often focuses on a small number of chemical series in which the molecules are structurally similar. Global models may lack the resolution required to distinguish between molecules with subtle differences and once *in vitro* data have been generated for a chemical series, the ability of the corresponding models to discriminate within the series should be tested. If a model is found to lack discrimination, it may be possible to develop a local model to improve resolution within this limited range of chemistry.

Local models may be integrated within StarDrop (see "Can we use our own models in StarDrop?" below) and the Auto-Modeler module provides clients with the ability to generate models using their own data.

#### Are your models predictive of our chemistry?

Each model result is provided with an estimate of the uncertainty in the prediction, in the form of a root mean square error (RMSE) for continuous data or a probability for classification data. The uncertainty is based on how similar the structure is to those molecules used when building the model. If your chemistry differs significantly from those used to build the models, it is less likely that the models will be predictive and this can be seen explicitly from the uncertainties returned (see "How do you estimate uncertainties?" below). As noted above, the models supplied with StarDrop are 'global' and should readily distinguish 'long range' trends across a wide range of chemistries, but they are less likely to be able to differentiate between similar molecules in the same series.



## FAQ

### StarDrop: ADME QSAR Models

#### How do you define the chemical space of a model?

The '**chemical space**' of a model represents the range of model descriptors that are well represented by the molecules in the training set for that model. This is defined using a Hotelling's  $T^2$  statistical test.

#### How do you estimate uncertainties?

Within the chemical space of the model, there is a statistically significant number of data points in the independent test set, and hence the performance of the model on this set provides a good estimate of the uncertainty. For compounds outside but 'close' to the chemical space of the model, extrapolation beyond the training set increases the uncertainties in these predictions. Where possible, this is estimated using the performance of the model on compounds in the independent test set that lie outside of the chemical space. For compounds that differ significantly from the training set, no valid estimate of the uncertainty can be made. The uncertainty in the predictions for these compounds is reported as maximal i.e. for continuous models an infinite uncertainty and for classification models an equal probability for each class.

#### Can we use our own models within StarDrop?

Yes. We provide example code and a technical document explaining how to set up a simple server which will enable all users of StarDrop within your organisation to run your own models. We recommend this is done by a software developer or someone with a strong computational background. Once a server is set up, users can easily change the settings in the StarDrop preferences to tell it to use this server to access and run your own models.

#### Can we use other third party models?

Yes, if they can be run from a command-line (rather than within a desktop application) and can therefore be accessed from the code we provide to set up and run your own additional model server (See "Can we use our own models within StarDrop?" above). A straightforward alternative is to import the prediction from your own models as a .csv or .sd file.

#### What information would third party models need to provide to employ the scoring capability of StarDrop?

The models would need to provide a measure of uncertainty and the user would need an understanding of the output in order to choose appropriate scoring thresholds.

#### Is it possible to integrate proprietary data into StarDrop models?

We offer to integrate proprietary data for additional compounds on a consultancy basis. When we do this, the new data will be checked to ensure the quality and compatibility with the data points on which the original model was built. Alternatively, our computational chemists can build a custom model on a client's data set. If the new data are derived from a limited range of chemical diversity, the resulting model will give optimal predictions for compounds of similar chemistry and should also provide greater resolution than a global model. However, the converse may also apply and the resulting model is likely to have reduced accuracy in other areas of chemistry.

The optional Auto-Modeler module provides clients with the ability to generate models using their own data.

#### Can you run ADME QSAR models without a server?

Yes, this module can run on a desktop/laptop machine without access to a server.