

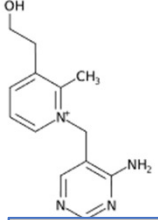
# Data Modeling

## Create a model

Retrieve a list of "Good" molecules from a:

- Search
- Saved search
- Collection

350 Selected: [Launch Vision](#) [Plot](#) [Export](#) [Add to collection](#) [Build model](#)

Select...	Molecule	Chemical Properties
all none		Molecular weight (g/mol) pKa pKa typ
<input checked="" type="checkbox"/>	CDD-2301082 	259.332 5.54 Basic

Build a Machine Learning Model

Training set baseline:  **Select protocol or collection for training set baseline**

"Good" molecules: Your 431 selected results

Name:  **Enter model name**

Project:  **Select project**

or

## Model are treated as Protocols

Protocol	Description	Category	Molecules
<a href="#">Good molecules</a>		Machine Learning Model	<a href="#">Explore</a>
<a href="#">Inhibitors data</a>			<a href="#">Explore</a>

**View**

- Run data
- ROC curve

**Category for models**

**View molecules in project with their scores**

## Readout definitions calculated

Name	Description
Score	Relative score (higher is better)
Applicability	Fraction of structural features shared with the training set
Maximum similarity	Maximum Tanimoto/Jaccard similarity to any "good" molecules

## Algorithm

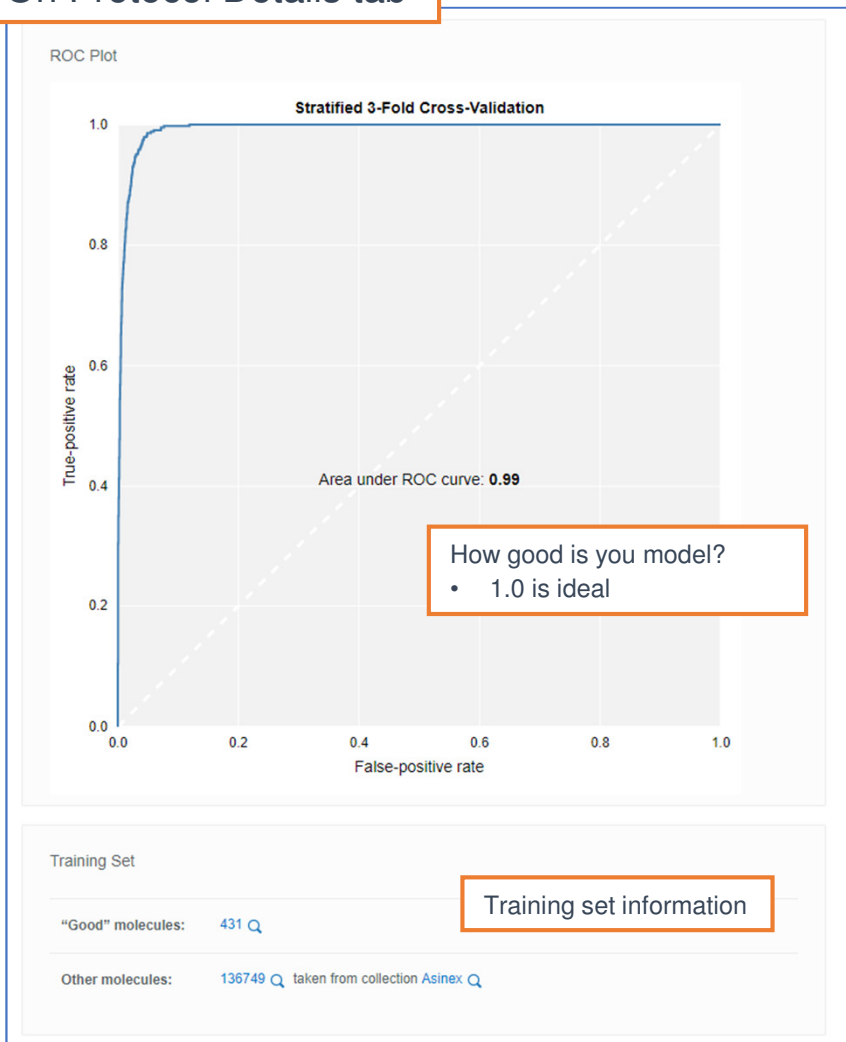
**Modified Bayesian model with FCFP6 descriptors**

<b>Model:</b>	CDD's open source <a href="#">modified Bayesian model</a> described in <a href="#">Xia et al. 2004</a> .
<b>Descriptors:</b>	CDD's open source <a href="#">FCFP6 fingerprints</a> described in <a href="#">Rogers and Hahn 2010</a> .

# Data Modeling

## Quality of model

On Protocol Details tab



## Use a model

Run Data Protocol Details **Projects 2** Files 0

Projects with access to Good molecules: **Molecules scored automatically**

Bayesian Remove

Internal data Remove

select project [v](#)

**Save changes** or **cancel**

Add model to the desired project  
*Note: All molecules in project are scored*

Recalculating statistics and calculations for the following protocols:

- Good molecules

Banner is displayed while molecules are being scored