GUSAR

The acronym GUSAR stands for "General Unrestricted Structure-Activity Relationships". The system has been developed by the team of Professor Vladimir Poroikov in Moscow, Russia.

The GUSAR software is designed to create reliable quantitative QSAR/QSPR models using a training set of 2D chemical structures and quantitative data on biological activities. These models can then be used for activity prediction.

GUSAR uses unique algorithms like self-consistent regression, and supports the use of nearest neighbors, consensus models and applicability domains.

PASS

The acronym PASS stands for „Prediction of Activity Spectra for Substances". If you want to predict many biological activities qualitatively at the same time, this is the software for you.

GUSAR allows to:

- Create QSAR/QSPR models;
- Predict biological activities based on these models;
- Calculate consensus models for prediction;
- Validate internal and external models;
- Interpret the results easily.

Recent GUSAR publications

Kokurkina G.V., Dutov M.D., Shevelev S.A., Popkov S.V., Zakharov A.V., Poroikov V.V. (2011) Synthesis, antifungal activity and QSAR study of 2-arylhydroxynitroindoles. Eur J Med Chem July 19, Epub ahead of print.

Lagunin A., Zakharov A., Filimonov D., Poroikov V. (2011) QSAR modelling of rat acute toxicity on the basis of PASS prediction. Mol Inform 30:241–250.

Filimonov D.A., Zakharov A.V., Lagunin A.A., Poroikov V.V. (2009) QNA based 'Star Track' QSAR approach. SAR QSAR Environ. Res. 20:679-709.

Lagunin A., Zakharov A., Filimonov D., Poroikov V. (2009). In silico assessment of acute toxicity in rodents. Toxicol. Lett. 189:S264.

More publications can be found on our web page.

GUSAR

About geneXplain

The geneXplain GmbH is a young company with a growing portfolio of useful software for life scientists. Whether you are working on genome, network or compound analysis, we have what you need!

Our **geneXplain platform** is the online toolbox and workflow management system for scientists in the fields of transcriptomics and proteomics. Here, you can store and analyze your experimental data (including raw microarray results), search for master regulator molecules, map to GO terms, and even add your own workflows and scripts.

We distribute software developed by in silico molecular biology, Inc., Japan, for geneticists. **In-Silico Molecular Cloning (IMC)** lets you handle annotated DNA, conduct cloning experiments in silico, map features and sequences, and compare and align genomes. **GenomeTraveler (GT)** takes this to the next level by adding the possibility to visualize and interpret your Next Generation Sequencing (NGS) data.

For scientists working on drug discovery and drug optimization, we distribute a range of tools from the Institute of Biomedical Chemistry (IBMC), Russia. **PASS (Prediction of Activity Spectra for Substances)** predicts whole bioactivity spectra for your compounds qualitatively, based only on 2D structural formulae. **PharmaExpert** is the additional software to help you choose the most suitable substance from a set of PASS predictions. **GUSAR (General Unrestricted Structure-Activity Relationships)** predicts biological activities quantitatively, based on (Q)SAR models, which can also be created with the software.

It is geneXplain's mission to provide the computational methodology required to achieve the goal of "personalized pharmacogenomics". We wish to help academic researchers in their daily work with easy-to-use tools that are compatible with the low-budget requirements of most academic groups. At the same time, we shall provide high-end technology platforms to fulfill bioinformatics requirements to industrial standards. Finally, we intend to offer partnerships for research and training in the area of our expertise.

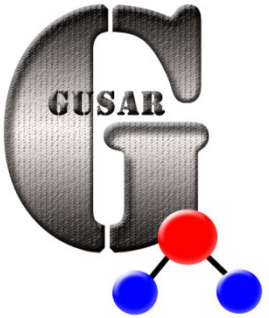
geneXplain GmbH
Am Exer 10b
D-38302 Wolfenbüttel, Germany

info@genexplain.com
www.genexplain.com
www.facebook.com/genexplain
www.linkedin.com/company/genexplain

Directors: E. Wingender, A. Kel • Commercial register:
HRB 202564, Amtsger. Braunschweig • VAT No.: DE271983408

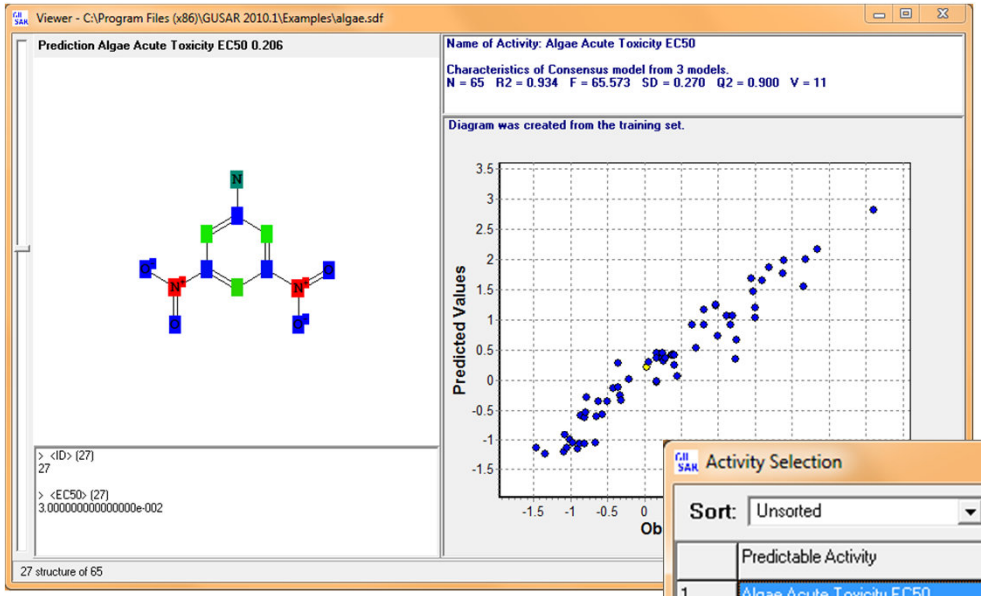
GUSAR

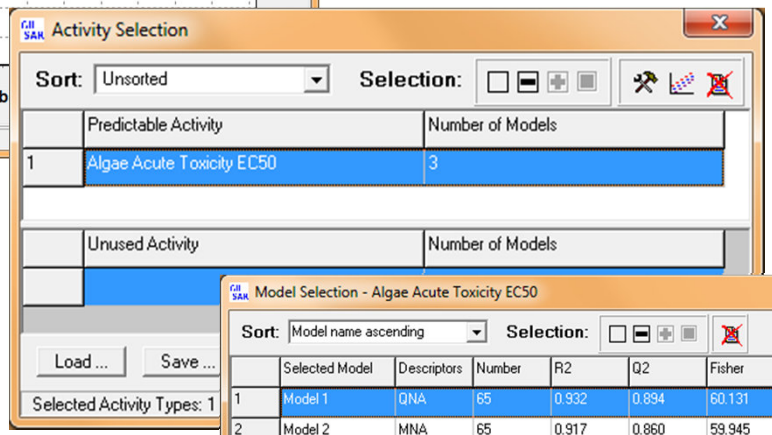
The tool to create models on quantitative structure- activity relationships

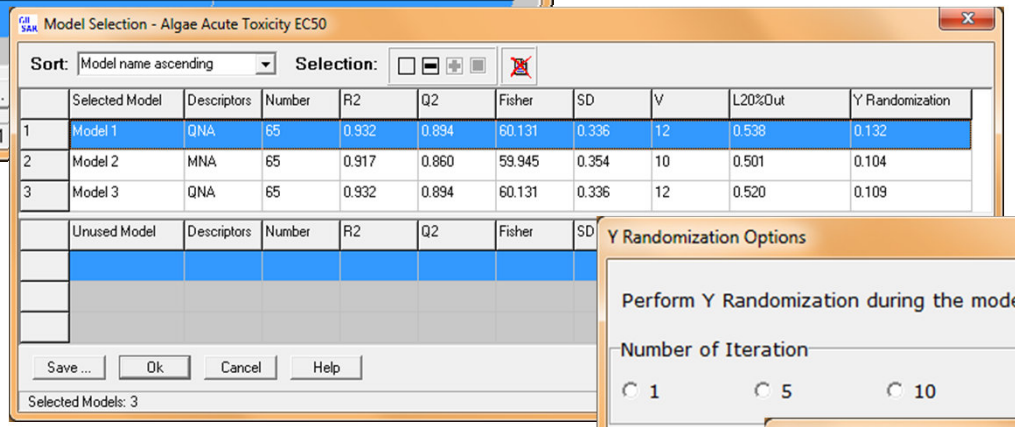


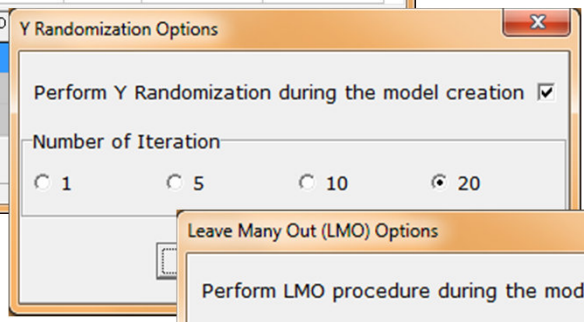
geneXplain

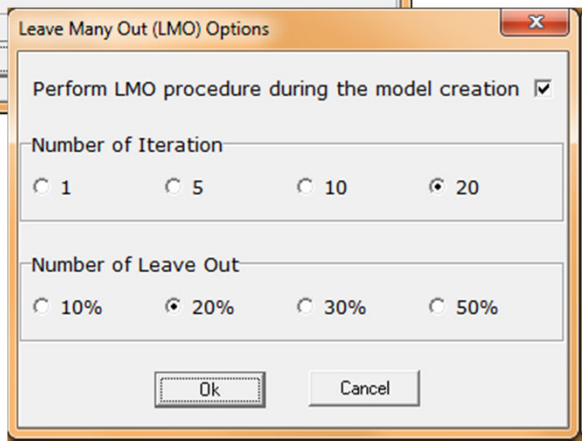
GUSAR – a software for the creation of quantitative SAR/SPR models











Chemical structure description in GUSAR

2D structures are represented by MNA (Multilevel Neighborhoods of Atoms) and/or QNA (Quantitative Neighborhoods of Atoms) descriptors and biological activity descriptors that are based on the **PASS** prediction results for more than 4,000 biological activities. QNA descriptors easily reflect the nature of intermolecular interactions. Models developed using biological activity descriptors enable to reveal key mechanisms of action of complex biological effects.

Key benefits

- Fast creation of high quality QSAR models
- Optimized lead identification through *in silico* screening with GUSAR models
- Creation of QSAR models for different biological activities, suitable for their simultaneous prediction
- Revealing key mechanisms of action when modeling complex biological effects like acute toxicity, carcinogenicity, hepatotoxicity etc.

Key features

- Unique descriptors and mathematical algorithms,
- Creation of QSAR models for large data sets, up to 30,000 chemical compounds,
- High speed of predictions,
- Easy-to-use interface,
- Selection of the most predictive models,
- Uploading of SD files for batch predictions,
- Saving GUSAR output predictions in SDF and CSV formats for subsequent analyses.

Mathematical algorithm

The unique algorithm of a self-consistent regression allows to select the best set of descriptors for a robust and reliable QSAR model. It is based on the statistical regularization of ill-posed problems, and uses the same data samples to estimate both the regression coefficients and the regularization parameters.

Ready-Trained Models

Ready-trained QSAR models for GUSAR are available on request.

The model on acute rat toxicity contains four activities (LD_{50} for different administration options) and is build from training sets with between 759 and 6280 structures.

The model on the affinity of substances to antitargets contains 32 activities (IC_{50}, K_i and K_{act}), with training sets including between 60 and 1366 structures.

Validation

In comparison with a number of 3D and 2D QSAR methods, the predictivity of GUSAR was superior to that of most other QSAR methods both on heterogeneous and on homogeneous data sets, as published in

Filimonov D.A., Zakharov A.V., Lagunin A.A., Poroikov V.V. (2009). QNA based 'Star Track' QSAR approach. SAR QSAR Environ. Res. 20: 679-709.

A slide show with screenshots can be found on our homepage (www.genexplain.com) and our Facebook account (www.facebook.com/genexplain).