Selected references on GUSAR

Gorla U.S., Rao K., Kulandaivelu U.S., Alavala R.R., Panda S.P. Lead Finding from Selected Flavonoids with Antiviral (SARS-CoV-2) Potentials Against COVID-19: An In-silico Evaluation. *Comb. Chem. High Throughput Screen.*, **2021**, 24(6): 879-890.

Background: COVID-19 is a pandemic respiratory contagious viral (SARS-CoV-2) disease associated with high morbidity and mortality worldwide. Currently, there are no effective preventive or treatment strategies for COVID-19 and it has been declared as a global health emergency by WHO. In silico molecular docking studies can be useful to predict the binding affinity between the phytocompound and the target protein and play a vital role in finding an inhibitor through structure-based drug design.

Objective: In this aspect, our objective was to screen essential flavonoids against possible protein targets such as SARS-CoV-2 spike glycoprotein receptor binding domain (RBD-S) and host Angiotensin Converting Enzyme-2 protease domain (PD-ACE-2) using in silico molecular docking studies.

Methods: Approximately 49 flavonoids were identified and were evaluated for their drug-likeness based on Lipinski rule, bioactivity scores, antiviral and toxicity profiles using SwissADME, Molinspiration, PASS and GUSAR online tools. The flavonoids that passed Lipinski rule were subjected to in silico analysis through molecular docking on RBD-S and PD-ACE-2 using Molegro Virtual Docker v6.0.

Results: The bioactive flavonoids that showed NIL violations and were found in compliance with Lipinski rule were selected for docking studies. In silico analysis reported that biochanin A and silymarin bind significantly at the active sites of RBD-S and PD-ACE-2 with a MolDock score of -78.41and -121.28 kcal/mol respectively. Bioactivity scores, antiviral potential and toxicity profiles were predicted for the top interacting phytocompounds and substantial relevant data was reported.

Conclusion: The current outcomes created a new paradigm for understanding biochanin A and silymarin bioflavonoids as potent inhibitors of RBD-S and PD-ACE-2 targets respectively. Further work can be extended to confirm their therapeutic potential for COVID-19.

Korotkevich E.I., Rudik A.V., Dmitriev A.V., et al. Prediction of metabolic stability of xenobiotics by the PASS and GUSAR programs. *Biomed. Khim.*, **2021**, 67(3): 295-299.

Metabolic stability refers to the susceptibility of compounds to the biotransformation; it is characterized by such pharmacokinetic parameters as half-life (T1/2) and clearance (CL). Generally, these parameters are estimated by in vitro assays, which are based on cells or subcellular fractions (mainly liver microsomal enzymes) and serve as models of the processes occurring in living organisms. Data obtained from the experiments are used to build QSAR (Quantitative Structure-Activity Relationship) models. More than 8000 compounds with known CL and/or T1/2 values obtained in vitro using human liver microsomes were selected from the freely available ChEMBL v.27 database. GUSAR (General Unrestricted Structure-Activity Relationships) and PASS (Prediction of Activity Spectra for Substances) softwares were used to make quantitative and classification models. The quality of the models was evaluated using 5-fold cross-validation. Compounds were subdivided into "stable" and "unstable" by means of the following threshold parameters: T1/2 = 30 minutes, CL = 20 ml/min/kg. The accuracy of the models ranged from 0.5 (calculated in 5-fold CV on the test set for the half-life prediction quantitative model) to 0.96 (calculated in 5-fold CV on the test set for the clearance prediction classification model).

Khairullina V., Safarova I., Sharipova G., et al. QSAR Assessing the Efficiency of Antioxidants in the Termination of Radical-Chain Oxidation Processes of Organic Compounds. *Molecules*, **2021**, 26(2): 421.

Using the GUSAR program, the quantitative structure-antioxidant activity relationship has been studied for 74 phenols, aminophenols, aromatic amines and uracils having lgk7 = 0.01-6.65 (where k7 is the rate constant for the reaction of antioxidants with peroxyl radicals generated upon oxidation). Based on the atomic descriptors (Quantitative Neighborhood of Atoms (QNA) and Multilevel Neighborhoods of Atoms (MNA)) and molecular (topological length, topological volume and lipophilicity) descriptors, we have developed 9 statistically significant QSAR consensus models that demonstrate high accuracy in predicting the lgk7 values for the compounds of training sets and appropriately predict lgk7 for the test samples. Moderate predictive power of these models is demonstrated using metrics of two categories: (1) based on the determination coefficients R2 (R2TSi, R20, Q2(F1), Q2(F2), RmTSi2¯) and based on the concordance correlation coefficient (CCC)); or (2) based on the prediction lgk7 errors (root mean square error (RMSEP), mean absolute error (MAE) and standard deviation (S.D.)) The RBF-SCR method has been used for selecting the descriptors. Our theoretical prognosis of the lgk7 for 8-PPDA, a known antioxidant, based on the consensus models well agrees with the experimental value measure in the present work. Thus, the algorithms for calculating the descriptors implemented in the GUSAR program allow simulating kinetic parameters of the reactions underling the liquid-phase oxidation of hydrocarbons.

Ahmed S., Moni D.A., Sonawane K.D. et al. A comprehensive *in silico* exploration of pharmacological properties, bioactivities and COX-2 inhibitory potential of eleutheroside B from *Eleutherococcus senticosus* (Rupr. & Maxim.) Maxim. *J. Biomol. Struct. Dyn.*, **2021**, 39(17): 6553-6566.

Eleutherococcus senticosus (Rupr. & Maxim.) Maxim., popularly known as 'Siberian ginseng', is an important medicinal plant. Pharmacologically active compounds of this plant are called eleutherosides and among them, eleutheroside B is the most prevalent. The E. senticosus has been reported to have many medicinal properties however; very few studies are reported to understand the medicinal properties of eleutheroside B. Consequently, in the present study various computational tools have been used to predict the drug-likeness, bioactivities, and pharmacokinetic properties of eleutheroside B. Besides, the inhibitory potential of eleutheroside B has been investigated against cyclooxygenase 2 (COX-2) enzyme. This study suggests that eleutheroside B is a drug-like compound with bioactivity score (-0.08 to 0.38), having satisfactory pharmacokinetic values. Metabolism and toxicities were further studied using FAME3, GLORY, pred-hERG and Endocrine Disruptome tools. No severe toxicities (Ames, hepatotoxicity, cardiotoxicity, skin sensitization) were predicted. Rat acute toxicity, ecotoxicity and cell line cytotoxicity were evaluated based on GUSAR and CLC-pred. The compound has been predicted as non-toxic (class 5), non-hERG inhibitor and less likely to cause adverse drug interactions. Molecular docking against COX-2 enzyme revealed strong hydrogen bonds (SER530, TYR355, LEU352, SER353, VAL349, TYR385, MET522) and hydrophobic interaction (LEU352) with eleutheroside B. The docking score (-6.97 kcal/mol) suggested that this molecule can be utilized as an anti-inflammatory agent as well as a potential anticancer drug in the future. Hence, this is a comprehensive integrated in silico approach to establish the anti-inflammatory mechanism of eleutheroside B in the background of its potential in future drug development.

Mansouri K., Kleinstreuer N., Abdelaziz A.M., et al. CoMPARA: Collaborative Modeling Project for Androgen Receptor Activity, *Environmental Health Perspectives*, **2020**, 128 (2), 027002-1 - 027002-17.

Background: Endocrine disrupting chemicals (EDCs) are xenobiotics that mimic the interaction of natural hormones and alter synthesis, transport, or metabolic pathways. The prospect of EDCs causing adverse health effects in humans and wildlife has led to the development of scientific and regulatory approaches for evaluating bioactivity. This need is being addressed using high-throughput screening (HTS) in vitro approaches and computational modeling.

Objectives: In support of the Endocrine Disruptor Screening Program, the U.S. Environmental Protection Agency (EPA) led two worldwide consortiums to virtually screen chemicals for their potential estrogenic and androgenic activities. Here, we describe the Collaborative Modeling Project for Androgen Receptor Activity (CoMPARA) efforts, which follows the steps of the Collaborative Estrogen Receptor Activity Prediction Project (CERAPP).

Methods: The CoMPARA list of screened chemicals built on CERAPP's list of 32,464 chemicals to include additional chemicals of interest, as well as simulated ToxCast™ metabolites, totaling 55,450 chemical structures. Computational toxicology scientists from 25 international groups contributed 91 predictive models for binding, agonist, and antagonist activity predictions. Models were underpinned by a common training set of 1,746 chemicals compiled from a combined data set of 11 ToxCast™/Tox21 HTS in vitro assays.

Results: The resulting models were evaluated using curated literature data extracted from different sources. To overcome the limitations of single-model approaches, CoMPARA predictions were combined into consensus models that provided averaged predictive accuracy of approximately 80% for the evaluation set.

Discussion: The strengths and limitations of the consensus predictions were discussed with example chemicals; then, the models were implemented into the free and open-source OPERA application to enable screening of new chemicals with a defined applicability domain and accuracy assessment. This implementation was used to screen the entire EPA DSSTox database of ~875,000 chemicals, and their predicted AR activities have been made available on the EPA CompTox Chemicals dashboard and National Toxicology Program's Integrated Chemical Environment.

Amiranashvili L., Nadaraia N., Merlani M., et al. Antimicrobial Activity of Nitrogen-Containing 5-Alpha-androstane Derivatives: In Silico and Experimental Studies. *Antibiotics (Basel)*, **2020**, 9(5): 224.

We evaluated the antimicrobial activity of thirty-one nitrogen-containing 5-alpha-androstane derivatives in silico using computer program PASS (Prediction of Activity Spectra for Substances) and freely available PASS-based web applications (www.way2drug.com). Antibacterial activity was predicted for 27 out of 31 molecules; antifungal activity was predicted for 25 out of 31 compounds. The results of experiments, which we conducted to study the antimicrobial activity, are in agreement with the predictions. All compounds were found to be active with MIC (Minimum Inhibitory Concentration) and MBC (Minimum Bactericidal Concentration) values in the range of 0.0005-0.6 mg/mL. The activity of all studied 5-alpha-androstane derivatives exceeded or was equal to those of Streptomycin and, except for the 3β-hydroxy-17α-aza-d-homo-5α-androstane-17-one, all molecules were more active than Ampicillin. Activity against the resistant strains of *E. coli*, *P. aeruginosa*, and methicillin-resistant *Staphylococcus aureus* was also shown in experiments. Antifungal activity was determined with MIC and MFC (Minimum Fungicidal Concentration) values varying from 0.007 to 0.6 mg/mL. Most of the compounds were found to be more potent than the reference drugs Bifonazole and Ketoconazole. According to the results of docking studies, the putative targets for antibacterial and antifungal activity are UDP-N-acetylenolpyruvoylglucosamine reductase and 14-alpha demethylase, respectively. In silico assessments of the acute rodent toxicity and cytotoxicity obtained using GUSAR (General Unrestricted Structure-Activity Relationships) and CLC-Pred (Cell Line Cytotoxicity Predictor) web-services were low for the majority of compounds under study, which contributes to the chances for those compounds to advance in the development.

Stolbov L.A., Druzhilovskiy D.S., Filimonov D.A., et al. (Q)SAR Models of HIV-1 Protein Inhibition by Drug-Like Compounds. *Molecules*, **2019**, 25(1): 87.

Despite the achievements of antiretroviral therapy, discovery of new anti-HIV medicines remains an essential task because the existing drugs do not provide a complete cure for the infected patients, exhibit severe adverse effects, and lead to the appearance of resistant strains. To predict the interaction of drug-like compounds with multiple targets for HIV treatment, ligand-based drug design approach is widely applied. In this study, we evaluated the possibilities and limitations of (Q)SAR analysis aimed at the discovery of novel antiretroviral agents inhibiting the vital HIV enzymes. Local (Q)SAR models are based on the analysis of structure-activity relationships for molecules from the same chemical class, which significantly restrict their applicability domain. In contrast, global (Q)SAR models exploit data from heterogeneous sets of drug-like compounds, which allows their application to databases containing diverse structures. We compared the information for HIV-1 integrase, protease and reverse transcriptase inhibitors available in the EBI ChEMBL, NIAID HIV/OI/TB Therapeutics, and Clarivate Analytics Integrity databases as the sources for (Q)SAR training sets. Using the PASS and GUSAR software, we developed and validated a variety of (Q)SAR models, which can be further used for virtual screening of new antiretrovirals in the SAVI library. The developed models are implemented in the freely available web resource AntiHIV-Pred.

Stolbov L., Druzhilovskiy D., Rudik A., et al. AntiHIV-Pred: web-resource for in silico prediction of anti-HIV/AIDS activity. *Bioinformatics*, **2020**, 36(3): 978-979.

Motivation: Identification of new molecules promising for treatment of HIV-infection and HIV-associated disorders remains an important task in order to provide safer and more effective therapies. Utilization of prior knowledge by application of computer-aided drug discovery approaches reduces time and financial expenses and increases the chances of positive results in anti-HIV R&D. To provide the scientific community with a tool that allows estimating of potential agents for treatment of HIV-infection and its comorbidities, we have created a freely-available web-resource for prediction of relevant biological activities based on the structural formulae of drug-like molecules.

Results: Over 50 000 experimental records for anti-retroviral agents from ChEMBL database were extracted for creating the training sets. After careful examination, about seven thousand molecules inhibiting five HIV-1 proteins were used to develop regression and classification models with the GUSAR software. The average values of R2 = 0.95 and Q2 = 0.72 in validation procedure demonstrated the reasonable accuracy and predictivity of the obtained (Q)SAR models. Prediction of 81 biological activities associated with the treatment of HIV-associated comorbidities with 92% mean accuracy was realized using the PASS program.

Martynova Y.Z., Khairullina V.R., Gimadieva A.R., Mustafin A.G. QSAR-modeling of desoxyuridine triphosphatase inhibitors in a series of some derivatives of uracil. *Biomed. Khim.*, **2019**, 65(2):103-113.

Due to the widespread prevalence, deoxyuridine triphosphatase (UTPase) is considered by modern biochemists and physicians as a promising target for the development of drugs with a wide range of activities. The therapeutic effect of these drugs will be due to suppression of DNA biosynthesis in various viruses, bacteria and protozoa. In order to rationalize the search for new dUTPase inhibitors, domestic and foreign researchers are actively using the QSAR methodology at the selection stage of hit compounds. However, the practical application of this methodology is impossible without existence of valid QSAR models. With the use of the GUSAR program, a quantitative analysis of the relationship between the structure and efficacy of 135 dUTPase inhibitors based on uracil derivatives was performed in the IC50 range of 30¸185000 nmol/L. Six statistically significant valid consensus models, characterized by high descriptive ability and moderate prognostic ability on the structures of training and test samples, are constructed. To build valid QSAR models for dUTPase inhibitors can use QNA or MNA descriptors and their combinations in a consensus approach.

Butina Y.V., Kudayarova T.V., Danilova E.A., Islyaikin M.K. The prediction of the spectrum of biological activity and antimicrobial properties of diaminoazoles. *Biomed. Khim.*, **2019**, 65(2): 99-102.

The study is devoted to predicting and studying biological properties of N-substituted analogs of 3,5-diamino-1,2,4-thiadiazole, which, in their turn, include in the composition of many drugs that exhibit a wide range of pharmacological actions. For searching of new alternative drugs with an antibacterial activity, but lacking resistance of microorganism strains to them, a computer screening of 2N-alkyl-substituted 5-amino-3-imino-1,2,4-thiadiazolines previously synthesized by us was carried out. The prediction of the spectrum of biological activity, as well as the determination of the probable toxicity of these compounds, was performed using the freely available computer programs PASS, Anti-Bac-Pred, and GUSAR. The study of the antibacterial activity in vitro against gram-positive (Staphylococcus aureus, Staphylococcus saprophyticus, Staphylococcus epidermidis) and gram-negative (Escherichia coli, Pseudomonas aeruginosae) bacterial strains was performed by the disco-diffusion method. Experimental data roughly correspond to the predictions.

Rudik A., Bezhentsev V., Dmitriev A., et al. Metatox - Web application for generation of metabolic pathways and toxicity estimation. *J. Bioinform. Comput. Biol.*, **2019**, 17(1):1940001.

Xenobiotics biotransformation in humans is a process of the chemical modifications, which may lead to the formation of toxic metabolites. The prediction of such metabolites is very important for drug development and ecotoxicology studies. We created the web-application MetaTox (http://way2drug.com/mg) for the generation of xenobiotics metabolic pathways in the human organism. For each generated metabolite, the estimations of the acute toxicity (based on GUSAR software prediction), organ-specific carcinogenicity and adverse effects (based on PASS software prediction) are performed. Generation of metabolites by MetaTox is based on the fragments datasets, which describe transformations of substrates structures to a metabolites structure. We added three new classes of biotransformation reactions: Dehydrogenation, Glutathionation, and Hydrolysis, and now metabolite generation for 15 most frequent

classes of xenobiotic's biotransformation reactions are available. MetaTox calculates the probability of formation of generated metabolite - it is the integrated assessment of the biotransformation reactions probabilities and their sites using the algorithm of PASS (http://way2drug.com/passonline). The prediction accuracy estimated by the leave-one-out cross-validation (LOO-CV) procedure calculated separately for the probabilities of biotransformation reactions and their sites is about 0.9 on the average for all reactions.

Nadaraia N.S., Amiranashvili L.S., Merlani M., et al. Novel antimicrobial agents' discovery among the steroid derivatives. *Steroids*, **2019**, 144: 52-65.

Fourteen steroid compounds were in silico evaluated using computer program PASS as antimicrobial agents. The experimental studies evaluation revealed that all compounds have good antibacterial activity with MIC at range of 0.003-0.96 mg/mL and MBC 0.06-1.92 mg/mL. Almost all compounds except of compound 4 (3β-acetoxy-1′-p-chlorophenyl-3′-methyl-5α-androstano[17,16-d]pyrazoline) were more potent than Ampicillin, and they were equipotent or more potent than Streptomycine. All compounds exhibited good antifungal activity with MIC at 0.003-0.96 mg/mL and MFC at 0.006-1.92 mg/mL but with different sensitivity against fungi tested. According to docking studies 14-alpha demethylase inhibition may be responsible for antifungal activity. Prediction of toxicity by PROTOX and GUSAR revealed that compounds have low toxicity and can be considered as potential lead compounds for the further studies.

Lagunin A.A., Geronikaki A., Eleftheriou P., et al. Rational Use of Heterogeneous Data in Quantitative Structure-Activity Relationship (QSAR) Modeling of Cyclooxygenase/Lipoxygenase Inhibitors. *J. Chem. Inf. Model.*, **2019**, 59 (2): 713-730.

Numerous studies have been published in recent years with acceptable quantitative structure-activity relationship (QSAR) modeling based on heterogeneous data. In many cases, the training sets for QSAR modeling were constructed from compounds tested by different biological assays, contradicting the opinion that QSAR modeling should be based on the data measured by a single protocol. We attempted to develop approaches that help to determine how heterogeneous data should be used for the creation of QSAR models on the basis of different sets of compounds tested by different experimental methods for the same target and the same endpoint. To this end, more than 100 QSAR models for the IC_{50} values of ligands interacting with cyclooxygenase 1,2 (COX) and seed lipoxygenase (LOX), obtained from ChEMBL database were created using the GUSAR software. The QSAR models were tested on the external set, including 26 new thiazolidinone derivatives, which were experimentally tested for COX-1,2/LOX inhibition. The IC_{50} values of the derivatives varied from 89 μM to 26 μM for LOX, from 200 μM to 0.018 μM for COX-1, and from 210 μM to 1 μM for COX-2. This study showed that the accuracy of the models is dependent on the distribution of IC_{50} values of low activity compounds in the training sets. In the most cases, QSAR models created based on the combined training sets had advantages in comparison with QSAR models, based on a single publication. We introduced a new method of combination of quantitative data from different experimental studies based on the data of reference compounds, which was called "scaling".

Martynova Y.Z., Khairullina V.R., Biglova Y.N., Mustafin A.G. Quantitative structure-property relationship modeling of the C60 fullerene derivatives as electron acceptors of polymer solar cells: Elucidating the functional groups critical for device performance. *J. Mol. Graph. Model.*, **2019**, 88: 49-61.

Using the GUSAR program, we have performed a quantitative analysis of the "structure-power conversion efficiency (PCE)" on the series of 100 methano[60]fullerenes previously tested as acceptor components of bulk-heterojunction polymer organic solar cells (PSCs) utilizing the same donor polymer, viz. poly(3-hexylthiophene). Based on the MNA and QNA descriptors and self-consistent regression implemented in the program, six statistically significant consensus models for predicting the PCE values of the methano[60]fullerene-based PSCs have been constructed. The structural fragments of the fullerene compounds leading to an increase in the device performances are determined. Based on these structural descriptors, we have designed the three methano[60]fullerenes included in the training sets and characterized by poor optoelectrical properties is performed. As a result, two new compounds with potentially moderate efficiency have been proposed. This result opens opportunities of using the GUSAR program for modeling of the "structure-PCE" relationship for diverse compounds (not only fullerene derivatives).

Lagunin A.A., Romanova M.A., Zadorozhny A.D., et al. Comparison of Quantitative and Qualitative (Q)SAR Models Created for the Prediction of Ki and IC50 Values of Antitarget Inhibitors. *Front. Pharmacol.*, **2018**, 9:1136.

Estimation of interaction of drug-like compounds with antitargets is important for the assessment of possible toxic effects during drug development. Publicly available online databases provide data on the experimental results of chemical interactions with antitargets, which can be used for the creation of (Q)SAR models. The structures and experimental Ki and IC50 values for compounds tested on the inhibition of 30 antitargets from the ChEMBL 20 database were used. Data sets with Ki and IC50 values including more than 100 compounds were created for each antitarget. The (Q)SAR models were created by GUSAR software using quantitative neighborhoods of atoms (QNA), multilevel neighborhoods of atoms (MNA) descriptors, and self-consistent regression. The accuracy of (Q)SAR models was validated by the fivefold cross-validation procedure. The balanced accuracy was higher for qualitative SAR models (0.80 and 0.81 for Ki and IC50 values, respectively) than for quantitative QSAR models (0.73 and 0.76 for Ki and IC50 values, respectively). In most cases, sensitivity was higher for SAR models than for QSAR models, but specificity was higher for QSAR models. The mean R 2 and RMSE were 0.64 and 0.77 for Ki values and 0.59 and 0.73 for IC50 values, respectively. The number of compounds falling within the applicability domain was higher for SAR models than for the test sets.

Khairullina V.R., Gimadieva A.R., Gerchikov A.Y., et al. Quantitative structure-activity relationship of the thymidylate synthase inhibitors of Mus musculus in the series of quinazolin-4-one and quinazolin-4-imine derivatives. *J. Mol. Graph. Model.*, **2018**, 85: 198-211.

A quantitative structure-activity relationship analysis of the 2-methylquinazolin-4-one and quinazolin-4-imine derivatives, well-known antifolate thymidylate synthase (TYMS) inhibitors, has been performed in the range IC_{50} = 0.4÷380000.0 nmoL/L using the GUSAR program. Based on the MNA and QNA descriptors using the self-consistent regression, 6 statistically significant consensus models for predicting the IC_{50} numerical values have been constructed. These models demonstrate high and moderate prognostic accuracies for the training and external validation test sets, respectively. The molecular fragments of TYMS inhibitors regulating their antitumor activity are identified. The obtained data open opportunities for developing novel promising inhibitors of TYMS.

Ivashchenko D.V., Rudik A.V., Poloznikov A.A., et al. Which cytochrome P450 metabolizes phenazepam? Step by step *in silico*, *in vitro*, and *in vivo* studies. *Drug Metab. Pers. Ther.*, **2018**, 33(2):65-73.

Background: Phenazepam (bromdihydrochlorphenylbenzodiazepine) is the original Russian benzodiazepine tranquilizer belonging to 1,4-benzodiazepines. There is still limited knowledge about phenazepam's metabolic liver pathways and other pharmacokinetic features. **Methods:** To determine phenazepam's metabolic pathways, the study was divided into three stages: in silico modeling, in vitro experiment (cell culture study), and in vivo confirmation. In silico modeling was performed on the specialized software PASS and GUSAR to evaluate phenazepam molecule affinity to different cytochromes. The in vitro study was performed using a hepatocytes' cell culture, cultivated in a microbioreactor to produce cytochrome P450 isoenzymes. The culture medium contained specific cytochrome P450 isoforms inhibitors and substrates (for CYP2C9, CYP3A4, CYP2C19, and CYP2B6) to determine the cytochrome that was responsible for phenazepam's metabolism. We also measured CYP3A activity using the 6-betahydroxycortisol/cortisol ratio in patients. **Results:** According to in silico and in vitro analysis results, the most probable metabolizer of phenazepam is CYP3A4. By the in vivo study results, CYP3A activity decreased sufficiently (from 3.8 [95% CI: 2.94-4.65] to 2.79 [95% CI: 2.02-3.55], p=0.017) between the start and finish of treatment in patients who were prescribed just phenazepam. **Conclusions:** Experimental in silico and in vivo studies confirmed that the original Russian benzodiazepine phenazepam was the substrate of CYP3A4 isoenzyme.

Podlewska S., Kafel R. MetStabOn-Online Platform for Metabolic Stability Predictions. *Int. J. Mol. Sci.*, **2018**, 19(4): 1040.

Metabolic stability is an important parameter to be optimized during the complex process of designing new active compounds. Tuning this parameter with the simultaneous maintenance of a desired compound's activity is not an easy task due to the extreme complexity of metabolic pathways in living organisms. In this study, the platform for *in silico* qualitative evaluation of metabolic stability, expressed as half-lifetime and clearance was developed. The platform is based on the application of machine learning methods and separate models for human, rat and mouse data were constructed. The compounds' evaluation is qualitative and two types of experiments can be performed-regression, which is when the compound is assigned to one of the metabolic stability classes (low, medium, high) on the basis of numerical value of the predicted half-lifetime, and classification, in which the molecule is directly assessed as low, medium or high stability. The results show that the models have good predictive power, with accuracy values over 0.7 for all cases, for Sequential Minimal Optimization (SMO), k-nearest neighbor (IBk) and Random Forest algorithms. Additionally, for each of the analyzed compounds, 10 of the most similar structures from the training set (in terms of Tanimoto metric similarity) are identified and made available for download as separate files for more detailed manual inspection. The predictive power of the models was confronted with the external dataset, containing metabolic stability assessment via the GUSAR software, leading to good consistency of results for SMOreg and Naïve Bayes (~0.8 on average).

Rudik A.V., Bezhentsev V.M., Dmitriev A.V., et al. MetaTox: Web Application for Predicting Structure and Toxicity of Xenobiotics' Metabolites. *J. Chem. Inf. Model.*, **2017**, 57(4): 638-642.

A new freely available web-application MetaTox (http://www.way2drug.com/mg) for prediction of xenobiotic's metabolism and calculation toxicity of metabolites based on the structural formula of chemicals has been developed. MetaTox predicts metabolites, which are formed by nine classes of reactions (aliphatic and aromatic hydroxylation, N- and O-glucuronidation, N-, S- and C-oxidation, and N- and O-dealkylation). The calculation of probability for generated metabolites is based on analyses of "structure-biotransformation reactions" and "structure-modified atoms" relationships using a Bayesian approach. Prediction of LD50 values is performed by GUSAR software for the parent compound and each of the generated metabolites using quantitative structure-activity relationship (QSAR) models created for acute rat toxicity with the intravenous type of administration.

Savjani J.K., Mulamkattil S., Variya B., Patel S. Molecular docking, synthesis and biological screening of mefenamic acid derivatives as anti-inflammatory agents. *Eur. J. Pharmacol.*, **2017**, 801: 28-34.

Drug induced gastrointestinal ulceration, renal side effects and hepatotoxicity are the main causes of numerous Non-Steroidal Anti-inflammatory Drugs (NSAIDs). Cyclooxygenase-2 (COX-2) inhibitors discovered to decrease the gastrointestinal issues, but unfortunately, most of them are associated with major cardiovascular adverse effects. Along these lines, various new strategies and frameworks were developed wherein basic alterations of the present medications were accounted for. The aim of the study was to prepare

derivatives of mefenamic acid to evaluate anti-inflammatory activity with fewer adverse reactions. In this study, molecular docking investigations of outlined derivatives were done utilizing Protein Data Bank (PDB ID-4PH9). Synthesis of heterocyclic compounds was carried out utilizing Dicyclohexylcarbodiimide/4-Dimethylaminopyridine (DCC/DMAP) coupling. Acute toxicity prediction was performed using free online GUSAR (General Unrestricted Structure-Activity Relationships) software. The study indicated most of the compounds under safe category. In-vitro pharmacological assessment of heterocyclic compounds was done for COX-1 and COX-2 enzymes for the determination of selectivity. In vivo pharmacological screening for anti-inflammatory activity and ED50 value were determined utilizing carrageenan induced rat paw edema. Gastro intestinal safety study was carried out on selected compounds and found to be devoid of any gastric ulcer toxicity. Most of the compounds indicated high scores as compared to standard during molecular modelling, analysis and displayed interactions with active amino acids of a COX-2 enzyme. The pharmacological screening uncovered that compound substituted with p-bromophenyl indicated maximum potency.

Ajeet V.M., Rani S., Kumar A. Antitarget Interaction, Acute Toxicity and Protein Binding Studies of Quinazolinedione Sulphonamides as GABA1 Antagonists. *Indian J. Pharm. Sci*., **2016**, 78(1): 48-53.

Diseases characterized by recurrent seizures are known as epilepsy. One of the most important mechanisms for handling it is GABA1 receptor mediated inhibition. In the same context while studying the treatment of epilepsy we observed significant effects by derivatives of sulfonamides, which prompted us to design novel derivatives by means of in silico resources with antiepileptic effects. Molecular docking approaches are routinely used in modern drug design to help understand drug-receptor interaction. This study has been performed with the help of Chemdraw Ultra 7.0, GUSAR online tool for IC50 and LD50 predictions, AutoDock Vina (Python Prescription 0.8), and PaDEL software. Results revealed that ligand-protein interaction affinity of all 10 designed molecules ranges from -5.7 Kcal/mol to -5.2 Kcal/mol, which is approximately comparable to pre-existing GABA1 inhibitor i.e. phenytoin (CID: 1775, ligand-protein interaction affinity is -6.5 Kcal/mol).

Zakharov A.V., Varlamova E.V., Lagunin A.A., et al. QSAR Modeling and Prediction of Drug-Drug Interactions. *Mol. Pharm*., **2016**, 13(2):545-56.

Severe adverse drug reactions (ADRs) are the fourth leading cause of fatality in the U.S. with more than 100,000 deaths per year. As up to 30% of all ADRs are believed to be caused by drug-drug interactions (DDIs), typically mediated by cytochrome P450s, possibilities to predict DDIs from existing knowledge are important. We collected data from public sources on 1485, 2628, 4371, and 27,966 possible DDIs mediated by four cytochrome P450 isoforms 1A2, 2C9, 2D6, and 3A4 for 55, 73, 94, and 237 drugs, respectively. For each of these data sets, we developed and validated QSAR models for the prediction of DDIs. As a unique feature of our approach, the interacting drug pairs were represented as binary chemical mixtures in a 1:1 ratio. We used two types of chemical descriptors: quantitative neighborhoods of atoms (QNA) and simplex descriptors. Radial basis functions with self-consistent regression (RBF-SCR) and random forest (RF) were utilized to build QSAR models predicting the likelihood of DDIs for any pair of drug molecules. Our models showed balanced accuracy of 72-79% for the external test sets with a coverage of 81.36-100% when a conservative threshold for the model's applicability domain was applied. We generated virtually all possible binary combinations of marketed drugs and employed our models to identify drug pairs predicted to be instances of DDI. More than 4500 of these predicted DDIs that were not found in our training sets were confirmed by data from the DrugBank database.

Goncharuk V.V., Buben A.L., Borisenok O.A., et al. Analgesic activity of some new decahydroquinoline derivatives. *Experimental and Clinical Pharmacology*, **2016**, 79 (11), 7-10.

The study was aimed at assessing in vivo the analgesic properties of ten decahydroquinoline derivatives (pharmacologically active substances, PAS) and deter- mining the role of opioid receptors in mechanism of their action. Among the derivatives studied, pronounced analgesic properties at a dose of 1/4 LD50 was observed for two compounds (PAS-70 and PAS-71), while four compounds (PAS-66, PAS-69, PAS-74, PAS-76) produced weak and short anesthetic effects. PAS-70 and PAS-71 showed analgesic action even in a dose of 1/8 LD50. The maximum effect of PAS-70 and PAS-71 was developed within 20 - 60 min after administration and lasted for two hours (PAS-70 in a dose of 1/4 LD50, PAS-71 in doses of 1/4 and 1/8 LD50). The analgesic effect of PAS-70 (at 1/4 LD5,) and PAS-71 (at 1/4 and 1/8 LD50) significantly exceeds that of reference drugs metamizol (1/4 LD5) and ketorolac (1/4 LD50). The mechanism of drug action is not related to opioid receptors.

Khayrullina V.R., Gerchikov A.Y., Lagunin A.A., Zarudii F.S. Quantitative analysis of structure-activity relationships of tetrahydro-2H-isoindole cyclooxygenase-2 inhibitors. *Biochemistry (Mosc).*, **2015**, 80(1):74-86.

Using the GUSAR program, structure-activity relationships on inhibition of cyclooxygenase-2 (COX-2) catalytic activity were quantitatively analyzed for twenty-six derivatives of 4,5,6,7-tetrahydro-2H-isoindole, 2,3-dihydro-1H-pyrrolyzine, and benzothiophene in the concentration range of 0.6-700 nmol/liter IC50 values. Six statistically significant consensus QSAR models for prediction of IC50 values were designed based on MNA- and QNA-descriptors and their combinations. These models demonstrated high accuracy in the prediction of IC50 values for structures of both training and test sets. Structural fragments of the COX-2 inhibitors capable of strengthening or weakening the desired property were determined using the same program. This information can be taken into consideration on molecular design of new COX-2 inhibitors. It was shown that in most cases, the influence of structural fragments on the inhibitory activity of the studied compounds revealed with the GUSAR program coincided with the results of expert evaluation of their effects based on known experimental data, and this can be used for optimization of structures to change the value of their biological activity.

Fedorova E.V., Buryakina A.V., Zakharov A.V., et al. Design, synthesis and pharmacological evaluation of novel vanadium-containing complexes as antidiabetic agents. *PLoS One*, **2014**, 9(7): e100386.

Based on the data about structure and antidiabetic activity of twenty seven vanadium and zinc coordination complexes collected from literature we developed QSAR models using the GUSAR program. These QSAR models were applied to 10 novel vanadium coordination complexes designed in silico in order to predict their hypoglycemic action. The five most promising substances with predicted potent hypoglycemic action were selected for chemical synthesis and pharmacological evaluation. The selected coordination vanadium complexes were synthesized and tested in vitro and in vivo for their hypoglycemic activities and acute rat toxicity. Estimation of acute rat toxicity of these five vanadium complexes was performed using a freely available web-resource (http://way2drug.com/GUSAR/acutoxpredict.html). It has shown that the selected compounds belong to the class of moderate toxic pharmaceutical agents, according to the scale of Hodge and Sterner. Comparison with the predicted data has demonstrated a reasonable correspondence between the experimental and predicted values of hypoglycemic activity and toxicity. Bis{tert-butyl[amino(imino)methyl]carbamato}oxovanadium (IV) and sodium(2,2'-Bipyridyl)oxo-diperoxovanadate(V) octahydrate were identified as the most potent hypoglycemic agents among the synthesized compounds.

Zakharov A.V., Peach M.L., Sitzmann M., Nicklaus M.C. QSAR modeling of imbalanced high-throughput screening data in PubChem. *J. Chem. Inf. Model.*, **2014**, 54(3): 705-712.

Many of the structures in PubChem are annotated with activities determined in high-throughput screening (HTS) assays. Because of the nature of these assays, the activity data are typically strongly imbalanced, with a small number of active compounds contrasting with a very large number of inactive compounds. We have used several such imbalanced PubChem HTS assays to test and develop strategies to efficiently build robust QSAR models from imbalanced data sets. Different descriptor types [Quantitative Neighborhoods of Atoms (QNA) and "biological" descriptors] were used to generate a variety of QSAR models in the program GUSAR. The models obtained were compared using external test and validation sets. We also report on our efforts to incorporate the most predictive of our models in the publicly available NCI/CADD Group Web services (http://cactus.nci.nih.gov/chemical/apps/cap).

Zakharov A.V., Peach M.L., Sitzmann M., Nicklaus M.C. A new approach to radial basis function approximation and its application to QSAR. *J. Chem. Inf. Model.*, **2014**, 54(3): 713-9.

We describe a novel approach to RBF approximation, which combines two new elements: (1) linear radial basis functions and (2) weighting the model by each descriptor's contribution. Linear radial basis functions allow one to achieve more accurate predictions for diverse data sets. Taking into account the contribution of each descriptor produces more accurate similarity values used for model development. The method was validated on 14 public data sets comprising nine physicochemical properties and five toxicity endpoints. We also compared the new method with five different QSAR methods implemented in the EPA T.E.S.T. program. Our approach, implemented in the program GUSAR, showed a reasonable accuracy of prediction and high coverage for all external test sets, providing more accurate prediction results than the comparison methods and even the consensus of these methods. Using our new method, we have created models for physicochemical and toxicity endpoints, which we have made freely available in the form of an online service at http://cactus.nci.nih.gov/chemical/apps/cap.

Lagunin A.A., Gloriozova T.A., Dmitriev A.V., et al. Computer evaluation of drug interactions with P-glycoprotein. *Bull. Exp. Biol. Med.*, **2013**, 154(4): 521-524.

The (Q)SAR models for evaluating the structure-property relationships, fit for prediction of drug interactions with P-glycoprotein as inhibitors or substrates, were constructed using PASS and GUSAR software. The models were constructed and validated on the basis of information on the structure and characteristics of 256 and 94 compounds used as P-glycoprotein substrates and inhibitors, respectively. The initial samples were divided 80:20 into training and test samples. The best prediction accuracy for the test samples was 78% for P-glycoprotein substrate prediction (PASS) and 89% for inhibitor prediction (GUSAR).

Masand V.H., Mahajan D.T., Patil K.N., et al. Optimization of antimalarial activity of synthetic prodiginines: QSAR, GUSAR, and CoMFA analyses. *Chem. Biol. Drug Des.*, **2013**, 81(4): 527-536.

In the present study, we have carried out extensive General Unrestricted Structure-Activity Relationships, conventional 3D-Quantitative Structure-Activity Relationships, and CoMFA analyses of synthetic prodiginines displaying moderate to high activities against Plasmodium Falciperum. 2D and 3D descriptors, various statistical parameters viz. R(2), R(2)(adj), standard error, Y-randomization, etc., were checked to build fruitful 3D-Quantitative Structure-Activity Relationships model. The best five parametric 3D-Quantitative Structure-Activity Relationships model is with R(2) = 0.924 and R(2)(pred) = 0.901. CoMFA was performed to check the electrostatic and steric regions, which affect the activity. The CoMFA model is graphically inferred using contour plots, which provide insight into the structural requirements for increasing the activity of a compound. The General Unrestricted Structure-Activity Relationships model, with R(2) = 0.940 and Q(2) = 0.912, suggests that the presence of F on aromatic ring is good for activity. The analyses reveal that lipophilicity plays a crucial role in deciding the activity for these molecules.

Zakharov A.V., Peach M.L., Sitzmann M., et al. Computational tools and resources for metabolism-related property predictions. 2. Application to prediction of half-life time in human liver microsomes. *Future Med. Chem.*, **2012**, 4(15): 1933-1944.

Background: The most important factor affecting metabolic excretion of compounds from the body is their half-life time. This provides an indication of compound stability of, for example, drug molecules. We report on our efforts to develop QSAR models for metabolic stability of compounds, based on in vitro half-life assay data measured in human liver microsomes.
Method: A variety of QSAR models generated using different statistical methods and descriptor sets implemented in both open-source and commercial programs (KNIME, GUSAR and StarDrop) were analyzed. The models obtained were compared using four different external validation sets from public and commercial data sources, including two smaller sets of in vivo half-life data in humans.
Conclusion: In many cases, the accuracy of prediction achieved on one external test set did not correspond to the results achieved with another test set. The most predictive models were used for predicting the metabolic stability of compounds from the open NCI database, the results of which are publicly available on the NCI/CADD Group web server (http://cactus.nci.nih.gov).

Zakharov A.V., Lagunin A.A., Filimonov D.A., Poroikov V.V. Quantitative prediction of antitarget interaction profiles for chemical compounds. *Chem. Res Toxicol.*, **2012**, 25(11): 2378-2385.

The evaluation of possible interactions between chemical compounds and antitarget proteins is an important task of the research and development process. Here, we describe the development and validation of QSAR models for the prediction of antitarget end-points, created on the basis of multilevel and quantitative neighborhoods of atom descriptors and self-consistent regression. Data on 4000 chemical compounds interacting with 18 antitarget proteins (13 receptors, 2 enzymes, and 3 transporters) were used to model 32 sets of end-points (IC(50), K(i), and K(act)). Each set was randomly divided into training and test sets in a ratio of 80% to 20%, respectively. The test sets were used for external validation of QSAR models created on the basis of the training sets. The coverage of prediction for all test sets exceeded 95%, and for half of the test sets, it was 100%. The accuracy of prediction for 29 of the end-points, based on the external test sets, was typically in the range of R(2)(test) = 0.6-0.9; three tests sets had lower R(2)(test) values, specifically 0.55-0.6. The proposed approach showed a reasonable accuracy of prediction for 91% of the antitarget end-points and high coverage for all external test sets. On the basis of the created models, we have developed a freely available online service for in silico prediction of 32 antitarget end-points: http://www.pharmaexpert.ru/GUSAR/antitargets.html.

Kokurkina G.V., Dutov M.D., Shevelev S.A., et al. Synthesis, antifungal activity and QSAR study of 2-arylhydroxynitroindoles. *Eur. J. Med. Chem.*, **2011**, 46(9): 4374-4382.

A series of 2-arylhydroxynitroindoles were prepared and tested for antifungal activity in vitro. The preliminary bioassays indicated that some compounds are comparable to the commercial fungicide (triadimefon). To further explore the structure-activity relationships, the data set of the seventeen structures and their quantitative values of antifungal activities were used for QSAR modeling. Based on the obtained QSAR models four new chemical compounds were designed, synthesized and tested in fungicidal assays. Reasonable correspondence between the experimental and predicted values of antifungal activity was observed.

Lagunin A., Zakharov A., Filimonov D., Poroikov V. QSAR Modelling of Rat Acute Toxicity on the Basis of PASS Prediction. *Mol. Inform.*, **2011**, 30(2-3): 241-250.

The method for QSAR modelling of rat acute toxicity based on the combination of QNA (Quantitative Neighbourhoods of Atoms) descriptors, PASS (Prediction of Activity Spectra for Substances) predictions and self-consistent regression (SCR) is presented. PASS predicted biological activity profiles are used as independent input variables for QSAR modelling with SCR. QSAR models were developed using LD50 values for compounds tested on rats with four types of administration (oral, intravenous, intraperitoneal, subcutaneous). The proposed method was evaluated on the set of compounds tested for acute rat toxicity with oral administration (7286 compounds) used for testing the known QSAR methods in T.E.S.T. 3.0 program (U.S. EPA). The several other sets of compounds tested for acute rat toxicity by different routes of administration selected from SYMYX MDL Toxicity Database were used too. The method was compared with the results of prediction of acute rodent toxicity for noncongeneric sets obtained by ACD/Labs Inc. The test sets were predicted with regards to the applicability domain. Comparison of accuracy for QSAR models obtained separately using QNA descriptors, PASS predictions, nearest neighbours' assessment with consensus models clearly demonstrated the benefits of consensus prediction. Free available web-service for prediction of LD50 values of rat acute toxicity was developed: http://www.pharmaexpert.ru/GUSAR/AcuToxPredict/.

Filimonov D.A., Zakharov A.V., Lagunin A.A., Poroikov V.V. QNA-based 'Star Track' QSAR approach. *SAR QSAR Environ. Res.*, **2009**, 20(7-8): 679-709.

In the existing quantitative structure-activity relationship (QSAR) methods any molecule is represented as a single point in a many-dimensional space of molecular descriptors. We propose a new QSAR approach based on Quantitative Neighbourhoods of Atoms (QNA) descriptors, which characterize each atom of a molecule and depend on the whole molecule structure. In the 'Star Track' methodology any molecule is represented as a set of points in a two-dimensional space of QNA descriptors. With our new method the estimate of the target property of a chemical compound is calculated as the average value of the function of QNA descriptors in the points of the atoms of a molecule in QNA descriptor space. Substantially, we propose the use of only two descriptors rather than more than 3000 molecular descriptors that apply in the QSAR method. On the basis of this approach we have developed the computer program GUSAR and compared it with several widely used QSAR methods including CoMFA, CoMSIA, Golpe/GRID, HQSAR and others, using ten data sets representing various chemical series and diverse types of biological activity. We show that in the majority of cases the accuracy and

predictivity of GUSAR models appears to be better than those for the reference QSAR methods. High predictive ability and robustness of GUSAR are also shown in the leave-20%-out cross-validation procedure.

Zakharov A.V., Lagunin A.A., Filimonov D.A., Poroikov V.V. Quantitative structure-activity relationships of cyclin-dependent kinase 1 inhibitors. *Biomed. Khim*., **2006**, 52(1): 3-18.
A new approach for quantitative structure-activity relationships based on MNA descriptors, fuzzy gradation method and self-consistent regression has been proposed. This approach has been realized in the computer program GUSAR (General Unrestricted Structure Activity Relationships). Our method has been validated on CDK1 inhibitors. Prediction accuracy is comparable with popular methods of 3D QSAR: CoMFA and CoMSIA. However, in contrast to CoMFA and CoMS1A, GUSAR approach does not require information about 3D structure of enzyme and ligand. Application of GUSAR method for heterogeneous training sets has been shown.

Poroikov V.V., Filimonov D.A., Borodina Y.V., et al. Quantitative Relationships Between Structure and Delayed Neurotoxicity of Chemicals Studied by the Self-Consistent Regression Method Using the Pass Program. *Pharmaceutical Chemistry Journal*, **2004**, 38, 188–190. doi: 10.1023/B:PHAC.0000038416.33964.dc. LINK

Filimonov D.A., Akimov D.V., Poroikov V.V. The Method of Self-Consistent Regression for the Quantitative Analysis of Relationships Between Structure and Properties of Chemicals. *Pharmaceutical Chemistry Journal*, **2004**, 38, 21–24. doi: 10.1023/B:PHAC.0000027639.17115.5d. LINK

For questions, you are welcome to address the geneXplain support team:
GeneXplain GmbH
Am Exer 10 b
D-38302 Wolfenbüttel
Germany
office: +49-5331-992200-0
fax: +49-5331-992200-22
e-mail: support@genexplain.com
URL: www.genexplain.com