



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



ORIGINAL ARTICLE

Prediction of inhibitory constants of compounds against SARS-CoV 3CLpro enzyme with 2D-QSAR model



Xinliang Yu

Hunan Provincial Key Laboratory of Environmental Catalysis & Waste Regeneration, College of Materials and Chemical Engineering, Hunan Institute of Engineering, Xiangtan, Hunan 411104, China

Received 1 April 2021; revised 3 May 2021; accepted 14 May 2021
Available online 21 May 2021

KEYWORDS

COVID-19;
QSAR;
SARS-CoV;
SARS-CoV-2

Abstract Developing broad-spectrum anti-coronavirus drugs is greatly important, since the novel SARS-CoV-2 has rapidly become a threat to the public health and economy worldwide. SARS-CoV 3-chymotrypsin-like protease (3CLpro), as highly conserved in betacoronavirus, is a viable target for anti-SARS drugs. A quantitative structure–activity relationship (QSAR) for inhibitory constants (pKi) of 89 compounds against SARS-CoV 3CLpro enzyme was developed by using support vector machine (SVM) and genetic algorithm. The optimal SVM model ($C = 90.2339$ and $\gamma = 1.19826 \times 10^{-5}$) based on six molecular descriptors has determination coefficients of 0.839 for the training set (65 compounds) and 0.747 for test set (24 compounds), and *rms* errors of 0.435 and 0.525, respectively. These results are accurate and acceptable compared with that in other models reported, although our SVM model deals with more samples in the data set. The SVM model could be beneficial for search of novel 3CLpro enzyme inhibitors against SARS-CoV.

© 2021 The Author(s). Published by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

COVID-19, caused by the SARS-CoV2 virus, has been found all over the world since its first outbreak in Wuhan City, China, in December 2019 [1,2]. COVID-19 is a potentially fatal disease and becomes a global public health concern [3,4]. As of 1 May 2021, 8:36 pm GMT + 8, 150,989,419 cases have been

E-mail address: yxl@hnie.edu.cn

Peer review under responsibility of King Saud University.



reported worldwide, resulting in 3,173,576 deaths (<https://www.who.int/>). Two human coronavirus diseases previously discovered in the 21st century, severe acute respiratory syndrome CoV (SARS-CoV) in 2002 and Middle East respiratory syndrome CoV (MERS-CoV) in 2012, respectively, infected at least 8422 and 1700 people. Their fatality rates were about 10% and 36%, respectively [5,6]. Although SARS-CoV disappeared mysteriously, the MERS-CoV has not been controlled so far. SARS-CoV2 is the 3rd human coronavirus disease discovered in the 21st century. At present, there are no specific and effective drugs for COVID-19.

SARS-CoV 3-chymotrypsin-like protease (3CLpro) is one of the major proteases produced by the 2019-nCoV, which plays a pivotal role in the replication of the virus [7,8]. Most

of the functional proteins of coronavirus are encoded by ORF1ab gene. The coding RNAs are translated into a proteome (7096aa), and then cut into multiple active proteins (e.g. viral replication protein RdRp) by 3CLpro. In addition, 3CLpro may cleave the intracellular protein NEMO and inhibit the activation of interferon signaling pathway. Therefore, SARS-CoV 3CLpro plays an important role in the virus life cycle and has become a viable target for anti-SARS drug development. There is >95% sequence similarity in RdRp and 3CLpro between SARS-CoV-2 and SARS-CoV [3–8]. Therefore, 3CLpro inhibitors developed on SARS-CoV could be effective against SARS-CoV-2 [7,8].

Quantitative structure–activity relationship (QSAR) models can be used for drug screening and mechanistic understanding of drug action. This technique has many advantages, such as lower-cost and higher speed, even can be used to evaluate drug candidates that have not been synthesized [9,10]. But only a few researchers have carried out QSAR studies for inhibitor activities against SARS-CoV.

Inhibitory constant (K_i) is the concentration of the inhibitor that is required in order to decrease the maximal rate of the reaction by half. Masand et al. introduced a QSAR model for activities (K_i) of SARS-CoV 3CLpro inhibitors [7]. A six-descriptor model was based on a training set (50 compounds) and evaluated with a test set (12 compounds). Although the data sets are relatively small, the coefficients of determination R^2 are 0.824 for the training set and 0.758 for the test set, which are accurate and satisfactory. Kumar and Roy built up an eight-descriptor model for inhibitory activities (IC_{50}) of 69 molecules against SARS-CoV 3CLpro enzyme [8]. The data set was divided into a training set (56 molecules) and a test set (13 molecules). Their coefficients of determination R^2 are 0.764 and 0.711, respectively.

The two QSAR models referred to were developed with multiple linear regression (MLR) analysis, which is suitable for linear relationships between dependent variables and independent variables. Generally, nonlinear regression techniques can improve the prediction performance of QSAR models. The aim of this study is to develop a six-descriptor QSAR for inhibitory constants (K_i) of 89 molecules against SARS-CoV 3CLpro enzyme, by applying support vector machine (SVM) technique. It is hoped that our SVM model will be beneficial for search of novel 3CLpro enzyme inhibitors against SARS-CoV.

2. Methods

2.1. Experimental data

Table S1 in Supplemental Information shows the SMILES notations and inhibitory constants (K_i) of 89 molecules against SARS-CoV 3CLpro, which were taken from the binding database [11] and references [12–14]. The experimental K_i values varied from 3 to 56,000 nM, and their pK_i ($= -\log K_i$) values were in the range of 8.523–4.252 by converting to negative logarithm of K_i . A larger pK_i value means a higher activity for the inhibitor. Inhibition constants, K_i , were obtained through measuring the apparent kinetic parameters at a constant substrate concentration (10 μ M) and different inhibitor concentrations (0–200 μ M) at 25 °C [12]. These experimental data were randomly split into a training set ($n = 65$ inhibitors) and a test

set ($n = 24$ inhibitors). QSAR models were developed with the training set and evaluated with test set.

2.2. Molecular descriptors

Besides three-dimensional (3D) QSAR methods based on ligand-receptor interactions, the two-dimensional (2D) QSAR models derived only from the ligand molecules can be used for describing the activity of biologically active compounds. The structural and physicochemical features of active compounds become the critical factors determining inhibitory constant (K_i) when the inhibitors have the same biological target (e.g. SARS-CoV 3CLpro enzyme). In this study, the structures of inhibitors were used to derive molecular descriptors for 2D-QSAR models of inhibitory constants (K_i). According to the SMILES notations in Table S1, molecular structures were drawn using ChemBioDraw Ultra 12.0 in ChemBioOffice 2010. Subsequently, 3D-structures were generated using ChemBio3D Ultra 12.0 and optimized with semi-empirical AM1 method in Gaussian 09. Finally, the optimized molecules were used to calculate molecular descriptors with Dragon 6.0 [15]. Totally, 648 descriptors were obtained when those molecular descriptors with high co-linearity ($|R| > 0.90$) or being a constant were removed.

2.3. SVM principle

For the nonlinear support vector regression machine, the low-dimensional data need be mapped to the high-dimensional space, from which the linearly separable hyperplane would be found. Finally, the hyperplane in the high-dimensional space should be mapped back to the low-dimensional space, so as to realize SVM regression or classification. However, mapping low-dimensional data to high-dimensional space and then performing regression analysis involve a great quantity of computations. Especially for high-dimensional data, the problem of over-fitting can occur. Kernel function is introduced to solve this problem. Replacing the linear terms in linear equations with kernel function can make the original linear algorithm nonlinear, that is to say, it can do nonlinear regression. Thus, the introduction of kernel function can achieve the purpose of increasing dimension and effectively control over fitting. Radial basis function was used in this work. The SVM parameters C (the regulation constant) and γ (the width of kernel function) need to be optimized with genetic algorithm, since too large or too small C (and γ) values may lead to over-fitting or under-fitting and reduce model's predictive power.

3. Results and discussion

3.1. Optimal descriptor subset

Stepwise multiple linear regression (MLR) analysis was performed using IBM SPSS Statistical 19 for 89 pK_i (in Table S1 in Supplemental Information) and 648 descriptors. The increment of determination coefficient $\Delta R^2 > 0.02$ was used as the criterion for introducing new variables. A subset consisting of six descriptors (MATS6m, MATS1s, MATS3s, nArNR2, C-028, and F10[N-O]) were obtained. The descriptor

definitions were listed in Table 1. We found that the determination coefficients of MLR models were improved when these descriptors were divided by their respective molecular weight (MW) and used as new independent variables for MLR model. Therefore, the new subset of descriptors (MATS6m/MW, MATS1s/MW, MATS3s/MW, nArNR2/MW, C-028/MW, and F10[N-O]/MW) (see Table S1 in Supplemental Information) were used to develop QSAR in this work. The MLR model has determination coefficients R^2 of 0.598 for the training set and 0.659 for test set.

The three descriptors, MATS6m, MATS1s and MATS3s, belong to Moran autocorrelation (MATSkw) [15], which are calculated with:

$$\text{MATSkw} = \frac{\frac{1}{2\Delta} \cdot \sum_{i=1}^{nAT} \sum_{j=1}^{nAT} \delta_{ij} \cdot (w_i - w_j)^2}{\frac{1}{nAT} \cdot \sum_{i=1}^{nAT} (w_i - w^-)^2} \quad (1)$$

where k denotes the lag value (from $k = 1$ to 8), w means the atomic property (viz. atomic mass (m), atomic van der Waals volume (v), atomic Sanderson electronegativity (e), atomic polarizability (p), atomic ionization potential (i), or intrinsic state (s)) used for weighting molecular graphs, nAT is the total number of molecule atoms, δ_{ij} denotes the Kronecker delta and Δ denotes the sum of the Kronecker deltas [15]. MATSkw are correlated with molecular size, geometry and symmetry [16,17].

The descriptor nArNR2 denotes the number of tertiary amines (aromatic), i.e. ArNYY (Ar being aromatic ring and Y being aromatic ring or aliphatic groups). The descriptor C-028 is defined as the number of R--CR--X group (R means any group linked to C atom; X denotes any electronegative atom, O, N, S, P, Se, or halogens). The descriptor F10[N-O] is the frequency of N-O at topological distance 10. F10[N-O] value is 1 or 0, representing the presence or absence of N-O atom pairs at topological distance 10.

Table 2 shows the characteristics of new descriptors (MATS6m/MW, MATS1s/MW, MATS3s/MW, nArNR2/MW, C-028/MW, and F10[N-O]/MW) in MLR model. As is shown in Table 2, the six new descriptors have low significance values (<0.05), which mean that they are important independent variables to correlate with the inhibitory constants (pKi). The variance inflation factors (VIF) are less than 10, implying no serious multicollinearity problem. By the t -test values in

Table 2, the $|t\text{-test}|$ values increase in the sequence: MATS1s/MW, MATS6m/MW, C-028/MW, nArNR2/MW, F10[N-O]/MW, and MATS3s/MW, and the relative importance of descriptors increases in the same sequence.

3.2. SVM model

The new descriptors (MATS6m/MW, MATS1s/MW, MATS3s/MW, nArNR2/MW, C-028/MW, and F10[N-O]/MW) were used as independent variables to develop SVM model for inhibitory constants pKi, by employing LibSVM and MATLAB R2014a [18]. Leave-one-out (LOO) cross-validation, together with genetic algorithm [19,20], was selected to train the QSAR models for pKi of 65 compounds in the training set. The ranges of C from 0 to 600 and γ from 0 and 1 were searched. In the end, the relatively optimal parameters C ($=90.2339$) and γ ($=1.19826 \times 10^{-5}$) were obtained. The optimal SVM ($C = 90.2339$ and $\gamma = 1.19826 \times 10^{-5}$) model was evaluated with 24 compounds in the test set. The pKi values calculated from the SVM model were listed in Table S1 in Supplemental Information and illustrated in Fig. 1. The determination coefficients of the training and test sets are 0.839 and 0.747, respectively; and their rms errors are 0.435 and 0.525, respectively. These results are accurate and acceptable compared with that (R^2 being 0.824 for the training set and 0.758 for the test set; rms being 0.433 and 0.527, respectively) from the QSAR model reported [7], although our SVM model dealt with more samples. According to the prediction results in test set, we can obtain some statistical parameters: $k = 0.974$, $k' = 1.021$, $q_{\text{ext}}^2 = 0.756$, $R^2 = 0.747$, $R_0^2 = 0.745$, $R_0'^2 = 0.696$, which satisfy the criteria: $0.85 \leq k$ (or $k' \leq 1.15$), $q_{\text{ext}}^2 > 0.5$; $R^2 > 0.6$; and $(R^2 - R_0^2)/R^2 < 0.1$ [21]. Therefore, the pKi SVM model built in this work is successful. Here k and k' are slopes of regression lines, q_{ext}^2 is external correlation coefficient, R^2 is square of correlation coefficient, R_0^2 and $R_0'^2$ determination coefficient. These definitions of above parameters can be found in references [21–23]. The test compounds have absolute prediction error of 0.400, less than $0.15 \times$ training set range ($=0.15 \times 4.271 = 0.641$), which indicates the model has good prediction [24]. In addition, the SVM model has determination coefficients of 0.839 for the training set and 0.747 for test set, which are higher than the results ($R^2 = 0.598$ and 0.659, respectively) of the MLR model in this study. Therefore, there are nonlinear relationships between the six molecular descriptors used and inhibitory constants pKi.

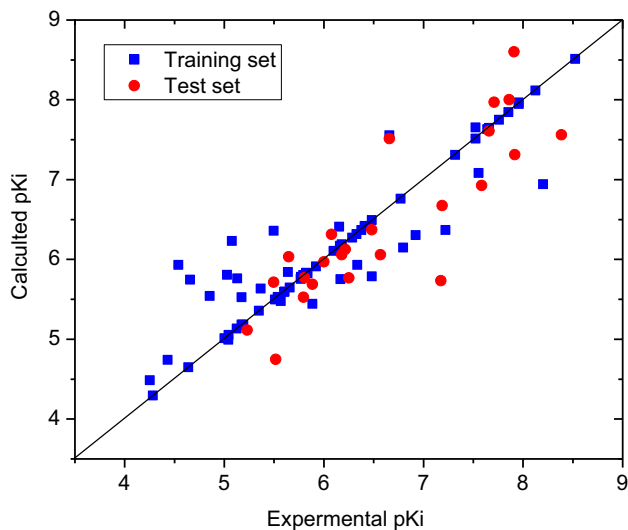
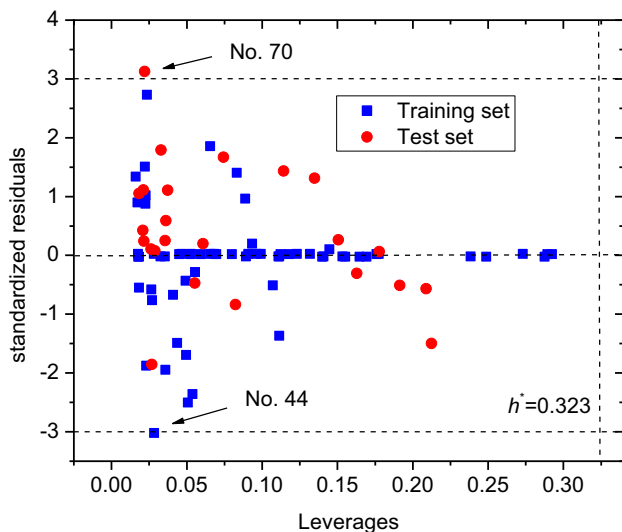
To make a statistical performance comparison between MLR and SVM algorithm, we develop another SVM model II for pKi of 62 compounds, which have been studied by Masand et al. with MLR [7]. The same six molecular descriptors were used to divide the 62 samples into a training set (50 compounds) and a test set (12 compounds) by employing the Kennard-Stone algorithm. The same procedures referred to were used to develop SVM model of pKi. Finally, the SVM model II ($C = 337.692$ and $\gamma = 0.0092516$) was obtained, which produces determination coefficient R^2 and rms error of 0.817 and 0.428 for the training set; and of 0.789 and 0.430 for the test set. Obviously, the rms errors from the SVM model II in this work are less than (0.435 and 0.525, respectively) in reference [7], although the data sets of samples and descriptors

Table 1 The definitions of molecular descriptors used.

Descriptor	Definition	Block
MATS6m	Moran autocorrelation of lag 6 weighted by mass	2D autocorrelations
MATS1s	Moran autocorrelation of lag 1 weighted by intrinsic state	2D autocorrelations
MATS3s	Moran autocorrelation of lag 3 weighted by intrinsic state	2D autocorrelations
nArNR2	Number of tertiary amines (aromatic)	Functional group counts
C-028	R--CR--X	Atom-centered fragments
F10[N-O]	Frequency of N-O at topological distance 10	2D Atom Pairs

Table 2 Characteristics of molecular descriptors in MLR model.

Descriptor	Coefficients	Std. Error	<i>t</i> -test	Sig.	VIF
Constant	5.968	0.264	22.647	0.000	/
MATS6m/MW	-0.010	0.003	-3.579	0.001	1.637
MATS1s/MW	-0.046	0.013	-3.479	0.001	1.768
MATS3s/MW	0.034	0.005	6.566	0.000	1.439
nArNR2/MW	0.002	0.000	4.403	0.000	1.058
C-028/MW	0.001	0.000	3.765	0.000	1.122
F10[N-O]/MW	-0.001	0.000	-6.348	0.000	1.211

**Fig. 1** Plot of experimental versus calculated pKi with SVM model.**Fig. 2** Williams plot with the warning leverage $h^* = 0.323$.

are same. Therefore, applying SVM algorithm to develop models for pKi is successful.

3.3. Applicability domain

Fig. 2 shows the Williams plot of the standardized residuals (σ) against the leverages (h) of both the training and test sets of the optimal SVM model. The predictions may be considered reliable if the samples fall into the domain of Williams plot [19]. In Fig. 2, we can observe that there are only two samples (Nos. 44 and 70) with $|\sigma| > 3$. Their molecular structures are shown in Fig. 3. The two samples have dissimilar structures with other compounds in the training set. Of course, the possibility exists that they have large experimental errors. In addition, we calculated the warning leverage h^* ($=0.323 = 3 \times (p + 1)/n = 3 \times (6 + 1)/65$, here p and n are, respectively, the numbers of independent variables and samples in training set). All the sample points have leverage values less than the warning leverage h^* , indicating the prediction results (except Nos. 44 and 70) from the optimal SVM model in this work reliable.

4. Conclusions

Six molecular descriptors, MATS6m/MW, MATS1s/MW, MATS3s/MW, nArNR2/MW, C-028/MW, and F10[N-O]/MW, were successfully used for developing a 2D-QSAR model of inhibitory constants pKi of 89 compounds against SARS-CoV 3CLpro enzyme, although many factors influence the inhibitory activity. The optimal SVM ($C = 90.2339$ and $\gamma = 1.19826 \times 10^{-5}$) model was based on 65 compounds in the training set and evaluated with 24 compounds in the test set. The optimal SVM model possesses good statistical performance compared with other models reported. There were non-linear relationships between the six descriptors used and inhibitory constants pKi of compounds against SARS-CoV 3CLpro enzyme. It was reasonable applying the SVM algorithm to establish this nonlinear model. The model developed in this work could be beneficial for search of novel 3CLpro enzyme inhibitors against SARS-CoV.

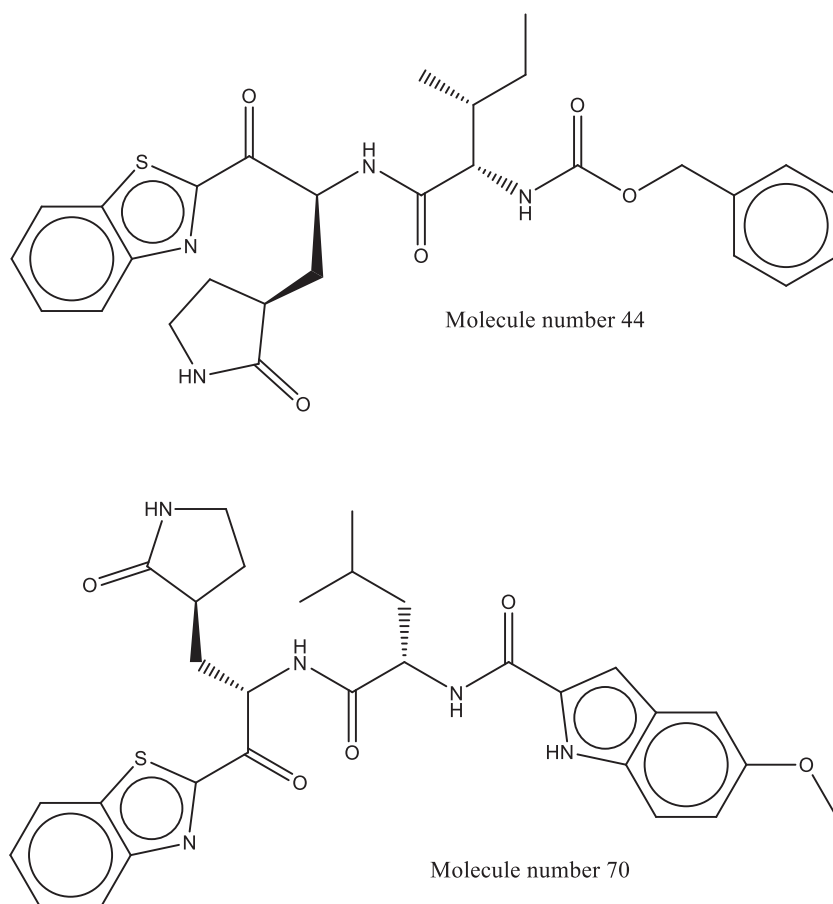


Fig. 3 Outliers based on the predictions of the optimal SVM model.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Open Project Program of Hunan Provincial Key Laboratory of Environmental Catalysis & Waste Regeneration (Hunan Institute of Engineering) (No. 2018KF11).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jscs.2021.101262>.

References

- [1] D. Kumar, R. Malviya, P.K. Sharma, Corona virus: a review of COVID-19 history and origin, *EJMO* 4 (2020) 8–25.
- [2] S.H. Nile, A. Nile, J. Qiu, L. Li, X. Jia, G. Kai, COVID-19: Pathogenesis, cytokine storm and therapeutic potential of interferons, *Cytokine Growth Fact. Rev.* 53 (2020) 66–70.
- [3] B. Shanmugaraj, K. Siri wattananon, K. Wangkanont, W. Phoolcharoen, Perspectives on monoclonal antibody therapy as potential therapeutic intervention for Coronavirus disease-19 (COVID-19), *Asian Pac. J. Allergy Immunol.* 38 (2020) 10–18.
- [4] C. Liu, Q. Zhou, Y. Li, L.V. Garner, S.P. Watkins, L.J. Carter, J. Smoot, A.C. Gregg, A.D. Daniels, S. Jerve, D. Albaiu, Research and development on therapeutic agents and vaccines for COVID-19 and related human coronavirus diseases, *ACS Central Sci.* 6 (2020) 315–331.
- [5] C. Wu, Y. Liu, Y. Yang, P. Zhang, W. Zhong, Y. Wang, Q. Wang, Y. Xu, M. Li, X. Li, M. Zheng, L. Chen, H. Li, Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods, *Acta Pharm. Sin. B.* 10 (2020) 766–788.
- [6] Y. Zhou, Y. Hou, J. Shen, Y. Huang, W. Martin, F. Cheng, Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2, *Cell Discov.* 6 (2020) 14.
- [7] V. Masand, A. Gandhi, V. Rastija, M.K. Patil, Structure Features of Peptide-Type SARS-CoV Main Protease Inhibitors: Quantitative Structure Activity Relationship Study, *ChemRxiv*, 2020, Preprint <https://doi.org/10.26434/chemrxiv.12196683.v1>.
- [8] V. Kumar, K. Roy, Development of a simple, interpretable and easily transferable QSAR model for quick screening antiviral databases in search of novel 3CLike protease (3CLpro) enzyme inhibitors against SARS-CoV diseases, *SAR QSAR Environ. Res.* 31 (2020) 511–526.
- [9] V.H. Masand, D.T. Mahajan, A.M. Alafeefy, S.N. Bukhari, N.N. Elsayed, Optimization of antiproliferative activity of substituted phenyl 4-(2-oxoimidazolidin-1-yl) benzenesulfonates: QSAR and CoMFA analyses, *Eur. J. pharm. Sci.* 77 (2015) 230–237.
- [10] V.H. Masand, D.T. Mahajan, P. Gramatica, J. Barlow, Tautomerism and multiple modelling enhance the efficacy of

- QSAR: antimalarial activity of phosphoramidate and phosphorothioamidate analogues of amiprofos methyl, *Med. Chem. Res.* 23 (2014) 4825–4835.
- [11] M.K. Gilson, T. Liu, M.G. Baitaluk, G. Nicola, L. Hwang, J. Chong, BindingDB: A public database for medicinal chemistry, computational chemistry and systems pharmacology, *Nucleic Acids Res.* 19 (2015) 1045–1053.
- [12] S. Konno, P. Thanigaimalai, T. Yamamoto, K. Nakada, R. Kakiuchi, K. Takayama, Y. Yamazaki, F. Yakushiji, K. Akaji, Y. Kiso, Y. Kawasaki, S.-E. Chen, E. Freire, Y. Hayashi, Design and synthesis of new tripeptide-type SARS-CoV 3CL protease inhibitors containing an electrophilic arylketone moiety, *Bioorgan. Med. Chem.* 21 (2013) 412–424.
- [13] P. Thanigaimalai, S. Konno, T. Yamamoto, Y. Koiwai, A. Taguchi, K. Takayama, F. Yakushiji, K. Akaji, S.-E. Chen, A. Naser-Tavakolian, A. Schön, E. Freire, Y. Hayashi, Development of potent dipeptide-type SARS-CoV 3CL protease inhibitors with novel P3 scaffolds: Design, synthesis, biological evaluation, and docking studies, *Eur. J. Med. Chem.* 68 (2013) 372–384.
- [14] P. Thanigaimalai, S. Konno, T. Yamamoto, Y. Koiwai, A. Taguchi, K. Takayama, F. Yakushiji, K. Akaji, Y. Kiso, Y. Kawasaki, S.-E. Chen, A. Naser-Tavakolian, A. Schön, E. Freire, Y. Hayashi, Design, synthesis, and biological evaluation of novel dipeptide-type SARS-CoV 3CL protease inhibitors: Structure–activity relationship study, *Eur. J. Med. Chem.* 65 (2013) 436–447.
- [15] Talet srl, DRAGON (Software for Molecular Descriptor Calculation) Version 6.0, 2012, <http://www.talet.mi.it/>.
- [16] X.L. Yu, R.M. Zhan, J.Y. Deng, X.W. Huang, Prediction of the maximum nonseizure load of lubricant additives, *J. Theor. Comput. Chem.* 16 (2017) 1750014.
- [17] X.L. Yu, L. Huang, Prediction of the onset temperature of decomposition of lubricant additives, *J. Therm. Anal. Calorim.* 130 (2017) 943–947.
- [18] C.C. Chang, C.J. Lin, LIBSVM: A library for support vector machines, *Acm. T. Intel. Syst. Tec.* 2 (2011) 27.
- [19] X. Yu, Quantitative structure-toxicity relationships of organic chemicals against *Pseudokirchneriella subcapitata*, *Aquat. Toxicol.* 224 (2020) 105496.
- [20] K. Khanand, K. Roy, Ecotoxicological QSAR modelling of organic chemicals against *Pseudokirchneriella subcapitata* using consensus predictions approach, *SAR QSAR Environ. Res.* 30 (2019) 665–681.
- [21] X. Yu, B. Yi, W. Yu, X. Wang, DFT-based quantum theory QSPR studies of molar heat capacity and molar polarization of vinyl polymers, *Chem. Pap.* 62 (2008) 623–629.
- [22] V.H. Masand, N.N.E. El-Sayed, M.U. Bambole, V.R. Patil, S. D. Thakur, Multiple quantitative structure-activity relationships (QSARs) analysis for orally active trypanocidal N-myristoyltransferase inhibitors, *J. Mol. Struct.* (2019, 1175,) 481–487.
- [23] V.H. Masand, N.N.E. El-Sayed, D.T. Mahajan, V. Rastija, QSAR analysis for 6-arylpyrazine-2-carboxamides as *Trypanosoma brucei* inhibitors, *SAR QSAR Environ. Res.* 28 (2017) 165–177.
- [24] K. Roy, P. Ambure, S. Kar, How precise are our quantitative structure–activity relationship derived predictions for new query chemicals, *ACS Omega* 3 (2018) 11392–11406.