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.

Chinese Journal of Analytical Chemistry 49 (2021) 63–73

Contents lists available at ScienceDirect

Chinese Journal of Analytical Chemistry

journal homepage: www.elsevier.com/locate/cjac

Molecular design, molecular docking and ADMET study of cyclic sulfonamide derivatives as SARS-CoV-2 inhibitors

Jian-Bo TONG[a,b,*], Xing ZHANG[a,b], Ding LUO[a,b], Shuai BIAN[a,b]

[a] College of Chemistry and Chemical Engineering, Shaanxi University of Science and Technology, Xi'an 710021, PR China
[b] Shaanxi Key Laboratory of Chemical Additives for Industry, Xi'an 710021, PR China

ARTICLE INFO

Keywords:
Cyclic Sulfonamide derivatives
SARS-CoV-2
Topomer CoMFA
HQSAR
ADMET

ABSTRACT

Severe acute respiratory syndrome coronavirus type 2 (SARS-CoV-2) continues to spread globally with more than 172 million confirmed cases and 3.57 million deaths. Cyclic sulfonamide derivative is identified as a successful compound and showed anti-SARS-CoV-2 activity. In this study, the structure and activity relationships of 35 cyclic sulfonamide compound inhibitors are investigated by using three-dimensional quantitative structure-activity relationship (3*D*-QSAR) and holographic quantitative structure-activity relationship (HQSAR). Two models with good statistical parameters and reliable predictive ability are obtained from the same training set, including Topomer CoMFA (q^2 = 0.623,r^2 = 0.938,r^2_{pred} = 0.893) model and HQSAR (q^2 = 0.704,r^2 = 0.958,r^2_{pred} = 0.779) model. The established models not only have good stability, but also show good external prediction ability for the test set. The contour and color code maps of the models provide a lot of useful information for determining the structural requirements which might affect the activity; this information paves the way for the design of four novel cyclic sulfonamide compounds, and predictes their pIC_{50} values. We explore the interaction between the newly designed molecule and SARS-CoV-2 3CLpro by molecular docking. The docking results show that GLU166, GLN192, ALA194, and VAL186 may be the potential active residues of the SARS-CoV-2 inhibitor evaluated in this study. Finally, the oral bioavailability and toxicity of the newly designed cyclic sulfonamide compounds are evaluated and the results show that the four newly designed cyclic sulfonamide compounds have major ADMET properties and can be used as reliable inhibitors against COVID-19. These results may provide useful insights for the design of effective SARS-CoV-2 inhibitors.

1. Introduction

Since the first case of pneumonia was reported in Wuhan, China in December 2019 [1], coronavirus disease 2019(COVID-19) has spread all over the world, causing serious negative impacts on the health of people in all countries. COVID-19 is lethal and highly infectious, and the international committee on taxonomy of viruses (ICTV) has named it severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2). As one of the deadliest viruses in the world, the virus has become an ongoing medical challenge for the world [2]. The most commonly used therapeutic drugs in clinical trials of antiviral research include remdesivir, ribavirin, favipiravir, etc. The U.S. food and drug administration (FDA) approved the emergency use of remdesivir in hospitalized patients with serious illnesses. Although this antiviral drug reduced the recovery time of surviving patients, it did not improve overall survival. Therefore, it is urgent to find new drugs that are more reliable and effective without any harmful side effects.

The genetic material of coronavirus is single-stranded RNA (diameter 65 ~ 125 nm, nucleic acid length is 2 ~ 32 kbs), and the genetic similarity with human SARS-CoV is 79 % [3], one third of the genome coding structure proteins (SPs), the remaining two-thirds of the genome encodes nonstructural proteins (nSPs). The main structural proteins include spike protein (S), envelope protein (E), membrane protein (M) and nucleocapsid protein (N) [4]. The spike protein has a coronal structure and consists of three identical chains, each of which has two subunits, S1 and S2 [5]. The n-terminus of the S1 subunit immediately follows the receptor-binding domain (RBD) region. The S2 subunit is responsible for the membrane fusion process. During virus infection, the target cell protease activates S protein by splitting S protein into S1 and S2 subunits, which is necessary for the activation of the membrane fusion domain after the virus enters the target cell and plays an important role in entering the host cell. Coronavirus produces a polypeptide that is hydrolyzed by 3-chymotrypsin-like protease (3CLpro) during genome transcription. 3CLpro cuts multiple proteins at 11 different sites to generate various

* Corresponding author at: College of Chemistry and Chemical Engineering, Shaanxi University of Science and Technology, Xi'an 710021, PR China.
E-mail address: jianbotong@aliyun.com (J.-B. TONG).

https://doi.org/10.1016/j.cjac.2021.09.006
Received 3 June 2021; Received in revised form 8 September 2021; Accepted 22 September 2021
Available online 29 September 2021

non-structural proteins which are important for virus replication. This major protease binds viral particles to the capsid protein shell and prevents the increase of viral load in the host cell, which is critical for the viral life cycle, making it as an attractive target for anti-SARS-CoV-2 inhibitors [6]. Therefore, 3CLPro is considered as a prime target for the treatment of SARS-CoV-2 infection.

As an important class of compounds, sulfonamides have a wide range of applications in medicine and pesticides. There are many sulfonamides in the market for the treatment of diseases with different properties, because they can identify various protein targets. Sulfonamide derivatives are an important part of many biologically active compounds and drug molecules, such as anti-bacterial [7], anti-cancer [8], anti-inflammatory [9], anti-tumor [10] and anti-malaria [11]. cyclic sulfonamide derivatives are known to have various pharmacological activities, such as analgesia [12], anti-inflammatory [13] and anti-diabetic [14]. Recently, there has been much increased interest in cyclic sulfonamide derivatives as they show potential inhibition of SARS CoV-2 3CLpro, and in this study we focuse on cyclic sulfonamide derivatives as inhibitors of SARS-CoV-2.

The establishment of a quantitative structure-activity relationship (QSAR) model can guide the modification of compound structures, design new and more active compounds and predict their activity. Commonly used QSAR models include 2*D*-QSAR and 3*D*-QSAR [15]. 3*D*-QSAR correlates biological properties with structural descriptors and is used to predict the activity value of non-synthetic molecules, which is an important method for molecular modeling. The techniques most commonly used in 3*D*-QSAR are comparative molecular field analysis (CoMFA) and comparative molecular similarity index analysis (CoMSIA) [16]. However, CoMFA has many limitations, which can be limited if the molecular structure is not three-dimensional in the database. Topomer comparative molecular field analysis (Topomer CoMFA) is the second-generation CoMFA method that overcomes many limitations of CoMFA and can predict the bioactivity of compounds in just a few minutes, which makes it more repeatable. 2*D*-QSAR is a method to quantitatively describe the relationship between the physicochemical properties and other measurable properties of a compound structure and its activity through a linear model or a nonlinear model [17]. HQSAR is a relatively new 2*D*-QSAR method, which realizes the need of molecular arrangement and conformation specification by converting the chemical representation of the molecule into its corresponding molecular hologram.

In this study, 35 SARS-CoV-2 inhibitors are studied using two methods, 2*D*-HQSAR and 3*D*-Topomer CoMFA, and key structural factors affecting the inhibitory activity are identified. Through molecular docking, the binding mechanism of the SARS-CoV-2 3CLpro and the cyclic sulfonamide compound is investigated. Besides some key amino acid residues at the active site of the 7JYC protein are identified. Finally, we utilize ADMET (absorption, distribution, metabolism, excretion, and toxicity) properties to evaluate the oral bioavailability of these novel cyclic sulfonamides and measure their toxicity.

2. Materials and methods

2.1. Computer simulations

The research route used in this research is shown in Fig. 1. The division of data sets, the construction of QSAR models, the design of new SARS-CoV-2 inhibitors through virtual screening, molecular docking and ADMET prediction are the main components of this research.

2.2. Data sources

Bioactivity and molecular structure of a series of cyclic sulfonamide compounds as inhibitors of SARS-CoV-2 are obtained by literature search and used as a data set for molecular modeling [18]. The chemical structure and activity data are shown in Table S1. Among the 35 inhibitors, the IC_{50} values range from 0.88 to 25 μM. Equation (1) is used to convert them into pIC_{50} values to provide a larger value as the dependent variable for the model construction, where the pIC_{50} values are obtained in the range of 4.60 ~ 6.06. Fig. 2 shows the distribution of pIC_{50} values for all inhibitors and indicates that the diversity of activities in the data set is sufficient to construct a stable QSAR model. According to the pIC_{50} value and molecular structure, 35 inhibitors are divided into two groups: 23 compounds are assigned to the training set for model establishment, and 12 compounds are allocated to the test set for model verification.

$$pIC_{50} = -\log\left(IC_{50}\right) \quad (1)$$

2.3. Establishment of 3D-QSAR model and HQSAR model

2.3.1. 3D-QSAR model

The two-dimensional structure of the SARS-CoV-2 inhibitor is drawn in Chem Draw Ultra v8.0.3 (PerkinElmer, Waltham, MA, USA) and converts into a 3*D* structure using SYBYL-X 2.0 (Tripos Software, Saint Louis, MO, USA) [19]. Tripos force field and Powell conjugate gradient algorithm with convergence criterion of 0.005 kcal/mol ∀ are adopted, and the maximum iteration set is set as 1000 [20]. The Gasteiger-Hückel method is used to calculate partial atomic charges of compounds under electrostatic interaction, and the remaining parameters are Topomer CoMFA default value to obtain the low energy conformation of each molecule (Pharmacodynamic conformation).

Structure alignment is one of the most important input variables in 3*D*-QSAR analysis, and the prediction accuracy of the model depends on the contour map of molecular structure comparison. Topomer CoMFA can simultaneously describe the changes of the electrostatic field and the steric field around the molecule, revealing the molecular structure information that affects the biological activity more comprehensively, and determining the group properties that affect the molecular activity. In this experiment, the molecule No. 33 with the highest activity value in the training set is used as the cutting template. The splitting mode of R group is very important for the generation of Topomer CoMFA model. Therefore, we compare two different split modes for modeling research (Fig. 3). The remaining few molecules that cannot be automatically recognized will be manually broken until all the molecules have been cut, and the resulting structural fragments will be used for subsequent analysis of model results and virtual screening research.

2.3.2. HQSAR model

The contribution map obtained by the HQSAR method can clearly show where the inhibitor molecule needs to be optimized; it can get the advantages and disadvantages of the atom or group on the molecular activity, and find a suitable replacement group for molecular design. Therefore, combining the Topomer CoMFA model with the HQSAR model and applying it for molecular modification can not only accurately locate the active site, but also determine the groups and atoms that affect the corresponding properties of the molecular activity.

HQSAR is a new QSAR technology based on fragments and has high predictive ability. There is no need to determine the 3*D* structure, molecular comparison and assumed conformation [21], which has irreplaceable advantages in molecular design. All possible molecular fragments (such as linearity, branching, overlap, hybridization, and chirality) can be encoded in molecular holograms. Each compound is divided into structural fragments defined by fragment discrimination (FD) parameters: atomic number (A) can distinguish atom types, bond type (B) can distinguish chemical bonds formed between atoms, and atomic connections (C) can distinguish the hybridization state of the atoms inside the fragment, hydrogen (H), chirality (Ch) can distinguish the chirality of the atom in the fragment and the stereochemical information of the chemical bond, and the donor/acceptor (DA) distinguishes the hydrogen bond donor or acceptor of the fragment. The cyclic redundancy check (CRC) algorithm is used to assign a specific positive integer to each segment, ranging from 0 to 231. Each of these integers is in correspondence

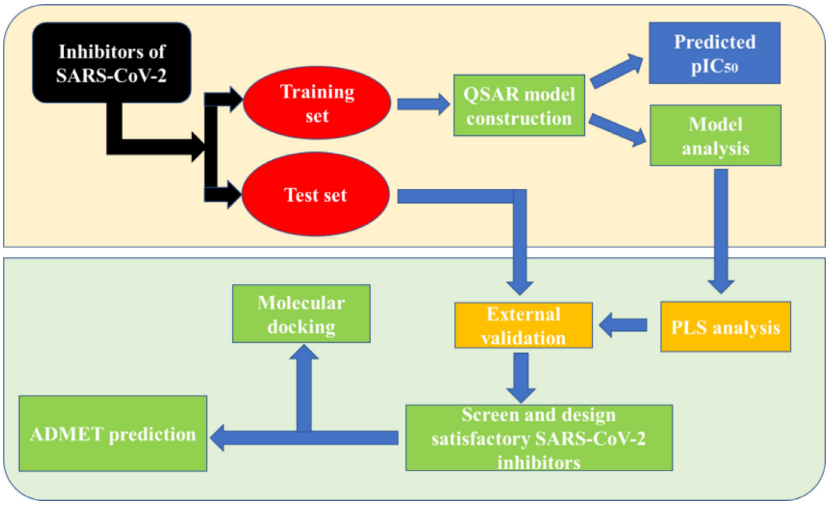

Fig. 1. Flowchart for computational drug design.

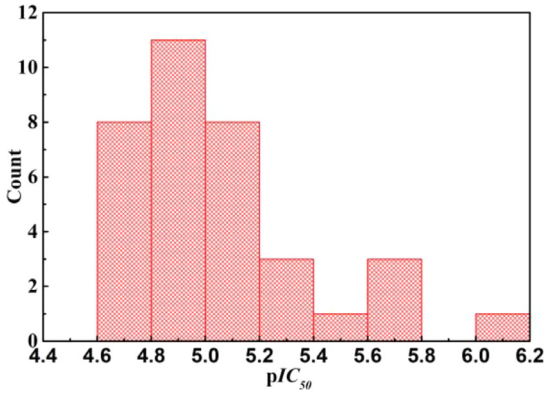

Fig. 2. Activity distribution range of pIC_{50}.

with a bin in an integer array of hologram length (HL, ranging from 53 to 401) and the bin occupancies of the molecular hologram are structural descriptors [22]. In the HQSAR method, there is a partial least squares (PLS) relationship between these descriptors and attribute values. Many parameters related to hologram generation, such as HL, fragment size (FS) and FD, will affect the quality of the HQSAR model [23].

The fragment parameters determine the topological information mapped in the molecular hologram, and the model can be optimized by changing the fragment parameters and fragment size. In the process of building the HQSAR model, we selected all 12 holographic lengths (HL) (53, 59, 61, 71, 83, 97, 151, 199, 257, 307, 353 and 401). The default FS (4-7) and different combinations of FD (atoms, bonds, chirality, connection, hydrogen bond donor and hydrogen bond acceptor) to generate the initial model. 25 HQSAR models were analyzed to study the influence of different FD combinations on the model, so as to preliminarily screen out the better model. Then, on the basis of the selected better model, different fragments lengths were selected to analyze the influence of different fragments lengths on the HQSAR analysis results, so as to obtain the optimal HQSAR model.

2.4. Partial least square (PLS) analysis

In 3*D*-QSAR research, the PLS method [24] is an extension of multiple regression analysis to analyze the relationship between quantitative descriptors and biological activity in the model. The established model descriptors (electrostatic field and stereo field parameters) are used as independent variables, and pIC_{50} is used as the dependent variable for PLS regression analysis. The leave-one-out method (LOO) cross-validation is one of the simplest methods for internal model verification [25]. PLS method is used for model fitting, and the LOO method is used to cross-validate and evaluate the predictive ability of the model's internal verification, and the optimal group score (NOC) is determined. At the same time, the cross-validation correlation coefficient (q^2), the standard error of estimation (SEE), the non-cross-validation correlation coefficient (r^2) and the Fischer ratio value (F) are calculated to verify the stability of the constructed model. Among them, r^2 and SEE

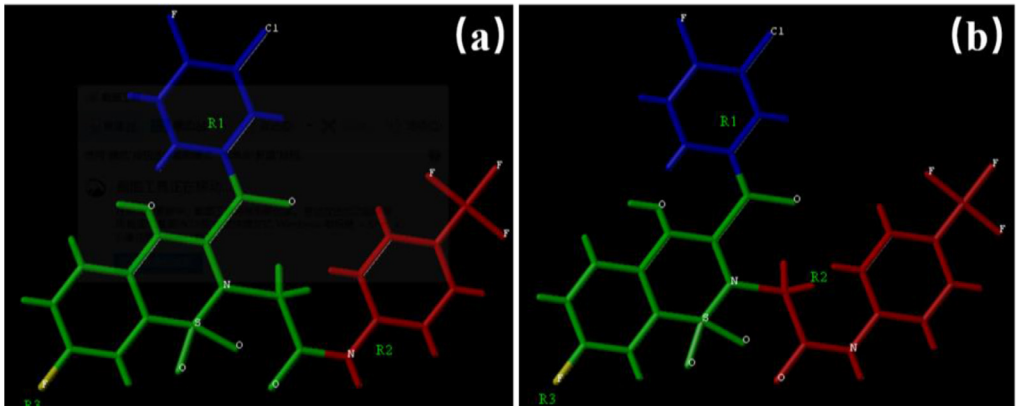

Fig. 3. Cutting method of model 1 (a) and model 2(b). Blue, red and yellow represents the R_1, R_2, R_3 groups, respectively. green represents the common skeleton.

are automatically generated by the system. The larger the r^2 and F values are, the smaller the SEE value is, which proves that the model's fitting ability is stronger, q^2: <0 (the model predictive ability is poor), 0.4 ~ 0.5 (the model can be considered), > 0.5 (a statistically significant prediction model); high q^2 and r^2 (q^2 >0.5, r^2>0.6) value can prove that the established 3*D*-QSAR model and HQSAR model have high predictive ability [26]. The q^2, r^2, SEE, and F are calculated for the data set as equations (2)-(5):

$$q^2 = 1 - \frac{\sum (Y_{pred} - Y_{exp})^2}{\sum (Y_{exp} - Y_{mean})^2} \tag{2}$$

$$r^2 = \sqrt{1 - \frac{\sum (Y_{cal} - Y_{exp})^2}{\sum (Y_{exp} - \bar{Y}_{exp})^2}} \tag{3}$$

$$SEE = \sqrt{\frac{\sum (Y_{cal} - Y_{exp})^2}{n - k - 1}} \tag{4}$$

$$F = \sqrt{\frac{r^2(n-k-1)}{k(1-r)^2}} \tag{5}$$

Where Y_{exp} is the experimental value of biological activity; Y_{cal} is the simulated fitting value of biological activity; n is the number of samples; k is the number of parameters used in modeling; Y_{pred} is the predicted activity of the test set compounds; Y_{mean} is the calculated average activity of the training set compounds.

2.5. External validation

Studies have shown that there is no correlation between internal prediction ability (q^2) and external prediction ability (r^2_{pred}). The q^2 obtained by the *LOO* method cannot be used to evaluate the external predictive ability of the model [27]. The established model has good internal prediction ability, but the external prediction ability may be very low, and vice versa. Therefore, the QSAR model must pass effective external validation to ensure the predictive ability of the model for external samples. International journals such as Food Chem, Chem Eng J, Eur J Med Chem and J Chem Inf Model explicitly state that each QSAR/QSPR paper must be externally verified. The best method for external validation of the model is to use a representative and large enough test set, and the predicted value of the test set can be compared with the experimental value. The prediction correlation coefficient r^2_{pred} (r^2_{pred} >0.6) [28] based on the test set is calculated according to equation (6):

$$r^2_{pred} = \frac{SD - PRESS}{SD} = 1 - \frac{\sum_{n=1}^{test} (y_i - \hat{y}_i)^2}{\sum_{n=1}^{test} (y_i - \bar{y})^2} \tag{6}$$

Where $\hat{y}_i$: test set activity prediction value, y_i: test set activity experimental value, $\bar{y}$: average value of training set experimental values, $\hat{y}$: average value of training set prediction values.

Using test sets and classic verification standards to test the external predictive ability of the developed QSAR model: the Golbraikh–Tropsha method [29]. The usual conditions of the 3*D*-QSAR models and HQSAR models with more reliable external verification capabilities must meet are: (1) q^2 >0.5, (2) r^2 >0.6, (3) $(r^2 - r_0^2)/r^2 < 0.1$ and $0.85 \le k \le 1.15$ or $(r^2 - r_0^{2'})/r^2 < 0.1$ and $0.85 \le k' \le 1.15$ and (4) $|r^2 - r_0^{2'}| < 0.1$.

$$r_0^2 = 1 - \frac{\sum (y_i - k \times \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2} \tag{7}$$

$$r_0^{2'} = 1 - \frac{\sum (\hat{y}_i - k' \times y_i)^2}{\sum (\hat{y}_i - \hat{y})^2} \tag{8}$$

$$k = \frac{\sum (y_i \times \hat{y}_i)}{\sum (\hat{y}_i)^2} \tag{9}$$

$$k' = \frac{\sum (y_i \times \hat{y}_i)}{\sum (y_i)^2} \tag{10}$$

Where r_0^2 and $r_0^{2'}$ are squared correlation coefficients of determination for regression lines through the origin between predicted (y) and observed (x) activities and the values of k and k' are the slopes of their models.

In addition, the rigorous and powerful statistical indicators proposed by Roy on the basis of the Golbraikh-Tropsha method are also applied: r_m^2.

$$r_m^2 = r^2\left(1 - \sqrt{\left|r^2 - r_0^2\right|}\right) \tag{11}$$

$$r_m^{2'} = r^2\left(1 - \sqrt{\left|r^2 - r_0^{2'}\right|}\right) \tag{12}$$

$$\Delta r_m^2 = \left|r_m^2 - r_m^{2'}\right| \tag{13}$$

For an acceptable model, r_m^2 value more than 0.5 (r_m^2>0.5) and $\Delta r_m^2 < 0.2$ show good external predictability of the models. In addition, other kinds of methods, Q^2_{F1}, Q^2_{F2} , RMSE -the root mean square error of training set and test set, CCC-the concordance correlation coefcient (CCC >0.85) [30], MAE -the mean absolute error, and RSS- the residual sum of squares, which is a new method created by Roy, are also calculated within this tool. The RMSE, MAE, RSS, and CCC are calculated for the data set as equations (14)-(19):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}} \tag{14}$$

$$MAE = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n} \tag{15}$$

$$RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{16}$$

$$CCC = \frac{2\sum_{i=1}^{n} (y_i - \bar{y})(\hat{y}_i - \hat{y})}{\sum_{i=1}^{n} (y_i - \bar{y})^2 + \sum_{i=1}^{n} (\hat{y}_i - \hat{y})^2 + n(\bar{y} - \hat{y})^2} \tag{17}$$

$$Q^2_{F1} = 1 - \frac{\sum_{n=1}^{test} (\hat{y}_i - y_i)^2}{\sum_{n=1}^{test} (y_i - \bar{y})^2} \tag{18}$$

$$Q^2_{F2} = 1 - \frac{\sum_{n=1}^{test} (\hat{y}_i - y_i)^2}{\sum_{n=1}^{test} (y_i - \bar{y}_{Test})^2} \tag{19}$$

2.6. Virtual screening of new novel SARS-CoV-2 inhibitors

The 3*D*-QSAR model of 35 cyclic sulfonamide compounds inhibitors is established by using Topomer CoMFA based on R group search technology. The molecules in the database are segmented into fragments, and the fragments are compared with the substituents in the data set, and the similarity degree of compound structure is evaluated by scoring function [31], so as to perform virtual screening of similar structure for the molecular fragments in the database. Therefore, after the Topomer CoMFA modeling, the Topomer CoMFA module in SYBYL-X 2.0 is used for Topomer Search technology to find new molecular substituents, which can efficiently, quickly and more economically design a large number of new compounds with better activity.

In this study, by searching the compound database of ZINC (2015) [32] (a source of molecular structure fragments), the topomer technique is used to compare and find the molecular fragments with similarity. The Topomer Distance (TOPDIST) and the contribution value of substituents are integrated and the established Topomer CoMFA model scores these fragments and performs virtual screening on the cleaved fragments to

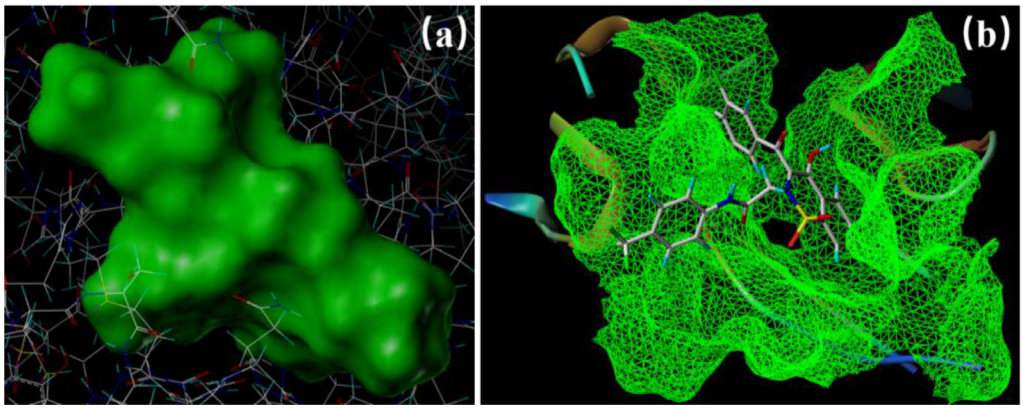

Fig. 4. (a): Prototype molecular generation diagram (Green area represents prototype molecule). (b): Compound 33 interacts with the active site of protein 7JYC.

obtain R_1, R_2 and R_3 substituents with higher contribution value. Then, SARS-CoV-2 inhibitor small molecules with better activity are obtained by splicing design.

2.7. Molecular docking study

Molecular docking is one of the most commonly used methods to study the mutual recognition process of geometric matching and energy matching in drug design. The principle of molecular docking is the "lock and key model" [33]. The lock is a macromolecular receptor with different structures, and the key is a small molecule ligand with a specific structure. When the macromolecular receptor and the small molecule ligand are complementary and matched in geometry, the electrostatic interactions, hydrogen bond interactions and hydrophobic interactions will occur. Then, in the process of binding, the conformation of the small molecule ligand and its surrounding amino acid conformation gradually change, adapt to each other and induce fit.

In order to exert its inhibitory activity against SARS-CoV-2, cyclic sulfonamide compounds need to have certain affinity with SARS-CoV-2 enzyme protein. After the two are sufficiently close to each other, they will combine with each other and interact with each other through appropriate conformational adjustment, finally forming a stable complex conformation [34]. Surflex-Dock takes polarity effect, hydrophobic effect and hydrogen bond effect into account to score the interaction between ligand and receptor, and the Total score is the dissociation constant (representing docking activity). We use SYBYL-X 2.0 (Surflex-Dock method) and Discovery Studio Visualization tool 2017 to study the molecular docking of the least active compound(2, 3, 7, 8, 25, 26, 27, 29) and the most active compound 33 with the 7JYC protein on the data set reported in the previous experimental studies to further analyze and verify the molecular structure of cyclic sulfonamide compounds [35]; and through the comparison of the two methods, the reason why compound 33 has a higher inhibitory activity against SARS-CoV-2 is explained. Finally, the four newly designed inhibitor molecules are docked to understand the antiviral mechanism of the designed compound.

The three-dimensional crystal structure of protease (7JYC) comes from the PDB database (http://www.rcsb.org/). Before molecular docking, the protein receptor molecules are pretreated, the required small molecule ligands are extracted from the macromolecular complexes, and the own ligands, metal ions, water molecules, and other residues and terminal residues of protein are removed. Polar hydrogen and point charges are added to expose the binding pocket (represented by the prototype molecule) [36]. The binding pocket is filled with hydrogen bond donors, hydrogen bond acceptors and hydrophobic site molecular probes. The interaction mode of the processed prototype small molecule and protein macromolecule is shown in Fig. 4(a). The crystal structure of the protein macromolecule is extracted from the eutectic ligand, and then the crystal structure is reconnected to the eutectic ligand by molecular docking technology. The original ligand is taken as a reference, as shown in Fig.S1, the conformation of the crystal structure and the conformation after ligand docking are almost overlapped, and the similarity is 0.751. The root mean square deviation (RMSD) is used to evaluate the quality of molecular docking. Generally speaking, when RMSD <2, molecular docking is considered to be successful. The RMSD of this study is 1.434, indicating that the docking method is reasonable and reliable, and protein 7JYC could be used for molecular docking of newly designed molecules.

2.8. Predicted pharmacokinetic and toxicity properties

In addition to forming a good interaction with the target, a good drug molecule should also have good pharmacokinetic characteristics and as few toxic and side effects as possible. Molecules with important biological activities and suitable pharmacokinetic properties are one of the most important issues in drug design [37]. Although many molecules have important biological activities, they cannot be used clinically due to poor absorption, distribution, metabolism and excretion (ADME) parameters or toxicity. In order to meet the requirements of drug design, ADMET prediction has become a hot research direction in the field of computer-aided drug molecular design in recent years [38], and great progress has been made. ADMET and toxicology screening systems can provide opportunities to predict in vivo performance in silicon. These data are very important for the availability of drugs and the design of new and more active molecules. Predetermining ADME parameters from molecular synthesis can significantly reduce failures due to inappropriate pharmacokinetic properties. Therefore, this study uses preADMET online server [39] to conduct ADMET evaluation of compounds to determine their various physicochemical properties, pharmacokinetic properties and molecular toxicological characteristics.

3. Results and discussion

3.1. 3D-QSAR and HQSAR results and analysis

3.1.1. Topomer CoMFA analysis

The template molecule (33) is specifically cut as shown in Fig.S2 and the cyclic sulfonamide derivatives inhibitor is cut into four parts. From Table S2, the number of principal components (N) of the two QSAR models are 4 and 2, respectively, r^2 is 0.938 and 0.837, both greater than 0.6; q^2 is 0.623 and 0.504, both greater than 0.5; r^2_{pred} are 0.893 and 0.770, respectively, both greater than 0.6. The results show that the Topomer CoMFA models constructed by the two cutting methods are both ideal and have good predictive ability and statistical parameters. According to the comprehensive analysis of Table S2, Model 1 not only has good external predictive ability, but also well retains the core skeleton of the inhibitor in the cutting method, which is conducive for

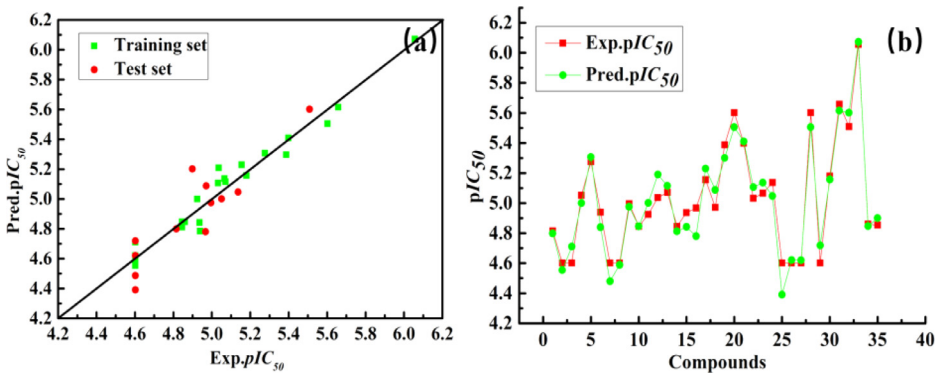

Fig. 5. Regression analysis graph (a) and line graph (b) of experimental activity and predicted activity of the data set of Topomer CoMFA model.

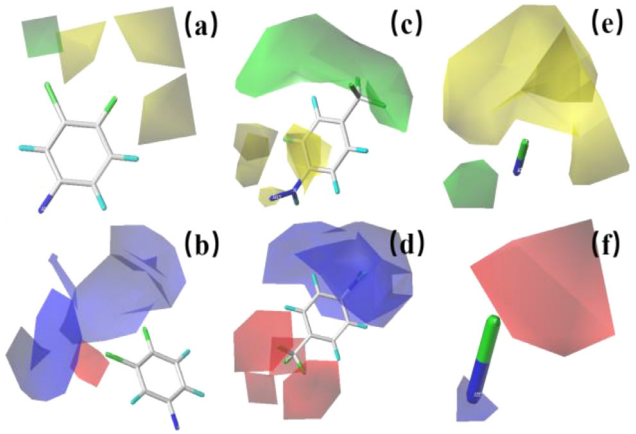

Fig. 6. 3*D* contour of the Topomer CoMFA model. (a) and (b): steric and electrostatic field maps of the R_1 fragment, respectively; (c) and (d): steric and electrostatic field maps of the R_2 fragment, respectively; (e) and (f): steric and electrostatic field maps of the R_3 fragment, respectively.

the selection of R groups. Therefore, model 1 is selected for subsequent research and molecular design.

The linear regression between the experimental and predicted values of the Topomer CoMFA model across the entire dataset is shown in Fig. 5(a) with all samples evenly distributed around 45° line. Fig. 5(b) shows that the predicted pIC_{50} values for these compounds are highly similar to the experimental values, indicating that the Topomer CoMFA model shows a satisfactory predictive ability for the low activity compounds (2, 3, 7, 8, 25, 26, 27, 29) and the highest activity compounds (33) in the whole dataset. These results confirm that Topomer CoMFA model has good predictive ability for cyclic sulfonamide derivatives. Therefore, the established 3*D*-QSAR model can be used for the screening and design of novel inhibitor molecules.

3.1.2. 3D contour maps analysis

The results of the Topomer CoMFA model are graphically interpreted using contour maps. Fig. 6 shows the calculated Topomer CoMFA electrostatic field and stereo field profile. In the stereo field map, the green part shows that increasing the volume of substituents is beneficial for the improvement of compound activity, while the yellow part shows the opposite. The presence of large yellow groups at the position 3 and 4 of R_1 group (-Cl, -F) could explain the higher activity of compound 19(pIC_{50} =5.387) with 3-Cl-Ph as R_1, while the lower activity of compound 18 (pIC_{50} =4.971) with 3-CN-Ph as R_1; the activity of compound 21(R_1=4-Cl-Ph, pIC_{50}=5.398) is higher than that of compound 22(R_1=4-CN-Ph, pIC_{50}=5.032) (compounds R_2 and R_3 have the same group). The R_3 sites of compound 34(pIC_{50}=4.860) and compound 35(pIC_{50} = 4.854) are replaced by template compound 33 (pIC_{50}= 6.056) with smaller substituents, and the activity is significantly improved, which is consistent with the contour map. The green (-CF_3 group in the C-4 position) and yellow (C-3, C-4 positions) polyhedrons in the R_2 group are distributed on both sides of the six-membered ring, and the green equipotential region is larger than the yellow equipotential region (Fig. 6(c)), which means that increasing the volume of this group will increase the activity of the compound. Comparing the chemical structures and pIC_{50} values of compound 5(R_2=4-CF_3-Ph, pIC_{50}= 5.276), compound 1(R_2= 3-F-Ph, pIC_{50}= 4.815) and compound 2(R_2= Ph, pIC_{50}=4.602), it's found that the R_2 group is consistent with the above conclusions.

In an electrostatic field, the red area indicates that introducing a negatively charged substituent or increasing the electronegativity of the group is beneficial for the improvement of compound activity, and the blue area indicates that introducing a positively charged substituent or reducing the electronegativity of the group is beneficial to increase the activity. As shown in Fig. 6(b), there are large blue outline near C-3(-Cl) and C-4(-F) positions on the benzene ring of R1 group. The electronegativity of -F group at C-3 position on the benzene ring of compounds 9 and 10 is less than that of -CN group, and the activity of compound 9(pIC_{50} =4.996) is greater than that of compound 10(pIC_{50} =4.845). For the R_2 group of the cyclic sulfonamide derivative, the 1,4- position of the benzene ring has the largest red and blue equipotential region. However, considering that the red equipotential region is closer to the benzene ring, we pay more attention to the influence of the negatively charged groups. Therefore, more consideration should be given on the introduction of electro negative groups to enhance compound activity in this model. The -CF_3 group near C-4 is surrounded by large red blocks, indicating that the bulkly and negatively charged group has a positive contribution for the activity, for example, compound 33 (pIC_{50} = 6.056) > compound 31 (pIC_{50} = 5.658), compound 4 (pIC_{50} = 5.051) > compound 3 (pIC_{50} = 4.602). As shown in Fig. 6(f), the red contour lines of the R_3 group indicate that it is advantageous to increase the electronegativity of the group here. Among the 35 compounds, compounds 31, 32, and 33 are compounds with fluorine atom of R_3, which have high inhibitory activity against SARS-CoV-2 (pIC_{50} value is 5.658, 5.509, 6.056, respectively). The activity of compound 33(pIC_{50} = 6.056, R_3=-F) is higher than that of compound 28 (pIC_{50} = 5.602, R_3=-H), and almost all compounds with negative R_3 groups show better inhibitory activity.

3.1.3. HQSAR analysis

The performance of the HQSAR model is affected by parameters such as HL (hologram length), FD (fragment discrimination type) and FS (fragment size), and these parameters need to be refined and optimized. We initially use the default FS (4-7), all HLs and different FD combinations to generate the model. Then selecting different FS to study its influence on the HQSAR analysis results and obtaining the optimal HQSAR model. The HQSAR model of 37 statistical parameters is shown in Table S3. The results show that the model generate when FD is "A + B + C + Ch" and FS is "4-7" is the best HQSAR model: 71 for hologram length and

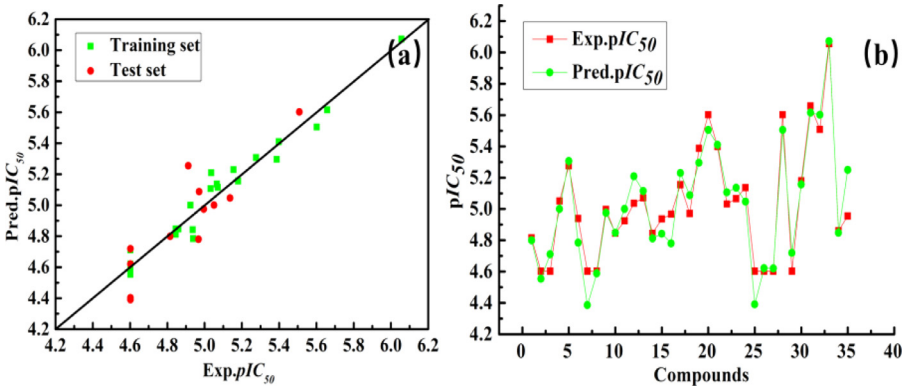

Fig. 7. Regression analysis graph (a) and line graph (b) of experimental activity and predicted activity of the data set of HQSAR model.

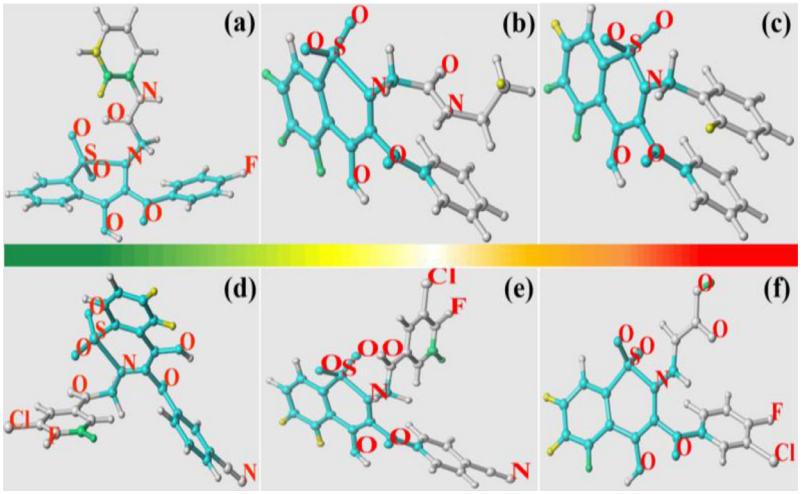

Fig. 8. HQSAR contribution maps of compound 3(a), 7(b), 25(c),26(d), 27(e) and 29(f). The red end of the spectrum (red, orange-red, and orange) reflects the negative contribution to the activity, the green end (yellow, blue and green) represents a positive effect, and the middle contribution is represented by white.

4–7 for fragment size, showing the highest q^2(0.704) and r^2(0.958) with 6 components and the standard error of 0.091.

Fig. 7(a) shows the pIC_{50} correlation diagram of the experimental and predicted values of the HQSAR model data set. All samples are evenly distributed near the Y=X line, showing a good linear relationship. Fig. 7(b) shows that the predicted pIC_{50} values of these compounds are almost in agreement with the experimental values. Both the low activity compounds (2,3,7,8,25,26,27,29) and the highest activity compounds (33) have good predictive ability, indicating that the HQSAR model has a satisfactory predictive ability. These results confirm that the HQSAR model has good predictive ability for cyclic sulfonamide derivatives. Therefore, the established HQSAR model can be used for the screening and design of novel inhibitor molecules.

3.1.4. Interpretation of HQSAR contribution map

HQSAR provides color-coded diagrams as direct evidence of the contribution of individual atoms to biological activity. In this study, the selected compound 33 with the best activity is taken as the representative for the color-coded HQSAR model analysis, and its single atomic contribution is shown in Fig. S3. Fig. 8 shows the atomic contribution diagrams (3, 7, 25, 26, 27, 29) of each series of representative molecules with lowest activity. It is worth noting that the common skeleton color of most compounds is blue, which is a common structural fragment of all molecules and acts in the same way for all inhibitors, indicating their important role in enhancing activity.

As shown in Fig. S3, most of the atoms of the R_1 group in the template molecule (33) are blue, indicating the importance of the R_1 group on the inhibitory activity. The -Cl atom and the -F atom at the C-3 and C-4 positions on the benzene ring are green, which are positive contributions to the activity of the cyclic sulfonamide derivative inhibitor, indicating that the introduction of the R_1 group here is beneficial for the improvement of the activity. The changes of C-2(-H),C-3(-Cl) and C-4(-CF_3) on the benzene ring of R_2 group are yellow, blue and green, indicating that the contribution of these fragments or atoms to the activity of the compound increase successively, and these groups should be retained when synthesized compounds that may have better biological activity. The F atom of the R_3 group appears green, which contributes positively to the improved activity of the compound; the C-1 and C-5 positions on the benzene ring of the R_2 group are white, indicating that the activity of the cyclic sulfonamide inhibitor is neutral or negative. It may be substituted by a substituent that can produce a stronger inhibitory effect, which is consistent with the results of Topomer CoMFA. In Fig. 8, almost the R_2 groups of the low-activity molecules appear white, the R_1 groups of the compounds 7, 25, 27 and 29 appear white, and the R_3 groups of the compounds 3 and 27 appear white, indicating that the inhibitor activity in these places is neutral or negative, and it may be substituted by substituents that can produce stronger inhibitory effects. This also explains the reason for the low activity of these molecules.

3.2. External validation analysis

The external prediction correlation coefficient, Golbraaikh-Tropsha method and r_m^2 (Roy) are used to verify the external prediction capabilities of the two models. The mathematical expressions of different statistics of HQSAR model and Tomoper CoMFA model are listed in Table S4. It can be seen from Table S4 that the established model satisfies both the Golbraaikh-Tropsha criterion and the Roy criterion. In addition, the calculated other indicators further show that our model has reliable predictive power and is acceptable. The QSAR models for the whole test set including 12 compounds give the r_{pred}^2 and r_m^2 values of 0.893

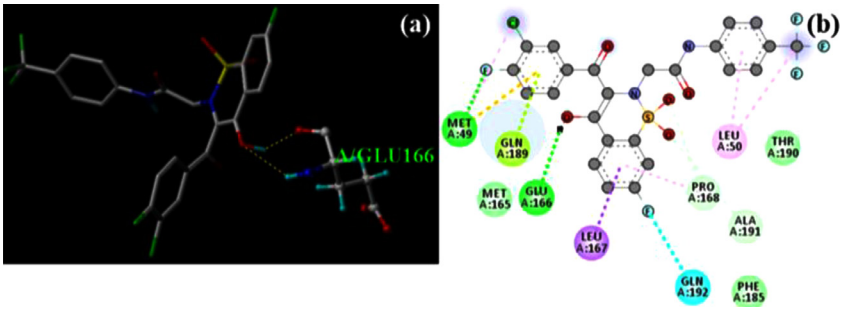

Fig. 9. Molecular docking results of template molecule. (a): 3D schematic representation of protein complex (The rod shape represents the small molecule ligand, the ball and stick shape represent the amino acid residue that forms the hydrogen bond, and the yellow dashed line represents hydrogen bonds); (b): 2D schematic representation of protein complex (The spheres represent the amino acid residues that form the forces, hydrogen bonds are shown as green dashed lines, and hydrophobic bonds are shown as pink dashed lines).

and 0.596 (Topomer CoMFA), and 0.779 and 0.504 (HQSAR), and high slope regression lines with k and k' values of 1.004 and 0.995 (Topomer CoMFA), 0.997 and 1.001 (HQSAR), r^2, r_0^2 and $r_0^{2'}$ values of 0.938, 0.834 and 0.878 (Topomer CoMFA), and 0.958 and 0.686 and 0.703 (HQSAR), respectively are obtained. Obviously, both models produce quite low RMSE, MAE and RSS values and high CCC values, the QSAR models yield RMSE, MAE, RSS and CCC values of 0.109, 0.090, 0.144 and 0.927 (Topomer CoMFA); 0.188, 0.136, 0.324 and 0.915 (HQSAR), indicating robust external predictive power of both models.

3.3. New compounds design and experimental activity

According to the constructed two models and related analysis results, the molecular structure is optimized using compound 33 as a template. The selected Topomer distance is close to 185, and the contribution value of each group exceeds the R group of the template molecule. As shown in Fig. S4, we select the R_3 group of the compound with the highest activity from the existing compounds, combine with two R_1 groups and two R_2 groups with high contribution values searched in the ZINC database, and designe four new compounds according to the permutation and combination principle. The previously established Topomer CoMFA and HQSAR models are used to predict the activity of these new compounds. The structure of the newly designed molecule and the predicted pIC_{50} value are shown in Table S5. The results show that the pIC_{50} values of the newly designed inhibitor molecules are better than those of compound 33 and could be used as candidate compounds against the new coronavirus. Among them, compound 1-03 has the highest activity. The four designed compounds can be further studied by ADMET to predict whether they have a good inhibitory effect on SARS-CoV-2.

3.4. Docking analyses

In order to understand the binding mode of cyclic sulfonamide derivatives and protease (PDB code: 7JYC, Resolution: 1.79 ∀) and the antiviral mechanism of the designed compound, and to further verify the results of the QSAR contour map, using SYBYL-X 2.0 software (Surflex-Dock) and discovery studio visualization tool 2017 evaluate their binding affinity to SARS-CoV-2 3CLpro. The docking scores of the compounds with the highest (33) activity and the lowest (2,3,7,8,25,26,27,29) and the interaction patterns of the newly designed compounds with 7JYC are studied. The docking scoring results of the compounds are listed in Table S6 and the scoring functions are used to select the best ligands and predict their binding mode. The higher scoring function value of Total-score, the better affinity between the small molecule ligand extracted from the macromolecular protein and the receptor; the closer absolute value of Crash is to zero, the smaller degree of inappropriateness between the ligand and the receptor extracted from the macromolecular protein. Polar is the score of the polarity function, which can be divided into binding sites located on the surface (the higher the score, the better) and the interior of the molecule (the lower the score, the better) [40]. When the Total-score is greater than 5.0, the output posture is excellent. In this work, the Total score method is used to screen the best posture.

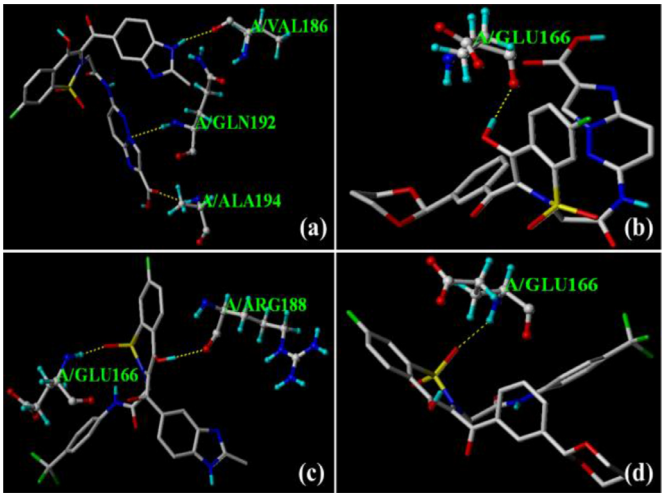

Fig. 10. Results of the newly designed molecular docking:(a)1-01, (b)1-02, (c)1-03, (d)1-04. The rod shape represents the small molecule ligand, the ball and stick shape represent the amino acid residue that forms the hydrogen bond, and the yellow dashed line represents hydrogen bonds.

Fig. 9 shows the docking of molecule No. 33 and the receptor protein. The hydrogen bond formed between the hydroxyl group and the amino acid residue GLU166 (N-H-O, 2.705 Å;O-H-O, 1.804 Å) in the small molecule plays an important role in the inhibitor activity. The hydrophobic channel is composed of Leu50, Met49 and Pro168. The docking score Total-score, Crash score and Polar score are 5.11, -1.33 and 0.84, respectively. Table S6 shows a summary of the scoring functions of all the interaction forces between the molecular ligands of the studied compounds and the proteins. The docking results show that all newly designed molecules (Total-score: 5.65-6.01) have a higher total score function than compound 33 (Total score: 5.11), indicating that the newly designed molecules have a good stability on the active site of the 7JYC protein. Compound 1-02 shows better docking score. Compounds 2,3,7,8,25,26,27,29 have low predicted activity, and the total scoring function is relatively low, indicating that theoretically these compounds have a low antiviral ability.

The same docking protocol is used to link all the designed molecules to the active site of the target protein. The orientation in the docking pocket and the hydrogen bonds formed with surrounding amino acids are shown in Fig. 10 and Fig. S5. The interaction between compound 1-01 and the active binding site of 7JYC is shown in Fig. 10(a). Compound 1-01 forms hydrogen bond donor interaction with GLN192 (N-H-N:2.545 ∀), ALA194 (O-H-N:2.034 ∀) and VAL186 (O-H-N:2.034 ∀); the hydrophobic channel consists of Met165, Pro168, Ala191, and Thr190. Total-score, Crash score and Polar score are 5.66, -1.38 and 1.30, respectively. When compound 1-02 interacts with the active region of the target protein (Fig. 10(b)), it is observed that it forms a hydrogen bond with GLU166 (O-H-O:1.825Å); it has a hydrophobic effect with Met165,

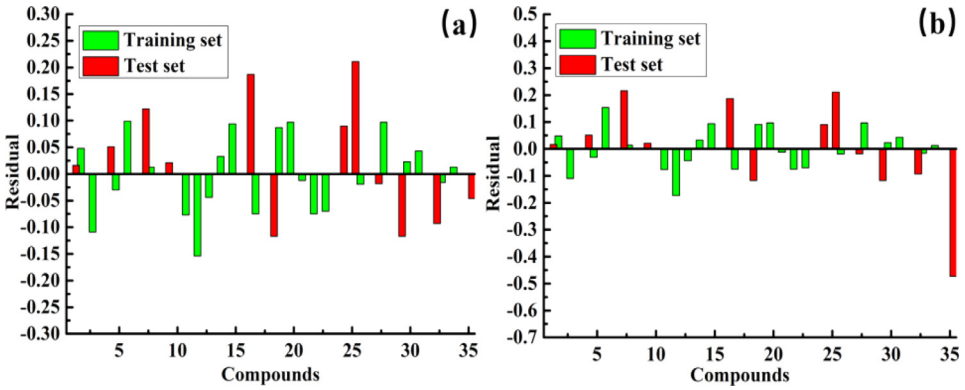

Fig. 11. Residual plots of Topomer CoMFA model (a) and HQSAR model (b).

His41, Met49, Leu167, and Pro168. Total-score, Crash score and Polar score are 6.01, -2.45 and 1.09, respectively.

In Fig. 10(c), compound 1-03 forms a hydrogen bond with GLU166 (NHO:1.827Å) and ARG188 (OHO:2.006Å); the hydrophobic channel is composed of Ala191, Leu167, Thr190 and His41. Total-score, Crash score and Polar score are 5.65, -1.37 and 1.75, respectively. In Fig. 10(d), compound 1-04 forms a hydrogen bond with GLU166 (N-H-O:2.123 Å), and forms highly hydrophobic interactions with residues Ala191, Leu167, Phe185, Pro168, and Met165. Total-score, Crash score and Polar score are 5.11, -1.33 and 0.84, respectively. It is found that the designed new compound is in good agreement with the observed biological activity data, and have a higher activity and Total-score, indicating that the compound is successfully designed.

3.5. Comparative analysis of model results

The predicted activity values and residual values of Tomoper CoMFA model and HQSAR model are shown in Table S7. The residual values of the QSAR model of cyclic sulfonamide derivatives are shown in Fig. 11(a) and Fig. 11(b) respectively. Comprehensive comparison, the Tomoper CoMFA model has smaller residuals than the HQSAR model and is a better model; compounds 1, 8, 10, 21, 26, 27, 33 and 34 obtain the best residual predictions in Topomer CoMFA and HQSAR analysis (residuals <0.02).

The two established models have good internal and external predictive capabilities (Table S8). The results of different models can be verified by each other. Combined with the contour map and color code map of compound 33, it shows a significant area that affects the inhibition of SARS-CoV-2 by cyclic sulfonamide derivatives. Although the two models have obvious differences in structure, the experimental results and predicted biological activities are consistent, indicating that the two models have reliable predictive power for the remodification of cyclic sulfonamide derivative inhibitors. By comparison, the analysis results of the HQSAR model are consistent with the analysis results of the Topomer CoMFA model. In addition, through the HQSAR model analysis, the -CF_3 group on the benzene ring of the R_2 fragment and the fluorine atom position of the R_2 fragment have a positive effect on the inhibitory activity. From the results of molecular docking, we believe that the formation of hydrogen bonds between the oxygen atoms in the common skeleton position of the newly designed molecules and GLU166 has a positive effect on the inhibitory activity. In terms of the structure-activity relationship, the following conclusions can be drawn: R_1 should be a small-volume group with low electronegativity, R_2 should be a large volume group with negative charge, and R_3 should be a small volume group with strong electronegativity, which also explains why compound 33 has the highest activity among all compounds.

3.6. Predicted pharmacokinetic and toxicity properties

The pharmacokinetic/Pre ADMET toxicity predictor (Table S9) is used to predict the main pharmacokinetic parameters (absorption, distribution, metabolism, and excretion) and toxicological properties (Table S10). Absorption is the process by which a drug enters the blood circulation from the site where the drug is used, and the drug can only exert its efficacy after being absorbed. Properties related to absorption include intestinal absorption value (HIA) of drug molecules, skin permeability (SP, log Kp), permeability of Caco-2 cells and permeability of MDCK cells. According to the results recorded in Table S9, it is found that the new design compound shows good intestinal absorption at all of the above sites: greater than 91.00 % (allowable limit: 70 ~100 % abs), and the skin permeability is also within the acceptable range. In addition, the observed moderate permeability through Caco-2 cells in vitrois 84.78 to 88.37 nm/sc, and the detection value of in vitro MDCK cells is low.

The treatment strategy of most drugs is to transport the drugs to the blood through oral administration, subcutaneous injection and transdermal administration, and then distribute them to various tissues. Plasma proteins can adsorb a considerable proportion of drug molecules, and the binding of drugs to plasma proteins (PPB) in the body will directly affect the efficacy and toxicity of the drugs. Central nervous system (CNS) drugs must cross the blood-brain barrier (BBB) to penetrate into brain tissue. The results show that all compounds exhibit strong PPB values (94.85 % ~ 98.93 %), long half-life and low brain penetration. Four compounds have low BBB values and have low permeability to prevent drug entry into CNS and are within the acceptable range of CNS active compounds (>0.4). The metabolic stability of drugs in the body is one of the main factors that affect the pharmacokinetic properties, and unstable compounds often lead to poor pharmacokinetic and pharmacodynamic properties in the body. Cytochrome P450 (CYP450) is the main metabolic enzyme of drugs and other internal and exogenous substances, and its activity can be inhibited or induced by certain drugs. Oxidative metabolism by CYP450 enzymes is the most common way of metabolism. For CYP450-mediated moderate to low levels of metabolism, in vitro metabolic stability is well correlated with in vivo clearance. Calculating the ability of the tested compounds as inhibitors of drug metabolizing enzymes CYP2C19, CYP2C9, CYP2D6 and CYP3A4. In addition, the determination of glycoprotein (P-gp) inhibition rate can predict the excretion performance of the target compound. The tested compounds exhibit good inhibitory effect on CYP2C9 and CYP3A4, but have no inhibitory effect on CYP2C19 and CYP3A4. All the compounds exhibit an inhibitory effect on P-gp.

Toxicity is the degree of damage to the body or cells and organs of a substance, and it is one of the most important reasons for the failure of the late drug development. Predicting the toxicological behavior of the test compound are measurecd by the AMES test, carcino-Mouse/Rate and hERG-inhibition (Table S10). In the AMES test, half of the compounds show non-mutagenic behavior. All compounds have negative carcinogenic effects in mice and rats, except for the low to medium risk cardiotoxic substances.

It is proved from the prediction of the ADMET performance of newly designed compounds that they may have good properties as lead compounds, and we can theoretically consider them to be the best SARS-

CoV-2 inhibitors. It is worth noting that the synthetic feasibility of the designed compounds has been estimated to be around 4.17 ~ 4.5 (difficulty scale is 1 ~ 10), which proves that the difficulty of synthesis is moderate, and the test compounds have great potential to be successfully developed as a drug with low toxicological risk characteristics.

4. Conclusion

In this study, 2*D* (HQSAR) and 3*D* (Topomer CoMFA) methods are used to model 35 cyclic sulfonamide derivatives with anti-SARS-CoV-2 activity, and the ideal and reliable models are obtained. Topomer CoMFA contour maps and HQSAR fragment contribution maps are analyzed to identify the modification sites of the inhibitors, revealing important sites that may significantly affect (increase or decrease) molecular activity. By using molecule 33 as a template for fragment search, four new compounds are designed as potential inhibitors of SARS-CoV-2, and their pIC_{50} values are predicted. Through molecular docking, the interaction between ligand and receptor protein is further explored, which verifies the rationality of the results of the previous structure-activity relationship analysis and the deduced model. The docking results show that the newly designed inhibitor molecule has a significant effect on the GLU166, GLN192, ALA194 and VAL186 sites of the protein(7JYC), and the formation of hydrogen bonds with ASN46 amino acid residues may be the main reason for its antiviral. ADMET prediction results have good pharmacokinetic properties, acceptable absorption, good metabolic conversion, low binding to hERG, no cytotoxicity, and can be used as a reliable SARS-CoV-2 inhibitor. In conclusion, we have used reliable computer-aided drug design methods to design more effective SARS-CoV-2 inhibitors. However, the accuracy of the prediction must be demonstrated experimentally. This study has implications for understanding the mechanism of inhibition of SARS-CoV-2 by cyclic sulfonamide derivatives, as well as the design and synthesis of inhibitors.

Declaration of Competing Interest

The authors declare no competing financial interest.

Acknowledgements

This work was supported by the National Natural Science Funds of China [No.21475081], the Natural Science Foundation of Shaanxi Province of China [No.2019JM237], and the Graduate Innovation Fund of Shaanxi University of Science and Technology.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cjac.2021.09.006.

References

[1] Li Q, Guan X, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. N Engl J Med 2020;382(13). doi:10.1056/NEJMoa2001316.
[2] Silva LR, Santos-Júnior PFDS, Brando J, et al. Druggable targets from coronaviruses for designing new antiviral drugs. Bioorg Med Chem 2020;28(22):115745. doi:10.1016/j.bmc.2020.115745.
[3] Jeong H, Rogers JA, Xu S. Continuous on-body sensing for the COVID-19 pandemic: Gaps and opportunities. Sci Adv 2020;6(36):eabd4794. doi:10.1126/sciadv.abd4794.
[4] Zhang Y, He X, Zhai J, et al. In silico binding profile characterization of SARS-CoV-2 spike protein and its mutants bound to human ACE2 receptor. Brief Bioinform 2021. doi:10.1093/bib/bbab188.
[5] Amin SA, Ghosh K, Singh S, et al. Exploring naphthyl derivatives as SARS-CoV papain-like protease (PLpro) inhibitors and its implications in COVID-19 drug discovery. Mol Divers 2021:1–14. doi:10.1007/s11030-021-10198-3.
[6] Aderibigbe AOO, Pandey P, Doerksen RJ. Negative allosteric modulators of cannabinoid receptor 1: Ternary complexes including CB1, orthosteric CP55940 and allosteric ORG27569. J Biomol Struct Dyn 2021(44):1–19. doi:10.1080/07391102.2021.1921032.
[7] Mao I, Furuyama W, Kuroda M, et al. A biaryl sulfonamide derivative as a novel inhibitor of filovirus infection. Antiviral Res 2020;183:104932. http://dx.org/10.1016/j.antiviral.2020.104932.
[8] Han T, Goralski M, Gaskill N, et al. Anticancer sulfonamides target splicing by inducing RBM39 degradation via recruitment to DCAF15. Science, 2017;356(6336):eaal3755. doi:10.1126/science.aal3755.
[9] Gundla R, Gorantla V, Jadav SS, et al. Molecular Hybrid Design, Synthesis and Biological Evaluation of N-Phenyl Sulfonamide Linked N-Acyl Hydrazone Derivatives Functioning as COX-2 Inhibitors: New Anti-Inflammatory, Anti-Oxidant and Anti-Bacterial Agents. New J Chem 2017. doi:10.1039/C7NJ03332J.
[10] Alessio N, Supuran CT. Carbonic anhydrase inhibitors as antitumor/antimetastatic agents: a patent review (2008–2018). Expert Opin Ther Pat 2018;28. doi:10.1080/13543776.2018.1508453.
[11] Boechat N, Pinheiro L, Santos-Filho OA, et al. Design and synthesis of new N-(5-trifluoromethyl)-1H-1,2,4-triazol-3-yl benzenesulfonamides as possible antimalarial prototypes. Molecules 2011;16(9):8083–97 https://doi.org/10.3390/molecules16098083.
[12] Ukrainets I, V, et al. New Synthesis and Analgesic and Diuretic Activity of Halo-Substituted 4-Hydroxy-1-Methyl-2,2-Dioxo-1H-2 lambda (6),1-Benzothiazine-3-Carboxanilides. Pharm Chem J 2016 0091-150X/16/5009-0589.
[13] Gannarapu MR, Vasamsetti SB, Punna N, et al. Synthesis of novel 1,2-benzothiazine 1,1-dioxide-3-ethanone oxime N-aryl acetamide ether derivatives as potent anti-inflammatory agents and inhibitors of monocyte-to-macrophage transformation. Eur J Med Chem 2014;75:143–50. doi:10.1016/j.ejmech.2013.12.053.
[14] Kim SH, Ramu R, Kwon SW, et al. Discovery of cyclicsulfonamide derivatives as 11β-hydroxysteroid dehydrogenase 1 inhibitors. Bioorg Med Chem Lett 2010;20(3):1065–9. doi:10.1016/j.bmcl.2009.12.035.
[15] ghoobi Marzieh Y ae, Frimayanti Neni, Chee Chin Fei, et al. QSAR, in silico docking and in vitro evaluation of chalcone derivatives as potential inhibitors for H1N1 virus neuraminidase. Med Chem Res 2016. doi:10.1007/s00044-016-1636-5.
[16] Karoll FP, Luis E, Jaime M. Insights into the Structural Requirements of Potent Brassinosteroids as Vegetable Growth Promoters Using Second-Internode Elongation as Biological Activity: CoMFA and CoMSIA Studies. Int J Mol Sci 2017;18(12):2734. doi:10.3390/ijms18122734.
[17] Faidallah HM, Panda SS, Serrano JC, et al. Synthesis, antimalarial properties and 2D-QSAR studies of novel triazole-quinine conjugates. Bioorg Med Chem 2016:3527–39 http://dx.doi.org/. doi:10.1016/j.bmc.2016.05.060.
[18] Shin YS, Lee JY, Noh S, et al. Discovery of cyclic sulfonamide derivatives as potent inhibitors of SARS-CoV-2. Bioorg Med Chem Lett 2020;31:127667. doi:10.1016/j.bmcl.2020.127667.
[19] None. GALAHAD Tripos, Inc. 1699 South Hanley Road, St. Louis, MO 63144-2319. www.tripos.comContact company for pricing information. J Am Chem Soc 2007;129(17):5780 -5780. doi:10.1021/ja069815j.
[20] Bai F, Liu X, Li J, et al. Bioactive conformational generation of small molecules: A comparative analysis between force-field and multiple empirical criteria based methods. BMC Bioinformatics 2010;11(1):545. doi:10.1186/1471-2105-11-545.
[21] Jiao L, Wang Y, Qu L, et al. Hologram QSAR study on the critical micelle concentration of Gemini surfactants. Colloids and Surfaces A Physicochemical and Engineering Aspects 2019;586:124226. doi:10.1016/j.colsurfa.2019.124226.
[22] Huang D, Liu Y, Shi B, et al. Comprehensive 3D-QSAR and binding mode of BACE-1 inhibitors using R-group search and molecular docking. J Mol Graph Model 2013;45C(18):65–83. doi:10.1016/j.jmgm.2013.08.003.
[23] Maltarollo VG, Honório KM, Emery FS, et al. Hologram quantitative structure–activity relationship and comparative molecular interaction field analysis of aminothiazole and thiazolesulfonamide as reversible LSD1 inhibitors. Future Medicinal Chemistry 2015;7(11):1381–94. doi:10.4155/fmc.15.68.
[24] Zhang H, Lv Q, Xu W, et al. 4D-QSAR studies of CB 2 cannabinoid receptor inverse agonists: a comparison to 3D-QSAR. Med Chem Res 2019. doi:10.1007/s00044-019-02303-x.
[25] Kiralj R, Ferreira M. Basic validation procedures for regression models in QSAR and QSPR studies: theory and application. J Braz Chem Soc 2009;20(4):770–87. doi:10.1590/S0103-50532009000400021.
[26] Golbraikh A, Tropsha A. Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. J Comput Aided Mol Des 2002;16(5):357–69. doi:10.1023/A:1020869118689.
[27] Aptula Aynur O, Jeliazkova Nina G, Schultz Terry W, et al. The Better Predictive Model: High q2 for the Training Set or Low Root Mean Square Error of Prediction for the Test Set? Mol Inf 2010;24(3):385–96. doi:10.1002/qsar.200430909.
[28] Golbraikh A, Tropsha A. Beware of q2!. J Mol Graph Model 2002;20(4):269–76. doi:10.1016/S1093-3263(01)00123-1.
[29] Patricia DCL, Golbraikh A, Oloff S, et al. Combinatorial QSAR Modeling of P-Glycoprotein Substrates. J Chem Inform Model 2006;46(3):1245–54. doi:10.1021/ci0504317.
[30] Chirico N, Gramatica P. Real external predictivity of QSAR models: how to evaluate it? Comparison of different validation criteria and proposal of using the concordance correlation coefficient. J Chem Inform Model 2011;51(9):2320–35. doi:10.1021/ci200211n.
[31] Patil V, Noonikara-Poyil A, Joshi SD, et al. Synthesis, molecular docking studies, and in vitro evaluation of 1,3,5-triazine derivatives as promising antimicrobial agents. J Mol Struct 2020;1220:128687. doi:10.1016/j.molstruc.2020.128687.
[32] Sterling T, Irwin JJ. ZINC 15 – Ligand Discovery for Everyone. J Chem Inf Model 2015:2324. doi:10.1021/acs.jcim.5b00559.
[33] Dudutien V, Zubrien A, Kairys V, et al. Isoform-Selective Enzyme Inhibitors by Exploring Pocket Size According to the Lock-and-Key Principle. Biophys J 2020. doi:10.1016/j.bpj.2020.08.037.

[34] Saadah LM, Deiab IA, Al-Balas Q, et al. Carnosine to Combat Novel Coronavirus (nCoV): Molecular Docking and Modeling to Cocrystallized Host Angiotensin-Converting Enzyme 2 (ACE2) and Viral Spike Protein. Molecules 2020;25(23):5605. doi:10.20944/preprints202010.0486.v1.

[35] Ambure PS, Sangamwar GAT. 3D-QSAR and molecular docking analysis of biphenyl amide derivatives as p38α mitogen-activated protein kinase inhibitors. Mol Divers 2012;16(2):377–88. doi:10.1007/s11030-011-9353-y.

[36] Tong JB, Luo D, Bian S, Zhang X. Structural investigation of tetrahydropteridin analogues as selective PLK1 inhibitors for treating cancer through combined QSAR techniques, molecular docking, and molecular dynamics simulations. J Mol Liq 2021. doi:10.1016/j.molliq.2021.116235.

[37] Prashantha CN, Gouthami K, Lavanya L, et al. Molecular screening of antimalarial, antiviral, anti-inflammatory and HIV protease inhibitors against spike glycoprotein of Coronavirus. J Mol Graph Model 2020;102. doi:10.1016/j.jmgm.2020.107769.

[38] Wang Y, Xing J, Xu Y, et al. In silico ADME/T modelling for rational drug design. Q Rev Biophys 2015;48(04):488–515. doi:10.1017/S0033583515000190.

[39] Zhang XZ, Xu Y, et al. Synthesis, in vitro assays, molecular docking, theoretical ADMET prediction, and evaluation of 4-methoxy-phenylthiazole-2-amine derivatives as acetylcholinesterase inhibitors. Med Chem Res 2019;28(10):1683–93. doi:10.1007/s00044-019-02405-6.

[40] Tong JB, Luo D, Zhang X, et al. Design of novel SHP2 inhibitors using Topomer CoMFA, HQSAR analysis, and molecular docking. Struct Chem 2020;32(3):1–16. doi:10.1007/s11224-020-01677-8.