



Rational design of novel phenol ether derivatives as non-covalent proteasome inhibitors through 3D-QSAR, molecular docking and ADMET prediction

Miao Yuan¹, Hanwen Ji¹, Fengxin Sun², Qiang Chen¹, Ping Cheng^{1*}

¹School of Materials and Chemistry, University of Shanghai for Science and Technology, Shanghai 200093, China

²China United Concrete Jining Co., Ltd, Jining 272000, Shandong, China

***Correspondence:** Ping Cheng, School of Materials and Chemistry, University of Shanghai for Science and Technology, 516 Jungong Road, Shanghai 200093, China. chengp@usst.edu.cn

Academic Editor: Walter Filgueira de Azevedo Jr., Pontifical Catholic University of Rio Grande do Sul, Brazil

Received: June 30, 2023 **Accepted:** August 29, 2023 **Published:** December 27, 2023

Cite this article: Yuan M, Ji H, Sun F, Chen Q, Cheng P. Rational design of novel phenol ether derivatives as non-covalent proteasome inhibitors through 3D-QSAR, molecular docking and ADMET prediction. *Explor Drug Sci.* 2023;1:435–53. <https://doi.org/10.37349/eds.2023.00029>

Abstract

Aim: The purpose of this paper is to use different structures and ligand-based drug design methods properly to provide theoretical guidance for the design of novel non-covalent proteasome inhibitors, and conduct theoretical analysis of the binding interaction mode between receptors and ligands. At the same time, the pharmacokinetic (PK) prediction, drug-likeness, and synthesis prediction were made for the screened novel drugs. Therefore, potentially attractive non-covalent proteasome inhibitors with low toxicity could be found as anticancer drugs.

Methods: In this work, computer-aided drug design methods, including quantitative structure-activity relationship (QSAR), molecular docking, absorption, distribution, metabolism, excretion, and toxicology (ADMET) prediction, and drug-likeness prediction methods were performed.

Results: In this study, the structure-activity relationship (SAR) of a series of non-covalent proteasome inhibitors were studied and the optimal comparative molecular field analysis (CoMFA; $Q^2 = 0.574$, $r^2 = 0.999$, $r^2_{\text{pred}} = 0.755$) and comparative molecular similarity indices analysis (CoMSIA)-SEHA ($Q^2 = 0.584$, $r^2 = 0.989$, $r^2_{\text{pred}} = 0.921$) models were obtained. According to the results of the QSAR model, some vital clues were found that would effectively enhance the biological activity of the compound. Based on these clues, 24 novel non-covalent proteasome inhibitors (D01–D24) were finally designed and screened. While the binding models between proteasome [protein data bank (PDB) code: 3MG6] and three representative compounds (15, 20, and D24) were also analyzed by using the molecular docking method. The results suggested that hydrogen bond and hydrophobic interaction played a key role in binding interaction between the receptor and ligand. In addition, the results of ADMET prediction indicated that the new designed compounds had reasonable PK parameters and drug-like properties.

Conclusions: These statistical results can provide theoretical guidance for structural optimization, design, and synthesis of more effective non-covalent proteasome inhibitors in the future.

© The Author(s) 2023. This is an Open Access article licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



Keywords

Non-covalent proteasome inhibitor, 3D-QSAR, CoMFA, CoMSIA, molecular docking, ADMET properties, drug-likeness

Introduction

The ubiquitin-proteasome system (UPS) is a major pathway for the selective degradation of proteins in eukaryotic cells [1]. UPS involves many biological processes such as cell cycle progression, transcription repair, cell proliferation, differentiation, and apoptosis. Thus, UPS plays an important role in transcription, protein degradation, and protein stability of eukaryotic cells. It is an intracellular non-lysosomal protein degradation pathway, which is composed of ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzyme (E2), ubiquitin ligase (E3), and the proteasome. Most of the substrates of this pathway can be recognized and degraded by 26S proteasome only after they are labeled by ubiquitin enzyme E1–E3, i.e., ubiquitination [2, 3]. The common 26S proteasome specifically degrades target proteins into polypeptides containing 7–9 amino acid residues in the form of ATP consumption [4]. The 26S proteasome mainly contains a 20S core particle (20S proteasome) and a 19S regulatory particle. The 20S proteasome consists of two α chains and two β chains respectively. Each α chain and β chain consists of seven subunits ($\alpha 1$ – $\alpha 7$) and seven subunits ($\beta 1$ – $\beta 7$), respectively. Among them, $\beta 1$, $\beta 2$, and $\beta 5$ subunits are active sites for protein degradation, which are responsible for the caspase-like (C-L), trypsin-like (T-L), and chymotrypsin-like (ChT-L) activities of the proteasome, respectively. Each of these three β -subunits has a catalytic site that cleaves the peptide bond using the nucleophilic γ -hydroxyl group of N-terminal threonine. According to the different composition sequences of $\beta 1$, $\beta 2$, and $\beta 5$ subunits, 20S proteasomes can be divided into immune 20S proteasomes ($\beta 1i$, $\beta 2i$, and $\beta 5i$) and constitutive 20S proteasomes ($\beta 1c$, $\beta 2c$, and $\beta 5c$), and different 20S proteasome inhibitors act on different subunit types and therefore have different pharmacological effects [5, 6]. UPS controls most biological functions within cellular mechanisms and handles 80% to 90% of intracellular protein degradation, therefore proteasomes have become attractive targets for the treatment of inflammatory, autoimmune, and neoplastic diseases [7].

It has been reported that the level and activity of proteasomes are higher than 90% in malignant tumors compared to normal cells, providing a survival advantage for tumor cells to continue proliferating [8]. Therefore, tumor cells will show greater sensitivity to proteasome inhibitors [9]. In 2003, bortezomib (the first-generation drug) became the first proteasome inhibitor approved by the Food and Drug Administration (FDA) for the treatment of multiple myeloma and mantle cell lymphoma, enabling the proteasome to be clinically validated as an oncology therapeutic target [10]. These were carfilzomib (the second-generation drug) [11] and ixazomib [12], which were also approved by FDA. These major 20S proteasome inhibitors reported in the literature are covalent inhibitors, which have highly reactive and unstable chemical groups [13]. They generally react covalently and irreversibly with the proteolytic sites of the proteasome, resulting in permanent blockage of the proteasome [14]. In addition, the reactive head groups lead to nonspecific binding to the active center of the proteasome and the conversion of many enzyme off-target activities into serious side effects. These may be the main cause of side effects, acquired resistance, and unsatisfactory pharmacokinetic (PK) properties of covalent proteasome inhibitors [15–17]. Hence scientists have turned their attention to the study of non-covalent proteasome inhibitors with lower toxicity in recent years. Although there is less research on non-covalent proteasome inhibitors than covalent inhibitors, it provides a promising alternative mechanism for inhibiting proteases [18]. Their potential advantages are high selectivity, moderate response, and reduced instability, which may reduce some side effects [19]. Considering these reasons, it is essential to continue the design and research of novel non-covalent proteasome inhibitors.

The purpose of this paper is to use advanced computer-aided drug design technology to study the structure-activity relationship (SAR) of a series of non-covalent proteasome inhibitors based on the research results of Yu et al. [20], so as to find their common backbones and do structural modifications. By enumerating a large number of novel compounds on the core skeleton, various physicochemical properties of new molecules

are predicted and analyzed, and then novel drug molecules with anticancer activity are screened out. In this study, the quantitative SAR (QSAR) model was constructed by SYBYL-X software to study SAR between the structural characteristics of non-covalent proteasome inhibitors and their biological activities. These models provided theoretical guidance for the design of new compounds; and then the molecular docking simulation was applied to analyze the binding mode and stability between the active pocket of proteasome and inhibitors. Moreover, the absorption, distribution, metabolism, excretion, and toxicology (ADMET) and drug-like properties of new non-covalent proteasome inhibitors with higher selectivity were also evaluated. As an effective, rapid, and economical tool, computer-aided drug design has been widely used in the development of new drugs. It not only improves the hit rate of drug candidates but provides new strategies for novel drug design [21]. For example, Xu et al. [22] discovered a new set of non-covalent proteasome inhibitors through a fragment-based approach to drug design (i.e., docking calculations). Li et al. [23] obtained 2,167 compounds by virtually screening the SPECS database through using non-covalent docking and a 20S proteasome-based pharmacophore model, and finally two hit compounds were selected after molecular dynamics simulations.

Materials and methods

Dataset selection and biological activities

A data set, containing twenty-eight phenol ether derivatives as non-covalent proteasome inhibitors, was collected from the same literature to build the QSAR models [20]. These compounds were randomly divided into two groups. The training set of 22 compounds was used to generate quantitative models. The other was a test set of 6 molecules, which was designed to verify the reliability of the created model. The half maximal inhibitory concentration (IC_{50}) of these compounds was known, which was converted to the pIC_{50} value [$pIC_{50} = -\log_{10}(IC_{50})$] as the dependent variable for QSAR modeling. The pIC_{50} values of 28 phenol ether derivatives covered a wide range from 5.229 to 7.310, which supplied extensive and homogenous data for QSAR analysis. The backbones of non-covalent proteasome inhibitors are shown in Figure 1, which are separated into two series. The backbone of the first series is shown in Figure 1A, and the compounds are from 1 to 16. The backbone of compounds 17–28 in Table 1 is illustrated in Figure 1B. The chemical structures of non-covalent proteasome inhibitors and their biological activity values are shown in Table 1.

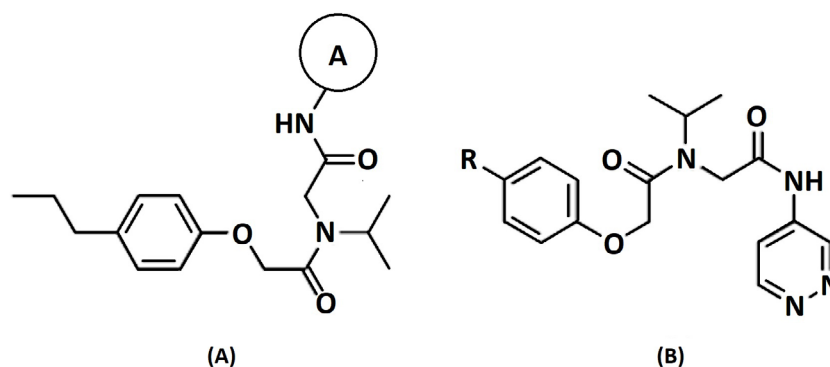


Figure 1. The chemical structure of non-covalent proteasome inhibitors. (A) The backbone of compound 1–16 in Table 1; (B) the backbone of compound 17–28 in Table 1

Table 1. The structures and activities of non-covalent proteasome inhibitors

No.	Group	IC_{50} (nmol/L)	pIC_{50}	CoMFA		CoMSIA	
				Pred.	Res.	Pred.	Res.
A-group							
1		1,682	5.774	5.793	-0.019	5.840	-0.066
2		836	6.078	6.093	-0.015	6.088	-0.010

Table 1. The structures and activities of non-covalent proteasome inhibitors (*continued*)




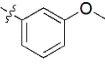
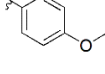
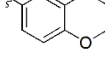
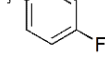
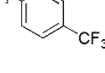
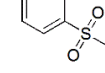
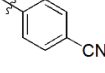


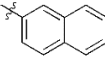
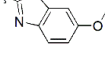


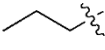
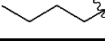
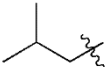
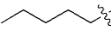

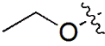


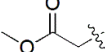
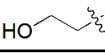
No.	Group	IC ₅₀ (nmol/L)	pIC ₅₀	CoMFA		CoMSIA	
				Pred.	Res.	Pred.	Res.
3		280	6.553	6.494	0.059	6.463	0.090
4*		395	6.403	6.285	0.118	6.285	0.118
5		89	7.051	7.071	-0.020	7.071	-0.020
6		2,000	5.699	5.705	-0.006	5.661	0.038
7		952	6.021	6.044	-0.023	5.956	0.065
8		4,726	5.326	5.317	0.009	5.370	-0.044
9*		1,197	5.922	6.137	-0.215	6.095	-0.173
10*		922	6.035	6.744	-0.709	6.345	-0.310
11		230	6.638	6.624	0.014	6.625	0.013
12		361	6.442	6.440	0.002	6.490	-0.048
13		3,863	5.413	5.423	-0.010	5.447	-0.034
14*		2,011	5.697	5.562	0.135	5.528	0.169
15		5,907	5.229	5.221	0.008	5.187	0.042
16		1,111	5.954	5.955	-0.001	5.977	-0.023
R-group							
17		481	6.318	6.334	-0.016	6.270	0.048
18		243	6.614	6.593	0.021	6.537	0.077
19*		89	7.051	6.837	0.214	6.887	0.164
20**		49	7.310	7.303	0.007	7.164	0.146

Table 1. The structures and activities of non-covalent proteasome inhibitors (*continued*)

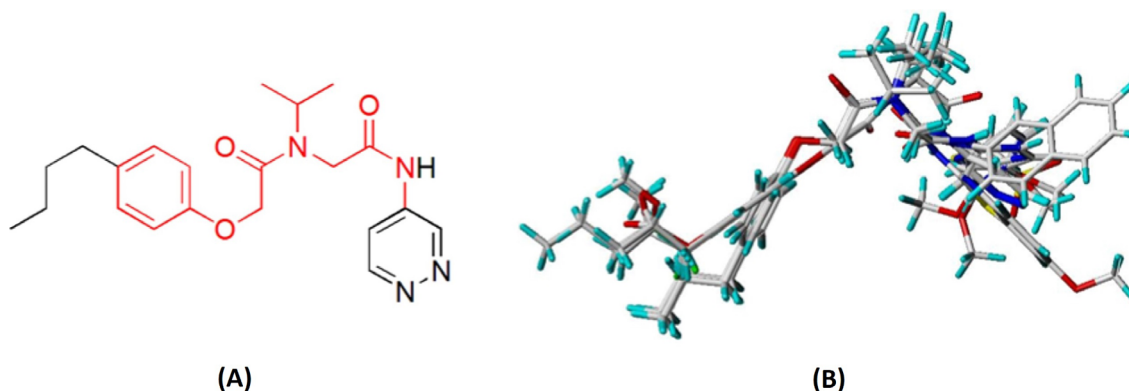
No.	Group	IC ₅₀ (nmol/L)	pIC ₅₀	CoMFA		CoMSIA	
				Pred.	Res.	Pred.	Res.
21		118	6.928	6.938	-0.010	6.978	-0.050
22		64	7.194	7.212	-0.018	7.304	-0.110
23		579	6.237	6.226	0.011	6.306	-0.069
24		259	6.587	6.592	-0.005	6.593	-0.006
25		952	6.021	6.015	0.006	6.096	-0.075
26*		405	6.393	6.099	0.294	6.281	0.112
27		1,561	5.807	5.821	-0.014	5.765	0.042
28		968	6.014	5.995	0.019	6.020	-0.006

*: test set compounds; **: the most active compound. CoMFA: comparative molecular field analysis; CoMSIA: comparative molecular similarity indices analysis; Pred.: predicted value; Res.: residual value

Minimization and molecular alignment

The three-dimensional (3D) structure of 28 phenol ether derivatives was constructed in the Sketch module of the SYBYL-X 2.0 software package (Tripos, St. Louis, USA) [24]. Each molecule was minimized to obtain the lowest energy conformation. Energy minimization of all compounds was performed using the Tripos force field with Powell method. Gasteiger Hückel method was used to calculate the atomic charge. In order to get stable conformations, the maximum number of iterations was set to 10,000 and the energy convergence gradient was set to 0.005 kcal/(mol Å) [25–27]. The other parameters adopt the system default.

Molecular alignment operation is not only the basic step to establish a reliable 3D-QSAR model, but the quality of molecular alignment directly affects the prediction ability of the constructed model [28, 29]. To get the best QSAR model, we selected the most active compound 20 (IC₅₀ = 49 nmol/L) as the template molecule for molecular alignment [30, 31]. The chemical structure of the template molecule is shown in Figure 2A and the red part represents the common skeleton. The result of alignment based on the common substructure is shown in Figure 2B.

**Figure 2.** The common substructure and the result of alignment. (A) Compound 20 was used as a template for the alignment. The common substructure was marked in red color; (B) the alignment diagram of the training set

CoMFA and CoMSIA models

CoMFA and CoMSIA are two general applied tools of 3D-QSAR [26]. These methods are based on descriptors of 3D structures to analyze the different contributions of steric (S), electrostatic (E), hydrophobic (H), H-bond donor (HD), and H-bond acceptor (HA) fields [32, 33]. These descriptors are directly related to the atomic properties of compounds and the spatial geometry of molecules. Potentials are reflected in their position and expansion in space and their intensity. The CoMFA and CoMSIA models are performed using the QSAR option of SYBYL-X 2.0 software with the default parameters.

The non-covalent proteasome inhibitors were placed in the spatial grid to establish the CoMFA model. A 3D cubic box with 2.0 Å grid spacing in the X, Y, and Z directions was generated by the default values. The spatial lattice was linked with 3D structure, bioactive data, and molecular potential energy, thus providing valuable information for molecular modifications. S and E fields were calculated at each grid point with Tripos force field using a carbon atom probe with sp³ hybridization by means of a van der Waals radius of 1.52 Å and a charge of +1.0. In addition, the default S and E energy cutoffs were both set to 30 kcal/mol to avoid the infinity of energy values inside a compound [34, 35].

CoMSIA and CoMFA applied for the same molecular alignment method, however, the results of CoMSIA were more robust than the CoMFA. CoMSIA model was not limited to S and E fields but covered H, HD, and HA fields, which compensated for some shortcomings of CoMFA model. CoMSIA model was calculated using a sp¹ hybridized carbon atom with the radius of 1.0 Å and net +1.0 charge as a probe atom. The column filter and attenuation factor were set to 2.0 kJ/mol and 0.3, respectively, to reduce noise and speed up analysis [36]. The other parameters took the system default.

Model performance and validation

The partial least-squares (PLS) method was a multi-regression analysis method based on the ideas of linear transformation [37]. The inhibitor activity was used as the dependent variable and descriptors of potential fields as independent variables for the 3D-QSAR model. In the cross-validation process, the optimum number of components (ONCs) and cross-validation correlation coefficient (Q^2) were calculated by leave-one-out (LOO) method. Then the non-cross-validation correlation coefficient (r^2), standard error of estimate (SEE), and F -test values (F) were further obtained in the non-cross-validation analysis. The reliable 3D-QSAR models should have low values of SEE and high values of Q^2 , r^2 , and F . These important statistical parameters reflected whether the constructed model was reliable and robust or not [38, 39].

The external validation was also a crucial step because it could guarantee the predictive accuracy of 3D-QSAR. A favorable model relied on the high predicted external validation correlation coefficient r^2_{pred} value ($r^2_{\text{pred}} > 0.6$) [40]. According to Roy et al. [41] and Mitra et al. [42], the other four statistical parameters also need to meet certain rules to ensure the robustness of the model. These rules were the following criteria:

$$R^2 > 0.6, 0.85 \leq K \leq 1.15, [(R^2 - R_0^2)/R^2] < 0.1, R_m^2 > 0.5$$

R^2 was the square correlation coefficient between the experimental and predicted bioactivity. The formula was as follows:

$$R^2 = \left| \frac{\sum (Y_{\text{test}} - \overline{Y_{\text{test}}}) (Y_{\text{pred}(\text{test})} - \overline{Y_{\text{pred}(\text{test})}})}{\sqrt{\sum (Y_{\text{test}} - \overline{Y_{\text{test}}})^2 \sum (Y_{\text{pred}(\text{test})} - \overline{Y_{\text{pred}(\text{test})}})^2}} \right|^2$$

K was the regression of experimental and Pred. and the regression line slope through the origin. The formula was as follows:

$$K = \frac{\sum (Y_{\text{test}} * Y_{\text{pred}(\text{test})})}{\sum (Y_{\text{pred}(\text{test})})^2}$$

R_0^2 was the square correlation coefficient by calculating the experimental *versus* Pred. The formula was as follows:

$$R_0^2 = 1 - \frac{\sum (Y_{pred(test)} - kY_{pred(test)})^2}{\sum (Y_{pred(test)} - \overline{Y_{pred(test)}})^2}$$

R_m^2 was another criterion to analyze the external predictability of the model. The formula was as follows:

$$R_m^2 = R^2 * (1 - \sqrt{R^2 - R_0^2})$$

Where Y_{test} , $\overline{Y_{test}}$, $Y_{pred(test)}$, and $\overline{Y_{pred(test)}}$ represented experimental, average, predictive and predictive average bioactivity values in the test set, respectively.

Molecular docking

Molecular docking is a general method of simulating ligand-protein interactions, through which some vital information about ligand at particular binding site can be provided. For example, we can know the overall geometric conformation of ligands and the microscopic interactions within the pockets [43]. In this study, the Autodock software was used for the docking process [44, 45]. Since the ligand in the crystal structure of 3MG6 [protein data bank (PDB) ID] are selective for the $\beta 5$ (ChT-L) site of the 20S core particle of the proteasome, our designed molecules have the same mechanism of selective inhibition of the $\beta 5$ site of the proteasome [46, 47]. Therefore, to elucidate the binding modes of a new series of non-covalent proteasome inhibitors with exquisite potency and selectivity for the 20S $\beta 5$ -subunit, molecular docking calculations were performed. The X-ray crystal structure of the proteasome (PDB ID: 3MG6) was downloaded from the PDB (<https://www.rcsb.org/>). There were two core operations in the docking process. The first one involved pretreatment of proteins, including the removal of water molecules and original ligand, modification of missing loop regions, the addition of missing atoms and polarized hydrogen, and charges calculations. Then the compounds were docked into the active pockets of protein, which was defined by a $45 \times 45 \times 45$ box centroid of the crystal ligand with a default grid space size of 0.375 \AA [48]. All other parameters remained the default values. Finally, ten different molecular conformations were generated for each compound. The docking results were analyzed using PyMol software.

Prediction of ADMET and drug-likeness

PKs and toxicology are important components of drug preclinical research. PK reflects the patterns of absorption, distribution, metabolism, and excretion of drugs, which are necessary processes for drug metabolism in the human body. According to statistics, most candidate drugs failed in drug development due to poor PK properties or excessive toxicity [49, 50]. Therefore, the ADMET comprehensive evaluation of candidate compounds is conducive to improving the success rate of new drug research and development [51]. In this study, ADMET and drug-like properties of newly designed phenol ether derivatives were obtained from pkCSM online server (<http://biosig.unimelb.edu.au/pkcsm/prediction>) [52] and SwissADME web tool (<http://www.swissadme.ch>) [53]. The predicted ADMET data will provide some theoretical support for further experimental verification.

Results

Statistical results of CoMFA and CoMSIA

In order to acquire the best statistical model, we combined different fields to build a wide variety of models, calculated their statistical data, and finally elected the best QSAR model. The results of the CoMFA model and some CoMSIA models utilizing different combination types of fields are shown in Table S1. The Q^2 and r^2 values of a good 3D-QSAR model should be greater than 0.5 and 0.9, respectively. For the PLS analysis results of CoMFA-SE (using both S and E fields to build models) model, Q^2 , ONC, r^2 , SEE, and F was 0.574, 10, 0.999, 0.026 and 1,084.404, respectively, suggesting that it had good internal validation capabilities. The predicted external validation correlation coefficient r_{pred}^2 value was 0.755. The contribution for S and E fields were 47.1% and 52.9%, respectively. All this statistical data indicated that this model was reliable and robust. Therefore, CoMFA-SE was selected as the final CoMFA model.

The CoMSIA-HAD model had the highest Q^2 value ($Q^2 = 0.647$) in all CoMSIA series, but it was not chosen as the final CoMSIA model. Because we found that the contribution rate of E field always had a high weight in CoMSIA series. For example, the contribution of E was up to 32.3% in the CoMSIA-SEHAD model with all fields. Therefore, we could not ignore the importance of E in the CoMSIA model. However, the contribution of the HD showed a downward trend with the increase in the number of fields. For instance, it contributed the least (8.1%) in the CoMSIA-SEHAD model. So, the role of HD was not considered at all. Meanwhile, the CoMSIA-SEHA model had the highest Q^2 value in these five models (CoMSIA-SEHD, CoMSIA-SEHA, CoMSIA-SEDA, CoMSIA-SHDA, CoMSIA-EHAD), hence CoMSIA-SEHA was selected as the CoMSIA model. The CoMSIA-SEHA model had a large Q^2 value of 0.584 and r^2 value of 0.989, lower SEE value of 0.077, and F value of 172.183, showing that this model was reliable. The r^2_{pred} value was 0.921, suggesting that this model had a strong external predictive capability. These statistical results demonstrated that the CoMSIA-SEHA model was robust. The contributions of S, E, H, and HA were 12.6%, 35.2%, 27.6%, and 24.5%, respectively. It showed that E played a vital role in this model. To sum up, CoMSIA-SEHA was selected as the best CoMSIA model. The statistical parameters of 3D-QSAR model after PLS analysis are shown in [Table 2](#).

Table 2. Statistical results of optimal CoMFA and CoMSIA models

Parameters	Threshold value	CoMFA-SE	CoMSIA-SEHA
Q^2 (cross-validate)	> 0.5	0.574	0.584
r^2	> 0.6	0.999	0.989
SEE	-	0.026	0.077
F	-	1,084.404	172.183
ONC	-	10	7
r^2_{pred}	> 0.6	0.755	0.921
R^2	> 0.6	0.684	0.897
K	$0.85 \leq K \leq 1.15$	0.996	1.005
R_0^2	Close to value of r^2	0.998	0.996
$(R^2 - R_0^2)/R^2$	< 0.1	-0.459	-0.110
R_m^2	> 0.5	0.559	0.614
Field contribution (%)	-	-	-
S	-	47.1	12.6
E	-	52.9	35.2
H	-	-	27.6
HA	-	-	24.5

-: not applicable

To test the predictive ability of the QSAR model, the test set of six compounds was selected for external verification to evaluate it. The r^2_{pred} of the CoMFA-SE (hereafter referred to CoMFA) and CoMSIA-SEHA (hereafter referred to as CoMSIA) models was greater than 0.6, indicating that the established QSAR model had a decent external predictive ability. The relevant statistical results are shown in [Table 2](#), and the other four statistical parameters calculated from the formulas of Roy et al. [41] and Mitra et al. [42] are also shown in [Table 2](#). The CoMFA and CoMSIA models had R^2 at 0.684 and 0.897, K at 0.996 and 1.005, $(R^2 - R_0^2)/R^2$ at -0.459 and -0.110, R_m^2 at 0.559 and 0.614, respectively. All four statistical parameters met the Roy et al. [41] and Mitra et al. [42] rules, which indicated the created model had stronger external verification. The above results suggested that the CoMFA and CoMSIA models had good external predictive capacity.

The predicted bioactivity values of QSAR models for the training set and test set are shown in [Table 1](#). The scatter plots of experimental and predicted pIC_{50} values are shown in [Figure 3](#). It could be seen from the picture that red triangle blocks and blue dots were close to the straight line ($Y = X$), which meant the actual values of whole molecules were almost consistent with the Pred. All these data clearly indicated the excellent stability and the highly predictive characteristics of 3D-QSAR models.

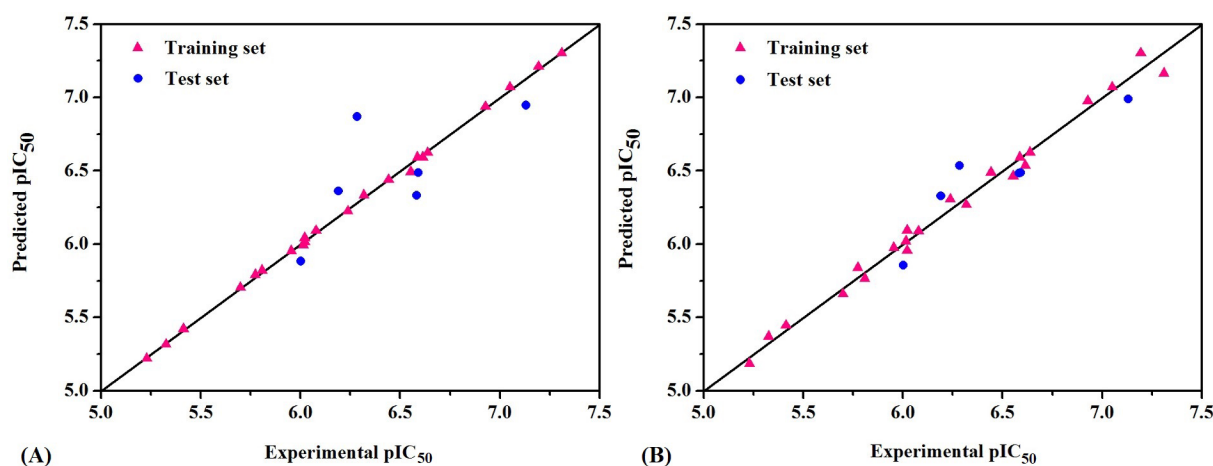


Figure 3. The scatter plots of experimental versus predicted pIC_{50} for CoMFA (A) and CoMSIA (B) models

Contour map analysis

By analyzing the contour maps of CoMFA and CoMSIA models, we can vividly study the physical and chemical properties of the compound to find the key information affecting the bioactivity, which will guide us to design novel molecules. In this study, we mainly discussed the 3D-QSAR in four potential fields. The favorable and unfavorable contributions of all fields default to 80% and 20%, respectively. To better explain the contour maps, the inhibitor of the highest activity (compound 20) was labeled as shown in [Figure 2A](#).

In the S contour maps, the green blocks indicated that the introduction of bulky groups in the region would improve the bioactivity, while the yellow regions mean the opposite. S fields for CoMFA and CoMSIA were shown in [Figures 4](#) and [5](#), respectively. As shown in [Figure 4A](#), a large-sized yellow block was displayed near the R^3 area, while a medium-sized green contour existed in the R^1 region. In addition, in [Figure 5A](#), compound 20 was surrounded by a large yellow contour and relatively small green block in the R^3 and R^1 regions, respectively, which was basically consistent with the CoMFA model analysis. It was worth mentioning that a large-sized yellow contour appeared at the R^3 opposite position in the CoMSIA model, but not in the CoMFA model. Overall, we found three key clues to enhance the activity of non-covalent proteasome inhibitors in the S fields: the bulky substituents at the R^1 of compound 20, and the small group at the R^3 and R^3 opposite position, respectively.

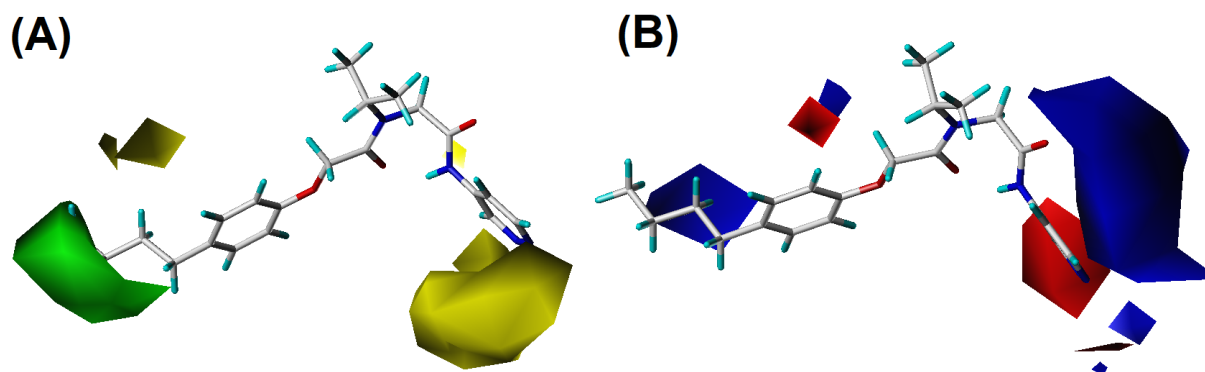


Figure 4. The CoMFA contour maps of template molecule 20. (A) The S field; (B) the E field

In E field contour maps shown in [Figures 4B](#) and [5B](#) for CoMFA and CoMSIA, respectively, the red and blue blocks indicated that negative and positive groups in these regions could strengthen the inhibition of the compound. A medium-sized blue and a large blue contour appeared at the R^1 and R^2 in [Figure 4B](#), which suggested that the original group replaced by the positive substitutions would increase the bioactivity. In the R^3 region, there was a medium-sized red contour indicating that the negative group here was beneficial to the inhibition of the compound. As for the CoMSIA, the situation was similar to the CoMFA model, so the details were not discussed.

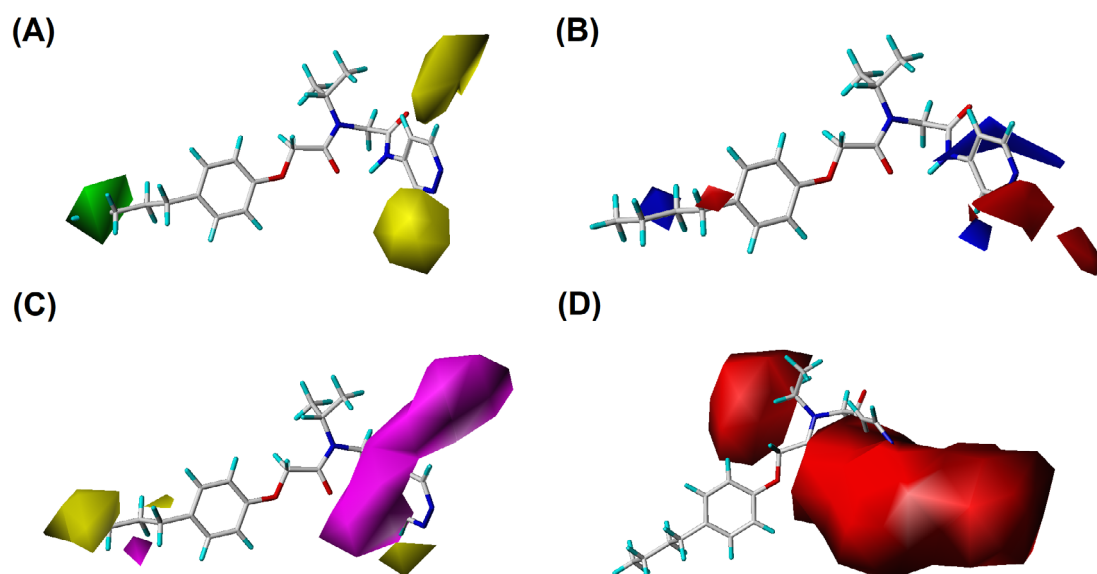


Figure 5. The CoMSIA contour maps of template compound 20. (A) The S field; (B) the E field; (C) the H field; (D) the HA field

For the CoMSIA model, the H field contour map was shown in [Figure 5C](#). The yellow block indicates that the H group introduced in the area can enhance the activity of compounds, while the magenta block indicates that the hydrophilic groups introduced in the area are helpful for increasing bioactivity. It could be seen from [Figure 5C](#) that a yellow contour was embedded in the R¹ area, indicating that the H group here was advantageous. There was a huge magenta block in the R³ area, showing that the hydrophilic group here was favorable.

In the HA field of the CoMSIA model shown in [Figure 5D](#), magenta blocks indicate that the introduction of HA groups is valuable to increase the activity of the compound, while red blocks are the opposite. There was a medium-sized red color block in the R² area, and a large red contour embedded in the R³ area, indicating that the HD groups in the R² and R³ regions were advantageous for improving the bioactivity of the compound.

Design of novel molecules

By analyzing the 3D-QSAR of a series of non-covalent proteasome inhibitors, some key clues about structural modifications to improve bioactivity were obtained: 1) The R¹ area was mainly surrounded by the green contour in the S fields of the CoMFA and CoMSIA models, the blue block in the E field, and the yellow block in the H field of the CoMSIA model. These results showed that the addition of bulky, positive, or H groups in the R¹ region was beneficial. 2) The R² region was mainly surrounded by a red contour in the HA field of the CoMSIA model, indicating that the introduction of the HD here was helpful. 3) The R³ area had the appearance of color block in four fields, namely, the yellow block in the S field, the red contour in the E field, the magenta block in the H field, and the red contour in the HA field, which represented the addition of small group, negative charge, hydrophilic, or HD group in this area was favorable. The main SAR information of non-covalent proteasome inhibitors is illustrated in [Figure 6](#).

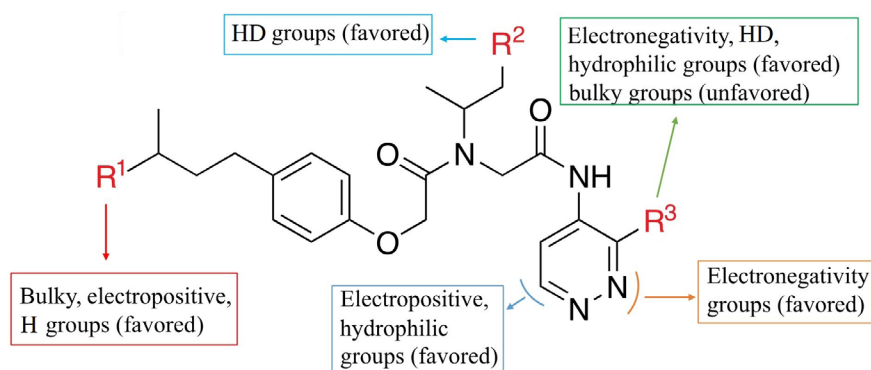
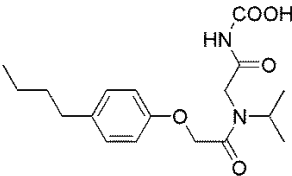


Figure 6. SAR information obtained from 3D-QSAR study

Based on the above clues, a series of novel non-covalent proteasome inhibitors were designed and evaluated using the highest bioactivity compound 20 as the template (Figure 6). Finally, twenty-four new inhibitors (compound D01–D24) were screened out with higher Pred. than the template. The chemical structures of these novel inhibitors and their pIC_{50} values predicted by the constructed QSAR models are shown in Table 3.

Table 3. The structures and activities of newly designed non-covalent proteasome inhibitors

ID	Substituent			Predicted pIC_{50}		Synthetic accessibility	Drug-likeness Lipinski
	R ¹	R ²	R ³	CoMFA	CoMSIA		
C20	H	H	H	7.303	7.164	3.02	Yes
D1	Me	H	H	7.311	7.306	3.14	Yes
D2	H	CHO	H	7.369	7.233	3.56	Yes
D3	H	COOH	H	7.322	7.233	3.63	Yes
D4	H	NH ₂	H	7.317	7.207	3.56	Yes
D5	H	H	F	7.394	7.190	3.20	Yes
D6	H	H	CHO	7.436	7.613	3.18	Yes
D7	H	H	COOH	7.435	7.652	3.29	Yes
D8				7.323	7.287	2.61	Yes



D9	Me	H	F	7.416	7.331	3.32	Yes
D10	Me	H	CHO	7.498	7.756	3.30	Yes
D11	Me	H	COOH	7.432	7.792	3.41	Yes
D12	H	COOH	F	7.425	7.258	3.79	Yes
D13	H	COOH	CHO	7.466	7.680	3.79	Yes
D14	H	COOH	COOH	7.458	7.718	3.89	Yes
D15	H	NH ₂	F	7.388	7.232	3.73	Yes
D16	H	NH ₂	CHO	7.437	7.649	3.72	Yes
D17	H	NH ₂	COOH	7.431	7.688	3.82	Yes
D18	Me	CHO	COOH	7.349	7.792	3.94	Yes
D19	Me	COOH	F	7.346	7.364	3.91	Yes
D20	Me	COOH	CHO	7.450	7.788	3.90	Yes
D21	Me	COOH	COOH	7.362	7.823	4.01	Yes
D22	Me	NH ₂	F	7.442	7.373	3.85	Yes
D23	Me	NH ₂	CHO	7.522	7.794	3.83	Yes
D24	Me	NH ₂	COOH	7.468	7.828	3.94	Yes

Molecular docking analysis

In order to better explain the relevant docking results, three representative compounds for detailed descriptions were selected. The results of docking are clearly illustrated in Figure 7. How the least biologically active compound 15, the most biologically active compound 20, and the newly designed compound with the highest predictive activity D24 bound with the receptor protein 3MG6 are shown in Figure 7A, B, and C, respectively. At the same time, their molecular docking scores were -6.3 kcal/mol (compound 15), -8.3 kcal/mol (compound 20), and -8.9 kcal/mol (compound D24), respectively. The lower the molecular docking score, the stronger the binding ability of the ligand to the receptor [54]. And it can be seen that the size of the binding force is consistent with the size of the biological activity value. The binding mode of compound 15 which is the least active molecule in the dataset is clearly shown in Figure 7A. This inhibitor afforded some interactions with the proteasome. It could be seen that the oxygen atom of compound 15 near the R² region presented a hydrogen bond with alanine 22 ($-O \cdots HN$, 3.0 Å). Furthermore, the oxygen atom was also

bound with the amino acid residue of asparagine 24 by the formation of a hydrogen bond ($-O\cdots HN$, 3.1 Å). In addition, there was a π - π stacking between tryptophane 25 and the benzene ring of compound 15. Several H interactions in the binding site were also found. However, when the docking result of compound 20 (the most active inhibitor) was analyzed, there were stronger interactions in the binding pocket which could explain why the activity of compound 20 was higher than compound 15. As could be seen from Figure 7B, the oxygen atom near the R² region of compound 20 (in the same position as compound 15) formed a hydrogen bond with tryptophane 25 ($-O\cdots HN$, 3.4 Å). The second one was that the NH group of compound 20 near the R³ region interacted with alanine 27 ($-O\cdots HN$, 3.8 Å). The third one was the oxygen atom bound with glycine 128 and there were other hydrogen bond interactions with key residues like serine 112, serine 118, and aspartic acid 114, respectively. Compared with compound 15, compound 20 had more hydrogen bond interactions with receptor which would be vital for the binding stability of inhibitor in the active site. Meanwhile, there was the same π - π stacking interaction as compound 15. In addition, the H interaction between compound 20 and key residues such as tryptophane 25, histidine 98, and aspartic acid 114 further enhanced the bioactivity. These multi-conjugate effects revealed that compound 20 (IC_{50} = 49 nmol/L) had stronger stability in the active pocket of the receptor and higher activity than compound 15 (IC_{50} = 5,907 nmol/L), which is consistent with the experimental result.

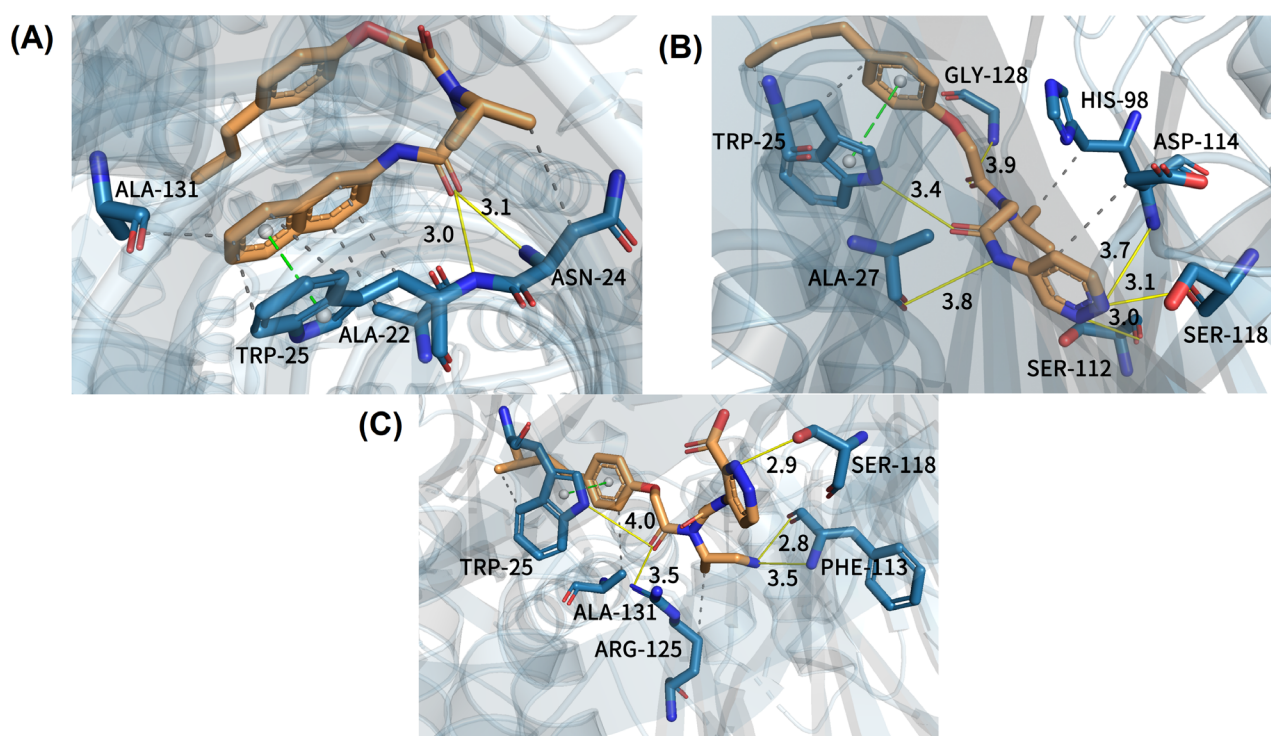


Figure 7. Docking results of three representative compounds in the binding site of the proteasome. (A) Least active compound 15; (B) most active compound 20; (C) new designed compound D24. Yellow solid line: hydrogen bond; green dotted line: π - π stacking; gray dotted line: H interaction

In order to better demonstrate that the new designed non-covalent proteasome inhibitors had higher biological activity than the template molecule (compound 20), the molecular docking method was applied for studying their binding mode with the receptor. New designed compound D24 with the predicted most biological active values was taken as an example. In Figure 7C, we noticed that compound D24 had the same hydrogen bond interaction as compound 20 where they interacted with the same residue (tryptophane 25), but the distinction between them was the hydrogen bond distance. For compound D24, the oxygen atom was far away from the surrounding residue, which was bad for ligand-protein interactions. It was noteworthy that there was a stronger 2.8 Å-hydrogen bond interaction between the NH group of compound D24 (at R² position) and phenylalanine 113 ($-O\cdots HN$, 2.8 Å), which made the ligand-protein interaction binding more stable. Additionally, the pyridazine ring of compound D24 formed another stable hydrogen bond with serine 118 ($-O\cdots HN$, 2.9 Å). Except that, compound D24 was bound with arginine 125 and phenylalanine 113,

respectively, to generate other hydrogen bond interactions. Generally speaking, compound D24 had a closer hydrogen bonding distance than the template molecule, so it could more tightly bind with the receptor to have better inhibitor activity. In addition, we found that all three representative compounds formed π - π stacking with an indole ring of tryptophane 25, which implied it may be a key residue in the active pocket. Of course, the H interactions also existed in the compound D24 to enhance the binding with the protein. In summary, all of these results clearly indicated that compound D24 had high stability at the binding site, followed by a high inhibitory activity.

Results of ADMET and drug-likeness

ADMET is a complex process and each link may have a great impact on the effectiveness and safety of drugs, so it is an important index to evaluate the drug properties of compounds in the process of new drug development and clinical use. Therefore, the online predictive evaluation of the newly designed compound (D01–D24) was performed. The results of statistical parameters for ADMET prediction of the newly designed non-covalent proteasome inhibitors are shown in Table 4. The results of drug-likeness and synthesis accessibility are shown in Table 3.

Table 4. ADMET properties of novel designed compounds

ID	Absorption	Distribution	Metabolism CYP (substrate)				Metabolism CYP (inhibitor)				Excretion		Toxicology		
			Intestinal absorption (human; % absorbed)	VD _{ss} (human; log L/kg)	2D6	3A4	1A2	2C19	2C9	2D6	3A4	Total clearance (log mL/min per kg)	AMES	Hepatotoxicity	Skin sensitization
C20	93.041	-0.45	No	Yes	No	Yes	Yes	No	No	Yes	1.354	No	Yes	No	
D1	92.849	-0.426	No	Yes	No	Yes	Yes	No	No	Yes	1.227	No	Yes	No	
D2	75.759	-0.506	No	Yes	No	No	No	No	No	Yes	1.421	No	Yes	No	
D3	53.994	-1.385	No	No	No	No	No	No	No	No	1.222	No	Yes	No	
D4	68.354	0.248	No	No	No	No	No	No	No	No	0.998	No	Yes	No	
D5	92.105	-0.551	No	Yes	No	Yes	Yes	No	No	Yes	1.206	No	Yes	No	
D6	70.044	-0.534	No	Yes	No	Yes	Yes	No	No	Yes	1.351	No	Yes	No	
D7	45.352	-1.07	No	No	No	No	No	No	No	No	1.109	No	Yes	No	
D8	50.399	-1.261	No	No	No	No	No	No	No	No	1.313	No	Yes	No	
D9	50.399	-1.261	No	No	No	No	No	No	No	No	1.313	No	Yes	No	
D10	70.203	-0.518	No	Yes	No	Yes	Yes	No	No	Yes	1.223	No	Yes	No	
D11	45.51	-1.066	No	No	No	No	No	No	No	No	0.981	No	Yes	No	
D12	50.741	-1.417	No	No	No	No	No	No	No	No	1.074	No	Yes	No	
D13	45.056	-1.442	No	No	No	No	No	No	No	No	1.218	No	Yes	No	
D14	20.297	-1.516	No	No	No	No	No	No	No	No	0.976	No	Yes	No	
D15	65.101	0.133	No	No	No	No	No	No	No	No	0.864	No	Yes	No	
D16	59.416	0.088	No	No	No	No	No	No	No	No	0.977	No	Yes	No	

Table 4. ADMET properties of novel designed compounds (*continued*)

ID	Absorption	Distribution	Metabolism CYP (substrate)				Metabolism CYP (inhibitor)				Excretion	Toxicology			
			2D6	3A4	1A2	2C19	2C9	2D6	3A4	Total clearance (log mL/min per kg)		AMES	Hepatotoxicity	Skin sensitization	
	Intestinal absorption (human; % absorbed)	VDss (human; log L/kg)													
D17	34.657	-0.65	No	No	No	No	No	No	No	No	No	0.811	No	Yes	No
D18	42.22	-1.202	No	No	No	No	No	No	No	No	No	1.048	No	Yes	No
D19	50.9	-1.39	No	No	No	No	No	No	No	No	No	0.946	No	Yes	No
D20	45.215	-1.41	No	No	No	No	No	No	No	No	No	1.09	No	Yes	No
D21	20.455	-1.477	No	No	No	No	No	No	No	No	No	0.848	No	Yes	No
D22	65.26	0.138	No	No	No	No	No	No	No	No	No	0.723	No	Yes	No
D23	59.575	0.098	No	No	No	No	No	No	No	No	No	0.837	No	Yes	No
D24	34.815	-0.648	No	No	No	No	No	No	No	No	No	0.671	No	Yes	No

CYP: cytochrome P450; VDss: volume of distribution steady-state

The intestinal absorptions of the newly designed compounds (D01–D24) are all greater than 30%, except compounds D14 and D21, which indicates that they will be highly absorbed by the human intestine. In particular, the highest intestinal absorption of compound D1 is 92.849%, suggesting that the novel non-covalent proteasome inhibitors have better absorption capacity as an oral drug. Enzymatic metabolism of the newly designed compound is assessed by whether it is a substrate or inhibitor of CYP. Considering that there are different enzymes in the family of CYP enzymes in human cells, CYP2D6 and CYP3A4 are two remarkable isoforms of P450 responsible for drug metabolism. As can be seen from Table 4, all the designed compounds are neither substrates nor inhibitors of CYP2D6. Compounds D1, D2, D5, D6, and D10 are substrates of CYP3A4, suggesting that they can be metabolized by CYP3A4. The other compounds are not CYP3A4 inhibitors, meaning that they may not affect normal metabolism. In terms of drug clearance, the Pred. indicates that all newly designed compounds could be cleared by combined hepatorenal clearance. Candidate drugs are supposed to be non-toxic to humans, and the AMES test is a bacteria short-term test that determines whether a chemical has mutagenic potential. Fortunately, none of the designed compounds are toxic to AMES and induce skin sensitization. The designed compounds are hepatotoxic, which may impair the normal function of the liver. According to the results of ADMET prediction, we may theoretically assume that the newly designed molecules have good PK properties.

Drug-likeness qualitatively assesses the likelihood of a compound becoming an oral drug from the perspective of bioavailability. As shown in Table 3, all novel compounds fulfill the Lipinski rule, which means they meet the criteria for being oral drugs. In addition, it is worth noting that the newly designed compounds D01–D24 have a synthetic accessibility range of 2.61–4.01, suggesting that they are not difficult to synthesize. These results could provide researchers with valuable information on drug synthesis and screening.

Discussion

A growing body of research suggests that proteasome inhibitors could be valuable drugs for the treatment of neoplastic diseases [55]. Most of the 20S proteasome inhibitors reported in the literature are peptide-based compounds with C-terminal electrophilic warheads that form covalent additions to the active site Thr10 γ . However, this mode of covalent action, together with the high reactivity of the compound, can lead to off-target interactions [56]. To overcome these shortcomings, researchers have looked for a number of inhibitors with different mechanisms of action, including non-covalent proteasome inhibitors. Non-covalent inhibitors may exhibit unique advantages over covalent inhibitors. For example, non-covalent inhibitors are usually reversible and reduce toxicity due to their non-cumulative proteasome inhibition [57, 58]. Off-target effects can be limited by non-covalent inhibition of regulatory enzyme activity, thereby reducing drug side effects [59]. Thus, there is a recent tendency to identify non-covalent inhibitors, mainly including peptides, pseudopeptides, and some organic compounds [60]. Compared with covalent inhibitors, our rationally designed novel phenol ether derivatives as non-covalent proteasome inhibitors have the effect of selectively inhibiting the β 5 site of the proteasome and they may provide an alternative mechanism for proteasome inhibition.

In this work, computer-aided drug design methods, including the QSAR model, molecular docking, ADMET prediction, and drug-likeness prediction, were used to systematically study the theoretical SARs of 28 phenol ether derivatives to design novel non-covalent proteasome inhibitors for tumor treatment. The optimal QSAR models were established (CoMFA-SE: $Q^2 = 0.574$, $r^2 = 0.999$, $r^2_{\text{pred}} = 0.755$; CoMSIA-SEHA: $Q^2 = 0.584$, $r^2 = 0.989$, $r^2_{\text{pred}} = 0.921$). In addition, the results showed that the created models were robust and could be used to predict the bioactivity of the newly designed compounds. By comprehensively analyzing the contour maps of CoMFA and CoMSIA models, some key information about the structural modifications that can significantly enhance the molecular activity were found: 1) introducing bulky, positive, or H groups in the R¹ region; 2) introducing HD group in R² area; 3) introducing small group, electronegative, hydrophilic, or HD groups in the R³ region. Based on the SAR information, we designed, predicted, evaluated, and finally screened 24 novel non-covalent proteasome inhibitors (D01–D24). The docking results between compound 15 (the least active molecule), 20 (the most active molecule), D24, and proteasome (PDB ID: 3MG6) showed that compound D24 had high stability in the binding site of the protein. In addition, the prediction results of ADMET and drug-likeness properties indicated that the novel-designed compounds had rational PKs and drug-likeness properties. In general, all newly designed non-covalent proteasome inhibitors had good bioactivity, binding stability, ADMET, and drug-likeness characteristics.

In our study, SAR on the phenol ether derivatives as non-covalent proteasome inhibitors has been extensively explored. The computational studies show a non-covalent binding mode. It also provides a new chemical template for non-covalent proteasome inhibitors, providing good insights for future research on structural modification and synthesis of more efficient and selective proteasome inhibitors to improve potency and subunit selectivity. Of course, the accuracy of these predictions also needs to be verified by experiments in the future. This study could lead to the discovery of anti-tumor drugs with higher inhibitory effects.

Abbreviations

3D: three-dimensional

ADMET: absorption, distribution, metabolism, excretion, and toxicology

CoMFA: comparative molecular field analysis

CoMSIA: comparative molecular similarity indices analysis

CYP: cytochrome P450

E: electrostatic

H: hydrophobic

HA: H-bond acceptor

HD: H-bond donor
IC₅₀: half maximal inhibitory concentration
ONCs: optimum number of components
PDB: protein data bank
PK: pharmacokinetic
PLS: partial least-squares
QSAR: quantitative structure-activity relationship
S: steric
SAR: structure-activity relationship
SEE: standard error of estimate
UPS: ubiquitin-proteasome system

Supplementary materials

The supplementary material for this article is available at: https://www.explorationpub.com/uploads/Article/file/100829_sup_1.pdf.

Declarations

Author contributions

MY: Conceptualization, Investigation, Validation, Writing—original draft, Writing—review & editing. HJ: Conceptualization, Data curation, Validation. FS: Conceptualization, Investigation. QC: Conceptualization, Investigation, Formal analysis. PC: Conceptualization, Project administration, Writing—review & editing, Supervision. All authors read and approved the submitted version.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Ethical approval

Not applicable.

Consent to participate

Not applicable.

Consent to publication

Not applicable.

Availability of data and materials

All relevant data is contained within the manuscript.

Funding

Not applicable.

Copyright

© The Author(s) 2023.

References

1. Tanahashi N, Murakami Y, Minami Y, Shimbara N, Hendil KB, Tanaka K. Hybrid proteasomes: induction by interferon- γ and contribution to ATP-dependent proteolysis. *J Biol Chem*. 2000;275:14336–45.
2. Voges D, Zwickl P, Baumeister W. The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu Rev Biochem*. 1999;68:1015–68.

3. Hölzl H, Kapelari B, Kellermann J, Seemüller E, Sümegi M, Udvardy A, et al. The regulatory complex of *Drosophila melanogaster* 26s proteasomes. Subunit composition and localization of a deubiquitylating enzyme. *J Cell Biol.* 2000;150:119–30.
4. Hough R, Pratt G, Rechsteiner M. Purification of two high molecular weight proteases from rabbit reticulocyte lysate. *J Biol Chem.* 1987;262:8303–13.
5. Inobe T, Genmei R. N-Terminal coiled-coil structure of ATPase subunits of 26S proteasome is crucial for proteasome function. *PLoS One.* 2015;10:e0134056.
6. Bard JAM, Goodall EA, Greene ER, Jonsson E, Dong KC, Martin A. Structure and function of the 26S proteasome. *Annu Rev Biochem.* 2018;87:697–724.
7. Adams J. The development of proteasome inhibitors as anticancer drugs. *Cancer Cell.* 2004;5:417–21.
8. Chen L, Madura F. Increased proteasome activity, ubiquitin-conjugating enzymes, and eEF1A translation factor detected in breast cancer tissue. *Cancer Res.* 2005;65:5599–606.
9. Voutsadakis IA. Proteasome expression and activity in cancer and cancer stem cells. *Tumor Biol.* 2017;39:1010428317692248.
10. Bilalis AG, Papadimitriou K, Pouli A, Papanastasiou K, Tsakanikas S, Stefanitsi P, et al. Bortezomib in multiple myeloma: treatment and retreatment. A single center experience. *Blood.* 2007;110:4819.
11. Kortuem KM, Stewart AK. Carfilzomib. *Blood.* 2013;121:893–7.
12. Richardson PG, Zweegman S, O'Donnell EK, Laubach JP, Raje N, Voorhees P, et al. Ixazomib for the treatment of multiple myeloma. *Expert Opin Pharmacother.* 2018;19:1949–68.
13. Borissenko L, Groll M. 20S proteasome and its inhibitors: crystallographic knowledge for drug development. *Chem Rev.* 2007;107:687–717.
14. Löwe J, Stock D, Jap B, Zwickl P, Baumeister W, Huber R. Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution. *Science.* 1995;268:533–9.
15. Prudhomme J, McDaniel E, Ponts N, Bertani S, Fenical W, Jensen P, et al. Marine actinomycetes: a new source of compounds against the human malaria parasite. *PLoS One.* 2008;3:e2335.
16. Lin G, Li D, de Carvalho LP, Deng H, Tao H, Vogt G, et al. Inhibitors selective for mycobacterial *versus* human proteasomes. *Nature.* 2009;461:621–6.
17. Zhou J, Chng WJ. Novel mechanism of drug resistance to proteasome inhibitors in multiple myeloma. *World J Clin Oncol.* 2019;10:303–6.
18. Maccari R, Ettari R, Adornato I, Naß A, Wolber G, Bitto A, et al. Identification of 2-thioxoimidazolidin-4-one derivatives as novel noncovalent proteasome and immunoproteasome inhibitors. *Bioorg Med Chem Lett.* 2018;28:278–83.
19. McDaniel TJ, Lansdell TA, Dissanayake AA, Azevedo LM, Claes J, Odom AL, et al. Substituted quinolines as noncovalent proteasome inhibitors. *Bioorg Med Chem.* 2016;24:2441–50.
20. Yu J, Xu L, Hong D, Zhang X, Liu J, Li D, et al. Design, synthesis, and biological evaluation of novel phenol ether derivatives as non-covalent proteasome inhibitors. *Eur J Med Chem.* 2019;161:543–58.
21. Pandey S, Singh BK. *De-novo* drug design, molecular docking and *in-silico* molecular prediction of AChEI analogues through CADD approaches as anti-Alzheimer's agents. *Curr Comput Aided Drug Des.* 2020;16:54–72.
22. Xu K, Wang K, Yang Y, Yan DA, Huang L, Chen CH, et al. Discovery of novel non-covalent inhibitors selective to the β5-subunit of the human 20S proteasome. *Eur J Med Chem.* 2015;98:61–8.
23. Li A, Sun H, Du L, Wu X, Cao J, You Q, et al. Discovery of novel covalent proteasome inhibitors through a combination of pharmacophore screening, covalent docking, and molecular dynamics simulations. *J Mol Model.* 2014;20:2515.
24. Yuan MA, Cheng P, Zhang SP. Structure-activity relationship analysis of a series of nonsteroidal analogues as androgen receptor antagonists. *New J Chem.* 2021;45:1176–86.

25. Clark M, Cramer RD III, Van Opdenbosch N. Validation of the general purpose tripos 5.2 force field. *J Comput Chem*. 1989;10:982–1012.
26. Gasteiger J, Marsili M. Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. *Tetrahedron*. 1980;36:3219–28.
27. Klebe G, Abraham U, Mietzner T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *J Med Chem*. 1994;37:4130–46.
28. Zhang Z, Cheng P, Zhu Y, Xia Q, Zhang S. 3D-QSAR analysis of a series of 1,2,3-triazole-chromenone derivatives as an acetylcholinesterase inhibitor against Alzheimer's disease. *Chin J Struc Chem*. 2020;39:1235–42.
29. Cho SJ, Tropsha A. Cross-validated R^2 -guided region selection for comparative molecular field analysis: a simple method to achieve consistent results. *J Med Chem*. 1995;38:1060–6.
30. Wang J, Yang Y, Li Y, Wang Y. Computational study exploring the interaction mechanism of benzimidazole derivatives as potent cattle bovine viral diarrhea virus inhibitors. *J Agric Food Chem*. 2016;64:5941–50.
31. Gao Y, Chen Y, Tian Y, Zhao Y, Wu F, Luo X, et al. *In silico* study of 3-hydroxypyrimidine-2,4-diones as inhibitors of HIV RT-associated RNase H using molecular docking, molecular dynamics, 3D-QSAR, and pharmacophore models. *New J Chem*. 2019;43:17004–17.
32. Mercado J, Gómez H, Vivas-Reyes R. Comparative molecular field analysis and comparative molecular similarity indices analysis studies of α -ketothiazole arginine analogues inhibitors of coagulation factor XIa. *New J Chem*. 2011;35:820–32.
33. Han W, Li J. Structure-activity relationship analysis of 3-phenylpyrazole derivatives as androgen receptor antagonists. *J Biomol Struct Dyn*. 2020;38:2582–91.
34. Stähle L, Wold S. 6 Multivariate data analysis and experimental design in biomedical research. *Prog Med Chem*. 1988;25:291–338.
35. Good AC, Peterson SJ, Richards WG. QSAR's from similarity matrices. Technique validation and application in the comparison of different similarity evaluation methods. *J Med Chem*. 1993;36:2929–37.
36. Balasubramanian PK, Balupuri A, Gadhe CG, Cho SJ. 3D QSAR modeling study on 7-aminofuro [2,3-c] pyridine derivatives as TAK1 inhibitors using CoMFA and CoMSIA. *Med Chem Res*. 2015;24:2347–65.
37. Wold S. Validation of QSAR's. *Quant Struc Act Rel*. 1991;10:191–3.
38. Cao D, Deng Z, Zhu M, Yao Z, Dong J, Zhao R. Ensemble partial least squares regression for descriptor selection, outlier detection, applicability domain assessment, and ensemble modeling in QSAR/QSPR modeling. *J Chemom*. 2017;31:e2922.
39. Kim JH, Jeong JH. Structure-activity relationship studies based on 3D-QSAR CoMFA/CoMSIA for thienopyrimidine derivatives as triple negative breast cancer inhibitors. *Molecules*. 2022;27:7974.
40. Hadni H, Elhallaoui M. 3D-QSAR, docking and ADMET properties of aurone analogues as antimalarial agents. *Heliyon*. 2020;6:e03580.
41. Roy PP, Paul S, Mitra I, Roy K. On two novel parameters for validation of predictive QSAR models. *Molecules*. 2009;14:1660–701.
42. Mitra I, Roy PP, Kar S, Ojha PK, Roy K. On further application of r_m^2 as a metric for validation of QSAR models. *J Chemom*. 2010;24:22–33.
43. Ferreira LG, Dos Santos RN, Oliva G, Andricopulo AD. Molecular docking and structure-based drug design strategies. *Molecules*. 2015;20:13384–421.
44. Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. *J Mol Biol*. 1997;267:727–48.
45. Eberhardt J, Santos-Martins D, Tillack AF, Forli S. AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings. *J Chem Inf Model*. 2021;61:3891–8.

46. Yang X, Liu H, Liu J, Li F, Li X, Shi L, et al. Rational selection of the 3D structure of biomacromolecules for molecular docking studies on the mechanism of endocrine disruptor action. *Chem Res Toxicol*. 2016;29:1565–70.
47. Blackburn C, Gigstad Kenneth M, Hales P, Garcia K, Jones M, Bruzzese Frank J, et al. Characterization of a new series of non-covalent proteasome inhibitors with exquisite potency and selectivity for the 20S β 5-subunit. *Biochem J*. 2010;430:461–76. Erratum in: *Biochem J*. 2010;431:433.
48. Solis FJ, Wets RJB. Minimization by random search techniques. *Math Oper Res*. 1981;6:19–30.
49. Waring MJ, Arrowsmith J, Leach AR, Leeson PD, Mandrell S, Owen RM, et al. An analysis of the attrition of drug candidates from four major pharmaceutical companies. *Nat Rev Drug Discov*. 2015;14:475–86.
50. Sun D, Gao W, Hu H, Zhou S. Why 90% of clinical drug development fails and how to improve it? *Acta Pharm Sin B*. 2022;12:3049–62.
51. Ferreira LLG, Andricopulo AD. ADMET modeling approaches in drug discovery. *Drug Discov Today*. 2019;24:1157–65.
52. Pires DE, Blundell TL, Ascher DB. pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J Med Chem*. 2015;58:4066–72.
53. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep*. 2017;7:42717.
54. Vidal-Limon A, Aguilar-Toalá JE, Liceaga AM. Integration of molecular docking analysis and molecular dynamics simulations for studying food proteins and bioactive peptides. *J Agric Food Chem*. 2022;70:934–43.
55. Li Z, Zou CB, Yao Y, Hoyt MA, McDonough S, Mackey ZB, et al. An easily dissociated 26 S proteasome catalyzes an essential ubiquitin-mediated protein degradation pathway in *Trypanosoma brucei**210. *J Biol Chem*. 2002;277:15486–98.
56. Groll M, Ditzel L, Löwe J, Stock D, Bochtler M, Bartunik HD, et al. Structure of 20S proteasome from yeast at 2.4Å resolution. *Nature*. 1997;386:463–71.
57. Gaczynska M, Osmulski PA. Characterization of noncompetitive regulators of proteasome activity. *Methods Enzymol*. 2005;398:425–38.
58. Anbanandam A, Albarado DC, Tirziu DC, Simons M, Veeraraghavan S. Molecular basis for proline- and arginine-rich peptide inhibition of proteasome. *J Mol Biol*. 2008;384:219–27.
59. Chitra S, Nalini G, Rajasekhar G. The ubiquitin proteasome system and efficacy of proteasome inhibitors in diseases. *Int J Rheum Dis*. 2012;15:249–60.
60. Kaffy J, Bernadat G, Ongeri S. Non-covalent proteasome inhibitors. *Curr Pharm Des*. 2013;19:4115–30.