eFood
Vol. **2**(**6**); December (2021), *pp.* 296–306
DOI: https://doi.org/10.53365/efood.k/147019; eISSN 2666-3066
https://efood.visagaapublishing.com/

Original Research

Xanthine oxidase targeted model setup and its application for antihyperuricemic compounds prediction by *in silico* methods

Chuanli Hou [1], Chuanchao Shi [1], Jiaoyan Ren [1,2,*]

[1] *School of Food Sciences and Engineering, South China University of Technology, China*
[2] *China-Singapore International Joint Research Institute, China*

ARTICLE INFO

Article History

Received 26 January 2022
Accepted 1 March 2022

Keywords

hyperuricemia
molecular docking
Multiple linear regressions (MLR)
Principal component regression (PCR)
Artificial neural network (ANN)

ABSTRACT

To achieve potential alternatives for hyperuricemia therapeutics, a novel structure-docking energy relationship model was established for high-throughput screening inhibitors of xanthine oxidase (XO). Molecular docking was performed between XO and 311 natural compounds from 6 traditional Chinese herbs. Then, structure-docking energy relationship model was simulated between molecular docking energy and 63 molecular descriptors by multiple linear regressions (MLR), principal component regression (PCR), and artificial neural network (ANN), respectively. The results showed that the ANN model was the best model to predict the docking energy of XO with the coefficient of determination (R^2) and mean squared error (MSE) at 0.8746 and 0.9414, respectively. The data of XO inhibitory activity were consistent with the prediction in vitro, which was also further confirmed by hyperuricemia cell model. The results suggested that the structure-docking energy relationship model provides a paradigm framework for the screening of XO inhibitors.

1. INTRODUCTION

Hyperuricemia is a metabolic disease characterized by high serum (or plasma) uric acid levels and is considered an important risk factor for gout [1]. In humans, uric acid is the end product of endogenous and dietary purines [2]. The associated metabolic pathway consists of multiple steps of reactions catalyzed by various enzymes, including adenosine deaminase, purine nucleoside phosphorylase and xanthine oxidase (XO) [3]. Among these, XO catalyzes the conversion of hypoxanthine to xanthine and then xanthine to uric acid [4, 5].

To date, studies have mainly focused on the inhibition of XO in lowering uric acid and XO inhibitors are the mainstay of therapy for reducing serum urate levels in patients with hyperuricemia or gout [6]. However, XO-inhibiting antihyperuricemic drugs result in renal impairment and hypersensitivity reactions, which could be fatal in some cases [7, 8]. Therefore, great diversity and novel chemical structure derived from natural products can be used as potential candidates. Most importantly, the screening of phytochemicals for inhibitors with fewer adverse effects and multi-targets has become the main direction of drug research and development.

Molecular docking is an intelligent calculation method based on the principle of geometric complementarity and energy complementation to evaluate the interaction of small molecule ligands and protein receptors [9]. It is done by setting multiple structural description parameters, simulating the binding sites of specific enzymes and compounds and the results are ranked, which is of great significance in enzyme research and drug design. Simultaneously, molecular docking can also speculate on the mechanism of binding sites, making it useful to screen functional compounds [10–12]. However, analyzing the results of stochastic search methods can sometimes be unclear. Additionally, in silico approaches are used to analyze the chemical structure of fictive molecules and compounds to establish a quantitative structure-activity relationship (QSAR) model, which can be used for high-throughput screening of active molecules [13]. QSAR model and molecular docking methods have been widely used in recent study of drug discovery and development due to their prominent advantages such as timesaving, cost-reduction, high efficiency in in silico screening and prediction of candidate drugs [14]. However, the structure-activity relationship model should be based on molecules with known chemical activity and cannot be applied to molecules with unknown chemical activity. Therefore, for molecules with unknown chemical activity, a suitable model for high-throughput prediction and screening is urgently needed.

In this paper, the interactions between XO and 311 compounds from six traditional Chinese herbal drugs were explored by molecular docking. The molecular descriptors of small molecules were evaluated by PaDEL-Descriptor (a Java software that can calculate molecular descriptors and fingerprints), and the

*Corresponding author. Emails: jyren@scut.edu.cn
Peer review under responsibility of the *International Association of Dietetic Nutrition and Safety*

representative molecular descriptors were screened. Multiple linear regressions (MLR), Principal component regression (PCR), and artificial neural network (ANN) were used to establish the structure-docking energy relationship model of XO. Then, XO inhibitory activity of herbal extracts was used to verify the predicted results. The binding activity of the functional compounds with XO be calculated and predicted directly by building the relationship model for each compound. Moreover, the model also provided some novel insight on the mechanism by which structures exert inhibitory activity, thereby offering guidance for high throughput screening and molecular design of uric acid lowering drugs.

2. MATERIALS AND METHODS

2.1. Dataset

To screen those compounds with unspecified antagonistic activity, a data set of 311 purine metabolic enzymes (XO) - inhibitor compounds were collected from Chinese pharmacopoeia and used for molecular docking and structure-docking energy relationship model in this study. These 311 compounds were from Apium graveolens L. (57), Puerariae Lobatae Radix (64), Taraxacum mongolicum Hand. -Mazz. (63), Cynara scolymus (31), Cirsium japonicum Fisch. ex DC. (39), and Periploca forrestii Schltr. (57) (Supplementary material). The molecular structures were built using ChemDraw Professional 15.1 and optimized using MM1 force field by Chem3D 15.1. The molecular docking-energy of these compounds was used as a dependent variable for further investigation. The whole data set was randomly divided into a training set of 233 compounds (75%) and a test set of 78 compounds (25%). The training set was used to construct structure-docking energy relationship models, and the test set was used as extra independent samples for the validation of the established models. The R language (Version 3.4.1) was used to analyze the data.

2.2. Materials

Human kidney cells (HK-2 cells) were acquired from Wuxi Puhe Biomedical Company (Wuxi, China). Analytical-grade xanthine, 4-aminoantipyrine, adenosine, inosine, hypoxanthine, uric acid, and adenine, were purchased from Aladdin Reagent Int. (Shanghai, China). Phosphate-buffered saline (PBS), 3 - [4, 5-Dimethylthiazol-2-yl]-2, 5-diphenyltetrazolium bromide (MTT), horseradish peroxidase and XO were purchased from Sigma-Aldrich. RMPI 1640, fetal bovine serum (FBS), penicillin, and streptomycin were bought from Gibco Life Technologies (Grand Island, NY).

2.3. Protease receptor selection and molecular docking studies

Molecular docking was carried out using XO as receptors XO. The three-dimensional crystal structures of the XO were downloaded from Protein Data Bank [https://www.rcsb.org/]. The information was listed in Table 1.

Table 1 Information of receptor protease

Name/ Abbreviation	Role in the process of purine metabolism	PDB ID	
Xanthine oxidase/ XO	Catalyze the formation of uric acid by hypoxanthine and xanthine	1N5X	[15, 16]

All non-polar hydrogen atoms (hydrogens attached to carbon atoms) of XO were merged, and partial atomic charges of the molecules were calculated. Heteroatoms and water molecules were eliminated. The grid box of XO of 40 × 40 × 40 points was defined by centering on the ligand in the active site (96.663, 54.963, 39.433). With the above settings, Autodock 4.2 and Autodock Vina were used simultaneously. Autodock Vina has a free energy scoring function, which can estimate the ligand binding orientation and affinity. For molecules with better Auto Dock Vina score, Auto Dock 4.2 and DS were used to study the site of action and the bonding mechanism.

2.4. Calculation of the molecular descriptors

Molecular descriptors are formal mathematical representations of a molecule obtained by a well-specified algorithm which is applied to a defined molecular representation or a well-specified experimental procedure [17]. In this paper, a good number of descriptors were generated for the lowest energy conformer using PaDEL-Descriptor followed by objective feature selection. It is obvious that not every molecular descriptor affects the binding activity of ligand molecules to the receptor protease, and there were numerous of redundant or unrelated descriptors. The existence of independent descriptors increases the complexity of the prediction model. Alternatively, it reduces the weight of key descriptors, which decreases model prediction accuracy, therefore it is necessary to have it filtered. Use the fhcor function in R, and this function can filter the molecular descriptors with high correlation coeffecientin data matrix. The correlation coefficient between molecular descriptors is generally 0.95, but we also set as 0.99 or 0.9, depended on the aim and condition of study. On the basis of the output file from PaDEL-Descriptor, sixty-three chemical descriptors were chosen (**Supplementary Table 1**).

2.5. Construction of structure-docking energy relationship model

For the construction of the structure-docking energy relationship model, MLR, PCA, and ANN were applied.

MLR is a statistical analysis method [18]. The purpose of MLR is to establish a mathematical function, which best depicts the desired activity Y (molecular docking binding energy) as a linear combination of the X-variables (the molecular descriptors). The equation is as follows:

$$Y=\beta_0+\beta_1X_1+\ldots+\beta_pX_p+\varepsilon$$

Where, Y = dependent variable, X_1, X_2, ... X_p are independent variables, i.e., 63 molecular descriptors, ε = stochastic error term, and β_0, β_1, β_2 are the model's parameters to be estimated. R language (Version 3.4.1) was used to analyze the relationship between the binding activities of XO and molecular descriptors, and multiple linear regression models were established. The model

equation was then tested by F (Fisher's statistics) statistic, the predicted residual diagram, Q-Q diagram, location scale diagram, and residual and leveraged graphs.

Principal component analysis (PCA) is a useful method in dealing with irrelevant information in original descriptor matrices [19]. The dataset of PCA procedure can summarize many dispersed continuous descriptors into a few summary PCs, which explains as much of the original data's variance as possible. This tutorial uses the prcomp () function from stats package to do the PCA. The dataset is Cereals. And it was computed the PCA manually to apply the Spectral decomposition theorem. 1) Standardize each column, i.e., subtract mean and divide by sd.2) Compute the correlation matrix for columns 3) Compute eigenvalues and eigenvectors for corr. matrix 4) Each eigenvalue represents the variance captured by the corresponding principal component 5) Each eigenvector represents the loading of the variable along the principal component.

Linear models are not sufficient to explain all the sources of variability due to the complex nature of the relationships between molecular structure and activity. Artificial neural networks (ANN) are a type of machine-learning prediction method with the ability to self-learn relationships from labeled experimental data and generalize to unlabeled situations [20]. One of the most popular types of ANN used in biological research is multilayer perceptron (MLP). The MLP is a feed-forward ANN model that consists of an input layer, hidden layer (s), and an output layer. The neural network is composed of many neurons connected to each other. Different neurons represent different nodes, and different nodes represent different output functions.

2.6. Assay for inhibition of XO activity

Inhibitory effects on XO activity were measured by a decrease in uric acid and in superoxide. Extracts obtained by boiling water extraction and ethanolic extraction from six Chinese traditional herbal drugs were investigated. In brief, for each test tube, 200 μL of various concentrations of extracts and 50 μL XO (0.52 μ/mL) were added and incubated for 10 minutes at 37°C, then 400 μL of 0.22 mmol/L xanthine and 3050 μL of coloration liquid (1 mmol 4-aminoantipyrine, 12 mmol phenol and 7500 U/L horseradish peroxidase dissolved in 50 mmol/l Tris-HCl buffer) were added to these samples and incubated for 20 minutes at 37°C, and reactions were terminated by adding 100 μL of 1 mol/L NaOH, and the absorbance at 508 nm was measured. The system without the addition of extract samples and enzymes is used as the blank control for zero adjustment, and the calculation formula is as follows,

$$\text{Inhibition rate (\%)} = \frac{Ac - (As - Ao)}{Ac} \times 100\%$$

Among them, Ac is the absorbance of the positive blank system without the addition of the extract sample, As is the absorbance of the sample group, and Ao is the absorbance of the sample blank system without the addition of xanthine solution. The IC50 values were calculated from percent inhibition of enzyme activity.

2.7. Detection of antihyperuricemic activity by hyperuricemic cell model

The uric acid-reducing activity of the water and alcohol extracts of six Chinese herbal medicine plants was tested using a hyperuricemia cell model in human proximal tubule cells (HK-2 cells). The establishment method of the cell model refers to our previous research foundation [21]. HK-2 cells were grown to subconfluence in normal growth medium and seeded into 24-well plate at a density of 1×10^5 cells/mL, and then the plates were incubated for 24 h at 37 °C. After that, cells were washed with PBS for three times then control and model groups were changed to new medium, while the sample groups were changed to medium with different concentrations of extracts. Three groups were pre-incubated for 24 h. Then PBS was provided to wash each well three times and adenosine (2.5 mmol/L) in serum-free medium was added in model and sample groups. The control group was maintained in fresh medium in absence of adenosine. After incubation for 30 h, 0.005 u/mg XO was supplied to each well. Cultural supernatants were collected at 12 h after treatment and the amount of UA was measured be HPLC method. The quantitative HPLC separations were conducted on a Microsorb-MV C18 column with photodiode array detection (PDA) at 254 nm. Mobile phase A comprised of 0.52 mM sodium 1-pentanesulfonate and 0.20 M monopotassium phosphate at pH 4.0 by adding phosphoric acid (HPLC-grade). Mobile phase B had the same final concentrations as mobile phase A, apart from the addition of 10% acetonitrile (v/v). The flow rate was 1.0 mL/min, and the system was usually equilibrated for a certain time at that flow rate before the first sample was injected. The injection volume of sample was 20 μL and total run time was 20 min.

2.8. Statistical Analysis

Data were shown as mean ± standard deviation (SD) for three replicates. A one-way analysis of variance (ANOVA) was performed to evaluate the significant differences between the groups using SPSS 20.0. $p<0.05$ was statistically significant.

3. RESULTS AND DISCUSSION

3.1. Molecular docking of XO with natural compounds

First, the molecular docking energy was considered a critical parameter for screening targeted compounds. For compounds with lower docking energy, it showed relatively higher inhibitory activity.

In terms of XO, 311 tested samples from 6 traditional Chinese herbs were subjected to molecular docking, in which, 59 small molecules with a lower docking energy than the positive ligand 1N5X of XO were screened. Of the 59 XO compounds, the best three hits (corylin, luteolin, luteolin-3'-O- beta-D-glucoside) were considered for further evaluation (Fig. 1A). The free energy of binding values for each ligand and their interaction with enzyme residues were presented in Table 2. The results obtained from the molecular docking of XO indicated that amino acid residues

involved in the formation of hydrogen bonds included ARG880, THR1010, ASN768, SER876, THR1010, ARG880, MOS3004, ALA1079, GLU1261, VAL1011, ASP872, and ASP872. The sources and classifications of XO inhibitors were shown in Fig. 1 B-C. Of these ligands, 36.11% were reported in pueraria lobata, 16.67% from dandelion and 16.67% were from cirsium japonicum. In terms of species distribution, 83.05% of them were flavonoids and their glycosides, and 11.86% were polyphenols.

Table 2 Docking Statistics of the Compounds with XO

Receptors	Ligands	Binding Energy(Kcal/mol)	Hydrogen Bond Resides
XO	TEI	-9.0	ARG880, THR1010, ASN768
	Corylin	-11.1	SER876, THR1010, ARG880
	Luteolin	-10.8	MOS3004, ARG880, ALA1079, THR1010, GLU1261, VAL1011
	Luteolin-3'-O-beta-D-glucoside	-10.7	ARG880, MOS3004, ASP872, ALA1079, SER876, THR1010, VAL1011

3.2. Description and analysis of variables

To investigate the effect of the selected molecular descriptors on statistical models, we considered the boxplots of statistical variables (Fig. 2). The original boxplot in Fig. 2A showed that a substantial number of data points were exceeding the upper whisker. There were extreme outliers in variables VR1_DT, VR2_DT, topoRadius and topoDiameter. The discrete anomalous values in the variables indicated that molecular descriptors were different in the different samples. To maintain data integrity and accuracy, it was not necessary to predict and adjust these values temporarily, but to adjust the variables according to the needs of the resulting model.

To further observe the relationship between the 63 independent variables and variables (docking energy with XO), the data were examined and identified using hierarchical cluster analysis (Fig. 2B). Trees indicating variable clustering be observed at the top and left of the diagram. The heatmap showed a data matrix where coloring gave an overview of the numeric differences. Blue indicated that the two sets of data are positively correlated, and green meant negative correlation. Based on the analysis in Fig. 2B, there was a strong positive correlation among a considerable number of variables, and a negative correlation between the response and other variables.

Ten of the 63 variables were randomly selected, and scatter plot matrix analysis was used to observe possible multiple collinearities between the variables. The ten variables were TopoPSA, MW, AMW, WTPT1, WTPT2, WTPT3, WTPT4, WPATH, WPOL, XlogP, and Zagreb. Each small chart showed the correlation between the given pairs of variables (one listed on the right, the other listed above) and contour lines were used for the boxes on the upper hand side of the matrix, and scatter plot for that on the lower left-hand side of the matrix, as shown in Fig. 2C. As shown in Fig. 2C, multiple collinearities existed between variables; for example, WTPT3 and TopoPSA, WTPT4 and TopoPSA, and WTPT3 and WTPT4. Consequently, we chose multiple linear MLR, PCR, and ANN to establish the structure-docking-energy relationship model, respectively.

3.3. Multilinear regression model (MLR)

The evaluation of statistical significance of the regression equations was performed on the basis of the values of the correlation coefficient (R^2), the mean squared error (MSE), Fisher's criterion (F) and the significance level (P). A higher correlation coefficient and a lower mean squared error indicated that the model was more reliable. A p-value smaller than 0.05 showed that the regression equation was statistically significant. The stepwise regression method was used to select the independent variables, and the Akaike information criterion (AIC) information statistic was the criterion. With the decrease in AIC value, the reliability of the model was improved. Variables were gradually deleted or added, the most significant independent variables of the samples were chosen to further modify the model, get the optimal regression equation. The coefficients of each variable in the regression models were shown in Table 3.

MLR has four prerequisites, including linearity, independence, normality, and homogeneity. In the process of constructing regression equations, the residual analysis was used to determine whether the data used for the experiment satisfied the above four conditions, and the residual analysis was also used to determine whether the model needed new independent variables to be introduced.

The residual points obeyed the random distribution, which implied that there was no heteroscedasticity (first) (Fig. 3A). Almost all the residual points were close to or on the Q-Q line, which suggested that residuals were normally distributed (second) (Fig. 3A). The random distribution of residual points around the horizontal band showed that the assumption of constant variance was satisfied (third) (Fig. 3A). There were two leverage points and according to cook's distance, it lies around 50% and 100% Cook's interval, indicating that they had a bit of influence on the model properties of XO model (fourth) (Fig. 3A). It was clear that XO model could explain most of the variability.

The results of the model were evaluated using the coefficient of determination (R^2) and the mean squared error (MSE). From the fitted multiple regression model, the R^2 values of XO models were 0.737. Additionally, based on the F test, the differences among groups were significant. The mean square error (MSE) for the stepwise regression analysis of the XO model was 1.053.

3.4. Principal component regression model (PCR)

PCA is a method that can be used to deal with irrelevant information, unfavorable ratio of the number of descriptors and collinearity among descriptors. PCA was done using the prcomp function in the stats package [22]. PCA revealed 25 components explaining 99.82% of the variance (Table 4).

As evidenced, the first four principal components explain more than 85% of total variance, and the first six principal components explained 90.57% of total variance. principal component regression models were used to analyze the first four components, six components and sixteen components respectively. It was found that the F-test of the regression analysis of the response variable

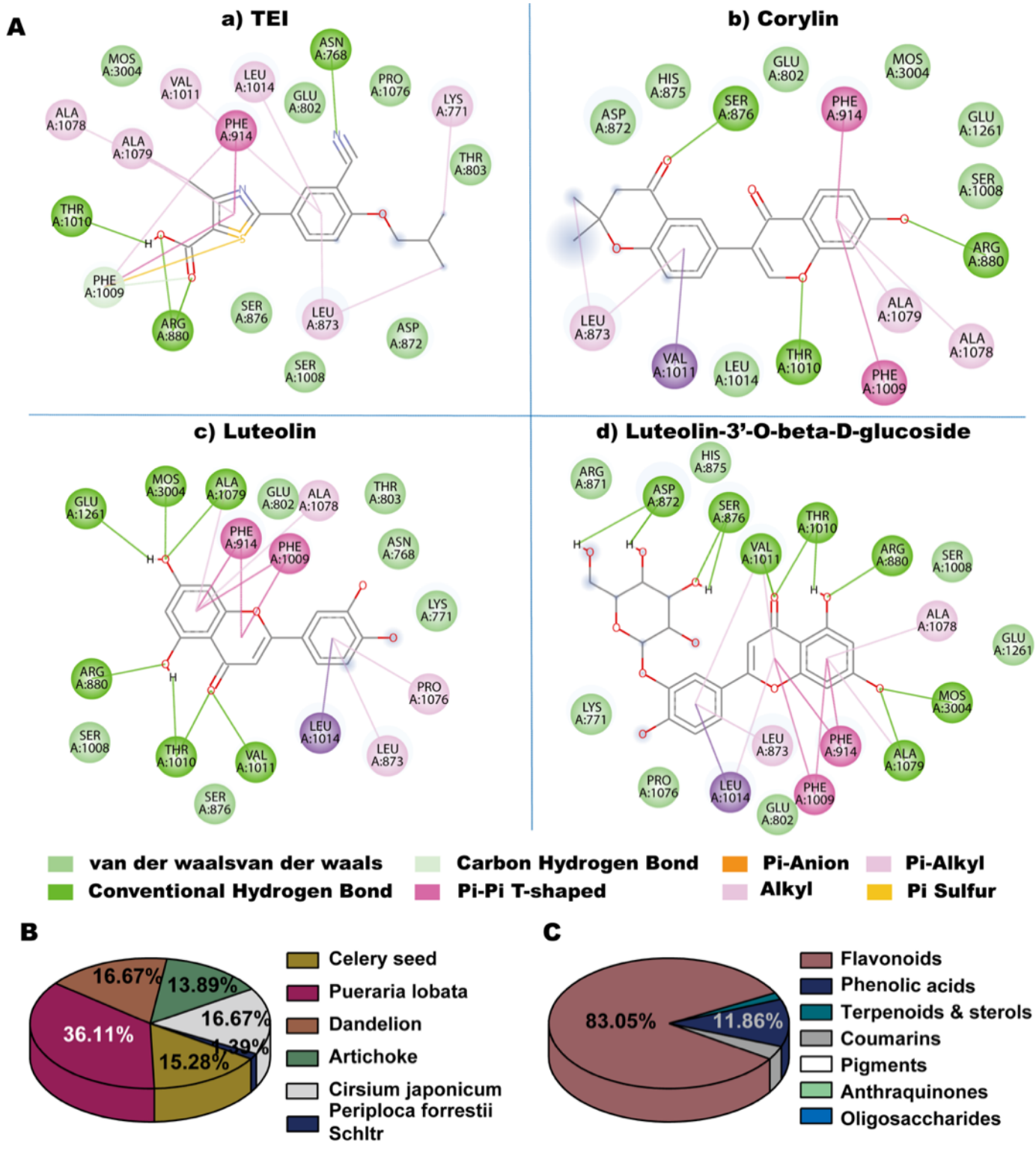

Figure 1 | **Molecular docking of XO with natural compounds.** (A) Molecular docking results of 1N5X (a), corylin (b), luteolin (c) and luteolin-3'-O-beta-D-glucoside (d) with XO.
Proportion of sources (B) and classification (C) of XO inhibitors.

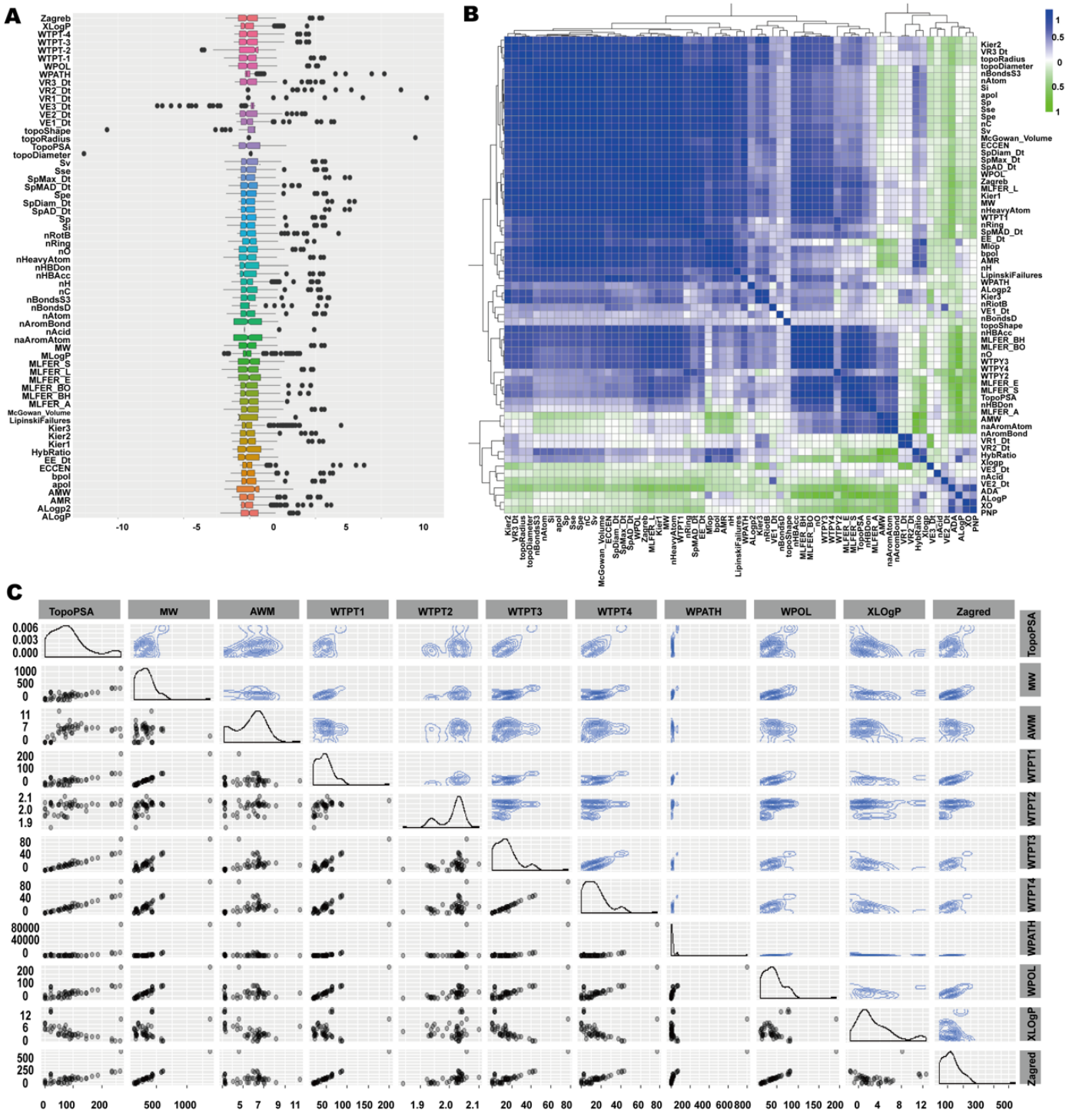

Figure 2 | **Description and analysis of variables.** (A) Box-plot representation of molecular descriptors
(B) Analysis of correlation between molecular descriptors
(C) Scatter plot matrix of ten molecular descriptors.

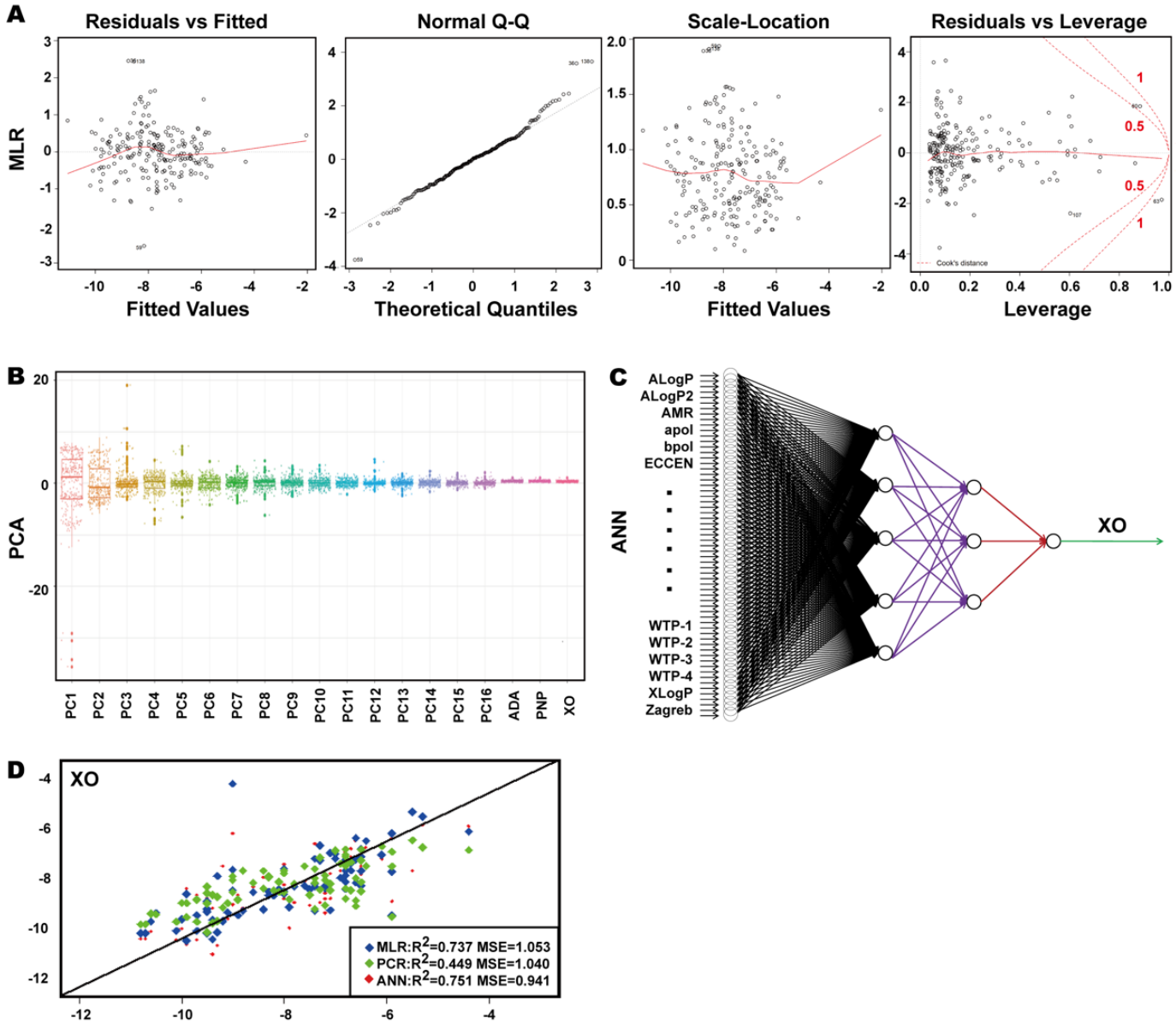

Figure 3 | stablishment of SDER model by MLR, PCR and ANN. (A) Residuals diagnostics plots for XO SDER model.
(B) Boxplot of the response variable and sixteen principal components.
(C) Artificial neural network structure.
(D) Comparison of prediction results and actual values of XO.

was significant and the R^2 increased with the increase of the number in components (Table 4). However, with the analysis of sixteen components (explaining 99.01% of total variance) by principal component regression, the R^2 value was only about 0.489. To observe the discretization of principal components and corresponding dependent variables, the corresponding variable was standardized, and the boxplot of the first sixteen principal components of the PCA was made (Fig. 3B). The distribution of the sample data of the main components PC1, PC2, PC3, PC4, and PC5 was discrete, and there were fewer outliers. Therefore, the principal component regression model was established with the first 16 components. The coefficient of each variable was counted in Table 5.

In the principal component regression of the XO model, a significant correlation with R^2 value was 0.489. As shown, the XO model obtained by conventional PCR had not sufficient quality for modeling. The reason may be that their components did not provide enough information for model building. The mean square error (MSE) for the principal component regression of the XO model was 1.040.

Table 4 | Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% Of variance	Cumulative %	Total	% Of variance	Cumulative %
1	5.979	56.739	56.739	5.979	56.739	56.739
2	3.327	17.565	74.304	3.327	17.565	74.304
3	2.055	6.703	81.007	2.055	6.703	81.007
4	1.706	4.618	85.625	1.706	4.618	85.625
5	1.274	2.574	88.199	1.27	2.574	88.199
6	1.222	2.371	90.570	1.22	2.371	90.570
7	1.045	1.732	92.301	1.044	1.732	92.301
8	0.980	1.524	93.825	0.980	1.524	93.825
9	0.928	1.366	95.192	0.928	1.366	95.192
10	0.847	1.139	96.331	0.847	1.139	96.331
11	0.642	0.654	96.985	0.642	0.654	96.985
12	0.619	0.607	97.593	0.619	0.607	97.593
13	0.558	0.494	98.087	0.558	0.494	98.087
14	0.506	0.407	98.493	0.506	0.407	98.493
15	0.420	0.280	98.773	0.420	0.280	98.773
16	0.385	0.236	99.009	0.385	0.236	99.009
17	0.326	0.169	99.178			
18	0.290	0.134	99.312			
19	0.282	0.126	99.438			
20	0.249	0.099	99.536			
21	0.214	0.073	99.609			
22	0.207	0.068	99.677			
23	0.183	0.053	99.730			
24	0.168	0.045	99.774			
25	0.160	0.041	99.815			

3.5. Artificial neural network (ANN)

In generally, normalization can significantly help to transposes the input variables into the data range that the Sigmoid [0,1] and/or Tanh [1] activation functions lied. In this paper, variable parameters were normalized by the Komaron formula [23]. The dataset of sixty-three independent variables was directly fed into the input layer of the multi layered perception model and the expected result was produced in the output layer (Fig. 3C). 239 of the 319 groups of samples were selected at random as training samples, and the prediction model was established. The remaining 80 groups were used as test sets to verify the generated model. A four-layer artificial neural network model was processed; among them the first layer (input layer) acted as a space for the inputs fed to the network with 63 neurons. The last layer was where the overall mapping of the network input was made available [24]. The second and third layers were hidden layers, and the number of neurons was 5 and 3, respectively. Training with R to achieve the specified convergence accuracy was used. ANN's efficiency was assessed for different MSE and R^2 values. It was found that the values of R^2 of XO model were 0.751, and the MSE value was 0.941.

3.6. Model Performance Criteria

To evaluate the predictive ability and to check the statistical significance of the developed model, the proposed model was used for the prediction of values of the docking energy for the test set (Fig. 3D). The ideal regression line (y = x) was plotted as reference. Corresponding statistical data were reported in Fig. 3D. According to the scatter plot, there was no significant difference between measured data and predicted data of docking energy in the MLR, PCR and ANN models. This figure clearly showed that the obtained MLR, PCR and ANN models worked well over the entire range of the docking energy values. Furthermore, corresponding values of the test set compounds close to the ideal regression line. In addition to a high R^2, a reliable model should also be characterized by a low mean square error (MSE) [18]. Compared to the multilinear regression model and principal component regression model of XO docking energy, the performance of the ANN model was better, with higher correlation coefficient of 0.751, and lower mean square error (0.941) of potential to extract chemically meaningful information (Fig. 3D). Based on the comparison of results from the models in this study, it could be deduced that the prediction of ANN was superior to MLR and PCR for XO docking energy.

3.7. Effects of plant extracts on XO activity and cell model of hyperuricemia

The present results showed that all extracts were effective in inhibiting uric acid production in a dose dependent manner (water extracts of celery seed and dandelion were excluded). IC_{50} values were calculated from linear or no-linear regression lines. The IC_{50} values of different extracts (Fig. 4A) showed that ethanol extract of pueraria lobata seems to have the strongest inhibitory effect ($p < 0.05$) with an IC_{50} of 6.14 ± 0.17 mg/mL. Overall, the XO-inhibitory activity of the ethanol extracts was higher than that of

Table 3 | Coefficient statistics of each variable in XO-MLR model

Variable	XO-MLR Model
Intercept	15.987
ALogP	0.351
ALogp2	0.043
AMR	0.086
apol	49.102
bpol	0.617
ECCEN	0.003
EE_Dt	-0.011
HybRatio	2.371
Kier2	0.593
Kier3	-0.372
LipinskiFailures	0.261
McGowan_Volume	74.308
MLFER_BO	0.817
naAromAtom	-0.870
nAromBond	4.767
nAtom	-46.580
nBondsD	2.179
nBondsS3	4.648
nC	24.346
nHBDon	0.583
nHeavyAtom	58.698
nO	-5.811
nRing	-1.706
nRotB	0.499
SpAD_Dt	0.090
SpDiam_Dt	0.021
SpMAD_Dt	0.312
SpMax_Dt	-0.203
Sse	69.207
Sv	-202.626
topoDiameter	-0.363
VE1_Dt	-4.171
VE2_Dt	50.457
VE3_Dt	0.008
VR3_Dt	0.031
WPATH	-0.0003
WPOL	0.076
WTPT2	-10.512
WTPT3	-1.274
XLogP	-0.242
AMW	NA
MLFER_A	NA
MLFER_S	NA
topoShape	NA
VR1_Dt	NA
WTPT1	NA
Zagreb	NA
MLFER_E	NA
MLFER_L	NA
nAcid	NA
topoRadius	NA
WTPT4	NA

Table 5 | Coefficient statistics of each component in the model

Variable	XO Model
Intercept	0.346
PC1	0.004
PC2	0.016
PC3	0.012
PC4	-0.026
PC5	0.006
PC6	0.042
PC7	-0.026
PC8	-0.006
PC9	-0.003
PC10	-0.012
PC11	0.012
PC12	0.029
PC13	-0.027
PC14	0.050
PC15	-0.003
PC16	0.050

water extracts.

Cell viability was tested in accordance with Amakye et al [25] by an MTT-reagent (Fig. 4B). The effect of sample concentration on the production of uric acid in hyperuricemia model was verified with each plant extract's cell survival rate reaching 85%. Uric acid concentration in the supernatant was determined by reversed-phase high-performance liquid chromatography. According to Fig. 4C, compared to the model group the ethanol extract of pueraria lobata significantly decreased the uric acid content in the supernatant. When the concentration of ethanol extract of pueraria was 0.5 mg/mL and 2.0 mg/mL, the yield of uric acid was 60.28% and 13.14% of the model group, respectively, which were shown in the chromatogram of cell supernatant of the extract.

4. CONCLUSIONS

In this study, molecular docking technique was used to calculate the inhibitory activity of natural product molecules on XO to predict the activity of those extracts, and a structure-docking energy relationship model was used to characterize the relevant parameters of the inhibitory activity. It was found that pueraria lobata had more inhibitory ligands for XO and the extract of pueraria lobata has a better effect on reducing uric acid. In this study, MLR, PCR and ANN models were applied to predict the binding energy and to find the most important descriptors that affect the docking energy. According to the different analyses, the ANN model was the best model to predict the binding energy of XO. Furthermore, in vitro chemical and cell models proved that ethanol extract of pueraria lobata had the best activity of XO inhibition and the effect of reducing uric acid in the cell model, which was in accordance with the molecular docking results. The model can therefore be used for high-throughput screening of XO inhibitors of many models from phytochemicals in an intelligent, fast and reproducible way.

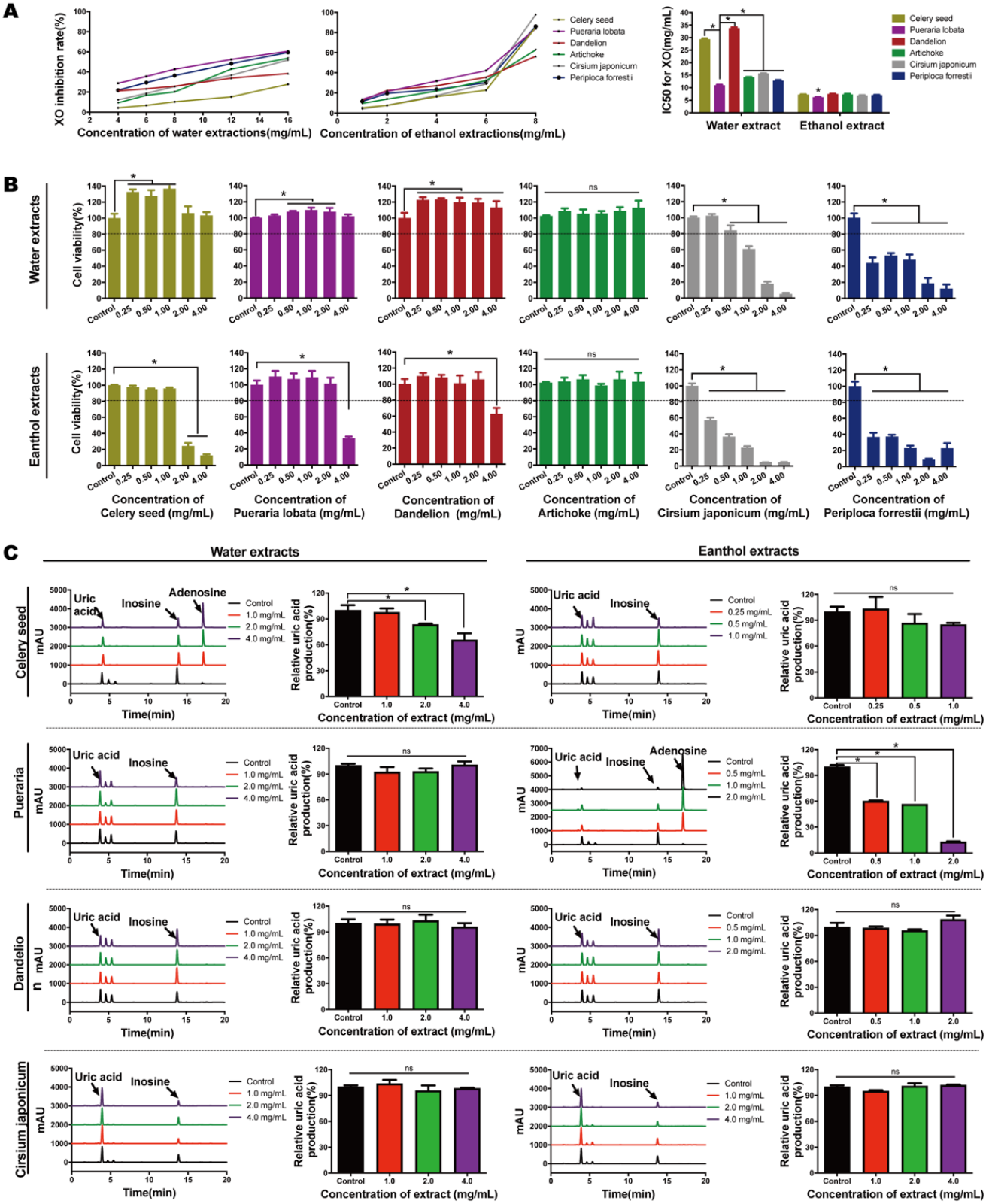

Figure 4 Effects of plant extracts on XO activity and hyperuricemic cell model. (A) XO Activity and IC_{50} values of plants extracts for inhibition of XO

(B) Effect of plant extracts on HK-2 cell viability

(C) Chromatogram of supernatant of plant extracts from celery seed, pueraria, dandelion, cirsium japonicum (left) and relative uric acid yield in cell model (right).

* Indicates $p < 0.05$, and ns indicates $p > 0.05$.

5. ACKNOWLEDGMENTS

The authors gratefully acknowledge the Natural Science Foundation of Guangdong Province of China (2019A1515012230), National Natural Science Foundation of China (No. 32072207).

6. CONFLICTS OF INTEREST

There are no conflicts of interest to declare.

7. SUPPLEMENTARY INFORMATION

Supplementary data to this article can be found online at https://doi.org/10.53365/efood.k/147019.

REFERENCES

[1] Isaka Y, Takabatake Y, Takahashi A, Saitoh T, Yoshimori T. Hyperuricemia-induced inflammasome and kidney diseases. 2016;31:890–896.
[2] Ishikawa T, Aw W, Kaneko K. Metabolic Interactions of Purine Derivatives with Human ABC Transporter ABCG2: Genetic Testing to Assess Gout Risk. Pharmaceuticals (Basel). 2013;6.
[3] Pedley AM, Benkovic SJ. A New View into the Regulation of Purine Metabolism: The Purinosome. Trends in Biochemical Sciences. 2017;42:141–154.
[4] Gliozzi M, Malara N, Muscoli S, Mollace V. The treatment of hyperuricemia. Int J Cardiol. 2016;213:23–27.
[5] Sarawek S, Butterweck V. Effects of various flavonoids on xanthine oxidase activities in vitro and on plasma uric acid levels in oxonate-induced rats. Planta Med. 2006;72:1067–1067.
[6] Rock KL, Kataoka H, Lai JJ. Uric acid as a danger signal in gout and its comorbidities. Nature Reviews Rheumatology. 2013;9:13–23.
[7] Bayat S, Aati O, Rech J, Sapsford M, Cavallaro A, Lell M, et al. Development of a Dual-Energy Computed Tomography Scoring System for Measurement of Urate Deposition in Gout. Arthritis Care Res (Hoboken). 2016;p. 68–68.
[8] Abramowicz M, Zuccotti G, Pflomm JM. Lesinurad/Allopurinol (Duzallo) for Gout-Associated Hyperuricemia; 2018.
[9] Sicho M, Svozil D. Molecular Docking as a Tool to Virtually Develop Drugs. Chem Listy. 2017;111:754–759.
[10] Basciu A, Pietrucci F, Bonvin AM, Vargiu AV. A new protocol to improve the predictive power of molecular docking. European Biophysics Journal with Biophysics Letters. 2017;46.
[11] Piotto S, Biasi LD, Fino R, Parisi R, Sessa L, Concilio S. Yada: a novel tool for molecular docking calculations. Journal of Computer-Aided Molecular Design. 2016;30:753–759.
[12] Ruyck JD, Brysbaert G, Blossey R, Lensink MF. Molecular docking as a popular tool in drug design, an in silico travel. Advances and applications in bioinformatics and chemistry: AABC 9. 2016;.
[13] Pradeep P, Povinelli RJ, White S, Merrill SJ. An ensemble model of QSAR tools for regulatory risk assessment. J Cheminform. 2016;8.
[14] Balupuri A, Balasubramanian PK, Cho SJ. Design of Novel Chemotherapeutic Agents Targeting Checkpoint Kinase 1 Using 3D-QSAR Modeling and Molecular Docking Methods. Curr Comput Aided Drug Des. 2016;12:302–313.
[15] Sato D, Kisen T, Kumagai M, Ohta K. Synthesis, structure-activity relationships, and mechanistic studies of 5-arylazo-tropolone derivatives as novel xanthine oxidase (XO) inhibitors. Biorg Med Chem. 2018;26.
[16] Song JU, Choi SP, Kim TH, Jung CK, Lee JY, Jung SH, et al. Design and synthesis of novel 2-(indol-5-yl)thiazole derivatives as xanthine oxidase inhibitors. Bioorg Med Chem Lett. 2015;25.
[17] Basak SC, Zhu QH, Mills D. Quantitative Structure-Activity Relationships for Anticancer Activity of 2-Phenylindoles Using Mathematical Molecular Descriptors. Curr Comput Aided Drug Des. 2011;7:98–108.
[18] Saghaie L, Sakhi H, Sabzyan H, Shahlaei M, Shamshirian D. Stepwise MLR and PCR QSAR study of the pharmaceutical activities of antimalarial 3-hydroxypyridinone agents using B3LYP/6-311++G** descriptors. Med Chem Res. 2013;22.
[19] Granato D, Santos JS, Escher GB, Ferreira BL, Maggio RM. Use of principal component analysis (PCA) and hierarchical cluster analysis (HCA) for multivariate association between bioactive compounds and functional properties in foods: A critical perspective. Trends Food Sci Technol. 2018;72.
[20] Millan-Ocampo DE, Parrales-Bahena A, Gonzalez-Rodriguez JG, Silva-Martinez S, Porcayo-Calderon J, Hernandez-Perez JA. Modelling of Behavior for Inhibition Corrosion of Bronze Using Artificial Neural Network (ANN). Entropy. 2018;20.
[21] Hou CL, Liu D, Wang M, Gong CC, Li YJ, Yang L, et al. Novel xanthine oxidase-based cell model using HK-2 cell for screening antihyperuricemic functional compounds. Free Radical Biology and Medicine. 2019;136:135–145.
[22] Terhoeven-Urselmans T, Vagen TG, Spaargaren O, Shepherd KD. Prediction of Soil Fertility Properties from a Globally Distributed Soil Mid-Infrared Spectral Library. Soil Sci Soc Am J. 2010;74:1792–1799.
[23] Namdarvand F, Jafarnejadi A, Sayyad G. Estimation of soil compression coefficient using artificial neural network and multiple regressions. Int Res J Appl Basic Sci. 2013;4:3232–3236.
[24] Sy NL. Modelling the infiltration process with a multi-layer perceptron artificial neural network. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques. 2006;51.
[25] Amakye WK, Hou C, Xie L, Lin X, Gou N, Yuan E, et al. Bioactive anti-aging agents and the identification of new anti-oxidant soybean peptides. Food Bioscience. 2021;.