Studies of febuxostat analogues as xanthine oxidase inhibitors through 3D-QSAR, Topomer CoMFA and molecular modeling

Hongjin Tang¹ · Dongsheng Zhao²

Received: 12 October 2018 / Accepted: 28 June 2019 / Published online: 2 July 2019
© Iranian Chemical Society 2019

Abstract

Xanthine oxidase, a complex molybdoflavoprotein, catalyzes the hydroxylation of xanthine to uric acid, which has emerged as an important target for gout and hyperuricemia. In this work, a combination of molecular modeling methods was performed on a series of febuxostat analogues as xanthine oxidase inhibitors to establish molecular models for new drug design, including three-dimensional quantitative structure–activity relationship, topomer comparative molecular field analysis (CoMFA), molecular docking and molecular dynamic simulations. The optimal CoMFA model yielded a leave-one-out correlation coefficient (q^2) of 0.841 and a non-validated correlation coefficient (r^2) of 0.985. The respective q^2 and r^2 of the best comparative molecular similarity indices analysis (CoMSIA) model were 0.794 and 0.972, respectively. The Topomer CoMFA model provided a q^2 of 0.915 and an r^2 of 0.977. 3D contour maps generated from CoMFA and CoMSIA have identified several key features responsible for the inhibition activity. Molecular modeling was taken to further elucidate the proposed binding conformations of the inhibitors to the protein. The obtained results can be served as a useful guideline for designing novel febuxostat derivatives with improved activity against xanthine oxidase.

Keywords Xanthine oxidase inhibitor · 3D-QSAR · Topomer CoMFA · Molecular modeling

Abbreviations

XO	Xanthine oxidase
ROS	Reactive oxygen species
Mo-Pt	Molybdopterin
3D-QSAR	Three-dimensional quantitative structure–activity relationship
SARs	Structure–activity relationships
PLS	Partial least-squares
CoMFA	Comparative molecular field analysis
CoMSIA	Comparative molecular similarity indices analysis
LOO	Leave-one-out

ONC	The number of optimal components
SEE	The standard error of estimate

List of symbols

q^2	Cross-validated correlation coefficient
r^2	Non-cross-validated correlation coefficient
r^2_{pred}	The predictive r^2 value
F	An F test value

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13738-019-01726-y>) contains supplementary material, which is available to authorized users.

✉ Hongjin Tang
tanghongjin@ahpu.edu.cn

¹ College of Biological and Chemical Engineering, Anhui Polytechnic University, Wuhu 241000, People's Republic of China

² College of Pharmacy, Shandong University of Traditional Chinese Medicine, Jinan 250355, People's Republic of China

Introduction

Gout, a common human metabolic disorder, is a chronic disease of deposition of monosodium urate crystals, which forms in the presence of increased urate concentrations [1]. Hyperuricemia is regarded as the primary pathogenic defect in gout and lowering serum uric acid level below deposition threshold is the main goal in management of gout [2]. Xanthine oxidase (XO), a versatile molybdoflavoprotein, is the key enzyme in the purine metabolic pathway, which catalyzes the conversion of hypoxanthine to xanthine and then to uric acid coupled with the production of reactive oxygen species (ROS) such as superoxide anions, H_2O_2 [3–5]. The generation of excess ROS can interfere in cell functions and lead to cellular damage, which is involved in the

pathological progression of many diseases [6]. Moreover, the overproduction of uric acid can cause hyperuricemia, which is the principal characteristic of gout [7]. Thus, XO has emerged as one of the most promising targets for controlling uric acid level and treating gout. In clinical therapy, two known XO inhibitors, including allopurinol and febuxostat (Fig. 1), have been widely prescribed in the treatment of hyperuricemia and gout for many years. However, to some extent, these drugs are still restrained from their clinical applications because of some severe side effects, such as skin problems, progressive renal failure, Stevens–Johnson syndrome and fulminant hepatitis [7, 8]. Under these limitations, there is a considerable demand to develop new XO inhibitors with higher activity and fewer adverse effects for the treatment of gout.

Rational drug design is of great importance for the development and optimization of drug candidates due to their outstanding advantages such as time and labor saving, high efficiency and reproducibility, which provides valuable information to guide drug design and accelerates significantly the drug discovery process [9–11]. Three-dimensional quantitative structure–activity relationship (3D-QSAR) has been widely applied to correlate structural descriptors from a series of derivatives to their biological activities [12, 13]. Also, molecular modeling serves as a simple and useful technique to investigate the possible binding mechanism between bioactive molecules and target protein [14].

In recent years, many studies including our previous works have concentrated on developing novel XO inhibitors with potential use [15, 16]. For example, Kim et al. [17] found compound **1a** (Fig. 1) as a potent XO inhibitor with IC_{50} of 3.5 nM, which could distinctly decrease plasma uric acid level. Zhang et al. [18] synthesized a library of 2-phenyl-4-methyl-1,3-selenazole-5-carboxylic acid derivatives, and compound **1b** (Fig. 1) showed the most potent inhibitory activity (IC_{50} = 5.5 nM). Among them, some excellent

non-purine XO inhibitors, such as febuxostat, Y-700 and topiroxostat (Fig. 1), have attracted extensive attention from medicinal chemists. Indeed, febuxostat received approval for the chronic management of hyperuricemia in patients with gout from the U.S. Food and Drug Administration in 2009, indicating that there was a great potential to further develop novel XO inhibitors based on the scaffold of febuxostat. However, to the best of our knowledge, few literatures systemically reported their SARs. Thus, in the present study, to provide a theoretical guidance for developing more desirable febuxostat-derived XO inhibitors, 3D-QSAR and Topomer CoMFA models were constructed to clarify their SARs. Meanwhile, molecular modeling studies were also performed to explore the proposed binding characteristic between the active molecules and XO.

Materials and methods

Dataset and biological activity

A dataset of 129 compounds was chosen from the related literatures with deleting inactive molecules [7, 8, 17, 19–24]. The biological activity expressed as IC_{50} value was converted to the corresponding pIC_{50} taking $-Lg(IC_{50})$ (Table S1) which was used as a dependent variable in all developed models. Taking structure diversity and bioactivity range into consideration, the dataset was divided into the training set (105 compounds, 80%) and the test set (24 compounds, 20%). The training set was used to build the comparative molecular field analysis (CoMFA), comparative molecular similarity indices analysis (CoMSIA) and Topomer CoMFA models. The test set was applied to validate the reliability and stability of the developed models as extra independent sample.

Fig. 1 The chemical structures of the known XO inhibitors

Molecular modeling and alignment

All molecular modeling calculations and optimization were performed on Sybyl X-2.0 (Tripos, Inc.). The 3D structures of all compounds were drawn by Chem3D Pro 14.0, partial atomic charges were calculated by Gasteiger–Hückel method, and the energy was minimized by the Tripos force field with the gradient convergence criterion and max iterations of 0.005 kcal/mol and 10,000, respectively. Compound **29** with the strongest XO inhibitory activity was used as the reference molecule to fit the remaining compounds.

Molecular alignment plays a key role in 3D-QSAR study. Herein, three superposition methods taking both ligand- and docking-based alignment rules were employed to construct the optimal 3D-QSAR models. The docking-based molecular alignment was completed by the Surflex-dock module. The X-ray co-crystal structure of XO-salicylic acid complex from bovine milk (PDB code: 1FIQ) was obtained from RCSB protein database bank. The protein was prepared by deleting waters, adding hydrogen atoms and repairing the missing and terminal residues of polypeptide chains. The small molecules were flexibly docked into the active pocket of the protein defined by protocol. The other parameters were set as defaults. The binding conformation of each ligand with the highest total score was automatically aligned together and used to build 3D-QSAR models.

CoMFA and CoMSIA field calculation

The CoMFA descriptor fields including steric and electrostatic fields at each lattice intersection were calculated separately using a sp^3 hybridized carbon atom served as the probe with a charge of +1.00 and scaled by the CoMFA-STD method with default energy of 30 kcal/mol. The van der Waals potentials and Coulombic terms were calculated by the Tripos force field [25]. The CoMSIA descriptor fields incorporated five different similarity fields including steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor fields calculated by the same probe atom as CoMFA [26]. In order to assess the reliability and predictive ability of the models generated by the CoMFA and CoMSIA, several common variables were determined including the cross-validated correlation coefficient (q^2), the non-cross-validated correlation coefficient (r^2), the predictive r^2 value (r^2_{pred}), the number of optimal components (ONC), the standard error of estimate (SEE) and an F test value (F). Among them, r^2_{pred} representing the predictive performance of a PLS model, was calculated by Eq. (1):

$$r^2_{pred} = \frac{SD - PRESS}{SD} \quad (1)$$

where SD is the sum of the squared deviations between the inhibitory activities of the test set and mean activity of the

training set, and PRESS is the sum of the squared deviations between predicted and actual activity values of the test set [27]. The results of CoMFA and CoMSIA models were shown as field contour maps, where the coefficients were generated by the field type “Stdev*Coeff”.

Topomer CoMFA field calculation

Topomer CoMFA has emerged as a widely used alignment-independent 3D-QSAR method because of its prominent advantages such as time-saving, high efficiency and convenience, which combines the Topomer technology with the conventional CoMFA method and overcomes the alignment problem of CoMFA [28]. Because of the automatic and user-inaccessible features of Topomer generation, a Topomer CoMFA prediction is almost entirely objective only relying on the 2D connectivity's, user-specified fragmentations and measured biological activity of the training set [29].

The dataset was classified as the method of CoMFA, and the divided fragments (R_1 and R_2) were topomerically aligned. The steric and electrostatic field descriptors were calculated from these alignment fragments using the sp^3 carbon probe. Partial least-squares (PLS) regression was taken to generate the resulting Topomer CoMFA model. It was used to analyze the structure-activity correlation which was shown as colored contour maps representing steric and electrostatic fields. Statistical parameters expressed by the established model were q^2 , r^2 , ONC and standard error that provided the highest q^2 results and y-intercept for PLS analysis. Contribution of each fragment toward the predicted activity was calculated by Topomer CoMFA method.

Molecular docking studies

Molecular docking was carried out to discuss the binding characteristic between the ligands and XO via AutoDock v4.2. The ligands and the protein were prepared as the method of “Molecular modeling and alignment”. The ligands were detected root and chosen torsions to determine the rotatable bonds with AutoDock tools. A grid box of $60 \times 60 \times 60$ Å with spacing of 0.375 Å was defined to enclose the active site of the protein, and the Lamarckian genetic algorithm (LGA) with runs of 100 was taken to conduct the docking calculation.

The other parameters were set as defaults. The complex conformation with the lowest binding energy and highest frequent locus was regarded as the most favorable binding mode, which was used to discuss the potential binding features through some factors, such as hydrogen bonds, π - π interactions, van der Waals and hydrophobic interactions. The graphic display was completed with PyMOL.

Molecular dynamic simulations

Molecular dynamic (MD) simulations were further conducted to evaluate the dynamics of the docked complex on Discovery Studio v2.5. MD simulations started from the previously docked XO-compound **29** complex with the CHARMM force field, a 12 Å non-bound spherical cutoff, the isothermal–isochoric ensemble (NVT) and distance-dependent dielectric implicit solvent model. To eliminate some possible poor contacts in the system, energy minimization of 5000 steps was taken before the simulations. The standard dynamics protocol including minimization, minimization 2, heating, equilibration and production was used to perform the MD simulations with equilibration at 300 K, 10 ns of production at 300 K and the electrostatics of particle mesh Ewald. The main parameters of minimization, minimization 2 and heating were set as follows, minimization: algorithm (Steepest Descent), max steps (5000), RMS gradient of minimization (0.1); minimization 2: algorithm (Conjugate Gradient), max steps (5000), RMS gradient of minimization (0.0001); heating: steps (5000), initial temperature (50), target temperature (300). The time step was set as 2 fs during all stages, and a snapshot was saved every 10 ps. The RMSD value of the backbone atoms referring to the initial docking complex was monitored during the process. Only production phases were regarded as a relatively stable conformation of the ligand–protein complex for further analysis.

Results and discussion

CoMFA, CoMSIA and Topomer CoMFA statistical results

3D-QSAR models were constructed to explore the SARs of februxostat analogues as XO inhibitors. Molecular alignment determines the predictive potency of established models due to the assumption that molecules share scaffold binding sites in common to the active site in a similar manner. Herein, to obtain the optimal 3D-QSAR statistical model, three different superposition rules were used to build the corresponding models. Firstly, the ligand-based superposition (superposition I) relying on database alignment technology was employed using compound **29** as the template. The common substructure was marked in bold (Fig. 2a), and the superposition I model is shown in Fig. 2b. Secondly, the docking-based alignment was adopted to build the superposition II model (Fig. 2c), which was used to obtain the CoMFA and CoMSIA models. Additionally, superposition III (Fig. 2d), similar to superposition I, was a common scaffold-based alignment method except with all molecular conformations obtained by the same manner as superposition II.

Meanwhile, several PLS statistical parameters including q^2 , ONC, r^2 , r^2_{pred} , SEE and F value were determined to assess the predictive potency of the established 3D-QSAR

Fig. 2 The alignment of all molecules in the dataset. **a** The common substructure is shown in bold atoms for the alignment of all training molecules using compound **29** as the template. **b–d** The alignments of the training molecules from superpositions I, II and III, respectively

models (Table 1). By comparing the CoMFA and CoMSIA models, superposition I showed better statistical results than superpositions II and III. Thus, the CoMFA and CoMSIA models based on superposition I were used to explore the SARs of febuxostat analogues. From Table 1, the CoMFA model derived from superposition I exhibited the q^2 value of 0.841 with 10 principal components, $r^2 = 0.985$, $r^2_{\text{pred}} = 0.838$, SEE = 0.182 and an F value of 597.696, revealing that the model had a good predictive ability due to $q^2 > 0.5$ and $r^2 > 0.6$. The CoMFA model presented that the contributions of steric and electrostatic fields were 61.4% and 38.6%, respectively, and the steric field was the main contribution in this model. For the CoMSIA model, the q^2 of 0.794, an ONC of 9, an r^2_{pred} value of 0.797, an r^2 of 0.972, an SEE of 0.197 and an F value of 454.237 were obtained. The contributions of the steric, electrostatic, hydrophobic, hydrogen bond acceptor and hydrogen bond donor field were 17.6%, 18.7%, 21.7%, 22.5% and 19.5%, respectively, suggesting that the electrostatic, hydrophobic and hydrogen bond acceptor fields were of great importance to this model.

In addition, the test set was adopted to estimate the predictive ability of the established models, and the structures, experimental and predictive pIC_{50} values of all compounds are outlined in Table S1. The predicted and actual pIC_{50} values of the training and test set molecules calculated by

the CoMFA and CoMSIA models were mapped (Fig. 3), and both presented the good linear correlation indicating the good predictive property.

Contour maps of CoMFA and CoMSIA

To visualize the field effects on the target molecule in 3D space, CoMFA, CoMSIA and Topomer CoMFA contour maps are graphically interpreted using field type “StDev*Coeff”. For making it easier to explain the contour maps, the most potent compound **29** was chosen as a reference structure to illustrate all contour maps of the models, explaining the key structural features required for XO inhibitory activity.

As shown in Fig. 4a, the sterically favored and disfavored regions are colored as green and yellow, respectively. The default values of 80% contribution for favored and 20% for disfavored regions were set for visualization of the contour maps. A large green contour was located near R_1 group of compound **29**, suggesting that bulky groups were favored at this position. For example, compounds **2–5**, **52–56** and **100–102** with a bulkier substituent at R_1 significantly increased the inhibitory activity compared to compounds **1**, **49** and **99**, respectively. One yellow contour surrounding region R_2 indicated that a bulky group at this position would decrease the potency. By checking the

Table 1 Statistical results for the CoMFA, CoMSIA and Topomer CoMFA models based on different superposition methods

PLS statistical parameters	Superposition methods						
	I			II		III	
	CoMFA	CoMSIA	Topomer	CoMFA	CoMSIA	CoMFA	CoMSIA
q^{2a}	0.841	0.794	0.915	0.541	0.418	0.398	0.443
r^{2b}	0.985	0.972	0.977	0.979	0.991	0.994	0.948
ONC ^c	10	9	9	8	10	10	5
SEE ^d	0.182	0.197	0.163	0.211	0.139	0.117	0.325
F^e	597.696	454.237	–	555.653	1037.696	1467.478	363.077
$r^2_{\text{pred}}^f$	0.838	0.797	0.755	–	–	–	–
Fraction of field contributions ^g							
Steric	0.614	0.176	–	0.346	0.081	0.462	0.112
Electrostatic	0.386	0.187	–	0.654	0.281	0.538	0.254
Hydrophobic	–	0.217	–	–	0.187	–	0.170
Hydrogen bond acceptor	–	0.225	–	–	0.242	–	0.240
Hydrogen bond donor	–	0.195	–	–	0.208	–	0.223

^aCross-validated correlation coefficient

^bNon-cross-validated correlation coefficient

^cOptimum number of components

^dStandard error of estimate

^e F test value

^fThe predictive r^2 value

^gField: steric, electrostatic, hydrophobic, hydrogen bond acceptor and hydrogen bond donor

Fig. 3 Plot of predicted versus experimental pIC_{50} values for all the molecules based on the CoMFA (a) and CoMSIA models (b)

Fig. 4 Steric contour maps for CoMFA (a) and CoMSIA (c), and electrostatic contour maps for CoMFA (b) and CoMSIA (d). The green color shows the favored steric area, and the yellow color shows

the disfavored steric area. The red color shows the favored electron-withdrawing area, and the blue color shows the favored electron-donating area

febuxostat analogues, it could be observed that compounds **30**, **38**, **53**, **55** and **71** ($R_2 = H$) showed higher inhibition activity than compounds **32**, **39**, **60**, **61** and **73**, respectively ($R_2 = CH_3$ or OCH_3).

As shown in Fig. 4b, for the electrostatic field of CoMFA model, a large red contour map surrounding group R_1 and one red contour around region R_2 suggested their affinity to electron-withdrawing substituents. If these positions were substituted by electron-donating groups, it would result in a decrease of inhibitory activity. For example, derivatives **65**, **68**, **70**, **82** and **83** ($IC_{50} = 3.9$ – 8.6 nM) bearing electron-withdrawing groups (i.e., Cl, CN, NO_2) at position R_1 presented distinct increase compared to compound **64** ($IC_{50} = 110$ nM) with $R_2 = H$. Additionally, the activities of compounds **116** and **117** ($IC_{50} = 0.02$ – 0.14 μM) were higher than those of compounds **113**, **114** and **115** ($IC_{50} = 0.95$ – 6.72 μM), which was mainly due to the presence of strong electron-withdrawing groups (i.e., CN, NO_2) at region R_1 .

The CoMSIA steric and electrostatic contour maps are depicted in Fig. 4c, d, respectively, which provided similar information as the CoMFA model. The hydrophobic field contour maps of the CoMSIA model are shown in Fig. 5a, where one yellow contour map near groups R_1 and R_2 , respectively, revealed that hydrophobic groups helped to increase the binding affinity. In contrast, the potency would decrease when positions R_1 and R_2 were substituted by hydrophilic groups. Compared to compound **88** ($IC_{50} = 233$ nM), the inhibition activity of compounds

90–93 substituted by Cl, Br, CN and NO_2 presented significant improvement with IC_{50} values ranging from 8.9 to 17 nM. Moreover, it was also in agreement with the fact that compounds **96–98**, **100–102** ($IC_{50} = 5.1$ – 9.4 nM) showed more potent inhibitory activity than the inhibitor **99** whose R_1 group was substituted by the hydrophilic group amino ($IC_{50} = 31.8$ nM). In addition, as shown in Fig. 5a, one yellow contour map was observed covering region R_2 , suggesting that the existence of hydrophobic substituents would be beneficial to such activity. It could be demonstrated that compounds **103–109** showed relatively weak inhibitory effect ($IC_{50} = 0.6$ – 16.52 μM) due to the presence of the hydrophilic groups OH and NH_2 on position R_2 .

The hydrogen bond donor field is depicted in Fig. 5b, in which the cyan contours refer to the favored hydrogen bond donor regions, whereas the purple contours represent where hydrogen bond donor disfavors the biological activity. Only one piece of a medium-sized region of the cyan contour was observed near position R_2 , suggesting that the existence of the hydrogen bond donor group (i.e., COOH) made key contribution to the inhibitory activity. The data were consistent with the statistical results (24.1% field contributions) of the CoMSIA model. As shown in Fig. 5c, the favored and disfavored regions for the hydrogen bond acceptor field are represented by magenta and red, respectively. One medium-sized region of magenta contour surrounded group R_2 and one small magenta contour map was located near position R_1 ,

Fig. 5 Hydrophobic (a), hydrogen bond donor field (b) and hydrogen bond acceptor (c) contour maps for CoMSIA. Yellow represents the favored hydrophobic area; white represents the favored hydrophilic

area; cyan and purple represents the favored and disfavored hydrogen bond donor area, respectively; magenta and red represent the favored and disfavored hydrogen bond acceptor area, respectively

manifesting that the introduction of hydrogen bond acceptor group also elicited an important effect on such activity, such as CN and NO₂.

With respect to the highest active molecule **29** and the lowest active molecule **124**, the SARs could be explained rationally by contour maps as follows: the introduction of the bulky group (OCH₂CH(CH₃)₂) on the phenyl ring of compound **29** would distinctly improve inhibitory potency, whereas the existence of halogen (Cl) on the aryl ring of compound **124** was not benefit for such activity. The results were in agreement with the observations of structure features on the 3D-QSAR contour maps. Moreover, for compound **29**, the generated hydrogen bonds and π -conjugated system of phenyl-pyrazole ring played a key role in its binding to the active site of XO as evinced by molecular docking, which further confirmed the importance of hydrogen bond fields resulting in the increment of inhibitory activity.

Contour maps of Topomer CoMFA

Topomer CoMFA is regarded as the second generation of CoMFA. In the Topomer CoMFA analysis, firstly, each of the training set structure was divided into two sets of fragments shown as R₁ (labeled red) and R₂ groups (labeled blue) (Fig. 6). All the models are investigated using the full cross-validated (q^2) PLS leave-one-out (LOO) method with CoMFA standard options for scaling of the variables. The established models provide the quantitative relationships between the chemical structure and activity.

Fig. 6 R₁ fragment is represented by the red color, and R₂ fragment is represented by the blue color

Fig. 7 Steric contour maps of Topomer CoMFA around R₁ (a) and R₂ (b), and electrostatic contour maps of Topomer CoMFA around R₁ (c) and R₂ (d) group. Green contours refer to the favored steric area,

As a result, the statistical parameters derived from this model were as follows: the q^2 of 0.915 with an ONC of 9 and an r^2 of 0.977, suggesting that the model also had the predictive ability (the standard for a reliable model was taken as $q^2 > 0.2$). The pIC₅₀ value of the test set was predicted with the r^2_{pred} value of 0.755. In addition, the standard error was found to be 0.43 for q^2 and 0.22 for r^2 with a y-intercept of 5.03.

The steric and electrostatic contour maps generated by the Topomer CoMFA model are shown in Fig. 7. For the steric field, a large green contour was mapped near group R₁ (Fig. 7a) and three areas with medium-sized yellow contours were located at group R₂ (Fig. 7b), suggesting that bulky groups were favored at region R₁ and disfavored at region R₂. As shown in Fig. 7c, d, for the electrostatic field, three red contours surrounding positions R₁ and R₂, respectively, showed the importance of electron-withdrawing substituents in imparting better biological activity. The result of Topomer CoMFA was almost similar to that of the CoMFA and CoMSIA models, which further confirmed the predictive potency of the CoMFA and CoMSIA models.

Molecular docking analysis

Molecular docking studies are widely used to discuss the proposed binding characteristic between bioactive molecules and protein. XO contains three major domains, namely [Fe-S] center, FAD and Molybdenum (Mo-pt) center domain. Among them, Mo-pt is the functional domain where the oxidation of xanthine takes place [30]. The well-known inhibitor febuxostat (compound **129**) was firstly docked into the active site of XO (Mo-pt domain). As shown in Figure S1, cluster analysis of 100 docking runs presented a total of 6 multimember conformational clusters at a RMSD tolerance of 2.0 Å and the cluster with the lowest binding free energy and highest frequent locus (labeled as red histogram) was chosen to further analyze. The results indicated that key

and yellow contours refer to the disfavored steric area. Red contours refer to the favored negative electrostatic area, and blue contours refer to the favored positive electrostatic area

Fig. 8 Docking results of compounds **29** (a), **52** (b) and **97** (c) in the binding site of XO (PDB code: 1FIQ). View of superimposed docking conformations for the inhibitors in the active pocket of XO (d).

Key residues are shown as stick model, and hydrogen bonds are shown as red dashed lines with the distance of Å. Compounds **29**, **52** and **97** are colored as cyan, yellow and magenta, respectively

residues Arg880 and Thr1010 (forming hydrogen bonds), Phe914 and Phe1009 (generating π -stacking effects) and hydrophobic interactions were extremely helpful to maintain the binding conformation of febuxostat in the active Mo-pt domain of XO.

The docking results of compounds **29**, **52** and **97** with XO are shown in Fig. 8. It was found that the three inhibitors presented the similar binding conformations in the active site of XO due to the approximate docking orientations and key interaction residues, including Leu648, Asn768, Leu873, Arg880, Phe914, Phe1009, Thr1010, Val1011, Phe1013 and Ala1079. For compound **29**, there were two hydrogen bonds generated with residue Arg880 at the distance of 1.68 and 1.93 Å. It could be observed that four hydrogen bonds formed between compound **52** and residues Arg880 (2.06, 2.20 Å) and Thr1010 (1.82, 2.70 Å). Also, the free carboxyl group of the inhibitor **97** generated three hydrogen bonds with residues Arg880 (1.89, 2.10 Å) and Thr1010 (2.03 Å). These results revealed that hydrogen bonds and the residues Arg880 and

Thr1010 made great contributions for the binding of such inhibitors to XO. Meanwhile, potential aromatic interaction (π - π effects) should not be ignored due to the fact that π -stacking interactions (face-to-face with Phe914 and face-to-edge with Phe1009, Phe1013) found for the original ligand salicylic acid were maintained by compounds **29**, **52** and **97**. Moreover, from Fig. 8d, it is clearly seen that the five-membered heterocycle groups of these inhibitors were all adjacent to the Mo-pt center and stretched into the innermost hydrophobic region of the active pocket, which was consistent with the analytical results of the CoMSIA model. In addition, according to the further analysis of the docking results, van der Waals interaction and some non-polar/polar interactions could be also found for their binding process. Thus, the obtained data not only provided deep insights into understanding the possible binding modes of these inhibitors on XO, but also suggested that hydrophobic interactions and hydrogen bonds played an indispensable role in stabilizing the docking complex conformation.

Fig. 9 The SARs of febxostat analogues revealed by 3D-QSAR and docking studies

Summary of SARs

Based on the analysis of 3D-QSAR and molecular modeling, the key structure features affecting the XO inhibitory activity, namely SARs could be concluded as follows (Fig. 9): (1) bulky groups on region R₁ and minor groups on R₂ help to increase the activity; (2) introducing electron-withdrawing substituents at positions R₂ and R₃ favored the inhibitory potency; (3) hydrophobic groups surrounding regions R₁ and R₂ would contribute to improving the activity; (4) the formation of hydrogen bonds with groups R₃ and R₄ is beneficial to stabilize the ligand–protein complex; (5) the generated π -conjugated system of phenyl–heteroaromatic ring is of great importance for the ligands binding to XO because of intermolecular interactions.

Molecular dynamic simulations

In order to complement the docking studies, 10 ns of MD simulations were performed based on the complex compound **29**–XO to refine the docking models. The RMSD value of the backbone atoms was monitored to investigate the stabilization of complex conformation during the MD simulations. As shown in Fig. 10a, the RMSD values gradually increased in the range of 0.15–0.30 nm from 0 to 8 ns, and then maintained at a relatively constant level within the rest of time, which suggested that the system and complex conformation were stable throughout simulations. Meanwhile, the root mean square fluctuation values (RMSF) of the backbone atoms were recorded. From Fig. 10b, the ligand binding pocket, residues 680–1080 for ligand–protein

showed the lowest RMSF values and several areas presented large fluctuation that were far from the active binding pocket, which suggested that the simulation system was stable. After 10-ns simulations, the superimposed conformations of the MD-simulated and original ligand structures were compared. Only a slight shifting for them was observed in the active pocket of XO, and they exhibited the similar binding characteristic due to the same key residues, such as Asn768, Arg880, Phe914, Phe1009, Thr1010, Val1011 and Phe1013. Meanwhile, hydrophobic interactions and aromatic π – π effects were still maintained for the MD-simulated complex. However, the number of hydrogen bond interactions increased although one more hydrogen bond formed with the distance of 2.4 Å between residue Asn768 and the cyan group of compound **29** (Fig. 10c). The MD simulation study further provided direct evidence for the potential binding properties of the active compound with XO and supported the observation of conformational changes within XO as affected by the inhibitor. All the explanations would be favored to the design and development of novel febxostat-derived drug candidates as potent XO inhibitors.

Design of novel febxostat derivatives

Taking the SARs and the feasibility of analogues synthesis into consideration, a small amount of novel febxostat derivatives were rationally designed and their pIC₅₀ values were predicted based on the established CoMFA and CoMSIA models (Table 2). The obtained data revealed that a majority of the designed derivatives showed more potent predicted pIC₅₀ values than the reference compound **29**,

Fig. 10 **a** RMSD values of the ligand–protein complex within 10 ns of simulation time. **b** RMSF values of the backbone atoms averaged for each residue of XO. **c** Interaction conformation of MD simula-

tions from compound 29-XO complex. Key residues are shown as ball models colored by interaction. Hydrogen bonds with the distance of Å are shown as blue and green dashed lines colored by interaction

which further confirmed the summarized SARs. Thus, these results provides useful information to guide the design of new potent XO inhibitors derived from febuxostat; however, the further synthetic study combined with the bioactivity evaluation *in vitro* and *in vivo* is also inevitable for them to develop novel drug leads for the treatment of gout and related complications.

Conclusions

In this study, an integrated method consisting of 3D-QSAR, Topomer CoMFA and molecular modeling was used to investigate the SARs for a library of febuxostat analogues as potent XO inhibitors. As demonstrated by the high cross-correlation coefficients, the developed models

could provide stable, reliable correlation and statistically significant predictive ability. The CoMFA, CoMSIA and Topomer CoMFA contour maps provided important information to understand the SARs and identify several primary structure features correlating with the inhibitory activity. Additionally, for a better understanding of the binding mechanisms of the inhibitors to XO, molecular docking and MD simulations were undertaken. As a result, several key residues (i.e., Arg880, Phe914, Thr1010), hydrophobic interactions and hydrogen bonds were demonstrated to be of great importance for their binding and the stability of the complex conformation. The overall obtained results not only provide useful information to recognize the proposed factors affecting activity, but also help to develop new XO inhibitors with improved activity based on febuxostat.

Table 2 Structures and predicted pIC_{50} values of compound **29** and newly designed febuxostat derivatives

<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;"> <p>1-2</p> </div> <div style="text-align: center;"> <p>3-4</p> </div> <div style="text-align: center;"> <p>5-7</p> </div> <div style="text-align: center;"> <p>8-10</p> </div> </div>				
No.	Structure		pIC_{50}^a	pIC_{50}^b
	R ₁	R ₂		
29^c	-	-	8.379	8.288
1		-	8.480	8.375
2		-	8.704	8.421
3		-	8.396	8.523
4		-	8.294	8.683
5		CN	8.557	8.492
6		CN	8.396	8.372
7		NO ₂	8.471	8.409
8		NO ₂	8.689	8.594
9		NO ₂	8.379	8.295
10		CN	8.673	8.537

^a Predicted pIC_{50} by CoMFA model.^b Predicted pIC_{50} by CoMSIA model.^c Experimental pIC_{50} for compound **29** as 8.620.

Acknowledgements This work was financially supported by the Natural Science Foundation of Anhui Province (No. 1908085QH346), the Introduction of Talent Research Start-up Fund of Anhui Polytechnic University (No. 2017YQQ017) and the Pre-research project of National Natural Science Foundation of Anhui Polytechnic University (No. 2019yyzr10).

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

References

1. N. Dalbeth, T.R. Merriman, L.K. Stamp, *Lancet* **388**, 2039 (2016)
2. G. Ragab, M. Elshahaly, T.J. Bardin, *J. Adv. Res.* **8**, 495 (2017)
3. K. Sugamura, J.F.J. Keaney, *Free Radic. Biol. Med.* **51**, 978 (2011)
4. R. Harrison, *Free Radic. Biol. Med.* **33**, 774 (2002)
5. D. Kumar, G. Kaur, A. Negi, S. Kumar, S. Singh, R. Kumar, *Bioorg. Chem.* **57**, 57 (2014)
6. J.H. Cheng, A.M. Huang, T.C. Hour, S.C. Yang, Y.S. Pu, C.N. Lin, *Eur. J. Med. Chem.* **46**, 1222 (2011)
7. S. Khanna, S. Burudkar, K. Bajaj, P. Shah, A. Keche, U. Ghosh, A. Desai, A. Srivastava, A. Kulkarni-Almeida, N.J. Deshmukh, A. Dixit, M.K. Brahma, U. Bahirat, L. Doshi, K.V. Nemmani, P. Tannu, A. Damre, C.B. Rao, R. Sharma, H. Sivaramakrishnan, *Bioorg. Med. Chem. Lett.* **22**, 7543 (2012)
8. S.J. Wang, J.F. Yan, J. Wang, J.R. Chen, T.J. Zhang, Y. Zhao, M.X. Xue, *Eur. J. Med. Chem.* **45**, 2663 (2010)
9. D. Singla, S.K. Dhanda, J.S. Chauhan, *Curr. Top. Med. Chem.* **13**, 1172 (2013)
10. A.M. Davis, D.J. Wood, *Mol. Pharm.* **10**, 1183 (2013)
11. S.J.Y. Macalino, V. Gosu, S. Hong, S. Choi, *Arch. Pharm. Res.* **38**, 1686 (2015)
12. C.L. Hung, C.C. Chen, *Drug Dev. Res.* **75**, 412 (2014)
13. M.H. Dong, Y.J. Ren, *RSC Adv.* **5**, 13754 (2015)
14. L.G. Ferreira, R.N. dos Santos, G. Oliva, A.D. Andricopulo, *Molecules* **20**, 13384 (2015)
15. H.J. Tang, X.W. Zhang, L. Yang, W. Li, J.H. Li, J.X. Wang, J. Chen, *Eur. J. Med. Chem.* **124**, 637 (2016)
16. H.J. Tang, W. Li, M. Zhou, L.Y. Peng, J.X. Wang, J.H. Li, J. Chen, *Eur. J. Med. Chem.* **151**, 849 (2018)
17. J.U. Song, S.P. Choi, T.H. Kim, C.K. Jung, J.Y. Lee, S.H. Jung, G.T. Kim, *Bioorg. Med. Chem. Lett.* **25**, 1254 (2015)
18. Q. Guan, Z.J. Cheng, X.X. Ma, L.J. Wang, D.J. Feng, Y.H. Cui, K. Bao, L. Wu, W.G. Zhang, *Eur. J. Med. Chem.* **85**, 508 (2014)
19. T.J. Zhang, Y.Y. Lv, Y. Lei, D. Liu, Y. Feng, J.X. Zhao, S.L. Chen, F.H. Meng, S.J. Wang, *Eur. J. Med. Chem.* **146**, 668 (2018)
20. T.J. Zhang, S.Y. Li, L. Wang, Q. Sun, Q.X. Wu, Y. Zhang, F.H. Meng, *Eur. J. Med. Chem.* **141**, 362 (2017)
21. J. Li, F.P. Wu, X.G. Liu, Y.K. Zou, H.X. Chen, Z. Li, L. Zhang, *Eur. J. Med. Chem.* **140**, 20 (2017)
22. S.L. Chen, T.J. Zhang, J. Wang, F.Y. Wang, H.D. Niu, C.F. Wu, S.J. Wang, *Eur. J. Med. Chem.* **103**, 343 (2015)
23. J.U. Song, J.W. Jang, T.H. Kim, H. Park, W.S. Park, S.H. Jung, G.T. Kim, *Bioorg. Med. Chem. Lett.* **26**, 950 (2016)
24. M.R. Ali, S. Kumar, O. Afzal, N. Shalmali, M. Sharma, S. Bawa, *Chem. Biol. Drug Des.* **87**, 508 (2016)
25. F.F. Wang, W. Yang, Y.H. Shi, G.W. Le, *Chem. Biol. Interact.* **240**, 324 (2015)
26. M.Y. Wang, Y.Y. Jin, H.Y. Wei, L.S. Zhang, *Eur. J. Med. Chem.* **103**, 91 (2015)
27. W.M. Ding, M. Sun, S.M. Luo, T. Xu, Y.B. Cao, X.F. Yan, Y. Wang, *Molecules* **18**, 10228 (2013)
28. S.D. Joshi, S.R. Dixit, U.A. More, R. Narayan, T.M. Aminabhavi, V.H. Kulkarni, *Med. Chem. Res.* **23**, 4542 (2014)
29. R.D. Cramer, *J. Comput. Aided Mol. Des.* **25**, 197 (2011)
30. K. Okamoto, K. Matsumoto, R. Hille, B.T. Eger, E.F. Pai, T. Nishino, *Proc. Natl. Acad. Sci.* **101**, 7931 (2004)