

# Design and prediction of new anticoagulants as a selective Factor IXa inhibitor via three-dimensional quantitative structure-property relationships of amidinobenzothiophene derivatives

Jia-Suo Gao,<sup>1,\*</sup> Xu-Peng Tong,<sup>2,\*</sup> Yi-Qun Chang,<sup>1</sup> Yu-Xuan He,<sup>1</sup> Yu-Dan Mei,<sup>1</sup> Pei-Hong Tan,<sup>1</sup> Jia-Liang Guo,<sup>1</sup> Guo-Chao Liao,<sup>3</sup> Gao-Keng Xiao,<sup>1</sup> Wei-Min Chen,<sup>1</sup> Shu-Feng Zhou,<sup>4</sup> Ping-Hua Sun<sup>1</sup>

<sup>1</sup>Guangdong Province Key Laboratory of Pharmacodynamic Constituents of Traditional Chinese Medicine and New Drugs Research, College of Pharmacy, Jinan University, Guangzhou, People's Republic of China; <sup>2</sup>College of Pharmacy, China Pharmaceutical University, Nanjing, People's Republic of China; <sup>3</sup>Department of Chemistry, Wayne State University, Detroit, Michigan, USA; <sup>4</sup>College of Pharmacy, University of South Florida, Tampa, FL, USA

\*These authors contributed equally to this work

**Abstract:** Factor IXa (FIXa), a blood coagulation factor, is specifically inhibited at the initiation stage of the coagulation cascade, promising an excellent approach for developing selective and safe anticoagulants. Eighty-four amidinobenzothiophene antithrombotic derivatives targeting FIXa were selected to establish three-dimensional quantitative structure-activity relationship (3D-QSAR) and three-dimensional quantitative structure-selectivity relationship (3D-QSSR) models using comparative molecular field analysis and comparative similarity indices analysis methods. Internal and external cross-validation techniques were investigated as well as region focusing and bootstrapping. The satisfactory  $q^2$  values of 0.753 and 0.770, and  $r^2$  values of 0.940 and 0.965 for 3D-QSAR and 3D-QSSR, respectively, indicated that the models are available to predict both the inhibitory activity and selectivity on FIXa against Factor Xa, the activated status of Factor X. This work revealed that the steric, hydrophobic, and H-bond factors should appropriately be taken into account in future rational design, especially the modifications at the 2'-position of the benzene and the 6-position of the benzothiophene in the R group, providing helpful clues to design more active and selective FIXa inhibitors for the treatment of thrombosis. On the basis of the three-dimensional quantitative structure-property relationships, 16 new potent molecules have been designed and are predicted to be more active and selective than Compound 33, which has the best activity as reported in the literature.

**Keywords:** CoMFA, CoMSIA, 3D-QSAR, 3D-QSSR, benzothiophene antithrombosis

## Introduction

Arterial and venous thromboses are still the major cause of mortality in clinic, although tremendous efforts are taken to prevent and treat thrombotic events. Current anticoagulants such as ticlopidine, clopidogrel, heparin, and warfarin have significant adverse effects like thrombocytopenia, bleeding, and need for careful patient monitoring, making their therapeutic utility restricted.<sup>1-3</sup> For the past decade, many Factor Xa (FXa) inhibitors targeting the specific enzymes of the extrinsic common pathway in the coagulation cascade have been advanced to clinical development, but they remain in the potential risk category of bleeding and therefore have a narrow therapeutic window.<sup>4-6</sup> Factor IXa (FIXa), a vitamin K-dependent blood coagulation factor, is essential for the amplification or consolidation phase of blood coagulation via activating Factor X (FX) to FXa in the formation of a thrombus with 50 times more efficiency than Factor VIIa/tissue factor.<sup>7-11</sup> The selective inhibition of FIXa is believed to be effective in intravascular anticoagulation with maintenance of extravascular hemostasis for FIX

Correspondence: Shu-Feng Zhou  
College of Pharmacy, University of South Florida, Tampa, FL 33612, USA  
Tel +1 813 974 5699  
Fax +1 813 9059890  
Email szhou@health.usf.edu

Ping-Hua Sun  
Guangdong Province Key Laboratory of Pharmacodynamic Constituents of Traditional Chinese Medicine and New Drugs Research, College of Pharmacy, Jinan University, Guangzhou 510632, People's Republic of China  
Tel +86 20 8522 1367  
Fax +86 20 8522 4766  
Email biochemdoctor@sina.com

activation, both by the stimulation of the intrinsic system and by maintaining a low level of tissue factor.<sup>5,12–14</sup> As an alternative therapeutic target, FIXa is specifically inhibited at the initiation stage of the coagulation cascade, promising an excellent approach for developing selective and safe anticoagulants. Recently, 2-amidinobenzothiophene derivatives have been reported as potent FIXa-selective ligands.<sup>15,16</sup>

Three-dimensional quantitative structure–property relationships (3D-QSPRs) including three-dimensional quantitative structure–activity relationship (3D-QSAR) and three-dimensional quantitative structure–selectivity relationship (3D-QSSR) techniques such as comparative molecular field analysis (CoMFA)<sup>17</sup> and comparative similarity indices analysis (CoMSIA)<sup>18</sup> are useful methods of ligand-based drug design used to correlate physicochemical descriptors from a set of related compounds to their known molecular activity or molecular property values.<sup>19</sup> These computational techniques combining 3D information for the ligands have proved particularly helpful in the design of novel, more potent inhibitors.<sup>20,21</sup> Hao et al<sup>22</sup> have done some 3D-QSAR work on the binding mode of benzothiophene analogs such as FIXa inhibitors, only by CoMFA method without using CoMSIA, and the predictive CoMFA model they developed with leave-one-out cross-validation  $q^2=0.52$ , conventional  $r^2=0.97$ , and  $r^2_{pred}=0.81$  still has room for improvement. To our knowledge, there has been no report concerning the application of a 3D-QSSR methodology to selectivity of FIXa, although the 3D-QSSR is also crucial for the development of these FIXa inhibitors. The systematic study on FIXa inhibitors of 3D-QSAR and 3D-QSSR methodologies obtained highly predictive models, by which 16 new potent molecules have been designed. The present paper also explored why and how the target that 2-amidinobenzothiophene derivatives selectively bind with is not FXa, but is actually FIXa. The satisfactory 3D-QSAR and 3D-QSSR models of 84 compounds as serine protease enzyme FIXa inhibitors provide a solid basis for future rational design of more active and selective anticoagulants.

## Materials and methods

### Database and biological activity

Eighty-four molecules selected for the present study were taken from the works of Wang et al.<sup>15,16</sup> In the 3D-QSAR analysis in the current study, 12 compounds were randomly chosen as the test set and the remaining 72 compounds as the training set. In the 3D-QSSR analysis, 73 compounds with the half maximal inhibitory concentration ( $IC_{50}$ ) values against FIXa and FXa were randomly divided into a test

set of nine compounds and a training set of the remaining 64 compounds. For convenience, the  $IC_{50}$  (FIXa) values were converted to their negative logarithm ( $pIC_{50}$ ) values. The  $\log IC_{50}$  (FXa/FIXa) that is the negative logarithm of  $\log IC_{50}$  (FIXa/FIXa), can be used as an index for the selectivity (FIXa/FIXa). The  $pIC_{50}$  values have a span of four log units, from 4.43 to 8.70, while the  $\log IC_{50}$  (FXa/FIXa) values have a span of five log units, from  $-2.80$  to  $2.73$ , indicating the reliable 3D-QSAR and 3D-QSSR models. The  $pIC_{50}$  (FIXa) and  $\log IC_{50}$  (FXa/FIXa) values and chemical structures of the compounds are shown in Tables 1 and 2.

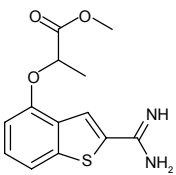
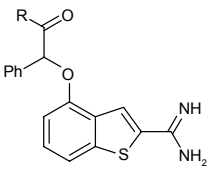
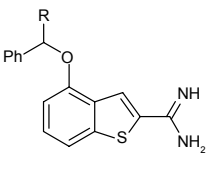
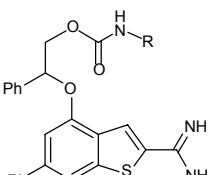
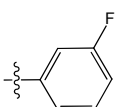

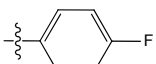
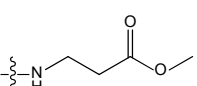
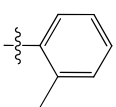
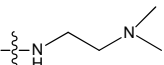
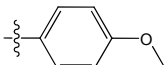
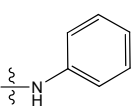
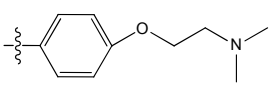
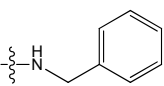
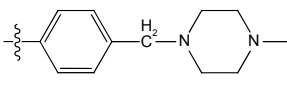
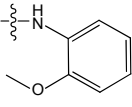
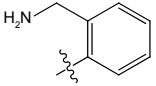
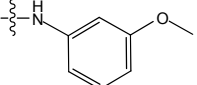
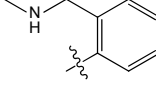
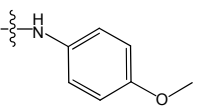
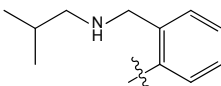
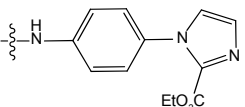
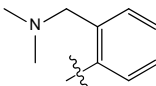
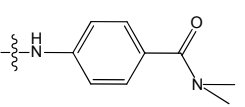
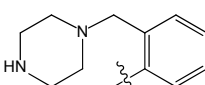
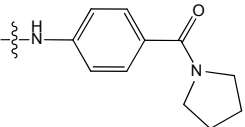
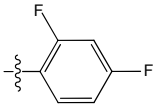
### Molecular structure building and alignment

Molecular alignment is considered to be one of the most essential elements that extensively influence the quality and the predictive ability of the models in CoMFA and CoMSIA analyses, although other factors such as lattice shifting step size, probe atom type, and orientation of the aligned compounds might have influence on the results as well.<sup>23,24</sup> A 2-amidinobenzothiophene ring with structural rigidity, comprising the common core of all molecules in the data set, was selected as the common substructure to align all of the molecules (Figure 1). Active conformation selection is a key step for 3D-QSAR analysis, and in the present study, the molecular conformation of Compound 33 (the most active compound) obtained from the molecular docking was used as the template to build molecular structures of all the compounds. It can be seen from Figure 1 that all the compounds studied have similar bioactive conformations.

### CoMFA methodology

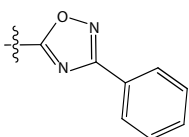
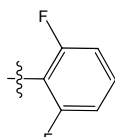

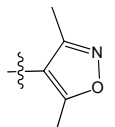
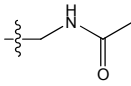
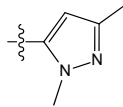
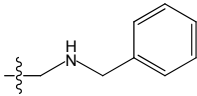
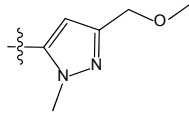
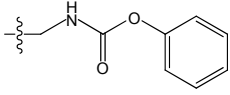
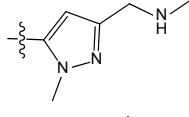
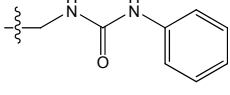
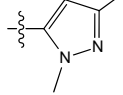

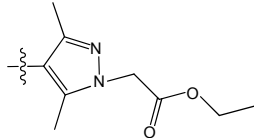
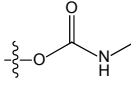
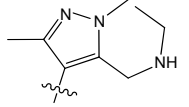
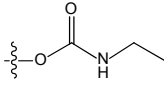
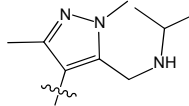
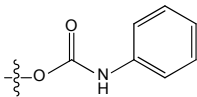
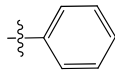
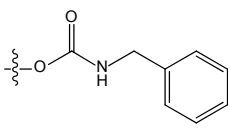
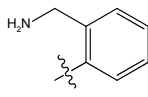
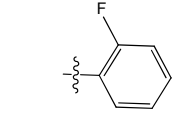
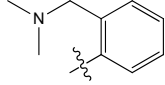
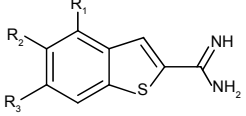
Despite being unable to describe all of the binding forces, CoMFA is still a widely useful tool for QSPR analysis at 3D level.<sup>25</sup> In the present study, the grid spacing was set at 2 Å in the x, y, and z directions, and the grid region was automatically generated to be a 3D cubic lattice that extended at least 4 Å beyond van der Waals volume of all aligned molecules on all axes. Lennard-Jones potential and Coulomb potential were employed to calculate steric and electrostatic energies of each molecule, respectively,<sup>19,20</sup> and an  $sp^3$ -hybridized carbon atom with a +1 charge was used as the probe atom to determine the magnitude of the field values. The column filtering value of 2.0 kcal/mol and the cutoff value of 30 kcal/mol improved the signal-to-noise ratio and reduced domination by large steric and electrostatic energies to a minimum.<sup>21</sup> The regression analysis was carried out using the partial least squares (PLS) method,<sup>20</sup> and the final model was obtained.

**Table I** Structures of 2-amidinobenzothiophene derivatives

<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  <p><b>I</b></p> </div> <div style="text-align: center;">  <p><b>2-12</b></p> </div> <div style="text-align: center;">  <p><b>13-24</b></p> </div> <div style="text-align: center;">  <p><b>25-48</b></p> </div> </div>				
Compd no	R	Compd no	R	R'
I	—	25		H
2*#		26		H
3		27		H
4		28		H
5*#		29		H
6		30		H
7		31*#		H
8		32		H
9		33		H
10		34		H
11		35		H
12		36		H

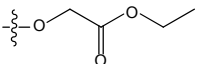
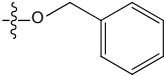
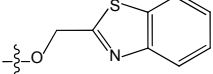
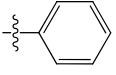
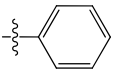
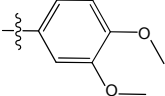
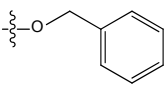
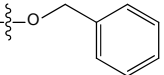
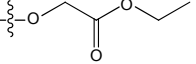
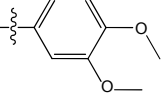
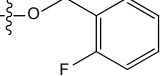
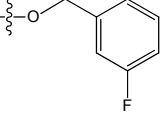
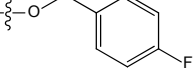
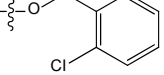
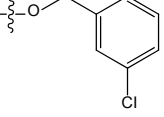
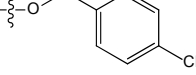
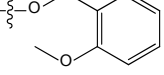
(Continued)

Table 1 (Continued)

Compd no	R	Compd no	R	R'
13		37		H
14		38*#		H
15		39		H
16		40		H
17		41		H
18		42		H
19		43		H
20*#		44		H
21		45		H
22		46		F
23		47		F
24		48		F
				
	49-84			

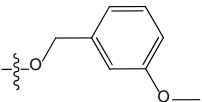
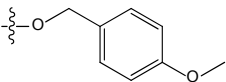
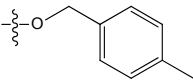
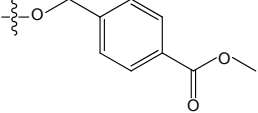
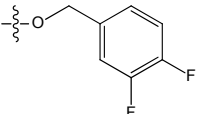
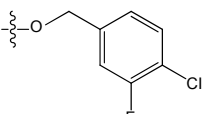
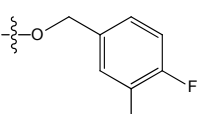
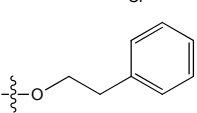
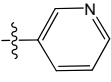
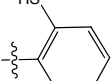
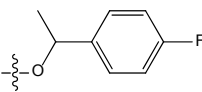
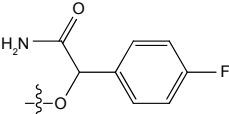
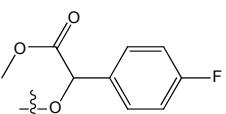
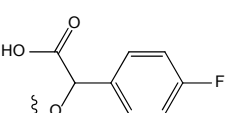
(Continued)

Table 1 (Continued)

	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>
49*:#	I	H	H
50		H	H
51		H	H
52		H	H
53		H	H
54	H		H
55	H		H
56	H		H
57*:#	H	H	
58	H	H	
59	H	H	
60*	H	H	
61	H	H	
62	H	H	
63	H	H	
64	H	H	
65	H	H	
66	H	H	

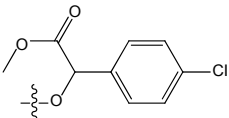
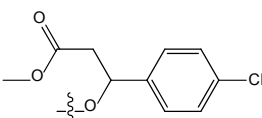
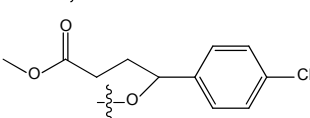
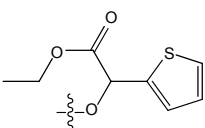
(Continued)

Table I (Continued)

	$R_1$	$R_2$	$R_3$
67	H	H	
68	H	H	
69	H	H	
70*	H	H	
71	H	H	
72	H	H	
73	H	H	
74	H	H	
75	H	H	
76	H	H	
77*	H	H	
78*#	H	H	
79	H	H	
80*#	H	H	

(Continued)

Table 1 (Continued)

	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>
81	H	H	
82	H	H	
83	H	H	
84	H	H	

**Notes:** \*Test samples for 3D-QSAR model validation; #test samples for 3D-QSSR model validation.

**Abbreviations:** 3D-QSAR, three-dimensional structure–activity relationship; 3D-QSSR, three-dimensional structure–selectivity relationship.

## CoMSIA methodology

CoMSIA and CoMFA are both based on the same assumption that changes in binding affinities of ligands are related to changes in molecular properties, represented by fields. However, CoMFA only describes the two binding forces of the steric field and electrostatic field, but three other different fields including hydrophobic, hydrogen bond donor, and hydrogen bond acceptor are calculated in CoMSIA.<sup>26</sup> For the distance dependence between the probe atom and the molecule atoms, a Gaussian function is used. Because of the different shape of the Gaussian function, the similarity indices can be calculated at all grid points, both inside and outside the molecular surface, while the calculations are not done inside the molecular surface in CoMFA. Equation 1 is used to calculate the similarity indices as follows:

$$A_{F,K(j)}^q = \sum_i W_{probe,k} W_{ik} e^{-\alpha r_{iq}^2} \quad (1)$$

where  $A$  is the similarity index at grid point  $q$ , summed over all atoms  $i$  of the molecule  $j$  under investigation.  $W_{probe,k}$  is the probe atom with radius 1 Å, charge +1, hydrophobicity +1, hydrogen bond donating +1, and hydrogen bond accepting +1.  $W_{ik}$  is the actual value of the physicochemical property  $k$  of atom  $i$ .  $r_{iq}^2$  is the mutual distance between the probe atom at grid point  $q$  and atom  $i$  of the test molecule.  $\alpha$  is the attenuation factor whose optimal value is normally between 0.2 and 0.4, with a default value of 0.3.<sup>27,28</sup>

## PLS analysis

To linearly correlate the CoMFA and CoMSIA fields to all types of molecular properties, PLS analysis<sup>29</sup> was carried out with the leave-one-out and leave-group-out (ten compound groups) cross-validation methods to determine cross-validated  $r^2$  ( $q^2$ ) values and the optimal number of components, which usually corresponds to the highest cross-validated coefficient ( $q^2$ ) and the smallest root mean predictive sum of squared errors for the final non-cross-validated models. To avoid over-fitting the models, a higher number of components was accepted only when the  $q^2$  differences between two components were larger than 10%. Bootstrapping analysis (ten runs) was also utilized to calculate confidence intervals for the  $r^2$  and standard errors of estimate to further assess the robustness of the derived models.<sup>20,29</sup> Equations 2 and 3 of  $q^2$  and standard errors of estimate, respectively, are given below.

$$q^2 = 1 - \frac{\sum (Y_{obs} - Y_{pre})^2}{\sum (Y_{obs} - Y_{mean})^2} \quad (2)$$

where  $Y_{obs}$ ,  $Y_{pre}$ , and  $Y_{mean}$  are experimental, predicted, and mean values of the target property for pIC<sub>50</sub> (FIXa) or logIC<sub>50</sub> (FXa/FIXa), respectively, and

$$SEE = \sqrt{\frac{PRESS}{n - c - 1}} \quad (3)$$

**Table 2** The experimental inhibitory activity, selectivity, and the predicted values

Compd no	Meas pIC <sub>50</sub> (FIXa)	Pred for 3D-QSAR				Meas logIC <sub>50</sub> (FXa/FIXa)	Pred for 3D-QSSR			
		CoMFA	Res	CoMSIA	Res		CoMFA	Res	CoMSIA	Res
1	5.28	5.27	-0.01	5.30	0.02	1.28	1.45	0.17	1.35	0.07
2*#	6.30	6.55	0.25	6.61	0.31	1.44	1.17	-0.27	1.19	-0.25
3	5.08	5.18	0.10	5.21	0.13	-	-	-	-	-
4	6.00	5.89	-0.11	6.29	0.29	0.53	0.39	-0.14	0.55	0.02
5*#	6.77	7.07	0.30	7.19	0.42	-0.05	-0.35	-0.30	-0.17	-0.12
6	6.41	6.54	0.13	6.66	0.25	0.39	0.40	0.01	0.43	0.04
7	6.85	6.69	-0.16	6.56	-0.29	-0.30	-0.41	-0.11	-0.25	0.05
8	6.89	7.01	0.12	7.00	0.11	0.35	0.15	-0.20	0.36	0.01
9	6.85	6.92	0.07	6.90	0.05	2.55	2.26	-0.29	2.25	-0.30
10	6.38	6.32	-0.06	6.24	-0.14	-2.21	-2.37	-0.16	-2.38	-0.17
11	6.43	6.75	0.32	6.81	0.38	-2.80	-3.42	-0.62	-2.76	0.04
12	6.54	6.43	-0.11	6.36	-0.18	-2.76	-3.21	-0.45	-3.01	-0.25
13	6.57	6.75	0.18	6.83	0.26	1.17	1.37	0.20	1.12	-0.05
14	6.80	6.73	-0.07	6.69	-0.11	1.44	1.68	0.24	1.39	-0.04
15	6.47	6.30	-0.17	6.35	-0.12	0.58	0.91	0.33	0.67	0.09
16	7.00	7.27	0.27	7.47	0.47	0.38	0.19	-0.19	0.28	-0.10
17	7.30	7.25	-0.05	7.05	-0.25	1.05	1.49	0.44	1.00	-0.05
18	7.54	7.33	-0.21	7.3	-0.24	1.52	1.21	-0.31	1.67	0.15
19	6.57	6.23	-0.34	6.33	-0.24	0.81	0.99	0.18	0.89	0.08
20*#	7.01	6.77	-0.24	6.48	-0.53	0.86	0.70	-0.16	0.78	-0.08
21	7.55	7.56	0.01	7.67	0.12	1.11	1.47	0.36	1.01	-0.10
22	8.00	8.12	0.12	8.34	0.34	2.26	2.86	0.60	2.33	0.07
23	7.21	7.53	0.32	7.48	0.27	1.13	1.29	0.16	1.32	0.19
24	8.30	8.29	-0.01	8.25	-0.05	2.04	2.18	0.14	1.95	-0.09
25	7.60	7.87	0.27	7.98	0.38	1.72	1.89	0.17	1.70	-0.02
26	8.10	8.06	-0.04	8.26	0.16	2.33	2.00	-0.33	2.35	0.02
27	8.10	8.19	0.09	8.01	-0.09	2.17	2.08	-0.09	2.14	-0.03
28	8.30	8.55	0.25	8.64	0.34	2.20	2.75	0.55	2.29	0.09
29	8.40	8.13	-0.27	8.08	-0.32	2.34	2.05	-0.29	2.43	0.09
30	8.00	8.46	0.46	8.61	0.61	2.19	1.84	-0.35	2.25	0.06
31*#	8.40	8.34	-0.06	8.47	0.07	2.31	2.66	0.35	2.34	0.03
32	8.52	8.22	-0.30	8.19	-0.33	2.58	2.47	-0.11	2.45	-0.13
33	8.70	8.96	0.26	9.24	0.54	2.56	2.72	0.16	2.64	0.08
34	8.10	7.84	-0.26	7.69	-0.41	2.01	2.42	0.41	2.22	0.21
35	8.52	8.56	0.04	8.69	0.17	2.07	1.95	-0.12	2.05	-0.02
36	8.15	8.20	0.05	8.38	0.23	2.33	2.71	0.38	2.41	0.08
37	8.15	8.02	-0.13	7.88	-0.27	2.27	2.11	-0.16	2.20	-0.07
38*#	7.92	7.49	-0.43	7.37	-0.55	2.32	2.04	-0.28	2.30	-0.02
39	8.00	8.08	0.08	8.27	0.27	2.31	2.13	-0.18	2.19	-0.12
40	8.05	7.90	-0.15	8.35	0.30	2.24	2.23	-0.01	2.25	0.01
41	8.30	8.23	-0.07	8.07	-0.23	2.40	2.44	0.04	2.34	-0.06
42	8.10	8.01	-0.09	7.91	-0.19	2.55	2.86	0.31	2.60	0.05
43	7.96	8.02	0.06	8.35	0.39	2.07	2.22	0.15	2.01	-0.06
44	8.30	8.15	-0.15	8.03	-0.27	2.68	3.06	0.38	3.00	0.32
45	8.52	8.44	-0.08	8.27	-0.25	2.73	2.61	-0.12	2.66	-0.07
46	7.81	7.98	0.17	8.08	0.27	2.01	1.90	-0.11	1.99	-0.02
47	8.36	8.43	0.07	8.55	0.19	2.68	2.45	-0.23	2.53	-0.15
48	8.10	7.99	-0.11	8.47	0.37	2.23	1.96	-0.27	2.03	-0.20
49*#	6.02	6.22	0.20	6.21	0.19	1.36	0.83	-0.53	1.01	-0.35
50	4.92	4.53	-0.39	4.58	-0.34	-	-	-	-	-
51	5.82	6.00	0.18	6.21	0.39	0.60	0.64	0.04	0.59	-0.01
52	6.10	6.34	0.24	6.29	0.19	0.51	0.66	0.15	0.55	0.04
53	5.80	5.68	-0.12	5.53	-0.27	-	-	-	-	-

(Continued)



Table 2 (Continued)

Compd no	Meas pIC <sub>50</sub> (FIXa)	Pred for 3D-QSAR				Meas logIC <sub>50</sub> (FXa/FIXa)	Pred for 3D-QSSR			
		CoMFA	Res	CoMSIA	Res		CoMFA	Res	CoMSIA	Res
54	5.10	5.23	0.13	5.35	0.25	—	—	—	—	—
55	5.23	5.23	0.00	5.29	0.06	0.30	0.4	0.10	0.41	0.11
56	5.72	5.90	0.18	6.09	0.37	0.29	0.37	0.08	0.30	0.01
57*:#	5.42	5.33	−0.09	5.55	0.13	0.59	0.91	0.32	0.82	0.23
58	5.42	5.44	0.02	5.51	0.09	−0.80	−0.93	−0.13	−0.77	0.03
59	4.62	4.11	−0.51	4.03	−0.59	—	—	—	—	—
60*	4.64	4.86	0.22	4.93	0.29	—	—	—	—	—
61	6.17	6.01	−0.16	6.12	−0.05	0.96	1.11	0.15	1.06	0.10
62	6.15	6.19	0.04	6.21	0.06	1.02	0.94	−0.08	0.98	−0.04
63	5.14	5.04	−0.1	5.14	0	—	—	—	—	—
64	6.05	6.25	0.2	6.22	0.17	0.68	0.77	0.09	0.70	0.02
65	6.35	6.46	0.11	6.5	0.15	0.81	1.08	0.27	0.88	0.07
66	4.43	4.03	−0.40	4.07	−0.36	—	—	—	—	—
67	6.35	6.51	0.16	6.53	0.18	0.98	0.82	−0.16	0.89	−0.09
68	6.60	6.61	0.01	6.44	−0.16	1.72	1.51	−0.21	1.56	−0.16
69	6.05	6.00	−0.05	6.5	0.45	0.94	1.35	0.41	1.04	0.10
70*	4.95	4.81	−0.14	4.83	−0.12	—	—	—	—	—
71	6.35	6.37	0.02	6.38	0.03	1.10	1.19	0.09	1.20	0.10
72	7.12	7.17	0.05	7.01	−0.11	0.15	0.22	0.07	0.24	0.09
73	7.51	7.47	−0.04	7.55	0.04	0.25	0.09	−0.16	0.16	−0.09
74	4.67	5.01	0.34	5.14	0.47	—	—	—	—	—
75	5.98	6.05	0.07	6.03	0.05	0.50	0.43	−0.07	0.49	−0.01
76	6.46	6.44	−0.02	6.52	0.06	0.52	0.61	0.09	0.60	0.08
77*	4.95	4.87	−0.08	4.69	−0.26	—	—	—	—	—
78*:#	5.70	5.73	0.03	5.81	0.11	1.36	1.98	0.62	1.64	0.28
79	7.40	7.29	−0.11	7.3	−0.1	0.04	0.01	−0.03	0.01	−0.03
80*:#	7.68	8.19	0.51	8.23	0.55	1.64	1.29	−0.35	1.05	−0.59
81	7.70	7.75	0.05	7.85	0.15	−0.07	−0.3	−0.23	−0.11	−0.04
82	7.55	7.53	−0.02	7.61	0.06	0.07	0.15	0.08	0.05	−0.02
83	7.44	7.41	−0.03	7.42	−0.02	0.01	0.10	0.09	0.10	0.09
84	8.52	8.50	−0.02	8.71	0.19	0.92	1.13	0.21	1.00	0.08

**Notes:** \*Test samples for 3D-QSAR model validation; #test samples for 3D-QSSR model validation.

**Abbreviations:** 3D-QSAR, three-dimensional quantitative structure–activity relationship; 3D-QSSR, three-dimensional structure–selectivity relationship; IC<sub>50</sub>, half-maximal inhibitory concentration; pIC<sub>50</sub>, negative logarithm of IC<sub>50</sub>; logIC<sub>50</sub>, negative logarithm of pIC<sub>50</sub>; FXa, Factor Xa; FIXa, Factor IXa; CoMFA, comparative molecular field analysis; CoMSIA, comparative similarity indices analysis; Res, residue; Pred, predicted activity; Meas, measured activity; Compd no, compound number.

where  $n$ ,  $c$ , and  $PRESS$  (predicted sum of squares) represent mean number of compounds, number of components, and  $\sum (Y_{obs} - Y_{pre})^2$ , respectively.

### Predictive correlation coefficient

The  $q^2$  has been a good indicator of the accuracy of actual predictions:  $q^2 > 0.6$  means a fairly good model,  $q^2 = 0.4–0.6$  means a questionable model, and  $q^2 < 0.4$  means a poor model.<sup>30</sup>  $q^2$  is a useful but not sufficient criterion for model validation, so external test sets<sup>23</sup> were needed to estimate the model's predictive ability. The predictive correlation coefficient ( $r^2_{pred}$ ), based on the molecules of the test set, was calculated by Equations 4 and 5, shown below:

$$r^2_{pred} = 1 - (PRESS/SD) \quad (4)$$

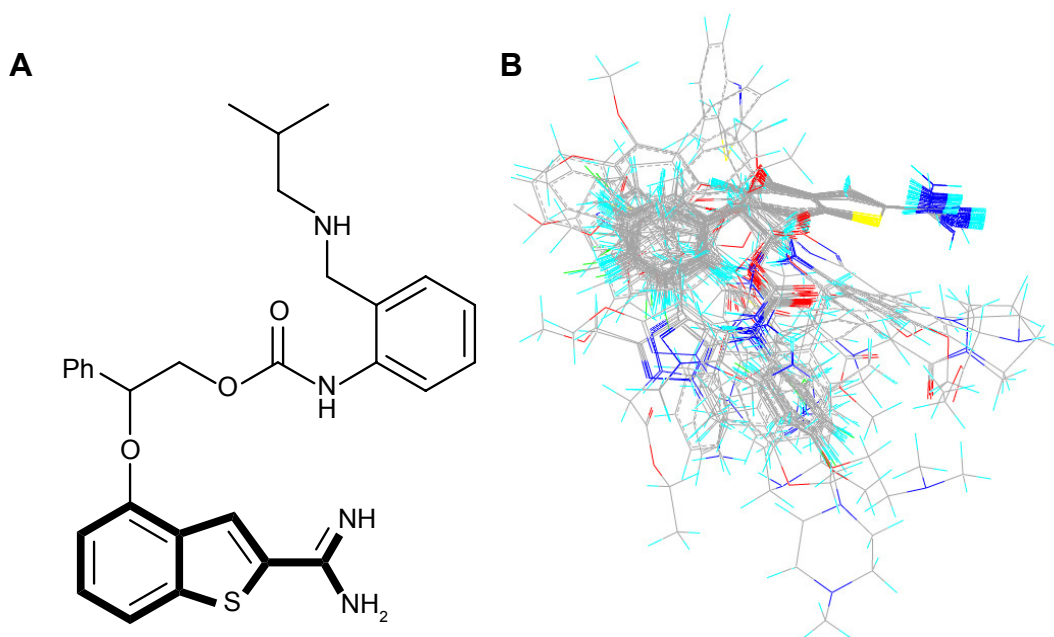
$$SD = \sum (Y_{obs(test)} - \bar{Y}_{training})^2 \quad (5)$$

where  $Y_{obs(test)}$  means the observed activity values of the test set, and  $\bar{Y}_{training}$  indicates the average activity values of the training set.

## Results and discussion

### CoMFA analysis for 3D-QSAR and 3D-QSSR models

The most active Compound 33 was selected as the template for alignment (Figure 1). The CoMFA model provided a cross-validation  $q^2$  value of 0.713 with six components, and the  $r^2$  value of 0.945 for the inhibitory activity as characterized by pIC<sub>50</sub> (FIXa), whereas a model with the  $q^2$  value of 0.689 for



**Figure 1** A 2-amidinobenzothiophene ring with structural rigidity, comprising the common core of all molecules in the data set, was selected as the common substructure to align all of the molecules **(A)** Template used for molecular alignment of benzothiophene derivatives (Compound 33). **(B)** Molecular alignment.

five components and the  $r^2$  value of 0.942 was obtained for the selectivity characterized by  $\log IC_{50}$  (FXa/FIXa) (Table 2).

After region focusing, the new models were generated with the predictive power ( $q^2$ ) value increased, the resolution enhanced, the grid spacing tighter, and the stability improved at a higher number of components. The CoMFA contours for Compound 33 after region focusing are shown in Figure 2. The internal validity were increased from 0.713 to 0.753 for 3D-QSAR, and from 0.689 to 0.770 for 3D-QSSR, and the group cross-validation was increased from 0.740 to 0.768 and 0.581 to 0.765, respectively, with the non-validated  $r^2$  increased from 0.942 to 0.965 for the 3D-QSSR model and a negligible effect for the 3D-QSAR model increasing its value from 0.945 to 0.940 (Table 3). The inhibitory activity values predicted for the tested compounds were in good agreement with the experimental values (Figure 3). The  $r^2_{pred}$  values of 0.757 (3D-QSAR) and 0.678 (3D-QSSR) further confirmed that the two models were both reliable and accurate, with higher predictive capacity. Therefore, the  $pIC_{50}$  (FIXa) and  $\log IC_{50}$  (FXa/FIXa) model will be used to predict the activity and guide future synthetic efforts on novel potent and selective FIXa inhibitors.

### CoMSIA analysis for 3D-QSAR and 3D-QSSR models

Eighteen and eleven CoMSIA models were generated for 3D-QSAR and 3D-QSSR, respectively, using combinations

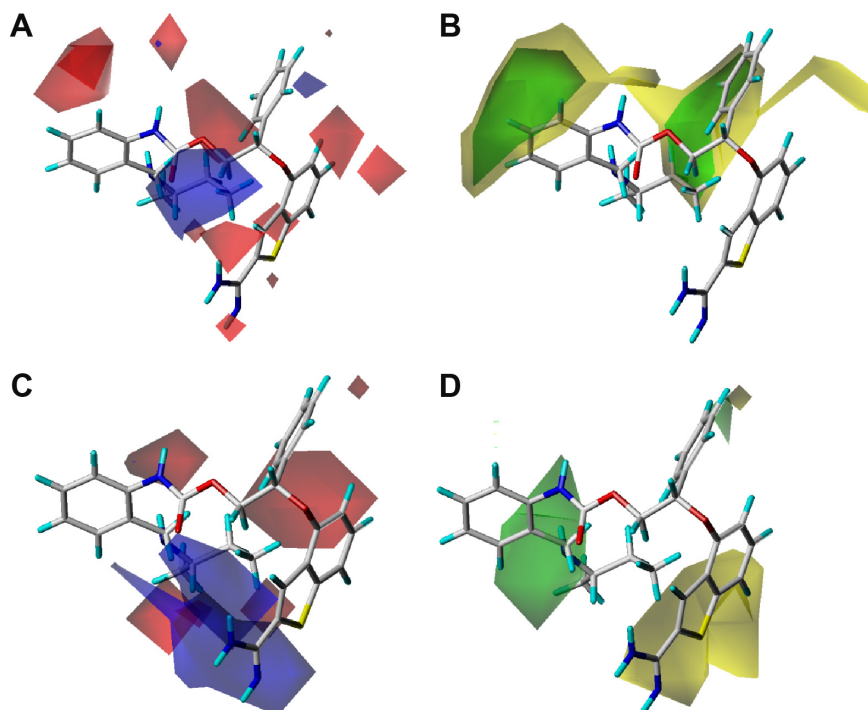
of two, three, four, and all five descriptor fields, as shown in Table 4. In 3D-QSAR, Model 11, which was based on steric, hydrophobic, and hydrogen-bond acceptor fields, was found to be the most accurate, yielding a  $q^2$  value of 0.735 and an  $r^2$  value of 0.966, and the group cross- $q^2$  value of 0.723 and bootstrapped value of  $0.980 \pm 0.005$  confirmed model accuracy. In comparison, Model 9 in 3D-QSSR based on steric, hydrophobic, and hydrogen-bond acceptor and donor fields, was the most accurate, with a  $q^2$  value of 0.837, an  $r^2$  value of 0.973, a group cross- $q^2$  value of 0.854, and a bootstrapped value of  $0.977 \pm 0.007$  (Table 5). The predicted values were also consistent with the experimental data (Figure 3). These two models were subsequently selected to generate the final CoMSIA models.

### 3D contour maps of QSAR and QSSR

Scatter plots gave a visual impression of the region as a whole, with the color of each point indicating the field intensity. The results of QSAR models are presented in the contour maps, as shown in Figures 2 and 4, while 3D-QSSR contour maps are shown in Figures 2 and 5.

#### CoMFA contour maps for $pIC_{50}$ (FIXa) in the 3D-QSAR model

Figure 2 shows the CoMFA steric and electrostatic fields of Compound 33 for the 3D-QSAR model. The green



**Figure 2** CoMFA fields for  $pIC_{50}$  (FIXa) model (**A** and **B**) and  $\log IC_{50}$  (FXa/FIXa) model (**C** and **D**) in combination with Compound 33 after region focusing. Electrostatic fields (**A** and **C**): blue fields indicate electropositive groups favored, red fields indicate electronegative groups favored. Steric fields (**B** and **D**): green fields indicate steric bulk favored, yellow fields indicate steric bulk disfavored.

**Abbreviations:**  $IC_{50}$ , half-maximal inhibitory concentration;  $pIC_{50}$ , negative logarithm of  $IC_{50}$ ;  $\log IC_{50}$ , negative logarithm of  $IC_{50}$  (FIXa/FXa); FXa, Factor Xa; FIXa, Factor IXa; CoMFA, comparative molecular field analysis; CoMSIA, comparative similarity indices analysis; 3D-QSAR, three-dimensional quantitative structure–activity relationship; 3D-QSSR, three-dimensional quantitative structure–selectivity relationship.

contours located around the two aromatic rings linked to the side chain of benzothiophene characterized the regions where bulky substituents would increase biological activity, and this could explain Compounds 24–47 consisting of two aromatic rings about 100 times more active than other compounds, such as Compounds 1–4 and Compounds 52–61 without the substituents. That is to say, the two aromatic rings are essential groups of 2-amidinobenzothiophene derivatives for its inhibitory activity on FIXa.

The red and blue contours in Figure 2 depict the favorable sites for electronegative and electropositive groups, respectively. A large blue contour located around the distal side chain at the 2'-position of R group (phenyl ring) in Compound 33 indicated regions where positively charged groups were favorable for the activity, and that is why Compound 22 was less active than the Compound 33 with isobutyl group at the site. Many red contours either encircling the benzothiophene ring or around the side chain indicated that electron-rich substituents for these positions were favorable for the activity. That is the reason why Compound 84, linked with the negatively charged group, showed better activity, only a little lower than the most active Compound 33.

### CoMSIA contour maps for $pIC_{50}$ (FIXa) in the 3D-QSAR model

The best CoMSIA model contour maps of the Compound 33, derived using steric, hydrophobic, and hydrogen bond acceptor fields, are shown in Figure 4. Its steric contour (Figure 4A) was more or less similar to that of CoMFA. Hydrophobic contour maps from CoMSIA are shown in Figure 4B, in which the purple and the gray contours depicted favored and disfavored sites for hydrophobic substituents, respectively. The Compound 33 was surrounded by the gray contours, suggesting that it is essential for the benzothiophene ring to link with some hydrophilic groups such as amino, carboxyl, and oxygen atoms. Therefore, Compounds 81–84 exhibited better activity than Compounds 54–61. Figure 4C shows the CoMSIA hydrogen-bond acceptor fields, denoted by magenta and cyan contours. Magenta contours represent regions where hydrogen-bond acceptor substituents were preferred, and cyan contours indicate unfavorable regions. One large magenta contour around the side chain indicates that the carboxyl group was necessary for potencies, and that is why Compounds 20–48 showed more potency than Compounds 62–73 without carboxyl groups at the side chain. The CoMSIA steric field, hydrogen-bond acceptor,

**Table 3** Statistical results of CoMFA for 3D-QSAR and 3D-QSSR models

	3D-QSAR		3D-QSSR	
	CoMFA-1 (before region focusing)	CoMFA-2 (after region focusing)	CoMFA-1 (before region focusing)	CoMFA-2 (after region focusing)
<b>PLS statistics</b>				
LOO cross $q^2$ /SEP	0.713/0.638	0.753/0.592	0.689/0.684	0.770/0.592
Group cross $q^2$ /SEP	0.740/0.608	0.768/0.607	0.581/0.794	0.765/0.599
Non-validated $r^2$ /SEE	0.945/0.278	0.940/0.292	0.942/0.295	0.965/0.232
F	205.046	215.821	224.464	309.797
$r^2_{\text{bootstrap}}$	0.970 $\pm$ 0.011	0.969 $\pm$ 0.011	0.961 $\pm$ 0.016	0.976 $\pm$ 0.007
$S_{\text{bootstrap}}$	0.255 $\pm$ 0.120	0.200 $\pm$ 0.110	0.236 $\pm$ 0.136	0.199 $\pm$ 0.111
N	6	6	6	6
<b>Field distribution, %</b>				
Steric	74.8	66.3	65.4	56.7
Electrostatic	25.2	33.7	34.6	43.3
$r^2_{\text{pred}}$		0.757		0.678

**Notes:** N is the optimal number of components  $r^2_{\text{bootstrap}}$  is the average of correlation coefficient for 100 samplings using the bootstrapping procedure.  $S_{\text{bootstrap}}$  is the average standard error of estimates for 100 samplings using the bootstrapping procedure.

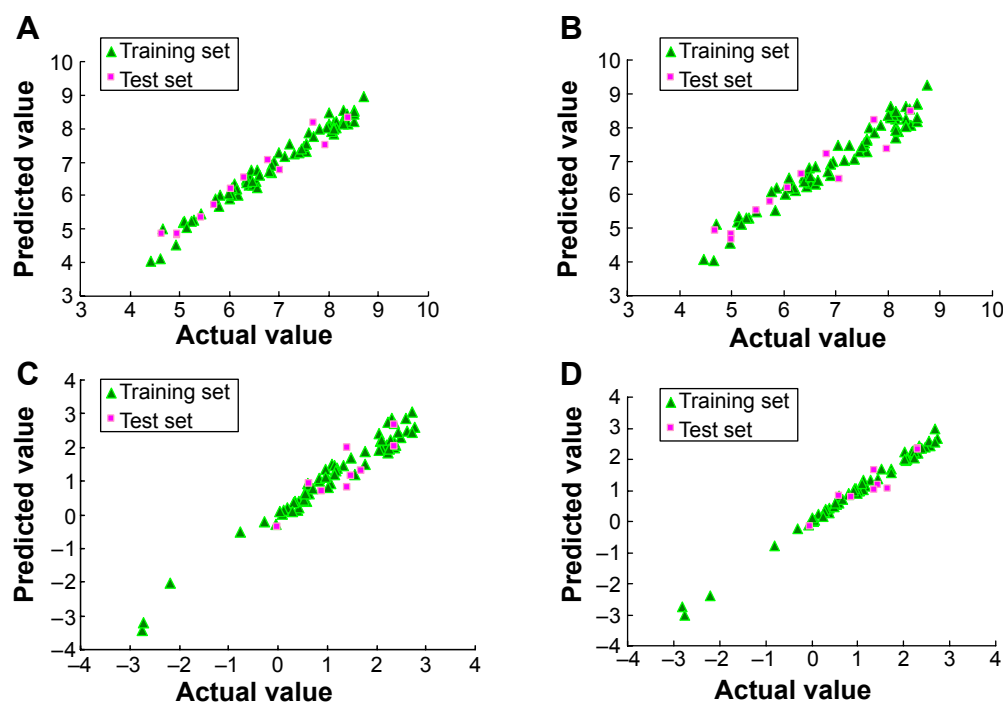
**Abbreviations:** CoMFA, comparative molecular field analysis; 3D-QSAR, three-dimensional structure–activity relationship; 3D-QSSR, three-dimensional structure–selectivity relationship; SEE, standard error of estimation; F, the F-test value;  $r^2_{\text{pred}}$ , the predictive correlation coefficient;  $q^2$ , the cross-validated coefficient;  $r^2$ , the non-cross-validated coefficient; PLS, partial least squares; LOO leave-one-out.

and hydrophobic field explained 27.0%, 26.5%, and 46.5% of the variances, respectively.

#### CoMFA contour maps for $\log IC_{50}$ (FXa/FIXa) in the 3D-QSSR model

The CoMFA steric contour map of the Compound 33 displayed a large green polyhedron at the 2'-position of the

phenyl in the R group, suggesting that bulky substituents linked to this position would improve the selectivity of compounds on FIXa against FXa (Figure 2). Compound 33 inserted a butylaminomethyl group at the 2'-position into the green area, and Compounds 46 and 47 similarly inserted an ethylaminomethyl and an *i*-propylaminomethyl group into the same area, respectively, and therefore showed the best



**Figure 3** Graphs of experimental versus predicted  $pIC_{50}$  (FIXa) via (A) CoMFA, and (B) CoMSIA for 3D-QSAR model, and graphs of experimental versus predicted  $\log IC_{50}$  (FXa/FIXa) via (C) CoMFA, and (D) CoMSIA for 3D-QSSR model.

**Abbreviations:**  $IC_{50}$ , half-maximal inhibitory concentration;  $pIC_{50}$ , negative logarithm of  $IC_{50}$ ; FIXa, Factor IXa; CoMFA, comparative molecular field analysis; CoMSIA, comparative similarity indices analysis; 3D-QSAR, three-dimensional quantitative structure–activity relationship;  $\log IC_{50}$ , negative logarithm of  $pIC_{50}$ ; FXa, Factor Xa; 3D-QSSR, three-dimensional quantitative structure–selectivity relationship.

**Table 4** Results of CoMSIA models using combinations of the five field descriptors for 3D-QSAR model

Model	Descriptors	LOO cross $q^2$ /SEP	Group cross $q^2$ /SEP	Bootstrapped $r^2$	Bootstrapped SEE	Non-validated $r^2$ /SEE
1	S and E	0.534/0.797	0.519/0.810	0.727±0.026	0.616±0.236	0.720/0.618
2	D and A	0.627/0.713	0.626/0.714	0.832±0.022	0.482±0.199	0.766/0.565
3	S, E, and H	0.573/0.763	0.581/0.756	0.810±0.078	0.508±0.310	0.735/0.601
4	S, E, and A	0.599/0.739	0.617/0.723	0.808±0.045	0.496±0.234	0.737/0.599
5	S, E, and D	0.646/0.695	0.644/0.697	0.836±0.035	0.477±0.204	0.766/0.565
6	D, A, and H	0.701/0.652	0.700/0.653	0.973±0.007	0.199±0.092	0.954/0.255
7	D, A, and S	0.707/0.641	0.710/0.639	0.975±0.010	0.193±0.104	0.971/0.232
8	D, A, and E	0.579/0.763	0.603/0.736	0.805±0.022	0.516±0.172	0.748/0.586
9	S, D, and H	0.716/0.636	0.705/0.639	0.932±0.015	0.308±0.139	0.960/0.240
10	E, D, and H	0.576/0.765	0.539/0.793	0.922±0.017	0.336±0.151	0.883/0.402
11	S, A, and H	0.735/0.613	0.723/0.628	0.980±0.005	0.173±0.091	0.966/0.220
12	E, A, and H	0.569/0.766	0.591/0.746	0.860±0.047	0.449±0.254	0.755/0.578
13	S, E, D, and A	0.643/0.698	0.636/0.704	0.819±0.035	0.511±0.246	0.764/0.567
14	S, E, D, and H	0.644/0.697	0.645/0.696	0.842±0.036	0.470±0.245	0.777/0.552
15	S, E, A, and H	0.637/0.713	0.646/0.704	0.958±0.012	0.245±0.131	0.933/0.307
16	D, A, H, and S	0.684/0.656	0.689/0.651	0.875±0.019	0.403±0.187	0.818/0.498
17	D, A, H, and E	0.610/0.730	0.612/0.728	0.851±0.042	0.451±0.238	0.781/0.547
18	S, E, D, A, and H	0.662/0.688	0.701/0.647	0.954±0.014	0.252±0.138	0.929/0.316

**Abbreviations:** S, steric field; E, electrostatic field; D, H-bond donor field; A, H-bond acceptor field; H, hydrophobic field; CoMSIA, comparative similarity indices analysis; 3D-QSAR, three-dimensional structure–activity relationship; LOO, leave-one-out cross-validation; SEE, standard errors of estimate; SEP, standard errors of prediction.

selectivity. A large yellow polyhedron around the benzo-thiophene ring indicated that bulk was disfavored here for the selectivity on FIXa, which could explain why the Compounds 51, 57, and 58 with some bulk groups substituted at the benzothiophene ring showed lower selectivity.

Figure 2 also indicated that the compounds with electropositive groups at the 2'-position increased not only the activity but also the selectivity. Both the Compound 32 with a methylaminomethyl group substituted at 2'-position of the phenyl ring and Compound 33 with a butylaminomethyl substituent showed relatively higher activity and selectivity. These results also could explain why the Compounds 44 and 47 had the best selectivity and much more activity.

#### CoMSIA contour maps for logIC<sub>50</sub> (FXa/FIXa) in the QSSR model

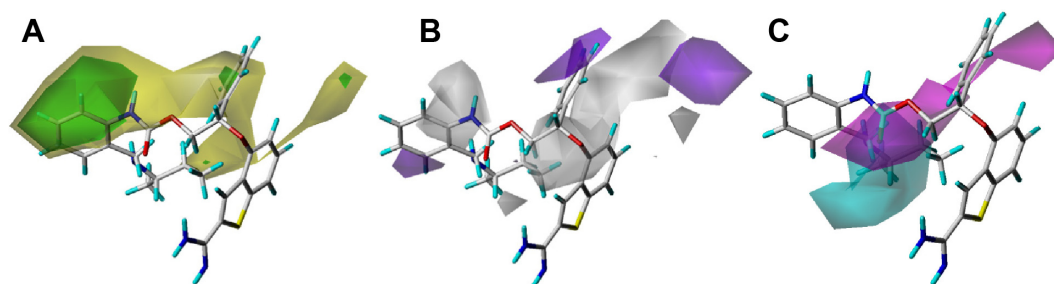
The best CoMSIA model contour maps, derived using steric, hydrophobic, and hydrogen bond donor and hydrogen bond acceptor fields, are shown in Figure 5. The steric contour plots (Figure 5A) correlated well with its CoMFA contour maps. The hydrophobic contour maps with the purple contour indicating favored sites and the gray contour indicating disfavored sites are shown in Figure 5B. The combined information from Figures 4B and 5B indicated that the hydrophobic substituents at the 3'-position and the hydrophilic groups at the 2'-position of the phenyl ring increased not only increased the activity, but also increased the selectivity. That is the reason

**Table 5** Results of CoMSIA models using combinations of the five field descriptors for the 3D-QSSR model

Model	Descriptors	LOO cross $q^2$ /SEP	Group cross $q^2$ /SEP	Bootstrapped $r^2$	Bootstrapped SEE	Non-validated $r^2$ /SEE
1	S and E	0.709/0.661	0.712/0.657	0.940±0.025	0.280±0.165	0.907/0.373
2	D and A	0.677/0.694	0.686/0.698	0.927±0.031	0.302±0.171	0.900/0.374
3	S, E, and H	0.789/0.567	0.756/0.610	0.971±0.012	0.198±0.089	0.953/0.269
4	S, E, and A	0.786/0.568	0.775/0.582	0.953±0.014	0.259±0.131	0.922/0.343
5	S, E, and D	0.742/0.627	0.724/0.649	0.966±0.013	0.208±0.124	0.941/0.300
6	S, E, D, and A	0.816/0.525	0.790/0.561	0.965±0.009	0.255±0.117	0.938/0.306
7	S, E, D, and H	0.790/0.566	0.788/0.568	0.974±0.009	0.190±0.102	0.961/0.244
8	S, E, A, and H	0.789/0.568	0.782/0.572	0.954±0.017	0.263±0.137	0.955/0.261
9	D, A, H, and S	0.837/0.499	0.854/0.472	0.977±0.007	0.169±0.075	0.973/0.204
10	D, A, H, and E	0.822/0.52	0.774/0.587	0.975±0.009	0.183±0.096	0.958/0.254
11	S, E, D, A, and H	0.820/0.523	0.808/0.541	0.970±0.010	0.182±0.101	0.961/0.245

**Abbreviations:** S, steric field; E, electrostatic field; D, H-bond donor field; A, H-bond acceptor field; H, hydrophobic field; CoMSIA, comparative similarity indices analysis; 3D-QSSR, three-dimensional structure–selectivity relationship; LOO, leave-one-out cross-validation; SEE, standard errors of estimate; SEP, standard errors of prediction.





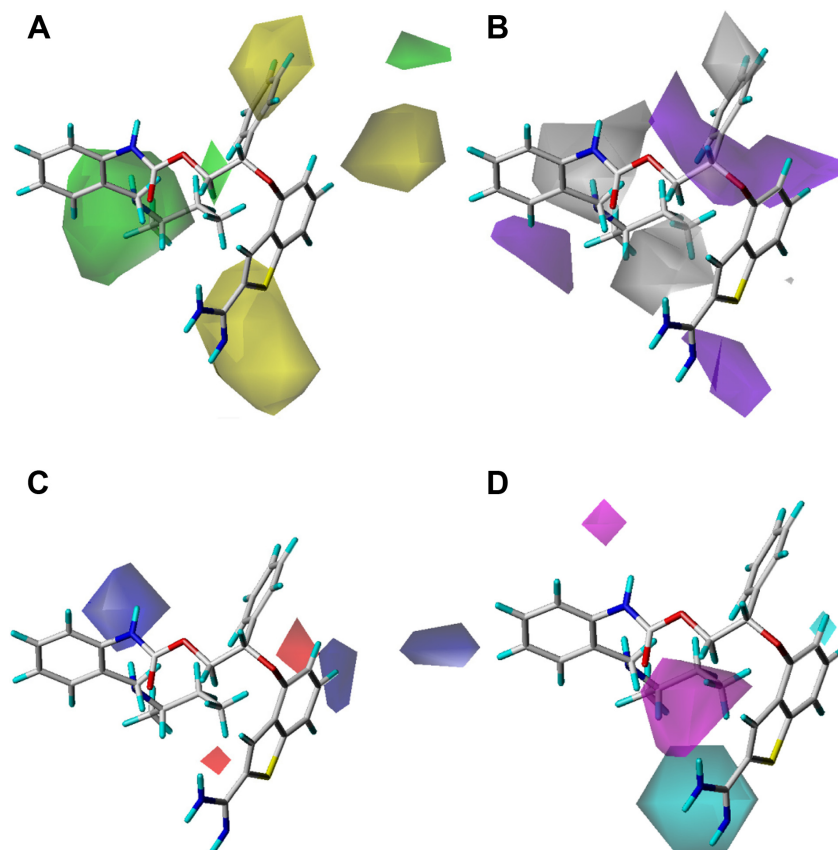
**Figure 4** CoMSIA fields for  $pIC_{50}$  (FIXa) model. The CoMSIA fields from Model 11 are shown with active Compound 33.

**Notes:** (A) Steric fields; green indicates steric bulk favored, yellow indicates steric bulk disfavored. (B) Hydrophobic fields; purple indicates hydrophobic groups favored, gray fields indicate hydrophilic groups favored. (C) H-bond acceptor fields; magenta indicates acceptor favored, cyan indicates acceptor disfavored.

**Abbreviations:**  $IC_{50}$ , half-maximal inhibitory concentration; CoMSIA, comparative similarity indices analysis;  $pIC_{50}$ , negative logarithm of  $IC_{50}$ ; FIXa, Factor IXa.

why Compounds 44 and 47, possessing alkylamino and methyl group at the 2', 3'-position, respectively, showed the highest selectivity and almost the strongest activity, nearly three times more than Compound 45 with a hydrophilic ethoxyl group at the 3'-position and a hydrophobic methyl group at the 2'-position. Figure 5C and D give

the combined information that the amide group at the side chain was an essential group for a high selectivity on FIXa. The contributions of the steric, hydrophobic, hydrogen bond donor, and hydrogen bond acceptor to the final model were 25.5%, 23.0%, 23.6%, and 27.9%, respectively.



**Figure 5** CoMSIA fields for  $\log IC_{50}$  (FXa/FIXa) model. The CoMSIA fields from Model 9 are shown with active Compound 33.

**Notes:** (A) Steric fields; green indicates steric bulk favored, yellow indicates steric bulk disfavored. (B) Hydrophobic fields; purple indicates hydrophobic groups favored, gray regions indicate hydrophilic groups favored. (C) H-bond donor fields; blue indicates donor favored, red indicates donor disfavored. (D) H-bond acceptor fields; magenta indicates acceptor favored, cyan indicates acceptor disfavored.

**Abbreviations:**  $IC_{50}$ , half-maximal inhibitory concentration; CoMSIA, comparative similarity indices analysis;  $\log IC_{50}$ , negative logarithm of  $pIC_{50}$ ; FXa, Factor Xa; FIXa, Factor IXa.

**Table 6** Structures and predicted  $pIC_{50}$  values of newly designed derivatives

Compound	Substitutes				Pred $pIC_{50}$ (FIXa)		Pred $\log IC_{50}$ (FXa/FIXa)	
	n	X	R <sub>1</sub>	R <sub>2</sub>	CoMFA	CoMSIA	CoMFA	CoMSIA
33	1	O		H	8.96	9.24	2.72	2.64
D1	1	O			9.01	9.35	3.48	2.97
D2	2	NH			10.32	10.59	4.33	4.26
D3	1	NH			9.55	9.67	4.04	4.09
D4	1	NH			9.33	9.52	4.11	4.23
D5	1	NH			9.40	9.48	3.95	3.79
D6	1	NH			9.20	9.23	4.12	4.00
D7	1	NH			9.09	9.44	4.12	3.99
D8	2	O			10.46	10.49	4.45	4.32
D9	1	NH			9.00	9.19	4.21	4.03
D10	1	NH			9.15	9.11	3.97	3.98
D11	1	O			9.75	9.80	4.04	3.97
D12	2	O			10.50	10.66	4.34	4.41
D13	1	O			9.07	9.01	3.89	3.77
D14	2	O			9.35	9.47	3.92	3.89
D15	1	O			9.39	9.51	4.22	4.05
D16	2	O			9.55	9.50	3.98	4.01

**Abbreviations:**  $IC_{50}$ , half-maximal inhibitory concentration;  $pIC_{50}$ , negative logarithm of  $IC_{50}$ ;  $\log IC_{50}$ , negative logarithm of  $pIC_{50}$ ; FXa, Factor Xa; FIXa, Factor IXa; CoMFA, comparative molecular field analysis; CoMSIA, comparative similarity indices analysis; Pred, predicted activity.

## Designing potent derivatives

Based on the 3D-QSARs and 3D-QSSRs revealed by the present study, 16 novel 2-amidinobenzothiophene derivatives were designed. These molecules were aligned to the database, and their activity and selectivity on FIXa were predicted by the best CoMFA and CoMSIA models previously established in the current study. All of the designed molecules showed better activity and selectivity compared with Compound 33, especially Compounds D2, D8, and D12, which showed 40 times more activity and 60 times more selectivity, respectively. The chemical structures and predicted  $pIC_{50}$  and  $logIC_{50}$  (FXa/FIXa) values of these compounds are shown in Table 6. The predictive results may validate the structure–activity and structure–selectivity relationships in this present work.

## Conclusion

Based on a series of 2-amidinobenzothiophenes with FIXa and FXa inhibitory activity, and the selectivity on FIXa, the predictive 3D-QSAR and 3D-QSSR models with high cross-validation correlation coefficient  $q^2$  and non-cross-validation correlation coefficient  $r^2$  values have been developed. Along with further testing, the obtained 3D-QSAR and 3D-QSSR models showed their significance in predicting the inhibitory activity and selectivity of amidinobenzothiophene derivatives on FIXa. It was concluded that both the electrostatic and hydrophobic fields should be considered more appropriately; in particular, the electropositive and hydrophilic substituents at the 2'-position and the hydrophobic groups at the 3'-position of the phenyl ring would improve the activity and selectivity of the compounds. On the basis of the 3D-QSPRs, 16 newly designed molecules predicted higher activity and selectivity than the best active Compound 33 described in the literature.

## Acknowledgments

This study was supported by a research grant from the National Natural Science Fund of China (No 81373285) and a research project grant from the Industry-Academy-research projects of Zhongshan (No 2012cxy020).

## Disclosure

The authors declare no conflicts of interest in this work.

## References

1. Feuerstein GZ, Nichols AJ, Ruffolo RR. Cardiovascular & renal: clopidogrel: a novel antiplatelet drug for prevention and treatment of thrombotic disorders. *Expert Opin Inv Drug*. 1995;4:425–430.
2. Feuerstein GZ, Ruffolo RR. The antiplatelet GPIIb/IIIa antibody: TE3 A pioneer drug for the treatment of arterial thrombotic disorders. *Expert Opin Inv Drug*. 1994;3:1163–1169.
3. Topol EJ. Toward a new frontier in myocardial reperfusion therapy: emerging platelet preeminence. *Circulation*. 1998;97:211–218.
4. Gómez-Outes A, Suárez-Gea ML, Lecumberri R, Rocha E, Pozo-Hernández C, Vargas-castrillón E. New parenteral anticoagulants in development. *Ther Adv Cardiovasc Dis*. 2011;5:33–59.
5. Hirsh J, O'Donnell M, Weitz JI. New anticoagulants. *Blood*. 2005;105:453–463.
6. Quan ML, Smallheer JM. The race to an orally active Factor Xa inhibitor: recent advances. *Curr Opin Drug Discov Devel*. 2004;7:460–469.
7. Frederick R, Pochet L, Charlier C, Masereel B. Modulators of the coagulation cascade: focus and recent advances in inhibitors of tissue factor, factor VIIa and their complex. *Curr Med Chem*. 2005;12:397–417.
8. Jesty J, Nemerson Y. The pathways of blood coagulation. In: Beutler E, Lichman MA, Collar BS, Kipps TJ, editors. New York: McGraw Hill; 1995:317–325.
9. Johnson DJ, Langdown J, Huntington JA. Molecular basis of factor IXa recognition by heparin-activated antithrombin revealed by a 1.7-Å structure of the ternary complex. *Proc Natl Acad Sci U S A*. 2010;107: 645–650.
10. Lawson JH, Mann KG. Cooperative activation of human factor IX by the human extrinsic pathway of blood coagulation. *J Biol Chem*. 1991; 266:11317–11327.
11. Neuenschwander P. Coagulation cascade: intrinsic factors. In: Laurent GL, Shapiro SD, editors. *Encyclopedia of Respiratory Medicine*. Oxford: Elsevier Ltd; 2006:509–514.
12. McKean ML, Adelman SJ. Future therapies for the prevention and treatment of venous and arterial thrombosis. *Expert Opin Investig Drugs*. 1998;7:687–690.
13. Mizurini DM, Francischetti IM, Andersen JF, Monteiro RQ. Nitrophenol 2, a factor IX(a)-directed anticoagulant, inhibits arterial thrombosis without impairing haemostasis. *Thromb Haemost*. 2010;104:1116–1123.
14. Spanier TB, Chen JM, Oz MC, et al. Selective anticoagulation with active site-blocked factor IXa suggests separate roles for intrinsic and extrinsic coagulation pathways in cardiopulmonary bypass. *J Thorac Cardiovasc Surg*. 1998;116:860–869.
15. Wang S, Beck R, Blench T, et al. Studies of benzothiophene template as potent factor IXa (FIXa) inhibitors in thrombosis. *J Med Chem*. 2010; 53:1465–1472.
16. Wang S, Beck R, Burd A, Blench T, Marlin F, et al. Structure based drug design: development of potent and selective factor IXa (FIXa) inhibitors. *J Med Chem*. 2010;53:1473–1482.
17. Cramer RD, Patterson DE, Bunce JD. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J Am Chem Soc*. 1998;110:5959–5967.
18. 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;37:4130–4146.
19. Srivastava V, Kumar A, Mishra BN, Siddiqi MI. CoMFA and CoMSIA 3D-QSAR analysis of DMDP derivatives as anti-cancer agents. *Bioinform*. 2008;2:384–391.
20. Hansch C, Smith N, Engle R, Wood H. Quantitative structure-activity relationships of antineoplastic drugs: nitrosoureas and triazenoimidazoles. *Cancer Chemother Rep*. 1972;56:443–456.
21. Xiao A, Zhang Z, An L, Xiang Y. 3D-QSAR and docking studies of 3-arylquinazolinethione derivatives as selective estrogen receptor modulators. *J Mol Model*. 2008;14:149–159.
22. Hao M, Zhang SW, Yang W. Investigation on the binding mode of benzothiophene analogues as potent factor IXa (FIXa) inhibitors in thrombosis by CoMFA, docking and molecular dynamic studies. *J Enzyme Inhib Med Chem*. 2011;26:792–804.
23. Cho SJ, Tropsha A. Cross-validated R2-guided region selection for comparative molecular field analysis: a simple method to achieve consistent results. *J Med Chem*. 1995;38:1060–1066.
24. Zheng J, Xiao G, Guo J, et al. Exploring QSARs for 5-lipoxygenase (5-LO) inhibitory activity of 2-substituted 5-hydroxyindole-3-carboxylates by CoMFA and CoMSIA. *Chem Biol Drug Des*. 2011; 78:314–321.



25. Huang X, Xu L, Luo X et al. Elucidating the inhibiting mode of AHPBA derivatives against HIV-1 protease and building predictive 3D-QSAR models. *J Med Chem*. 2002;45:333–343.
26. Klebe G, Abraham U. Comparative molecular similarity index analysis (CoMSIA) to study hydrogen-bonding properties and to score combinatorial libraries. *J Comput Aided Mol Des*. 1999;13:1–10.
27. Böhm M, St rzebecher J, Klebe G. Three-dimensional quantitative structure-activity relationship analyses using comparative molecular field analysis and comparative molecular similarity indices analysis to elucidate selectivity differences of inhibitors binding to trypsin, thrombin, and factor Xa. *J Med Chem*. 1999;42:458–477.
28. Woodard B, Saleh MA. Three-dimensional quantitative structure activity relationships of quorum-sensing and biofilm inhibitors in gram-negative bacteria. *J Environ Sci Health B*. 2008;43:281–287.
29. Cramer RD 3rd, Patterson DE, Bunce JD. Recent advances in comparative molecular field analysis (CoMFA). *Prog Clin Biol Res*. 1989; 291: 161–165.
30. Clark M, Cramer RD 3rd, Jones DM, Patterson DE, Simeroth PE. Comparative molecular field analysis (CoMFA). 2. Toward its use with 3D-structural databases. *Tetrahedron Comput Method*. 1990;3:47–59.

### Drug Design, Development and Therapy

### Publish your work in this journal

Drug Design, Development and Therapy is an international, peer-reviewed open-access journal that spans the spectrum of drug design and development through to clinical applications. Clinical outcomes, patient safety, and programs for the development and effective, safe, and sustained use of medicines are a feature of the journal, which

Submit your manuscript here: <http://www.dovepress.com/drug-design-development-and-therapy-journal>

has also been accepted for indexing on PubMed Central. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Dovepress