Published in final edited form as:

J Chemom. 2017 December ; 31(12): . doi:10.1002/cem.2934.

Quinolone carboxylic acid derivatives as HIV-1 integrase inhibitors: Docking-based QSAR and topomer CoMFA analyses

Jianbo Tong¹, Pei Zhan^{1, ID}, Xiang Simon Wang², and Yingji Wu¹

¹College of Chemistry and Chemical Engineering, Shaanxi University of Science and Technology, Xi'an 710021, China

²Department of Pharmaceutical Sciences, College of Pharmacy, Howard University, Washington, District of Columbia 20059, USA

Abstract

Quinolone carboxylic acid derivatives as inhibitors of HIV-1 integrase were investigated as a potential class of drugs for the treatment of acquired immunodeficiency syndrome (AIDS). Hologram quantitative structure-activity relationships (QSAR) and translocation comparative molecular field vector analysis (topomer CoMFA) were applied to a series of 48 quinolone carboxylic acid derivatives. The most effective QSAR model was obtained using atoms and bonds as fragment distinctions: cross-validation $q^2 = 0.796$, standard error of prediction $SD_{CV} = 0.36$, the non-cross-validated $r^2 = 0.967$, non-cross validated standard error $SD = 0.17$, the correlation coefficient of external validation $Q_{ext}^2 = 0.955$, and the best hologram length $HL = 180$. topomer CoMFA models were built based on different fragment cutting models, with the most effective model of $q^2 = 0.775$, $SD_{CV} = 0.37$, $r^2 = 0.967$, $SD = 0.15$, $Q_{ext}^2 = 0.915$, and $F = 163.255$. These results show that the models generated from QSAR and topomer CoMFA were able to effectively predict the inhibitory potency of this class of compounds. The molecular docking method was also used to study the interactions of these drugs by docking the ligands into the HIV-1 integrase active site, which revealed the likely bioactive conformations. This study showed that there are extensive interactions between the quinolone carboxylic acid derivatives and THR80, VAL82, GLY27, ASP29, and ARG8 residues in the active site of HIV-1 integrase. These results provide useful insights for the design of potent new inhibitors of HIV-1 integrase.

Keywords

3D-QSAR; QSAR; molecular docking; quinolone carboxylic acid derivatives; topomer CoMFA

1 INTRODUCTION

Acquired immunodeficiency syndrome (AIDS) is a disease of the immune system, caused by the human immunodeficiency virus (HIV) in humans.¹ There are 2 species of HIV which

Correspondence Jianbo Tong, College of Chemistry and Chemical Engineering, Shaanxi University of Science and Technology, Xi'an 710021, China. jianbotong@aliyun.com.

ID ORCID

Pei Zhan <http://orcid.org/0000-0002-0446-1985>

infect humans: HIV-1 and HIV-2. HIV-1 is the more virulent form because it is easily transmitted.² In the early stage of infection, the envelope glycoprotein (gp120 and gp41) is mediated, and HIV binds to CD4 positive T lymphocytes through which the virus subsequently enters. Following the initial infection, 3 key enzymes play essential roles in the HIV-1 life cycle. (1) Reverse transcriptase (RT), an RNA-guidance DNA polymerase, which is used by the viral genomic RNA to generate a proviral DNA for incorporation into the host cell DNA. (2) Integrase (IN), by which the proviral DNA is inserted into the host cell genome. (3) HIV protease (PR), which is the enzyme necessary for the processing and packaging of new virulent viral particles that exit the host cell.^{3,4} Because of the drug resistance of mutation-prone HIV, more than 20 antiretroviral drugs that have been approved for treating HIV-infected patients are becoming ineffective.⁵

Recently, HIV-1IN has been identified as a potential target for antiretroviral therapy,^{6–9} because it plays an important role in viral replication and productive infection.¹⁰ HIV-1 IN can mediate the integration of HIV-1 DNA into host chromosome in 2 steps. The first step is 3'-end processing, in which 2 dinucleotide units are removed from 3'-end of the viral HIV-1 DNA, and the terminal 3'-OH is exposed. The second step is DNA strand transfer, in which the viral HIV-1 DNA is inserted into the host cell DNA. Thus, the HIV-1 IN is an attractive target for antiretroviral drug design.

Quinolone carboxylic acid derivatives have been reported as inhibitors of HIV-1 IN. In this study, 48 quinolone carboxylic acid derivatives were analyzed by hologram quantitative structure-activity relationships (HQSAR) and topomer CoMFA to reveal structural-activity factors. Molecular docking was also applied to study the mechanism of drug action. This work will be helpful in guiding the synthesis of new quinolone carboxylic acid derivatives.

2 COMPUTATIONAL METHODS

2.1 Software and hardware

The SYBYL-x 2.0 was used to construct a 3D-QSAR model. Chemical structures of 48 quinolone carboxylic acid derivatives were built using the Sketch Molecule of SYBYL-x 2.0. All molecules were optimized using the tripos force field,¹¹ and charges were calculated using the Gasteiger-Hückel. To obtain stable conformations, the energy gradient limit was set to 0.05 kJ/mol, and the maximum iterations were set to 2000.¹² AutoDock Tools were used in molecular docking, and Discovery Studio 3.5 Visualizer was used to analyze receptor-ligand interactions.

2.2 QSAR modeling data set and biological properties

The structures of 48 quinolone carboxylic acid are shown in Figure 1 and Table 1, which also include their bioactivities obtained from the literatures.¹³ The 48 compounds were grouped randomly into a training set of 36 compounds used for model generation, and a test set of 12 compounds used for model validation. The biological activity was characterized by the half maximal (50%) inhibitory concentration (IC_{50}). For QSAR analysis, the IC_{50} values were transformed into pIC_{50} values (ie, $-\log IC_{50}$).¹⁴

2.3 HQSAR studies

HQSAR is a QSAR method in which each molecule in the data set is generated from sub-structural fragments, utilizing the standard parameters of Sybyl-x 2.0. In our study, the atoms of these fragments were encoded with different colors in holograms to reflect their contribution to biological activity. Negative, intermediate, and positive contributions are represented by red, white, and green, respectively.¹⁵ Each fragment is defined by its distinctive property parameters: atomic number (A), bond type (B), atomic connections (C), hydrogens (H), chirality (Ch), and participation as a donor or acceptor (D). In HQSAR methodology, the related value of these descriptors was analyzed by partial least square (PLS).

2.4 topomer CoMFA studies

topomer CoMFA is the combination of CoMFA and topomer technologies that overcomes the alignment problem of CoMFA. A topomer is a molecular fragment, which was prescribed by both its conformation and its position,¹⁶ and indicates any structure that contains at least 1 open valence. topomer CoMFA is the second CoMFA, which is a rapid fragment-based 3D-QSAR method to predict the biological activity of compounds, and the bioactivity values were identified with automated rules.¹⁷ topomer CoMFA used the 3D-QSAR technique to relate the biological activity of a series of molecules with their steric and electrostatic fields. The steric field and electrostatic field were calculated using the standard Tripos force field method and sp³ carbon probe atom with a +1 charge.¹⁸ topomer CoMFA model was built using PLS regression, and steric and electrostatic descriptors were used as independent variables; pIC₅₀ value was used as a dependent variable.¹⁹ The optimum number of components (C) was determined by leave-one-out cross-validated squared correlation coefficient (namely statistical q^2 parameter),²⁰ the validity of a model depends on non-cross-validated correlation coefficient (namely statistical r^2 parameter), and the robustness was evaluated by external validation (a test set, not a training set).²¹

The SYBYL-x 2.0 software was applied for topomer CoMFA studies. In the topomer CoMFA model, the molecular structure is split into 2 parts to generate 3D conformation. In the training set, a high bioactivity compound was selected as the template to generate 2 fragments, and the rest of the molecules were divided automatically. The fragments were adjusted by experience rules to produce a 3D-QSAR model.²²

2.5 Docking studies

In this study, the crystal structure of HIV-1 integrase (PDB ID: 5cok) was retrieved from the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank, and 5cok was prepared by adding hydrogen and charges, treating the terminal residues and extracting the ligand.^{23,24} Thus, the biomacromolecule 3D molecular model of HIV-1 integrase was generated. The ligand was prepared in accordance with the method used to training molecules, and then added Gasteiger-Hückel charge. Energy minimization was performed using a Tripos Force Field by the Powell method. The docking area was defined by analyzing the interaction of ligand and the active residue. To test this, 3D structure of the ligands (compounds: 42, 25, 2) was flexibly docked into the active pocket of 5cok.

3 RESULTS AND DISCUSSION

3.1 HQSAR results and analysis

The HQSAR model can be optimized by varying parameters, which include holographic length, fragment size, and fragment distinctive parameters. In general, 6 hologram lengths (97, 151, 199, 257, 307, and 353 bins) and fragment size (4–7) were defaulted parameters of HQSAR analysis. The best model was determined as the one with the lowest standard error. The HQSAR models of the 36 compounds in the training set were built using different fragment distinctives. The results of the different models are listed in Table 2 which show that the most effective combination of fragment distinctive was A/B/H/Ch/DA (shown in bold in Table 2). For this HQSAR model, the $PCs = 5$; $q^2 = 0.794$; $SD_{CV} = 0.363$; $r^2 = 0.954$; $SD = 0.117$, and $HL = 199$.

To improve the statistical power, more tests were conducted. The results of these additional tests show (see Table 3) that the most effective model (shown in bold font in Table 3) was the one with fragment size of 6–7 and A/B/H/Ch/DA, which gave $PCs = 5$; $q^2 = 0.796$; $SD_{CV} = 0.360$; $r^2 = 0.967$; $SD = 0.165$, and $HL = 180$. The values predicted by the model for the 12 compounds in the test set are shown in Table 5, with an external validation (Q_{ext}^2) value of 0.955.

The linear correlation regression between experimental pIC_{50} and predicted pIC_{50} for 2 sets is shown in Figure 2A, which proves that the model has both favorable estimation stability and good predictive ability.

3.2 topomer CoMFA results and analysis

The cutting mode is the key element in the topomer CoMFA model. In order to obtain a good predictive model, many cutting modes were applied. Forty-eight compounds were randomly divided into 36 training sets and 12 test sets in models. A highly active compound (compound 42) was selected as the template. The results of topomer CoMFA models that had 3 fragmentation patterns for R groups are listed in Table 4. The equations for the various parameters are as follows (also shown in Table 3):

$$F = \frac{\frac{1}{m} \sum_{i=1}^n (\bar{Y} - \bar{Y}_{training})}{\frac{1}{n-m-1} \sum_{i=1}^n (Y_{obs} - \bar{Y})^2} \quad (1)$$

$$Q_{ext}^2 = 1 - \frac{\sum_{i=1}^{test} (Y_{obs} - Y_{test})^2}{\sum_{i=1}^{test} (Y_{obs} - \bar{Y}_{training})^2} \quad (2)$$

$$SD = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (Y_{obs} - \bar{Y})^2} \quad (3)$$

where, $\bar{Y}_{training}$ is the mean value in the training data set, Y_{obs} is the observed value, and Y_{pred} is the predicted value. $\bar{Y}$ is the average of the experimental values, and n is number of compounds. A good model should have large values for q^2 , r^2 , Q_{ext}^2 , and F , and the value for SD should be small.

In our case, the data in Table 4 show that the topomer CoMFA 1 model is the most suitable, demonstrated by $PCs = 5$; $q^2 = 0.775$; $SD_{CV} = 0.37$; $r^2 = 0.967$; $SD = 0.15$; and $F = 163.255$. These results prove that the topomer CoMFA method is an accurate and reliable 3D-QSAR method. When the 12 values predicted by the model are applied to the test set compounds, a Q_{ext}^2 value of 0.915 is yielded, which indicates that the model has a good predictive ability.

The pIC_{50} of all molecules of the topomer CoMFA model are shown in Table 5. When the predicted values are plotted against the experimentally determined value, a linear regression is observed in Figure 2B, which proves that the model has both favorable estimation stability and good predictive ability.

3.3 HQSAR contribution maps and topomer CoMFA contour maps

The HQSAR contribution map of the best model is shown in Figure 3. The contribution to activity for each atom of a molecule is depicted by a coloring code: negative, neutral, and positive contributions are represented as red, white, and green, respectively; the unfavorable contribution was increased from orange to red, and the favorable contribution was increased from yellow to green.

Compound 42 was used as the template in the model. Figure 3 shows that the color changed from yellow to green in the region around the N of the hydroquinoline ring (ie, the N of the hydroquinoline ring, as well as the C connected to the N of the hydroquinoline ring and the phenyl group), which indicates positive contributions to biological activity. These groups should be retained when synthesizing compounds that might have even better biological activity. The atoms of other substituents in the compound are presented in white, indicating that they make a neutral contribution to the biological activity and might be replaced by substituents that could generate more potent inhibitory effects.

In the topomer CoMFA contour maps (Figure 4), Figure 4A,B are steric field contour maps of R1 and R2 groups. Unfavorable areas (shown in yellow) indicate that bulky substituents decrease biological activity, and favorable areas (shown in green) indicate that bulky substituents increase biological activity. Figure 4C,D show the electrostatic field contour map of R1 and R2 groups, in which negative and positive charges favoring areas are represented by red and blue contours, respectively.

Again, compound 42 was selected as the template in the model. In Figure 4A,B, steric field contour maps show a large green area located at the N-position (R_6) of the hydroquinoline

ring, R_2 , and nearby R_9 substituent. The biological activity should be improved when large substituents are located at those positions. Comparing compounds 2 ($R_2 = H$; $pIC_{50} = 5.79$) and 3 ($R_2 = Cl$; $pIC_{50} = 6.10$), and 4 ($R_2 = H$; $pIC_{50} = 6.39$) and 11 ($R_2 = Cl$; $pIC_{50} = 7.16$), one can conclude that the bulky substituents of R_2 (ie, Cl) do increase the inhibition. Similarly, comparing compounds 13 ($R_6 = Me$; $pIC_{50} = 7.08$), 14 ($R_6 = Et$; $pIC_{50} = 7.51$), 15 ($R_6 = Pr$; $pIC_{50} = 7.26$), and 16 ($R_6 = i\text{-}Pr$; $pIC_{50} = 7.59$) proves that bulkier groups at R_6 also increase inhibition.

In the unfavorable (yellow) field located at the R_1 and R_9 positions, small substituents should increase the biological activity. This effect is observed in compound 5 ($R_1 = F$; $pIC_{50} = 6.30$) which is more active than compound 8 ($R_1 = CF_3$; $pIC_{50} = 6.14$). The activity of compound 45 ($R_9 = t\text{-}Bu$; $pIC_{50} = 8.24$) is also higher than that of compound 46 ($R_9 = \text{cyclohexyl}$; $pIC_{50} = 8.17$). However, it should be noted that the series of compounds 36–38 show that bulky groups at R_9 can increase inhibitory activity, as indicated by the green contour near the R_9 position.

In Figure 4C,D, electrostatic field contour maps show several large blue areas at R_4 and R_8 substituents, where positive charges increase the inhibitory. The red contours appear at R_2 substituents denoting those areas where a negative charge (ie, ketone group of the hydroquinoline ring, the hydroxyl group or and carboxyl group) can increase inhibition. This property can explain the better inhibitory activity of compound 3 ($R_2 = Cl$; $pIC_{50} = 6.10$) than that of compound 2 ($R_2 = H$; $pIC_{50} = 5.79$).

HQSAR contribution maps (Figure 3) show that the color changed from yellow to green at the N-position of the hydroquinoline ring. Furthermore, the topomer CoMFA contour maps (Figure 4) show a large green area located at the N-position of the hydroquinoline ring, indicating that the N-position of the hydroquinoline ring is important to inhibitory activity. In short, these 2 methods generate consistent results.

3.4 Validation of the docking method

To validate the docking accuracy, the crystal structure of protein (5cok) with the cognate ligand was re-docked. As the reference ligand, the cognate ligand was taken out of its protein-ligand complex (5cok) and re-docked into its binding site.^{25,26} As shown in Figure 5, the re-docked ligands (3 conformations) and the reference ligand can be almost completely superimposed (RMSD = 1.9 Å, 1.4 Å, and 2.1 Å), and their rotational tendency is similar. This proves that the docking method is rational and reliable.

3.5 Docking results and analysis

The docking conformations show the manner in which the ligand binds to the active residues in the predefined hydrophobic binding pocket. High inhibitory compound 42, low inhibitory compound 2, and neutral inhibitory compounds 25 were chosen to investigate the binding mode between the receptor and the inhibitors. The results are shown in Figures 6–8. In Figure 6, the ligand (compound 42) is closest to these residues: THR80, VAL82, GLY27, ASP29, ARG8, ASP30, VAL32, ILE47, and GLY48. Further, the key amino acids GLY27 and ASP29 in chain A and ARG8, THR80, and VAL82 in chain B engage in hydrogen bonding at the binding site. The distances of hydrogen bond are 2.0 Å (ARG8/B-H...O-),

2.3 Å (ASP29/A-N-H...H-O-), 1.9 Å (GLY27/A-O...H-O-), 2.2 Å (Asp30/A-O...H-N-), 2.3 Å (Gly48/A-HN...O-), 2.5 Å (Asp25/B-O...H-N-), 2.6 Å (Asp25/B-NH...O-), and 2.0 Å (Asp29/B-N-H...O-).

Figure 7 shows that compound 2 was most closely associated with these residues: ILE84, THR80, VAL82, PRO81, ILE50, ARG8, GLY49, and GLY48. The key amino acids GLY49 and ILE50 are hydrogen bonded at the binding site to the ligand (compound 2) with the following hydrogen bond distances: 2.1 Å (GLY49/-N-H...O-) and 2.0 Å (ILE50/-N-H...O-).

Similarly, Figure 8 shows that compound 25 was closest to these residues: ASP30, GLY48, VAL32, ASP29, ARG8, ALA28, ILE84, GLY49, PRO81, VAL82, GLY27, ILE50, and GLY51. The amino acids ASP30, ASP29, GLY49, and ILE50 participate in hydrogen bonding with the ligand (compound 25), and the distances of hydrogen bond are 2.1 Å (ASP30/-N-H...-H-O-), 2.5 Å (ASP29/-N-H...H-O-), 2.5 Å (ASP29/-N-H...-O), 1.7 Å (GLY49/-N-H...-H-O-), and 1.7 Å (ILE50/-N-H...H-O-).

These interactions can enhance the bonding strength of ligand and receptor, thereby inhibiting the action of the protein. Therefore, the ketone group of the hydroquinoline ring, the hydroxyl group, and the carboxyl group increase inhibitory effect of the compounds. The molecular docking results are in agreement with the topomer CoMFA and HQSAR analyses.

4 CONCLUSIONS

A series of 48 quionolone carboxylic acid derivatives were selected and randomly assigned to a training set of 36 compounds and a test set of 12 compounds. QSAR models were built using HQSAR and topomer CoMFA methods. The 2 resulting models displayed a statistically strong ability to predict inhibitory activity. The accuracy of the molecular docking method was verified, and the results showed that the docking method can predict the combination mode of ligand and 5ock protein. The QSAR simulation method showed that the substituent, such as the ketene group of the hydroquinoline ring, the hydroxyl group, and the carboxyl group, increases the potency of compounds. The results of the QSAR model and docking analyses provide valuable insights to design potent new inhibitors for fighting HIV.

Acknowledgments

This work was supported by the National Natural Science Funds of China (21475081), the Natural Science Foundation of Shaanxi Province of China (2015JM2057), and the Graduate Innovation Fund of Shaanxi University of Science and Technology. We are thankful to Dr Donald G. Barnes for carefully reviewing this manuscript.

Funding information

Natural Science Foundation of Shaanxi Province of China, Grant/Award Number: 2015JM2057; National Natural Science Funds of China, Grant/Award Number: 21475081; Graduate Innovation Fund of Shaanxi University of Science and Technology

References

1. Gonzalez SF, Martin GJ. The neuropathogenesis of AIDS. *Nat Rev Immunol*. 2005; 5:69–81. [PubMed: 15630430]
2. Reeves JD, Doms RW. Human immunodeficiency virus type 2. *J Gen Virol*. 2002; 83:1253–1265. [PubMed: 12029140]
3. Johnston MI, Hoth DF. Present status and future prospects for HIV therapies. *Science*. 1993; 260:1286–1293. [PubMed: 7684163]
4. Craigie R. HIV integrase, a brief overview from chemistry to therapeutics. *J Bio Chem*. 2001; 276:23213–23216. [PubMed: 11346660]
5. Thompson MA, Aberg JA, Hoy JF, et al. Antiretroviral treatment of adult HIV infection: 2012 recommendations of the international antiviral society-USA panel. *J Am Med Assoc*. 2012; 308:387–402.
6. Grobler JA, Stillmock K, Binghua H, et al. Diketo acid inhibitor mechanism and HIV-1 integrase: implications for metal binding in the active site of phosphotransferase enzymes. *Proc Natl Acad Sci USA*. 2002; 99:6661–6666. [PubMed: 11997448]
7. Barreca ML, Ferro S, Rao A, et al. Pharmacophore-based design of HIV-1 integrase strand-transfer inhibitors. *J Med Chem*. 2005; 48:7084–7088. [PubMed: 16250669]
8. Zhuang L, Wai JS, Embrey MW, et al. Design and synthesis of 8-hydroxy-[1,6] naphthyridines as novel inhibitors of HIV-1 integrase in vitro and in infected cells. *J Med Chem*. 2003; 46:453–456. [PubMed: 12570367]
9. Angelo JD, Mouscadet JF, Desmaele D, Zouhiri F, Leh H. HIV-1 integrase: the next target for AIDS therapy. *PatholBiol*. 2001; 49:237–246.
10. La Femina RL, Schnider CL, Robbins HL, et al. Requirement of active human immunodeficiency virus type 1 integrase enzyme for productive infection of human T-lymphoid cells. *J Virol*. 1993; 66:7414–7419.
11. Swalina CW, Zauhar RJ, Degrazia MJ, Moyna G. Derivation of ¹³C chemical shift surfaces for the anomeric carbons of oligosaccharides and glycopeptides using ab initio methodology. *J Biomol Nmr*. 2001; 21:49–61. [PubMed: 11693568]
12. Wu L, Wang Y, Liu Y, et al. A central role for TRPS1 in the control of cell cycle and cancer development. *Oncotarget*. 2014; 5:7677–7690. [PubMed: 25277197]
13. Sato M, Kawakami H, Motomura T, et al. Quinolone carboxylic acids as a novel monoketo acid class of human immunodeficiency virus type 1 integrase inhibitors. *J Med Chem*. 2009; 52:4869–4882. [PubMed: 19719237]
14. Gangjee A, Adair O, Queener SF. *Pneumocystis carinii* and *Toxoplasma gondii* dihydrofolate reductase inhibitors and antitumor agents: synthesis and biological activities of 2, 4-diamino-5-methyl-6-[(monosubstituted anilino) methyl] pyrido [2, 3-d] pyrimidines. *J Med Chem*. 1999; 42:2447–2455. [PubMed: 10395486]
15. Winker D, Burden F. Holographic QSAR of benzodiazepines. *QSAR & Combin Sci*. 1998; 7:224–231.
16. Le BM, Polanski J, Zouhiri F, et al. Use of the Kohonen neural network for rapid screening of ex vivo anti-HIV activity of styrylquinolines. *J Med Chem*. 2002; 45:4647–4654. [PubMed: 12361391]
17. Gangjee A, Adair OO, Pagley M, Queener SF. Corrections to N9-substituted 2, 4-diaminoquinazolines: synthesis and biological evaluation of lipophilic inhibitors of *Pneumocystis carinii* and *Toxoplasma gondii* dihydrofolate reductase. *J Med Chem*. 2008; 51:6195–6200. [PubMed: 18771252]
18. Ferreira MMC. Multivariate QSAR. *J Braz Chem Soc*. 2002; 13:742–753.
19. Bush BL, Nachbar RB. Sample-distance partial least squares: PLS optimized for many variables, with application to CoMFA. *J Comput Aided Mol Des*. 1993; 7:587–619. [PubMed: 8294948]
20. Hulland J. Use of partial least squares (PLS) in strategic management research: a review of four recent studies. *SMJ*. 1999; 20:195–204.

21. Chirico N, Gramatica P. Real external predictivity of QSAR models. Part 2. New intercomparable thresholds for different validation criteria and the need for scatter plot inspection. *J Chem Inf Model.* 2012; 52:2044–2058. [PubMed: 22721530]
22. Liu YL, Li YT, Shi ZB, et al. Using topomer CoMFA and Surflex-dock to analyze 3D-QSAR of GSK-3inhibitors and their mechanism of action. *Sci China Chem.* 2013; 43:198–208.
23. Yu SL, Yuan JT, Shi JH, et al. HQSAR and topomer CoMFA for predicting melanocortin-4 receptor binding affinities of trans-4-(4-chlorophenyl) pyrrolidine-3-carboxamides. *Chemometr Intell Lab.* 2015; 146:34–41.
24. Liu HC, Tang SZ, Lu S, Ran T, Wang J. Studies on [5,6]-fused bicyclic scaffolds derivatives as potent dual B-RafV600E/KDR inhibitors using docking and 3D-QSAR approaches. *Int J Mol Sci.* 2015; 31:24451–24474.
25. Shen MY, Yu HD, Li Y, et al. Discovery and optimization of triazine derivatives as ROCK1 inhibitors: molecular docking, molecular dynamics simulations and free energy calculations. *J Mol Bio.* 2013; 9:361–374.
26. Shen MY, Zhou SY, Li YY, Pan PC, Zhang LL, Hou T. Discovery of rho-kinase inhibitors by docking-based virtual screening. *J Mol Bio.* 2013; 9:1511–1521.

FIGURE 1.
The structure of quionolone carboxylic acid derivatives

FIGURE 2.

Plots of the experimental activity value and predictive activity value for the 48 molecules from (A) the best HQSAR model and (B) the best topomer CoMFA model

FIGURE 3.
Contribution map for typical molecules obtained from the best HQSAR mode

FIGURE 4.
3D contour maps of the best topomer CoMFA model of compound 42

FIGURE 5.
Superimposition of CO-crystallized and redocked conformation by multiple docking in the active site

FIGURE 6.

A, Ribbon display of the 5cok structure in complex with the compound 42 (shown in rod like). B, The 3D model structure of the sample 42 binding model with the 5cok complex, the green dotted line represents the hydrogen bonding. C, Ligand interaction diagram of compound 42 and 5cok using Discovery Studio 3.5 Visualizer with the amino acid residues at the binding site depicted in circles. The purple circles show the amino acids that participate in hydrogen bonding or electrostatic or polar interactions, while the green circles show the amino acids that participate in van der Waals interactions. D, Aubergin represents H-bond donors, and green represents H-bond acceptors

FIGURE 7.

A, The 3D model structure of sample 2 binding model with 5cok complex. Green dotted line represents the hydrogen bonding. B, Aubergin represents H-bond donors, and green represents H-bond acceptors

FIGURE 8.

A, The 3D model structure of the sample 25 binding model with the 5cok complex, the green dotted line represents the hydrogen bonding. B, Aubergin represents H-bonds donor, and green represents H-bonds acceptor

TABLE 1

The structures and biological activities of quionolone carboxylic acid derivatives

No.	R ₁	R ₂	R ₃ /R ₇	R ₄ /R ₈	R ₅ /R ₆	R ₉	pIC ₅₀ Exp ^{1/2}
1	-	-	-	-	-	-	7.30
2	H	H	H	H	-	-	5.79
3 ^a	H	Cl	H	H	-	-	6.10
4	Cl	H	H	H	-	-	6.39
5	F	H	H	H	-	-	6.30
6 ^a	Me	H	H	H	-	-	5.97
7	OMe	H	H	H	-	-	5.93
8	CF ₃	H	H	H	-	-	6.14
9	Cl	H	H	Cl	-	-	6.43
10 ^a	H	Cl	H	Cl	-	-	6.60
11	Cl	Cl	H	H	-	-	7.16
12	H	H	-	-	Me	-	5.64
13	Cl	Cl	-	-	Me	-	7.08
14	Cl	Cl	-	-	Et	-	7.51
15 ^a	Cl	Cl	-	-	Pr	-	7.26
16	Cl	Cl	-	-	<i>i</i> -Pr	-	7.59
17	Cl	Cl	-	-	Bu	-	7.19
18	Cl	Cl	-	-	CH ₂ CO ₂ H	-	7.50
19 ^a	Cl	Cl	-	-	CH ₂ CH ₂ CO ₂ H	-	7.42
20	Cl	Cl	-	-	CH ₂ CONH ₂	-	7.46
21	Cl	Cl	-	-	CH ₂ CH ₂ CONH ₂	-	6.94
22	Cl	Cl	-	-	CH ₂ CH ₂ NH ₂	-	6.67
23 ^a	Cl	Cl	-	-	CH ₂ CH ₂ OH	-	7.68
24	Cl	Cl	-	-	CH ₂ CH ₂ CH ₂ OH	-	7.11
25	Cl	F	-	-	CH ₂ CH ₂ OH	-	7.36

No.	R ₁	R ₂	R ₃ /R ₇	R ₄ /R ₈	R ₅ /R ₆	R ₉	pIC ₅₀ Exp ¹²
26	F	Cl	-	-	CH ₂ CH ₂ OH	-	7.62
27 ^a	Cl	Cl	F	H	H	-	7.08
28	Cl	Cl	H	H	F	-	7.60
29	Cl	Cl	H	F	H	-	7.47
30	Cl	Cl	H	H	OMe	-	7.92
31 ^a	Cl	Cl	H	H	Cl	-	7.37
32	Cl	Cl	H	H	Me	-	7.39
33	Cl	Cl	H	H	CF ₃	-	6.17
34	Cl	Cl	H	H	CN	-	7.30
35 ^a	F	Cl	H	H	OMe	-	8.05
36	-	-	-	-	H	(S)-Me	7.83
37	-	-	-	-	H	(R)-Me	7.42
38	-	-	-	-	H	(S)-Et	8.05
39 ^a	-	-	-	-	H	(S)-Pr	8.09
40	-	-	-	-	H	(S)- <i>i</i> -Pr	8.09
41	-	-	-	-	H	(S)- <i>i</i> -Bu	8.22
42	-	-	-	-	H	(S)-cyclohexyl	8.25
43 ^a	-	-	-	-	H	(S)-Ph	8.01
44	-	-	-	-	OMe	(S)-Pr	8.24
45	-	-	-	-	OMe	(S)- <i>i</i> -Pr	8.14
46	-	-	-	-	OMe	(R)- <i>i</i> -Pr	7.84
47 ^a	-	-	-	-	OMe	<i>i</i> -Bu	8.24
48	-	-	-	-	OMe	Cyclohexyl	8.17

^aTest sets.

TABLE 2

Statistical results of the HQSAR model using different fragment distinctions

Fragment distinction	Fragment size	PCs	q^2	SD_{CV}	r^2	SD	HL
A	4-7	4	0.677	0.448	0.878	0.275	257
A/B	4-7	3	0.687	0.436	0.859	0.292	353
A/B/C	4-7	4	0.708	0.427	0.917	0.227	199
A/B/C/H	4-7	4	0.728	0.411	0.925	0.216	199
A/B/C/H/Ch	4-7	4	0.723	0.415	0.925	0.215	199
A/B/C/H/Ch/DA	4-7	3	0.691	0.450	0.972	0.135	353
A/C/H/Ch/DA	4-7	6	0.688	0.452	0.980	0.114	257
A/B/H/Ch/DA	4-7	5	0.794	0.363	0.954	0.117	199
A/B/C/Ch/DA	4-7	5	0.713	0.428	0.968	0.143	307
A/B/C/H/DA	4-7	6	0.723	0.426	0.969	0.143	97
A/B/H	4-7	3	0.696	0.430	0.864	0.287	353
A/B/C/Ch	4-7	4	0.708	0.427	0.917	0.227	199
A/C/Ch/DA	4-7	6	0.691	0.450	0.972	0.135	353

HL, best hologram length; *PCs*, number of principal components; q^2 , LOO cross-validated correlation coefficient; r^2 , non-cross-validated correlation coefficient; SD_{CV} , standard error of prediction; *SD*, non-cross validated standard error.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE 3

Statistical results of the HQSAR modeling using different fragment sizes

Fragment distinction	Fragment size	PCs	q^2	SD_{CV}	r^2	SD	HL
A/B/H/Ch/DA	2-5	5	0.674	0.456	0.957	0.166	257
A/B/H/Ch/DA	3-6	4	0.737	0.405	0.921	0.222	151
A/B/H/Ch/DA	4-7	5	0.794	0.363	0.954	0.171	199
A/B/H/Ch/DA	5-8	4	0.763	0.384	0.941	0.191	257
A/B/H/Ch/DA	6-9	4	0.731	0.409	0.937	0.199	257
A/B/H/Ch/DA	7-10	6	0.705	0.439	0.957	0.174	97
A/B/H/Ch/DA	8-11	6	0.758	0.398	0.948	0.189	151
A/B/H/Ch/DA	9-12	4	0.689	0.461	0.934	0.203	307
A/B/H/Ch/DA	3-4	1	0.704	0.415	0.756	0.377	351
A/B/H/Ch/DA	4-5	3	0.724	0.410	0.898	0.249	307
A/B/H/Ch/DA	5-6	4	0.744	0.359	0.949	0.179	199
A/B/H/Ch/DA	6-7	5	0.796	0.360	0.967	0.165	180
A/B/H/Ch/DA	7-8	4	0.763	0.384	0.941	0.192	257
A/B/H/Ch/DA	8-9	4	0.743	0.400	0.935	0.202	257
A/B/H/Ch/DA	9-10	9	0.705	0.439	0.953	0.174	97

TABLE 4

Results from 3 topomer CoMFA models

	QSAR model		
	topomer CoMFA 1	topomer CoMFA 2	topomer CoMFA 3
Cutting mode
<i>PCs</i>	5	3	4
q^2	0.775	0.728	0.700
SD_{CV}	0.37	0.43	0.50
r^2	0.967	0.894	0.890
SD	0.15	0.25	0.31
F	163.255	118.549	133.145
<i>Intercept</i>	6.29	6.59	6.12

TABLE 5Experimental and predicted pIC₅₀ (M) and residual values of all compounds

No.	pIC ₅₀ Exp	The best topomer CoMFA		The best HQSAR	
		pIC ₅₀ Pred	Residual	pIC ₅₀ Pred	Residual
1	7.30	7.32	-0.02	7.12	0.18
2	5.79	5.83	-0.04	6.15	-0.36
3 ^a	6.10	6.79	-0.69	6.30	-0.20
4	6.39	6.29	0.10	6.30	0.09
5	6.30	6.45	-0.15	6.23	0.07
6 ^a	5.97	5.96	0.01	6.11	-0.14
7	5.93	6.08	-0.15	6.13	-0.20
8	6.14	6.05	0.09	5.90	0.24
9	6.43	6.34	0.09	6.48	-0.05
10 ^a	6.60	6.41	0.19	6.34	0.26
11	7.16	6.97	0.19	6.97	0.19
12	5.64	5.60	0.04	6.06	-0.42
13	7.08	7.14	-0.06	6.88	0.20
14	7.51	7.39	0.12	7.22	0.29
15 ^a	7.26	7.28	-0.02	7.23	0.03
16	7.59	7.60	-0.01	7.51	0.08
17	7.19	7.17	0.02	7.23	-0.04
18	7.50	7.56	-0.06	7.26	0.24

No.	pIC ₅₀ Exp	The best topomer CoMFA		The best HQSAR	
		pIC ₅₀ Pred	Residual	pIC ₅₀ Pred	Residual
19 ^a	7.42	7.22	0.20	7.20	0.22
20	7.46	7.54	-0.08	7.54	-0.08
21	6.94	6.89	0.05	7.20	-0.26
22	6.67	7.24	-0.57	7.22	-0.55
23 ^a	7.68	7.47	0.21	8.11	-0.43
24	7.11	7.17	-0.06	7.24	-0.13
25	7.36	7.32	0.04	7.32	0.04
26	7.62	7.62	0.00	7.31	0.31
27 ^a	7.08	7.13	-0.05	7.51	-0.43
28	7.60	7.53	0.07	7.48	0.12
29	7.47	7.51	-0.04	7.43	0.04
30	7.92	7.69	0.23	7.80	0.12
31 ^a	7.37	7.38	-0.01	7.38	-0.01
32	7.39	7.48	-0.09	7.48	-0.09
33	6.17	6.12	0.05	6.12	0.05
34	7.30	7.31	-0.01	7.31	-0.01
35 ^a	8.05	8.02	0.03	7.93	0.12
36	7.83	7.83	0.00	8.03	-0.20
37	7.42	7.47	-0.05	7.31	0.11

No.	pIC ₅₀ Exp	The best topomer CoMFA		The best HQSAR	
		pIC ₅₀ Pred	Residual	pIC ₅₀ Pred	Residual
38	8.05	8.00	0.05	8.06	-0.01
39 ^a	8.09	7.83	0.26	8.13	-0.04
40	8.09	8.11	-0.02	7.97	0.12
41	8.22	8.17	0.05	8.20	0.02
42	8.25	7.98	0.27	7.92	0.33
43 ^a	8.01	8.08	-0.07	8.00	0.01
44	8.24	8.18	0.06	8.18	0.06
45	8.14	8.23	-0.09	8.18	-0.04
46	7.84	7.85	-0.01	7.83	0.01
47 ^a	8.24	8.45	-0.21	8.32	-0.08
48	8.17	8.31	-0.14	8.35	-0.18

^aTest sets.