

Quantitative structure activity relationship (QSAR) studies on nitazoxanide-based analogues against *Clostridium difficile* *In vitro*

Han Zhang, Xiwang Liu, Yajun Yang and Jianyong Li*

Key Laboratory of New Animal Drug Project, Gansu Province; Key Laboratory of Veterinary Pharmaceutical Development, Ministry of Agriculture; Lanzhou Institute of Husbandry and Pharmaceutical Sciences of CAAS, Lanzhou, China

Abstract: Quantitative structure activity relationship (QSAR) has been established between the various physiochemical parameters of a series of nitazoxanide-based analogues and its antibacterial activity against *Clostridium difficile*. Genetic function approximation (GFA) and comparative molecular field analysis (CoMFA) techniques were used to identify the descriptors that have influence on biological activity. The most influencing molecular descriptors identified in 2D-QSAR include spatial, topological, and electronic descriptors, while electrostatic and stereoscopic fields were the most influencing molecular descriptors identified in 3D-QSAR. Statistical qualities (r^2 , q^2) indicated the significance and predictability of the developed models. The study indicated that antibacterial activity of *Clostridium difficile* can be improved by increasing molecular connectivity index, local charge surface index, sharp index and decreasing molecular flexibility index.

Keywords: Nitazoxanide-based analogues; antibacterial activities; QSAR; *Clostridium difficile*

INTRODUCTION

Parasitic and bacterial infections were the major causes of morbidity and mortality in the world especially in developing countries. Diseases caused by intestinal parasites affected billions of people every year and there had been few drug innovations for treating these infections in the past three decades. Fortunately, nitazoxanide (NTZ) named as 2-acetylloxy-N-(5-nitro-2-thiazolyl) benzamide was developed as a promising compound to treat these diseases in the future. NTZ was first synthesized by Romark Laboratory in 1975 (Rossignol *et al.* 1975) and showed activities against *Taenia saginata* and *Hymenolepis nana* (Rossignol *et al.* 1984). Further studies reported that NTZ exhibited an unusual broad spectrum of activities against various parasites in human beings and animals (*Giardia lamblia*, *Cryptosporidium parvum*, *Entamoeba histolytica*, *Neospora caninum*, *Trichuris trichiura*, *Ascaris lumbricoides*, *Enterobius vermicularis*) (Rossignol *et al.* 1984; Rossignol *et al.* 2001; Adagu *et al.* 2002; Esposito *et al.* 2007; Elvia *et al.* 2003). In 1996, NTZ had been marketed in Latin America and then approved by US Food and Drug Administration (FDA) as an agent treating infections caused by parasites (Gilles *et al.* 2002; Anderson *et al.* 2007). NTZ also showed broad antimicrobial activities for anaerobic bacteria (*Clostridium difficile*, *Helicobacter pylori*) as well as notable anti-virus actions on influenza and hepatitis (Catherine *et al.* 2000; Guttner *et al.* 2003; Rossignol *et al.* 2009; Korba *et al.* 2008).

In contrast to metronidazole, NTZ has been shown non-mutagenic, which suggested that the two compounds have fundamental differences with the mode of action. In fact, recent studies revealed that NTZ inhibited pyruvate ferredoxin oxidoreductase (PFOR), a key enzyme of central intermediary metabolism in anaerobic organisms. NTZ appeared to interact with PFOR and the nitro group was not reduced. The mechanism was supposed that amide anion (NTZ⁻) may couple directly with thiamine pyrophosphate (TPP), which was a cofactor of PFOR (Hoffman *et al.* 2007). Recently, a number of nitazoxanide derivatives were synthesized and their activities were also investigated. The results demonstrated that the nitro group may be essential for the antimicrobial efficacy with a minimum inhibitory concentration of organisms (MIC₉₀) of 0.06-4.00mg·L⁻¹ (Sisson *et al.* 2002; Glenn *et al.* 2006). But notably, specific replacements of the nitro group were also shown significant activities against intracellular parasites (Esposito *et al.* 2007) while invalidity for intestinal parasite *Giardia lamblia*. On the other hand, modifications of the salicylate moiety may reduce the activity even abrogate it (Müller *et al.* 2006). This indicated that other enzymes beside PFOR might also be relevant.

In order to prevent resistant emergence of NTZ and develop more effective drugs, a quantitative structure activity relationship (QSAR) study on nitazoxanide-based analogues was a useful tool. QSAR methodology is used to predict the activity of novel molecules by mathematical equations, which deduce the relationship (s) between a chemical structure and its biological activity (Kamalakaran *et al.* 2009). QSAR models are pointers to design effective drugs.

*Corresponding author: e-mail: lijy1971@163.com

In this study, QSAR study has been investigated on 28 nitazoxanide-based analogues against *Clostridium difficile* *in vitro*. From the results obtained we established a functional predictive model, which can be further developed to aid in the design of novel derivatives with improved *in vitro* potency against *Clostridium difficile* and other anaerobic organisms including parasites.

MATERIALS AND METHODS

The antibacterial data of nitazoxanide-based analogues were used for QSAR study. These selected compounds should have enough variable grad in antibacterial activity and were divided into training set (23 compounds) and test set (5 compounds including 1a, 1, 8, 10, 17) as structural diversity with 80% rate. Moreover, the compounds in training set had enough quantity and included all structural information. The biological activity data were minimal inhibitory concentration (MIC) for *Clostridium difficile* *in vitro* and converted into pMIC by equation $\text{pMIC} = -\log \text{MIC}$. All compounds are shown in table 1 serially citing their structures and biological activities. All calculations were carried out with Discovery Studio 3.1 (DS3.1, Accelrys Software Inc., USA) QSAR software and default values were used in all parameters without specialized exception. The general structure of nitazoxanide-based analogues was followed as fig. 1.

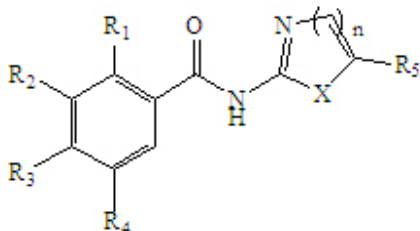


Fig. 1: The general structure of nitazoxanide-based analogues

2D-QSAR model

Initial conformation energy optimization

2D structures of all compounds were drawn with Chem Bio Draw 11.0 software and then imported into DS3.1 QSAR software. After pMICs were imported and molecular optimized parameters were set in Minimize Ligands, initial conformation energy optimization was carried out with second-generation force field MMFF.

Descriptors calculation

In the Calculate Molecular Properties, 98 Molecular descriptors for all compounds were selected. They included hydrophobic parameters, electronic parameters, space configuration parameters, the molecular character parameters, topological parameters and charge related parameters, such as, AlogP, Molecular Properties, Molecular Property Count, Surface Area and Volume, Topological Descriptors, Dipole, Jurs Descriptors,

Molecular Properties, Principal Moments of Inertia, Shadow Indices, Surface Area and Volume.

2D-QSAR model establishment

Using genetic function approximation (GFA) method, the relationships between structure and biological activity of target compounds (training set compound) were established and analyzed. The generated model as planned contained 4~5 item parameters with the initial population numbers for 100, the iterative 5000 times. Such parameters produced were the no cross validation correlation coefficient r^2 , the adjusted coefficient of determination r^2 (adj), prediction correlation coefficient r^2 (pred), RMS Residual (RMS Residual Error), etc.

2D-QSAR model validation

In order to test the model's repeatability and accuracy and judge its predictive ability, active predictions of sample molecules in training set and test set were carried out by the GFA mode established with training set. In the DS 3.1 QSAR software, Calculate Molecular Properties module was run, and at the same time, test set and prediction set file data, GFATempModel-1 model were selected.

3D-QSAR model

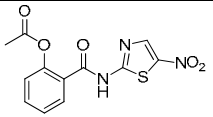
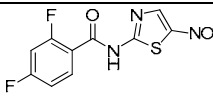
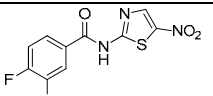
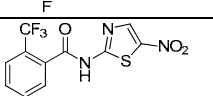
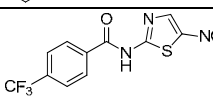
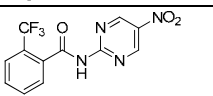
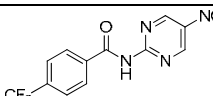
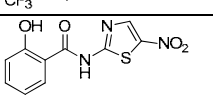
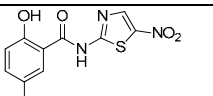
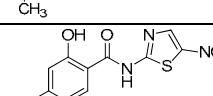
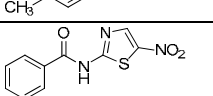
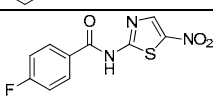
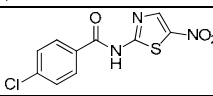
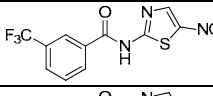
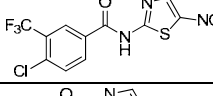
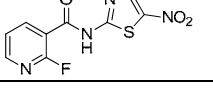
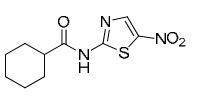
Molecular optimization and overlay

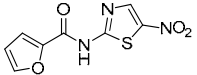
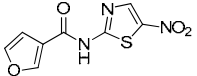
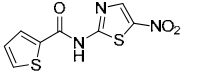
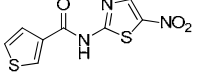
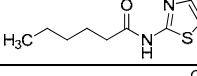
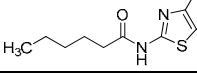
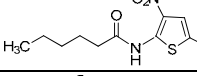
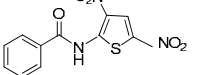
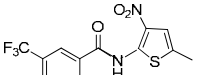
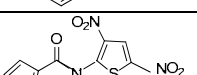
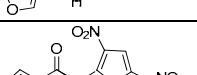
After energy optimization were carried out for three-dimensional structure of small molecular compound in DS with molecular dynamics program such as Ligands Minimize, lowest energy conformation of each molecule was obtained. Through the most commonly used overlay method based on Molecular field, compound 1e with best activity was selected as template molecule. Then Molecular Overlay was started and at the same time all molecular force field with the same orientation was ensured.

3D-QSAR model establishment and validation

In the overlayed molecule surrounding, molecular field space range was defined and rectangular space was used in the software. The defined space evenly divided in accordance with 1.5Å step length produced grid point. Such probe molecules as van der Waals carbon atom and H^+ were used to calculate the interaction of stereo energy and electrostatic energy between molecule and lattice point probe atom. After molecular field characteristics on lattice point were evaluated, molecular force field parameters (structural parameters) can be produced. The molecular fields of molecule at each grid point were calculated, and then the molecular field parameters as the independent variable and molecular activity value as the dependent variable. The model best principal component number n and cross validation correlation coefficient q^2 were determined by partial least squares (PLS) combined with interactive test. Then based on the best principal component, the relationship between structure and biological activity for goal compounds was established by PLS.

Table 1 Structures and activities (pMIC) of the nitazoxanide-based analogues

Compound no.	Structure	MIC ($\mu\text{g/mL}$)	pMIC	Ref
1a* (NTZ)		0.25	0.60206	Liu <i>et al.</i> 2014
1b		0.5	0.30103	Liu <i>et al.</i> 2014
1c		1	0	Liu <i>et al.</i> 2014
1d		0.5	0.30103	Liu <i>et al.</i> 2014
1e		0.125	0.90309	Liu <i>et al.</i> 2014
17d		4	-0.60206	Liu <i>et al.</i> 2014
17e		2	-0.30103	Liu <i>et al.</i> 2014
1* (TIZ)		0.5	0.30103	Glenn <i>et al.</i> 2006
2		0.5	0.30103	Glenn <i>et al.</i> 2006
3		0.25	0.60206	Glenn <i>et al.</i> 2006
4		1.50	-0.17609	Ballard <i>et al.</i> 2011
5		0.88	0.055517	Ballard <i>et al.</i> 2011
6		0.23	0.63827	Ballard <i>et al.</i> 2011
7		0.16	0.79588	Ballard <i>et al.</i> 2011
8*		0.49	0.30980	Ballard <i>et al.</i> 2011
9		0.75	0.12494	Ballard <i>et al.</i> 2011
10*		1.51	-0.17898	Ballard <i>et al.</i> 2011

Compound no.	Structure	MIC ($\mu\text{g/mL}$)	pMIC	Ref
11		0.57	0.24413	Ballard <i>et al.</i> 2011
12		0.41	0.38722	Ballard <i>et al.</i> 2011
13		0.38	0.42022	Ballard <i>et al.</i> 2011
14		0.74	0.13077	Ballard <i>et al.</i> 2011
15		1.99	-0.29885	Ballard <i>et al.</i> 2011
16		8	-0.90309	Ballard <i>et al.</i> 2010
17*		7.99	-0.9026	Ballard <i>et al.</i> 2010
18		8.01	-0.9036	Ballard <i>et al.</i> 2010
19		1.52	-0.18184	Ballard <i>et al.</i> 2010
20		0.99	0.004365	Ballard <i>et al.</i> 2010
21		5.98	-0.7767	Ballard <i>et al.</i> 2010

*: Test set compounds.

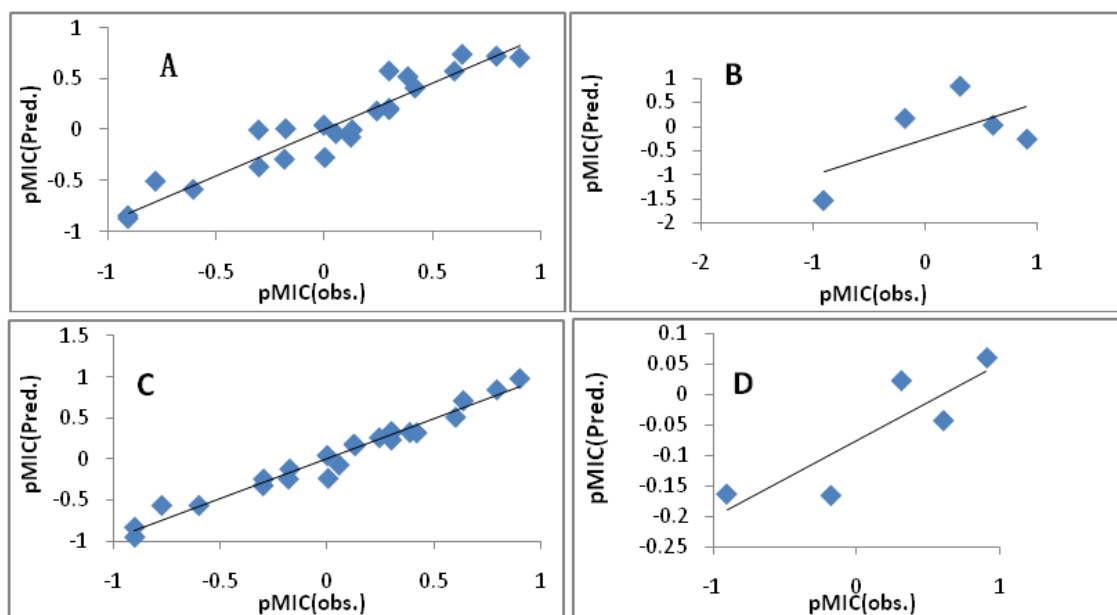


Fig. 2: Plots of predicted versus observed values for 2D-QSAR and 3D-QSAR model. A was 2D-QSAR model training set, B was 2D-QSAR model test set, C was 3D-QSAR model training set and D was 3D-QSAR model test set.

Table 3: The active data and molecular descriptors of Nitazoxanide derivatives

Compounds	pMIC	X ₁	X ₂	X ₃	X ₄
1a*(NTZ)	0.60206	0.59777	-1.41991	3.75111	4.47758
1b	0.30103	0.57306	-1.38743	2.9496	3.58002
1c	0	0.57578	-1.51676	2.9496	3.58002
1d	0.30103	0.66182	-1.7325	3.4101	4.00186
1e	0.90309	0.68415	-1.69419	3.60966	4.00186
17d	-0.60206	0.49131	-1.79074	3.6214	4.12649
17e	-0.30103	0.51364	-1.71325	3.83057	4.12649
1*(TIZ)	0.90309	0.52006	-1.5119	2.7621	3.4179
2	0.30103	0.68672	-1.4884	2.99282	3.65045
3	0.60206	0.68672	-1.4895	2.99282	3.65045
4	-0.17609	0.46667	-1.33713	2.71121	3.21631
5	0.055517	0.52967	-1.6469	2.9434	3.3967
6	0.63827	0.65565	-1.39687	3.10779	3.65628
7	0.79588	0.68415	-1.68329	3.60966	4.00186
8*	0.3098	0.83664	-1.78409	3.79652	4.44497
9	0.12494	0.49776	-1.34423	2.71932	3.34753
10*	-0.17898	0.53557	-1.14007	3.07112	3.79993
11	0.24413	0.44226	-1.24819	2.31785	2.74979
12	0.38722	0.46667	-1.09226	2.31785	2.74979
13	0.42022	0.56012	-1.34781	2.4976	3.03742
14	0.13077	0.46667	-1.33018	2.4976	3.03742
15	-0.29885	0.40448	-1.11752	4.37829	5.00198
16	-0.90309	0.51007	-1.23746	4.37088	5.46379
17*	-0.9026	0.52246	-1.5974	4.46661	5.6714
18	-0.9036	0.58465	-1.9168	3.10843	3.9363
19	-0.18184	0.80213	-2.35659	3.99636	4.72686
20	-0.7767	0.6781	-1.94689	2.87553	3.75762
21	0.004365	0.58465	-1.62298	2.70279	3.46217

*: Test set compounds.

In the Create 3D QSAR Model modules, setting step length was 1.5Å and high correlation value described was 0.9. After the establishment of 3D-QSAR model, at the same time test set data were input to validate model. The steric field and electrostatic field to activity contribution were intuitively reflected with three dimensional force field coefficient equipotential diagrams and other molecular modified information can be got.

RESULTS

2D-QSAR model

The structure and biological activity relationship model of target compounds was constructed by GFA in DS software. The optimal 2D-QSAR model with relevant activity data was finally determined after each parameter was investigated, which was expressed as regression equation (1) through GFA in training set molecule.

$pMIC = 0.75552 + 4.9471 * X_1 + 1.397 * X_2 + 1.7507 * X_3 - 1.8442 * X_4$ (model 1)
 $n = 23$, $r^2 = 0.9052$, $r^2(\text{adj}) = 0.8842$, $r^2(\text{pred}) = 0.8402$, RMS Residual Error = 0.1733, Friedman L.O.F. = 0.08764, S.O.R p-value = $1.65 * 10^{-0.9}$.

X₁: molecular connectivity index, X₂: molecular local charge surface parameter, X₃: molecular shape index, X₄: molecular flexibility index.

3D-QSAR model

CoMFA method was chose for 3D-QSAR analysis in this study. Under unknowing receptor structure, by using two kinds of probe molecule (van der Waals C atom and H⁺) to calculate electrostatic field and steric field and energy as a descriptor, CoMFA method studied the relationships between activity (y) and molecular field (x) and at the same time gave some suggestions for the existing molecular modification.

The results revealed that the non-cross validation correlation coefficient r^2 for the optimal model of 3D-QSAR was 0.967. The r^2 was greater than 0.8, which showed that the model had good correlation. In addition, its adjustment decision coefficient $r^2(\text{adj})$ was 0.964, which indicated that the model had a strong predictability. RMS residual error was smaller as 0.09005, indicating the model with good repeatability (table 4). The observed and

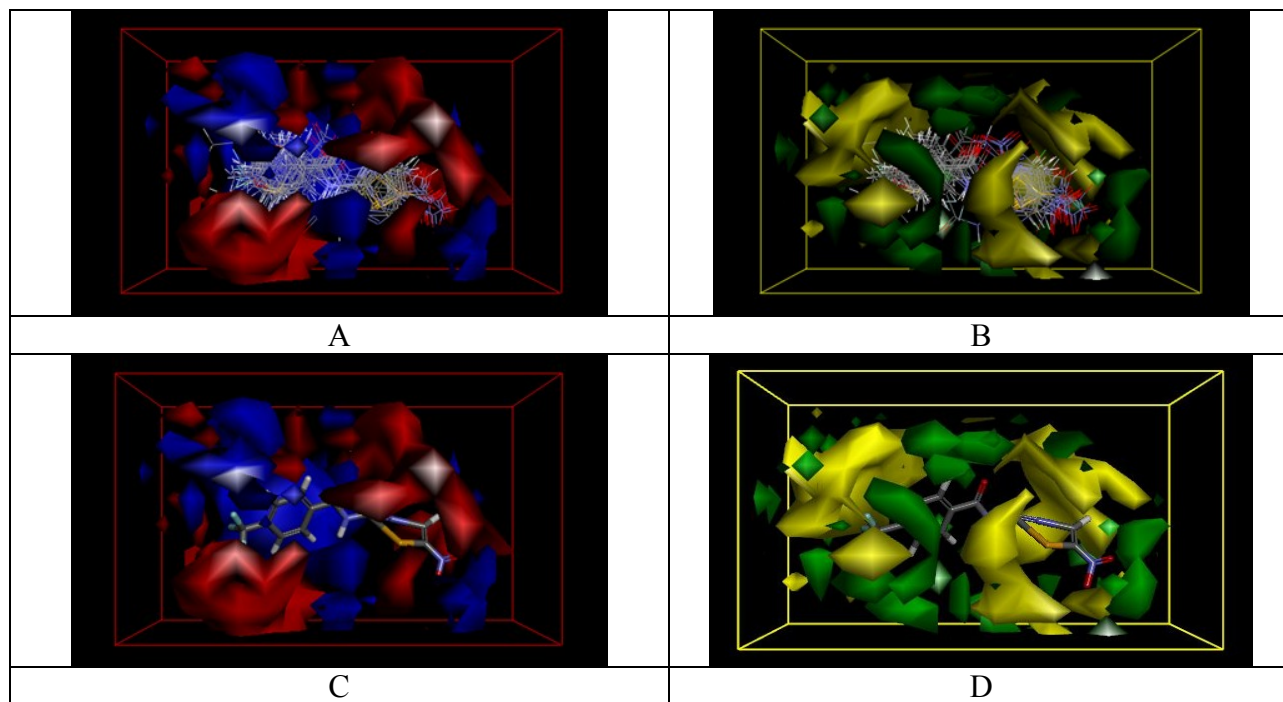


Fig. 3: Steric and electrostatic equipotential diagrams. A and C for 3D electrostatic force field coefficient equipotential diagram, B and D for 3D steric force field coefficient equipotential diagrams, A and B for 3D-QSAR model equipotential diagrams of the electrostatic field and steric field, C and D for 3D equipotential diagram for the highest active compounds **1e**.

predicted MICs of the compounds against difficult *Clostridial* were followed as table 2.

3D-QSAR model validation

By the best principal component number to get CoMFA model and with CoMFA model to forecast the molecular activity in the sample concentration for testing the model's repeatability and accuracy, the cross validation correlation coefficient for 0.744 ($q^2 > 0.5$) showed that its predictability was good. The correlation curve fig. for the predicted and observed value of the compound in training set and testing set was followed as fig. 3(C, D). From fig. 3 (C, D), the linear relationship between predicted and observed value in model training set was good, and all data were concentrated near the trend line, which showed that this model had high repeatability and internal forecast ability. Therefore, 3D-QSAR model can be used to predict the compound activity with similar structure against *C. difficile*.

Equipotential diagram analysis of 3D-QSAR model

The 3D force field coefficient equipotential diagrams by CoMFA method implied some important information for molecular modification. The 3D equipotential diagrams of the electrostatic field and steric field were as shown in fig. 3. In the diagram A and C for 3D electrostatic force field coefficient equipotential diagram, red represented that the stronger the electro negativity of the area substituent, the

better the compound activity. However, blue represented that the weaker the electro negativity of the area substituent, the better the compound activity. In the diagram B and D for 3D steric force field coefficient equipotential diagrams, yellow represented that the less the volume of the area substituent, the better the compound activity. However, green represented that the bigger the volume of the area substituent, the better the compound activity. The 3D equipotential diagram for the highest active compounds **1e** was followed as fig. 3 (C and D).

DISCUSSION

In constructing a statistically significant predictive ability of the model, the model validation was indispensable. Non cross validation correlation coefficient r^2 was an important basis for testing the quality of QSAR model and the most commonly used internal inspection index, which can explain the deviation between molecular predicted value and observed value of training set. When r^2 was greater than 0.8 and the more close to 1, it was generally thought that the model has good predictability (Hansch and Verma, 2009). In the QSAR model, cross validation correlation coefficient q^2 was the statistical index of model predictability and usually used to judge the merits of the model, which meant that the higher the q^2 , the stronger the predictability. When $q^2 > 0.3$, the

established model in the 5% significant level has statistical significance, and while $q^2 > 0.5$, the established model has certain fitting ability and very remarkable statistical significance (Xu *et al*, 2004).

In constructing the 2D-QSAR model, 23 training set compounds and 98 Molecular descriptors such as hydrophobic parameters, electronic parameters, space configuration parameters, molecular character parameter, topological parameters and charge related parameters were used. Because of 23 compounds in training set, the QSAR model with 4~5 item parameters was constructed. After the initial population numbers was designed for 100 and the iterative 5000 times, 10 GFA models were produced at the end of calculation. Compared with other models, r^2 , r^2 (adj), r^2 (pred), RMS Residual Error, Friedman L.O.F. and S.O.R p-value for the selected optimal 2D-QSAR model has obvious improvement. From the formula (1), r^2 was 0.9052 (>0.8), which indicated that this model could explain 90% of the variable. Adjustment coefficient r^2 (adj) and prediction correlation coefficient r^2 (pred) were more than 0.8, which indicated that the model has good internal forecast ability.

In order to validate the actual predictability of the 2D-QSAR model, 5 compounds in test set were used to forecast and evaluate the model. From predicted value, the model can predict accurately their activities. Moreover, the cross validation correlation coefficient q^2 for 0.371 (>0.3) indicated that the model credibility was more than 95%. Observed value of 28 compounds' activity and predicted value with GFA model were as shown in table 2, while linear relationship of observed value and predicted value in training set and test set were as shown in fig. 2 (A, B). The linear relationships between observed value and predicted value in training set and test set were fitted well and their slopes were 0.905 and 0.751, which showed that GFA model selected had good internal predictability.

In constructing the 2D-QSAR model, over fitting phenomenon quite often happened. This usually resulted from too many parameters introduced, so only part parameters from the physical and chemical parameters can be selected. It is proposed that the ratio between the sample number and the number of parameters selected should be more than or equal to 5 in the less sample, while their ratio should be more than 2^n (n =parameter number taken) in the more sample, only so the appearance of over fitting phenomenon can be effectively avoided (Sharma *et al*, 2009). The model met the above principle, for example, the training set of sample number was for 23 and the model contained four parameters. In equation 1, all the parameters belonged to topological parameters and electrical parameters. Topological parameters reflected the characteristics of the molecular topology structure, generally including graphical description of molecular

structure, the matrixing and quantification of graphics structure.

X_1 was molecular connectivity index, according to the atomic arrangement or connection order of each frame in the molecules to describe the nature of the molecular structure and with the reciprocal of square root for dot grade product to represent. It was the most representative topology parameter and reflected the molecular connectivity and branch nature. For the coefficient of the parameters, pMIC can be positive correlation with X_1 , that was, the greater the molecular connectivity index, the better the compound activity, the smaller the MIC. From table 2, pMIC of compound 1e (0.68415), 3 (0.68672), 7 (0.68415), 8 (0.83664) with large X_1 were more than 0.3908. Since X_1 coefficient was 4.9471 and one of the biggest coefficient in equation 1, X_1 should be one of the most important parameters for activity. X_2 was molecular local charge surface parameter, combining molecular shape information with molecular charge information, and its numerical value was equal to polarizable surface area of all the negatively charged atomic solvent divided with all the negative charge, which was correlated to intestinal absorption of drug molecules. pMIC can be positive correlation with X_2 , that was, the greater the molecular local charge surface parameter, the better the compound activity. For example, X_2 value of compound 9, 11, 12 were as followed: $-1.34423 < -1.24819 < -1.09226$, and its activity also increased in turn. X_3 was molecular shape index and positive correlation with pMIC. Because only substituent (F, CF₃) and different position differences in benzene ring of compound 1c (2.9496), 1d (3.4101), 1e (3.60966) structure led to the shape index changes, the greater the X_3 , the better the compound activity, so were compound 14, 15, 17. X_4 was molecular flexible index and negatively correlated with pMICs, which showed that the greater the X_4 , the less the activity. For example, X_4 of compound 12, 14, 15, 16 were as followed: $2.74979 < 3.03742 < 5.00198 < 5.46379$, and the pMIC were as followed: $0.38722 > 0.13077 > 0.29885 > 0.90309$. In addition, X_4 of compound 1, 6, 1d, 8, 19, 17 were as followed: $3.4179 < 3.65628 < 4.00186 < 4.44497 < 4.72686 < 5.6714$, and the pMIC were as followed: $0.90309 > 0.63827 > 0.30103 > 0.3098 > -0.18184 > -0.9026$, which had the same rule.

3D-QSAR indirectly reflected nonbonding interaction characteristic between the ligand and receptor in the drug molecule action process, and had the rich connotation of physical chemistry. Before constructing the 3D-QSAR model, different compounds were needed to molecule overlap. In the molecular field analysis, different molecular overlap method will lead to different calculation results, and it directly decided the credibility of the model. After comprehensive investigation of the molecular structure on this study system, the molecular field and static electric field were selected as overlap

method. For model validation, when best principal component number for PLS was 3, maximum value for cross validation correlation coefficient q^2 was 0.744 and RMS error was minimum, which indicated that the model was highly credible and predictable.

From electrostatic field equipotential diagrams C, red area was distributed in the neighbouring of R_1 , R_3 and R_4 and the surface of thiazole ring, and this indicated that electro negativity group in R_1 , R_3 and R_4 can improve activity. For example, the pMIC of compound 1b greater than compound 5 was due to F as R_1 of compound 1b stronger electro negativity than H as R_1 of compound 5. The pMICs of compound 4, 6, 1e for 0.17609, 0.63827, 0.90309 increased in turn. However, from their structure analysis, only a difference of R_3 for H, Cl, CF_3 in their structure and electro negativity increases successively showed that the activity of compounds increased with the enhancement of a substituent electro negativity in R_3 . The activity of compound 1b (R_1 , R_3 =F, R_2 , R_4 =H) was more than compound 1c (R_1 , R_2 =H, R_3 , R_4 =F) activity, and compound 1e (R_3 = CF_3 , R_1 , R_2 , R_4 =H) activity was more than compound 1d (R_1 , R_2 = CF_3 , R_3 , R_4 =H) activity. On the structure, the differences of compound 1b, 1c and 1d, 1e were F and CF_3 in the different position, respectively. This showed that electro negativity F as R_1 had more biological activity than as R_4 , and electro negativity CF_3 as R_3 had more biological activity than as R_1 . It was concluded that the stronger electro negativity in R_3 can improve more biological activity than in R_1 and at the same time it in R_1 was better than in R_4 . Therefore, stronger electronegative substituent into R_3 , R_1 , R_4 in turn should be taken into account in designing ideal activity of thiazole nitazoxanide.

From steric field equipotential diagrams D, green area was mainly distributed in the neighbouring of R_1 , R_4 and R_5 , and this indicated that bigger volume than hydrogen atom in R_1 , R_4 and R_5 can improve activity. For example, the only differences of compound 1a (R_1 =OAc), 1 (R_1 =OH), 4 (R_1 =H) in structure were different substituent in R_1 . As the substituent volume in R_1 became smaller, the pMIC of three compounds (0.60206, 0.30103, 0.17609) also gradually became less and the activity weaker. Compound 7 (R_3 =H) activity more than compound 8 (R_3 =Cl) showed that larger volume Cl in R_3 can decrease the activity, and compound 1b (R_1 =F, R_3 =F) activity more than compound 5 (R_1 =H, R_3 =F), namely, was bulky F in the R_1 with improved activity.

In sum for steric field, stronger electro negativity and larger volume group in R_1 , weaker electro negativity and smaller volume groups in R_2 , stronger electro negativity and smaller volume group in R_3 , can help to enhance compound activity against *C. difficile*. From table 1, H as R_1 , R_2 , R_3 and R_4 in benzene ring of compound 4 didn't enhance electro negativity in R_1 , R_3 and R_4 position, and

nor increased a steric hindrance in R_1 and R_4 position. Because of the worst accordance with CoMFA model, compound 4 was eventually one of 13 thiazole nitazoxanide derivatives with less activity.

CONCLUSION

The 2D and 3D QSAR for nitazoxanide-based analogues were studied with DS-QSAR software. Through the genetic function approximation method, a simple 2D-QSAR equation was generated:

$$\text{pMIC} = 0.75552 + 4.9471 * X_1 + 1.397 * X_2 + 1.7507 * X_3 - 1.8442 * X_4 \text{ (model 1)}$$

$$n = 23, r^2 = 0.9052, r^2 (\text{adj}) = 0.8842, r^2 (\text{pred}) = 0.8402$$

The results showed that nitazoxanide-based analogues against *C. difficile* activity can be increased mainly through such as increase for the molecular connectivity index, increase for the molecular local charge surface parameters of the compound, increase for the compound molecular shape index and decrease for the compound molecular flexible index. The 3D model for $r^2=0.967$, $q^2=0.744$ was generated through CoMFA method. R_1 , R_3 and R_4 substituent were identified as the main factors for influencing compound activity by analyzing the equipotential diagram, and this was consistent with the results of 2D-QSAR model.

In a word, 2D and 3D-QSAR models with better predictability were constituted. They described the structure-activity relationship of nitazoxanide-based analogues and provided theoretical guidance for further designing new compounds with better activity.

REFERENCES

- Adagu IS, Nolder D, Warhurst D and Rossignol JF (2002). *In vitro* activity of nitazoxanide and related compounds against isolates of *Giardia intestinalis*, *Entamoeba histolytica* and *Trichomonas vaginalis*. *J. Antimicrob. Chemother.*, **49**: 103-111.
- Anderson VR and Curran MP (2007). Nitazoxanide: A review of its use in the treatment of gastrointestinal infections. *Drugs*, **67**: 1947-1967.
- Ballard T E, Wang X, Olekhovich I, Seymour C, Salamoun J, Warthan M, Hoffman PS and Macdonald TL (2011). Development of isoxazoline-containing peptidomimetics as dual $\alpha\beta3$ and $\alpha5\beta1$ integrin ligands. *Chem. Med. Chem.*, **6**: 362-377.
- Ballard TE, Wang X, Olekhovich I, Koerner T, Seymour C, Hoffman PS and Macdonald TL (2010). Biological activity of modified and exchanged 2-amino-5-nitrothiazole amide analogues of nitazoxanide. *Bioorg. Med. Chem. Lett.*, **20**: 3537-3539.
- Catherine SM and Rial DR (2000). *In vitro* and *in vivo* activities of nitazoxanide against *Clostridium difficile*. *Antimicrob. Agents Chemother.*, **44**: 2254-2258.

- Elvia D, Jaime M, Enrique R and Rosamaria B (2003). Epidemiology and control of intestinal parasites with nitazoxanide in children in Mexico. *Am. J. Trop. Med. Hyg.*, **68**: 384-385.
- Esposito M, Moores S, Naguleswaran A, Müller J and Hemphill A (2007). Induction of tachyzoite egress from cells infected with the protozoan *Neospora caninum* by nitro- and bromo-thiazolides, a class of broad-spectrum anti-parasitic drugs. *Int. J. Parasitol.*, **37**: 1143-1152.
- Esposito M, Müller N and Hemphill A (2007). Structure-activity relationships from *in vitro* efficacies of the thiazolide series against the intracellular apicomplexan protozoan *Neospora caninum*. *Int. J. Parasitol.* **37**: 183-190.
- Gilles HM & Hoffman PS (2002). Treatment of intestinal parasitic infections: A review of nitazoxanide. *Trends Parasitol.* **18**: 95-97.
- Glenn AP and Peter CA (2006). Activities of tizoxanide and nitazoxanide compared to those of five other thiazolides and three other agents against anaerobic species. *Antimicrob. Agents Chemother.*, **50**: 1112-1117.
- Guttner Y, Windsor HM, Viiala CH, Dusci L & Marshall BJ (2003). nitazoxanide in treatment of helicobacter pylori: a clinical and *in vitro* study. *Antimicrob. Agents Chemother.*, **47**: 3780-3783.
- Hansch C and Verma RP (2009). Diacylhydrazine derivatives as novel potential chitin biosynthesis inhibitors: Design, synthesis, and structure-activity relationship. *Eur. J. Med. Chem.*, **44**(1): 274-279.
- Hoffman PS, Sisson G, Croxen MA, Welch K, Harman WD, Cremades N and Morssh MG (2007). Antiparasitic drug nitazoxanide inhibits the pyruvate oxidoreductases of *Helicobacter pylori*, selected anaerobic bacteria and parasites, and *Campylobacter jejuni*. *Antimicrob. Agents Chemother.*, **51**: 868-876.
- Kamalakaran Anand Solomon, Srinivasan Sundararajan and Veluchamy Abirami (2009). QSAR Studies on *N*-aryl Derivative Activity Towards Alzheimer's Disease. *Molecules*, **14**: 1448-1455.
- Korba BE, Montero AB, Farrar K, Gaye K, Mukerjee S, Ayers MS and Rossignol JF (2008). Nitazoxanide, tizoxanide and other thiazolides are potent inhibitors of hepatitis B virus and hepatitis C virus replication. *Antiviral Research*, **77**: 56-63.
- Liu X W, Zhang H, Yang YJ, Li JY, Li B, Zhou XZ and Zhang JY (2014). Synthesis, Antibacterial Evaluation and Molecular Docking Study of Nitazoxanide Analogues. *Asia J. Chem.*, **26**(10): 2921-2926.
- Müller J, Rühle G, Müller N, Rossignol JF and Hemphill A (2006). A. *In vitro* Effects of Thiazolides on *Giardia lamblia* WB Clone C6 Cultured Axenically and in Coculture with Cac02 Cells. *Antimicrob. Agents Chemother.*, **50**: 162-170.
- Rossignol JF and Cavier R (1975). New derivative of 2-benzamino-5-nitrothiazoles. *Chem. Abstr.*, **83**: 28216.
- Rossignol JF and Maisonneuve H (1984). Nitazoxanide in the treatment of *Taenia saginata* and *Hymenolepis nana* infections. *Am. J. Trop. Med. Hyg.*, **33**: 511-512.
- Rossignol JF, Ayoub A and Ayers MS (2001). Treatment of diarrhea caused by *Cryptosporidium parvum*: A prospective randomized, double-blind, placebo-controlled study of Nitazoxanide. *J. Infect. Dis.*, **184**: 103-106.
- Rossignol JF, Frazia SL, Chiappa L, Ciucci A and Santoro MG (2009). Thiazolides, a new class of anti-influenza molecules targeting viral hem agglutinin at the post-translational level. *J. Biological Chem.*, **284**: 29798-29808.
- Rossignol JF, Maisonneuve H and Cho YW (1984). Nitroimidazoles in the treatment of trichomoniasis, giardiasis and amebiasis. *Int. J. Clin. Pharmacol. Ther. Toxicol.*, **22**: 63-72.
- Sharma D, Narasimhan B, Kumar P and Jalbout A (2009). Synthesis and QSAR evaluation of 2-(substituted phenyl)-1H-benzimidazoles and [2-(substituted phenyl)-benzimidazol-1-yl]-pyridin-3-yl-methanones. *Eur. J. Med. Chem.*, **44**(3): 1119-1127.
- Sisson G, Goodwin A, Raudonikiene A, Hughes NJ, Mukhopadhyay AK, Berg DE and Hoffman PS (2002). Enzymes associated with reductive activation and action of nitazoxanide, nitrofurans and metronidazole in *Helicobacter pylori*. *Antimicrob. Agents Chemother.* **46**: 2116-2123.
- Xu X, Hou T, Qiao X and Zhang W (2004). Computer aided drug design. Chemical industry press, Beijing, pp.281-286.