© 2007 🔇 SCIENCE IN CHINA PRESS

QSAR analysis of substituted benzylamino- and heterocyclylmethylamino-carbodithioate derivatives of 4-(3H)-quinazolinone using CoMFA and SCORE2.0

LIU Feng^{1,2}, LIU ShiYing¹, LIU Yan³, HAN DaXiong³, JIANG YuYang^{1,2†}, CAO ShengLi^{4†} & ZHAO YuFen^{2,3}

¹ Key Laboratory of Chemical Biology, Guangdong Province, Division of Life Science, Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China;

² Key Laboratory of Bioorganic Phosphorus Chemistry, Ministry of Education, Department of Chemistry, School of Life Science and Engineering, Tsinghua University, Beijing 100084, China;

³Key Laboratory for Chemical Biology of Fujian Province, Department of Chemistry, Xiamen University, Xiamen 361005, China;

⁴ Department of Chemistry, Capital Normal University, Beijing 100037, China

Thymidylate synthase (TS) is a critical enzyme for DNA biosynthesis and many nonclassical lipophilic antifolates targeting this enzyme are quite efficient and encouraging as antitumor drugs. In this paper, the binding model of 14 antifolates of substituted benzylamino- and heterocyclylmethylamino-carbodithioate derivatives of 4-(3H)-quinazolinone with TS is examined using molecular simulation methods — FlexiDock and SCORE2.0. The resulting conformation and orientation of these antifolates are directly applied to CoMFA study. The robust QSAR model, its three-dimensional contour map, and binding score of these antifolates derived from SCORE2.0 provide guidelines for structural optimization of current antifolates. The experiment indicates that deletion of cancer chemopreventive structure of dithiocarbamate is unfavorable for interaction between TS and antifolates.

thymidylate synthase, antifolate, FlexiDock, QSAR, SCORE

Thymidylate synthase (TS) catalyzes the conversion of dUMP to dTMP using the cofactor CH₂THF, providing the sole *de novo* means for synthesizing dTMP. TMP is required for DNA synthesis and inhibition of TS has proven to be an effective target for anticancer drug design^[1]. Structural modification of folic acid led to the discovery of a number of antifolates as efficient anticancer agents, for example, Raltitrexed (ZD1694, D16414), an antifolate, has been registered widely for the first-line treatment of advanced colorectal cancer^[1-8].</sup> However, these classical antifolates containing L-glutamic acid moiety in molecule have shortcomings such as drug resistance and liver toxicity^[2,3]. One strategy to overcome these shortcomings is to design nonclassical lipophilic inhibitors of folate requiring enzymes by deleting or modifying L-glutamic acid component from the

folate analogues^[9]. Cao et al.^[9–11] incorporated the cancer chemopreventive structure of dithiocarbamate moiety into 4-(3H)-quinazolinone, synthesized several series of compounds and found their *in vitro* antitumor activity. In this paper, the binding model of 14 antifolates of substituted benzylamino- and heterocyclylmethylaminocarbodithioate derivatives of 4-(3H)-quinazolinone synthesized by Cao et al. is examined using molecular simulation methods — Flexi-Dock and SCORE2.0, then to quantitatively disclose the relationship between activity and structure, we developed a 3D-QSAR model

www.scichina.com

Chinese Science Bulletin | December 2007 | vol. 52 | no. 23 | 3200-3206

Received March 20, 2007; accepted May 15, 2007

doi: 10.1007/s11434-007-0470-z

[†]Corresponding authors (email: jiangyy@sz.tsinghua.edu.cn; sl_cao@sohu.com) Supported by the National Natural Science Foundation of China (Grant No. 20572060) and Guangdong Fundamental Research Foundation of China (Grant No. 2005CCAO3400)

with these aligned antifolates using CoMFA. The final results provided some guidelines for structural optimization of current antifolates.

1 Materials and methods

1.1 Materials

A series of 14 substituted benzylamino- and heterocyclylmethylamino-carbodithioate derivatives of 4-(3H)quinazolinone used in this study are listed in Table 1. Eleven of these compounds were selected as the training set, the rest labeled with "a)" as the test set for the validation of the final model. The biological activity of each compound expressed as IC₅₀, which is the inhibitory concentration of compounds needed to inhibit 50% of K562 proliferation, was converted to pIC₅₀ (–lgIC₅₀) for the 3D-QSAR analysis. The inhibitory activities shown as IC₅₀ values were tested through the MTT^[12].

1.2 Molecular docking and alignment

Three-dimensional structure building and all modeling were performed using the Sybyl7.1 (Created Jun 17, 2005) program package. The crystal structure of the complex of human TS with D16414 (ZD1694, Tomudex) was recovered from the Brookhaven Protein Data Bank (http://www.rcsb.org/pdb/) (entry code 1HVY)^[13]. The potential of the 3D structure of TS was assigned according to the Tripos Force Field^[14] with Kollmanall-atom^[15] charges in Sybyl7.1. The antifolates were built in the sketch molecule package. Each structure was fully geometry optimized using Powell method in Tripos force field with a distance-dependent dielectric function and a 0.005 kcal/(mol·Å) energy gradient convergence criterion. Partial atomic charges were calculated using Gasteiger-Hückel method^[16].

CoMFA results may be extremely sensitive to a number of factors, such as alignment rules, overall orientation of aligned compounds, lattice shifting, step size and the probe atom type^[17]. The selection of bioactive conformer and ascertainment of alignment rule are the most critical factors to 3D-QSAR. In principle, the actual bioactive conformation and the best alignment can only be derived from the complex structure of ligand and receptor. So we performed a docking-guided conformation selection by FlexiDock^[18,19] in Sybyl7.1. The Lamarckian genetic algorithm (LGA)^[20] was applied to dealing with the inhibitor-enzyme interactions. Replac-



LIU Feng et al. Chinese Science Bulletin | December 2007 | vol. 52 | no. 23 | 3200-3206

3201

ing the D16414 with these antifolates one by one into the complex structure 1HVY, structurally optimized receptor, the initial conformations of these antifolates were docked into the binding site among the 3D-structural models of TS. Then the docked ligands were subjected to flexible docking calculation employing FlexiDock in Sybyl7.1. All the single bonds of ligand and the side chains of the amino acid residues around the ligand, as well as the orientation of the ligand were taken as variables within the interaction region for flexible docking calculation. Then, we obtained the initial structures of the ligand-receotor complexes, which were successively refined using minimization. The binding energy (E_{bind}) of the ligand was calculated using the following formula:

$$E_{\text{bind}} = E_{\text{complex}} - E_{\text{ligand}} - E_{\text{receptor}}; \qquad (1)$$

where E_{ligand} is the energy of the ligand corresponding to the lower energy conformation and E_{receptor} is the energy of the receptor.

The conformer with the lowest binding energy was extracted from the docking result file and used for alignment. Aligned molecular aggregates were obtained by the Sybyl7.1 routine "Align Database". TMA was chosen as the template and the rest of training set molecules were aligned to it by the common substructure labeled with * in Figure 1 for the CoMFA^[21] study.



Figure 1 Template molecule TMA.

1.3 CoMFA and 3D-QSAR model

For the CoMFA calculation, the aligned molecules were placed one by one into a 3D cubic lattice with a 2Å spacing. The default sp3 carbon atom with +1 charge was selected as the probe atom for the calculations of the steric (Lennard-Jones 6–12 potential) and electrostatic fields (Coulombic potential) around the aligned molecules with a distance-dependent dielectric constant at each lattice point. A 30 kcal/mol energy cutoff was applied, which means steric and electrostatic energy greater than 30 kcal/mol was truncated to that value. Steric and electrostatic fields generated were scaled by the CoMFA-STD method in Sybyl7.1. The correlation of the CoMFA descriptor with corresponding pIC₅₀ was evaluated by the PLS regression analysis module^[22–24] in Sybyl7.1. To avoid overfitted 3D-QSAR, the opti-

mum number of components (N) to be used in the model derivation was chosen from the analysis with the highest cross-validated correlation coefficient (q^2) . The crossvalidated q^2 quantifies the predictive ability of the model and is determined by a leave-one-out (LOO) procedure of cross-validation in which each compound is successively removed from the model derivation and its pIC_{50} value can be predicted using the model built from the remaining compounds. The optimum number of components was determined in such a way that each additional component should increase the q^2 by at least $5\%^{[25]}$. In order to speed up the analysis and reduce noise, the column filtering was set to 1.0 kcal/mol, so that only those steric and electrostatic energy with values greater than 1.0 kcal/mol was considered in the PLS analysis. The CoMFA descriptors were used as independent variables, and the pIC₅₀ values were used as dependent variables in PLS regression analyses to derive the 3D-QSAR model using the standard implementation in the Sybyl7.1. The predictive value of model was evaluated first by leave-one-out cross-validation. The cross-validation coefficient, q^2 , was calculated by using relative equation^[26]. Cho et al.^[17] reported that q^2 value was sensitive to the orientation of aligned molecules on the computer terminal and might vary with the orientation by as much as 0.5 q^2 units. So in our CoMFA analysis, all-orientation and all-placement searching was performed by rotating the molecular aggregate systematically every 30° along the X, Y, and Z axes and translating it every 0.2 Å, at every place, the cross-validated q^2 of PLS calculated, until obtaining a maximal q^2 value. In this paper, the maximal q^2 is obtained by translating aligned molecules 1 Å along the horizontal axis at two directions respectively at 2 Å step size and 1 kcal/mol column filtering.

1.4 Score

The 3D-structure of antifolates extracted from ligandreceptor complexes after FlexiDock was scored using SCORE2.0^[27–29].

$$pK_{d} = 2.254+(0.916)MB-(0.168)VB+(0.141)WHB +(0.216)MHB+(0.593)SHB +(0.327)WWH -(0.708)MWH+(0.291)SWH+(1.178)HM -(0.169)RT.$$
 (2)

MB, coordinate bonding with metal ion; VB, VDW bump; WHB, weak H-bond; MHB, moderate H-bond; SHB, strong H-bond; WWH, weak water H-bond; MWH, moderate water H-bond; SWH, strong water

3202

ORGANIC CHEMISTRY

H-bond; HM, hydrophobic matching; RT, rotatable single bond.

2 Results and discussion

2.1 Molecular docking

The initial conformation of these antifolates was employed to perform flexible docking calculation by the FlexiDock program in Sybyl7.1. Twenty solutions for each antifolate were obtained after calculation and the complex structure with the lowest energy was chosen for study. The conformation and orientation of each antifolate were extracted from the complex structure for further research. The binding mode of substituted benzylamino- and heterocyclylmethylamino- carbodithioate derivatives of 4-(3H)-quinazolinone with TS could be explained explicitly with the 3D structures of these antifolate-TS complexes. In the 3D structure of docked complex, the quinazolinone and the lipophilic moieties of the 14 antifolates were in the hydrophobic groove formed by hydrophobic residues of TS, such as Phe80, Val79, Phe225, Gly83, Val84, Leu221, Asn226, Ile108, Met311, Pro224, Val313, His196, Gly222, Trp109 and Leu192. Hydrogen bonding (Table 2) was the important intermolecular interaction.

 Table 2
 Hydrogen bonding between these antifolates and TS

-	2	U	0			
0	Compound	Receptor (No.)	Residue	Donor (No.)	Distance (Å)	Angle (°)
	Asp218	CO	2FB	N(3)	2.408	131.2
	ADM	O(1)	Asn112	$-NH_2$	1.473	163.98
	Leu221	CO	TMA	NH(16)	2.210	160.36
	Asp218	CO	QPM	NH(3)	1.787	143.82

For example, the distance between the H of NH(16) of TMA and CO of the residue Leu221 side chain was 2.210 Å, shown with a dash line in red (Figure 2).

Table 3 lists the binding energy data of the 14 antifolate-TS complexes. A classical QSAR was performed to explore whether the inhibitory activities of these antifolates could be correlated with the binding energy. Employing the PLS method in Sybyl7.1, we calculated the regression equation for the inhibitory activities (pIC₅₀), using the binding energy E_{bind} as the sole descriptor variable. A good correlation was found between the inhibitory activities and binding energy.

$$pIC_{50} = -1.580 + 0.002 E_{bind},$$
 (3)

which is an indirect proof for the reasonability of the 3D structures of antifolate-TS complexes predicted by our

modeling method.

Table 3	Binding energy of 14 antifolate-TS complexes	5

	-			1	
Compound	pIC ₅₀	ΔE (kcal/mol)	Compound	pIC ₅₀	ΔE (kcal/mol)
2FB	-2.59	-481.44	AMB	-2.40	-532.25
4CB	-2.58	-486.19	QAT	-2.08	-407.71
4FB	-2.56	-474.10	QMT	-2.26	-367.52
4MB	-2.61	-480.26	QPF	-2.63	-474.65
ADM	-2.20	-467.32	QPM	-2.56	-464.78
DMB	-2.35	-365.77	TFA	-2.30	-468.55
OMA	-2.55	-448.94	TMA	-2.43	-436.95

2.2 3D-QSAR model and CoMFA analysis

Eleven of 14 antifolates constitute the training set, the rest labeled with "a" the test set. The aligned molecular aggregates are shown in Figure 3. A QSAR model was obtained using the methods of docking-guided conformer selection and all orientation and all-placment searches. The PLS statistics of the CoMFA analyses are summarized in Table 4. The cross-validated q^2 value was 0.534 with 6 components and noncross-validated conventional r^2 value was 0.999 with a standard error of estimate (SEE) value of 0.007. The relative contribution between steric and electrostatic fields for CoMFA model was 0.669 and 0.331 respectively.

Table 4	The PLS statistics of the CoMFA	
		-

Cross-	validation	Non-cross-validation			
$R_{\rm cross}^2$	Component	R^2	S	F	
0.534	6	0.999	0.007	1367.666	

The contours of the steric and electrostatic fields were displayed. The steric contours (Figure 4) are displayed in yellow and green and the electrostatic contours (Figure 5) in red and blue.

Introduction of either the electron-donating groups ($-OCH_3$) or the electron-withdrawing groups (-F, -Cl and -COOH) at the 4'-position of the phenyl ring did not cause the increase in bioactivity compared with the parent compound BZA^[9], whose pIC₅₀ was 2.40. Replacement of the phenyl group in compound BZA by 2-thiophenyl, and 2-tetrahydrofuryl, resulting in compounds TMA and TFA, also leads to the decrease in activity.

2.3 Validation of the QSAR model

To test the stability and predictive ability of the 3D-QSAR model, 3 antifolates, which were not included in the CoMFA model, were selected as a set for valida-

tion. The residual between experimental and predicted pIC_{50} for this test set is listed in Table 5. The cor-



Figure 2 3D structure of the TMA complex with TS and the binding score of TMA. (a) TMA complex with TS (one hydrogen bond shown with a dash line in red); (b) binding score of TMA.



Figure 3 The aligned molecular aggregates.



Figure 4 The steric field contours of the QSAR model. The yellow contours indicate regions of negative steric potential, while the green contours the regions of positive steric potential.

Compound	Actual	Predicted	Residual
4MB	-2.61	-2.598	-0.012
AMB	-2.4	-2.39	-0.01
QPF	-2.63	-2.626	-0.004

relation between experimental and predicted pIC₅₀ for



Figure 5 The electrostatic field contours of the QSAR model. The red contours indicate regions of negative electrostatic potential and the blue contours the regions of positive electrostatic potential.

training set is shown in Figure 6. It was shown that the model was stable and had robust predictive ability.

2.4 Score

Performing scoring for every antifolate within the docked complex, we obtained a score for every atom of every antifolate and assessed the contribution of every atom to the interaction between antifolates and TS (Ta-

3204

ble 6). The atoms, with score above 0.1, were favorable for the interaction, and those with score below -0.1 were unfavorable. The common favorable atoms of these 14 antifolates were phenyl moiety of quinazolinone and the phenyl moiety connected with the $-CH_2$. The common unfavorable atoms of these 14 antifolates were $-CH_2$ or -NH connected with the phenyl moiety.



Figure 6 Predicted versus experimental inhibitory activities for training set.

Table 6Binding score of TMA

No.	Туре	Score	No.	Туре	Score	No.	Туре	Score
1	0.2	0.000	9	C.ar	0.343	17	S.2	-0.336
2	C.3	0.557	10	C.ar	0.258	18	C.3	-0.169
3	N.2	0.000	11	C.ar	0.343	19	C.ar	0.258
4	C.3	0.268	12	C.ar	0.343	20	C.ar	0.343
5	N.3	0.000	13	C.ar	0.343	21	C.ar	0.000
6	C.2	0.000	14	S.3	0.439	22	C.ar	0.000
7	C.2	0.000	15	C.2	-0.169	23	S.2	0.000
8	C.ar	0.343	16	N.3	-0.028			





The subscript number represents atom number.

- Marriott J H, Neidle S, Matusiak Z, et al. Chemoenzymatic preparation of the novel antifolate thymidylate synthase inhibitor *N*-(4-{*N*-[(6*S*)-2-methyl-4-oxo-3,4,7,8-tetrahydro-6*H*-cyclopenta[*g*] quinazolin-6-yl]-*N*-(prop-2-ynyl)amino}-benzoyl)-L-glutamic acid and its glutamyl cleavage product. J Chem Soc, Perkin Trans, 1999, 1: 1495
- 2 Kelland L R, Kimbell R, Hardcastle A, et al. Relationships between resistance to cisplatin and antifolates in sensitive and resistant tumour cell lines. Eur J Cancer, 1995, 3IA: 981–986
- 3 Jackman A L, Farrugia D C, Gibson W, et al. ZD1694 (Tomudex): A

3 Conclusion

We applied FlexiDock, CoMFA, and SCORE2.0 to research the binding mode of a set of antifolates to TS receptor. After flexible docking, the hydrogen bonds between antifolates and the residues of TS were discovered. FlexiDock results indicated that the binding energy of the antifolates correlates well with the experimental inhibitory activities against TS, and the modeling results provided satisfactory explanation for the binding mode of these antifolates with TS. The probable conformations were selected by the docking-guided conformer selection method, and the 3D-QSAR model was obtained by the method of all-orientation and all-placement, whose statistical parameter (PLS) was ideal. Comparing to the compounds of FUR, DET, MOR, PDQ, etc.^[9], the bioactivities of these antifolates of ADM, TMA, TFA, 4MB, DMB, 4FB, 2FB, 4CB, and AMB in this paper with -CH₂ connected to the phenyl were much low. The scores of -CH₂ of those compounds mentioned above in this study were all below -0.1 using SCORE2.0, so introduction of -CH2 into these compounds was unfavorable. The bioactivities of another 5 antifolates of QAT, QMT, OMA, QPM, and QPF deletion of dithiocarbamate were also much low, the scores of -N which connect with the phenyl moiety were all below -0.1, so deletion of cancer chemopreventive structure of dithiocarbamate was unfavorable.

Based on the result of 3D-QSAR, the relative contribution between steric and electrostatic fields for the CoMFA model was 0.669 and 0.331 respectively, so steric field played dominant role in antifolate-TS interaction. In this study, we combined the results of 3D-QSAR with SCORE2.0, which leads to a better understanding of important antifolate-TS interactions and thus provided guidelines for ligand design plus a predictive model for scoring novel synthetic candidates.

new thymidylate synthase inhibitor with activity in colorectal cancer. Eur J Cancer, 1995, 31A: 1277–1282

- Peters G J, Smitskamp-Wilms E, Smid K, et al. Determinants of activity of the antifolate thymidylate synthase inhibitors tomudex (ZD1694) and GW1843U89 against mono- and multilayered colon cancer cell lines under folate-restricted conditions. Cancer Res, 1999, 59: 5529-5535
- 5 Lewis N L, Scher R, Gallo J M, et al. Phase I and pharmacokinetic study of irinotecan in combination with raltitrexed. Cancer Chemoth

Pharm, 2002, 50: 257-265

- 6 Bavetsias V, Clauss R, Henderson E A. Antifolate chemistry: Synthesis of 4-{N-[(6RS)-2-methyl-4-oxo-3,4,7,8-tetrahydro-6H-cyclo-penta[g]quinazolin-6-yl]-N-(prop-2-ynyl)amino}benzoic acid via a (propargyl)Co2(CO)6+ complex. Org Biomol Chem, 2003, 1: 1943-1946
- 7 McGuire J J. Anticancer antifolates: Current status and future directions. Curr Pharm Des, 2003, 9: 2593-2613
- 8 Hughes L R, Jackman J A L, Oldfield J. Quinazoline antifolate thymidylate synthase inhibitors: Alkyl, substituted alkyl, and aryl substituents in the C-2 position. J Med Chem, 1990, 33: 3060-3067
- 9 Cao S L, Feng Y P, Jiang Y Y. Synthesis and *in vitro* antitumor activity of 4(3H)-quinazolinone derivatives with dithiocarbamate side chains. Bioorg Med Chem Lett, 2005, 15: 1915-1917
- 10 Cao S L, Feng Y P, Zhang X L. Synthesis of substituted benzylaminoand heterocyclylmethylamino carbodithioate derivatives of 4-(3H)-quinazolinone and their cytotoxic activity. Arch Pharm Chem Life Sci, 2006, 339: 250-254
- 11 Liu S Y, Liu F, Yu X Q. The 3D-QSAR analysis of 4-(3H)quinazolinone derivatives with dithiocarbamate side chains on thymidylate synthase. Bioorgan Medic Chem, 2006, 14: 1425-1430
- 12 Mosmann T. Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. J Immunol Methods, 1983, 65: 55
- 13 Phan J, Koli S, Minor W. Human thymidylate synthase is in the closed conformation when complexed with dUMP and raltitrexed, an antifolate drug. Biochemistry, 2001, 40: 1897
- 14 Clark M, III Cramer R D, Van Opdenbosch O N J. The Tripos force field. J Comput Chem, 1989, 10: 982-1012
- 15 Weiner S J, Kollman P A, Case D A. A new force field for molecular mechanical simulation of nucleic acids and proteins. J Am Chem Soc, 1984, 106: 765-784
- 16 Tripos: St. Louis MO. Details of the implementation are given in

Sybyl6.5. Theory Manual, 1998, 69

- 17 Cho S J, 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
- 18 Tripos Inc. 1699 South Hanley Rd, St. Louis, MO, USA. Sybyl6.5 Manual, 1999
- Judson R. Reviews in Computational Chemistry. Lipkowitz K B, Boyd D B, ed. New York: VCH Publishers, 1997. 10
- 20 Morris G M, Goodsell D S, Halliday R S. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem, 1998, 19: 1639
- 21 Cramer R D, III David R D, Patterson E. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. J Am Chem Soc, 1988, 110: 5959-5967
- 22 Stahle L, Wold S, Ellis G P. Progress in Medicinal Chemistry. Amsterdom: Elsevier, 1988. 292
- 23 Solis F J, Wets R J B. Minimization by random search techniques. Math Oper Res, 1981, 6: 19-30
- 24 Bush B L, Jr Nachbar R B. Sample-distance partial least squares: PLS optimized for many variables, with application to CoMFA. J Comput Aided Mol Des, 1993, 7: 587-619
- 25 Gohlke H, Schwarz S, Gundisch D. 3D QSAR analyses-guided rational design of novel ligands for the $(\alpha 4)_2(\beta 2)_3$ nicotinic acetylcholine receptor. J Med Chem, 2003, 46: 2031–2048
- 26 Andrew R L. Molecular Modeling Principles and Applications. Dorchester: Henry Ling Ltd. Press, 2001. 695-702
- 27 Wang R X, Liu L, Lai L H. SCORE: A new empirical method for estimating the binding affinity of a protein-ligand complex. J Mol Model, 1998, 4: 379-394
- 28 Li C H, Ma X H, Chen W Z. The study of the scoring function in protein-protein docking. Acta Biophys Sin, 2003, 19: 47-52
- 29 Liu L, Wang X R, Lai L H. 3D-QSAR and pharmacophore modeling of growth hormone secretagogues. Acta Phys-Chim Sin, 1997, 13: 1090-1096