ARTICLE

3D-QSAR Study of 7,8-Dialkyl-1,3-diaminopyrrolo-[3,2-f] Quinazolines with Anticancer Activity as DHFR Inhibitors

Jin-can Chen[a,b], Lan-mei Chen[c], Si-yan Liao[b], Li Qian[b], Kang-cheng Zheng[b*]

a. Analysis Centre of Guangdong Medical College, Zhanjiang 524023, China
b. MOE Laboratory of Bioinorganic and Synthetic Chemistry, School of Chemistry and Chemical Engineering, Sun Yat-Sen University, Guangzhou 510275, China
c. School of Pharmacy, Guangdong Medical College, Zhanjiang 524023, China

(Dated: Received on October 22, 2008; Accepted on January 20, 2009)

A three-dimensional quantitative structure-activity relationship (3D-QSAR) study of a series of 7,8-dialkyl-1,3-diaminopyrrolo-[3,2-f] quinazolines with anticancer activity as dihydrofolate reductase (DHFR) inhibitors was carried out by using the comparative molecular field analysis (CoMFA), on the basis of our reported 2D-QSAR of these compounds. The established 3D-QSAR model has good quality of statistics and good prediction ability; the non cross-validation correlation coefficient and the cross-validation value of this model are 0.993 and 0.619, respectively, the F value is 193.4, and the standard deviation SD is 0.208. This model indicates that the steric field factor plays a much more important role than the electrostatic one, in satisfying agreement with the published 2D-QSAR model. However, the 3D-QSAR model offers visual images of the steric field and the electrostatic field. The 3D-QSAR study further suggests the following: to improve the activity, the substituent R′ should be selected to be a group with an adaptive bulk like Et or *i*-Pr, and the substituent R should be selected to be a larger alkyl. In particular, based on our present 3D-QSAR as well as the published 2D-QSAR, the experimentally-proposed hydrophobic binding mechanism on the receptor-binding site of the DHFR can be further explained in theory. Therefore, the QSAR studies help to further understand the "hydrophobic binding" action mechanism of this kind of compounds, and to direct the molecular design of new drugs with higher activity.

Key words: Dihydrofolate reductase, Quinazoline, 3D-QSAR, Comparative molecular field analysis

I. INTRODUCTION

Dihydrofolate reductase (DHFR) catalyzes the reduction of 7, 8-dihydrofolate to 5, 6, 7, 8-tetrahydrofolate using nicotinamide adenine dinucleotide phosphate (NADPH) as a cofactor [1]. Since DHFR plays an important role in the DNA synthetic pathway, its prominence as a target in antibacterial, antiparasitic, and antineoplastic chemotherapy is unsurpassed [2]. The classical inhibitor of DHFR is methotrexate (MTX) (**1**). MTX shows certain effectivity in cancer chemotherapy, but its hydrophilicity is greater and it is the hydrophilicity that restricts its distribution to various body tissues, and prevents it from entry into cells by diffusion [3].

In recent years, quinazolines with anticancer activity have been studied as inhibitors of DHFR, since they have strong lipophilicity than MTX, and they can improve the resistance to tumor cells. Many studies of structure-activity relationships and molecular modifications have been implemented, but the results are not very satisfying [4,5]. Recently, some potent DHFR inhibitors, *i.e.*, a series of 7,8-dialkyl-1,3-diaminopyrrolo-[3,2-f] quinazolines have been designed and reported by Kuyper *et al.* [6]. They are relatively small and compact in molecular size but show high affinity against *C. albicans* and an array of tumor cell lines. Recently, we have investigated the 2D-QSAR of this series of compounds [7]. In this published work, it is very interesting to find that the established best QSAR equation of the studied series involves only two descriptors: lipophilicity indexes CLog$P$ and (CLog$P$)^2, and they only depend on one variable. However, such descriptors can quite well describe a significant statistic quality and have remarkable predictive activity according to the square of adjusted correlation coefficient ($R^2_{\rm A}$=0.937) and the square of cross-validated coefficient (q^2=0.911) of the established equation. Obviously, such an unusual 2D-QSAR must relate to the particular action mechanism of this kind of compound, that is, the experimentally proposed hydrophobic binding mechanism on the receptor-binding site of the DHFR. To theoretically

*Author to whom correspondence should be addressed. E-mail: ceszkc@mail.sysu.edu.cn

clarify such a particular action mechanism, a further investigation, a three-dimensional quantitative structure-activity relationship (3D-QSAR) study is necessary and significant.

3D-QSAR method was used to build a QSAR model with high predictive ability and map the characteristics of a receptor active site by analyzing the characteristics of structure of a series of compounds on the 3D level. Comparative molecular field analysis (CoMFA) is the most widely applied method for the study of quantitative structure-activity relationships on the 3D level [8-11]. In CoMFA studies, steric and electrostatic properties are calculated according to Lennard-Jones and Coulombic potentials, respectively. Then the relationship between independent variables (steric and electrostatic potentials) and biological activity is developed by partial least square (PLS). It is a useful method in understanding the pharmacological properties of studied compounds, because not only the 3D models are vivid and robust, but also the obtained color contour maps may further help to understand the nature of the interaction on the ligand with the active site of the receptor [12,13].

In this work, a 3D-QSAR study of a series of 7,8-dialkyl-1,3-diaminopyrrolo-[3,2-f] quinazolines with anticancer activity was carried out using CoMFA method. The purpose of this work is to establish an optimal 3D-QSAR model, and hereby further reveal the factors, which may be related to the particular action mechanism of such a compound. We believe that a combined 3D and 2D-QSAR investigation must be more advantageous to understanding the interaction mechanism between the drug molecule and the DHFR, and to directing the molecular design of new drugs with higher activity.

II. METHODS

All CoMFA studies were performed using SYBYL v6.9 molecular modeling software (Tripos Inc., St. Louis, MO.) running on SGI R2400 workstations with default values except where specially explained below.

A. Compounds and biological data

All 19 members of 7,8-dialkyl-1,3-diaminopyrrolo-[3,2-f] quinazoline derivatives are selected from Ref.[6]. This series of compounds have the same skeleton and general structural diagram, shown in Fig.1. Anticancer activity is generally expressed by IC_{50} value which refers to the milli-molar concentration of the compound required for 50% inhibition against an HCT-8 cell (human ileocecal adenocarcinoma). These values are transformed to pIC_{50} (negative logarithm of IC_{50}) in CoMFA studies.

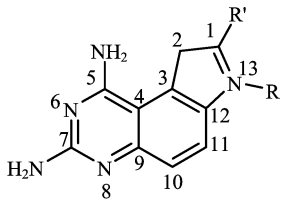

FIG. 1 Structural diagram of 7,8-dialkylpyrrolo-[3,2-f] quinazolines (the substituents of R and R′ are shown in Table I).

TABLE I CoMFA results of this series of compounds.

No.	Structure		pIC_{50}		ΔpIC_{50}
	R′	R	Expt.	Pred.	
1	H	H	5.839	5.815	0.024
2	H	Me	6.237	6.294	–0.057
3*	Me	Me	7.071	7.178	–0.107
4	Et	Me	8.097	8.043	0.054
5	*t*-Bu	H	8.481	8.494	–0.013
6	Et	Et	9.229	9.270	–0.041
7	Me	*i*-Pr	8.131	8.045	0.086
8	H	$CHEt_2$	8.222	8.196	0.026
9*	*i*-Pr	Et	8.854	9.053	–0.199
10	*n*-Pr	Et	8.509	8.542	–0.033
11	Et	*t*-Bu	8.886	8.920	–0.034
12	Me	$C(Me)_2Et$	8.678	8.697	–0.019
13*	Et	CH(Me)Et	9.086	8.871	0.215
14	Me	$CHEt_2$	9.131	9.104	0.027
15	*t*-Bu	*i*-Pr	8.886	8.899	–0.013
16	Et	$CHEt_2$	9.244	9.224	0.020
17	*n*-Bu	*t*-Bu	7.284	7.317	–0.033
18*	*n*-Pr	$CHEt_2$	8.357	8.458	–0.101
19	*i*-Pr	$CHEt_2$	9.125	9.130	–0.005

* selected at random as the test set.

B. Conformational analysis and alignment rule

Prior to the construction of a CoMFA model, the determination of the active conformation is the crucial step in CoMFA analysis. We obtained the structure of GW345 (compound **8** in this work) from the DHFR-NADPH-GW345 ternary complex crystal [14] and the structure of GW345 was optimized by using MM+ method. Comparing these two conformations obtained from experiment and calculation, no obvious difference is found. This suggests that the conformation of GW345 in crystal, *i.e.* the bioactive conformation, is similar to that of the calculated energy minimum. The superimposition diagram of the two conformations is shown in Fig.2. The pentyl moiety adopts such a conformation in which its carbon chain orientation is almost perpendicular to the pyrroloquinazoline ring sys-

DOI:10.1088/1674-0068/22/03/285-289

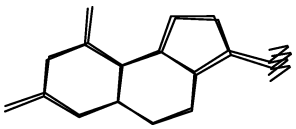

FIG. 2 Superimposition diagram of the energy-minimized configuration and crystal configuration of compound **8**.

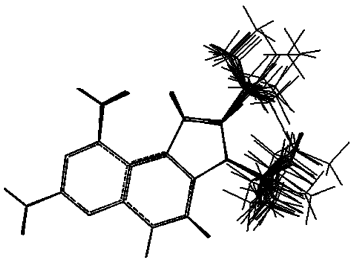

FIG. 3 Superimposition diagram of 7,8-dialkylpyrrolo-[3,2-f] quinazolines.

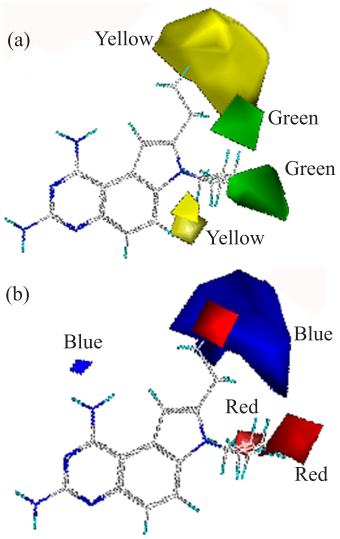

FIG. 4 Steric (a) and Electrostatic (b) field contour map of CoMFA. For interpretation of the color in this figure legend, the reader can refer to the web version of this article.

tem. The crystal structure of DHFR-NADPH-GW345 ternary complex comes from the Protein Data Bank under the entry 1AOE.

All 19 compounds were built by using the SYBYL Build program with the default SYBYL settings. The fully geometrical optimization of 19 compounds were performed using the standard Tripos force field including the electrostatic term calculated from Gasteiger and Hckel atomic charges. The methods of steepest descent and conjugate gradients were used for the energy minimization and the cut-off value was set to 0.21 kJ/mol.

Structural alignment is also a critical step in CoMFA studies [15,16]. In this work, the lowest energy conformation of compound **16** (the most potent activity compound) was selected as the template structure for molecular alignment. The skeletal atoms numbered 1 to 13 in Fig.1 were selected for the alignment of all compounds. The alignment of all the compounds with bioactive conformation is shown in Fig.3, in which pyrimidine ring, pyrrole ring and benzene ring are almost in the same plane.

C. CoMFA and partial least square (PLS) analysis

The CoMFA studies were performed with the QSAR module of Sybyl 6.9. The steric and electrostatic energies were calculated using sp^3 carbon probe with charge +1, with 0.2 nm grid spacing. Maximum energy cut-off for steric and electrostatic energies was 124 kJ/mol. Four compounds (marked * in Table I) were selected at random as the test set and the remaining compounds were used as the training set to build the CoMFA model. The PLS method was used to construct and validate the CoMFA models, and the column filtering value was set to 0 kJ/mol. Cross-validation was performed with the leave-one-out procedure. The optimum number of components in the final PLS model was determined by the highest cross-validation value (q^2) obtained from the leave-one-out cross-validation technique.

III. RESULTS AND DISCUSSION

A. CoMFA model

The CoMFA of this series of anticancer agents gives a good q^2 of 0.619 with an optimized component number of 3. The square of non-cross-validation correlation r^2 is 0.993, F is 193.4, and the estimated standard error is 0.208. For a reliable predictive model, q^2 should be greater or equate to 0.5, so we see that the established CoMFA model has a good predictive ability. The analysis results are presented in Table I, in which the residual values (ΔpIC_{50}) between experimental and calculated activity values in the training set are –0.005-0.086, and those in the test set are –0.101-0.215.

The established 3D-QSAR model shows that the steric field descriptor of the CoMFA model gives 89% of the total contribution, while the electrostatic descriptor only gives the remaining 11%. This proportion indicates that the steric field is absolutely predominant, in good agreement with the study of 2D-QSAR. The further discussion of 3D and 2D-QSAR will be carried out in a later section below.

B. CoMFA contour maps

The steric and electrostatic contour maps are displayed in Fig.4 respectively. In the contour maps, the green-colored regions indicate the areas where steric

DOI:10.1088/1674-0068/22/03/285-289

bulk increases activity, while the yellow contours indicate the regions where steric bulk is detrimental to the biological activity. Blue-colored regions show the areas where electropositive-charged group enhances activity, while red regions represent the regions where electronegative-charged group improves the activity.

As shown in Fig.4(a), a large region of yellow contour is located on the top of substitutent R′ and a large middle region of green contour is located on the underside of substitutent R′. This suggests that substituent R′ should be selected to be a group with an adaptive bulk like Et or *i*-Pr, not to be smaller than Et, and not to be *n*-Pr group, because the former acclimatizes itself to the structural characteristic of the steric map and the latter does not. For example, the activity of compound **4** (R′=Et, pIC_{50}=8.097) is higher than those of complexes **3** (R′=Me, pIC_{50}=7.071) and **2** (R′=H, pIC_{50}=6.237); the activity of compound **19** (R′=*i*-Pr, pIC_{50}=9.125) is quite higher than that of compound **18** (R′=*n*-Pr, pIC_{50}=8.357), and that of compound **9** (R′=*i*-Pr, pIC_{50}=8.854) is quite higher than that of compound **10** (R′=*n*-Pr, pIC_{50}=8.509). Molecular modeling also suggests that this class of inhibitor (binds to DHFR) with R′ occupies the large hydrophobic cleft in a manner similar to the dimethoxybenzyl group of piritrexim (PTX, a potent lipophilic inhibitor of DHFR) [6]. According to "lock-key theory", moderate R′ with appropriate shape can enter into the enzyme binding pocket and display higher affinity for DHFR. This suggestion is consistent with the 2D-QSAR results mentioned above. In addition to the above examples, we can also see that compound **11** (R′=Et) shows higher activity (pIC_{50}=8.886) than compound **17** (R′=*n*-Bu, pIC_{50}=7.284). Complex **16** (R′=Et, pIC_{50}=9.224) has the highest activity among complexes **14** (R′=Me, pIC_{50}=9.131), and **19** (R′=*i*-Pr, pIC_{50}=9.125), among which the structures are the same except for R′. This is because the R′ of complex **16** has an more adaptive bulk than those of complexes **14** (Me) and **19** (*i*-Pr) and thus easily enters into the enzyme binding pocket and displays higher inhibition affinity towards DHFR.

From Fig.4(a), we can also see that there is a green region outside of R, suggesting that the enzyme active site imposes fewer restrictions on size for substituent R and thus a larger R group is generally favored. For example, compound **8** (R=$CHEt_2$) displays higher activity (pIC_{50}=8.222) than compounds **1** (R=H, pIC_{50}=5.839) and **2** (R=Me, pIC_{50}=6.237).

The electrostatic contour map (Fig.4(b)) shows that blue or red regions are mainly located in positions outside of R and R′. Analysis with the GRID program [17] using a methyl group probe shows the interaction energy information between a charged methyl group, which comes from substructures Met-25, Ile-33, Phe-36, Met-54, etc., and subsituent R′ (or R) of enzyme-bound pyrroloquinazoline [6]. From Fig.4(b), we can see a large blue region around the terminal of R′ (Et, here). This fact can be used to further explain why the complexes with R′ of Et have higher activity, because the H atoms (positive-charged) of Et just fall into the blue area in addition to Et with optimal steric factor. For example, complexes **4** (pIC_{50}=8.097), **6** (9.229), **11** (8.886), and **16** (9.244) all have correspondingly high activities. In addition, we can also see that two smaller red regions exist near the R chain. Since the substituent R is generally selected to be alkyl, and thus the difference of substituent R on charge is very small, it can be predicted that the difference in the electrostatic interaction between the ligand and receptor must also be very small. Therefore, the selection of R is almost independent of electrostatic factor.

In short, to improve the activity, the substituent R′ should be selected to be a group with an adaptive bulk like Et or *i*-Pr, and the substituent R should be selected to be an alkyl with greater bulk.

C. Inquiry of action mechanism

According to our previous work [7], although a great number of parameters of electronic structures and molecular properties as descriptors have been considered, the best QSAR model for this series of compounds is only with two selected parameters CLog$P$ and (CLog$P$)^2 by stepwise multiple linear regression, and substantially, only one variable (CLog$P$) is involved. Obviously, this is very interesting because this model is quite different from general QSAR models in which some parameters of electronic structures are usually included [18-21]. However, the established 2D-QSAR model surely explains 93.7% variance in anticancer activity, and has a statistical significance and high predictive ability. This model describes a parabolic correlation between pIC_{50} and CLog$P$, suggesting that optimal pIC_{50} values require a suitable range of lipophilicity (around CLog$P$=4.43), that is, a compound with very high or very low hydrophobicity must be disadvantageous to improving the cytotoxicity.

In the present work, the established 3D-model has also good statistical quality and prediction ability: the q^2 and r^2 of this model are 0.619 and 0.993, respectively, and the F value is 193.4 and the standard deviation SD is 0.208. This model indicates that the steric field factor plays a more important role than the electrostatic one, in that the steric field descriptor of the CoMFA model explains 89% of the total contribution, which means that the 3D-QSAR model is in satisfying agreement with the 2D-QSAR model. However, the 3D-QSAR model offers visual images of the steric field and electrostatic field. From such visual steric and electrostatic contour maps, we can see how the substituents (R and R′) affect the anticancer activities of these compounds due to their different steric bulk or the related hydrophobicity. Moreover, we can also see that the electrostatic contour map shows the electrostatic fac-

DOI:10.1088/1674-0068/22/03/285-289

tor plays a very small role because the substituents R and R′ of this kind of compound are all alkyls. These results further confirm the reliability of the 2D-model and offer more insight into understanding the action mechanism.

Based on our present 3D-QSAR model along with the published 2D-QSAR model, the experimentally proposed hydrophobic binding mechanism on the receptor-binding site of the DHFR, can be theoretically understood. That CLog$P$ is the only important factor affecting the bio-activity of this kind of compound can be explained as follows: Since the interaction between the compound as ligand and the bio-receptor happens in receptor-binding pocket of the DHFR, the most important factor playing a decisive role is naturally the ability to enter the receptor-binding pocket. The very low hydrophobicity must result in a smaller steric interaction between the compound and receptor due to the longer distance between one another, although these small molecules can enter the pocket. On the other hand, a very high hydrophobicity may require large aliphatic hydrocarbon substituents (R and R′), and it is the large bulk that results in the large molecules having difficulty entering the pocket. Therefore, the action mechanism of this kind of the compounds can be reasonably explained, based on the 2D and 3D-QSAR studies.

The action mechanisms of anticancer agents are very complicated and most are not understood in a satisfying way. Different anticancer agents may obey different action mechanisms, and moreover, many factors may affect the biological activity. Here, our work further offers an accepted 3D-QSAR on the basis of the published 2D-QSAR, and offers a deeper insight into understanding the action mechanism of this kind of compound with potential anticancer activity.

IV. CONCLUSION

A 3D-QSAR model with good predictive ability for a series of 7,8-dialkyl-1,3-diaminopyrrolo-[3,2-f] quinazolines with anticancer activity was established by using CoMFA method. This model indicates that the steric field factor plays a much more important role than the electrostatic one, in that the steric field descriptor of the CoMFA model explains 89% of the total contribution, in satisfying agreement with the published 2D-QSAR model. However, the 3D-QSAR model offers visual images of the steric field and the electrostatic field. The 3D-QSAR study further suggests the following: to improve the activity, the substituent R′ should be selected to be a group with an adaptive bulk like Et or i-Pr, and the substituent R should be selected to be an alkyl with larger bulk. In particular, based on our present 3D-QSAR as well as the published 2D-QSAR, the experimentally-proposed hydrophobic binding mechanism on the receptor-binding site of the DHFR can be further explained in theory.

V. ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (No.20673148). We also heartily thank the College of Life Sciences, Sun Yat-Sen University for the SYBYL 6.9 computation environment support.

[1] A. H. Hitchings, Angew. Chem. Int. Ed. Engl. **28**, 879 (1989).
[2] C. D. Selassie, W. X. Gan, L. S. Kallander, and T. E. Klein, J. Med. Chem. **41**, 4261 (1998).
[3] J. Jolivet, K. H. Cowan, G. A. Curt, N. J. Clendeninn, and B. A. Chabner, N. Engl. J. Med. **309**, 1094 (1983).
[4] Z. Y. Cui, W. Z. Zhao, and R. L. Li, Prog. Nat. Sci. **12**, 1209 (1999).
[5] M. Tobe, Y. Isobe, H. Tomizawa, T. Nagasaki, F. Obara, and H. Hayashi, Bioorg. Med. Chem. **11**, 609 (2003).
[6] L. F. Kuyper, D. P. Baccanari, M. L. Jones, R. N. Hunter, R. T. Tansik, S. S. Joyner, C. M. Boytos, S. K. Rudolph, V. Knick, H. Robert Wilson, J. Marc Caddell, H. S. Friedman, J. C. W. Comley, and J. N. Stables, J. Med. Chem. **39**, 892 (1996).
[7] J. C. Chen, L. Qian, Y. Shen, L. M. Chen, and K. C. Zheng, Chin. J. Chem. **24**, 1531 (2006).
[8] S. Y. Liao, L. Qian, J. C. Chen, Y. Shen, and K. C. Zheng, J. Theor. Comput. Chem. **7**, 287 (2008).
[9] L. Qian, Y. Shen, J. C. Chen, Y. Y. Wang, X. T. Wu, T. J. Chen, and K. C. Zheng, QSAR Comb. Sci. **27**, 984 (2008).
[10] S. Y. Liao, L. Qian, H. L. Lu, Y. Shen, and K. C. Zheng, QSAR Comb. Sci. **27**, 740 (2008).
[11] J. C. Chen, Y. Shen, L. Qian, L. M. Chen, and K. C. Zheng, Chin. J. Chem. Phys. **20**, 135 (2007).
[12] Y. D. Aher, A. Agrawal, P. V. Bharatam, and P. Garg, J. Mol. Model. **13**, 519 (2007).
[13] S. Janardhan, P. Srivani, and G. N. Sastry, QSAR Comb. Sci. **25**, 860 (2006).
[14] M. Whitlow, A. J. Howard, D. Stewart, K. D. Hardman, L. F. Kuyperi, D. P. Baccanari, M. E. Fling, and R. L. Tansik, J. Biol. Chem. **272**, 30289 (1997).
[15] X. J. Zou, L. H. Lai, and G. Y. Jin, Chin. J. Chem. **23**, 1120 (2005).
[16] M. S. Castilho, M. P. Postigo, C. B. V. de Paula, C. A. Montanari, G. Oliva, and A. D. Andricopulo, Bioorg. Med. Chem. **14**, 516 (2006).
[17] P. J. Goodford, J. Med. Chem. **28**, 849 (1985).
[18] X. H. Li, X. Z. Zhang, X. L. Cheng, X. D. Yang, and Z. L. Zhu, Chin. J. Chem. Phys. **19**, 143 (2006).
[19] W. J. Wu, J. C. Chen, K. C. Zheng, and F. C. Yun, Chin. J. Chem. Phys. **18**, 936 (2005).
[20] W. J. Wu, J. C. Chen, L. Qian, and K.C. Zheng, J. Theor. Comput. Chem. **6**, 223 (2007).
[21] S. Y. Liao, J. C. Chen, L. Qian, Y. Shen, and K. C. Zheng, QSAR Comb. Sci. **27**, 280 (2008).

DOI:10.1088/1674-0068/22/03/285-289