ARTICLE

# 3D-QSAR of Benzothiazole Derivatives as Potent Anticancer Agents

Jin-can Chen, Yong Shen*, Li Qian, Lan-mei Chen, Kang-cheng Zheng

*School of Chemistry and Chemical Engineering, Zhongshan (Sun Yat-Sen) University, Guangzhou 510275, China*

(Dated: Received on June 1, 2006; Accepted on July 17, 2006)

Comparative molecular field analysis (CoMFA) method was applied to study three-dimensional quantitative structure activity relationship (3D-QSAR) of a series of benzothiazole derivatives as potent anticancer agents. The CoMFA model of cross-validation and the partial-least-square (PLS) model of non cross-validation have been well established. The best CoMFA model gives a good cross-validation coefficient of 0.642 and a conventional correlation coefficient of 0.976. Moreover, the estimated standard error is 0.161 and the statistical square deviation ratio $F_{(3,20)}$ is 111.4. The statistical parameters of the best CoMFA model show this model is reasonable and has predictive ability. The CoMFA results suggest that an electron-withdrawing group or atom (e.g. F atom) linking to the first atom ($C_{19}$) of substituent R can increase the positive charges of $C_{19}$ and its $\beta$-site atoms, which lie in the blue-colored regions in the electrostatic field contour map of CoMFA, and thus can improve the activity of the compound. Meanwhile, selecting an R with an appropriate volume is also advantageous for improving the activity.

**Key words:** Benzothiazole, Anticancer, 3D-QSAR, Comparative molecular field analysis

## I. INTRODUCTION

Cancer is a terrible disease which is the leading death of the human population in some areas of the world. It is the second leading cause of death, behind cardiovascular disease, in the United States [1]. At present, there are three main methods of cancer treatment: surgery, radiation therapy, and chemotherapy. With the development of molecular biology, chemotherapy is becoming a more important therapeutic method [2]. Therefore, designing new anticancer drugs with high-efficiency and broad-spectrum activity is a significant study area today.

The quantitative structure-activity relationship method (QSAR), especially the 3D-QSAR method, is a widely accepted technique used in drug design [3,4]. QSAR studies have been extensively used to correlate the molecular structures to their biological activities as a computer design aid [5-9]. Comparative molecular field analysis (CoMFA) [10] is the most broadly-applied method for the study of quantitative structure-activity relationships at the 3D level. Unlike the traditional Hansch analysis relying on substituent parameters, CoMFA relates the biological activity of a series of drugs with their steric and electrostatic fields sampled at grid points defining a large 3D box around the molecule [11]. CoMFA calculates steric and electrostatic properties according to Lennard-Jones and Coulombic potentials, respectively. Then partial least square (PLS) (a regression method) is used to develop the relationship between independent variables (steric and electrostatic potentials) and biological activity. Based on the CoMFA results (model and isocontour maps), the drug's molecular structure can be modified to improve the activity.

In order to develop new anticancer drugs, Yoshida *et al.* have designed and synthesized a series of promising anticancer agents: benzothiazole derivatives [12]. These compounds have overcome the shortcomings of lead compounds (i.e. metabolic instability and unsuccessful anticancer test *in vivo*) and showed a potent cytotoxicity against a tumorigenic cell line, WI-38 VA-13 subline 2RA (VA-13). In this research, 3D-QSAR of the benzothiazole derivatives was studied by using CoMFA method and the factors affecting the anticancer activity were further analyzed on the basis of the obtained reasonable 3D model. The results can offer theoretical reference for designing new anticancer drugs with potent and selective cytotoxicity.

## II. COMPUTATIONAL METHOD

CoMFA studies were performed using SYBYL v6.9 molecular modeling software (Tripos Inc., St. Louis, MO.) running on an SGI R2400 workstation. All parameters used in SYBYL were default except as explained below.

### A. Compounds and biological activities

All 25 benzothiazole derivatives were selected from literature [12]. This series of compounds have the same skeleton (see Fig.1). Anticancer activity data ($EC_{50}$) refers to the milli-molar concentration of the compound required for 50% inhibition of a tumorigenic cell line

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

 ©2007 Chinese Physical Society

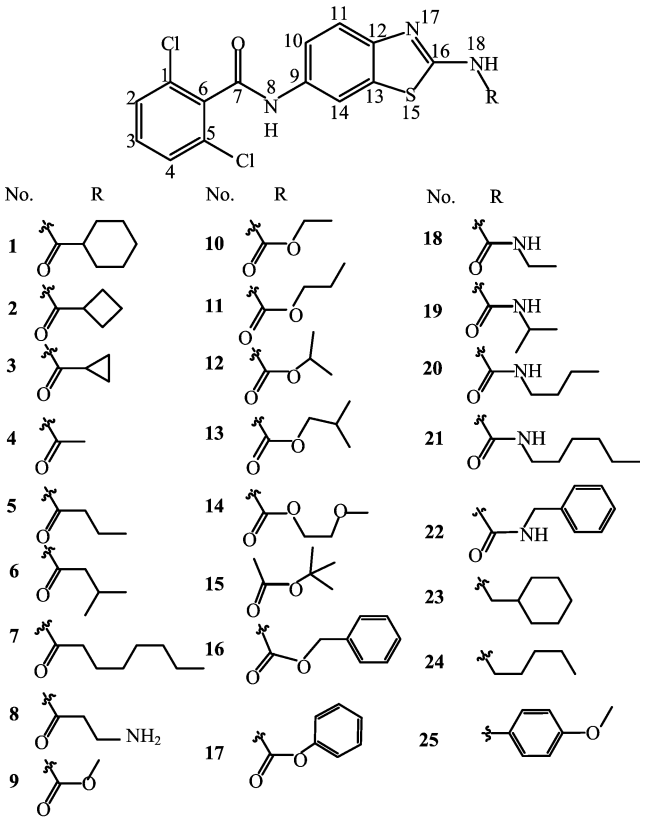


FIG. 1 The molecular structures of benzothiazole derivatives.

WI-38 subline 2RA (VA-13). These values were transformed to $-\lg(EC_{50})$ (negative logarithm of $EC_{50}$) in CoMFA studies.

### B. Conformation determination and molecular alignment

Prior to the construction of a CoMFA model, the determination of the active conformation is the crucial step in 3D-QSAR analyses. In the absence of any experimental knowledge of this bioactive conformation, determining the low-energy conformation is the first step. In this study, the highest active compound **11** was fully optimized using the DFT method [13-15] at B3LYP/6-31G(d,p) level [13,16-17] by use of the Gaussian 98 program-package [18]. Then the models of all the other compounds were obtained from the optimized conformation of compound **11** (see Fig.2) by modifying corresponding atoms or groups. In addition, a conformation search was performed to discover the global energy-minimum conformation by using Chemoffice 8.0 software for the compounds with rotatable bonds.

The fully geometrical optimization of 25 compounds was performed using the standard Tripos force field including the electrostatic term calculated from Gasteiger and Hückel atomic charges. The methods of the steepest descent and conjugate gradients were used for energy minimization and the cut-off value was 0.21 kJ/mol.

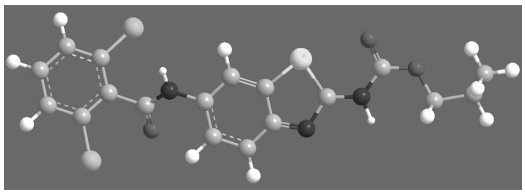

FIG. 2 Conformation of compound **11** (with the highest anticancer activity) fully optimized using the DFT method.

Structural alignment is also a critical step in CoMFA study, since much experience shows that the resulting CoMFA model is often sensitive to the particular alignment scheme [19]. In this study, the lowest energy conformation of compound **11** with the most potent activity was selected as the template structure for the molecular alignment (see Fig.2). The sketch atoms numbered from **1** to **18** in Fig.1 were selected for the alignment of all compounds. The alignment of the bioactive conformations is shown in Fig.3.

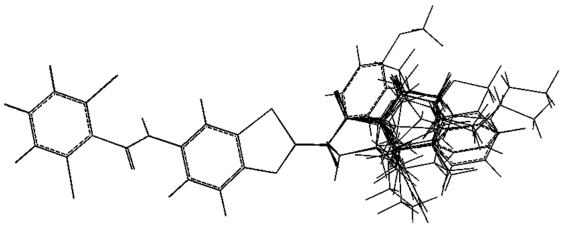

FIG. 3 Superimposition diagram of benzothiazole derivatives.

### C. Analysis of CoMFA and PLS

The CoMFA studies were performed with the QSAR module of Sybyl 6.9. Calculated steric and electrostatic cutoff were set as 124 kJ/mol. The steric (Lennard-Jones) and electrostatic (Coulombic) fields were calculated at each grid point using a $sp^3$ carbon probe with +1 charge, with 0.2 nm grid spacing. The PLS method was used to construct and validate the CoMFA models, and the column filtering value was set as 0. Cross-validation was performed with the leave-one-out procedure. The optional number of components $N$ for final PLS analysis was defined as the one yielding the highest cross-validation coefficient ($q^2$).

## III. RESULTS AND DISCUSSION

### A. CoMFA model

Two models were obtained from the CoMFA analysis, and their statistic parameters of the CoMFA models are listed in Table I. For a reliable predictive model,

ISSN 1003-7713/DOI:10.1360/cjcp2007.20(2).135.5 ©2007 Chinese Physical Society

TABLE I Statistical parameters of two CoMFA models

| CoMFA | N | $q^2$ | $r^2$ | $F$ | $S$ | Contribution/% | | Outlier |
|---|---|---|---|---|---|---|---|---|
| | | | | | | Steric | Electrostatic | |
| Model 1 | 2 | 0.420 | 0.909 | 30.111 | 0.345 | 75.7 | 24.3 | None |
| Model 2 | 3 | 0.642 | 0.976 | 110.448 | 0.161 | 82.1 | 17.9 | **8**, **25** |

N: the optimum number of components; $q^2$: the square of cross-validation correlation coefficient; $r^2$: the square of non-cross-validation correlation coefficient; $S$: the standard error of estimation.

TABLE II CoMFA results of this series of compounds

| No. | $EC_{50}$/(nmol/L) | $-\lg EC_{50}$ | | $\Delta pEC_{50}$ | No. | $EC_{50}$/(nmol/L) | $-\lg EC_{50}$ | | $\Delta pEC_{50}$ |
|---|---|---|---|---|---|---|---|---|---|
| | | Expt. | Calc. | | | | Expt. | Calc. | |
| **1** | 15 | 7.82 | 8.09 | −0.27 | **14** | 11 | 7.96 | 7.95 | 0.01 |
| **2** | 13 | 7.89 | 7.99 | −0.10 | **15** | 27 | 7.57 | 7.56 | 0.01 |
| **3** | 13 | 7.89 | 7.75 | 0.14 | **16** | 180 | 6.74 | 7.00 | −0.26 |
| **4** | 600 | 6.22 | 6.36 | −0.14 | **17** | 1100 | 5.96 | 6.00 | −0.04 |
| **5** | 10 | 8.00 | 7.65 | 0.35 | **18** | 32 | 7.49 | 7.55 | −0.06 |
| **6** | 12 | 7.92 | 7.90 | 0.02 | **19** | 30 | 7.52 | 7.39 | 0.13 |
| **7** | 180 | 6.74 | 6.82 | −0.08 | **20** | 28 | 7.55 | 7.58 | −0.03 |
| **8** | 5300 | 5.28 | – | – | **21** | 290 | 6.54 | 6.49 | 0.05 |
| **9** | 9.5 | 8.02 | 7.92 | 0.10 | **22** | 150 | 6.82 | 6.67 | 0.15 |
| **10** | 2.6 | 8.59 | 8.69 | −0.10 | **23** | 4300 | 5.37 | 5.37 | 0.00 |
| **11** | 2.5 | 8.60 | 8.48 | 0.12 | **24** | 270 | 6.57 | 6.51 | 0.06 |
| **12** | 2.6 | 8.59 | 8.56 | 0.03 | **25** | 870 | 6.06 | – | – |
| **13** | 3.5 | 8.46 | 8.56 | −0.10 | | | | | |

the square of cross-validation coefficient $q^2$ should be $\geq$0.5. In Table I, $q^2$ of Model 1 (obtained for all 25 compounds) is 0.450 (<0.5), so Model 1 is not a satisfying predictive model. Analyzing the relationship between structures and activities of compounds in Fig.1 as well as the alignment of the bioactive conformations (Fig.3), we can clearly find that compound **8** has a similar structure to compound **5**, **6** and **7** but shows the lowest activity. Compound **25** shows the greatest deviation in molecular alignments, so compounds **8** and **25** should be considered as two outliers. After omitting compound **8** and **25**, Model 2 with significant statistical quality and predictive ability is established, and its $q^2$=0.642, $r^2$=0.976, $SD$=0.161, $F_{(3,20)}$=111.4. Therefore, the following discussion is carried out based on Model 2. The experimental and calculated activities as well as residual values ($\Delta pEC_{50}$) of the title compounds are listed in Table II.

The steric field descriptor of the best CoMFA model explains 82.1% of the total variance, while the electrostatic descriptor explains the remaining 17.9%. This indicates that the steric field is more predominant.

### B. 3D contour maps of CoMFA

The steric contour map is displayed in Fig.4. The green-colored region indicates an area where the steric bulk of the compound can increase its biological activity, whereas the yellow-colored region indicates the area where the steric bulk is detrimental to the biological activity. There are three yellow contours around substituent R, which suggests that small substituents in size in these positions will be advantageous for improving the biological activity. For example, compound **5**, **6**, **10**-**14** with smaller substituents in the yellow positions show high activity. In contrast, the substituents R of compound **17** and **23** are bigger and lie in the yellow positions, hence they show lower activity. In addition, the two green contours near R indicate that steric bulky substituents in these positions will significantly improve the activity. For instance, compounds **1**-**3** with bulky groups in these positions have very potent activities ($-\lg EC_{50}$=7.82-7.89); however, compound **4**

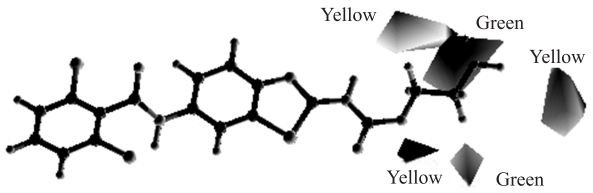


FIG. 4 Steric field contour map of CoMFA.

ISSN 1003-7713/DOI:10.1360/cjcp2007.20(2).135.5 ©2007 Chinese Physical Society

with small substituent $-COCH_3$ does not reach the green positions, hence, the activity of compound **4** ($-\lg EC_{50}$=6.22) is significantly lower than the activities of compounds **1-3**. In the same way, comparing the structure of compound **9** with the structures of **10-13**, in which methyl, ethyl, dimethyl and isopropyl are added onto the end of R chain respectively, compounds **10-13** show higher activities ($-\lg EC_{50}$=8.59-8.46) than that of compound **9** ($-\lg EC_{50}$=8.02). The analysis of the steric field contour map of CoMFA (Fig.4) suggests that compounds with high activities must select a moderate-size R.

The electrostatic contour map is displayed in Fig.5. Blue-colored regions show the areas where electropositive-charged groups enhance the biological activity of the compound, whereas red regions represent the areas where electronegative-charged groups improve the activity. In Fig.5, there is a large region of blue polyhedral surface covering over the end of substituent R, which indicates that more positive charge groups (or atoms) on this surface will significantly improve the biological activity. We find that compounds **9-15** have substituents R with –NCOO– and compounds **18-20** have substituents R with –NCON–. Since the electronegativity of N and O atoms is greater than that of C, according to the law of polarity alternation [20-21] and the idea of polarity interference [22-25], the schematic maps of polarity interference are drawn as shown in Fig.6. From Fig.6, we can clearly see that either the real-line arrowheads (expressing the polarity direction of direct bonds) or the dashed-line arrowheads (expressing the polarity direction of induced bonds) all depart from $C_{19}$. This must lead the $C_{19}$ atom to carry more positive charges, and lead to an increase in the positive charges of atoms at $\beta$-sites of $C_{19}$, i.e. the atoms binding to $O_{21}$ and $N_{21}$. Since these atoms at $\beta$-sites of $C_{19}$ just fall into the blue-colored position, these compounds certainly have high activities. On the contrary, since compounds **23** and **24** do not have substituents R with the above two groups, the positive charges of the first atom ($C_{19}$) of R are surely smaller, and thus the positive charges of the atoms at $\beta$-sites of $C_{19}$ do not increase. That is why the activities of compounds **23** and **24** are so low.

Analysis of the electrostatic field contour map of CoMFA shown in Fig.5 suggests that introducing an electron-withdrawing atom or group to the first atom

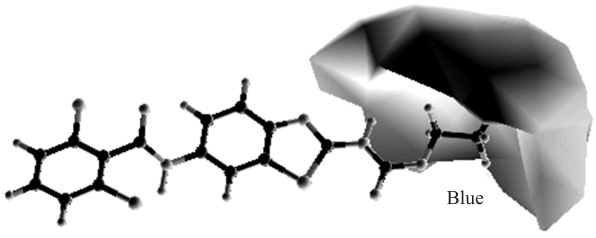


FIG. 5 Electrostatic field contour map of CoMFA

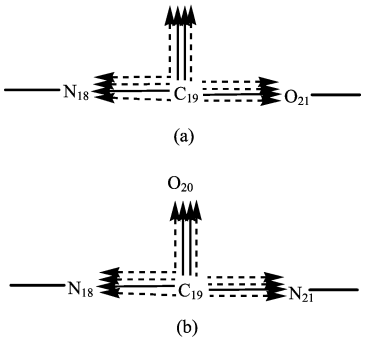


FIG. 6 Schematic maps of polarity interference for R of compound **9-15** (a) and **18-20** (b).

($C_{19}$) of R (e.g. F atom) will increase the positive charges of $C_{19}$, and thus the charges of atoms at the $\beta$- site of $C_{19}$ will also increase. As a result, the activity of the compound will be improved since these atoms at the $\beta$- site of $C_{19}$ fall only into the blue area.

## IV. CONCLUSION

Adopting the conformation of compound **11** fully optimized by the DFT method as a starting template and going through the structural optimization by using the standard Tripos force field, a CoMFA model with a good predictive ability was established. The analysis of the CoMFA model suggests that introducing electron-withdrawing atoms or groups at the first atom ($C_{19}$) of substituent R can lead the positive charges of $C_{19}$ to increase, and thus lead the positive charges of atoms at $\beta$-site of $C_{19}$ to increase. As a result, the activity of the compound will be improved since these atoms at $\beta$-site of $C_{19}$ fall only into the blue area. Moreover, selecting an R with an appropriate volume to fall into the green positions can also improve the activity. Therefore, the obtained 3D-QSAR model can offer a useful guide for the molecular design of new drugs with higher activity and help us understand the mechanism of the action of drugs. As the next step in our research, we will focus on the design of the new drugs with higher activity according to this model and the law of polarity alternation [20-21] and the idea of polarity interference [22-25] which can guide us to design proper substituent atoms with proper charge to fit the electrostatic contour map.

## V. ACKNOWLEDGMENTS

This work was supported by the Natural Science Foundation of Guangdong Province of China (No. 05003264) and the Natioal Natural Science Foundation of China (No.90608012). The computation environment support for SYBYL 6.9 of Department of Biochemistry, College of Life Sciences, Sun Yat-Sen University, are gratefully acknowledged.

ISSN 1003-7713/DOI:10.1360/cjcp2007.20(2).135.5 ©2007 Chinese Physical Society

[1] J. B. Gibbs, Science **287**, 1969 (2000).
[2] Z. M. Wang, L. N. Ji, Prog. Chem. **14**, 296 (2002).
[3] X. M. Luo, H. L. Jiang, J. H. Shen and K. X. Chen, Chin. Acad. Sci. **4**, 255 (2003).
[4] R. L. Li, *Drug Structure-Activity Relationship*, 1st Ed. Beijing: Chinese Medicine Science and Technology Press, 477 (2004).
[5] X. H. Li, X. Z. Zhang, X. L. Cheng, X. D. Yang and Z. L. Zhu, Chin. J. Chem. Phys. **19**, 143 (2006).
[6] W. J. Wu, J. C. Chen, K.C. Zheng and F. C. Yun, Chin. J. Chem. Phys. **18**, 936 (2005).
[7] W. J. Wu, R. Lai, K. C. Zheng and F. C. Yun, Acta. Phys. Chim. Sin. **21**, 28 (2005).
[8] X. F. Han, Y. Liu, Y. Gao and L. H. Lai, Acta. Chim. Sin. **61**, 1136 (2003).
[9] J. F. Xiao, Z. R. Guo, Y. S. Guo, F. M. Zhu and P. Y. Sun, Acta. Chim. Sin. **63**, 757 (2005).
[10] R. D. Cramer, D. E. Patterson and J. D. Bunce, J. Am. Chem. Soc. **110**, 5959 (1988).
[11] X. J. Zou, L. H. Lai and G. Y. Jin, Chin. J. Chem. **23**, 1120 (2005).
[12] M. Yoshida, I. Hayakawa, N. Hayashi, T. Agatsuma, Y. Oda, F. Tanzawa, S. Iwasaki, K. Koyama, H. Furukawa, S. Kurakatad and Y. Sugano, Bioorg. Med. Chem. Lett. **15**, 3328 (2005).
[13] A. D. Becke, J. Chem. Phys. **98**, 1372 (1993).
[14] A. Görling, Phys. Rev. A **54**, 3912 (1996).
[15] J. B. Foresman, Æ. Frisch, *in Exploring Chemistry with Electronic Structure Methods*, 2nd Edn., Pittsburgh, PA: Gaussian Inc, (1996).
[16] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[17] G. A. Petersson, A. Bennett, T. G. Tensfeldt, M. A. Al-Laham, W. A. Shirley and J. Mantzaris, J. Chem. Phys. **89**, 2193 (1988).
[18] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr. R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, N. Rega, P. Salvador, J. J. Dannenberg, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle and J. A. Pople, *Gaussian 98 Revision A.*,11.4, Pittsburgh, PA: Gaussian, Inc., (2002).
[19] M. S. Castilho, M. P. Postigo, C. B. V. de Paula, C. A. Montanari, G. Olivaa and A. D. Andricopulo, Bioorg. Med. Chem. **14**, 516 (2006).
[20] J. A. Pople and M. S. Gordon, J. Am. Chem. Soc. **89**, 4253 (1967).
[21] S. Fliszar, J. Am. Chem. Soc. **94**, 1068 (1972).
[22] K. C. Zheng, F. He, Z. T. Xu and F. C. Yun, Acta Phys. Chim. Sin. **15**, 698 (1999).
[23] K. C. Zheng, J. P. Wang, W. L. Peng, X. W. Liu and F. C. Yun, J. Phys. Chem. A **105**, 10899 (2001).
[24] K. C. Zheng, J. P. Wang, Y. Shen, W. L. Peng and F. C. Yun, J. Comput. Chem. **23**, 4369 (2002).
[25] S. Shi, J. Liu, J. Li, K. C. Zheng, C. P. Tan, L. M. Chen and L. N. Ji, J. Chem. Soc. Dalton. Trans. **11**, 2038 (2005).

ISSN 1003-7713/DOI:10.1360/cjcp2007.20(2).135.5 ©2007 Chinese Physical Society