Pergamon

Bioorganic & Medicinal Chemistry Letters 11 (2001) 2041–2044

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

CoMFA of Artemisinin Derivatives: Effect of Location and Size of Lattice

Mankil Jung* and Hanjo Kim

Department of Chemistry, Yonsei University, Seoul 120-749, South Korea

Received 26 February 2001; accepted 21 May 2001

Abstract—A CoMFA study of artemisinin derivatives with changes of the location and the number of lattice points was performed. The location of probe atoms in a large lattice has practically no effect on the cross-validated r^2 value (r^2_{cv}). The selection of only 18 probe atoms around the peroxide bond, considering the action mechanism of artemisinin, provided a less time-demanding and more reliable CoMFA model, which forecasts better than the large lattice model despite the lower r^2_{cv} value. Only 1 Å displacement of the small lattice caused a reduction of cross-validated r^2 value of more than 50%, which indicates the lattice location played an important role in this small lattice model.

Introduction

The naturally occurring endoperoxide sesquiterpene, artemisinin,[1] is the most promising drug[2] for treatment of malaria, an infectious disease caused by a protozoan parasite, *Plasmodium*. It is estimated that between 250 and 500 million people have malaria, with two or three million of those dying each year.[3] The most malignant of four related protozoa is *Plasmodium falciparum*. Artemisinin has been intensely studied because it has shown low toxicity and high activity against drug-resistant strains of *P. falciparum*. In addition, artemisinin or its derivatives are the most rapidly acting of all antimalarial drugs (Fig. 1).[2] Since 1972, the first isolation from *Artemisia annua*, many derivatives of artemisinin have been synthesized. The synthetic efforts have been focused on the modification of C-10. The second-generation artemisinin derivatives are the derivatives of dihydroartemisinin, the lactol form of artemisinin.[4] Deoxoartemisinin derivatives that have no C–O bond at C-10 are more stable in acidic conditions[5] and usually more active than dihydroartemisinin derivatives.[6] In addition, other derivatives having 1,2,4-trioxane moieties have been synthesized.[7]

Comparative Molecular Field Analysis (CoMFA) developed by Cramer III[8] can be used when drug receptor information is unavailable as in this case. It is known that the CoMFA contours from different locations of the lattice points appear different even if the statistics of the CoMFA correlations are very similar and the locations can also affect the number of latent variables in the final CoMFA model.[9] The CoMFA model will be better fitted when the probe atom is at the position that better mimics the locations of similar atoms in the target macromolecule. Several CoMFA studies on the artemisinin derivatives were reported before,[10] but the practical aspects of CoMFA were not studied. We report the effect of location and size of lattice on CoMFA of artemisinin derivatives as an example of the practical aspects in this letter.

CoMFA study

A dataset of 124 compounds including deoxoartemisinin (their observed activities) has been taken

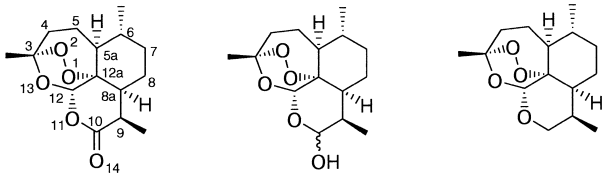

Figure 1. Structure and the numbering system of artemisinin and its derivatives.

*Corresponding author. Tel.: +82-2-2123-2648; fax: +82-2-364-7050; e-mail: mkjung@alchemy.yonsei.ac.kr

0960-894X/01/$ - see front matter
PII: S0960-894X(01)00361-4

from published results.[11] The log value of the relative activity (RA) of these compounds was used and defined as

$$\text{Log (RA)} = \log [(\text{artemisinin } IC_{50}/\text{analogue } IC_{50})]$$

So, the compounds more active than artemisinin have a positive sign of log (RA) value. Log (RA) of all the desoxy derivatives are assigned to −4 (10,000 times less active)[12] because the peroxide bond is essential for the antimalarial activities of artemisinin derivatives. Molecular models of artemisinin derivatives were built and fully minimized by the Tripos force field. Atomic charges were calculated using the Gasteiger–Hückel protocol in Sybyl 6.3.[13] The molecules were aligned to minimize the root-mean-square (rms) coordinate differences of 1,2,4-trioxane ring atoms between analogue molecules and template (deoxoartemisinin in here). Desoxy compounds that have an ether bond instead of a peroxide bond were aligned to minimize the rms difference of C4–O3–C5–C6.

The Effect of Lattice Location

We derived the initial CoMFA model, using all probes of interest, at 2 Å spacing with 13 different lattice locations as shown in Table 1. The CoMFA12 shows the best statistics and is used for more detailed studies. The difference of highest (CoMFA12—0.805) and lowest (CoMFA9—0.740) value of r^2_{cv} is 0.065 (about 8% of highest value).

An initial CoMFA run using 10 components and leave-one-out cross-validation with SAMPLS indicated that the optimal number of components was 3, as the addition of one more component did not increase 10% of the initial r^2_{cv}.

Figures 2 and 3 show the steric and electrostatic contour map of CoMFA12. A yellow contour and a red contour are placed around the peroxide bond. The yellow contour behind the peroxide bond may remind that the antimalarial activity of C-10β compounds is generally higher than that of C-10α compounds.

Table 1. Effects of changing the lattice offset on CoMFA models of artemisinin derivatives

CoMFA model	Lattice offset (x, y, z)	r^2_{cv}	Standard error	Number of components
1	(0, 0, 0)	0.775	0.757	3
2	(−0.5, 0, 0)	0.778	0.752	3
3	(0, −0.5, 0)	0.792	0.728	3
4	(0, 0 ,−0.5)	0.783	0.743	3
5	(−0.5, −0.5, 0)	0.783	0.744	3
6	(−0.5, 0,−0.5)	0.749	0.801	3
7	(0, −0.5, −0.5)	0.784	0.742	3
8	(0.5, 0, 0)	0.766	0.773	3
9	(0, 0.5, 0)	0.740	0.814	3
10	(0, 0, 0.5)	0.803	0.709	3
11	(0.5, 0.5, 0)	0.771	0.765	3
12	(0.5, 0, 0.5)	0.805	0.705	3
13	(0, 0.5, 0.5)	0.798	0.719	3

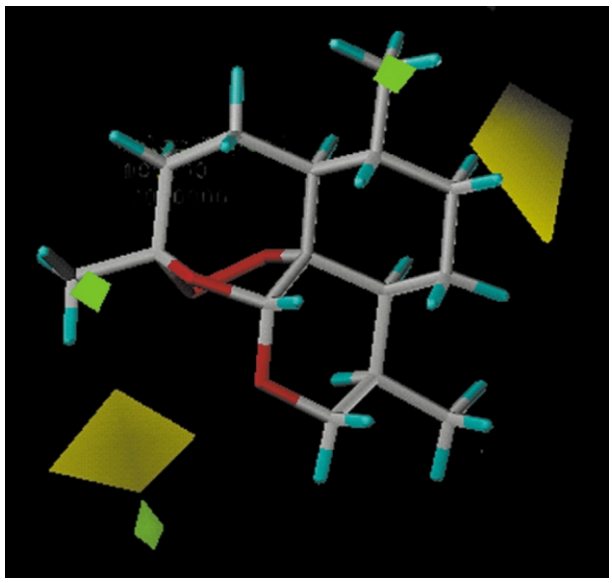

Figure 2. The CoMFA contour map: Steric map indicating areas where steric bulk is predicted to increase (green) or decrease (yellow) activity.

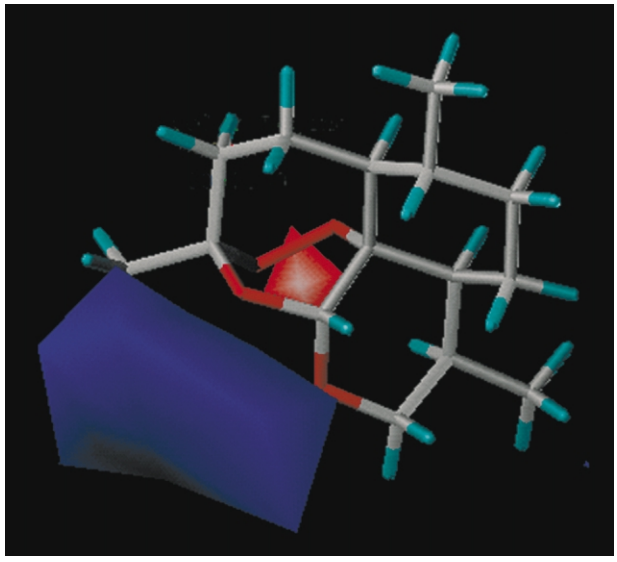

Figure 3. The CoMFA contour map: Electrostatic map indicating red contours around regions where high electron density (negative charge) is expected to increase activity, and blue contours represent areas where low electron density (partial positive charge) is expected to increase activity.

The Effect of Lattice Size

In the bio-system, the peroxide bond of artemisinin is broken to generate the carbon-centered radical essential to the antimalarial activity.[14] According to many experimental proofs,[15] it may be the Fe^{2+} ion or heme that breaks this peroxide bond. Thus, the chemical circumstances of the peroxide bond in artemisinin derivatives are important in the in vitro system while the stability and the solubility of the compound may be more important in the in vivo systems. Therefore, there should be high relevancy between the steric and electrostatic

field around the peroxide bond and the antimalarial activity. The structural deviation in artemisinin derivatives changes the chemical circumstance of the peroxide bond, which alters the activity.

The decrease of number of lattice points increases the chance of finding a CoMFA model and makes the calculations run faster. Moreover, the sensitivity of PLS is diluted by including a large number of descriptors with small signal-to-noise ratios.[9] Therefore, we were tempted to improve this CoMFA model by eliminating descriptors that have small correlations with antimalarial activity and performed the CoMFA with a far smaller region containing only 18 probe atoms around the peroxide bond. The statistical indices of resulting CoMFA models (CoMFA14 and CoMFA15) are listed in Table 2. The r^2_{cv} of CoMFA14 is 0.653, which is 81% of one of CoMFA12. The energy fields from these 18 probe atoms contribute more than 80% of the information for the CoMFA model.

Table 2. Effects of changing the lattice offset on CoMFA models of artemisinin derivatives

CoMFA model	Lattice offset (x, y, z)	r^2_{cv}	Standard error	Number of components
14	(0, 0, 0)	0.653	0.941	3
15	(1, 0, 0)	0.282	1.348	2

Table 3. The forecasting capability of CoMFA models

Compound[a]	Observed activity [log (RA)]	Forecast of CoMFA12	Forecast of CoMFA14
4a[b]	0.372	0.549	−0.053
4b[b]	0.176	0.537	−0.053
4c[b]	0.104	−0.568	1.261
4d[b]	0.041	0.814	−0.769
5a[b]	0.85	1.249	0.25
5b[b]	0.28	1.182	0.193
5c[b]	0.061	1.206	0.188
5d[b]	−0.004	1.694	0.778
6[b]	0.288	0.866	0.03
7a[b]	0.333	−0.031	−0.446
7b[b]	−0.209	0.082	−0.884
7c[b]	0.023	−0.140	−0.463
7d[b]	0.034	−1.065	−0.357
8[b]	0.394	0.284	−0.665
9a[b]	−0.046	−0.029	−0.97
9b[b]	0.077	−0.063	−0.963
9c[b]	0.071	−0.069	−1.003
5a alpha[c]	−1.078	−0.269	−1.009
5a beta[c]	−0.616	−1.475	−1.743
5b alpha[c]	−0.917	−0.364	−0.988
5b beta[c]	−0.869	−1.551	−1.721
5c alpha[c]	−1.267	−0.453	−1.055
5c beta[c]	−0.68	−1.684	−2.037
5d alpha[c]	−0.726	−0.401	−1.044
5d beta[c]	−0.777	−1.574	−1.753
5f alpha[c]	−1.814	0.073	−0.515
5g alpha[c]	−0.928	−0.421	−1.002
5g beta[c]	−0.212	−1.288	−1.68
5h alpha[c]	−0.627	−0.434	−0.994
5h beta[c]	−0.744	−1.275	−1.674
5i alpha[c]	−0.934	−0.582	−1.032
5j alpha[c]	−0.68	−0.728	−1.07
5j beta[c]	−0.337	−1.491	−1.712
5k alpha[c]	−0.66	−0.002	−0.995
5k beta[c]	−0.398	−1.571	−1.765
5l alpha[c]	−0.849	−0.361	−1.07
5l beta[c]	−0.513	−1.533	−1.753
5m alpha[c]	−1.032	−1.083	−1.518
5m beta[c]	−0.568	−2.192	−1.967
5n alpha[c]	−0.627	−1.072	−1.197
5n beta[c]	−0.761	−2.176	−1.859
Dihydroartemisinin	0.55	−0.543	−0.659

[a]The numbers of compounds stand for the same numbers in the reference. Alpha and beta reveals the stereochemistry at C-3 position of compounds.
[b]Antimalarial activity data taken from ref 16.
[c]Antimalarial activity data taken from ref 17.

In the case of this CoMFA model of a far smaller number of descriptors, the locations of the probe atoms alters the correlation greatly. When we used far more descriptors (in Table 1), the locations of probe atoms have little effect on the r^2_{cv} value as stated before but, in these models (Table 2), the locations have much more effect such that only 1 Å deviation of region gives cross-validated r^2 0.282, only about a half of that of CoMFA14. The removed descriptors have low correlation with antimalarial activity, but they play a role of buffer, which decreases the fluctuation of r^2_{cv} according to the changes of locations of probe atoms. So, when we use probe atoms selected in correlation of activity, we should choose the location of the lattice more carefully.

Table 3 shows the forecasting capability of these CoMFA models, CoMFA12 and CoMFA14. This test set includes deoxoartemisinin derivatives that have an aryl group at C-10 of artemisinin and 1,2,4-trioxane derivatives that have an aryl group at the C-3 position. All of these compounds were synthesized and the antimalarial activities toward the chloroquine-sensitive NF54 parasite were determined by Posner's group.[16,17] CoMFA14 predicts slightly better than CoMFA12 even though its cross-validated r^2 value is lower than those of CoMFA12.

Conclusion

The location and the size of lattice alter the final CoMFA model. When we used the region that can accommodate all the compounds, the location of the probe atoms had practically no effect on the statistical indices. All of the cross-validated r^2 values of 13 CoMFA models were higher than 0.740. We selected the probe atoms around the peroxide bond and performed CoMFA with this reduced region. The energy fields of only 18 probe atoms could give enough information on the QSAR. The location of lattice played a more important role and should be chosen carefully. The 1 Å deviation of the lattice lowered the r^2_{cv} more than 50%. Though the r^2_{cv} value was lower than the previous model (CoMFA12), CoMFA14 had slightly better forecasting capability. Therefore, we could get a more reliable CoMFA model using a small lattice containing only 18 probe atoms instead of large lattice that have generally several hundreds of probe atoms. These results show that the mechanism-based selection of lattice points can reduce the calculation time without reduction of the forecasting ability of CoMFA model.

Acknowledgements

This work was supported by Korea Research Foundation Grant (KRF-2000-A1004-DP0267). Special thanks are due to Dr. Gary Posner et al. for making it possible for us to use the antimalarial activity data in Table 3 in this paper taken from refs 16 and 17 published by them.

References and Notes

1. Klayman, D. L. *Science* **1985**, *223*, 1049.
2. White, N. J. *Trans. R. Soc. Trop. Med. Hyg.* **1994**, *88* (Suppl. 1), 3.
3. Vroman, J. A.; Alvim-Gaston, M.; Avery, M. A. *Curr. Phar. Des.* **1999**, *5*, 101.
4. Lin, A. J.; Lee, M.; Klayman, D. L. *J. Med. Chem.* **1989**, *32*, 1249.
5. Jung, M.; Lee, S. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1003.
6. Jung, M. *Curr. Med. Chem.* **1994**, *1*, 35.
7. Jefford, C. W.; Velarde, J.; Bernardnelli, G. *Tetrahedron Lett.* **1989**, *30*, 4485.
8. Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
9. Martin, Y. C.; Kim, K.-H.; Lin, C. T. In *Advances in Quantitative Structure–Property Relationships*; Charton, M., Ed.; JAI: Greenwich, 1996; pp 2–52.
10. (a) Grigorov, M.; Weber, J.; Tronchet, J. M. J.; Jefford, C. W.; Milhous, W. K.; Maric, D. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 124. (b) Avery, M. A.; Gao, F.; Chong, W. K. M.; Mehrotra, S.; Milhous, W. K. *J. Med. Chem.* **1993**, *36*, 4264. (c) Tonmunphean, S.; Kokpol, S.; Parasuk, V.; Wolschann, P.; Winger, R. H.; Liedl, K. R.; Rode, B. M. *J. Comput. Aided Mol. Des.* **1998**, *12*, 397.
11. Antimalarial activity: (a) ref 5. (b) Jung, M.; Freitas, A. C. C.; McChesney, J. D.; ElSohly, H. N. *Heterocycles* **1994**, *39*, 23. (c) Pu, Y. M.; Ziffer, H. *J. Med. Chem.* **1995**, *38*, 613. (d) Avery, M. A.; Fan, P.; Karle, J. M.; Bonk, J. D.; Miller, R.; Goins, D. K. *J. Med. Chem.* **1996**, *39*, 1885. (e) Posner, G. H.; Park, S. B.; Gonzàlez, L.; Wang, D.; Cumming, J. N.; Klinedindst, D.; Shapiro, T. A.; Bachi, M. D. *J. Am. Chem. Soc.* **1996**, *118*, 3537. (f) Haynes, R. K.; Vonwiller, S. C.; Wang, H. J. *Tetrahedron Lett.* **1995**, *36*, 4641. (g) Torok, D. S.; Ziffer, H.; Meshinick, S. R.; Pan, X. Q.; Ager, A. *J. Med. Chem.* **1995**, *38*, 5045. (h) Kamchonwongpaisan, S.; Paitayatat, S.; Thebtaranonth, Y.; Wilairat, P.; Yuthavong, Y. *J. Med. Chem.* **1995**, *38*, 2311. (i) Jung, M.; Li, X.; Bustos, D. A.; ElSohly, H. N.; McChesney, J. D.; Milhous, W. *J. Med. Chem.* **1990**, *33*, 1516. (j) Jung, M.; Li, X.; Bustos, D. A.; ElSohly, H. N.; McChesney, J. D.; Milhous, W. *Heterocycles* **1989**, *29*, 2273. (k) Venugopulan, B.; Baput, C. P.; Karnik, P. J.; Chartterjee, D. K.; Iyer, N.; Lepcha, D. *J. Med. Chem.* **1995**, *38*, 1922. (l) Pu, Y. M.; Torok, D. S.; Ziffer, H.; Pan, X. Q.; Meshinick, S. R. *J. Med. Chem.* **1995**, *38*, 4120. (m) Posner, G. H.; McGarvey, D. J.; Oh, C. H.; Kumar, N.; Meshnick, S. R.; Asawamahasadka, W. *J. Med. Chem.* **1995**, *38*, 607. (n) Jefford, C. W.; Velarda, J. A.; Bernardinelli, G.; Bray, D. H.; Warhurst, D. C.; Milhous, W. K. *Helv. Chim. Acta* **1993**, *76*, 2775. (o) Yang, Y. H.; Li, Y.; Shi, Y. L.; Yang, J. D.; Wu, B. A. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 1791.
12. Woolfrey, J. R.; Avery, M. A.; Doweyko, A. M. *J. Comput. Aided Mol. Des.* **1998**, *12*, 165.
13. *SYBYL 6.3*; Tripos Associates: St. Louis, MO, USA.
14. Posner, G. H.; Oh, C. H. *J. Am. Chem. Soc.* **1992**, *114*, 8328.
15. Robert, A.; Meunier, B. *J. Am. Chem. Soc.* **1997**, *119*, 5968.
16. Posner, G. H.; Parker, M. H.; Northrop, J.; Elias, J. S.; Ploypradith, P.; Xie, S.; Shapiro, T. A. *J. Med. Chem.* **1999**, *42*, 300.
17. Posner, G. H.; Cumming, J. N.; Woo, S.-H.; Ploypradith, P.; Xie, S.; Shapiro, T. A. *J. Med. Chem.* **1998**, *41*, 940.