ELSEVIER

QSAR analyses on ginkgolides and their analogues using CoMFA, CoMSIA, and HQSAR

Weiliang Zhu,[a,*] Gang Chen,[b] Lihong Hu,[c] Xiaomin Luo,[a] Chunshan Gui,[a,b] Cheng Luo,[a] Chum Mok Puah,[b] Kaixian Chen[a] and Hualiang Jiang[a,*]

[a]Drug Discovery and Design Centre, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica,
Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 555 Zuchongzhi Road, Pudong, Shanghai 201203, PR China
[b]Technology Centre for Life Sciences, Singapore Polytechnic, 500 Dover Road, Singapore 139651
[c]The National Center for Drug Screening, 189, Guo Shou Jing Road, Shanghai 201203, PR China

Abstract—Ginkgolides, isolated from ginkgo balba leaves, were found to be powerful as natural antagonists of human platelet activating factor (PAF) in treatment of some diseases such as acute inflammation, tissue rejection, asthma, and ischemic injury. Ginkgolides have a cage skeleton consisting of six five-membered rings, therefore, are very tough to be synthesized. For finding new powerful substitutes of the natural ginkgolides for treating those diseases, three methods, viz. CoMFA, CoMSIA, and HQSAR, were used to investigate the relationship between 117 ginkgolide analogues with great structural diversity and their bioactivities against PAF receptor. The high q^2 released from the different QSAR methods, ranging from 0.583 to 0.684, suggests that three rational and predictive QSAR models were successfully built. These models also show clearly how steric, electrostatic, hydrophobicity, and individual atom affect molecular bioactivity as antagonists of PAF. These results could also be used to account for the unusually higher bioactivity of ginkgolide B than other ginkgolides. The possible binding mechanism between ginkgolides and human PAF receptor was also deduced based on the QSAR models. Therefore, this study should be very helpful in discovering new drugs as PAF antagonists in fighting against various diseases related to PAF and PAF receptor.
© 2004 Elsevier Ltd. All rights reserved.

1. Introduction

Ginkgo is a kind of traditional Chinese herb, which has been used for disease treatment for more than 2,500 years in China.[1] Pharmacological studies indicated that ginkgolides, isolated from ginkgo leaves, are potent antagonists of platelet activating factor (PAF, 1-*O*-alkyl-2-acetyl-*sn*-glycero-3-phosphorylcholine). PAF is a bioregulator that appears to play a key role in many diseases, such as acute inflammation, tissue rejection, asthma, and ischemic injury.[2,3] Furthermore, PAF has also been suggested to play roles in disease of aging and in initiating human immunodeficiency virus (HIV)-related neuropathogenesis.[4,5] Therefore, the study and application of ginkgolides hold great appeal to medical chem-

Keywords: QSAR; CoMFA; CoMSIA; HQSAR; Ginkgolide; PAF; PAFR; Antagonist.
* Corresponding authors. Tel.: +86 21 50807188; fax: +86 21 50807088; e-mail: hljiang@mail.shcnc.ac.cn

ists. Experimental results showed that ginkgolide B (BN 52021, Fig. 1) is the most potent PAF antagonist among all isolated ginkgolides, viz., ginkgolide A, B, C, J, and M. However, ginkgolides B (GB, hereinafter) is present in ginkgo leaves in minute amount and its absorption of oral administration is poor. In the laboratory, it was tough work to synthesize a structure with six condensed five-membered rings that is the skeleton of all
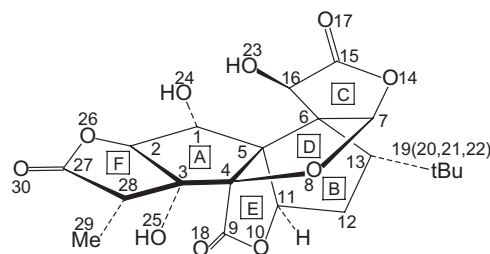

Figure 1. The chemical structure of ginkgolide B.

ginkgolides. Hence, it is of importance to develop new ginkgolide analogues that are not only as active as the natural ginkgolides, but easier to be synthesized as well.

In our previous study, we had performed a CoMFA[6] investigation on 25 ginkgolides and analogues for discovering more powerful PAF antagonists.[7] Based on that CoMFA result, three compounds were designed and synthesized. Indeed, pharmacological assay had demonstrated that the designed compounds were more active than ginkgolide B. However, the structures of those three compounds are still complicated, leading to little application of these structures in pharmaceuticals. Thereby, to simplify the condensed ring structure becomes a key step toward discovering new applicable PAF antagonists. Obviously, our earlier CoMFA model was not competent in this task, because it was derived from ginkgolides and structurally very similar analogues. On the other hand, we could not explain why ginkgolide B is much more powerful than other ginkgolides based on that QSAR model. In order to build predictive QSAR models for the further development of new PAF antagonists, more than 60 new compounds, which have obvious structural diversity with 3–6 condensed rings, have been synthesized.[8,9] Some other analogues have also been reported by other researchers.[10,11] Bioassay results revealed that the bioactivities of the newly synthesized analogues cover a great range.[8–11] The activity was expressed as IC_{50} value. It is the concentration required to inhibit platelet (3.5×10^5 cells/mm^3) aggregatory response by 50%, which was induced by C-18-PAF (1.5×10^{-8} M).[9] Hence, it is the right time now to construct new QSAR models to get the further insight into the relationship between molecular structures and bioactivities, for designing new PAF antagonists with simpler structure. In this paper, we report our new QSAR models, including CoMFA, CoMSIA,[12] and HQSAR,[13] based on 117 ginkgolides and their analogues. The objective of the CoMFA is to discover highly predictive models based on these 117 compounds. The main aim of CoMSIA study is to study the effect of molecular hydrophobic and hydrophilic fields on molecular biological activity. The purpose of the HQSAR study is to explore individual atomic contribution to molecular bioactivity with visual display of active centers in a compound. Furthermore, HQSAR result can be also somewhat used as control to assess our CoMFA and CoMSIA results for HQSAR technique does not require molecular superposition. Therefore, we believe these three QSAR models would provide some new useful information for designing new PAF antagonists with simpler structure. Based on the new QSAR models, we also deduce the possible binding mechanism between PAF receptor and its ligands.

2. Computational details

2.1. Materials

117 compounds were employed in this study, which are from Refs. 7–11. Table 1 listed their structures and bio-

activities. For 3D QSAR analyses, 94 (unasterisked molecules in Table 1) compounds were selected as training set for model construction, and the rest 23 compounds (asterisked molecules in Table 1) as testing set for model validation. The method for selection of testing set in this study is fixed interval sampling method. We randomly selected No. 5 molecule in the molecular spreadsheet as starting point. Then, any molecules with number of $5n + 5$ ($n = 1, 2, 3, \ldots, 22$) would be included in the testing set.

2.2. Molecular structures

The structure of ginkgolide B was constructed based on its X-ray crystallography data[14] followed by a molecular mechanics (MM) optimization using standard TRIPOS force field and Gasterger–Marsili charge,[15] with a energy gradient convergence criterion of 0.001 kcal/mol and a distance-dependent dielectric constant. All the other 116 3-D molecular structures were constructed based on the skeleton of ginkgolide B, followed by fully geometrical optimization with the same method as above. For those molecules having rotatable bonds, systematic conformational searches were carried out to find the lowest energy structures.
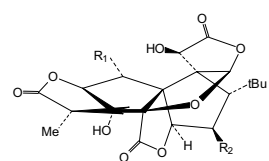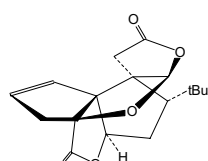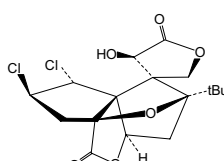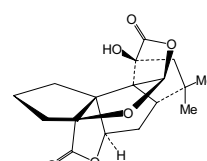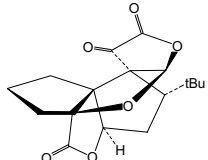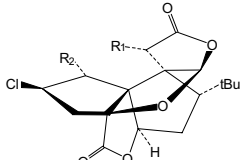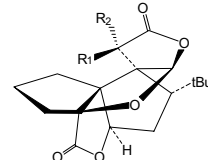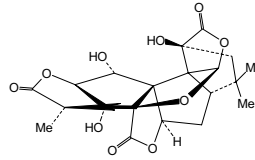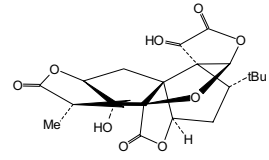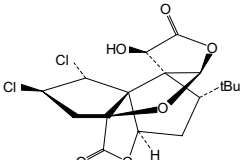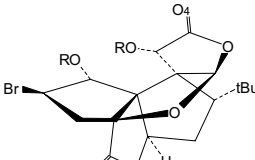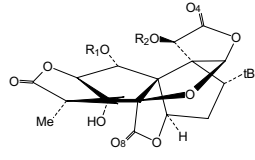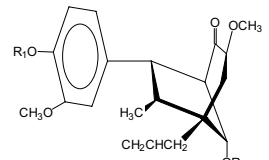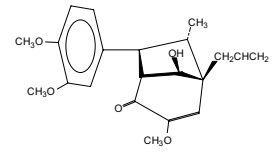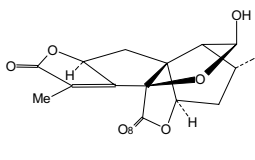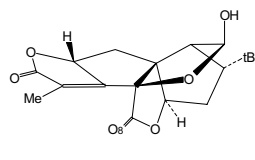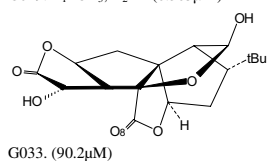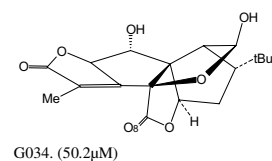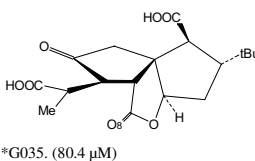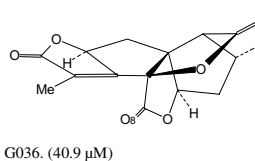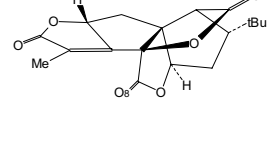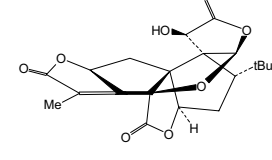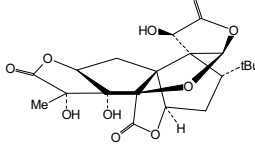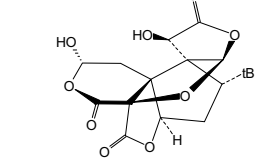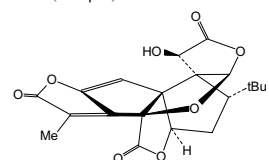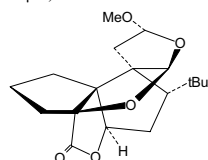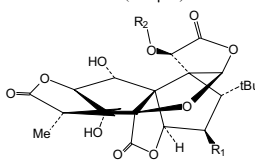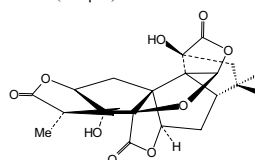
2.3. Molecular alignment

Molecular alignment may have significant influence on CoMFA and CoMSIA results. We noticed that all the 94 structures in the training set have rigid condensed ring structure. Therefore, the condensed ring structure of ginkgolide B (GB) was chosen as template for the structural alignment of the 94 molecules. Furthermore, two key atoms, ^{17}O and ^{18}O of GB in Figure 1 that exist in most of the studied molecules, were also taken into account during the alignment. For the molecules G27 to G30 that have great structural diversity from ginkgolides, the alignment was carried out based on the following considerations: (1) their allyl group corresponds to the *tert*-butyl of ginkgolide B; (2) their C=O group matches the ^{17}O of ginkgolide B; and (3) their aromatic ring fits the ring F of ginkgolide B. Regarding those molecules with long side chains, in order to align them well to each other, we adjust their orientation by rotating their rotatable bonds. However, all these rotations are confined within ∼5 kcal/mol to its lowest energy. Figure 2 depicts the aligned structures.

2.4. CoMFA

Steric and electrostatic interactions were calculated using a sp^3 carbon atom as steric probe and a +1 charge as electrostatic probe with Tripos force field. The CoMFA grid spacing is 2.0 Å in the x, y, and z directions, and the grid region is automatically generated by the CoMFA routine to encompass all molecules with an extension of 4.0 Å in each direction. The default value of 30 kcal/mol was set as the maximum steric and electrostatic energy cutoff. Minimum-sigma (column filtering) was set to be 2.0 kcal/mol to improve the signal-to-noise ratio by omitting those lattice points, which energy variation is below this threshold. Regres-

W. Zhu et al. / Bioorg. Med. Chem. 13 (2005) 313–322

315

Table 1. The structures and IC$_{50}$ values of the 117 ginkgolides and their analogues (*: Compounds in test set for model validation; all biological activity data were scaled to the same standard: IC$_{50}$ of ginkgolide B is 0.128 μM)

G001. R$_1$=H, R$_2$=H (3.79μM) (GA)
G002. R$_1$=OH, R$_2$=H (0.128μM) (GB)
G003. R$_1$=OH, R$_2$=OH (3.62μM) (GC)
G004. R$_1$=H, R$_2$=OH (27.6μM) (GJ)

*G005. (17.06μM)

G006. (8.10μM)

G007. (1.96μM)

G008. (3.84μM)

G009. R$_1$=H, R$_2$=H(0.235μM)
*G010. R$_1$=CH$_2$OMe, R$_2$=OCH$_2$OMe(0.064μM)
G011. R$_1$=H, R$_2$=Cl (0.235 μM)
G012. R$_1$=CH$_2$OCH$_3$,R$_2$=Cl(0.0.064μM)

G013. R$_1$=H, R$_2$=OH (16.2μM)
G014. R$_1$=H, R$_2$=OCOCH$_3$ (2.77μM)
*G015. R$_1$=H, R$_2$=OCH$_2$CH$_3$ (2.0 μM)
G016. R$_1$=OH, R$_2$=H (2.77 μM)
G017. R$_1$= OCOCH$_3$, R$_2$=H (2.35μM)
G018. R$_1$= OCH$_2$OCH$_3$, R$_2$=H (4.48μM)

G019. (0.149μM)

*G020. (0.832μM)

G021. (0.277μM)

G022. R=CH$_2$OCH$_3$ (0.128μM)
G023. R=H (2.99μM)

G024. R$_1$=H, R$_2$=PhCH$_2$OCH$_2$ (0.0404μM)
*G025. R$_1$=PhCH$_2$OCH$_2$, R$_2$=H (0.0588μM)
G026. R$_1$=Cl-PhCH$_2$OCH$_2$, R$_2$=H (0.0289μM)

G027. R$_1$=CH$_3$, R$_2$=COCH$_3$ (0.282μM)
G028. R$_1$=H, R$_2$=COCH$_3$ (0.485μM)
G029. R$_1$=CH$_3$, R$_2$=H (0.305μM)

*G030. (3.05μM)

G031. (54.0μM)

G032. (44.8μM)

G033. (90.2μM)

G034. (50.2μM)

*G035. (80.4μM)

G036. (40.9 μM)

G037. (17.8 μM)

G038. (17.8μM)

G039. (77.7μM)

*G040. (87.4μM)

G041 (74.1μM)

G042. (25.56μM)

G043. R$_1$=H, R$_2$=PhCOPhCH$_2$ (0.0343μM)
G044. R$_1$=OH, R$_2$=PhCOPhCH$_2$ (0.485μM)
*G045. R$_1$=H, R$_2$=(F$_3$C)(N$_2$)CPhCH$_2$ (0.0343μM)
G046. R$_1$=OH, R$_2$=(F$_3$C)(N$_2$)PhCH$_2$ (0.107μM)
G047. R$_1$=H, R$_2$=2,3,5,6-4F,4-N$_3$-PhCH$_2$ (0.0206μM)
G048. R$_1$=OH, R$_2$=2,3,5,6-4F,4-N$_3$-PhCH$_2$ (0.181μM)
G049. R$_1$=H, R$_2$=PhCOPhCO (0.0297μM)

*G050, (0.536μM)

Table 1 (continued)

G051, (1.89μM)   G052, (0.547μM)   G053, (112μM)   G054, (1.00μM)

*G055, R=PhCH₂,(65.9Mm)
G056, R=Ph,(1.56μM)
G057, R=CH₃,(80.8μM)   G061: R=H, (13.2μM)   G063, R1=H, R2=OH, (29.9μM)   *G065, (89.7μM)
G058, R=H, (123μM)   G062, R=OH, (6.34μM)   G064, R1=OH, R2=H, (3.85μM)
G059, R=OH,(68.2μM);
*G060, NHR=N(C2H4)(C2H4)O,
        (7.97μM)

G066, R=OMe, (4.02μM)   G069, R=OMe, (7.58μM)   G072, R=OMe, (123μM)   G074, R1=R2=OMe, (11.2μM)
G067, R=OEt, (7.19μM)   *G070, R=OEt, (71.9μM)   G073, R=OH, (101μM)   *G075, R1=R2=OEt, (112μM)
G068, R=O-n-Bu, (4.62μM)   G071, R=O-n-Bu, (109μM)   G076, R1=OH, R2=OEt, (4.62μM)

G077, (98.0μM)   G078, (50.5μM)   G079, (164μM)   *G080, (0.232μM)

G081. R₁=H,R₂=CH₃, (0.571μM)
G082, R₁=H,R₂=CH₃OCH₂, (1.68μM)

G083, R₁=H,R₂=PhCH₂OCH₂,
        (0.0588μM)
G084, R₁=H,R₂=CH₂=CHCH₂,
        (1.09μM)

*G085. R₁=H,R₂=C₂H₅OOCCH₂,(0.0429μM)
G086. R₁=H,R₂=PhCOCH₂, (0.121μM)
G087. R₁=H, R₂=p-MeOC₆H₄COCH₂, (0.106μM)
G088. R₁=H,R₂=PhCH₂, (0.0404μM)
G089. R₁=H,R₂=PhCH(CH₃), (1.63μM)
*G090. R₁=H,R₂=p-O₂NC₆H₄CH₂, (0.108μM)
G091. R₁=H,R₂=p-ClC₆H₄CH₂, (0.0289μM)
G092. R₁=H,R₂=p-FC₆H₄CH₂, (0.0440μM)
G093. R₁=H,R₂=CH₃CH₂OC, (0.426μM)
G094. R₁=CH₃,R₂=H, (0.344μM)   G099, (1.00μM)   *G100, (22.8μM)
*G095. R₁=CH₃OCH₂,R₂=H, (0.203μM)
G096. R₁=PhCH₂OCH,R₂= H, (0.0404μM)
G097. R₁=C₂H₅OOCCH₂,R₂=H, (0.213μM)
G098. R₁=R₂=CH₃OCH₂, (2.06μM)

G101, R1=R2=Ac, R3=H,(83.3μM)   G102, (139μM)   G103, R1=R2=H, R3=n-Bu,(646μM)   G104. (14.2μM)

*G105, (16.3μM)   G106, R=α-OAc, (1.14μM)   G108, R=α-OCOCH₂Ph, (0.349μM)   *G115, R=α-OH, (0.087μM)
                G107, R=α-OH, (0.620μM)   G109, R=α-N₃, (0.080μM)   G116, R=β-OH, (0.243μM)
                        *G110, R=α-F, (0.144μM)   G117, R=α-F, (0.015μM)
                        G111, R=α-Cl, (0.016μM)
                        G112, R=α-NHMe, (0.089μM)
                        G113, R=α-NHEt, (0.024μM)
                        G114, R=α-NH₂, (1.26μM)

sion analysis was performed using partial least square (PLS) algorithm.[16] The final model was developed with the optimum number of components equal to that yielding the highest q^2.
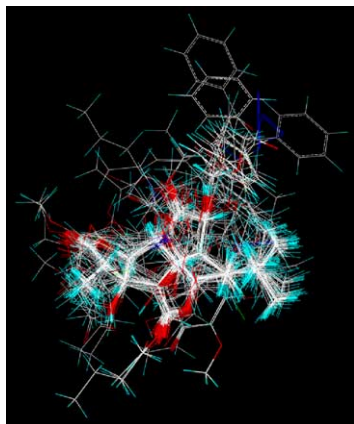
Figure 2. The alignment of the 94 structures in training set.

2.5. CoMSIA

The alignment model used in CoMFA study was adopted for CoMSIA investigation. Three physicochemical properties, viz. steric, electrostatic, and hydrophobic fields, were evaluated. The steric contribution was reflected by the third power of the atomic radii of the atoms. The electrostatic properties have been introduced as Gasteiger–Masili charges, and an atom based hydrophobicity is estimated according to the parameterization developed by Ghose and co-workers.[17,18] The lattice dimensions were selected with a sufficiently large margin (4 Å) to enclose all aligned molecules as in CoMFA. Any singularities were avoided at atomic positions in CoMSIA fields because a Gaussian type distance dependence of the physicochemical properties was adopted; thus, no arbitrary cutoffs were required. Similarity indices were computed using a probe with a charge of +1, a radius of +1 Å, a hydrophobicity of +1, and 0.3 as attenuation factor α for the Gaussian-type distance. The statistical evaluation for the CoMSIA analyses were performed in the same way as described for CoMFA.

In general, the CoMSIA contours are more compacted and centered on the ligand atoms, while CoMFA contours are scattered at the edge of the ligand surface. Therefore, CoMSIA result will be helpful in designing new ligand while CoMFA result should be useful in exploring the complementary features between ligand and its receptor active site.

2.6. HQSAR

Hologram QSAR (HQSAR) is an emerging QSAR technique. With the transformation of the chemical representation of a molecule into its corresponding molecular hologram, this method requires no explicit 3D information for the ligands.[13] In this study, the HQSAR method with its default parameters compiled in SYBYL6.9 was performed by the following procedure: at first, a predefined set of rules was used to hash a molecule into a molecular fingerprint that encoded the frequency of occurrence of various molecular fragment types; then, the molecular fingerprint was cut into

strings at a fixed interval as specified by a hologram length (HL) parameter; and at last, all of the generated strings were hashed into a fixed length array. The SYBYL line notation for each string was mapped to a unique integer in the range of 0–231 using a cyclic redundancy check algorithm. This numerical representation of molecules was exploited by a subsequent correlation analysis with PLS method. The optimal HQSAR model was finally constructed by screening the 12 default HL values ranging from 53 to 401.

3. Results and discussion

3.1. CoMFA

The CoMFA result was summarized in Table 2. The cross-validated value, q^2, is 0.665, with an optimum number of components is 6. The non-cross-validated PLS analysis produced a r^2 of 0.902. The estimated F value is 133 and standard error is 0.394. These statistical indexes are reasonably high, indicating that the new CoMFA model has a strong predictive ability. It is even better than what we had built based on 25 compounds in terms of q^2.[7] Table 3 summarized the PLS analysis results. Figure 3 depicted the correlation between experimental results and predicted values of the 94 studied compounds. Table 3 and Figure 3 demonstrate that the predicted values using the newly constructed CoMFA model are in well agreement with experimental data, suggesting that the new CoMFA model is reliable.

The data in Table 2 also showed that the CoMFA steric field descriptor explains 50.1% of the variance, while the electrostatic descriptor explains the rest 49.9%. These steric and electrostatic fields were presented as contour plots in Figures 4a and b, respectively. Ginkgolide B was displayed in the maps to aid in visualization. Two main sterically unfavorable regions (Fig. 4a, yellow contours) are located near to the atoms C29 and O25, and atom O17 (refer to Fig. 1 for atomic numbering), respectively, suggesting that bulky groups in these areas would lower the activity. Indeed, molecules G074–G076, G101, and G103 have bulky substituents around these two

Table 2. Statistical indexes of CoMFA, CoMSIA, and HQSAR models based on 94 compounds

	CoMFA	CoMSIA	HQSAR
PLS statistics			
q^2	0.665	0.669	0.591
r^2	0.902	0.906	0.798
S	0.394	0.384	0.562
F	133.038	168.747	353(BL)
Optimal components	6	5	5
Contribution			
Steric	0.501	0.143	
Electrostatic	0.499	0.450	
Hydrophobic		0.407	
Testing set			
r^2	0.811	0.541	0.666
S	0.657	0.993	0.847

Table 3. Predicted activities (PA) versus experimental activities (pIC_{50}) with residues (Δ) by CoMFA, CoMSIA, and HQSAR

Compd	PIC_{50}	CoMFA PA	Δ	CoMSIA PA	Δ	HQSAR PA	Δ
G001	5.42	5.49	−0.07	5.42	0.00	5.38	0.04
G002	6.89	6.11	0.78	6.06	0.83	5.81	1.08
G003	4.56	5.80	−1.24	5.78	−1.22	5.63	−1.07
G004	5.44	5.21	0.23	5.20	0.24	5.16	0.28
*G005	4.77	4.82	−0.05	4.61	0.16	5.68	−0.91
G006	5.09	5.05	0.04	6.01	−0.92	4.93	0.16
G007	5.71	6.08	−0.37	5.76	−0.05	5.90	−0.19
G008	5.42	5.28	0.14	5.94	−0.52	5.72	−0.30
G009	6.63	5.96	0.68	5.78	0.86	5.67	0.96
*G010	7.19	6.73	0.46	5.97	1.22	6.77	0.42
G011	6.63	6.54	0.09	6.30	0.33	6.59	0.04
G012	7.19	7.29	−0.10	6.45	0.74	6.95	0.24
G013	4.79	5.07	−0.28	5.36	−0.57	5.20	−0.41
G014	5.56	5.78	−0.22	5.33	0.23	5.16	0.40
*G015	5.70	6.14	−0.44	5.85	−0.15	5.56	0.14
G016	5.56	5.17	0.39	5.17	0.39	5.20	0.36
G017	5.63	5.30	0.33	5.33	0.30	5.16	0.47
G018	5.35	5.33	0.02	5.22	0.13	5.56	−0.21
G019	6.83	6.87	−0.04	6.57	0.26	6.50	0.33
*G020	6.08	6.58	−0.50	6.57	−0.49	6.16	−0.08
G021	6.56	6.53	0.03	6.48	0.09	6.59	−0.03
G022	6.89	6.89	0.00	6.58	0.31	6.65	0.24
G023	5.52	5.99	−0.47	6.49	−0.97	5.87	−0.35
G024	7.39	7.36	0.03	7.41	−0.02	6.92	0.47
*G025	7.23	7.57	−0.34	7.07	0.16	6.73	0.50
G026	7.54	7.61	−0.07	8.04	−0.50	6.83	0.71
G027	6.55	6.02	0.53	6.54	0.01	6.28	0.27
G028	6.31	6.07	0.24	6.48	−0.17	6.28	0.03
G029	5.84	5.95	−0.11	6.04	−0.20	6.36	−0.52
*G030	5.52	6.21	−0.69	5.68	−0.16	6.36	−0.84
G031	4.27	4.06	0.21	3.92	0.35	3.93	0.34
G032	4.35	4.31	0.04	3.89	0.46	3.93	0.42
G033	4.04	4.60	−0.56	4.38	−0.34	4.01	0.03
G034	4.30	4.37	−0.07	4.72	−0.42	4.13	0.17
*G035	4.09	4.36	−0.27	5.61	−1.52	4.16	−0.07
G036	4.39	4.46	−0.07	4.50	−0.11	4.27	0.12
G037	4.75	4.36	0.39	4.30	0.45	4.27	0.48
G038	4.75	5.07	−0.32	5.04	−0.29	5.43	−0.68
G039	4.11	5.04	−0.93	4.75	−0.64	5.35	−1.24
*G040	4.06	5.01	−0.95	5.40	−1.34	5.12	−1.06
G041	4.13	4.23	−0.10	4.43	−0.30	4.32	−0.19
G042	4.59	4.45	0.14	4.32	0.27	5.26	−0.67
G043	7.46	7.11	0.35	7.13	0.34	7.91	−0.45
G044	6.88	7.05	−0.17	6.93	−0.05	7.73	−0.85
*G045	7.46	6.88	0.58	7.53	−0.07	7.59	−0.13
G046	6.97	6.81	0.16	6.95	0.02	7.41	−0.44
G047	7.69	7.35	0.34	6.94	0.75	7.13	0.56
G048	6.74	7.08	−0.34	6.62	0.12	6.96	−0.22
G049	7.53	7.42	0.11	7.69	−0.16	7.44	0.09
*G050	6.27	5.56	0.71	4.29	1.98	6.08	0.19
G051	5.72	6.32	−0.60	6.30	−0.58	6.50	−0.78
G052	6.26	6.06	0.20	5.49	0.77	6.16	0.10
G053	3.95	3.89	0.06	4.03	−0.08	5.21	−1.26
G054	6.00	5.44	0.56	5.89	0.11	6.42	−0.42
*G055	4.18	4.84	−0.66	5.15	−0.97	4.71	−0.53
G056	5.81	5.54	0.27	4.99	0.82	4.76	1.05
G057	4.09	4.53	−0.44	4.49	−0.40	4.15	−0.06
G058	3.91	4.66	−0.75	4.51	−0.60	4.13	−0.22
G059	4.17	4.42	−0.25	4.15	0.02	4.16	0.01
*G060	5.10	3.87	1.23	4.82	0.28	4.29	0.81
G061	4.88	4.56	0.32	4.46	0.42	4.64	0.24
G062	5.20	5.05	0.15	5.13	0.07	4.96	0.24
G063	4.52	5.13	−0.61	5.19	−0.67	5.08	−0.56
G064	5.41	5.32	0.09	5.01	0.40	5.08	0.33

Table 3 (continued)

Compd	PIC_{50}	CoMFA PA	Δ	CoMSIA PA	Δ	HQSAR PA	Δ
*G065	4.05	4.19	−0.14	4.61	−0.56	4.63	−0.58
G066	5.40	5.51	−0.11	5.19	0.21	5.22	0.18
G067	5.14	5.02	0.12	5.13	0.01	5.20	−0.06
G068	4.42	4.44	−0.02	5.04	−0.62	5.15	−0.73
G069	5.12	5.12	0	4.47	0.65	4.75	0.37
*G070	4.14	4.63	−0.49	4.80	−0.66	4.72	−0.58
G071	3.96	4.02	−0.06	4.45	−0.49	4.68	−0.72
G072	3.91	4.03	−0.12	4.00	−0.09	4.06	−0.15
G073	4.00	4.35	−0.35	3.97	0.03	4.07	−0.07
G074	4.95	4.16	0.79	4.29	0.66	4.38	0.57
*G075	3.95	3.62	0.33	3.90	0.05	4.34	−0.39
G076	5.34	5.06	0.28	4.47	0.87	4.37	0.97
G077	4.01	4.21	−0.2	4.14	−0.13	4.54	−0.53
G078	4.30	4.21	0.09	3.87	0.43	4.41	−0.11
G079	3.79	3.84	−0.05	4.80	−1.01	4.31	−0.52
*G080	6.63	6.40	0.23	6.56	0.07	6.58	0.05
G081	6.24	6.27	−0.03	6.16	0.08	6.20	0.04
G082	5.77	5.85	−0.08	6.06	−0.29	6.17	−0.40
G083	7.23	7.52	−0.29	6.61	0.62	6.92	0.31
G084	5.96	6.20	−0.24	6.51	−0.55	6.16	−0.20
G085	7.37	6.76	0.61	6.15	1.22	6.15	1.22
G086	6.92	7.03	−0.11	7.16	−0.24	6.76	0.16
G087	6.97	6.69	0.28	6.79	0.18	6.83	0.14
G088	7.39	7.15	0.24	6.84	0.55	7.08	0.31
G089	5.79	5.89	−0.1	6.10	−0.31	7.04	−1.25
*G090	6.97	7.29	−0.32	5.45	1.52	7.28	−0.31
G091	7.54	7.28	0.26	7.01	0.53	7.18	0.36
G092	7.36	7.36	0	6.93	0.43	7.05	0.31
G093	6.37	6.32	0.05	6.58	−0.21	5.82	0.55
G094	6.46	6.24	0.22	6.89	−0.43	6.23	0.23
*G095	6.69	6.33	0.36	6.47	0.22	6.19	0.50
G096	7.39	7.33	0.06	7.75	−0.36	6.93	0.46
G097	6.67	6.82	−0.15	6.33	0.34	6.15	0.52
G098	5.69	5.77	−0.08	6.26	−0.57	6.56	−0.87
G099	6.00	5.92	0.08	6.11	−0.11	6.42	−0.42
*G100	4.64	4.70	−0.06	5.66	−1.02	6.73	−2.09
G101	4.08	3.85	0.23	4.66	−0.58	4.39	−0.31
G102	3.86	3.73	0.13	3.95	−0.09	3.54	0.32
G103	3.19	3.26	−0.07	4.47	−1.28	4.89	−1.70
G104	4.85	4.27	0.58	4.83	0.02	5.25	−0.40
*G105	4.79	4.02	0.77	4.60	0.19	5.67	−0.88
G106	5.94	6.59	−0.65	6.10	−0.16	5.41	0.53
G107	6.21	6.33	−0.12	6.22	−0.01	5.47	0.74
G108	6.46	6.41	0.05	7.16	−0.70	5.96	0.50
G109	7.10	7.07	0.03	6.80	0.30	6.85	0.25
*G110	6.84	6.34	0.5	6.42	0.42	6.31	0.53
G111	7.80	6.71	1.09	6.61	1.19	6.92	0.88
G112	7.05	6.95	0.1	6.38	0.67	6.66	0.39
G113	7.62	7.22	0.4	6.35	1.27	6.79	0.83
G114	5.90	6.52	−0.62	6.06	−0.15	6.49	−0.59
*G115	7.06	7.21	−0.15	6.87	0.19	6.91	0.15
G116	6.61	7.73	−1.12	7.37	−0.76	6.91	−0.30
G117	7.82	7.15	0.67	8.01	−0.19	7.58	0.24

* The compounds in test set for model validation.

regions, and are weaker in bioactivities in comparison with their analogues. There are two small yellow polyhedra close to tertiary butyl, indicating that bulky substituent in this area is not favorable to molecular bioactivity. A huge green polyhedron is near to the atoms O23 and O24, indicating a bulky substituent favorable region. This is in agreement with our earlier predicted and experimental results that bulkier
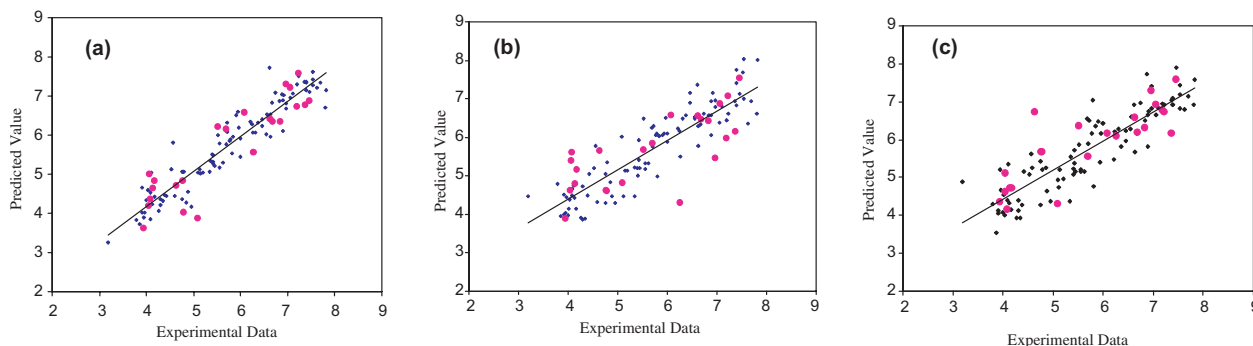
Figure 3. Plot of the observed pIC$_{50}$ versus the predicted activity (the data for the training set are marked as ♦, and testing set as ●). (a) Result from CoMFA. (b) Result from CoMSIA. (c) Result from HQSAR.
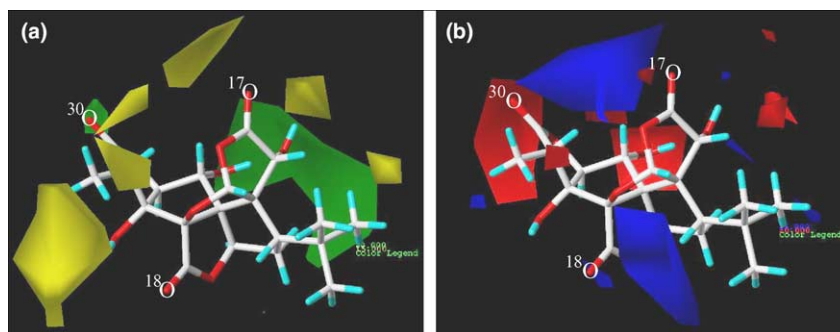

Figure 4. CoMFA contour maps in combination with ginkgolide B. (a) Regions where increasing the molecular volume increases bioactivity are in green and regions where increasing the volume decreases the activity are in yellow. (b) Regions where increasing the positive charges increases activity are in blue and regions where increasing the negative charges increases the activity are in red.

substitutions around these areas significantly improve molecular bioactivity.[7]

Shown in Figure 4b is CoMFA electrostatic contour map. A big blue polyhedra very close to the atoms O26, O30 and the area above them suggests that positively charged substituents in this region lead to an increase in ligand bioactivity. Another blue polyhedra is near the center of the area encompassed by atoms O8, C13, and O18, where positive charge should be favorable to bioactivity. There are two main red polyhedra, one near the atom O24, the other is located at the area below the atoms O26 and O30, indicating that more negative charges in these areas should play a favorable role in improving activity.

3.2. CoMSIA

The CoMSIA results were also summarized in Table 2. A cross-validated value, q^2 of 0.669 and a conventional r^2 of 0.906 were obtained. The F value and standard error are 169 and 0.384, respectively. These data also indicate that a reliable CoMSIA model was successfully constructed. The steric field descriptor explains 14.3% of the variance, while the proportion of electrostatic descriptor account for 45.0%, and the hydrophobic field explains the rest 40.7%. Therefore, the electrostatic and hydrophobic fields have greater effluence than steric. Figure 5 depicted the CoMSIA coefficient contour maps with ginkgolide B displayed for visualization.
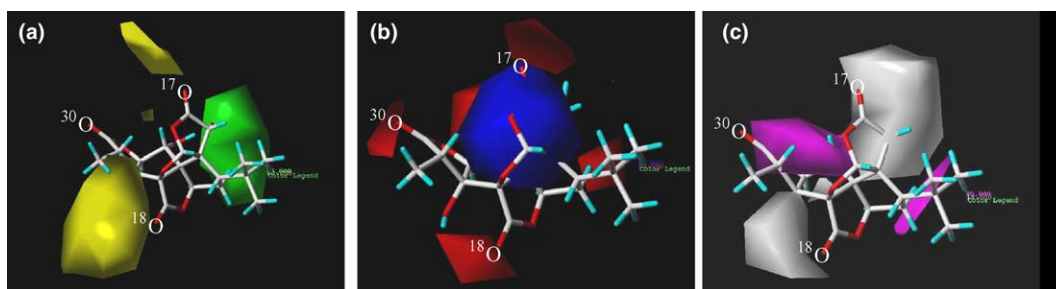
Figure 5. CoMSIA contour maps in combination with ginkgolide B. (a) Regions where increasing the molecular volume increases bioactivity are in green and regions where increasing the volume decreases the activity are in yellow. (b) Regions where increasing the positive charges increases activity are in blue and regions where increasing the negative charges increases the activity are in red. (c) Regions where improving hydrophilicity increases molecular activity are in white, and regions where improving hydrophobicity improve molecular activity are in magenta.

Very similar to the CoMFA steric Std* coeff contour map, CoMSIA also has two big yellow polyhedra: one is near the atoms C29, O25, and O18; the other is close to O17 (Fig. 5a). Therefore, bulky substituents around these regions would significantly lower molecular bioactivity. Meanwhile, in comparison with CoMFA some small polyhedra are disappeared in CoMSIA contour map (refer to Figs. 4a and 5a). Similar to CoMFA result, a big green polyhedron close to O24 in Figure 5a represents a region where steric bulk is favorable for biological activity.

There is only one blue region in CoMSIA electrostatic std* coeff contour map as indicated in Figure 5b, which is encompassed by the rings of A, D and C (Fig. 1) where more positive charge is favorable for enhancing biological potency. While there are five red contour polyhedra scattered around ligands, indicating regions where more negative charge is favored. Figure 5b shows that all the red contours are close to oxygen atoms except a polyhedron that is close to the α hydrogen of C12. Indeed, molecule G111 with higher activity than GB has a negatively charged α-Cl at this position, indicating a promising site for designing more powerful PAF antagonists.

One advantage of CoMSIA is that it introduces hydrophobic descriptor to explore the effect of molecular hydrophobicity on molecular bioactivity. The hydrophobic similarity index field of 94 compounds in the training set is showed in Figure 5c. White contours indicate that improving hydrophilicity around these areas improve molecular activity. There is a medium white polyhedron around O25 hydroxyl group, and a big one around O23 hydroxyl group, suggesting these two regions should be kept in hydrophilic for higher molecular bioactivity. Magenta contour indicates the places where improving hydrophobicity improve molecular activity. The region around ring A is surrounded by a big magenta polyhedron, suggesting an area where hydrophobic substituents are beneficial to molecular activity. This is in agreement with the previous experimental result that substituting the H atom of O24 or of O23 with phenyl group would improve molecular bioactivity.[7] A small magenta region close to α and β hydrogen atoms of C12 represents the preferred hydrophobic interaction around there. Therefore substitution of the H with OH in this region would reduce molecular bioactivity. This could be, at least, one of the reasons why ginkgolide B is stronger in bioactivity than ginkgolide C, G107 and G114. In contrast, the substitution of the α hydrogen atom with chlorine results in more powerful structure, G111.

3.3. HQSAR

The results of the HQSAR analysis on the training set were listed in Table 2 as well. The cross-validated PLS analysis yielded a q^2 of 0.591 with five optimal components, and conventional PLS analysis gave a r^2 value of 0.798 with a standard error of 0.562. The hologram that gives the lowest standard error has a length of 353. The predicted bioactivities were summarized in

Table 3, and the correlation between experimental data and HQSAR predicted values was depicted in Figure 3c. All these results suggest a good HQSAR model was successfully built. The model should be good enough to at least rank molecular activities for further drug discovery.

Figure 6 depicts the individual atomic contribution of GB to its molecular bioactivity. The colors at the red end of the spectrum (red, red orange, and orange) reflect poor (or negative) contributions, while colors at the green end (yellow, green blue, and green) indicate favorable (positive) contributions. Atoms with intermediate contributions are colored in white. Figure 6 shows that the condensed rings A, D, C, and E (refer to Fig. 1) are favorable for molecular bioactivity, while further bioactivity improvement could be achieved by modifying atoms C12 and C19 that are colored in orange. The atoms of ring F are colored in white, suggesting a possible site for simplifying the condensed ring structure.

3.4. Validation of the 3D-QSAR models

To verify the stability and predictive ability of the constructed 3D-QSAR models based on the training set, 23 compounds were selected as testing set for validation. The predicted bioactivities for the 23 compounds based on the above models were also shown in Table 3 (labeled with star). The correlation between the predicted and experimental results was depicted in Figure 3 (labeled as ●). The PLS analysis on the testing set released r^2 values of 0.811, 0.541, and 0.666; and s values of 0.657, 0.993, and 0.847 for CoMFA, CoMSIA, and HQSAR, respectively, demonstrating again that the CoMFA, CoMSIA, and HQSAR models are fairly reliable and have rather strong predictive ability, and should be useful in designing new PAF antagonists with simple structures.

3.5. Elucidating binding mechanism

Ginkgolides should bind with PAF receptor (PAFR) as they are PAF antagonists. Therefore, binding character-
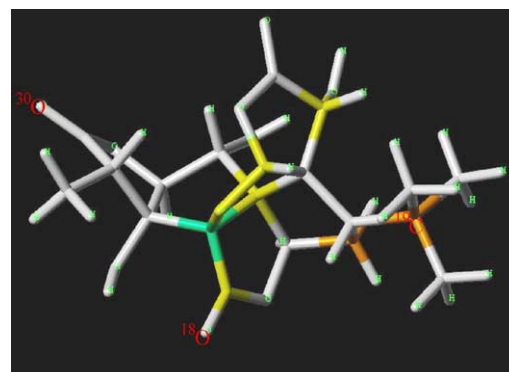

Figure 6. The individual atomic contributions to activity. The colors at red end of the spectrum reflect poor contributions, while colors at green end indicate favorable contributions.

istic between PAFR and ginkgolides should be helpful in better understanding how PAF antagonists bind with PAFR, and how to design or discover new PAF antagonists. Unfortunately, the 3D structure of PAFR is not available to date. As we know, a pharmacophoric model of possible electrostatic, H-bonding, and hydrophobic interaction sites could be deduced from CoMFA and CoMSIA models, which are useful in new drug discovery. In order to deduce this mechanism, the possible acceptor and donor sites of H-bonding interaction between ligands and PAFR binding site were explored by using the same parameters as in CoMSIA. Figure 7 depicted the favorable and unfavorable sites as H-Bonding donor and acceptor.

Figure 7a shows that ligands would have three H-bonding acceptor sites. The first one is around atom O17; the second is close to O18; and the last is near to O25 (refer to Fig. 1 for atomic No.). It was found that three histidine residues of PAFR (His-188, His-248, and His-249) would interact with a zinc ion to form a complex.[19] This

finding indicates that these three histidines should adjoin and form a binding pocket. Therefore, the three possible H-bonding acceptor sites in Figure 7a is in agreement with the characteristic of the binding pocket that the binding pocket with three histidines as H-bonding donor could form three H-bonds with ginkgolide B, demonstrating that our CoMSIA models are reasonable.

Figure 7b depicted the possible H-bonding donor sites. A major polyhedron in cyan is close to H atom of O25 (refer to Fig. 1 for atom numbering). It suggests that, as H-bond donor, the hydroxyl group HO^{25}—is important to ligand binding with PAFR. Overall, this group could form two H-bonds with the receptor binding pocket. Hence, this hydroxyl group is essential to the bioactivities of ginkgolides. Indeed, the compound G043, which has no such functional group, is much weaker in binding affinity than ginkgolide A (G001), GB, and ginkgolide C (G003). Another major polyhedron in cyan is close to the H atom of C11, indicating that modifying this area with H-bond donor
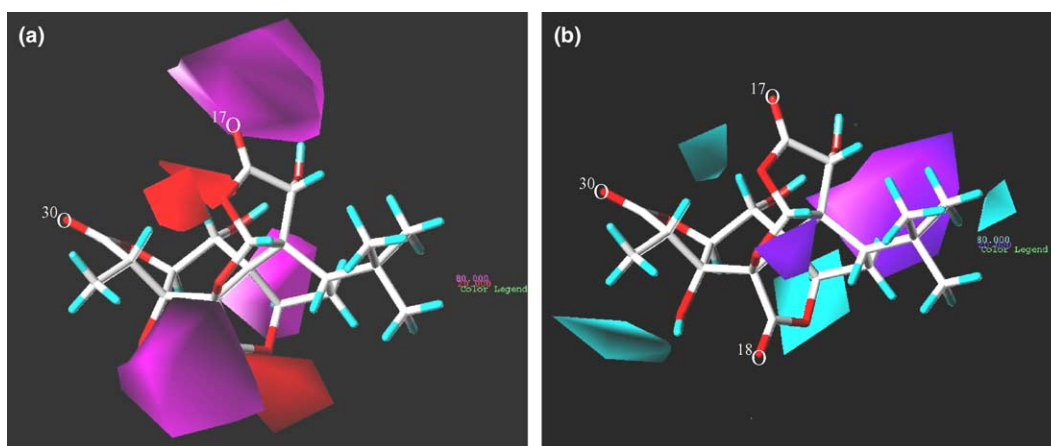

Figure 7. H-bonding sites for the binding of ligands with PAFR deduced from CoMSIA results with ginkgolide B displayed. (a) H-bonding acceptor sites. Polyhedra in magenta indicate those favored areas as H-bonding acceptor while polyhedra in red are disfavored; (b) H-bonding donor sites. Polyhedra in cyan color indicate those favored areas as H-bonding donor while polyhedra in purple are disfavored.
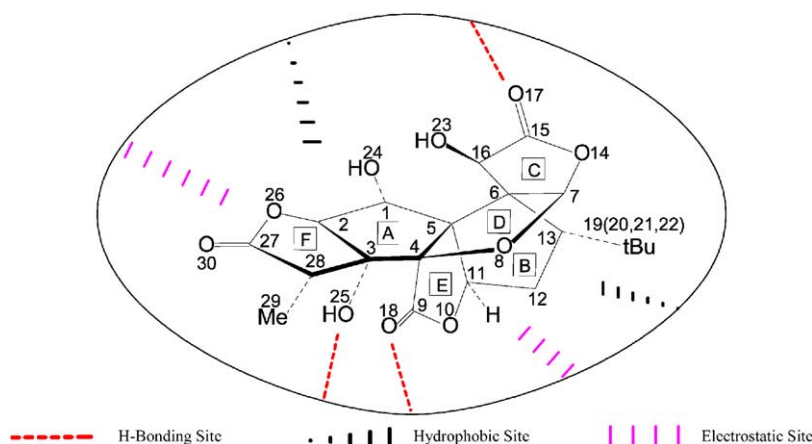

Figure 8. The possible binding mechanism between ligands and PAFR with ginkgolide B displayed for visualization.

group may further improve the binding affinity of ligand to PAFR.

Taking into account all the information of electrostatic, hydrophobic, H-bonding acceptor site, and H-bonding donor site from CoMFA and CoMSIA, we deduced the possible binding mechanism between PAFR and its ligands as shown in Figure 8.

4. Conclusions

Three QSAR methods, CoMFA, CoMSIA, and HQSAR, are used to investigate the relationship between the structures of 94 antagonists of PAF and their activities in order to get new clues to develop new potent PAF antagonists. The high q^2 released from these different QSAR methods suggest that we successfully got three rational and predictive QSAR models. The good correlation between experimental and predicted bioactivities for 23 compounds in testing set further verified the reliability of the constructed QSAR models. These models could be used to account for why ginkgolide B has unusual higher bioactivity than other ginkgolides. In addition, this study also reveals how individual atom and molecular hydrophobicity affect the molecular bioactivity against PAF receptor. The deduced binding mechanism between ligands and the binding site of PAFR is in agreement with experimental result that three histidines of PAFR could form H-bond with ligands, which further support the reliability of our QSAR models. Thus, the QSAR study result also cast a light on how ginkgolides binds to PAF receptor. All this information could be helpful in simplifying the complicated structure of ginkgolides to design new easily synthesized PAF antagonists that could be the drugs for curing diseases related to PAF or PAF receptor.

References and notes

1. Drieu, K.; Jaggy, H. In *Medicinal and Aromatic Plants-Industrial Profiles: Ginkgo biloba*; van Beek, T. A., Ed.; Harwood Academic: Amsterdam, 2000; Vol. 12, pp 267–277.
2. Shen, T. Y.; Hwang, S. B.; Doebber, T. W.; Robbins, J. C. In *Platelet-Activating Factor and Related Lipid Mediators*; Snyder, F., Ed.; Plenum Press: New York, 1987; pp 153–200.
3. Hanahan, D. J. *Ann. Rev. Biochem.* **1986**, *55*, 483.
4. Kroegel, C.; Kortsik, C.; Kroegel, N.; Matthys, H. *Drugs Aging* **1992**, *2*, 345.
5. Perry, S. W.; Hamilton, J. A.; Tjoelker, L. W.; Dbaibo, G.; Dzenko, K. A.; Epstein, L. G.; Hannun, Y.; Whittaker, J. S.; Dewhurst, S.; Gelbard, H. A. *J. Biol. Chem.* **1998**, *273*, 17660.
6. Clerk, M.; Cramer, R. D., III; Jones, D. M.; Patterson, D. E.; Simeroth, P. E. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
7. Chen, J. Z.; Hu, L. H.; Jiang, H. L.; Gu, J. D.; Zhu, W. L.; Chen, Z. L.; Chen, K. X.; Ji, R. Y. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1291.
8. Hu, L.; Chen, Z.; Xie, Y.; Jiang, Y.; Zhen, H. *Bioorg. Med. Chem.* **2000**, *8*, 1515.
9. Hu, L.; Chen, Z.; Xie, Y.; Jiang, Y.; Zhen, H. *J. Asian Nat. Prod. Res.* **2001**, *3*, 219.
10. Strømgaard, K.; Saito, D. R.; Shindou, H.; Ishii, S.; Shimizu, T.; Nakanishi, K. *J. Med. Chem.* **2002**, *45*, 4038.
11. Vogensen, S. B.; Strømgaard, K.; Shindou, H.; Jaracz, S.; Suehiro, M. I.; Ishii, S.; Shimizu, T.; Nakanishi, K. *J. Med. Chem.* **2003**, *46*, 601.
12. Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, *37*, 4130.
13. Tong, W.; Lowis, D. R.; Perkins, R.; Chen, Y.; Welsh, W. J.; Goddette, D. W.; Heritage, T. W.; Sheehan, D. M. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 669.
14. Dupont, L.; Dideberg, O.; Germain, G.; Braquet, P. *Acta Crystallogr.* **1986**, *C42*, 1759.
15. SYBYL [computer program]. Version 6.8. St Louis (MO): Tripos Associates, 2001.
16. Dunn, W. J.; Wold, S.; Edlund, U.; Hellberg, S.; Gasteiger, J. *Quant. Struct.–Act. Relat.* **1984**, *3*, 31.
17. Ghose, A.; Crippen, G. *J. Comput. Chem.* **1986**, *7*, 565.
18. Viswanadhan, V. N.; Ghose, A. K.; Revankar, G. R.; Robins, R. K. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 163.
19. Ishii, I.; Izumi, T.; Tsukamoto, H.; Umeyama, H.; Ui, M.; Shimizu, T. *J. Biol. Chem.* **1997**, *272*, 7846.