Chemometrics and Intelligent Laboratory Systems 131 (2014) 22–30

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemolab

Molecular docking and QSAR analysis on maleimide derivatives selective inhibition against human monoglyceride lipase based on various modeling methods and conformations

Jingheng Wu, Yaxue Wang, Yong Shen *

School of Chemistry and Chemical Engineering of Sun Yat-sen University, 510275 Guanzhou, PR China

ARTICLE INFO

Article history:
Received 4 August 2012
Received in revised form 4 November 2013
Accepted 24 November 2013
Available online 10 December 2013

Keywords:
Human MGL
Maleimide derivatives
Molecular docking
QSAR
Self-adaptive GA-ANN

ABSTRACT

Inhibitory effect to endocannabinoid system-related human monoglyceride lipase (MGL) and selectivity toward fatty acid amid hydrolase of promising maleimide derived inhibitors were investigated by molecular docking and QSAR study. The essential roles of Ala61, Ser132 and His279 related hydrogen bonds and Tyr204 involved π–π interaction, were emphasized by the docking analysis, which were in good agreement with the experimental observations by far. By performing our new developed self-adaptive genetic algorithm (GA) and artificial neural network (ANN) combined method, as well as multiple linear regression and least squares support vector machine based GA method, significant descriptors were selected to build linear and non-linear models. Strong internal and external validations proved the robustness and effectiveness of docking conformation derived models and that importing descriptors from unrealistic conformations based on geometry optimization is not always appropriate for non-linear modeling. Besides, good linear relation between predicted activities and experimental ones towards rat MGL implicates human MGL and rat MGL may share similar inhibitory mechanism.

1. Introduction

Anandamide (AEA) and 2-arachidonoyl glycerol (2-AG) are two known primary endocannabinoids greatly in response to various pathological events related with human neuron system [1]. In their generation and degradation regulating, fatty acid amid hydrolase (FAAH) and monoglyceride lipase (MGL, a.k.a. monoacylglycerol lipase, MAGL) are proposed to be the most common pathways [2]. FAAH has been in interest for a long time as an AEA-degrading enzyme, while MGL was assumed the primary route in neuronal 2-AG inactivation [3,4]. As more and more observations supported that 2-AG plays a more important role in the endocannabinoid system, MGL inhibition was believed to be significant in the therapy of pain, inflammation, and neurodegenerative disorders [5]. However, currently found MGL inhibitors are still not available for their potentials and selectivity. Some inhibitors of human MGL (hMGL) showed reversible interaction effect or questionable inhibitory action [5,6]. Besides, some MGL inhibitors can also interact with FAAH leading to low selectivity toward FAAH as the binding pockets of MGL and FAAH may share some structural similarities [7,8]. Therefore, selective MGL inhibitors still need further development [9].

N-arachidonylmaleimide (NAM) was brought up to be a potential MGL inhibitor with favorable selectivity in recent years [10]. Since NAM shares structural similarity with natural MGL substrate 2-AG, its mechanism of activity is full of interest. In a recent study, a new set of maleimide derivatives based on the structure of NAM with available activity and selectivity in hMGL inhibition was presented [11]. But far more details about how they act on those two enzymes and, the most interesting, their inhibitory selectivities still remain unclear. In order to design more rational drug candidates, further improvement on NAM-like hMGL inhibitors is in demand.

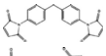
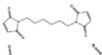
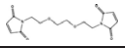
Quantitative structure–activity relationship (QSAR) is one of the effective approaches to discover the underlying inhibitory mechanism of NAM derivatives in hMGL. It can dig up important structural features relevant to molecular activities and screen for potential drugs by predicting their biological activities. Selecting significant descriptors is one of the key steps in QSAR analysis [12]. However, structure–activity relationships other than linear were rarely considered since they are usually complex [13]. Although developing non-linear models like artificial neural network (ANN) and support vector machine (SVM) is commonly used, it may take the risk mapping the original space into a higher-dimensional space. Proper feature selection method should be introduced in the case of non-linear modeling. Recently, our group developed a genetic algorithm (GA) and ANN combined descriptors selecting method, namely self-adaptive GA-ANN method [14]. This method takes the advantage of strong global searching ability of GA and effective unbiased error estimate in cross-validation for feature

* Corresponding author. Tel.: +86 20 8411 0696.
E-mail address: cessy@mail.sysu.edu.cn (Y. Shen).
URL: http://ce.sysu.edu.cn/en/facultystaff/faculty/13793.html (Y. Shen).

0169-7439/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.chemolab.2013.11.011

selection. It has been known that the ANN models evaluated by cross-validation are at the risk of over-training, which may result in over-fitted models [15]. In the self-adaptive GA-ANN method, the evaluated deviation of the left sample in leave-one-out cross-validation (LOO-CV) was employed to GA estimate function and only the model with the minimum value of the fitness function was used. As a rational step of training was introduced, some influence from inherent defects in non-linear modelling like over-fitting is able to be avoided in our feature selection when ANN encounter over-training with its estimated error ascending [14]. LOO-CV estimator was also a good choice in variable selection when least squares SVM (LS-SVM) is employed for modelling, which is a simple and efficient modified version of standard SVM algorithm [16]. In this study, molecular docking, together with 2D-QSAR, was employed to shed light on details in biological action of maleimide derivatives. As to QSAR, molecular conformations obtained by density functional theory (DFT) based quantum mechanics (QM) geometry optimization is commonly utilized, whereas docking-guided molecular conformations were considered to be more similar to the active conformations and help saving time as well [17]. These two types of molecular conformations were both involved and discussed in this work. Multiple linear regression (MLR), LS-SVM and ANN based descriptor selections using GA method with LOO-CV helped find significant descriptors to buid QSAR models with good predictability. To ensure the model validity, all established models were strictly verified by internal/external validations, ranking method and randomization tests.

Table 1
Set of maleimides derivatives and their experimental activities, selectivity ratios toward hMGL and hFAAH from Ref. [11] and rescored values toward hMGL by ChemScore.

No.	Substituent (**1**–**46**)/Structure (**47**–**53**)	Full IUPAC name	pIC_{50}		Selectivity ratio[a]	Docking score
			hMGL	FAAH		
1	Arachidonyl	*N*-arachidonylmaleimide	−0.049	−0.338	1.95	13.2
2	Phenyl	*N*-phenylmaleimide	−1.201	−2.360	14.4	8.4
3	*o*-Tolyl	1-(2-Methylphenyl)maleimide	−0.430	–	>200	10.3
4	*m*-Tolyl	1-(3-Methylphenyl)maleimide	−0.439	–	>200	9.5
5	*p*-Tolyl	1-(4-Methylphenyl)maleimide	−1.149	–	>150	8.2
6	2-Ethylphenyl	1-(2-Ethylphenyl)maleimide	−0.301	–	>150	6.6
7	3-Ethylphenyl	1-(3-Ethylphenyl)maleimide	−0.670	−2.013	22	7.1
8	4-Ethylphenyl	1-(4-Ethylphenyl)maleimide	−0.450	−2.212	58	5.0
9	4-Propylphenyl	1-(4-Propylphenyl)maleimide	−0.770	−2.270	32	7.3
10	2-Isopropylphenyl	1-(2-Isopropylphenyl)maleimide	−0.895	−1.813	9	3.0
11	4-Isopropylphenyl	1-(4-Isopropylphenyl)maleimide	−0.459	−2.079	40	6.0
12	4-Heptylphenyl	1-(4-Heptylphenyl)maleimide	−0.326	−1.940	41	9.9
13	4-Hydroxyphenyl	1-(4-Hydroxyphenyl)maleimide	−1.111	−2.459	30	4.5
14	2-Methoxyphenyl	1-(2-Methoxyphenyl)maleimide	−0.840	–	78	6.7
15	3-Methoxyphenyl	1-(3-Methoxyphenyl)maleimide	−0.820	–	>200	7.3
16	4-Methoxyphenyl	1-(4-Methoxyphenyl)maleimide	−0.760	–	>200	2.7
17	2-Fluorophenyl	1-(2-Fluorophenyl)maleimide	−0.650	−2.470	63	5.9
18	3-Fluorophenyl	1-(3-Fluorophenyl)maleimide	−0.378	−1.949	37	1.7
19	4-Fluorophenyl	1-(4-Fluorophenyl)maleimide	−0.714	−2.079	23	4.8
20	2-Chlorophenyl	1-(2-Chlorophenyl)maleimide	−0.620	−2.330	77	5.4
21	3-Chlorophenyl	1-(3-Chlorophenyl)maleimide	−0.450	−2.009	39	−3.6
22	4-Chlorophenyl	1-(4-Chlorophenyl)maleimide	−0.860	−2.049	7	8.0
23	2-Bromophenyl	1-(2-Bromophenyl)maleimide	−0.465	–	75	−2.4
24	3-Bromophenyl	1-(3-Bromophenyl)maleimide	−0.630	−2.179	27	−3.1
25	4-Bromophenyl	1-(4-Bromophenyl)maleimide	−0.640	−1.708	50	6.2
26	2-Iodophenyl	1-(2-Iodophenyl)maleimide	−0.700	−2.301	37	10.1
27	3-Iodophenyl	1-(3-Iodophenyl)maleimide	−0.350	−1.792	45	7.2
28	4-Iodophenyl	1-(4-Iodophenyl)maleimide	−0.637	−1.778	14	4.7
29	[1,1′-Biphenyl]-2-yl	1-Biphenyl-2-ylmaleimide	−0.190	−2.255	116	8.6
30	[1,1′-Biphenyl]-3-yl	1-Biphenyl-3-ylmaleimide	−0.369	−1.690	21	8.2
31	[1,1′-Biphenyl]-4-yl	1-Biphenyl-4-ylmaleimide	−0.210	−0.770	4	11.4
32	Benzyl	1-Benzylmaleimide	−0.711	−2.270	36	9.4
33	Phenethyl	1-Phenethylmaleimide	−0.881	−2.396	33	11.8
34	Phenylbutyl	1-(4-Phenylbutyl)maleimide	−0.330	−1.929	39	15.8
35	Ethylbenzyl	1-(4-Ethylbenzyl)maleimide	−0.550	−2.090	35	10.8
36	[1,1′-Biphenyl]-4-ylmethyl	1-Biphenyl-4-ylmethylmaleimide	0.102	−1.220	73	21.5
42	Methyl	*N*-methylmaleimide	−1.389	–	–	0.0
43	Ethyl	*N*-ethylmaleimide	−1.220	–	–	2.0
44	Nonyl	*N*-nonylmaleimide	−0.550	−1.610	12	10.7
45	Hexadecyl	*N*-hexadecylmaleimide	−0.430	–	–	N.A.[b]
46	Oleyl	*N*-oleylmaleimide	0.032	–	–	N.A.[b]
47		*N*,*N*′-(1,3-phenylene)dimaleimide	−0.307	−1.571	18	−1.0
48		*N*,*N*′-(1,4-phenylene)dimaleimide	−0.346	−1.130	6	12.2
49		1,4-Bismalimidoxylene	−0.230	−1.220	10	6.7
50		1,1′-(Methylenedi-4,1-phenylene)bismaleimide	−0.350	−1.029	5	5.2
51		1,4-Bismaleimidobutane	**−1.459**	−0.920	0.3	9.8
52		1,6-Bismaleimidohexane	−0.230	−1.260	11	9.5
53		1,8-Bis-maleimidodiethyleneglycol	−0.290	–	–	8.0

[a] The selectivity ratio was given by the original literature which was calculated by IC_{50} value of FAAH divided by IC_{50} value of hMGL.
[b] N. A.: scoring was not available.