Int. J. Mol. Sci. **2003**, *4*, 249-262

International Journal of
Molecular Sciences
ISSN 1422-0067

www.mdpi.org/ijms/

Quantitative Structure-Activity Relationships of Noncompetitive Antagonists of the NMDA Receptor: A Study of a Series of MK801 Derivative Molecules Using Statistical Methods and Neural Network

M. Elhallaoui[1,*], M. Elasri[1], F. Ouazzani[1], A. Mechaqrane[2] and T. Lakhlifi[3]

[1,2] Département de chimie et Département de Physique, Faculté des Sciences et Techniques Fès-Saïss Université Sidi Mohamed Ben Abdellah B.P.2202 Route D'Immouzzer, Fès, Morocco
[3] Département de chimie Faculté des sciences Université Moulay Ismail B.P.4010 Bni M'Hamed Meknes, Morocco
* Author to whom the correspondence should be addressed. E-mail: m.elhallaoui@caramail.com

Received: 12 December 2001 / Accepted: 8 October 2002 / Published: 15 April 2003

Abstract: From a series of 50 MK801 derivative molecules, a selected set of 44 compounds was submitted to a principal components analysis (PCA), a multiple regression analysis (MRA), and a neural network (NN). This study shows that the compounds' activity correlates reasonably well with the selected descriptors encoding the chemical structures. The correlation coefficients calculated by MRA and there after by NN, r = 0.986 and r = 0.974 respectively, are fairly good to evaluate a quantitative model, and to predict activity for MK801 derivatives. To test the performance of this model, the activities of the remained set of 6 compounds are deduced from the proposed quantitative model, by NN. This study proved that the predictive power of this model is relevant.

Keywords: structure-activity relationships, noncompetitive antagonists, MK801 derivatives, NMDA receptor, principal components analysis (PCA), multiple regression analysis (MRA), neural network (NN).

Introduction

The excitatory amino acids' receptors are implicated in the pathology of neurological and neurodegenerative disorders such as epilepsy, Huntington's and Alzheimer's diseases, schizophrenia,

etc. [1-3]. Among the attempts to discover neuroprotective agents, one is directed towards the search of noncompetitive antagonists of the NMDA receptor. The discovery that MK801 is a selective noncompetitive antagonist of the NMDA subclass of receptors, for the excitatory amino acid L-glutamic acid, in brain tissue, has provided insight into the underlying mechanism of the anticonvulsant action [4]. Based on the hypothesis that there exists an active common structure in the central nervous system consisting of an aromatic group and nitrogen atom, several geometric models have been proposed. In these models, the receptor sites have been localized and several authors tried to describe geometrically the interaction mode [5-12]. In a previous study we have proved that the pharmacophore is conform to one of MK801 configurations [13].

In this work we attempt to establish a quantitative structure - activity relationship for noncompetitive antagonists of NMDA receptor by studying a selected series of 44 MK801 derivatives (dibenzo [a,d] cycloalkenimines) from 50 compounds [4]. We accordingly propose a quantitative model, and we try to interpret the activity of the compounds relying on the multivariate statistical analyses. The principal components analysis (PCA) has served to classify the compounds according to their activities and to give an estimation of the values of the pertinent descriptors that govern this classification. The multiple regression analysis (MRA) has served to select the descriptors used as the input parameters for a back propagation network (NN). This linear method (MRA) has served also to predict activities, but when compared with the results given by the NN, we realized that the predictions fulfilled by this latter were more effective. To test the performance of this model we have used the cross validation method, thereafter, the activities of the remained set of 6 compounds are deduced from the proposed quantitative model with NN.

Methods and Equipment

So as to determine a quantitative structure - activity relationship for noncompetitive antagonists of NMDA receptor, we achieved our study on a series of 50 molecules that have been synthesized and evaluated for their ability to displace MK801from its specific binding site on rat cortical membranes (Ki) and for their antagonist activity to the NMDA receptor, as demonstrated by Thompson et al. [4]. 44 molecules are selected to propose the quantitative model (training set), and 6 compounds that have been selected randomly, have served to test the performance of the proposed model (test set). All these MK801 derivative molecules (Figure 1) are described by their substituents R1, R2, R3, R4, R5, R6 and R7, (Table 1 and Table 5). In reality, Thompson *et al.* proposed 73 compounds. The remained compounds have structures different of that required for this study (Figure 1).

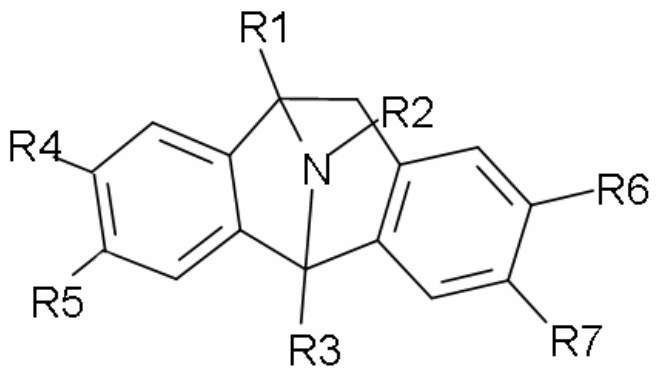

Figure 1. The general structure of studied compounds' series (MK801 derivatives).

In this work the activity will be expressed in the logarithmic form (logKi). The pharmacological activity Ki has been expressed in μM [4]. In the cases where Ki has been given in an interval, we have retained the minimum value. Values of logKi are divided into three parts, the activity is considered raised, when logKi ranged from -2 to -1, average when logKi ranged from -1 to 0.114, and weak when logKi is > 0.114. The study that we have achieved consists of a principal components analysis (PCA) with the aid of a software called STATLAB.2 [14], a multiple regression analysis (MRA) available in a software called SYSTATW5 [15], and a neural network available in a software called MATLAB [16].

To describe molecules we have chosen the properties that could have a role in the interaction of a molecule with the site receptor. For example the size can be a determinant factor of the activity for a molecule; in fact if the size of a molecule is not suitable for the site receptor the interaction with the receptor becomes impossible. The hydrogen bonding and the electronegativity are also important for the activity because to interact with a receptor, a hydrogen bonding can be formed. An interaction between phenyls is also possible using the π electrons of phenyls. The lipophlicity as it is known is an important property for the activity.

The physico-chemical parameters used then to describe molecules are as follow: the molecular weight (MW), the Van der Waals volume (VW) [17], the substituent length (L), electronic parameters as the electronegativity (EN) [18], the molar refraction (MR) [19], hydrogen bonding acceptor (HBA) [20] and donor (HBD) [20], the lipophilicity represented by fragmental constants (fi) [21-23], and finally the functionality Fpi [24,25]. These 9 physico-chemical parameters were calculated for the 7 substituents (R1,…,R7), for each molecule (63 descriptors). Five descriptors, Fpi(R1, R2, R4, R6, and R7), having a singular matrix (all their values are null or, are all the same), were eliminated. Overall, 58 descriptors were exploited to study and explain the structure-activity relation.

Table 1. The chemical structure of the studied compounds, the values of observed logKi (logKi-obs) corresponding to reference [4], and the values of predicted logKi (logKi-mra and logKi-nn) calculated using MRA and NN respectively.

N°	R1	R2	R3	R4	R5	R6	R7	ki. μM	logKi-obs	logKi-mra	logKi-nn
1	H	CH_3	CH_3	H	H	H	H	0.610	-0.215	-0.004	-0.219
2	H	H	CH_3	H	H	H	H	0.056	-1.252	-1.605	-1.232
3	CH_3	CH_3	CH_3	H	H	H	H	0.710	-0.149	-0.210	-0.138
4	H	OH	CH_3	H	H	H	H	19.000	1.279	1.411	1.279
5	H	H	CH_2CH_3	H	H	H	H	0.045	-1.347	-1.299	-1.200
6	H	CH_2CH_3	H	H	H	H	H	24.000	1.380	1.313	1.375
7	OH	OH	CH_2CO_2ET	H	H	H	H	4500.000	3.653	3.525	3.653
8	OH	H	CH_2CO_2ET	H	H	H	H	3.600	0.556	0.509	0.735
9	H	H	CH_2CH_2OH	H	H	H	H	0.260	-0.585	-0.850	-0.485
10	H	H	CH_2CO_2ET	H	H	H	H	0.550	-0.260	-0.263	-0.637
11	H	H	$CH(OH)CH_2OH$	H	H	H	H	0.320	-0.495	0.288	-0.118
12	H	H	CH_2OH	H	H	H	H	0.350	-0.456	-1.143	-0.580
13	H	H	CH_2F	H	H	H	H	0.160	-0.796	-0.882	-0.590
14	H	H	CH_2CH_2F	H	H	H	H	0.174	-0.759	-0.583	-1.005
15	H	H	$CH_2SC_6H_5$	H	H	H	H	73.000	1.863	2.107	1.778
16	H	H	$CH_2S(O)C_6H_5$	H	H	H	H	160.000	2.204	2.101	2.492
17	OH	H	CH_3	H	H	H	H	0.077	-1.114	-0.833	-0.944
18	F	H	CH_3	H	H	H	H	0.930	-0.032	-0.195	0.017
19	H	H	$CH=CH_2$	H	H	H	H	0.087	-1.060	-1.066	-1.167
20	OH	H	CH_2CO_2H	H	H	H	H	280.000	2.447	2.340	2.377
21	OH	H	CH_2CONH_2	H	H	H	H	6.400	0.806	0.892	0.506
22	Cl	H	CH_2CO_2H	H	H	H	H	4500.000	3.653	3.710	3.594
23	Cl	H	CH_2CONH_2	H	H	H	H	90.000	1.954	2.262	2.032
24	H	H	$CH(OH)CO_2H$	H	H	H	H	1000.000	3.000	2.694	2.859
25	Cl	H	CH_2CH_2Cl	H	H	H	H	53.000	1.724	1.606	1.671
26	H	H	CH_3	H	H	Cl	H	0.084	-1.076	-1.260	-1.407
27	H	H	CH_3	H	Cl	H	H	0.011	-1.959	-1.605	-1.232
28	H	H	CH_3	H	H	Br	H	0.180	-0.745	-0.815	-0.745
29	H	H	CH_3	H	H	OCH_3	H	0.036	-1.444	-1.305	-1.407
30	H	H	CH_3	H	H	OH	H	0.023	-1.638	-1.445	-1.407
31	H	H	CH_3	H	NH_2	H	H	0.027	-1.569	-1.485	-1.407
32	H	H	CH_3	H	Br	H	H	0.080	-1.097	-1.605	-1.232
33	H	H	CH_3	H	I	H	H	0.011	-1.959	-1.605	-1.232
34	H	H	CH_3	H	OCH_3	H	H	0.046	-1.337	-0.861	-1.406
35	H	H	CH_3	H	OH	H	H	0.018	-1.745	-1.637	-1.407
36	H	H	CH_3	H	CH_2OH	H	H	0.137	-0.863	-1.113	-1.406
37	H	H	CH_3	H	CH_3	H	H	0.034	-1.469	-1.233	-1.407
38	H	H	CH_3	H	$(CH_2)_3CH_3$	H	H	1.250	0.097	-0.117	-0.699
39	H	H	CH_3	H	C_6H_5	H	H	0.032	-1.495	-1.491	-0.699
40	H	H	CH_3	OCH_3	H	H	H	0.610	-0.215	-0.272	-1.232
41	H	H	CH_3	H	H	H	OCH_3	0.033	-1.481	-1.605	-1.232
42	H	H	CH_3	OH	H	H	H	0.277	-0.558	-0.487	-1.232
43	H	H	CH_3	H	H	H	OH	0.049	-1.310	-1.605	-1.232
44	H	H	CH_3	H	F	F	H	0.031	-1.509	-1.425	-1.407

Results and Discussion

Principal Components Analysis

The totality of the 58 descriptors (variables) coding the 44 molecules was submitted to a principal components analysis (PCA). Thirty-five principal components were obtained. The first three axes F1, F2, and F3 contributing respectively 25.4%, 15.7% and 13.29% to the total variance, were sufficient to describe the information represented by the data set. Table 2 shows the descriptor's contributions to F1, F2, and F3. Except for VW(R2) and MR(R2), which do not contributed to any principal component, all descriptors contribute to F1. The descriptors of R3, R1, and R5 have the most significant contributions to F1. On the other hand, there was no significant difference between descriptors' contributions to F2 and F3 except for the descriptors of R5 that contribute to F3 but not to F2.

Table2. The descriptor's contributions to the first three principal components F1, F2, and F3. The contributions are classified in increasing order for F1.

Descriptor	F1	F 2	F 3	Descriptor	F 1	F 2	F 3
EN(R3)	0.073	0.000	0.000	*Fi(R5)*	0.007	0.000	0.022
L(R3)	0.069	0.000	0.000	*Fpi(R5)*	0.006	0.000	0.025
MW(R3)	0.069	0.000	0.000	*MR(R6)*	0.006	0.000	0.000
VW(R3)	0.063	0.000	0.000	*MW(R6)*	0.006	0.000	0.001
HBA(R3)	0.062	0.000	0.000	*HBA(R7)*	0.006	0.090	0.042
MR(R3)	0.056	0.000	0.000	*HBA(R4)*	0.006	0.067	0.085
EN(R1)	0.044	0.000	0.000	*EN(R7)*	0.005	0.090	0.042
HBA(R1)	0.043	0.000	0.000	*EN(R4)*	0.005	0.067	0.065
VW(R1)	0.037	0.000	0.000	*MW(R7)*	0.005	0.089	0.041
MW(R1)	0.035	0.000	0.000	*MW(R4)*	0.005	0.066	0.085
Fi(R1)	0.035	0.000	0.000	*L(R7)*	0.005	0.088	0.041
HBD(R1)	0.034	0.000	0.000	*L(R4)*	0.005	0.065	0.064
L(R1)	0.033	0.000	0.000	*VW(R7)*	0.005	0.083	0.038
Fpi(R3)	0.030	0.000	0.000	*VW(R4)*	0.005	0.061	0.060
MR(R1)	0.027	0.000	0.000	*MR(R7)*	0.005	0.079	0.037
EN(R5)	0.021	0.000	0.055	*MR(R4)*	0.005	0.058	0.057
VW(R5)	0.020	0.000	0.055	*L(R6)*	0.004	0.000	0.000
MR(R5)	0.017	0.000	0.055	*HBD(R5)*	0.003	0.000	0.004
HBD(R3)	0.017	0.000	0.000	*HBD(R7)*	0.002	0.030	0.014
MW(R5)	0.012	0.000	0.038	*HBD(R4)*	0.002	0.022	0.022
L(R5)	0.008	0.000	0.046	*EN(R2)*	0.002	0.000	0.000
Fi(R3)	0.008	0.000	0.000	*Fi(R7)*	0.002	0.024	0.011
HBA(R5)	0.008	0.000	0.016	*Fi(R4)*	0.002	0.017	0.017
EN(R6)	0.008	0.000	0.001	*Fi(R6)*	0.002	0.000	0.000
HBD(R2)	0.008	0.000	0.000	*HBD(R6)*	0.001	0.000	0.000
HBA(R2)	0.008	0.000	0.000	*MW(R2)*	0.001	0.000	0.000
VW(R6)	0.008	0.000	0.001	*L(R2)*	0.001	0.000	0.000
Fi(R2)	0.007	0.000	0.000	*VW(R2)*	0.000	0.000	0.000
HBA(R6)	0.007	0.000	0.001	*MR(R2)*	0.000	0.000	0.000

In the projection of the compounds in the plane of the two first axes F1 and F2 (Figure 2), the compounds are distributed in three regions. Region 1 contains compounds having a higher activity, logKi is included in [-2,-1]. Region 2 contains those having an average activity, logKi is included in [-1,0.114]. In region 3 the activity is weak, logKi > 0.0114. The projection of the compounds on the first and the third principal components F1-F3, did not add any further or significant information.

In examining the descriptors evolution in the three regions, we noticed that the size, the electronegativity and the lipophilicity governed, with a great precision, the distribution of the compounds in each region. Indeed, in region 1 are situated the compounds whose L(R3) lies between 2 and 3. In this region the activity is higher when EN(R3) is included between 2.54 and 5.62 and fi(R3) is included between 0.7 and 2.56 as for compounds **2** and **5** having respectively R3=CH3 and R3=CH2-CH3 (Table 1). Otherwise, when fi(R3)<0.7 the activity is average (region 2), and when fi(R3)<0.54 the activity becomes weak (region 3). Furthermore, the activity seems to be linked also to the electronic nature of the radical R5. Thus when EN(R5)>0 and HBD(R5)>0, the activity is higher and when EN(R5)=2.54 and HBD(R5)=0 the activity declines. Similarly, for the radical R2, the size,

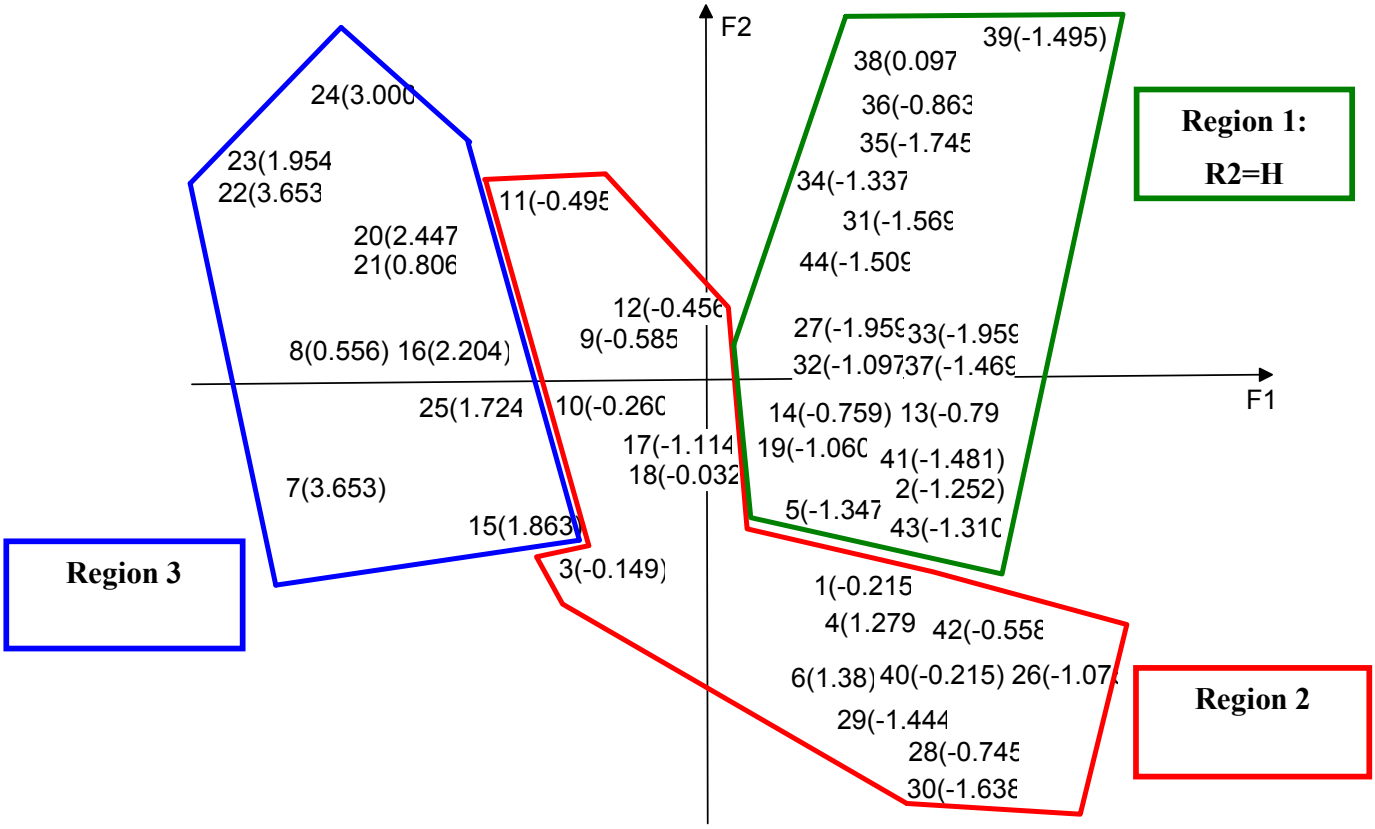

Figure 2. The graphic projection of the points representing MK801 derivatives on the first two principal components F1-F2. logKi is given into brackets. Region 1, contains compounds having higher logKi. Those having average ones are in region 2 and the lower ones are in region 3

notably the length, and the electronegativity play a very important role on the activity. So, in region 1, are situated only compounds having L(R2)=1 and EN(R2)=0, this means that R2 would be a monovalent and non electronegative atom, it is only the hydrogen atom that satisfy these criterions (R2=H). However all compounds having L(R2)>1 and EN(R2)>0 (R2≠H), are situated in regions 2 and 3 as for compound **4** in which R3=CH3 is satisfying for the activity but R2=OH is not, so with an electronegative group in this position a molecule looses its activity.

In attempting to analyze these results we conclude that to predict activity for any molecule, we must respect some physico-chemical properties of the substituents. The size, the electronegativity and the lipophilicity of R3 seem to be linked to each other. Therefore, in order to obtain increased activity for a molecule, we must satisfy, at the same time, the three descriptor's norms and respect their values as it is shown above. In Table 2 we have noted that the descriptors of R3 have the largest contributions, so it is reasonable to conclude that R3's descriptors have the largest impact on the prediction of the activity. The same, we have to respect the electronic nature of R5, and finally, the nitrogen substituent's structure (R2), determined by the length and the electronegativity of this substituent. So when R2=H the activity is raised and when R2≠H the activity declines. As Thomson et al [4] have concluded, the substitution on the ring nitrogen, is not tolerated. We note that the nitrogen atom is considered as an essential element for the activity in this receptor, it interacts with the receptor with the formation of hydrogen bonding [13].

Multiple Regression Analysis

In order to propose a mathematical model and to evaluate quantitatively the substituent's physico-chemical effects on the pharmacological activity of the totality of the set of these 44 molecules, we submitted the data matrix constituted obviously from the 58 physico-chemical variables corresponding to the different substituents and the 44 molecules, to a progressive multiple regression analysis. This method used the coefficients r, r^2, and the t-values to select the best regression performance. The best results were obtained with 14 descriptors MW(R1), MW(R6), VW(R1), VW(R2), L(R3), L(R5), EN(R2), EN(R3), HBD(R3), HBD(R5), HBA(R4), Fpi(R3), Fpi(R5), fi(R3).

It results then in the following equation:

LogKi = -0.915 + 0.099MW(R1) + 0.010MW(R6) - 0.155VW(R1) - 0.091VW(R2) - 0.744L(R3) + 0.372L(R5) + 0.998EN(R2) + 0.285EN(R3) + 2.381HBD(R3) - 0.505HBD(R5) + 0.860HBA(R4) + 0.132Fpi(R3) - 0.229Fpi(R5) + 0.614fi(R3)

n = 44 r = 0.986 s = 0.317 (1)

The t-values are shown in Table 3. The values of predicted logKi (logKi-mra) calculated from equation (1), and observed logKi values (logKi-obs) [4] are given in Table 1. The correlation of predicted logKi and observed logKi are illustrated in figure 3.

Table 3. t-values of regression equation (1) (ratio of the parameter's regression coefficient and the standard error)

Descriptor	Coefficient	Standard error	t-value
MW(R1)	0.099	0.014	7.280
MW(R6)	0.010	0.004	2.753
VW(R1)	-0.155	0.046	-3.337
VW(R2)	-0.091	0.036	-2.541
L(R3)	-0.744	0.180	-4.135
L(R5)	0.372	0.073	5.113
EN(R2)	0.998	0.115	8.665
EN(R3)	0.285	0.049	5.801
HBD(R3)	2.381	0.293	8.130
HBD(R5)	-0.505	0.335	-1.506
HBA(R4)	0.860	0.168	5.132
Fpi(R3)	0.132	0.063	2.091
Fpi(R5)	-0.229	0.069	-3.311
Fi(R3)	0.614	0.125	4.911

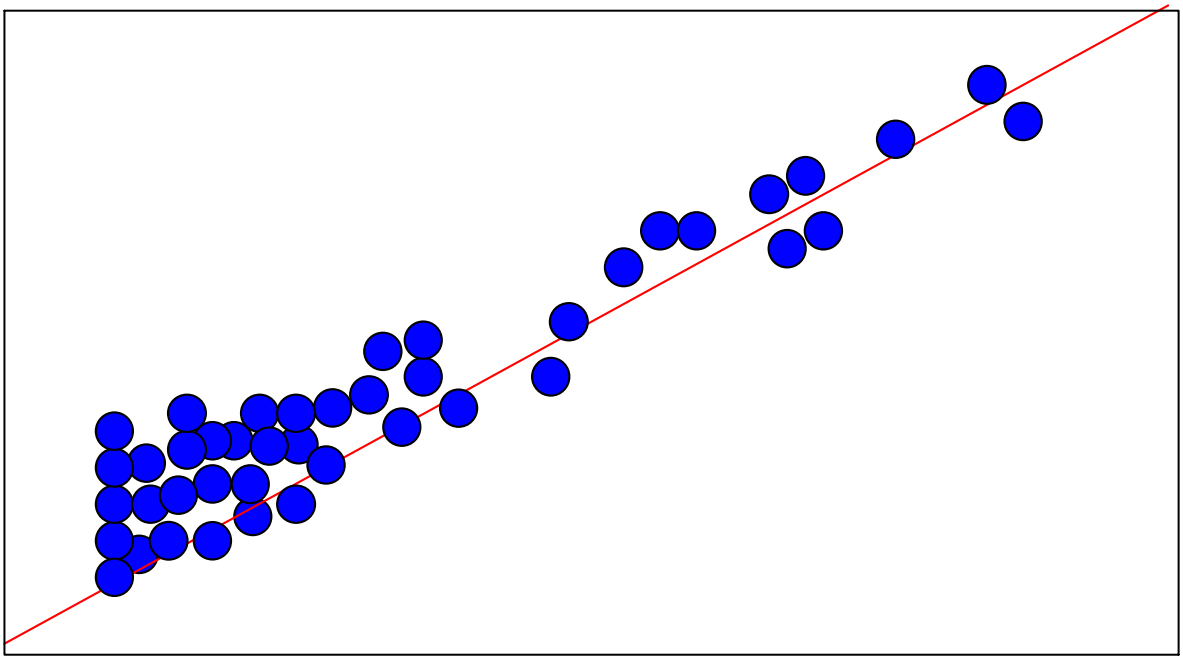

Figure 3. The correlation of predicted logKi (logKi-mra) calculated using MRA and observed logKi (logKi-obs).

The correlation coefficient is r = 0.986, the square r^2 = 0.972, and the standard error s = 0.317. These values are relevant to evaluate the quantitative model. In equation (1) we noticed an important contribution of R3, R5 and R2, whose regression coefficients are clearly raised when compared to the other radicals. Electronic parameters and the size of the radical R3 ; notably the factor of hydrogen bonding donor HBD(R3), its length L(R3), and the lipophilicity fi(R3) seem to play a very important role in this model. Similarly, for the radicals R2 and R5 electronic factors, HBD (R5) and electronegativity EN(R2) have a great impact on activity. This confirms very clearly the notable participation of these variables in the distribution of the compounds on F1-F2 plane, and the interpretations made with the aid of the PCA. The descriptors proposed in equation (1) by MRA were, therefore, used as the input parameters in NN.

Neural Network

Despite the good results obtained by the Multiple Linear Regression Analysis, notably the good correlation coefficient and the best predictions of the logKi shown in figure 3, it is always probable that a non linear relation may occur. The Neural Network (NN) is a suitable concept to achieve this goal. Several studies of QSAR have indeed been fulfilled using NN[26-29].

Training

In this work, we submitted the training set to a feed-forward network with three layers and complete connections between neurons. The input layer is constituted by the 14 descriptors proposed in equation (1), the hidden layer is selected with 2 tansig neurons, and the output layer is a linear neuron (14-2-1). The correlation coefficient obtained is 0.995. The number of nodes in the hidden layer is an important factor determining the network's performance. It was found that too many nodes cause the network to memorize the data set (overfitting). However, networks with few nodes may be insufficient to use all the information of the data set (underfitting) and generalization is poor. Previous studies conducted to determine the appropriate number of hidden units suggest that ρ, the ratio of the number of data points to the number of adjustable weights in the neural network, should have a value between 1.8 and 2.3 [27,30,31]. For a network with a 14-2-1 configuration, the number of weights is 33, therefore ρ = 1.33, far from the optimal values. In the attempt to propose a model with less descriptors than proposed by MRA, the elimination of five descriptors (HBD(R3), HBD(R5), HBA(R4), Fpi(R3), Fpi(R5)) allowed to an acceptable network with a 9-2-1 architecture, ρ = 1.9. In examining the data matrix of these descriptors we found that the majority of their values are null, so we expect that their elimination will not have a concrete impact on the model.

So, the final model was proposed with 9 descriptors (MW(R1), MW(R6), VW(R1), VW(R2), L(R3), L(R5), EN(R2), EN(R3), fi(R3)). Values of logKi calculated with 9-2-1 network (logKi-nn) are given in table 1 and the correlation obtained is illustrated in figure 4. The correlation coefficient is r = 0.974, the square r^2 = 0.949, the standard error s = 0.226.

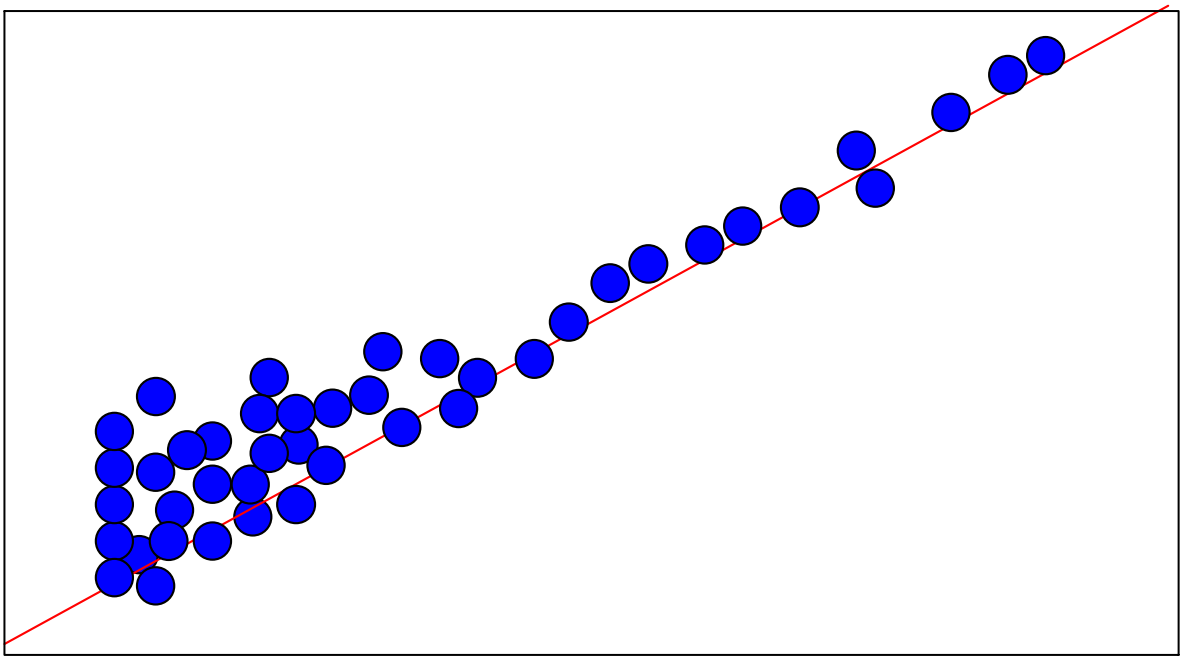

Figure 4. The correlation of predicted logKi (logKi-nn) calculated using NN and observed logKi (logKi-obs).

In comparison with MRA method, the NN method allowed us to propose an acceptable model with less descriptors than proposed with MRA (there is not an overfitting), with good correlation coefficient and good standard deviation.

As we have showed above, the elimination of 5 descriptors have not a concrete impact on the model proposed and the correlation coefficient remained good, so we can conclude that these 5 descriptors (HBD(R3), HBD(R5), HBA(R4), Fpi(R3), Fpi(R5)) are not really pertinent descriptors. In the other hand the attempts done with the elimination of one of the 9 descriptors finally proposed, lead to very law correlation coefficients, so with these results we conclude that these 9 descriptors (MW(R1), MW(R6), VW(R1), VW(R2), L(R3), L(R5), EN(R2), EN(R3), fi(R3)) are the really pertinent descriptors.

Test

We have used the cross validation method with 'leave one out' procedure [32], for the aim of testing the performance of the NN and the validity of the choice of our descriptors. The calculated logKi values (logKi-cv) are given in table 4, and the correlation obtained is illustrated in figure 5. The cross validation coefficient is $r_{cv} = 0.926$, the square $r^2 = 0.857$, the standard error $s = 0.244$.

In the aim to test the predictive power of our model, the activities of the remained set of 6 compounds (test set) are deduced from the quantitative model proposed with the 44 molecules (training set) by NN. Their structures and the observed and calculated logKi values (logKi-obs and logKi-test) are given in table 5.

Table 4. Values of observed logKi (logKi-obs) corresponding to reference [4], and values of predicted logKi (logKi-cv) calculated using the cross validation method with 'leave one out' procedure.

N°	logKi-obs	logKi-cv	N°	logKi-obs	logKi-cv
1	-0.215	-0.304	**23**	1.954	3.307
2	-1.252	-1.334	**24**	3.000	2.857
3	-0.149	-0.093	**25**	1.724	0.600
4	1.279	1.283	**26**	-1.076	-1.143
5	-1.347	-1.177	**27**	-1.959	-1.295
6	1.380	1.825	**28**	-0.745	-0.652
7	3.653	3.912	**29**	-1.444	-1.236
8	0.556	2.012	**30**	-1.638	-1.374
9	-0.585	-0.354	**31**	-1.569	-1.302
10	-0.260	0.023	**32**	-1.097	-1.235
11	-0.495	-0.484	**33**	-1.959	-1.309
12	-0.456	-1.142	**34**	-1.337	-1.293
13	-0.796	-0.853	**35**	-1.745	-1.268
14	-0.759	-0.708	**36**	-0.863	-1.143
15	1.863	1.540	**37**	-1.469	-1.303
16	2.204	3.383	**38**	0.097	-1.225
17	-1.114	-1.261	**39**	-1.495	-0.172
18	-0.032	-1.057	**40**	-0.215	-1.277
19	-1.060	-1.174	**41**	-1.481	-1.349
20	2.447	2.913	**42**	-0.558	-1.205
21	0.806	0.117	**43**	-1.310	-1.265
22	3.653	3.065	**44**	-1.509	-1.408

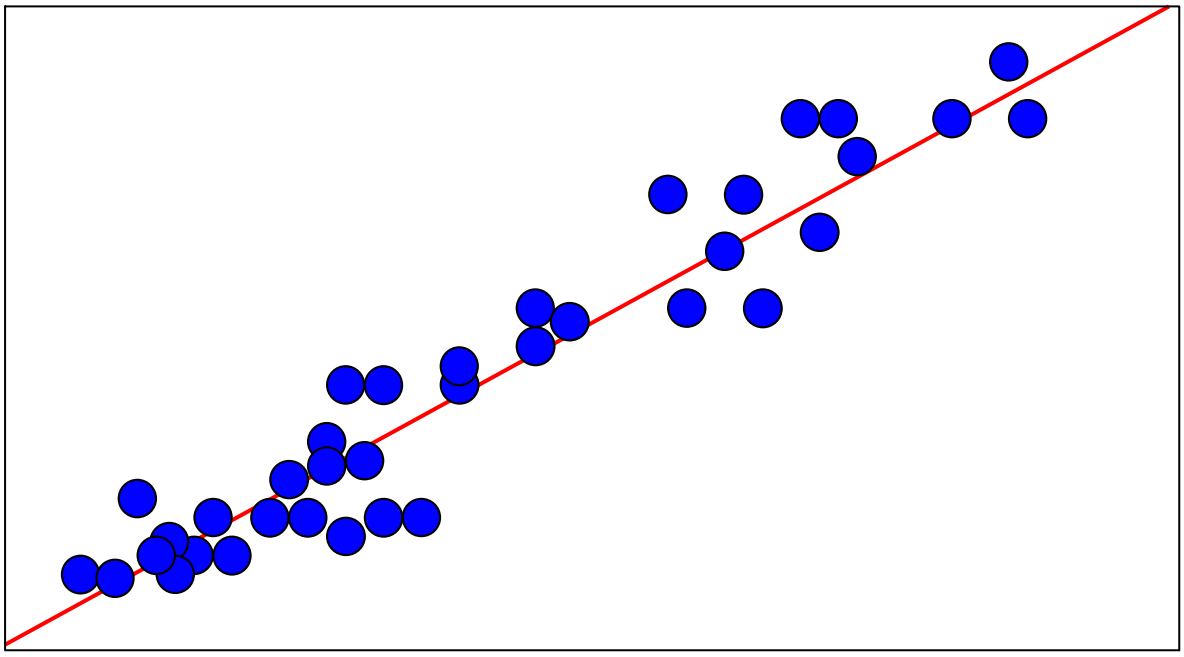

Figure 5. The correlation of predicted logKi (logKi-cv) Calculated using the cross validation method with 'leave one out' procedure and observed logKi (logKi-obs).

Table 5. The chemical structure of the 6 tested compounds (test set). Values of observed logKi (logKi-obs) corresponding to reference [4] and values of predicted logKi (logKi-test) of the 6 compounds, calculated using NN.

Compound N°	R1	R2	R3	R4	R5	R6	R7	Ki, μM	logKi-obs	logKi-test
45	H	CH_3	H	H	H	H	H	12.000	1.081	0.972
46	H	CH_2CH_2OH	H	H	H	H	H	71.000	1.854	2.066
47	H	H	$CH_2CH_2CH_3$	H	H	H	H	8.600	0.921	0.750
48	OH	H	CH_2CH_2OH	H	H	H	H	0.074	-1.328	-1.229
49	H	H	$CH(OH)CO_2ET$	H	H	H	H	0.390	-0.408	-0.547
50	Cl	H	CH_2CH_2OH	H	H	H	H	0.280	-0.553	-0.542

The comparison of the values of logki-test to logki-obs shows that a good prediction has been obtained for the 6 compounds. The good results obtained with the cross validation and with the prediction of activities of the 6 compounds, shows that the model proposed in this paper is able to predict activity with a great performance, and that the selected descriptors are pertinent.

The model was cross validated using "leave one out" procedure because we have a small set, so we have not fit only 6 point data but all the data points (44 compounds) are fitted with the cross validation method. With this method we have deduced the activity of each compound of the 44 compounds and we have obtained a good correlation (r=0.926). This result is very sufficient to conclude the performance of the model. The test done with the 6 compounds confirms the performance of the model. Even if it is possible that this good prediction is found by chance (especially that with only 6 compounds) we can claim that it is a positive result and it is a further confirmation of the results found by the cross validation. So, this model could be applied to all MK801 derivatives accordingly to figure 1 and could add further knowledge in the improvement of the search in the domain of non competitive antagonists of NMDA receptor and their interaction with the receptor.

Conclusion

The statistical analysis that we have undertaken to establish a structure-activity relationship for the antagonists of the NMDA receptor, showed that the activity of the MK801 derivatives is closely linked to the physico-chemical descriptors of radical R3, R5, and R2. Thus the size, the electronegativity, and the lipophilicity are estimated as relevant factors for this model. These descriptors selected automatically by MRA showed a high correlation of predicted logKi calculated by NN and observed ones. The test of the performance of this quantitative relationship, by the cross validation method (r = 0.926) and the prediction of activities of 6 compounds (r = 0.992), showed that the model proposed in this paper is able to predict activity with a great performance, and that the selected descriptors are pertinent.

References

1. Olney, J.W. *Annu. Rev. Pharmacol*. **1990**, *30*, 47.
2. Meldrum, B.S.; Garthwaite, J. *Trends Pharmacol. Sci*. **1990**, *11*, 379.
3. Andrianov, G.N.; Bretschneider, F.; Peters, R.C. *Neuroscience* **1997**, *79*, 1231.
4. Thompson, W. J.; Anderson, P. S.; et al. *J. Med. Chem*. **1990**, *33*, 789.
5. Kroemer, R. T.; Koutsilieri, E.; Hecht, P.; Liedl, K.R.; Riederer, P.; Kornhuber, J. *J. Med. Chem*. **1998**, *41*, 393.
6. Cone, E. J.; Mcquinn, R.L.; Shannon, H.E.; *J. Pharmacol*. Exp. Ther. **1983**, *228*, 147.
7. Lloyd, E.J.; Andrews, P.R. *J. Med. Chem*. **1986**, *29*, 453.
8. Andrews, P.R.; Lloyd, E.J. *Prog. Med. Chem*. **1986**, *23*, 91.
9. Andrews, P.R.; Lloyd, E.J.; Martin, J.L.; Munro, S.L.A. *J. Mol. Graphics* **1986**, *4*, 41.
10. Wong, M.G.; Defina, J.A.; Andrews, P.R. *J. Med. Chem*. **1986**, *29*, 562.
11. Laguerre, M.; Elhallaoui, M.; Carpy, A. *Molecular modeling of noncompetitive antagonists of NMDA receptor*; In: *Trends in QSAR and Molecular Modeling 92 ESCOM*; Wermuth, C.G., ed., 1992, pp. 444-445.
12. Elhallaoui, M. *Modélisation moléculaire d'antagonistes non compétitifs du récepteur NMDA*. Doctoral dissertation, Bordeaux II, France, 1994.
13. Elhallaoui, M.; Carpy, A.; Laguerre, M.; Ouazzani, C.F. *J. Mol. Model.* **2002**, *8*, 65-72.
14. STATLAB ^{TM}, Logiciel d'Analyse statistique de données, version 2, développée par DELBEC T., JAMBU M. and all, 1992.
15. SYSTAT version 5.02 for windows by SYSTAT, INC., 1800 sherman Ave. Evanston, IL USA 60201.
16. MATLAB, the language of technical computing, version 5.2.0.3084, January 17, 1998.
17. Bondi A. *J. Phys. Chem*. **1964**, *68*, 441.
18. Pauling L. *The Nature Of Chemical Bond*, Third Edition, Cornell University Press, 1960, p. 85.
19. Hanch, C.; Leo, A.; Unger, S.H.; Kim, K.H.; Nikaitani, D.; Lien, E.J. *J. Med. Chem.* **1973**, *16*, 1207.
20. Chastrette, M.; Zakarya, D.; Elmouaffek, A. *Eur. J. Med. Chem. –Chim. Ther*. **1986**, *21*, 505.
21. Nys, G.G.; Rekker, R.F. *Chim. Ther*. **1973**, *8*, 521.
22. Nys, G.G.; Rekker, R.F. *Eur. J. Med. Chem*. **1974**, *9*, 361.
23. Rekker, R.F. *The Hydrophobic Fragmental Constant*, Elseiver: Amsterdam, **1977**, 31.
24. Broto, P.; Moreau, G.; Vandycke, C. *Eur. J. Med. Chem. –Chim. Ther*. **1984**, *19*, 61.
25. Broto, P., Moreau, G.; Vandycke, C. *Eur. J. Med. Chem. –Chim. Ther*. **1984**, *19*, 66.
26. Breindl, A.; Beck, B.; Clark, T.; Robert, C.; Glen, R.C. *J. Mol. Model*. **1997**, *3*, 142.
27. So, S.-S.; Richards, W. G. *J. Med. Chem*. **1995**, *35*, 3201.

28. Oinuma, H.; Miyako, K.; Yamanaka, M.; Nomoto, K.I.; Katoh, H.; Sawada, K.; Shino, M.; Mamano, S. *J. Med. Chem*. **1990**, *33*, 905.
29. Aoyama, T.; Suzuki, Y.; Ichikawa, H. *J. Med. Chem*. **1990**, *33*, 2583.
30. Lopez-Rodriguez, M.L.; Morcillo, M.J.; Fernandez, E.; Rosado, M.L.; Pardo, L.; Schaper, K.J. *J. Med. Chem*. **2001**, *44(2)*, 186-197.
31. Andrea, T.A.; Kalayeh, H. *J. Med. Chem*. **1991**, *34*, 2824-2836.
32. Cramer, R. D., III; Bunce, J. D.; Patterson, D. E. *Quant. Struct. -Act. Relat*. **1988**, *7*, 18.