Journal of Computational Science & Engineering 1 (2011) 48-54

Journal of Computational Science & Engineering
Available at http://icse.qdu.edu.cn ISSN 1710-4068

# Quantitative structure-activity relationship study for 2, 3-Benzodiazepin-4 (thi) one- and 1, 2-phthalazine-related negative allosteric modulators of AMPA receptor by heuristic and support vector machine method

Yuming Dong[1*], Chenjun Liu[1] and Hongzong Si[2*]

[1]*School of Pharmacy, Lanzhou University, Lanzhou 730000, China;*
[2]*Institute for Computational Science and Engineering Qingdao University, Qingdao, Shandong 266071, China*

Article Information

Article history:
Received 24 December 2010
Revised 6 January 2011
Accepted 18 January 2011
Available online 25 January 2011

**Keywords:**
Support vector machine
Benzodiazepines
Quantitative structure-property
Negative allosteric modulators
HM

**Abstract**

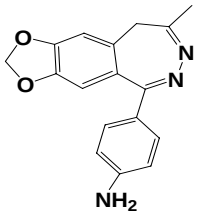
A set of sixty-one 2,3-Benzodiazepin-4 (thi) one- and 1, 2-phthalazine-related negative allosteric modulators of AMPA receptor was modeled using the descriptors, which were calculated from the molecular structure alone with a quantitative structure-activity relationship technique. The heuristic method and support vector machine were utilized to construct the linear and nonlinear prediction models, lead to correlation coefficients were 0.7930 and 0.7054, respectively. The specific information described by the heuristic linear model could provide some insights into the factors be used drug design process, the prediction results of the linear heuristic method model seemed well than those of support vector machine. Quantitative structure-activity relationship can be used to predict the ED50 of the compounds.

## 1. Introduction

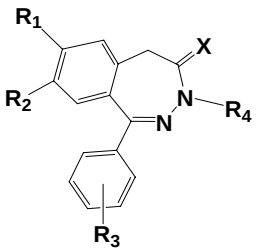
More recently, Chimirri and co-workers have reported a quantitative structure-activity relationship (QSAR) on a set of 61 negative allosteric modulators of a-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptor by partial least-squares (PLS) analysis [1]. Furthermore, Buchward investigated the QSAR of the 61 compounds based on Chimirri's study by structural descriptors-based reassessment method and discussed the Chimirri's study in detail [2]. AMPA receptors, together with N-methyl-D-aspartate (NMDA) and kainicacid (KA) receptors are the major ionotropic glutamate receptor subtypes, and ligands of the AMPA receptor are of considerable therapeutic interest [3]. The Chimirri and co-workers study was based on in vivo anticonvulsant activity as measured by the median effective doses (ED50) required to prevent clonic and tonic phases of audiogenic seizures in DBA/2 mice for 2, 3-benzodiazepin-4 (thi) ones (2-35), cyclofunctionalized 2,3-benzodiazepines (36-45), 1, 2-phthalazines (46-56), and isoquinolines (57-61) that mostly have been synthesized and tested in their own laboratories (Figure 1, Table 1). In an attempt to

* Corresponding author. E-mail: *dongym@lzu.edu.cn*

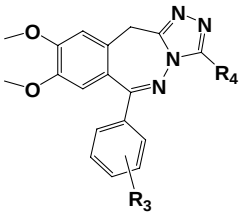
obtain a statistically more stable and chemically more meaningful quantitative structure-activity relationship, Buchward undertook a reassessment of the factors influencing anticonvulsant activity in this family of AMPA antagonists by using additional quantum chemical and more carefully selected structure-derived descriptors. Although Buchward's investigation improved prediction and intuitive quantitative interpretation of this set of allosteric AMPA, the method was still multiple linear regression.



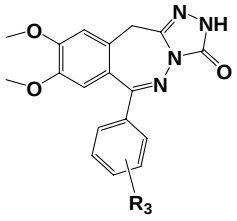
(1)



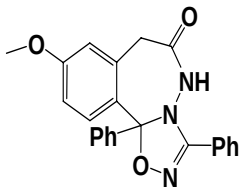
(2-35)



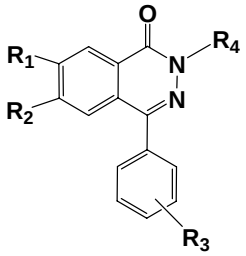
(36-38)



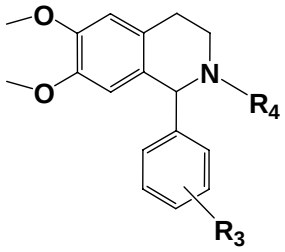
(39-44)



(45)



(46-56)



(57-61)

Figure 1. Structures of the 61 negative allosteric modulators of AMPA receptor. Substitute R was list in Table 1.

Heuristic methods provide a way to approach difficult combinatorial optimization problems. Combinatorial search gives us a method to construct possible solutions and find the best one, given a function that measures how good each candidate solution is. However, there may be no algorithm to find the best solution short of searching all configurations.

The foundation Support Vector Machine (SVM) was developed by Vapnik, and is gaining popularity due to its many attractive features and promising empirical performance [4, 5]. Compared with traditional neural networks, SVM possesses prominent advantages: (1) strong theoretical background provides SVM with high

generalization capability and can avoid local minima. (2) SVM always has a solution, which can be quickly obtained by a standard algorithm (quadratic programming). (3) SVM need not determine network topology in advance, which can be automatically obtained when training process ends. (4) SVM builds a result based on a sparse subset of training samples, which reduce the workload [6].

Originally, SVM was developed for pattern recognition problems. Now, with the introduction of ε-insensitive loss function, SVM has been extended to solve nonlinear regression problems and time series prediction [7]. SVM can be applied to regression problems by the introduction of an alternative loss function.

The purpose of this work is establishing a model to predict the ED50 of the analogs of this family of AMPA antagonists by SVM and HM and compared the results of the two methods for the first time.

Table 1 Structures, observed anticonvulsant activities (expressed as ln 1/ED50)

| NO | $R^1$ | $R^2$ | $R^3$ | $R^4$ | X | lnED50 |
|---|---|---|---|---|---|---|
| 1 | | | | | | -3.58 |
| 2 | MeO | H | H | H | O | -4.32 |
| 3 | MeO | H | H | Me | O | -4.76 |
| 4 | MeO | MeO | H | H | O | -3.52 |
| 5 | MeO | MeO | H | Me | O | -3.63 |
| 6 | MeO | MeO | H | Ac | O | -4.62 |
| 7 | $-OCH_2O-$ | | H | H | O | -3.77 |
| 8 | $-OCH_2O-$ | | H | Me | O | -4.41 |
| 9 | $-OCH_2O-$ | | H | H | S | -3.17 |
| 10 | MeO | MeO | H | H | S | -2.98 |
| 11 | MeO | MeO | H | Me | S | -3.42 |
| 12 | MeO | MeO | 4-F | H | O | -4.36 |
| 13 | $-OCH_2O-$ | | 4-F | H | O | -4.5 |
| 14 | MeO | MeO | 4-Cl | H | O | -4.63 |
| 15 | MeO | MeO | 4-Cl | Me | O | -3.72 |
| 16 | MeO | MeO | 4-Cl | H | S | -4.41 |
| 17 | MeO | MeO | 4-Br | H | O | -4.7 |
| 18 | MeO | MeO | 4-Br | Me | O | -4.14 |
| 19 | MeO | MeO | 4-Br | H | S | -3.95 |
| 20 | MeO | MeO | 4-Br | Me | S | -4.65 |
| 21 | $-OCH_2O-$ | | 4-OH | H | O | -3.96 |
| 22 | MeO | MeO | $3-NH_2$ | H | O | -2.96 |
| 23 | MeO | MeO | $3-NH_2$ | Me | O | -3.61 |
| 24 | MeO | MeO | $3-NH_2$ | H | S | -2.93 |
| 25 | $-OCH_2O-$ | | $3-NH_2$ | H | O | -2.89 |
| 26 | MeO | MeO | $4-NH_2$ | H | O | -2.71 |
| 27 | MeO | MeO | $4-NH_2$ | Me | O | -3.92 |
| 28 | MeO | MeO | $4-NH_2$ | H | S | -1.84 |
| 29 | MeO | MeO | $4-NH_2$ | Me | S | -3.39 |
| 30 | $-OCH_2O-$ | | $4-NH_2$ | H | O | -2.73 |
| 31 | $-OCH_2O-$ | | $4-NH_2$ | Me | O | -3.72 |
| 32 | $-OCH_2O-$ | | $4-NH_2$ | CONHMe | O | -2.52 |
| 33 | MeO | MeO | $4-NO_2$ | H | S | -4.54 |
| 34 | MeO | MeO | $4-NO_2$ | Me | S | -4.63 |
| 35 | MeO | MeO | $4-NAc_2$ | Ac | O | -4.03 |
| 36 | | | $4-NH_2$ | H | | -4.05 |
| 37 | | | $4-NH_2$ | Me | | -4.09 |
| 38 | | | $4-NO_2$ | Me | | -4.20 |
| 39 | | | H | | | -3.8 |
| 40 | | | 4-F | | | -3.47 |
| 41 | | | $3-NH_2$ | | | -3 |
| 42 | | | $4-NH_2$ | | | -2.78 |
| 43 | | | $3-NO_2$ | | | -4.03 |
| 44 | | | $4-NO_2$ | | | -4.31 |
| 45 | | | | | | -4.78 |
| 46 | $-OCH_2O-$ | | H | H | | -3.6 |
| 47 | MeO | MeO | $3-NH_2$ | H | | -4.22 |
| 48 | MeO | MeO | $3-NH_2$ | Me | | -4.15 |
| 49 | $-OCH_2O-$ | | $3-NH_2$ | H | | -3.15 |
| 50 | MeO | MeO | $4-NH_2$ | H | | -4.1 |
| 51 | $-OCH_2O-$ | | $4-NH_2$ | H | | -3.05 |
| 52 | $-OCH_2O-$ | | $4-NH_2$ | CONHMe | | -3.15 |
| 53 | $-OCH_2O-$ | | $4-NH_2$ | CONHnPr | | -2.76 |
| 54 | MeO | MeO | 4-NHAc | H | | -4.13 |
| 55 | MeO | MeO | $3-NO_2$ | H | | -4.4 |
| 56 | MeO | MeO | $3-NO_2$ | Me | | -4.7 |
| 57 | MeO | MeO | 4-Cl | H | | -3 |
| 58 | MeO | MeO | 4-Br | H | | -3.42 |
| 59 | MeO | MeO | $3-NO_2$ | H | | -2.96 |
| 60 | MeO | MeO | $4-NO_2$ | H | | -4.61 |
| 61 | MeO | MeO | $3-NO_2$ | Ac | | -3.62 |

## 2. Experiment section

### *2.1 Data set*

The data set of this investigation consisted of a set of 61compounds. The initial pool of substances is taken from the original paper [1]. A complete list of the compounds' structure and its ED50 was shown in Figure 1. and Table 1. The data set was randomly divided into two subsets: a training set of 46 compounds and a test set of 15 compounds. The training set was used to build HM and SVM models, and the test set was used to evaluate their prediction ability for both methods.

### *2.2 Calculation of the descriptors*

All molecules were drawn into Hyperchem [20] and pre-optimized using MM+ molecular mechanics force field. A more precise optimization was done with semi-empirical AM1 method in MOPAC [21]. The molecular structures were optimized using the Polak-Ribiere algorithm till the root mean square gradient was 0.01. The MOPAC output files were used in CODESSA program to calculate five classes of descriptors: constitutional (number of various types of atoms and bonds, number of rings, molecular weight, etc.); topological (Wiener index, Randic indices, Kier–Hall shape indices, etc.); geometrical (moments of inertia, molecular volume, molecular surface area, etc.); electrostatic (minimum and maximum partial charges, polarity parameter, charged partial surface area descriptors, etc.); and quantum chemical (reactivity indices, dipole moment, HOMO and LUMO energies, etc.) [22].

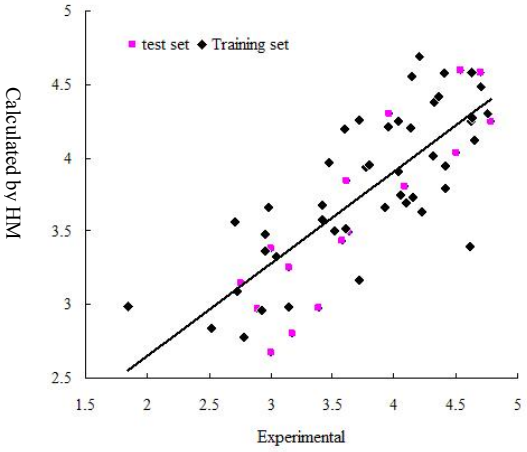



Figure 2 Predicted vs. experimental ED50 by HM

### *2.3 Development of linear model by heuristic method (HM) [22]*

Once molecular descriptors are generated, the heuristic method in CODESSA was used to accomplish the pre-selection of the descriptors and build the linear model. Its advantages of HM are the high speed and no software restrictions on the size of the data set. The heuristic method can either quickly give a good estimation about what quality of correlation to expect from the data, or derive several best regression models. Besides, it will demonstrate which descriptors have bad or missing values, which are insignificant (from the standpoint of a single-parameter correlation), and which are highly inter-correlated. This information will be helpful in reducing the number of descriptors involved in the search for the best QSAR/QSPR model. First of all, all descriptors are checked to ensure that values of each descriptor are available for each structure and that there is a variation in these values. Descriptors for which values are not available for every structure in the data in question are discarded. Descriptors having a constant value for all structures in the data set are also discarded. Thereafter all possible one-parameter regression models are tested and the insignificant descriptors are removed. As a next step, the program calculates the pair correlation matrix of descriptors and further reduces the descriptor pool by eliminating highly correlated descriptors.

Once molecular descriptors are generated, the heuristic method in CODESSA was used to accomplish the pre-selection of the descriptors and build the linear model. Its advantages of HM are the high speed and no software restrictions on the size of the data set. The heuristic method can either quickly give a good estimation about what quality of correlation to expect from the data, or derive several best regression models. Besides, it will demonstrate which descriptors have bad or missing values, which are insignificant (from the standpoint of a single-parameter correlation), and which are highly inter-correlated. This information will be helpful in reducing the number of descriptors involved in the search for the best QSAR/QSPR model. First of all, all descriptors are checked to ensure that values of each descriptor are available for each structure and that there is a variation in these values. Descriptors for which values are not available for every structure in the data in question are discarded. Descriptors having a constant value for all structures in the data set are also discarded. Thereafter all possible one-parameter regression models are tested and the insignificant descriptors are removed. As a next step, the program calculates the pair correlation matrix of

descriptors and further reduces the descriptor pool by eliminating highly correlated descriptors. The details of validating inter-correlation are (a) all quasi–orthogonal pairs of structural descriptors are selected from the initial set. Two descriptors are considered orthogonal if their inter-correlation coefficient rij is lower than 0.1. (b) CODESSA uses the pairs of orthogonal descriptors to compute the biparametric regression equations. The most significant 10 pairs of molecular descriptors are used in the third step. (c) To an MLR model containing n descriptors a new descriptor is added to generate a model with n+1 descriptors if the new descriptor is not significantly correlated with the previous n descriptors (inter-correlation coefficient lower than 0.8). Step (c) is repeated until MLR models with a prescribed number of descriptors are obtained. The goodness of the correlation is tested by the coefficient regression (R2), the F-test (F), and the standard deviation (s2). From the above processes, five descriptors were selected from 554 descriptors and the linear model was produced by HM. The heuristic method usually produces correlations 2–5 times faster than other methods with comparable quality.

Table 2. The involved parameters and the statistical of the HM model

| Descriptors | Cofficient | t-test |
|---|---|---|
| Intercept | 6.8686e+01 | 4.8377 |
| Min partial charge for a N atom[Zefirov's PC] | -2.4425e+01 | -8.1113 |
| Avg bond order of a O atom | -5.2267e+00 | -3.0150 |
| Max valency of a O atom | -2.4647e+01 | -4.9114 |
| Max coulombic interaction for a C-C bond | -2.5064e+00 | -3.9871 |
| Min net atomic charge for a C atom | 6.1413e+00 | 3.5044 |
| FNSA-2 Fractional PNSA (PNSA-2/TMSA) [Quantum-Chemical PC] | -1.1648e+00 | -2.7531 |
| R=0.7930 F=15.24 $S^2$=0.1954 | | |

## 3. Result and discussion

### *3.1 The results of HM*

About 500 descriptors were calculated for all the compounds. Using the method HM, the pool of the descriptors was reduced. A variety of subset sizes were investigated to determine the optimum number of the descriptor in a model. When adding another descriptor it didn't improve the statistics of the model significantly. It was proved that the optimum model was found. In order to avoid the "over-parametrization" of the model, an increase of the R2 value of less than 0.02 was chosen as the breakpoint criterion. Based on the selected descriptors, a linear regression model was constructed for the training set. The statistical analysis results of the six-parameter model and the involved molecular descriptors as well as the corresponding meaning were summarized in Table 2. In six-parameter model, there were one constitutional descriptor, three electrostatic descriptors, one topological descriptor and one quantum chemical descriptors. Figure 2. showed the experimental versus the predicted ED50 using HM.

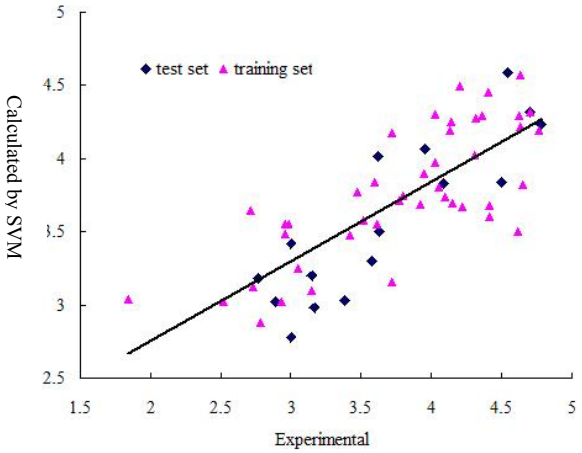



Figure 3. Predicted vs. experimental ED50 by SVM

### *3.2 Result of SVM*

In order to validate the model by HM, SVM was used to develop nonlinear models based on the same descriptors. Similar to other statistical methods, the performance of SVM for regression depend on the combination of several parameters. They are capacity parameter C, ε–insensitive loss

function, the kernel type K, and its corresponding parameters. C is a regularization parameter that controls the tradeoff between maximizing the margin and minimizing the training error. If C is too small, insufficient stress will be placed on

Table 3. Predictive results of ED50 by SVM and HM

| NO | ln1/ED50 | HM | Residue | SVM | Residue |
|---|---|---|---|---|---|
| 1* | -3.58 | -3.4356 | 0.1444 | -3.2981 | 0.281 |
| 2 | -4.32 | -4.3778 | -0.0578 | -4.2746 | 0.045 |
| 3 | -4.76 | -4.3007 | 0.4593 | -4.1959 | -0.056 |
| 4 | -3.52 | -3.5038 | 0.0162 | -3.576 | 0.124 |
| 5* | -3.63 | -3.4932 | 0.1368 | -3.5057 | 0.325 |
| 6 | -4.62 | -4.2508 | 0.3692 | -4.295 | 0.055 |
| 7 | -3.77 | -3.938 | -0.168 | -3.7142 | 0.729 |
| 8 | -4.41 | -3.9435 | 0.4665 | -3.6807 | 0.188 |
| 9* | -3.17 | -2.8057 | 0.3643 | -2.9819 | -0.575 |
| 10 | -2.98 | -3.6659 | -0.6859 | -3.5553 | -0.055 |
| 11 | -3.42 | -3.5758 | -0.1558 | -3.4755 | 0.065 |
| 12 | -4.36 | -4.4137 | -0.0537 | -4.2949 | 0.658 |
| 13* | -4.5 | -4.0324 | 0.4676 | -3.8416 | 0.409 |
| 14 | -4.63 | -4.2697 | 0.3603 | -4.2202 | -0.452 |
| 15 | -3.72 | -4.2549 | -0.5349 | -4.172 | 0.806 |
| 16 | -4.41 | -3.7918 | 0.6182 | -3.6033 | 0.379 |
| 17* | -4.7 | -4.5784 | 0.1216 | -4.3201 | -0.111 |
| 18 | -4.14 | -4.5515 | -0.4115 | -4.2512 | 0.051 |
| 19 | -3.95 | -4.2062 | -0.2562 | -3.899 | 0.825 |
| 20 | -4.65 | -4.1206 | 0.5294 | -3.8249 | -0.101 |
| 21* | -3.96 | -4.2977 | -0.3377 | -4.0617 | -0.588 |
| 22 | -2.96 | -3.4806 | -0.5206 | -3.5485 | 0.056 |
| 23 | -3.61 | -3.5213 | 0.0887 | -3.554 | -0.093 |
| 24 | -2.93 | -2.9613 | -0.0313 | -3.0232 | -0.132 |
| 25* | -2.89 | -2.9705 | -0.0805 | -3.0226 | 0.933 |
| 26 | -2.71 | -3.5628 | -0.8528 | -3.6436 | 0.236 |
| 27 | -3.92 | -3.6601 | 0.2599 | -3.6835 | -1.202 |
| 28 | -1.84 | -2.9883 | -1.1483 | -3.0422 | 0.359 |
| 29* | -3.39 | -2.981 | 0.409 | -3.0305 | -0.393 |
| 30 | -2.73 | -3.0903 | -0.3603 | -3.1235 | 0.561 |
| 31 | -3.72 | -3.1649 | 0.5551 | -3.1588 | -0.498 |
| 32 | -2.52 | -2.8451 | -0.3251 | -3.018 | -0.043 |
| 33* | -4.54 | -4.5996 | -0.0596 | -4.5839 | 0.056 |
| 34 | -4.63 | -4.5846 | 0.0454 | -4.5739 | -0.27 |
| 35 | -4.03 | -4.2502 | -0.2202 | -4.3007 | 0.241 |
| 36 | -4.05 | -3.7496 | 0.3004 | -3.8086 | 0.259 |
| 37* | -4.09 | -3.8107 | 0.2793 | -3.8301 | -0.27 |
| 38 | -4.2 | -4.6886 | -0.4886 | -4.4963 | -0.296 |
| 39 | -3.8 | -3.9514 | -0.1514 | -3.7446 | 0.055 |
| 40 | -3.47 | -3.9661 | -0.4961 | -3.7684 | -0.298 |
| 41* | -3 | -2.673 | 0.327 | -2.7762 | 0.233 |
| 42 | -2.78 | -2.7804 | -0.0004 | -2.8746 | -0.094 |
| 43 | -4.03 | -3.9045 | 0.1255 | -3.9743 | 0.055 |
| 44 | -4.31 | -4.0138 | 0.2962 | -4.0274 | 0.282 |
| 45* | -4.78 | -4.246 | 0.534 | -4.2336 | 0.564 |
| 46 | -3.6 | -4.1934 | -0.5934 | -3.8425 | -0.242 |
| 47 | -4.22 | -3.6351 | 0.5849 | -3.6659 | 0.554 |
| 48 | -4.15 | -3.732 | 0.418 | -3.6932 | 0.456 |
| 49* | -3.15 | -3.2556 | -0.1056 | -3.1952 | -0.045 |
| 50 | -4.1 | -3.6913 | 0.4087 | -3.735 | 0.365 |
| 51 | -3.05 | -3.329 | -0.279 | -3.2522 | -0.202 |
| 52 | -3.15 | -2.9827 | 0.1673 | -3.0942 | 0.055 |
| 53* | -2.76 | -3.1441 | -0.3841 | -3.1799 | -0.419 |
| 54 | -4.13 | -4.203 | -0.073 | -3.4202 | 0.709 |
| 55 | -4.4 | -4.5766 | -0.1766 | -4.1887 | 0.211 |
| 56 | -4.7 | -4.4803 | 0.2197 | -4.4559 | 0.244 |
| 57* | -3 | -3.3833 | -0.3833 | -4.3184 | -1.318 |
| 58 | -3.42 | -3.6769 | -0.2569 | -4.0144 | -0.594 |
| 59 | -2.96 | -3.3699 | -0.4099 | -3.4759 | -0.515 |
| 60 | -4.61 | -3.3977 | 1.2123 | -3.4866 | 1.123 |
| 61* | -3.62 | -3.8465 | -0.2265 | -3.5047 | 0.115 |

HM: R=0.7930 $S^2$=0.1950; SVM: R=0.7054 $S^2$=0.2436

fitting the training data. If C is too large, the algorithm will over fit the training data. However, ref [11] indicated that the prediction error was scarcely influenced by C. To make the learning process stable, a large value should be set up for C.

The kernel type is another important parameter. For regression tasks, the Gaussian kernel is commonly used. In our research, we used the grid research to find the optimum value for γ and ε. Figure.3 showed the experimental versus the predicted ED50 using SVM. Using the optimum value, we get a better non-linear regression model. The results of this model are shown in Table 3 and Figure. 3.

Comparing the correlation models and the

predictive ED50values obtained by HM and SVM, it can be concluded seen that the performance of HM is better the SVM. HM can find the most important factors which influence surface tension significantly. This investigation provides a new method to investigate and predict the ED50 of the family of AMPA receptor antagonists. The compare of the predictive power of both methods was list in Table 3.

## 4. Conclusion

QSAR models for the prediction of ED50 of diversity of compounds using the heuristic method and support vector machine based on descriptors calculated from molecular structure alone have been developed. We have attained satisfactory results using the proposed model. It will give some insight into the factors that influence ED50 of these diverse compounds by discussing the descriptors selected by HM. Using the same set of descriptors, non-linear regression model is constructed by SVM Our investigation first used SVM to research the relationship between structural descriptors and ED50 of the investigated antagonists, and obtained a promising result. This investigation provides a new and effective method to predict the ED50 for the family of AMPA receptor anagtonists. Furthermore, the proposed method can also be extended to other investigations into the property of the compounds and the other property.

## Acknowledgement

The authors are grateful for the financial support provided by the Natural Science Foundation of Gansu Province (3ZS061-A25-088), China.

## References

[1] A. Macchiarulo, L. De Luca, G. Costantino, M. L. Barreca, R. Gitto, R. Pellicciari, A. Chimirri, J Med Chem, 47 (2004) 1860-1863.
[2] P. Buchwald, B. Einstein, N. Bodor, QSAR & Combinatorial Science, 24 (2005) 325-331.
[3] G.J. Lees, Drugs 59 (2000) 33-78.
[4] C.J.C. Burges, A tutorial on support vector machines. Data Min Knowl Disc 2 1998 121-167.
[5] N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machines, Cambridge University Press,Cambridge, 2000 UK.
[6] S.R. Gunn, M. Brown, K.M. Bossley, Lecture Notes Comput Sci, 1280 (1997) 313-323.
[7] H.X. Liu, R.J. Hu, R.S. Zhang, X.J. Yao, M.C. Liu, Z.D. Hu, B.T. Fan, J Comput Aid Mol Des, 19 (2005) 33-46.
[8] A.R. Katritzky, V.S. Lobanov, M. Karelson, CODESSA: Training Manual. University of Florida, Gainesville, 1995 Florida.
[9] A.R. Katritzky, V.S. Lobanov, M. Karelson, CODESSA: Reference Manual. University of Florida, Gainesville, 1994 Florida.
[10] A.R. Katritzky, V.S. Lobanov, M. Karelson, Comprehensive Descriptors for Structural and Statistical Analysis, Reference Manual, 1994 Version 2.0.
[11] W.J. Wang, Z.B. Xu, W.Z. Lu, X.Y. Zhang, Neurocomputing, 55 (2003) 643-663.