Structure-Activity Relationship Study of Nonpeptide δ-Opioid Receptor Ligands

by **Joaquín Lahsen***[a]), **Helmut Schmidhammer**[b]), and **Bernd M. Rode**[a])

[a]) Department of Theoretical Chemistry, Institute of General, Inorganic, and Theoretical Chemistry, University of Innsbruck, Innrain 52a, A-6020 Innsbruck

[b]) Institute of Pharmacy, University of Innsbruck, Innrain 52a, A-6020 Innsbruck

A series of 21 δ-opioid receptor ligands were studied to find a common active center related to the affinity towards the δ-opioid receptor, based on AM1 molecular-orbital calculations. Multivariate adaptive regression splines (MARS) analysis was used to model the active center by atomic charges of selected sites as electronic descriptors. The results show that MARS is reliable in terms of PRESS. The resulting model is based on O(18), N(17), C(3), C(6), C(12), C(13), and C(16), indicating that the charges on these atoms have the strongest influence on the biological activity.

Introduction. – Multiple opioid receptors have been demonstrated through pharmacological and binding studies. Presently, the existence of at least three major types (μ, κ, δ) of opioid receptors is generally acknowledged. Opioid receptors are involved in the modulation of a variety of physiological effects. Recent advances in the understanding of the biological functions of the δ-opioid receptors have highlighted the important role of this receptor in the regulation of pain, and the search for δ receptor ligands has been pursued in numerous laboratories [1–3]. Because peptides are metabolically labile and have problems being orally absorbed, the design of nonpeptide ligands is desirable. The development of such ligands has improved our understanding of δ-opoid receptors and facilitated the identification of new drugs. A δ-selective agonist, BW373U86, a piperazine derivative, *e.g.*, is reported to be a potent analgesic that does not produce physical dependence [4]. Another possible application of δ-opioid receptor agonists is a new therapeutic approach in *Parkinson*'s disease [5]. The chief criterion for the classification of an agonist effect as being opioid-receptor-mediated is the ability of naloxone or naltrexone to reversibly antagonize this effect in a competitive fashion. In addition to their use as pharmacological tools, selective nonpeptide opioid antagonists may have potential clinical applications in the treatment of a variety of disorders where endogenous opioids play a modulatory role. These include, *e.g.*, disorders of food intake, shock, constipation, mental disorders, CNS injury, alcoholism, drug addiction, and immune function [6].

Quantitative structure-activity relations (QSAR) have first been investigated systematically *via* the *Hansch* analysis [7], and they have increasingly proven their importance in chemistry and medicine [8]. Besides the more empirical descriptors commonly used in classical QSAR, quantum-chemically derived descriptors are an excellent alternative to the experiment-based ones [9], and their use in drug design is increasing. Atomic net charges obtained from semiempirical quantum methods have

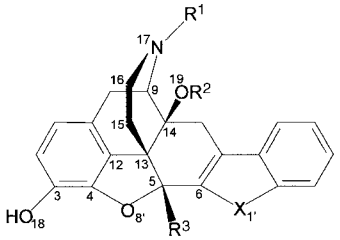
been employed in describing chemical and pharmacological activity in QSAR [10–12]. Searching for a model connecting the net atomic charges with pharmacological activity has led to the use of several methods of regression analysis ranking from parametric to nonparametric methods. One of the most promising nonparametric methods is multivariate adaptive regression splines (MARS) introduced by *Friedman* [13]. This regression tool is a generalization of adaptive regression spline methods, and it builds flexible models by fitting linear regressions piecewise. Application of MARS to chemical studies was introduced by *De Veaux et al.* [14] and, due to its accuracy, MARS could also be a common statistical tool to be chosen in drug design. However, to our knowledge, only one study has been published applying it in this field [15]. In that report, the authors compared MARS with the traditional approach, multiple linear regression (MLR), and two other nonparametric nonlinear methods, alternating conditional expectations (ACE), and projection pursuit regression (PPR) in a QSAR study of dihydroartemisinin derivates. They found that the QSARs derived from MARS method are the most satisfactory predictive models, and that the artemisinin pharmacophore identification is in agreement with previous experimental findings. In the light of their results, we apply in this study multivariate adaptive regression splines to a quantitative electronic structure-activity analysis of a series of 21 δ opioid ligands, agonists, and antagonists, in order to find a model that accounts for the receptor affinity. Their structure, **1**–**21**, and opioid-receptor binding affinities to the δ sites are shown in *Fig. 1* and *Table 1* [16].

Materials and Methods. – *Structures and Data*. The syntheses of compounds **1** and **5** [17] and **14** [18] were reported earlier, while the syntheses of all the other compounds will be presented in a forthcoming publication [16]. In the analysis of quantitative structure-activity relationships, ln K was used as index of the receptor affinity, where K is the inhibitory constant. The molecular geometries of the studied molecules were optimized with the MM+ force field implemented in HyperChem software (HyperChem, *Hypercute*, *Inc.*, Ontario, Canada) and the atomic net charges (*Table 2*) were calculated by the semiempirical molecular-orbital method AM1 and *Mulliken* population analysis implemented in GAUSSIAN 98 [19].

Multivariate Adaptive Regression Splines. MARS Analysis is a local nonparametric regression method that builds up a set of tensor-product splines basis functions. It models the true function $f(x)$ by *Eqn. 1*,

$$\hat{f}(x) = a_0 + \sum_{m=1}^{M} a_m \prod_{k=1}^{K_m} B_{km}(x_{v(k,m)}) \tag{1}$$

where $v(k,m)$ is the index of the factor used as argument of B_{km}. K_m is a parameter that limits the order of interactions. The basis functions B_{km} are first-order truncated power splines defined by $B_{km}(x) = \pm (x - t_{km})_+$, where t_{km} is a knot location chosen from the observed values of the corresponding components, and the function $(\ )_+$ equals its argument for positive arguments and zero otherwise. The key concept underlying the spline is the



1-21

Fig. 1. *Basic structure of the δ-opioid receptor ligands studied*. For R^1, R^2, and R^3, see *Table 1*.

Table 1. *Chemical Structure and Receptor Affinity*

		Binding affinities K_i[nM]
1	R^1 = CPM[a]), R^2 = Et, R^3 = Me, X = NH	4.4
2	R^1 = $PhCH_2CH_2$, R^2 = Et, R^3 = Me, X = NH	31
3	R^1 = allyl, R^2 = Me, R^3 = H, X = NH	3.8
4	R^1 = allyl, R^2 = Et, R^3 = H, X = NH	2.6
5	R^1 = Me, R^2 = Et, R^3 = Me, X = NH	3.9
6	R^1 = R^2 = R^3 = Me, X = NH	3.95
7	R^1 = R^3 = Me, R^2 = $PhCH_2$, X = NH	73
8	R^1 = R^3 = Me, R^2 = 2-$MeC_6H_4CH_2$, X = NH	18.8
9	R^1 = R^3 = Me, R^2 = Pr, X = NH	5.3
10	R^1 = allyl, R^2 = Me, R^3 = H, X = MeN	1.97
11	R^1 = Pr, R^2 = Me, R^3 = H, X = MeN	11.5
12	R^1 = R^2 = allyl, R^3 = H, X = (allyl)N	7.3
13	R^1 = CPM[a]), R^2 = Pr, R^3 = H, X = PrN	1.28
14	R^1 = CPM[a]), R^2 = Et, R^3 = H, X = O	1.18
15	R^1 = allyl, R^2 = 2-$ClC_6H_4CH_2$, R^3 = H, X = 2-$ClC_6H_4CH_2N$	151
16	R^1 = CBM[b]), R^2 = Et, R^3 = Me, X = NH	12.6
17	R^1 = CHM[c]), R^2 = Me, R^3 = H, X = NH	370
18	R^1 = allyl, R^2 = Et, R^3 = H, X = O	1.2
19	R^1 = R^3 = Me, R^2 = $Me_2CHCH_2CH_2$, X = NH	12.8
20	R^1 = R^3 = Me, R^2 = $PhCH_2CH_2CH_2$, X = NH	3.7
21	R^1 = CPM[a]), R^2 = Pr, R^3 = Me, X = NH	5.3

[a]) CPM = cyclopropylmethyl. [b]) CBM = cyclobutylmethyl. [c]) CHM = cyclohexylmethyl.

knot *t*. A knot marks the end of one region of data, and the beginning of another. The knot, thus, is where the behavior of the function changes. In a classical spline, the knots are predetermined and evenly spaced, whereas in MARS, the knots are determined by a search procedure. The MARS algorithm can be summarized as follows:

1. *Forward Stepwise:* It starts with one basis function, $B_o(x) = 1$. At each $(M+1)^{th}$ iteration, two new basis functions that have the strongest effect in decreasing the residual sum of squares are added to the current model at the same time

$$B_{2M+1}(x) = B_{1(M+1)}(x)[+(x_{v(M+1)} - t_{M+1}]_+$$

$$B_{2M+2}(x) = B_{1(M+1)}(x)[-(x_{v(M+1)} - t_{M+1}]_+$$

This procedure is continued until the maximum number of basis functions has been reached.

2. *Backward Stepwise:* The least-important basis functions are eliminated, leading to the final model selection. The optimal MARS model is the one with the lowest generalized cross-validation measure *GCV* according to *Eqn. 2*, where $C(M)$ is the number of parameters being fitted. The numerator is the average residual squared error, and the denominator is a penalty term that reflects model complexity. Once the final model is found, a procedure called ANOVA decomposition is applied, so that the model can reveal information about the predictive relationship between the response Y and X. The MARS model can be recast into the form of *Eqn. 3*, where the first sum is over all basis functions that involve only a single variable, the next is over all basis functions involving two variables, representing two-variable interactions, and so on. If the crossed terms in *Eqn. 3* are neglected and the basis functions were replaced by scalar parameters, the response variable would be simply modeled by a linear combination of the predictor variables (multivariate linear regression). The representation given in *Eqn. 3* identifies the variables that enter into the model, either additively or being involved in interactions. More information about MARS can be found in [13][20].

$$GCV = \frac{1}{n}\sum_{i=1}^{N}(y_i - \hat{f}(x))^2 / \left(1 - \frac{C(M)}{n}\right) \quad (2)$$

$$\hat{f}(x) = a_o + \sum_{K=1} f_i(x_i) + \sum_{K=2} f_{ij}(x_i, x_j) + \cdots \quad (3)$$

Predictive Residual Sum of Squares. A criterion for choice of the best model for predicting well a new sample is the predictive residual sum of squares (PRESS) statistics. Cross-validation is carried out by dividing the data set into a number of groups (here leave-one-out method) and then creating a number of models from the reduced data set with one of the groups omitted. Then, the omitted group is used as a test. This is repeated for each modified data set and the squared differences between predicted and actual response values are summarized to form PRESS, which is defined as $\text{PRESS} = \sum_y (y_{\text{pred}} - y_{\text{actual}})^2$. The model with highest predictive power will be the one that has the smallest PRESS value.

Results and Discussion. – The data set consists of 21 δ-ligands with 15 predictors, which are atomic net charges (*Table 2*), and the natural logarithm of the observed receptor affinity constant K, as the response. The MARS analysis [21] was applied to the data set. The models were obtained by varying the values of the control parameters, namely the maximum number of basis functions (12–16), the maximum number of observations between knots (0–2), and the maximum number of interaction terms (0–1). Then, for the models with the highest values of r^2, PRESS was used as a further criterion for choosing the best of them. *Table 3* shows some models thus obtained. Among them, the fitting quality ($r^2 = 0.98$) and the predictive power (PRESS = 11.9) of *Model 1* are the best. *Model 1* contains the following atoms as predictors: O(18), N(17), C(3), C(6), C(12), C(13), and C(16). Molecule **21** was omitted for PRESS because of the difference between the predicted (−13) and the observed ln K (1.67), and the discrepancy of the atomic net charges for some predictors in comparison with other compounds. The relative variable importance is shown in *Table 4*. These values were normalized so that the most important predictor has a value of 100. Atom C(12) is the most important predictor, followed by C(16) and C(3). The last predictor is attached to O(18), indicating H-bonding to the receptor. The incorporation of N(17) and C(16) into the model reflects the well-known concept that the basic N-function in the phenylpiperidine fragment is essential for the opioid ligands, as proposed by *Portoghese* [22]. C(12) and C(13) are likely to undergo hydrophobic interactions. In the group of molecules studied here, C(6) is attached to either a N- or O-atom, and, due to the lesser importance of this predictor in the model, replacement of N by O does not much influence the observed affinity values (K_i(**4**) = 2.6; K_i(**18**) = 1.2). C(5) is absent in the model, as is explained by a recent study that suggests that a 5-methyl group is not necessary for high δ opioid-receptor antagonism and selectivity [3].

The basis functions and coefficients are shown in *Table 5*. The plot of predicted affinity values obtained by cross-validation *vs.* observed affinity (ln K) is depicted in *Fig. 2*, and it has the regression value $r^2 = 0.78$, the associated PRESS value is 11.9. Due to its fitting and predictive ability, MARS can be considered a most agreeable statistical tool to be used in drug design of the group of molecules investigated. The atoms taken as predictors embrace the region formed by the chain from O(18) to N(17) and, hence, this region could be regarded as the active center of the δ-opioid ligand group studied here.

We also applied multivariate linear regression to the same data set, and we found that this statistical method predicts a similar active center, namely the chain from C(3) to O(19). However, it has a much less predictive power with a PRESS value of 69.6.

Table 2. *Atomic Net Charges Used in This Study*

	O(18)	O(8′)	X(1′)	N(17)	O(19)	C(3)	C(4)	C(5)	C(6)	C(9)	C(12)	C(13)	C(14)	C(15)	C(16)
1	−0.340406	−0.207245	−0.34478	−0.360059	−0.336209	0.128807	0.04388	0.137452	0.02765	−0.007472	−0.110789	−0.069624	0.11534	−0.2266	−0.103185
2	−0.339571	−0.20596	−0.34422	−0.381969	−0.335078	0.13142	0.045388	0.136734	0.029463	−0.008704	−0.104267	−0.071018	0.121325	−0.230657	−0.09297
3	−0.339401	−0.193298	−0.344468	−0.369569	−0.331275	0.130634	0.044272	0.0249	0.027949	−0.001568	−0.130996	−0.065807	0.101595	−0.250223	−0.114204
4	−0.339357	−0.193179	−0.344087	−0.370767	−0.329691	0.130581	0.043953	0.024698	0.021873	−0.001919	−0.131377	−0.064877	0.104209	−0.251856	−0.114305
5	−0.340319	−0.207104	−0.344724	−0.366795	−0.335267	0.129234	0.043502	0.137196	0.027767	−0.011748	−0.10842	−0.069171	0.118125	−0.225981	−0.103048
6	−0.340026	−0.206253	−0.344819	−0.390027	−0.33473	0.129696	0.046652	0.138177	0.027816	0.003188	−0.119329	−0.069797	0.109437	−0.239193	−0.093601
7	−0.339996	−0.207292	−0.34536	−0.362969	−0.328181	0.129875	0.043107	0.13998	0.013483	−0.015084	−0.10975	−0.068625	0.115233	−0.22862	−0.108373
8	−0.33995	−0.207601	−0.345133	−0.362751	−0.329549	0.129716	0.043236	0.140507	0.012537	−0.014851	−0.110112	−0.068521	0.116393	−0.228428	−0.107935
9	−0.340347	−0.206976	−0.344886	−0.366415	−0.336204	0.129195	0.043541	0.137135	0.027903	−0.011895	−0.108588	−0.069403	0.118611	−0.22633	−0.103106
10	−0.339889	−0.19613	−0.29	−0.369202	−0.331552	−0.130128	−0.044638	−0.025277	−0.042757	−0.001956	−0.130573	−0.066095	0.101606	−0.250571	−0.114175
11	−0.340925	−0.197374	−0.289523	−0.362552	−0.337521	0.127708	0.043199	0.025077	0.042182	0.011258	−0.134782	−0.065258	0.096236	−0.249327	−0.104195
12	−0.33813	−0.193999	−0.270963	−0.415753	−0.388982	0.133053	0.048405	0.023368	0.040344	−0.013306	−0.133404	−0.053815	0.178072	−0.228577	−0.09562
13	−0.334013	−0.149442	−0.270725	−0.409655	−0.299304	0.12951	0.051859	0.041077	0.066774	−0.010462	−0.141052	−0.052193	0.117242	−0.227521	−0.090808
14	−0.33662	−0.180611	−0.097768	−0.401314	−0.326394	0.132319	0.043736	0.049201	0.028764	0.015057	−0.13462	−0.068057	0.102584	−0.246586	−0.09212
15	−0.33678	−0.185928	−0.290389	−0.395635	−0.308179	0.12825	0.046902	0.030917	0.038413	0.007024	−0.134155	−0.067685	0.103103	−0.24858	−0.099733
16	−0.340361	−0.207147	−0.344734	−0.359311	−0.335483	0.128934	0.044078	0.137594	0.02759	−0.008469	−0.111518	−0.069835	0.115767	−0.22749	−0.10216
17	−0.338453	−0.192741	−0.344482	−0.355809	−0.293582	0.132909	0.045602	0.024841	0.027379	−0.00802	−0.131615	−0.070496	0.129571	−0.22446	−0.114369
18	−0.336291	−0.179521	−0.097748	−0.370306	−0.329958	0.133436	0.044373	0.048805	0.029407	0.000306	−0.135038	−0.068042	0.102164	−0.250384	−0.113233
19	−0.340177	−0.206342	−0.344168	−0.37485	−0.315359	0.130506	0.046682	0.135672	0.034574	−0.034433	−0.110819	−0.041634	−0.104207	−0.22333	−0.105615
20	−0.339076	−0.206589	−0.33896	−0.381233	−0.311254	0.132622	0.046938	0.133231	0.038324	−0.029575	−0.107036	−0.047203	0.100135	−0.222244	−0.110981
21	−0.340936	−0.298541	−0.364958	−0.408104	−0.307355	0.114528	0.100006	0.139017	−0.096918	−0.018271	−0.183774	−0.11358	0.095396	−0.23552	−0.085324

Table 3. *Best MARS Models*

Model	Predictors	r^2	PRESS
1	O(18), N(17), C(3), C(6), C(12), C(13), C(16)	0.98	11.9
2	O(8′), O(18), O(19), N(17), C(5), C(13), C(15)	0.97	33.2
3	O(18), N(17), C(3), C(12), C(13), C(16)	0.96	38.6

Table 4. *Relative Predictor Importance for MARS* Model 1

Predictor	C(12)	C(16)	C(3)	N(17)	O(18)	C(13)	C(6)
Importance	100.000	88.785	82.743	71.413	43.833	38.023	31.831

Table 5. *Basis Functions and Coefficients for MARS* Model 1

$B_1 = \max(0, \mathrm{N}(17) + 0.360059)$
$B_2 = \max(0, -0.360059 - \mathrm{N}(17)$
$B_3 = \max(0, \mathrm{C}(16) + 0.108373)$
$B_4 = \max(0, -0.108373 - \mathrm{C}(16))$
$B_5 = \max(0, \mathrm{C}(12) + 0.108420)$
$B_6 = \max(0, -0.108420 - \mathrm{C}(12)$
$B_7 = \max(0, \mathrm{O}(18) + 0.339571)$
$B_9 = \max(0, \mathrm{C}(3) - 0.128250)$
$B_{10} = \max(0, 0.128250 - \mathrm{C}(3))$
$B_{11} = \max(0, \mathrm{C}(13) + 0.065258)$
$B_{12} = \max(0, -0.065258 - \mathrm{C}(13))$
$B_{13} = \max(0, \mathrm{C}(16) + 0.107935)$
$B_{15} = \max(0, \mathrm{C}(6) + 0.096918)$

$$\ln K = 8.130531 + 1173.588623 \cdot B_1 + 77.540588 \cdot B_2 - 2753.047119 \cdot B_3 - 844.440369 \cdot B_4 + 1317.470947 \cdot B_5 + 202.326111 \cdot B_6 - 571.058289 \cdot B_7 - 728.092163 \cdot B_9 - 3284.472168 \cdot B_{10} + 118.505096 \cdot B_{11} + 614.327209 \cdot B_{12} + 2361.817627 \cdot B_{13} - 48.347370 \cdot B_{15}$$

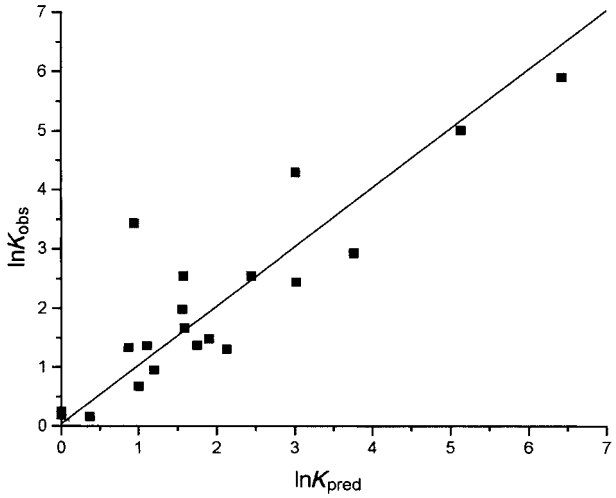


Fig. 2. *Scatter plot for the observed and predicted* (cross-validated) *values of ln* K *of the MARS* Model 1

REFERENCES

[1] P. S. Portoghese, *J. Med. Chem.* **1992**, *35*, 1927.
[2] X. Zhang, K. Rice, S. N. Calderon, H. Kayakiri, L. Smith, A. Coop, A. E. Jacobson, R. B. Rothman, P. Davis, C. M. Dersch, F. Porreca, *J. Med. Chem.* **1999**, *42*, 5455.

[3] H. Schmidhammer, R. Krassnig, E. Greiner, J. Schütz, A. White, I. P. Berzetei-Gurske, *Helv. Chim. Acta* **1998**, *81*, 1064.
[4] P. H. Lee, R. W. McNutt, K. J. Chang, *J. Pharmacol. Exp. Ther*. **1993**, *267*, 883.
[5] M. P. Hill, C. J. Hille, J. M. Rochie, *Drug News Perspect*. **2000**, *13*, 261.
[6] P. S. Portoghese, *J. Med. Chem*. **1991**, *34*, 1757; H. Schmidhammer in 'Progress in Medicinal Chemistry', Eds. G. Ellis, D. K. Luscombe, and A. W. Oxford, Elsevier, Amsterdam, 1998, Vol. 35, p. 83.
[7] C. Hansch, T. Fujita, *J. Am. Chem. Soc*. **1963**, *86*, 1616.
[8] C. Hansch, A. Leo, 'Exploring QSAR. Fundamentals and Applications in Chemistry and Biology', American Chemical Society, Washington DC, 1995.
[9] M. Karelson, V. Lobanov, A. Katritzy, *Chem. Rev*. **1996**, *96*, 1027.
[10] C. T. Supuran, B. W. Clare, *Eur. J. Med. Chem*. **1995**, *30*, 687.
[11] I. A. Doichinova, R. N. Natcheva, D. N. Mihailova, *Eur. J. Med. Chem*. **1994**, *29*, 133.
[12] M. Cocchi, M. C. Menziani, F. Fanelli, P. de Benedetti, *J. Mol. Struct. (Theochem)* **1995**, *331*, 79.
[13] J. H. Friedman, *Annals of Statistics a* **1991**, *19*, 1.
[14] R. D. De Veaux, D. C. Psichogios, L. H. Ungar, *Comput. Chem. Eng.* **1993**, *17*, 819.
[15] V. Nguyen-Cong, G. Van Dang, B. M. Rode, *Eur. J. Med. Chem*. **1996**, *31*, 797.
[16] H. Schmidhammer, R. Meditz, M. Koch, R. Krassnig, M. Spetea, E. Greiner, R. Horel, C. M. Dersch, R. B. Rothman, in preparation.
[17] H. Schmidhammer, D. Daurer, M. Wieser, K. Monory, A. Borsodi, J. Elliot, J. R. Traynor, *Bioorg. Med. Chem. Lett*. **1997**, *7*, 151.
[18] H. Schmidhammer, P. Schwarz, Z.-Y. Wei, *Helv. Chim. Acta* **1998**, *81*, 1215.
[19] 'Gaussian 98', Revision A.7, Gaussian Inc., Pittsburgh, PA, 1998.
[20] 'User Guide MARS', Salford Systems.
[21] 'MARS', Version 1.0, Salford Systems 1999.
[22] P. S. Portoghese, *J. Med. Chem*. **1965**, *8*, 609.

Received January 26, 2001