Available online at www.sciencedirect.com
SCIENCE DIRECT®
Talanta 67 (2005) 182–186
Talanta
www.elsevier.com/locate/talanta

Simple and highly predictive QSAR method: application to a series of (*S*)-*N*-[(1-ethyl-2-pyrrolidinyl)methyl]-6-methoxybenzamides

Matheus P. Freitas, José A. Martins *

EMS Sigma Pharma – R&D, Rodovia SP 101 Km 08, Hortolândia SP 13186-401, Brazil

Received 25 November 2004; received in revised form 2 February 2005; accepted 14 February 2005
Available online 23 March 2005

Abstract

A simple quantitative structure–activity relationship (QSAR) method of analysis used to predict biological activity for congeneric series of compounds is reported. This method is based on the application of bilinear or multilinear partial least squares regression to a data set, which is a binary matrix representing the substituents of a framework. It is appraised here to a series of (*S*)-*N*-[(1-ethyl-2-pyrrolidinyl)methyl]-6-methoxybenzamides, compounds with affinity towards the dopamine D_2 receptor subtype and showed high predictive ability, even when compared to a refined three-dimensional (3D) approach.

Keywords: QSAR analysis; PLS; (*S*)-*N*-[(1-ethyl-2-pyrrolidinyl)methyl]-6-methoxybenzamides

1. Introduction

Quantitative structure–activity relationship (QSAR) analyses are mostly performed by applying methodologies based on two-dimensional (2D) and three-dimensional (3D) computing approaches to predict biological activities. These have provided outstanding results, specially when using the known CoMFA [1], CoMSIA [2] and GRID [3] models to generate descriptors, in detriment of testing experimentally and sometimes to predict intuitively which compound could have a particular activity, allowing then to develop new drugs in a faster and lesser costly manner [4]. However, such computational methods require an exhaustive conformational and alignment treatment. Accordingly, alternative methods, which provide rapid analysis and results as reliable as the most sophisticated methodologies available today, but which are inexpensive and facile to handle, is of immediate and worldwide interest.

Promising attempts to simplify and offer advantages over 2D and 3D methods have been emerging [5], as they were developed in the past [6], but obtainment of comparable results is still an unsolved problem. A QSAR method based on multivariate image analysis developed by us has presented encouraging results as a 2D technique [7], and it was an effort to open precedents in the course of developing new, simple QSAR methods. The present paper is focused on the development of a simple method for QSAR analysis, capable to predict biological activity, or any other parameter, with high correlation coefficients and good statistics. Two partial least squares regression methods, namely PLS [8] and N-PLS [9], were used here to assess the calibration models. Using the PLS regression method, an unfolded **X**-matrix, where each row contains the variables (the binary descriptors) describing each molecule, is decomposed into a score vector ($\boldsymbol{s_1}$) and a weight vector ($\boldsymbol{w}_1$), and $\boldsymbol{s_1}$ is determined to have the property of maximum covariance with the dependent variable $\boldsymbol{y}$. The scores vectors then replace the original variables as regressors. Using the multilinear PLS (N-PLS), the unfolding step is not necessary and the decomposition is performed directly in the three-way matrix. It was shown that N-PLS is more stable than bilinear PLS, i.e. traditional PLS, since it is supposed to increase the predictive ability and improve the interpretation of the results, in 3D QSAR. A detailed account of N-PLS regression method may be found elsewhere [9].

* Corresponding author. Tel.: +55 19 3887 9353; fax: +55 19 3887 9889.
E-mail address: jose.martins@ems.com.br (J.A. Martins).

0039-9140/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2005.02.016

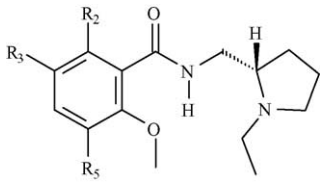


Fig. 1. Structure of (*S*)-*N*-[(1-ethyl-2-pyrrolidinyl)methyl]-6-methoxybenzamides (**1**-**58**). R_2 = H and OH; R_3 = H, Me, Et, *n*-Pr, *n*-Bu, OMe, SMe, NO_2, F, Cl, Br and I; R_5 = H, Me, Et, *n*-Pr, OMe, NO_2, F, Cl and Br.

Our proposed method is appraised in detail here to a series of (*S*)-*N*-[(1-ethyl-2-pyrrolidinyl)methyl]-6-methoxybenzamides (Fig. 1), compounds that possess high affinity and selectivity towards dopamine D_2 receptors, and which may be used in the treatment of psychiatric disorders, such as schizophrenia. The results from our modeling are compared with those from a reported study [10], where a multilinear PLS algorithm was utilized as the regression method in 3D QSAR and the GRID program [3] was used for generation of descriptors. That work rendered notorious results of correlation, and thus is a suitable reference for comparison with our purpose.

2. Methods

The key to this analysis is to build the matrices. Descriptors for each substituent of (*S*)-*N*-[(1-ethyl-2-pyrrolidinyl)methyl]-6-methoxybenzamides were built and inserted in the corresponding position of a template table of 9 rows by 38 columns. This table reserves the first 10 columns to the 2-position substituent, the next 14 columns to the 3-position substituent and the remaining 14 columns to the 5-position substituent, and it may be built by using a Word processor and the data may then be treated using any appropriate software (for example, Matlab [11]). The tables corresponding to the 58 compounds are grouped to form a calibration set of $40 \times 9 \times 38$ and a test set of $18 \times 9 \times 38$ (N-PLS model), which are subsequently unfolded to form a new calibration set of 40×342 and a test set of 18×342 (PLS model). In the present analysis, the only data preprocessing applied on our data was column mean-centering. For the three-way array, the matrix was unfolded, column mean-centered and then back-folded before regression analysis.

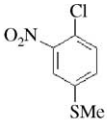
The quality of the calibrations was quantified with R^2 and the external predictions with Q^2, the squared correlation coefficients of the linear regression of experimental versus predicted pIC_{50} plots for the calibration and test sets, respectively. The *F*-statistic and *t*-test values were also obtained to evaluate the model quality. The leave-one-out cross-validations were performed with the NIPALS algorithm [12] and the minimum cumpress (cumulative predictive residual error sum of squares) values were estimated for 5 PLS and 7 N-PLS latent variables (LV).

3. Results and discussion

A priority needed to perform computer-assisted QSAR analysis is to calculate descriptors suitable to correlate structures with the corresponding activities. Three-dimensional descriptors, which describe non-covalent interactions, are the ones most usually applied for such purpose. However, their generation requires previous molecular treatment, such as conformational screening and ligand alignment. On the other hand, our proposed methodology avoids this time-consuming requirement and may be easily performed. Here, the descriptors are a suggestive binary table, where the ones give origin to the element symbol or group representation, while the remaining blanks of table are filled out by zeros. Each molecule of a congeneric series owns a table of *m* rows by *n* columns, and the columns are divided according to the number of substituents. As an instance, for a three-substituted aromatic ring, e.g. 1-chloro-2-nitro-4-methyltio-benzene, a 9×42 table may be built, whose columns are divided in 3 sets of 14 columns, as illustrated in Fig. 2.

A similar procedure is described here to a series of (*S*)-*N*-[(1-ethyl-2-pyrrolidinyl)methyl]-6-methoxybenzamides, and the provided data are then matricized to two-way and three-way arrays, in order to run bilinear (PLS) and multilinear (N-PLS) partial least squares regression, respectively, and consequently appraise the predictive ability of the model by correlating the fitted and predicted activities (pIC_{50}) with the corresponding experimental values. Comparison with predictions from literature [10] is also provided to better evaluate the quality of our model.

A 9×38 table (10 columns for the 2-position substituent and 14 columns for both the 3- and 5-position substituents) is built analogously to Fig. 2 for 58 (*S*)-*N*-[(1-ethyl-2-pyrrolidinyl)methyl]-6-methoxybenzamides, which were divided into a calibration set of 40 compounds and a test set of 18 compounds, identically as previously reported [10]. Two models are considered here: Model 1, consisting of three-way matrices to run N-PLS, i.e. a $40 \times 9 \times 38$ matrix for the calibration set and a $18 \times 9 \times 38$ matrix for the test set, and Model 2 consisting of unfolded matrices to run PLS, i.e. a



R1	R2	R4
0 1 1 1 0 1 0 0 0 0 0 0 0 0	1 0 0 0 1 0 1 1 1 0 0 0 0 0	0 1 1 1 0 1 0 0 0 1 0 0 0 0
1 0 0 0 1 1 0 0 0 0 0 0 0 0	1 1 0 0 1 1 0 0 0 1 0 0 0 0	1 0 0 0 1 1 1 0 1 1 0 0 0 0
1 0 0 0 0 1 0 0 0 0 0 0 0 0	1 0 1 0 1 1 0 0 0 1 0 0 0 0	1 0 0 0 0 1 0 1 0 1 0 0 0 0
1 0 0 0 0 1 0 0 0 0 0 0 0 0	1 0 1 0 1 1 0 0 0 1 0 0 0 0	1 0 0 0 0 1 0 1 0 1 0 0 0 0
1 0 0 0 0 1 0 0 0 0 0 0 0 0	1 0 1 0 1 1 0 0 0 1 1 1 1 1	0 1 1 1 0 1 0 1 0 1 0 1 1 0
1 0 0 0 0 1 0 0 0 0 0 0 0 0	1 0 1 0 1 1 0 0 0 1 0 0 0 1	0 0 0 0 1 1 0 0 0 1 1 0 0 1
1 0 0 0 0 1 0 0 0 0 0 0 0 0	1 0 1 0 1 1 0 0 0 1 1 1 1 1	0 0 0 0 1 1 0 0 0 1 1 1 1 1
1 0 0 0 1 1 0 0 0 0 0 0 0 0	1 0 0 1 1 1 0 0 0 1 1 0 0 0	1 0 0 0 1 1 0 0 0 1 1 0 0 0
0 1 1 1 0 1 0 0 0 0 0 0 0 0	1 0 0 0 1 0 1 1 1 0 1 1 1 1	0 1 1 1 0 1 0 0 0 1 0 1 1 1

Fig. 2. Table of descriptors for 1-chloro-2-nitro-4-methyltio-benzene. Note that ones designate, as a drawing, element symbols and groups, while zeros fill out the blanks.

Table 1
Aromatic substitution patterns and activities of compounds **1-58**

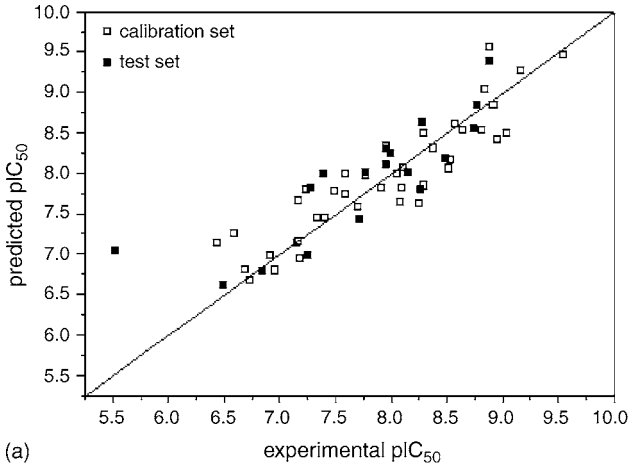
Compounds	R_2	R_3	R_5	Activity (pIC50)			Compounds	R_2	R_3	R_5	Activity (pIC_{50})		
				Exp[a]	Fitted (PLS/N-PLS)	Predicted (PLS/N-PLS)					Exp[a]	Fitted (PLS/N-PLS)	Predicted (PLS/N-PLS)
1	OH	Cl	Cl	7.49	7.77/7.76	–	**30**	OH	Br	NO_2	6.73	6.67/6.78	–
2	OH	OMe	Cl	7.15	7.13/7.15	–	**31**	OH	I	H	8.52	8.05/8.21	–
3	H	Br	Br	8.10	7.80/7.87	–	**32**	OH	Me	Cl	7.59	7.74/7.73	–
4	H	Et	Br	7.96	8.34/8.48	–	**33**	OH	Me	Me	8.11	8.07/7.53	–
5	H	I	OMe	9.17	9.26/9.28	–	**34**	OH	Et	H	8.54	8.16/8.15	–
6	OH	*n*–Pr	Me	8.30	8.50/8.34	–	**35**	OH	Et	F	8.82	8.52/8.53	–
7	OH	Cl	*n*–Pr	6.96	6.80/7.13	–	**36**	OH	Et	Cl	9.04	8.42/8.64	–
8	OH	H	Et	6.91	6.97/7.15	–	**37**	OH	Et	Br	8.64	8.52/8.68	–
9	OH	I	OMe	9.54	9.45/9.49	–	**38**	OH	*n*–Pr	H	8.30	7.84/8.04	–
10	H	SMe	OMe	8.96	8.50/8.40	–	**39**	OH	OMe	H	6.69	6.80/6.66	–
11	OH	Et	OMe	8.89	9.56/9.43	–	**40**	OH	OMe	Br	7.17	7.16/7.20	–
12	H	*n*-Bu	OMe	8.57	8.60/8.80	–	**41**	H	Br	OH	8.00	–	8.24/8.47
13	H	*n*-Pr	H	7.17	7.65/7.84	–	**42**	OH	*n*-Pr	Cl	8.49	–	8.19/8.62
14	H	Cl	H	6.59	7.25/7.06	–	**43**	OH	Me	Br	8.26	–	7.78/7.85
15	H	Cl	Cl	7.70	7.58/7.55	–	**44**	H	Me	OMe	8.28	–	8.63/8.40
16	H	Cl	Br	8.25	7.61/7.60	–	**45**	OH	Br	Et	7.77	–	8.01/8.29
17	H	Br	H	7.34	7.44/7.33	–	**46**	OH	Et	Et	8.75	–	8.54/8.90
18	H	Br	OMe	8.92	8.84/8.62	–	**47**	H	Et	OMe	8.89	–	9.38/9.31
19	H	Et	Cl	8.38	8.31/8.43	–	**48**	H	H	OMe	7.28	–	7.81/7.63
20	OH	H	Cl	7.19	6.93/6.96	–	**49**	H	Et	H	7.40	–	7.99/8.02
21	OH	H	OMe	8.06	8.00/7.76	–	**50**	OH	H	H	6.50	–	6.61/6.56
22	OH	F	H	6.44	7.13/7.00	–	**51**	OH	H	Br	7.25	–	6.97/7.09
23	OH	Cl	H	7.41	7.44/7.27	–	**52**	OH	Cl	Me	7.96	–	8.11/7.65
24	OH	Cl	Br	7.24	7.80/7.80	–	**53**	OH	Cl	OMe	8.77	–	8.85/8.64
25	OH	Cl	Et	7.92	7.81/7.94	–	**54**	OH	Br	F	8.15	–	8.00/8.01
26	OH	Br	H	8.08	7.63/7.54	–	**55**	OH	Br	Me	7.96	–	8.30/7.92
27	OH	Br	Cl	7.77	7.96/8.03	–	**56**	OH	Me	H	7.72	–	7.42/7.32
28	OH	Br	Br	7.59	7.99/8.07	–	**57**	OH	Me	*n*-Pr	6.85	–	6.78/7.18
29	OH	Br	OMe	8.85	9.03/8.82	–	**58**	OH	NO_2	H	5.52	–	7.03/6.83

[a] Experimental values are obtained from Ref. [10].

40×342 matrix for the calibration set and a 18×342 matrix for the test set. This was done in order to compare the traditional PLS regression method with the N-PLS approach, since N-PLS is supposed to demonstrate several advantages as compared to PLS [10].

The results for calibration and validation are presented in Table 1 and illustrated in Fig. 3, where good performances in prediction are shown for both PLS and N-PLS models. The predicted pIC_{50} values of Fig. 3 are referred to 5 PLS and 7 N-PLS latent variables, since these presented the minimum compress values in the leave-one-out cross-validation experiments. In order to show that the relationships did not result from happenstance and to assure that the calibration was not a fortuitous correlation, we scrambled the Y-block (the activities block) and no predictive relationship was found from the modeling (R^2 of 0.42 for 5 PLS LV's and 0.45 for 7 N-PLS LV's), as supposed if we consider that a set of compounds with no modeling capability is taken. The largest deviation in prediction, as can easily be seen in Fig. 3 (the point corresponding to the smallest experimental pIC_{50}), refers to the compound containing a NO_2 group bonded to the 3-position of the aromatic ring. No compounds with this connectivity were used in the calibration set, explaining the observed outlier result, but it should be born in mind that similar deviations usually occur when any other QSAR model is applied to predict properties in "non-trained situations". Also, since this point corresponds to an extreme value of experimental pIC_{50}, it is not advisable to belong to an external validation set, but it was left in to ends of comparison with the literature results [10] (when including this point into the calibration set, the predicted values are improved to a Q^2 of 0.76 for 5 PLS LV's and 0.78 for 7 N-PLS LV's). Tables 2 and 3 show the statistical parameters obtained from calibration and validation, as well as the correlation results from literature [10]. Our proposed method presented improving effects in many aspects in comparison with the results from the 3D approach of literature [10], such as better predictive power using a smaller number of data. The N-PLS regression method has slightly improved the predictive ability as compared to the bilinear PLS for the set of compounds studied here, while the latter was a bit more parsimonious.

Loadings analysis was performed, by using the chemometric Pirouette 3.11 Software [13], in order to identify important variables in the PLS model and also to reach some interpretation of the binary descriptors. Exclusion of variables with low loadings from the model, according to factor



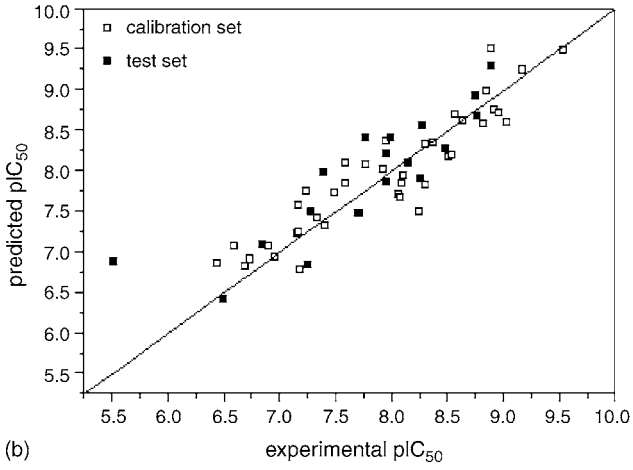


Fig. 3. Experimental vs. predicted pIC_{50} values for (a) PLS model and (b) N-PLS model.

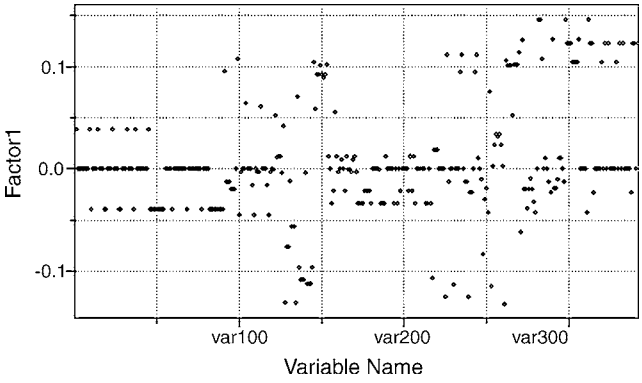


Fig. 4. Loadings plot, using factor number one for the PLS model.

number one (the factor that retains the most significant variance in the data—see loadings plot using factor one in Fig. 4), did not improve the calibration quality. However, some interpretability from the weights illustrated in Fig. 4 may be given for the calibration set. The last rows of the last columns, considering the binary tables for each molecule, corresponding to the last variables in the **X**-matrix, have a positive effect on the biological activity when populated by "ones", e.g. OMe and SMe substituents at the R_5 position. On the other hand, the first rows, according to factor number one, have small significance on the biological activities.

Other data sets have been analyzed in our laboratory using this approach, including a series of potential anxiolytic agents, which are some 5-HT_{2C} receptor antagonists. For this series, our cross-validated predictions gave a Q^2_{CV} of 0.60 for 5 PLS LV's, in agreement with the results reported from a CoMFA analysis [14], where a Q^2_{CV} of 0.656 for 4 PLS LV's was obtained.

Table 2
Statistical parameters of calibration and validation for the QSAR model (using PLS) and literature data

#LV	R^{2a}	Q^{2b}	S.D.$_e$[c]	S.D.$_p$[d]	S.E.$_e$[e]	S.E.$_p$[f]	F_e[g]	F_p[h]	t_e[i]	t_p[j]	R^{2k}	Q^2_{CV}[l]	R^{2l}	Q^{2l}
1	0.46	0.19	0.59	0.82	0.34	0.62	−5.7	−7.4	85.1	40.9	0.47	0.29	0.44	0.36
2	0.56	0.32	0.53	0.72	0.28	0.48	11.4	−7.4	94.6	46.4	0.59	0.44	0.64	0.45
3	0.68	0.43	0.46	0.67	0.20	0.41	42.6	−0.2	110.0	50.4	0.69	0.52	0.73	0.64
4	0.77	0.62	0.38	0.54	0.14	0.27	91.1	13.4	130.3	62.2	0.75	0.56	0.79	0.56
5	**0.82**	**0.71**	**0.34**	**0.47**	**0.11**	**0.20**	**137.5**	**24.4**	**147.2**	**71.3**	0.79	0.56	0.81	0.62
6	0.83	0.70	0.33	0.48	0.10	0.21	154.5	24.7	152.9	70.6	0.81	0.57	0.87	0.62
7	0.84	0.65	0.32	0.52	0.10	0.25	171.7	17.4	158.5	64.8	0.85	0.56	0.88	0.57
8	0.85	0.67	0.31	0.51	0.09	0.24	190.2	19.0	164.2	66.4	0.87	0.62	0.90	0.63
9	0.86	0.69	0.30	0.49	0.09	0.22	206.2	15.7	169.1	68.2	–	–	0.91	0.62
10	0.87	0.69	0.29	0.50	0.08	0.22	221.1	11.7	173.5	67.8	–	–	0.92	0.59

[a] Squared correlation coefficient of estimate.
[b] Squared correlation coefficient of prediction.
[c] Standard deviation of estimate.
[d] Standard deviation of prediction.
[e] Standard error of estimate.
[f] Standard error of prediction.
[g] *F*-test value of estimate.
[h] *F*-test value of prediction.
[i] *t*-Test value of estimate.
[j] *t*-Test value of prediction.
[k] R^2 and Q^2_{CV} from literature (Ref. [10]) for model with 26,400 variables.
[l] R^2 and Q^2 from literature (Ref. [10]) for model with 2940 variables.

Table 3
Statistical parameters of calibration and validation for the QSAR model (using N-PLS) and literature data

#LV	R^{2}[a]	Q^{2}[b]	S.D.$_{e}$[c]	S.D.$_{p}$[d]	S.E.$_{e}$[e]	S.E.$_{p}$[f]	F_{e}[g]	F_{p}[h]	t_{e}[i]	t_{p}[j]	R^{2}[k]	Q^{2}_{CV}[l]	R^{2}[l]	Q^{2}[l]
1	0.44	0.20	0.60	0.81	0.35	0.62	−8.7	−7.4	83.6	41.0	0.47	0.29	0.44	0.36
2	0.50	0.28	0.56	0.75	0.31	0.51	1.0	−7.8	88.9	45.0	0.59	0.44	0.64	0.45
3	0.60	0.33	0.50	0.74	0.25	0.50	20.7	−3.6	99.5	45.6	0.69	0.52	0.73	0.64
4	0.70	0.50	0.44	0.62	0.19	0.36	54.0	1.0	115.1	53.8	0.75	0.56	0.79	0.56
5	0.79	0.68	0.37	0.50	0.13	0.22	106.0	15.3	136.0	67.8	0.79	0.56	0.81	0.62
6	0.81	0.76	0.35	0.44	0.12	0.17	132.1	32.6	145.3	77.5	0.81	0.57	0.87	0.62
7	**0.84**	**0.75**	**0.32**	**0.43**	**0.10**	**0.17**	**168.5**	**35.9**	**157.5**	**77.1**	0.85	0.56	0.88	0.57
8	0.85	0.74	0.31	0.44	0.10	0.18	179.5	28.0	160.9	75.6	0.87	0.62	0.90	0.63
9	0.85	0.75	0.31	0.44	0.09	0.18	186.0	28.3	163.0	76.4	–	–	0.91	0.62
10	0.86	0.76	0.30	0.43	0.09	0.17	203.2	32.0	168.2	78.1	–	–	0.92	0.59

[a] Squared correlation coefficient of estimate.
[b] Squared correlation coefficient of prediction.
[c] Standard deviation of estimate.
[d] Standard deviation of prediction.
[e] Standard error of estimate.
[f] Standard error of prediction.
[g] *F*-test value of estimate.
[h] *F*-test value of prediction.
[i] *t*-Test value of estimate.
[j] *t*-Test value of prediction.
[k] R^{2} and Q^{2}_{CV} from literature (Ref. [10]) for model with 26,400 variables.
[l] R^{2} and Q^{2} from literature (Ref. [10]) for model with 2940 variables.

With this novel appliance, a simple and very accessible method for QSAR analysis, one can test molecules with different substituents and evaluate which type of them and in which position they can influence dependent variables. Models can be built rapidly, depending only on availability of biological data. We do not have alignment problems, of course, as using a single binary table instead of 3D descriptors, but having results as good as those provided from 3D approaches.

4. Conclusions

The proposed QSAR methodology for congeneric series of compounds allows the construction of much smaller matrices than the usually applied 3D approaches, thus requiring low computational cost, and provided high predictive ability for a series of (*S*)-*N*-[(1-ethyl-2-pyrrolidinyl)methyl]-6-methoxybenzamides, when using both PLS and N-PLS regression methods. Actually, the presented descriptors act as codes, though they do not have a direct physicochemical meaning, and were generated here to didactically represent the substituents, but any plausible and systematic representation for the substituents should be possible to guarantee similar modeling capability.

References

[1] R.D. Cramer III, D.E. Patterson, J.D. Bunce, J. Am. Chem. Soc. 110 (1988) 5959–5967.
[2] G. Klebe, U. Abraham, T. Mietzner, J. Med. Chem. 37 (1994) 4130–4146.
[3] P.J. Goodford, GRID, University of Oxford, Oxford, UK, 1995.
[4] P. Rees, Sci. Comput. World July/August (2003) 16–18.
[5] S.L. Dixon, K.M. Merz Jr., J. Med. Chem. 44 (2001) 3795–3809.
[6] S.M. Free, J.W. Wilson, J. Med. Chem. 7 (1964) 395–399.
[7] M.P. Freitas, S.D. Brown, J.A. Martins, J. Mol. Struct. 738 (2005) 149–154.
[8] P. Geladi, B.R. Kowalski, Anal. Chim. Acta 185 (1986) 1–17.
[9] R. Bro, J. Chemometrics 10 (1996) 47–61.
[10] J. Nilsson, E.J. Homan, A.K. Smilde, C.J. Grol, H. Wikström, J. Comput. Aided Mol. Des. 12 (1998) 81–93.
[11] Matlab Version 6.5.1, MathWorks Inc., Natick, MA, 1993.
[12] H. Wold, in: K.R. Krishnaiah (Ed.), Multivariate Analysis, Academic Press, New York, 1966, pp. 391–420.
[13] Pirouette Version 3.11, Infometrix Inc., Woodinville, WA, 1990–2003.
[14] S.M. Bromidge, S. Dabbs, D.T. Davies, D.M. Duckworth, I.T. Forbes, P. Ham, G.E. Jones, F.D. King, D.V. Saunders, S. Starr, K.M. Thewlis, P.A. Wyman, F.E. Blaney, C.B. Naylor, F. Bailey, T.P. Blackburn, V. Holland, G.A. Kennett, G.J. Riley, M.D. Wood, J. Med. Chem. 41 (1998) 1598–1612.