Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters 14 (2004) 3077–3079
Bioorganic & Medicinal Chemistry Letters

A TOPS-MODE approach to predict adenosine kinase inhibition

Maykel Pérez González[a,b,*] and Maria del Carmen Terán Moldes[c]

[a]*Service Unit, Experimental Sugar Cane Station 'Villa Clara-Cienfuegos', Ranchuelo, Villa Clara, C.P. 53100, Cuba*
[b]*Chemical Bioactive Center, Central University of Las Villas, Santa Clara, 54830 Villa Clara, Cuba*
[c]*Department of Organic Chemistry, Vigo University, Vigo, C.P. 36200, Spain*

Received 9 March 2004; revised 7 April 2004; accepted 10 April 2004

Abstract—The TOPological Sub-Structural Molecular Design (TOPS-MODE) approach has been applied to the study of the adenosine kinase inhibitory activity of pyrrolo[2,3-*d*]pyrimidine nucleoside analogues. A model capable of describing around 77% of the variance in the experimental activity of 32 analogues of these compounds was developed with the use of the mentioned approach. In contrast, no one of nine different approaches, including the use of Constitutional, Topological, BCUT, 2D autocorrelations, geometrical, RDF, 3D Morse, WHIM, and GETAWAY descriptors were able to explain more than 70% of the variance in the mentioned property with the same number of descriptors. Although, statistically significant models were derived containing other descriptors than spectral moments still the best one fitted out model was find with these variables.

1. Introduction

Endogenously produced adenosine (ADO) serves a number of roles in the body, but it is especially important as an extracellular messenger where it acts at specific receptors on the cell surface to modulate neuronal activity and inflammation.[1,2] In spite of extensive effort, the direct pharmacological modulation of adenosine receptors with agonists has not yielded useful drug candidates for human use, due to the prevalence of mechanism based side effects (prominently hemodynamic effects). A rationale for a therapeutic approach targeting an indirect modulation of ADO receptors has been proposed as providing 'site and event selectivity', with an enhanced therapeutic window.[3]

ADO kinase (AK) is the key intracellular enzyme in order to regulate intra- and extracellular ADO concentrations. Inhibition of AK produces marked increases in extracellular ADO levels that are localized to cells and tissues undergoing accelerated ADO release.[4,5] Thus, AK inhibition represents a mechanism to selectively enhance the protective actions of ADO during tissue trauma without producing nonspecific effects associated with the systemic administration of ADO receptor agonists.[6]

During the last few years, specific inhibitors of AK have been developed.[7] AK inhibitors have been shown to demonstrate potential clinical utility in animal models of epilepsy, ischemia, pain, and inflammation.[5–8] This study may be fruitfully facilitated by a quantitative structure–activity relationship (QSAR) study, which investigates the properties of molecules that can lead to strongly bind with the receptors (i.e. hydrophobicity, polarizability, polar surface, etc.).

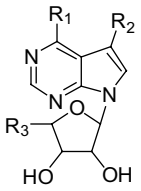
2. Data sets and computational strategies

In an attempt to identify the chemical structural features required and/or responsible for AK inhibition, QSAR studies were undertaken using the TOPS-MODE approach[9–11] and Dragon Software[12,13] on the 32 pyrrolo[2,3-*d*]pyrimidine nucleoside analogues reported by Srikanth et al.[14] The AK inhibition of pyrrolo[2,3-*d*]pyrimidine nucleoside analogues is listed in Table 1. IC_{50} refers to the micromolar concentration of the compound required for 50% inhibition of the enzyme activity. IC_{50} values are transformed to pIC_{50} (negative logarithm of IC_{50}) to get the linear relationship in the equation.

Keywords: QSAR; Adenosine kinase; Topological indices; TOPS-MODE.
* Corresponding author. Tel.: +53-42-281473; fax: +53-42-281130; e-mail: mpgonzalez76@yahoo.es

0960-894X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2004.04.040

Table 1. Structures and AK inhibition activities of pyrrolo[2,3-*d*]-pyrimidine nucleoside analogues



Compound	R^1	R^2	R^3	pIC_{50}
1	NH_2	I	CH_2OH	1.59
2	NH_2	I	CH_3	2.05
3	NH_2	Br	CH_3	1.40
4	NH_2	Br	CH_2OH	0.92
5	NH_2	Cl	CH_2OH	0.68
6	Cl	I	CH_2OH	1.62
7	Cl	I	CH_3	2.52
8	Cl	Br	CH_3	1.30
9	Cl	CH_3	CH_3	0.27
10	Cl	SCH_3	CH_3	1.15
11	Cl	I	CH_2N_3	2.05
12	Cl	Br	CH_2N_3	1.00
13	Cl	I	CH_2OCH_3	0.35
14	NH_2	CH_3	CH_3	−1.22
15	NH_2	SCH_3	CH_3	−0.13
16	NH_2	COOEt	CH_3	0.00
17	NH_2	I	CH_2N_3	1.46
18	NH_2	Br	CH_2N_3	1.20
19	NH_2	I	CH_2OCH_3	−0.08
20	NH_2	I	$CH{=}CH_2$	1.00
21	NH_2	I	C_2H_5	1.22
22	NH_2	Br	C_2H_5	0.30
23	$NHCH_3$	I	CH_2OH	−0.08
24	SH	I	CH_3	−0.81
25	SCH_3	I	CH_3	1.35
26	NH_2	I	CH_2NH_2	3.22
27	NH_2	Br	CH_2NH_2	3.70
28	Cl	I	CH_2NH_2	4.00
29	NH_2	CN	CH_2OH	0.51
30	NH_2	CN	CH_3	0.51
31	NH_2	$CONH_2$	CH_2OH	0.33
32	NH_2	$CONH_2$	CH_3	0.34

In the case of TOPS-MODE approach,[15] the polar surface, the hydrophobicity, and the Gasteiger–Marsili charges were used as bond weight for making differentiations of heteroatoms.[11] A total of 48 descriptors were used in order to obtain this model. On the other hand, we carry out geometry optimization calculations for each compound used in this study using the semi-empirical method AM1[16] included in MOPAC 6.0.[17] Other nine models were developed using the computer software Dragon, calculating the Constitutional, Topological, BCUT, 2D autocorrelations, Geometrical, RDF, 3D-MORSE, WHIM, and GETAWAY descriptors. The statistical processing to obtain the QSAR models was carried out using the forward stepwise regression methods. The statistical significance of the models was determined by examining the regression coefficient, the standard deviation, the number of variables, the cross validation leave-one-out statistics, and the proportion between the cases and variables in the equation.

3. Quantitative structure–activity relations

The best QSAR model obtained with the TOPS-MODE descriptors is given below together with the statistical parameters of the regression.

$$-\log(IC_{50}) = 16.34 - 0.166 \cdot \mu_1\mu_3^{GM} + 1.10 \times 10^{-11} \cdot \mu_0\mu_{15}^{H} - 0.028 \cdot \mu_5^{H} - 1.981 \times 10^{-4} \cdot \mu_3^{PS} + 5.518 \times 10^{-6} \cdot \mu_4^{PS} \quad (1)$$

$$N = 32 \quad S = 0.615 \quad R^2 = 0.769 \quad F = 17.37$$

$$p < 0.0001 \quad S_{cv} = 0.784 \quad q^2 = 0.734$$

where N is the number of compounds included in the model, R^2 is the correlation coefficient, S the standard deviation of the regression, F the Fisher ratio, q^2 the correlation coefficient of the cross-validation, S_{cv} the standard deviation of the cross-validation, and p is the significance of the variables in the model.

The variables included in the model were designed as follow: the sub-index represents the order of the spectral moment and the super-index the type of bond weight used, that is, PS for polar surface, GM for Gasteiger–Marsili charges, and H for hydrophobicity.

From the statistical point of view this model is a robust one as can be seen from the statistical parameters of the cross-validation.

One of the objectives of the current work is to compare the reliability of the TOPS-MODE approach to describe the property under study as compared with other different descriptors and methods. Consequently, we have developed other nine models using the same data set that was included in the TOPS-MODE QSAR model. The results obtained with the use of Constitutional, Topological, BCUT, 2D autocorrelations, Geometrical, RDF, 3D Morse, WHIM, and GETAWAY descriptors are given in Table 2.

As can be seen there are remarkable differences concerning the explanation of the experimental variance given by these models compared to the TOPS-MODE one. While the TOPS-MODE QSAR model explains more than 76% of activity the rest of the models is unable to explain more than 70% of such variance, besides of present important statistic parameters of higher quality that all the models obtained, such as the Fischer ratio (F) and the standard deviation (S).

The TOPS-MODE model not only overtakes the other nine models in the statistical parameters of the regression but more importantly in the stability to the inclusion–exclusion of compounds as measured by the correlation coefficient and standard deviation of the cross-validation. These statistics of the leave-one-out cross validation might be considered as a good measurement of the predictability of the models. As can be seen in the Table 2 the value of the determination coefficient of leave-one-out cross validation for the

Table 2. The statistical parameters of the lineal regressions models obtained for the 10 kinds of descriptors

Descriptors	Variables	S	R	R^2	F	q^2	S_{cv}
Spectral moments	$\mu_1\mu_3^{GM}$, $\mu_0\mu_{15}^{H}$, μ_5^{H}, μ_3^{PS}, μ_4^{PS}	0.615	0.877	0.769	17.37	0.714	0.784
Constitutional	Me, nDB, nC, nN, nX	0.934	0.685	0.469	4.609	0.416	1.215
Topological	Jhetv, PJI2, BIC0, BIC2, T(S..I)	0.810	0.774	0.599	7.815	0.525	1.114
BCUT	BELm4, BELm8, BEHv6, BELv8, BELe8	0.715	0.830	0.689	11.535	0.594	0.984
2D autocorrelations	ATS2p, MATS3v, MATS8e, GATS1v, GATS2e	0.774	0.797	0.635	9.059	0.528	1.112
Geometrical	MEcc, FDI, PJI3, SEig, G(N..N)	0.745	0.814	0.663	10.221	0.564	1.002
RDF	RDF020u, RDF020m, RDF065m, RDF070m, RDF055v	0.723	0.825	0.681	11.157	0.591	0.991
3D-MORSE	Mor09u, Mor12u, Mor16u, Mor26u, Mor32m	0.659	0.837	0.701	14.485	0.598	0.988
WHIM	E3v, G2p, E1p, Dm, Dv	0.891	0.719	0.517	5.572	0.421	1.207
GETAWAY	H4u, H0v, H4e, R1m+, RTe	0.819	0.769	0.591	7.536	0.519	1.093
Mixed model	BELm8, BEHv6, Mor12u, RDF065m, R1m+	0.645	0.849	0.721	13.37	0.644	0.872

model obtained with the spectral moments ($q^2 = 0.714$) was the highest for the all analyses model proving the high predict power of this approach and the high stability of the model. In addition this model presents the lowest value of the standard deviation of the cross-validation ($S_{cv} = 0.784$), therefore the previous criteria is reaffirmed.

However, in all previous studies we only consider models with a specified family of molecular descriptors. Thence, in order to complete the demonstration of the potentialities of TOPS-MODE over the remnant ones mixed models considering all the molecular descriptors at the same time must be developed. The total number of molecular descriptors considered here is higher than 1000. Thus, a strategy for feature selection is necessary. In this sense, we performed a genetic algorithm previous to forward stepwise regression analysis. Table 2 depicts the results of this study. It is interesting to note that the model obtained is not superior from statistical point of view at model reported for the spectral moments. For that reasons, these results have shown that the TOPS-MODE approach not only explains the experimental data, but seems to be the best one in doing so.

4. Concluding remarks

We have shown that the TOPS-MODE approach is able to explain the adenosine kinase inhibitory activity of pyrrolo[2,3-*d*]-pyrimidine nucleoside analogues. In fact, we have developed a model for predicting the inhibitory activity of a data set of 32 pyrrolo[2,3-*d*]-pyrimidine nucleoside, which is both statistically and chemically sounded. This model explains more than 76% of the variance in the experimental activity with a good predictive power. These features are significantly better than that obtained from nine other different methodologies contained in the Dragon Software. For these reasons, we can assert that the TOPS-MODE approach may be used as an efficient alternative to develop a new analogues of the adenosine kinase inhibition.

Acknowledgements

We thank the Universidad de Vigo and the Xunta de Galicia (PGIDT01PX130114PR) for financial support. In addition, the authors would like to express their more sincerely gratitude to unknown referees and the editor Prof. Dr. Dale L. Boger by useful comments and kind attention.

References and notes

1. Jacoboson, K. A.; van Galen, P. J. M.; Williams, M. J. *J. Med. Chem.* **1992**, *35*, 407.
2. Marlon, C.; Chih-Hung, L.; Gregory, A. G.; Erol, K. B.; Shripad, S. B.; Stewart, H. Y.; Kathy, L. K.; Steve, M.; Carol, T. W.; Joseph, M.; Chang, Z.; Karen, M. A.; Michael, F. J.; Elizabeth, A. K. *Bioorg. Med. Chem. Lett.* **2001**, *1*, 83.
3. Marangos, P. *Med. Hypotheses* **1990**, *32*, 45.
4. Chih-Hung, L.; Jerome, F. D.; Meiqun, J.; Haixia, Y.; Kathy, L. K.; Karen, A.; Michael, F. J.; Elizabeth, L. K.; Shripad, S. B. *Bioorg. Med. Chem. Lett.* **2001**, *17*, 2419.
5. Guo, Z. Z.; Chih-Hung, L.; John, K. P.; Richard, J. P.; Mei, Q. J.; Arthur, G.; Mark, M.; Yui, M.; John, R. K.; Ki, H. K.; Steve, M.; Haixia, Y.; Kathy, K.; Karen, M. A.; Steve, M.; Katharine, L. C.; Carol, T. W.; Joseph, M.; Michael, F. J.; Kennan, M.; Elizabeth, A. K.; Shripad, S. B.; Andrew, O. S. *Bioorg. Med. Chem. Lett.* **2001**, *6*, 2071.
6. Jarvis, M. F.; Yu, H.; Kohlhaas, K.; Alexander, K.; Lee, C. H.; Jiang, M.; Bhagwat, S. S.; Williams, M.; Kowaluk, E. A. *J. Pharmacol. Exp. Ther.* **2000**, *295*, 1156.
7. Kowaluk, E. A.; Mikusa, J.; Wismer, C. T.; Zhu, C. Z.; Schweitzer, E.; Lynch, J. J.; Lee, C. H.; Jiang, M.; Bhagwat, S. S.; Gomtsyan, A.; McKie, J.; Cox, B. F.; Polakowski, J.; Reinhart, G.; Williams, M.; Jarvis, M. F. *J. Pharmacol. Exp. Ther.* **2000**, *295*, 1165.
8. Cronstein, B. N.; Naime, D.; Firestein, G. *Arthritis Rheum.* **1995**, *38*, 1040.
9. Estrada, E.; Peña, A. *Bioorg. Med. Chem.* **2000**, *8*, 2755.
10. Estrada, E.; Uriarte, E. *SAR and QSAR Environ. Res.* **2001**, *12*, 309.
11. González, M. P.; González, H. D.; Molina, R. R.; Cabrera, M. A.; Ramos, R. A. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1192.
12. Todeschini, R., Consonni, V., Pavan, M. Dragon Software version 2.1, 2002.
13. Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors*; Wiley-VCH: Weinheim, Germany, 2000.
14. Srikanth, K.; Bikash, D.; Tarun J. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 899.
15. Gutierrez, Y., Estrada, E. Modes Lab®, 2002, version 1.0 b.
16. Michael, J. S.; Dewar, E.; Zoebisch, G.; Eamonn, F.; Stewart, J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
17. Stewar, J. J. P., 1990; MOPAC manual, 6th ed.; p 189, Frank J. Seiler Research Laboratory, U.S. Air Force academy, Colorado Springs, CO.