

## Exploring Potency and Selectivity Receptor Antagonist Profiles Using a Multilabel Classification Approach: The Human Adenosine Receptors as a Key Study

Lisa Michielan,<sup>†</sup> Federico Stephanie,<sup>‡</sup> Lothar Terfloth,<sup>⊥</sup> Dimitar Hristozov,<sup>⊥</sup> Barbara Cacciari,<sup>§</sup> Karl-Norbert Klotz,<sup>||</sup> Giampiero Spalluto,<sup>‡</sup> Johann Gasteiger,<sup>⊥</sup> and Stefano Moro<sup>\*,†</sup>

Molecular Modeling Section (MMS), Dipartimento di Scienze Farmaceutiche, Università di Padova, via Marzolo 5, I-35131 Padova, Italy, Dipartimento di Scienze Farmaceutiche, Università degli Studi di Trieste, Piazzale Europa 1, I-34127 Trieste, Italy, Dipartimento di Scienze Farmaceutiche, Università degli Studi di Ferrara, Via Fossato di Mortara 17-19, I-44100 Ferrara, Italy, Institut für Pharmakologie und Toxikologie, Universität Würzburg, Versbacher Strasse 9, D-97078 Würzburg, Germany, and Molecular Networks GmbH, Henkestraße 91, D-91052 Erlangen, Germany

Received August 19, 2009

Nowadays, in medicinal chemistry adenosine receptors represent some of the most studied targets, and there is growing interest on the different adenosine receptor (AR) subtypes. The AR subtypes selectivity is highly desired in the development of potent ligands to achieve the therapeutic success. So far, very few ligand-based strategies have been investigated to predict the receptor subtypes selectivity. In the present study, we have carried out a novel application of the multilabel classification approach by combining our recently reported autocorrelated molecular descriptors encoding for the molecular electrostatic potential (*auto*MEP) with support vector machines (SVMs). Three valuable models, based on decreasing thresholds of potency, have been generated as *in series* quantitative sieves for the simultaneous prediction of the hA<sub>1</sub>R, hA<sub>2A</sub>R, hA<sub>2B</sub>R, and hA<sub>3</sub>R subtypes potency profile and selectivity of a large collection, more than 500, of known inverse agonists such as xanthine, pyrazolo-triazolo-pyrimidine, and triazolo-pyrimidine analogues. The robustness and reliability of our multilabel classification models were assessed by predicting an internal test set. Finally, we have applied our strategy to 13 newly synthesized pyrazolo-triazolo-pyrimidine derivatives inferring their full adenosine receptor potency spectrum and hAR subtypes selectivity profile.

### INTRODUCTION

Adenosine receptors are widely considered interesting and promising therapeutic targets. In the past decade, the growing knowledge about the different adenosine receptor subtypes has inspired the development of potent and selective ligands.<sup>1,2</sup> Nowadays, this objective represents one of the most ambitious challenges in the field of medicinal chemistry research. During the optimization step of the drug discovery process, the general aim is to design drugs more effective in the therapeutic treatment, but with minimum side effects. If compounds do not differentiate between receptor subtypes, their therapeutic application might be accompanied by efficacy problems or side effects. Therefore, after the single receptor subtype that is responsible for a particular function is identified, the drug candidates may be sifted out to select the compounds attaining high potency profile and subtype selectivity.

So far, very few ligand-based strategies have been investigated to predict the receptor subtypes selectivity.<sup>3,4</sup> On the other hand, various *in silico* tools have been explored for the prediction of the receptor–ligand binding affinity.

In more detail, the generation of quantitative structure–activity relationships (QSARs) represents one of the most applied and reliable approaches to mathematically correlate the molecular properties and the corresponding receptor–ligand binding affinity to precisely predict the binding affinity of new compounds.<sup>5,6</sup> However, only few pioneer studies suggest an integration of both traditional classification and regression analysis as a useful filtering strategy to select potent and selective ligands.<sup>7</sup>

In the past few years, an intensive exploration of the chemical space has been pursued to discover new highly potent and selective adenosine receptors (ARs) antagonists. Briefly, the adenosine receptor family belongs to GPCR family A, including four different subtypes, referred to as A<sub>1</sub>, A<sub>2A</sub>, A<sub>2B</sub>, and A<sub>3</sub>, which are widely but differentially distributed in the tissues.<sup>1,2</sup> Moreover, they have been cloned from various mammalian species, where they differentiate for both their pharmacological profile and effector coupling.<sup>8</sup> Diverse potent and selective ligands for each subtype have demonstrated the potential therapeutic role of the adenosine receptor in several physiopathological processes.<sup>9–17</sup> In particular, A<sub>1</sub>R selective antagonists have shown anxiolytic effects, and they have been reported as promising candidates for the treatment of cognitive disorders, such as dementia. The antagonism selectivity for hA<sub>1</sub>R is also the proposed mechanism for some diuretic agents, which are considered effective in congestive heart failure and in edema.<sup>1,9</sup> A<sub>2A</sub>R antagonists have a neuroprotective activity during ischemic

\* Corresponding author phone: +39 049 8275704; fax: +39 049 8275366; e-mail: stefano.moro@unipd.it.

<sup>†</sup> Università di Padova.

<sup>‡</sup> Università degli Studi di Trieste.

<sup>§</sup> Università degli Studi di Ferrara.

<sup>||</sup> Universität Würzburg.

<sup>⊥</sup> Molecular Networks GmbH.

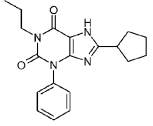
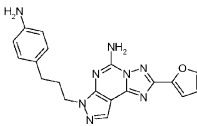
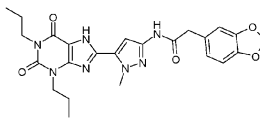
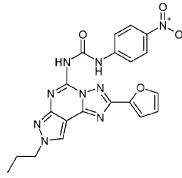
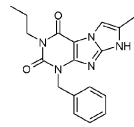
processes and seem to play a role in the reduction of neuronal damage in Parkinson's or Huntington's diseases.<sup>1,12–14</sup> A potential therapeutic activity in the asthma disease has been discovered for  $A_{2B}R$  selective antagonists or mixed antagonists to  $hA_{2B}Rs$  and  $hA_3Rs$ .<sup>1</sup>  $A_{2B}R$  antagonists are also studied as hypoglycaemic agents in diabetes, while  $A_3R$  antagonists have a potential application in tumor growth inhibition and in the treatment of glaucoma.<sup>1,15–17</sup> On the basis of different molecular scaffolds, both receptor-based and ligand-based drug design approaches have been applied for the discovery of more potent and selective hARs. The xanthine and pyrazolo-triazolo-pyrimidine scaffolds have been properly modified to introduce novelty in the chemical space of known adenosine receptors antagonists.<sup>18–20</sup> In particular, improving selectivity for the human  $A_1R$  subtype has been obtained by decorating the classical xanthine scaffold with 8-aryl or 8-cycloalkyl substituents.<sup>21</sup> On the other hand,  $hA_{2B}R$  selective antagonists have been developed with different substitutions at the  $N^1$ ,  $N^3$ , and 8 positions in the same xanthine scaffold.<sup>22–24</sup> Regarding the pyrazolo-triazolo-pyrimidine derivatives, position  $N^7$  has been suggested to be crucial for the selectivity to the human  $A_{2A}R$  subtype.<sup>18</sup> Conversely, proper substituents at the  $N^5$  and  $N^8$  positions shift the antagonism toward the human  $A_3R$  subtype.<sup>20</sup> Furthermore, an alternative imidazo[2,1-f]purinone scaffold has been discovered to improve potency and selectivity for the human  $A_3R$  antagonists.<sup>25</sup> This structural information is summarized in Table 1.

Unfortunately, a limited number of known AR antagonists has been synthesized and tested on all four human AR (hAR) subtypes to determine their binding affinity. In the past, most of the literature partially reported the binding affinity to some hAR subtypes or the data set was obtained by using ARs cloned from other mammalian species. In Figure 1, the potency and selectivity spectrum of known hAR antagonists, considering only homogeneous data on human ARs, is summarized.

Considering the limited information on  $hA_1R$  subtype and the available experimental data on hAR antagonists, it is very difficult to correctly predict the complete hAR potency profiles and infer the selectivity of novel xanthine and pyrazolo-triazolo-pyrimidine derivatives.

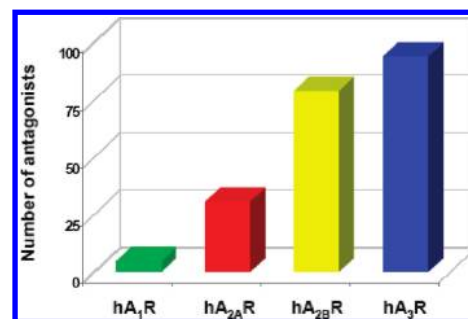
In the present Article, we would like to demonstrate how a novel application of the multilabel classification approach by combining our well-performing autocorrelated molecular descriptors encoding for the molecular electrostatic potential (*autoMEP*) vectors with support vector machine (SVM) analysis can represent a very powerful tool to simultaneously describe the  $hA_1R$ ,  $hA_{2A}R$ ,  $hA_{2B}R$ , and  $hA_3R$  potency profiles and identify the possible subtype selectivity for hAR antagonists.<sup>26–30</sup> SVM is widely applied as a supervised learning technique to solve both classification and regression problems.<sup>31,32</sup> In the last years, various classification results have been reported in several papers.<sup>33–36</sup> Very recently, we have developed an integrated SVM-SVR method by using the *autoMEP* molecular descriptors to discriminate  $A_{2A}R$  versus  $A_3R$  antagonists and to predict the binding affinity to the corresponding receptor subtype.<sup>7</sup> This work has clarified that the classification approach is a valuable tool to predict the receptor subtypes selectivity. However, in the traditional single-label classification, classes are considered mutually exclusive. In the present classification task, some samples

**Table 1.** Examples of Potent and Selective AR Antagonists to the Four Subtypes  $A_1$ ,  $A_{2A}$ ,  $A_{2B}$ , and  $A_3$

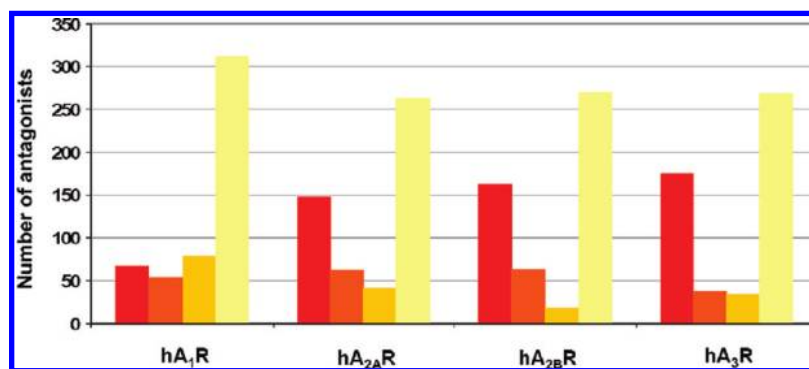
Structures	Exp. binding affinity data	References
<i>A<sub>1</sub> adenosine receptor antagonists</i>		
	$hA_1R$ $K_i$ = 7.1 nM $hA_{2A}R$ $K_i$ = 1200 nM $hA_{2B}R$ $K_i$ = 625 nM $hA_3R$ $K_i$ = 395 nM	21
<i>A<sub>2A</sub> adenosine receptor antagonists</i>		
	$hA_1R$ $K_i$ = 2160 nM <b><math>hA_{2A}R</math> <math>K_i</math> = 0.22 nM</b> $hA_{2B}R$ $K_i$ > 10,000 nM $hA_3R$ $K_i$ > 10,000 nM	18
<i>A<sub>2B</sub> adenosine receptor antagonists</i>		
	$hA_1R$ $K_i$ = 566 nM $hA_{2A}R$ $K_i$ > 1000 nM <b><math>hA_{2B}R</math> <math>K_i</math> = 18 nM</b> $hA_3R$ $K_i$ > 1000 nM	24
<i>A<sub>3</sub> adenosine receptor antagonists</i>		
	$hA_1R$ $K_i$ = 1214 nM $hA_{2A}R$ $K_i$ = 1115 nM $hA_{2B}R$ $K_i$ = 305 nM <b><math>hA_3R</math> <math>K_i</math> = 0.81 nM</b>	20
	$hA_1R$ $K_i$ > 1,000 nM $hA_{2A}R$ $K_i$ > 1,000 nM $hA_{2B}R$ $K_i$ > 1,000 nM <b><math>hA_3R</math> <math>K_i</math> = 0.8 nM</b>	25

belong to multiple classes, because the hAR antagonists may present a good potency profile for more subtypes. The multilabel classification analysis seems to be appropriate, whether our data set deals with nonmutually exclusive and overlapping classes. In this field, a novel multilabel classification technique, cross-training with support vector machines (ct-SVMs), has been recently reported.<sup>37,38</sup>

In the present study, the combination of *autoMEP* vectors with ct-SVM analysis (*autoMEP*/ct-SVM) represents a novel strategy for the prediction of the complete hARs potency profiles and infer hAR subtypes selectivity of known xanthine and pyrazolo-triazolo-pyrimidine analogues. Interestingly, our *autoMEP*/ct-SVM approach has been extended to all four hAR subtypes. In more detail, a large collection of hAR



**Figure 1.** Distribution representation of the known selective and potent ( $K_i \leq 100$  nM) hAR antagonists to the different subtypes ( $A_1$ ,  $A_{2A}$ ,  $A_{2B}$ , and  $A_3$ ).



**Figure 2.** Distribution representation of the experimental  $K_i$  binding affinity data of hAR antagonists in our data set (514 molecules), including training set, validation set, and *internal* test set. The number of hAR antagonists are reported in different ranges of binding affinity for each subtype:  $K_i \leq 100$  nM (red),  $100 \text{ nM} < K_i \leq 250$  nM (orange),  $250 \text{ nM} < K_i \leq 500$  nM (dark yellow), and  $K_i > 500$  nM (light yellow). The multiple classes hA<sub>1</sub>R, hA<sub>2A</sub>R, hA<sub>2B</sub>R, and hA<sub>3</sub>R are completely overlapping.

**Table 2.** Data Distribution Indicating the Number of Selective and Potent, Not Selective and Potent, and Not Selective and Not Potent hAR Antagonists in the Training, Validation, and Test Sets, According to the  $K_i$  Value for Each hAR Subtype

subtypes	training set			validation set			internal test set		
	selective and potent <sup>a</sup>	not selective but potent <sup>b,c</sup>	not selective and not potent <sup>d</sup>	selective and potent <sup>a</sup>	not selective but potent <sup>b,c</sup>	not selective and not potent <sup>d</sup>	selective and potent <sup>a</sup>	not selective but potent <sup>b,c</sup>	not selective and not potent <sup>d</sup>
hA <sub>1</sub> R	3	86		1	15		1	42	
hA <sub>2A</sub> R	22	94		3	23		6	48	
hA <sub>2B</sub> R	48	72		9	16		22	22	
hA <sub>3</sub> R	64	84		10	20		20	42	
total per column	137	148	33	23	36	6	49	67	15
total		318			65			131	

<sup>a</sup> hAR antagonists with a difference of at least 2 orders of magnitude between the corresponding  $K_i$  values and the minimum  $K_i$  value lower than 100 nM to the corresponding subtype. <sup>b</sup> Not selective antagonists belong to more than one class/subtype. Consequently, the sum of not selective hAR antagonists for all classes within a column is higher than the total hAR antagonists for the training set, the validation set, and the test set. <sup>c</sup> Not selective hAR antagonists with a difference lower than 2 orders of magnitude between the corresponding  $K_i$  values and all  $K_i$  values lower than 500 nM. <sup>d</sup> Not selective hAR antagonists with a difference lower than 2 orders of magnitude between the corresponding  $K_i$  values and all  $K_i$  values higher than 500 nM.

antagonists has been utilized to carry out and validate three *autoMEP/ct-SVM* models. They have been applied *in series* as quantitative sieves, based on decreasing thresholds of potency (500, 250, and 100 nM), corresponding to different binding affinity  $K_i$  values. For the further validation of our strategy, we have synthesized 13 new pyrazolo-triazolo-pyrimidine derivatives to inspect their A<sub>1</sub>R, A<sub>2A</sub>R, A<sub>2B</sub>R, and A<sub>3</sub>R potency profiles. The prediction results confirm the *autoMEP/ct-SVM* strategy as a valuable tool to select potent and selective hAR antagonists.

## MATERIALS AND METHODS

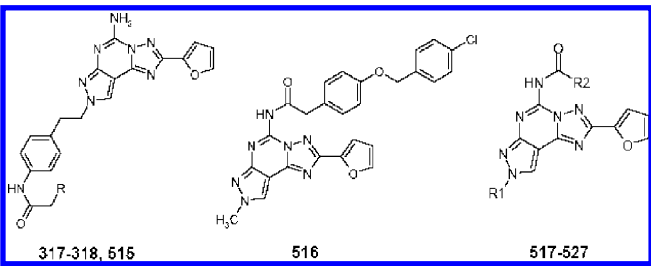
**Data Set.** The same collection of 514 compounds, comprising 512 known and two newly synthesized human A<sub>1</sub>, A<sub>2A</sub>, A<sub>2B</sub>, and A<sub>3</sub> adenosine receptor antagonists, has been used to derive and validate our three *autoMEP/ct-SVM* multilabel classification models.<sup>7,18–25,28,30,39–43</sup> This large data set was split into training set, validation set, and *internal* test set. The potency profiles distribution for each hAR subtype and the information on selectivity and potency are summarized in Figure 2 and Table 2, respectively.

A similar distribution of the corresponding potency profiles intervals is present for hA<sub>2A</sub>R, hA<sub>2B</sub>R, and hA<sub>3</sub>R subtypes, with a noteworthy number of potent antagonists having a binding affinity  $K_i$  value lower than or equal to 100 nM to the distinct subtypes (Figure 2). Concerning the hA<sub>1</sub>R subtype, our collection is short of potent compounds, while many hA<sub>1</sub>R antagonists have a binding affinity  $K_i$  value

higher than 250 nM. Few hA<sub>1</sub>R selective antagonists are also reported in Table 2. A low rate “number of selective and potent antagonists/number of not selective but potent antagonists” can be observed for the hA<sub>1</sub>R class in comparison to the other subtypes. In the final collection, 209 hAR antagonists (41%) are selective for one hAR subtype, with the corresponding  $K_i$  values lower than 100 nM. As anticipated in Figure 1, considering these 209 compounds, five are selective for hA<sub>1</sub>R (2%), 31 are selective for hA<sub>2A</sub>R (15%), 79 are hA<sub>2B</sub>R selective antagonists (38%), and the remaining 94 are hA<sub>3</sub>R selective antagonists (45%). Our data set is quite unbalanced; however, this distribution reflects the state of the art about the investigations on hAR antagonists. In the following, the training set, validation set, and test sets are separately presented.

**Training Set.** A collection of 318 xanthine derivatives, N<sup>7</sup>- and N<sup>8</sup>-substituted pyrazolo-triazolo-pyrimidine analogues (**1–318**), has been selected as training set in the *autoMEP/ct-SVM* models.<sup>7,18–25,30,39–43</sup> In particular, two N<sup>8</sup>-substituted new pyrazolo-triazolo-pyrimidine derivatives (**317**, **318**) have been synthesized. See the Supporting Information for details.

**Validation Set.** A collection of 65 xanthine derivatives, N<sup>7</sup>- and N<sup>8</sup>-substituted pyrazolo-triazolo-pyrimidine derivatives (**319–383**), hA<sub>1</sub>R, hA<sub>2A</sub>R, hA<sub>2B</sub>R, and hA<sub>3</sub>R selective and nonselective antagonists with different potency profiles, has been used as validation set of the three *autoMEP/ct-SVM* models.<sup>7,18–25,40–43</sup> See the Supporting Information for details.

**Table 3.** Biological Profile at the Four Human Adenosine Receptor Subtypes of Synthesized Compounds (**317**, **318**, **515**–**527**)


compd	R	R <sub>1</sub>	R <sub>2</sub>	hA <sub>1</sub> (K <sub>i</sub> nM) <sup>a</sup>	hA <sub>2A</sub> (K <sub>i</sub> nM) <sup>b</sup>	hA <sub>2B</sub> (K <sub>i</sub> nM) <sup>c</sup>	hA <sub>3</sub> (K <sub>i</sub> nM) <sup>d</sup>
<b>317</b>	4-CH <sub>3</sub> O-Ph			1221 (1710–2400)	21.2 (13.4–33.4)	>10 000	1760 (904–3420)
<b>318</b>	4-Cl-Ph-O			187 (133–262)	67.3 (62.4–72.7)	>10 000	1490 (991–2240)
<b>515</b>	α-naphthyl			1250 (1080–1450)	87.3 (78.2–97.6)	3700 (2283–6000)	1537 (1210–1960)
<b>516</b>				1710 (1500–1970)	2,520(2,030–3120)	>10 000	25.5 (13.0–49.8)
<b>517</b>	CH <sub>3</sub>		CH <sub>2</sub> -O-Ph	42 (374–523)	15.1 (11.5–19.8)	>10 000	692 (612–782)
<b>518</b>	CH <sub>2</sub> CH <sub>3</sub>		CHPh <sub>2</sub>	156 (102–239)	131 (127–135)	248 (160–3383)	0.98 (0.54–1.75)
<b>519</b>	CH <sub>2</sub> CH <sub>2</sub> CH <sub>3</sub>		CH <sub>2</sub> -2-thienyl	228 (205–254)	6.21 (3.97–9.72)	2853 (2073–3935)	255 (233–280)
<b>520</b>	CH <sub>2</sub> CH <sub>2</sub> CH <sub>3</sub>		CH <sub>2</sub> -3-thienyl	170 (145–199)	4.51 (3.25–6.27)	2820 (2190–3800)	3.14 (2.35–4.19)
<b>521</b>	CH <sub>2</sub> CH <sub>2</sub> CH <sub>2</sub> CH <sub>3</sub>		CH <sub>2</sub> -β-naphthyl	113 (106–122)	7.94 (4.32–14.6)	1200 (684–2106)	174 (152–199)
<b>522</b>	CH <sub>2</sub> CH <sub>2</sub> CH(CH <sub>3</sub> ) <sub>2</sub>		CH <sub>2</sub> -4-OCH <sub>3</sub> -Ph	22.6 (18.1–28.1)	1.95 (1.01–3.71)	460 (243–870)	359 (190–678)
<b>523</b>	CH <sub>2</sub> CH <sub>2</sub> CH(CH <sub>3</sub> ) <sub>2</sub>		CH <sub>2</sub> -2-thienyl	20.5 (14.6–28.8)	2.03 (1.57–2.63)	2030 (1715–2410)	308 (214–444)
<b>524</b>	CH <sub>2</sub> CH <sub>2</sub> CH(CH <sub>3</sub> ) <sub>2</sub>		CH <sub>2</sub> -3-thienyl	31.5 (24.5–40.6)	3.88 (2.55–5.89)	2487 (1660–3735)	540 (481–607)
<b>525</b>	CH <sub>2</sub> CH <sub>2</sub> CH(CH <sub>3</sub> ) <sub>2</sub>		CH <sub>2</sub> -O-Ph-4-Cl	155 (103–233)	16.1 (13.9–18.7)	>10 000	2306 (2000–2660)
<b>526</b>	CH <sub>2</sub> CH <sub>2</sub> Ph		CH <sub>2</sub> -O-Ph-4-Cl	199 (173–228)	31.7 (24.3–41.5)	>30 000	3251 (1830–5770)
<b>527</b>	CH <sub>2</sub> CH <sub>2</sub> CH <sub>2</sub> Ph		CH <sub>2</sub> -4-OCH <sub>3</sub> -Ph	44.8 (40.3–49.9)	8.93 (5.86–13.6)	2690 (2225–3243)	120 (108–133)

<sup>a</sup> Displacement of specific [<sup>3</sup>H]-CCPA binding at human A<sub>1</sub> receptors expressed in CHO cells (*n* = 3–6). <sup>b</sup> Displacement of specific [<sup>3</sup>H]-NECA binding at human A<sub>2A</sub> receptors expressed in CHO cells. <sup>c</sup> K<sub>i</sub> values of the inhibition of NECA-stimulated adenylyl cyclase activity in CHO cells expressing hA<sub>2B</sub> receptors. <sup>d</sup> Displacement of specific [<sup>3</sup>H]-NECA binding at human A<sub>3</sub> receptors expressed in CHO cells. Data are expressed as geometric means, with 95% confidence limits.

**Test Sets.** An *internal* test set of 131 N<sup>7</sup>- and N<sup>8</sup>-substituted pyrazolo-triazolo-pyrimidine analogs and xanthine derivatives (**384**–**514**) has been selected to validate our three *autoMEP*/ct-SVM models.<sup>7,18–25,28,39–43</sup> See the Supporting Information for details.

Finally, an additional (*external*) test set of 13 newly synthesized compounds having the pyrazolo-triazolo-pyrimidine scaffold decorated in N<sup>8</sup> and 5 positions (**515**–**527**) has been selected for the further validation of our three *autoMEP*/ct-SVM models. The structures and pharmacological profiles of all new synthesized antagonists are reported in Table 3.

**Computational Methodologies.** All modeling studies were carried out on a linux cluster running under open-Mosix architecture.<sup>44</sup> Molecular structure building and autocorrelation MEP descriptors have been carried out using *ADRIANA.Code* software (version 2.2).<sup>45</sup> Cross-training SVM (ct-SVM) multilabel classification models were generated with the R software (package e1071).<sup>46,47</sup>

**Molecular Structure Building.** 3D models of all hA<sub>1</sub>R, hA<sub>2A</sub>R, hA<sub>2B</sub>R, and hA<sub>3</sub>R antagonists in the training set, validation set, *internal*, and *external* test set were obtained using the 3D structure generator Corina, which is an integral part of *ADRIANA.Code*, setting parameters to standard values.<sup>45</sup> Conformer selection is one of the most crucial steps in every approach considering 3D molecular descriptors. Unfortunately, the information about the possible binding mode of all human A<sub>1</sub>, A<sub>2A</sub>, A<sub>2B</sub>, and A<sub>3</sub> receptor antagonists is very limited, and, consequently, we have decided to select the energetically most stable conformers produced by the software conformational analysis. However, we are intensively working in modeling the 3D structure of all human adenosine receptor subtypes with the aim to integrate the receptor-based conformer selection (using

conventional docking approaches) as direct input to MEP descriptors calculation. For these reasons, we decided to first validate our novel classification approach, to use for the lowest energy conformation selected by Corina as prototype of the “bioactive” conformation. We perfectly understand the limits of our selection criteria, but we can consider this a reasonable compromise to standardize the conformational selection.

Only few analyzed molecules are affected by dissociation processes (11 of 514). Only for these protonation states selected in agreement with the corresponding predicted pK<sub>a</sub> at the physiological pH value (7.4 unit).

**Molecular Electrostatic Potential (MEP) Calculation.** Autocorrelation MEP vectors have been introduced by Gasteiger and collaborators as molecular descriptors computed on the molecular surface.<sup>48</sup> In the present work, MEPs derive from a classical point charge model: the electrostatic potential for each molecule is obtained by moving a unit positive point charge across the molecular surface, and it is calculated at various points *j* on this surface by the following equation:

$$V_j = \frac{1}{4\pi\epsilon_0\epsilon_r} \sum_i^{\text{atoms}} \frac{q_i}{r_{ij}}$$

where  $\epsilon_0$  is vacuum permittivity,  $\epsilon_r$  is the relative permittivity or dielectric constant,  $q_i$  represents the partial charge of each atom *i*, and  $r_{ij}$  is the distance between points *j* and atom *i*. Starting from the 3D model of a molecule and its partial atomic charges, the electrostatic potential is calculated for points on the molecular surface. Partial atomic charges were calculated by the PEOE (partial equalization of orbital electronegativity) method and its extension to conjugated systems implemented in *ADRIANA.Code*.<sup>45,49,50</sup> Connolly's



solvent accessible surface with a solvent radius of 2.0 Å has been used to project the corresponding MEP. Once the autocorrelation function has been applied, the autocorrelation vector is derived. Connolly's solvent accessible surfaces and the corresponding MEPs have been calculated by *ADRIANA.Code*.<sup>45,49,50</sup>

**Autocorrelation MEP (autoMEP) Vectors.** The autocorrelation function transforms the constitution of a molecule into a fixed length representation. First investigated by Moreau and Broto, this concept was introduced in chemistry to analyze the properties of different molecules without molecular superimposition.<sup>51,52</sup> They considered that a certain property  $p$  of an atom  $i$  can be correlated with the corresponding property  $p$  of atom  $j$  and the products of  $p$  values can be summed over all atom pairs having a certain topological distance  $d$ . Each component of the autocorrelation vector is consequently calculated as follows:

$$A(d) = \sum_{ij} p_i p_j \delta(d_{ij}, d) \quad \delta = \begin{cases} 1 & \forall d_{ij} = d \\ 0 & \forall d_{ij} \neq d \end{cases}$$

where  $A$  is the autocorrelation coefficient referring to atom pairs  $i, j$  at the  $i, j$  topological distance  $d$  and  $p_i$  is the atomic property.<sup>51,52</sup>

Ligands and proteins interact through molecular surfaces, and, therefore, representations of molecular surfaces have also to be sought in the endeavor to understand the difference in the potency profiles to distinct receptor subtypes. Again, we are under the restriction of having to represent molecular surfaces of different size; thus Gasteiger and collaborators employed the autocorrelation concept to achieve this goal.<sup>48,53,54</sup> Starting from the topological autocorrelation examples of Moreau and Broto, a set of randomly distributed points on the molecular surface have to be generated first. Next, all distances between the surface points are calculated and sorted into the preset intervals  $d_{\text{lower}}-d_{\text{upper}}$ . So, the autocorrelation coefficients are computed:

$$A(d_{\text{lower}}, d_{\text{upper}}) = \frac{1}{2} \sum_{j=1}^N \sum_{i=1}^N p_i p_j \delta(d_{ij}, d_{\text{lower}}, d_{\text{upper}})$$

$$\delta = \begin{cases} 1 & \forall d_{\text{lower}} < d_{ij} \leq d_{\text{upper}} \\ 0 & \forall d_{ij} \leq d_{\text{lower}} \vee d_{ij} > d_{\text{upper}} \end{cases}$$

The component of the autocorrelation vector  $A(d_{\text{lower}}, d_{\text{upper}})$ , referring to the  $i, j$  distance  $d$  in the interval  $d_{\text{lower}}-d_{\text{upper}}$ , is the sum of all products of the properties  $p_i$  and  $p_j$  for atoms  $i$  and  $j$ . We consider  $L$  a parameter representing the total number of distances in the interval  $d_{\text{lower}}-d_{\text{upper}}$ . The application of this concept made possible the comparison of different molecular properties, as this 3D descriptor represents a compressed expression of the distribution of the property  $p$  on the molecular surface.

For the calculation of the autocorrelation coefficient, we have applied the default values for parameters computation, because no improvement in statistical model capability was observed by changing them in a various way. Default parameters values are the following:  $d_{\text{lower}} = 1$  Å;  $d_{\text{upper}} = 13$  Å;  $L = 12$ ; point density = 10 points/Å<sup>2</sup>; vdW radius reduction factor = 1.000. Consequently, we derived 12 autocorrelation vectors per molecule, computed at the 12 ( $L$  value) distances in the interval from 1 to 13 Å with a step width of 1 Å. By considering the size of the molecules in

our data set, we decided that the step width of 1 Å, derived from the partition in 12 intervals of the global distance 1–13 Å, was sufficient to describe in an accurate way the distribution of the MEP property on the molecular surface. This transformation produces a molecular descriptor, which is a unique fingerprint of each molecule under consideration. Autocorrelation vectors have been calculated by *ADRIANA.Code*.<sup>45</sup>

**Multilabel Classification.** In the present Article, a powerful methodology addressing the multilabel classification approach was investigated in combination with the *autoMEP* descriptors reported above. In particular, we have used cross-training with the recently developed support vector machines (ct-SVM) technique.<sup>37,46,47</sup>

**Support Vector Machine (SVM).** Support vector machines (SVMs) are widely applied supervised learning systems originated from the statistical learning theory, proposed by Vapnik.<sup>55,56</sup> First developed for pattern recognition, here SVMs are utilized to solve function approximation problems, where the training set is represented by pairs of samples,  $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i)\}$ , where  $x_i$  is the input data and  $y_i$  is the corresponding desired value, directly observed from the unknown function. In a binary classification problem, usually  $x_i \in \mathbb{R}^n$ , while  $y_i \in \{-1, +1\}$ . However, the aim of the learning system is to find out in a set of functions, that is, the hypothesis space, the optimal hypothesis function  $f(x)$ , able to approximate the desired responses  $y_i$  by minimizing the risk functional  $R$ . In more detail,  $R$  weights the cost of the approximation, given by a loss or cost function that measures the distance between  $y_i$  and  $f(x)$ , with the probability to observe the input–output couple  $(x_i, y_i)$ . We would like the function  $f(x)$  to be a reasonable estimate of the functional relation between input–output pairs (prediction or generalization property).

The linear SVM is based on: (i) linear hypotheses corresponding to separating hyperplanes, that is,  $f(x) = w \cdot x + b = \sum_{i=1}^n w_i x_i + b$ , where  $\cdot$  is the dot product between vectors; (ii) the solution of a quadratic optimization problem that represents a trade-off between the minimization of the empirical error, that is, the error over the training set, and the maximization of the smoothness of  $f(x)$ .<sup>56</sup> Nonlinear versions of SVM can be obtained by the introduction of a kernel.<sup>57</sup> An example of kernel function is the Gaussian kernel  $k(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2}$ . In support vector classification, binary classifiers are introduced to discriminate a set of compounds in two classes. The standard formulation for SVM is derived by using the Hinge loss function and slack variables  $\xi_i$ :

$$\min_{w, b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i$$

subject to:  $\forall i \in \{1, \dots, n\} y_i(w \cdot x + b) \geq 1 - \xi_i$  and  $\xi_i \geq 0$ , where we recall that  $y_i \in \{-1, +1\}$ ,  $w$  and  $b$  are the parameters that control the function  $f(x)$ , and the constraints are satisfied with zero error when it is possible to find a function that is able to “classify” any positive example ( $y_i = +1$ ) by returning a positive value, that is,  $f(x) \geq 1$ , and any negative example ( $y_i = -1$ ) returning a negative value, that is,  $f(x) \leq -1$ . If such function does not exist, then errors need to be compensated by choosing nonzero values for the corresponding slack variables  $\xi_i$ . The trade-off between the

minimization of the norm of the weight vector and the empirical error is given by the constant  $C$ .

The above quadratic constrained minimization problem can be more easily solved by resorting to the corresponding dual problem:

$$\max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (x_i \cdot x_j)$$

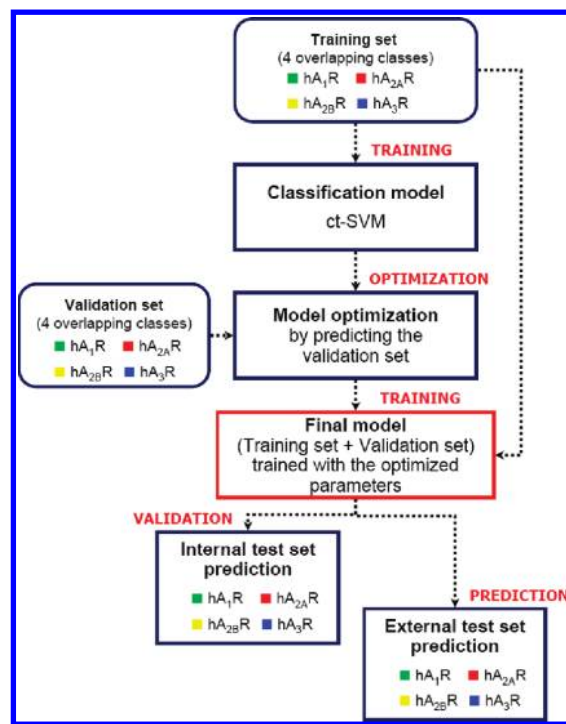
subject to:  $\sum_{i=1}^n y_i \alpha_i = 0$ , and  $\forall i \in \{1, \dots, n\} 0 \leq \alpha_i \leq C$ , where  $\alpha_i$  are called the dual variables. The input vectors  $x_i$  for which the corresponding dual variables satisfy  $\alpha_i^0 > 0$  are referred to as support vectors, where  $\alpha^0$  is the optimal solution vector of the dual problem. Finally, the decision rule is given by  $\text{sgn}(f(x))$ .

Nonlinearity of the boundary separating positive from negative samples is achieved by projecting the input vector into a higher dimensional feature space. In this way, the dot product  $x_i \cdot x_j$  is replaced by a kernel function  $k(x_i, x_j)$ , representing the dot product in the transformed space, that is,  $k(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$ . The decision function takes the final form:

$$f(x) = \text{sgn} \left( \sum_{i=1}^n \alpha_i y_i k(x_i, x) + b \right)$$

**Cross-Training with Support Vector Machines (ct-SVMs) Analysis.** The concept of cross-training was introduced by Boutell and collaborators.<sup>37</sup> They turned from the previously less performing attempted strategies for multilabel data to present a new classification method suitable for multiple and overlapping classes tasks, with samples simultaneously associated with more than one class. In the “cross-training” approach, the multilabel data are used more than once when training the classification model. Moreover, each sample is assigned a positive label for each actual class to which it belongs.<sup>37</sup> ct-SVM represents a novel application of SVMs analysis, when classes overlap in the feature space. In ct-SVM technique, the output real-valued scores of the trained binary classifiers for each class are transformed into the final labels according to different testing criteria, as recently reported.<sup>37,38</sup>

In the present study, three ct-SVM multilabel classification models have been developed to be applied *in series*. For each model, we have considered  $n$  binary classifiers,  $n$  corresponding to the number of the adenosine receptor  $A_1$ ,  $A_{2A}$ ,  $A_{2B}$ ,  $A_3$  subtypes ( $n = 4$ ), and a radial basis function (RBF) kernel has been utilized. The parameters for each classifier ( $C$  and  $\gamma$ ) were automatically optimized on the training set during the learning process by using a 10-fold cross validation and predicting a small validation set.<sup>38</sup> After the cross-training approach was applied, the real-valued scores were obtained. Recently, three different testing criteria (P, T, and C) have been proposed.<sup>37</sup> The P-criterion assigns to the samples all labels corresponding to a positive SVM score. If none of the scores is positive, the sample is classified as “unknown”. The T-criterion uses the closed world assumption (CWA), according to which all samples belong to at least one class. In our classification problem, the classes are overlapping and all SVM scores might be negative, if the samples are correctly assigned to none of the classes. As previously demonstrated, the C-criterion provides the best



**Figure 3.** Workflow for the generation of *autoMEP*/ct-SVM multilabel classification models and the prediction of the hAR antagonists potency profiles.

trade-off between the performance of the classifier on single-label data and multilabel data.<sup>37</sup> In fact, the C-criterion considers SVM scores without any sign, and the decision depends on the closeness between the top SVM scores. In this work, a validation set has been used to select the closeness between two scores. Once each binary classifier was optimized by predicting the validation set, the training and the validation sets were merged, and the new model was computed by using the previously optimized parameters, as illustrated in Figure 3.

Our three final multilabel classification models have been developed for the prediction of an *internal* test set to evaluate their statistical robustness (Figure 3). Moreover, an *external* test set has been predicted to further validate our approach.

**Modeling Results Evaluation.** The evaluation of a multilabel classification model performance is more complicated in comparison with the statistical quality of a single-label classification model.<sup>37</sup> In the present study, the confusion matrix was extracted from the predictions of the validation set and the *internal* test set to assess the robustness of our *autoMEP*/SVM models. In particular, the accuracy is referred to the overall performance on the testing data set, while the number of true positives (TP), false positives (FP), true negatives (TN), false negatives (FN), the recall, and precision are base-class measures, calculated for each class after the comparison between actual and predicted labels by our *autoMEP*/ct-SVM models.<sup>37,38</sup> In the analysis of the *external* test set prediction, we have computed the accuracy for each compound as percentage (%) of correctly predicted labels by MODELS 1, 2, and 3 to the total number of actual labels (12, one for each hAR classifier). In ct-SVM analysis, the ranking process provides a function to order the labels for each sample and to assign scores to the samples. Several ranking-based performance measures have been mathematically defined.<sup>37,38</sup> One-error represents the ratio of the

number of not top-ranked labels to the total number of actual labels. It can take on values between zero and one, and values close to zero indicate a good performance. Coverage measures how far one needs, on average, to go down the list of labels to cover all actual labels. The coverage interval is between one and the number of the classes; then, the best performance corresponds to a value of one. Average precision, which refers to the whole system, reflects the effectiveness of the label ranking and indicates the frequency of the top-ranking for the actual labels. The extreme values are zero and one, and perfect performance is achieved when the average precision is equal to one.

**Chemistry.** The desired compounds (**317**, **318**, **515**–**527**) have been prepared by acylation with the appropriate arylacetyl chloride (1.3 equiv) of the well-known and previously reported unsubstituted derivatives **528**–**534**<sup>41</sup> in dry THF at reflux (18 h) in the presence of triethylamine (1.3 equiv).<sup>58</sup> When not commercially available, the acyl chlorides were prepared from the corresponding acid by treatment with thionyl chloride in the presence of a catalytic amount of dimethyl formamide.<sup>59</sup> Instead, derivatives **317**–**318**, **515**, were prepared by condensation in dry DMF in the presence of EDC of the amino derivatives **535**<sup>39</sup> with commercially available acids.<sup>42</sup>

**General.** Reactions were routinely monitored by thin-layer chromatography (TLC) on silica gel (precoated F<sub>254</sub> Merck plates). Infrared spectra (IR) were measured on a Jasco FT-IT instrument. <sup>1</sup>H NMR spectra were determined in CDCl<sub>3</sub> or DMSO-*d*<sub>6</sub> solutions with a Varian Gemini 200 spectrometer, peaks positions are given in parts per million ( $\delta$ ) downfield from tetramethylsilane as internal standard, and *J* values are given in Hz. Light petroleum ether refers to the fractions boiling at 40–60 °C. Melting points were determined on a Büchi–Tottoli instrument and are uncorrected. Flash chromatography was performed using Merck 60–200 mesh silica gel. Elemental analyses were performed by the microanalytical laboratory of Dipartimento di Chimica, University of Trieste, and were within  $\pm 0.4\%$  of the theoretical values for C, H, and N.

**General Procedures for the Preparation of 5-[(Aryl)carbonyl]amino-8-(ar)alkyl-2-(2-furyl)-pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidine (**516**–**527**).** Amino compound (**528**–**534**, 10 mmol) was dissolved in freshly distilled THF (15 mL), and appropriate acid chloride (13 mmol) and triethylamine (13 mmol) were added. The mixture was refluxed under argon for 18 h. Next, the solvent was removed under reduced pressure, and the residue was dissolved in EtOAc (30 mL) and washed twice with water (15 mL). The organic phase was dried on Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The residue was purified by flash chromatography (EtOAc–light petroleum 4:6) to afford the desired compounds (**516**–**527**).

**5-[[4-(4-Chlorobenzoyloxy)]phenyl-acetyl]amino-8-methyl-2-(2-furyl)-pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidine (**516**).** Yield 78%, yellow solid; mp 140 °C (EtOAc–light petroleum). IR (KBr): 3250–2990, 1685, 1610, 1560, 1500 cm<sup>−1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>)  $\delta$ : 4.18 (s, 3H); 3.27 (s, 3H); 4.44 (s, 2H); 4.98 (s, 2H); 6.59 (dd, 1H, *J* = 2, *J* = 4); 6.93 (d, 2H, *J* = 9); 7.17 (d, 1H, *J* = 4); 7.29 (d, 2H, *J* = 9); 7.33 (s, 4H); 7.82 (d, 1H, *J* = 2); 8.19 (s, 1H); 9.11 (bs, 1H).

**5-[(Phenoxymethyl)carbonyl]amino-8-methyl-2-(2-furyl)-pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidine (**517**).** Yield 74%, brown solid; mp 140 °C (EtOAc–light petroleum). IR (KBr): 3235–2975, 1674, 1618, 1592, 1510 cm<sup>−1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>)  $\delta$ : 4.18 (s, 3H); 5.01 (s, 2H); 6.65 (dd, 1H, *J* = 2, *J* = 4); 6.85–7.08 (m, 3H); 7.15–7.40 (m, 3H); 7.99 (d, 1H, *J* = 2); 8.80 (s, 1H); 10.77 (bs, 1H).

**5-[(Diphenylmethyl)carbonyl]amino-8-ethyl-2-(2-furyl)-pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidine (**518**).** Yield 92%, white solid; mp 205 °C (EtOAc–light petroleum). IR (KBr): 3234–2977, 1676, 1612, 1588, 1515 cm<sup>−1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>)  $\delta$ : 1.42 (t, 3H, *J* = 6.7), 4.21 (q, 2H, *J* = 6.7); 5.64 (s, 1H); 6.68 (dd, 1H, *J* = 2, *J* = 4); 7.04–7.43 (m, 11H); 8.00 (d, 1H, *J* = 2); 8.82 (s, 1H); 11.40 (bs, 1H).

**5-[(2-Thienylmethyl)carbonyl]amino-8-*n*-propyl-2-(2-furyl)-pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidine (**519**).** Yield 75%, hygroscopic brown solid (EtOAc–light petroleum). IR (KBr): 3237–2980, 1672, 1614, 1565, 1505 cm<sup>−1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>)  $\delta$ : 0.91 (t, 3H, *J* = 6.7); 1.81–2.01 (m, 2H); 4.21 (s, 2H); 4.38 (t, 2H, *J* = 6.7); 6.68 (dd, 1H, *J* = 2, *J* = 4); 7.01 (dd, 1H, *J* = 2, *J* = 4); 7.03 (d, 1H, *J* = 2); 7.21 (d, 1H, *J* = 4); 7.39 (d, 1H, *J* = 4); 7.97 (d, 1H, *J* = 2); 8.85 (s, 1H); 11.07 (bs, 1H).

**5-[(3-Thienylmethyl)carbonyl]amino-8-*n*-propyl-2-(2-furyl)-pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidine (**520**).** Yield 63%, hygroscopic brown solid (EtOAc–light petroleum). IR (KBr): 3230–2975, 1670, 1618, 1587, 1508 cm<sup>−1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>)  $\delta$ : 0.87 (t, 3H, *J* = 6.7); 1.78–1.98 (m, 2H); 4.01 (s, 2H); 4.37 (t, 2H, *J* = 6.7); 6.71 (dd, 1H, *J* = 2, *J* = 4); 7.05 (d, 1H, *J* = 4); 7.13 (d, 1H, *J* = 4); 7.39 (s, 1H); 7.49 (d, 1H, *J* = 4); 7.98 (d, 1H, *J* = 2); 8.83 (s, 1H); 10.98 (bs, 1H).

**5-[( $\beta$ -Naphthylmethyl)carbonyl]amino-8-*n*-butyl-2-(2-furyl)-pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidine (**521**).** Yield 88%, pale yellow solid; mp 170 °C (EtOAc–light petroleum). IR (KBr): 3225–2985, 1673, 1611, 1580, 1510 cm<sup>−1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>)  $\delta$ : 0.86 (t, 3H, *J* = 6.7); 1.14–1.31 (m, 2H); 1.79–2.01 (m, 2H); 4.17 (s, 2H); 4.36 (t, 2H, *J* = 6.7); 6.77 (dd, 1H, *J* = 2, *J* = 4); 7.20 (d, 1H, *J* = 4); 7.40–7.59 (m, 3H); 7.81–7.99 (m, 4H); 8.01 (d, 1H, *J* = 4); 8.84 (s, 1H); 11.09 (bs, 1H).

**5-[(4-Methoxybenzyl)carbonyl]amino-8-isopentyl-2-(2-furyl)-pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidine (**522**).** Yield 80%, yellow solid; mp 75 °C (EtOAc–light petroleum). IR (KBr): 3225–2978, 1676, 1620, 1589, 1512 cm<sup>−1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>)  $\delta$ : 0.84 (d, 6H, *J* = 7); 1.37–1.53 (m, 1H); 1.62–1.80 (m, 2H); 3.65 (s, 3H); 3.89 (s, 2H); 4.35 (t, 2H, *J* = 6.7); 6.69 (dd, 1H, *J* = 2, *J* = 4); 6.85 (d, 2H, *J* = 9); 7.22 (d, 1H, *J* = 4); 7.26 (d, 2H, *J* = 9); 7.97 (d, 1H, *J* = 2); 8.85 (s, 1H); 11.03 (bs, 1H).

**5-[(2-Thienylmethyl)carbonyl]amino-8-isopentyl-2-(2-furyl)-pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidine (**523**).** Yield 71%, hygroscopic brown solid (EtOAc–light petroleum). IR (KBr): 3225–2975, 1675, 1613, 1590, 1520 cm<sup>−1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>)  $\delta$ : 0.82 (d, 6H, *J* = 7); 1.33–1.52 (m, 1H); 1.67–1.83 (m, 2H); 4.25 (s, 2H); 4.39 (t, 2H, *J* = 6.7); 6.67 (dd, 1H, *J* = 2, *J* = 4); 7.02 (dd, 1H, *J* = 2, *J* = 4); 7.07 (d, 1H, *J* = 2); 7.19 (d, 1H, *J* = 4); 7.43 (d, 1H, *J* = 4); 7.99 (d, 1H, *J* = 2); 8.82 (s, 1H); 11.13 (bs, 1H).

**5-[(3-Thienylmethyl)carbonyl]amino-8-isopentyl-2-(2-furyl)-pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidine (**524**).** Yield 65%, hygroscopic brown solid (EtOAc–light petroleum). IR



(KBr): 3230–2980, 1673, 1615, 1585, 1510  $\text{cm}^{-1}$ .  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$ : 0.85 (d, 6H,  $J = 7$ ); 1.33–1.54 (m, 1H); 1.65–1.83 (m, 2H); 4.05 (s, 2H); 4.37 (t, 2H,  $J = 6.7$ ); 6.68 (dd, 1H,  $J = 2$ ,  $J = 4$ ); 7.05 (d, 1H,  $J = 4$ ); 7.21 (d, 1H,  $J = 4$ ); 7.39 (s, 1H); 7.42 (d, 1H,  $J = 4$ ); 7.99 (d, 1H,  $J = 2$ ); 8.81 (s, 1H); 11.12 (bs, 1H).

5-[(4-Chlorophenoxymethyl)carbonyl]amino-8-isopentyl-2-(2-furyl)-pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidine (**525**). Yield 88%, pale yellow solid; mp 220 °C (EtOAc–light petroleum). IR (KBr): 3225–2977, 1682, 1618, 1578, 1510  $\text{cm}^{-1}$ .  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$ : 0.82 (d, 6H,  $J = 7$ ); 1.31–1.53 (m, 1H); 1.67–1.80 (m, 2H); 4.38 (t, 2H,  $J = 6.7$ ); 5.11 (s, 2H); 6.67 (dd, 1H,  $J = 2$ ,  $J = 4$ ); 7.02 (d, 2H,  $J = 9$ ); 7.21 (d, 1H,  $J = 4$ ); 7.33 (d, 2H,  $J = 9$ ); 7.99 (d, 1H,  $J = 2$ ); 8.81 (s, 1H); 11.11 (bs, 1H).

5-[(4-Chlorophenoxymethyl)carbonyl]amino-8-(2-phenylethyl)-2-(2-furyl)-pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidine (**526**). Yield 88%, pale yellow solid; mp 232 °C (EtOAc–light petroleum). IR (KBr): 3220–2977, 1676, 1610, 1575, 1510  $\text{cm}^{-1}$ .  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$ : 3.43 (t, 2H,  $J = 7$ ); 4.02 (s, 2H); 4.63 (t, 2H,  $J = 7$  Hz); 6.64 (dd, 1H,  $J = 2$ ,  $J = 4$ ); 7.02–7.49 (m, 10H); 7.94 (d, 1H,  $J = 2$ ); 8.67 (s, 1H); 11.04 (bs, 1H).

5-[(4-Methoxybenzyl)carbonyl]amino-8-(3-phenylpropyl)-2-(2-furyl)-pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidine (**527**). Yield 80%, pale yellow solid; mp 78 °C (EtOAc–light petroleum). IR (KBr): 3230–2995, 1683, 1615, 1580, 1510  $\text{cm}^{-1}$ .  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$ : 2.23 (t, 2H,  $J = 7$ ); 2.52–2.65 (m, 2H); 3.27 (s, 3H); 3.90 (s, 2H); 4.40 (t, 2H,  $J = 7$ ); 6.71 (dd, 1H,  $J = 2$ ,  $J = 4$ ); 6.87 (d, 2H,  $J = 9$ ); 7.05–7.21 (m, 8H); 7.94 (d, 1H,  $J = 2$ ); 8.85 (s, 1H); 10.96 (bs, 1H).

**General Procedures for the Preparation of 5-Amino-8-[[2-(4-aryl)acetylamino-phenyl]ethyl]-2-(2-furyl)-pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidines (**317**–**318**, **515**).** A solution of the amino derivative (**536**) (227 mg, 0.58 mmol), the appropriate arylacetic acid (3.47 mmol), EDC (677 mg, 3.47 mmol), DMAP (32.5 mg, 0.26 mmol), and triethylamine (0.87 mL, 6.3 mmol) in 18 mL of dry DMF was stirred at room temperature for 24 h. The solution was then concentrated at reduced pressure, and water (10 mL) was added. The solution was extracted with EtOAc (15 mL  $\times$  3), and the recombined organic layers were dried and concentrated. The residue was then purified by flash chromatography (EtOAc–light petroleum 1:1) to afford the desired product (**317**–**318**, **515**) as a solid.

5-Amino-8-[[2-(4-methoxyphenyl)acetylamino-phenyl]ethyl]-2-(2-furyl)-pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidine (**317**). Yield 85%, white solid; mp 170 °C (EtOAc–light petroleum). IR (KBr): 3260–2995, 1680, 1615, 1550, 1505  $\text{cm}^{-1}$ .  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$ : 3.20 (t, 2H,  $J = 7$ ); 3.64 (s, 2H); 3.78 (s, 3H); 4.44 (t, 2H,  $J = 7$ ); 4.98 (s, 2H); 6.12 (bs, 2H); 6.58 (dd, 1H,  $J = 2$ ,  $J = 4$ ); 6.87 (d, 2H,  $J = 9$ ); 6.92 (d, 2H,  $J = 9$ ); 7.09 (d, 1H,  $J = 4$ ); 7.17 (d, 2H,  $J = 9$ ); 7.26 (d, 2H,  $J = 9$ ); 7.30 (bs, 1H); 7.60 (d, 1H,  $J = 2$ ); 7.76 (s, 1H).

5-Amino-8-[[2-(4-chlorophenoxy)acetylamino-phenyl]ethyl]-2-(2-furyl)-pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidine (**318**). Yield 83%, white solid; mp 185 °C (EtOAc–light petroleum). IR (KBr): 3255–2990, 1678, 1610, 1555, 1515  $\text{cm}^{-1}$ .  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$ : 3.26 (t, 2H,  $J = 7$ ); 4.49 (t, 2H,  $J = 7$ ); 4.54 (s, 2H); 5.94 (bs, 2H); 6.57 (dd, 1H,  $J = 2$ ,  $J = 4$ ); 6.89 (d, 2H,  $J = 9$ ); 7.01 (d, 2H,  $J = 9$ ); 7.17 (d, 1H,  $J =$

4); 7.26 (d, 2H,  $J = 9$ ); 7.43 (d, 2H,  $J = 9$ ); 7.60 (d, 1H,  $J = 2$ ); 7.78 (s, 1H); 8.13 (bs, 1H).

5-Amino-8-[[2-(4- $\alpha$ -naphthyl)acetylamino-phenyl]ethyl]-2-(2-furyl)-pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidine (**515**). Yield 86%, white solid; mp 189 °C (EtOAc–light petroleum). IR (KBr): 3250–2993, 1681, 1613, 1552, 1511  $\text{cm}^{-1}$ .  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$ : 3.17 (t, 2H,  $J = 7$ ); 4.14 (s, 2H); 4.41 (t, 2H,  $J = 7$ ); 6.00 (bs, 2H); 6.58 (dd, 1H,  $J = 2$ ,  $J = 4$ ); 6.88 (d, 2H,  $J = 9$ ); 6.99 (d, 1H,  $J = 4$ ); 7.13–7.52 (m, 7H); 7.60 (d, 1H,  $J = 2$ ); 7.74 (s, 1H); 7.80–8.00 (m, 3H).

**Biology. Binding at Human  $A_1$ ,  $A_{2A}$ , and  $A_3$  Adenosine Receptors.** All pharmacological methods followed the procedures as described earlier.<sup>60</sup> In brief, membranes for radioligand binding were prepared from CHO cells stably transfected with human adenosine receptor subtypes in a two-step procedure. In a first low-speed step (1000g), cell fragments and nuclei were removed. The crude membrane fraction was sedimented from the supernatant at 100 000g. The membrane pellet was resuspended in the buffer used for the respective binding experiments, frozen in liquid nitrogen, and stored at –80 °C. For the measurement of adenylyl cyclase activity, only one high speed centrifugation of the homogenate was used. The resulting crude membrane pellet was resuspended in 50 mM Tris/HCl, pH 7.4, and immediately used for the cyclase assay.

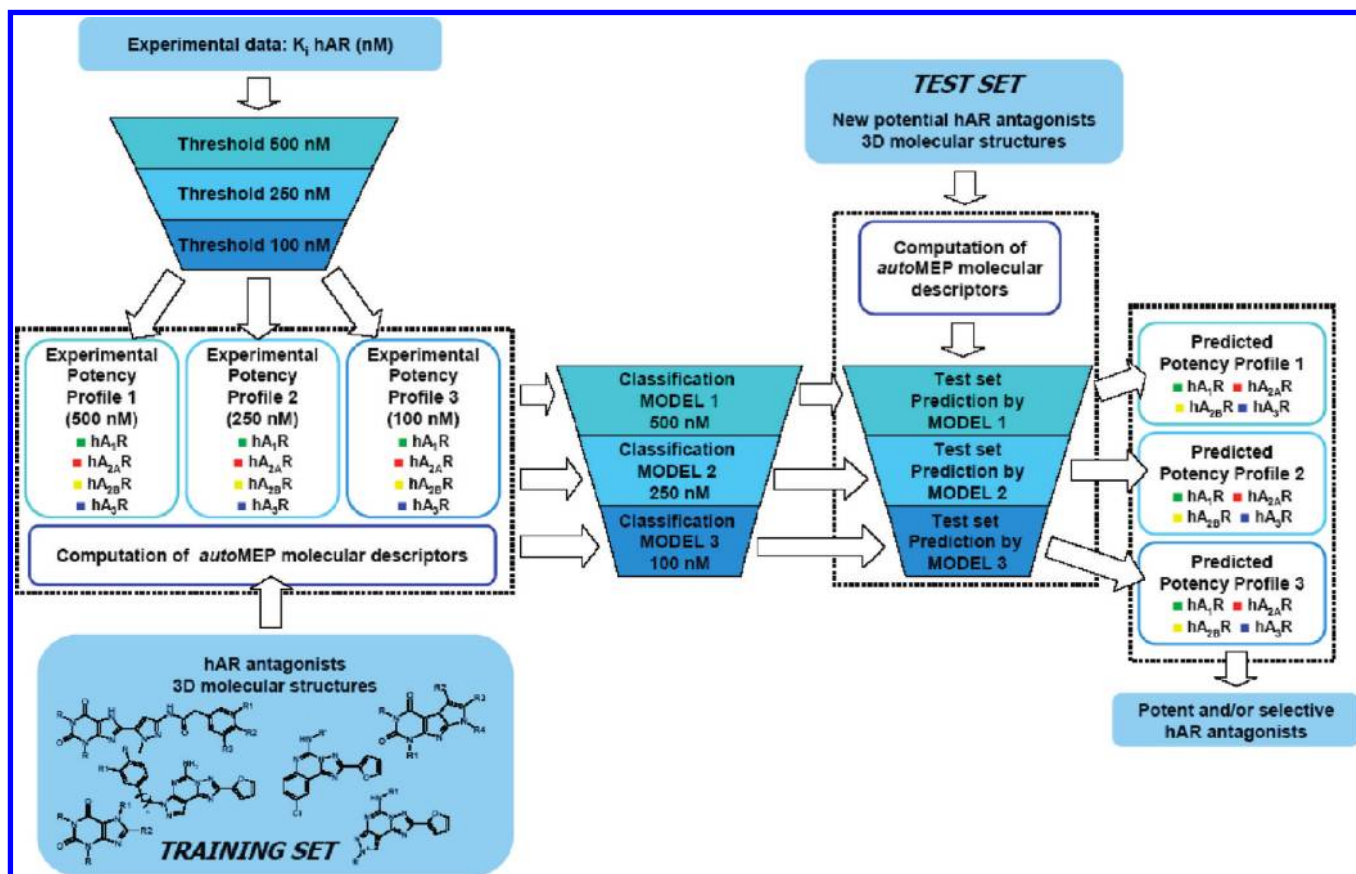
For radioligand binding at  $A_1$  adenosine receptors 1 nM [ $^3\text{H}$ ]CCPA was used, whereas 30 and 10 nM [ $^3\text{H}$ ]NECA were used for  $A_{2A}$  and  $A_3$  receptors, respectively. Nonspecific binding of [ $^3\text{H}$ ]CCPA was determined in the presence of 1 mM theophylline; in the case of [ $^3\text{H}$ ]NECA, 100  $\mu\text{M}$  R-PIA was used.<sup>61</sup>

**Adenylyl Cyclase Activity.** Because of the lack of a suitable radioligand, the potency of antagonists was determined in adenylyl cyclase experiments. The procedure was carried out as described previously with minor modifications.<sup>60,62</sup> Membranes were incubated with about 150 000 cpm of [ $\alpha$ - $^{32}\text{P}$ ]ATP for 20 min in the incubation mixture as described without EGTA and NaCl.<sup>60,62</sup> For agonists, the  $\text{EC}_{50}$ -values for the stimulation of adenylyl cyclase were calculated with the Hill equation. Hill coefficients in all experiments were near unity.  $\text{IC}_{50}$  values for concentration-dependent inhibition of NECA-stimulated adenylyl cyclase caused by antagonists were calculated accordingly. Dissociation constants ( $K_i$ ) for antagonist were then calculated with the Cheng and Prusoff equation.<sup>63</sup>

## RESULT AND DISCUSSION

**General Strategy.** In the last years, the support vector machine method has shown good generalization performance and high accuracy as a supervised learning technique in many classification tasks, such as active compounds selection.<sup>31,32,64</sup> Unfortunately, the traditional classification models do not give any quantitative information about the biological affinity of the compounds. Moreover, in chemoinformatics the classical single-label classification considers mutually exclusive classes, while in some cases compounds are simultaneously labeled in multiple classes. In the present work, hAR antagonists may present multiple molecular properties of multiple hAR subtypes. Because our data set deals with nonmutually exclusive and overlapping classes, the alternative novel ct-SVM multilabel classification provides an appropriate approach, as recently reported.<sup>37,38</sup>





**Figure 4.** Flowchart of the *autoMEP*/ct-SVM filtering approach for the selection of new potent and selective human A<sub>1</sub>R, A<sub>2A</sub>R, A<sub>2B</sub>R, or A<sub>3</sub>R antagonists.

**Table 4.** Data Distribution in the *autoMEP*/ct-SVM MODEL 1 (Threshold = 500 nM)<sup>a</sup>

AR subtype	training set			validation set			internal test set		
	single-label	multilabel	no-label	single-label	multilabel	no-label	single-label	multilabel	no-label
hA <sub>1</sub> R	4	122		1	20		1	53	
hA <sub>2A</sub> R	28	126		6	25		12	55	
hA <sub>2B</sub> R	56	104		15	17		23	29	
hA <sub>3</sub> R	23	129		5	26		11	53	
total (514)	111	175	32	27	32	6	47	68	16

<sup>a</sup> Single-label antagonists occur only once in each class; multilabel antagonists belong to more than one class. Consequently, the sum of multilabel antagonists for all classes within a column is higher than the number of multilabel antagonists for the training set, the validation set, or the *internal* test set. No-label antagonists belong to none of the classes.

**Table 5.** *autoMEP*/ct-SVM MODEL 1: Performance Measures after the Prediction on the Validation Set and the *Internal* Test Set Are Summarized

model prediction	accuracy <sub>ML(α=1)</sub>	one-error <sup>a</sup>	coverage <sup>a</sup>	average precision <sup>a</sup>
validation set	0.75	0.10	2.22	0.92
test set	0.86	0.06	2.19	0.96

<sup>a</sup> Parameters have been calculated after discarding compounds that do not belong to any class in the validation and test set.

Over the past few years, we have claimed that both topological and electrostatic complementarities are two key concepts in the molecular recognition processes. We consider this concept extremely crucial in describing the receptor subtypes selectivity. Gasteiger and collaborators investigated the MEP on the molecular surface as a particularly useful method for rationalizing the interactions between molecules

and molecular recognition processes.<sup>48,53,54</sup> Here, the distribution of the MEP on the molecular surface represents an important determinant for the receptor subtypes selectivity. Furthermore, the introduction of the autocorrelation vector allows then for overcoming the MEP information inconvenience to be reliant on the spatial rotation and translation of the molecule. We have already demonstrated that the *autoMEP* vectors can be used as interesting molecular descriptors in different 3D-QSAR applications.<sup>26–30</sup>

Very recently, we have also reported that the pyrazolo-triazolo-pyrimidine scaffold can be properly modified to obtain hA<sub>2A</sub>R or hA<sub>3</sub>R selectivity.<sup>7</sup> In particular, pyrazolo-triazolo-pyrimidine derivatives bearing specific substitutions at the N<sup>5</sup> and N<sup>8</sup> positions have been described as highly potent and selective human A<sub>3</sub>R antagonists, while the position N<sup>7</sup> shifts the selectivity profile to the human A<sub>2A</sub>R subtype.<sup>18,20</sup> On the other hand, the xanthine scaffold has

**Table 6.** Data Distribution in the *autoMEP*/ct-SVM MODEL 2 (Threshold = 250 nM)<sup>a</sup>

AR subtype	training set			validation set			internal test set		
	single-label	multilabel	no-label	single-label	multilabel	no-label	single-label	multilabel	no-label
hA <sub>1</sub> R	5	72		1	11		1	31	
hA <sub>2A</sub> R	30	93		7	23		12	46	
hA <sub>2B</sub> R	65	82		17	13		26	24	
hA <sub>3</sub> R	37	95		6	20		16	39	
total (514)	137	131	50	31	27	7	55	53	23

<sup>a</sup> Single-label antagonists occur only once in each class; multilabel antagonists belong to more than one class. Consequently, the sum of multilabel antagonists for all classes within a column is higher than the number of multilabel antagonists for the training set, the validation set, or the *internal* test set. No-label antagonists belong to none of the classes.

**Table 7.** *autoMEP*/ct-SVM MODEL 2: Performance Measures after the Prediction on the Validation Set and the *Internal* Test Set Are Summarized

model prediction	accuracy <sub>ML(α=1)</sub>	one-error <sup>a</sup>	coverage <sup>a</sup>	average precision <sup>a</sup>
validation set	0.70	0.15	2.02	0.91
test set	0.81	0.05	1.99	0.96

<sup>a</sup> Parameters have been calculated after discarding compounds that do not belong to any class in the validation and test set.

been investigated to develop highly potent and selective hA<sub>1</sub>R or A<sub>2B</sub>R antagonists.<sup>19,21</sup> However, the observation of the scaffold and its substitutions is not an unfailing strategy to correctly assign the potency profile covering all hAR subtypes of new human pyrazolo-triazolo-pyrimidine and xanthine AR antagonists.

In the present Article, we have developed a “sieve system” formed by three *in series* ct-SVM multilabel classification models using as input matrix our *autoMEP* vectors. The workflow of the procedure we have applied in this study is shown in Figure 4.

Our *autoMEP*/ct-SVM models have been derived after the selection of different thresholds of binding affinity, corresponding to three diverse *K<sub>i</sub>* values: 500 nM for MODEL 1, 250 nM in MODEL 2, and 100 nM for MODEL 3, as summarized in Figure 4. In this way, the selected thresholds act as meshes of our “sieve system”, able to filter hAR antagonists with increasing potency. The principal aim is to introduce the autocorrelated molecular descriptors encoding for the molecular electrostatic potential (*autoMEP*) of each pyrazolo-triazolo-pyrimidine and xanthine antagonist to best approximate the experimental A<sub>1</sub>R, A<sub>2A</sub>R, hA<sub>2B</sub>R, and A<sub>3</sub>R subtypes potency profile using three independent ct-SVM classification models (*autoMEP*/ct-SVM). In particular, once the statistical robust MODELS 1, 2, and 3 have been generated, they can be utilized *in series* to assign the complete potency profile of hAR antagonists and infer their potential subtype selectivity. Interestingly, following our previous valuable experience in the prediction of receptor subtype selectivity, our models can provide at the same time quantitative information about the binding affinity *K<sub>i</sub>* values to all hAR subtypes. Consequently, starting from the calculation of the corresponding *autoMEP* vectors of new pyrazolo-triazolo-pyrimidine and xanthine analogues, our MODELS 1, 2, and 3 are able to select both potent and selective hAR antagonists.

***autoMEP*/ct-SVM Multilabel Classification Models.** Information encoded by the *autoMEP* vectors has been used as input matrix for three ct-SVM multilabel classification

models. Four binary classifiers, corresponding to the hAR subtypes (hA<sub>1</sub>R, hA<sub>2A</sub>R, hA<sub>2B</sub>R, and hA<sub>3</sub>R), have been independently generated by using a Gaussian radial basis function (RBF) kernel for the *autoMEP*/ct-SVM MODELS 1, 2, and 3. To derive our *autoMEP*/ct-SVM models, we have selected a collection of 318 pyrazolo-triazolo-pyrimidine and xanthine derivatives (1–318), and we have defined them as our training set. Furthermore, we have considered 65 additional pyrazolo-triazolo-pyrimidine and xanthine analogues (319–383) as validation set for each model to find the optimal parameters of the four binary classifiers.

The actual labels for the samples in our *autoMEP*/ct-SVM models are assigned by selecting a different binding affinity *K<sub>i</sub>* value as threshold. In more detail, this binary criteria assign “1”, if the hAR subtype binding affinity *K<sub>i</sub>* value is lower than the selected threshold (500, 250, or 100 nM according to the MODELS 1, 2, or 3, respectively). Conversely, “0” is assigned if the hAR antagonists are less potent than the threshold; that is, the corresponding hAR subtype binding affinity *K<sub>i</sub>* value is higher than 500, 250, or 100 nM in the corresponding MODELS 1, 2, and 3. We have considered as selectivity criterion a difference of at least 2 orders of magnitude between the corresponding hAR subtypes *K<sub>i</sub>* values, with the lower receptor subtype *K<sub>i</sub>* value ≤ 100 nM. In our classification problem, hAR antagonists might have a negative potency profile (all zero values) for all subtypes, if all corresponding binding affinity *K<sub>i</sub>* values are lower than the selected threshold. We have applied the best performing C-criterion to transform the real-valued scores assigned by the classifiers into labels. By considering the classification results in the validation step, we have decided not to select a threshold to judge the closeness of two SVM scores. Thus, in this modified C-criterion, all SVM scores are converted into labels, regardless of their difference, and the value of α has been set to 1.

To derive the final *autoMEP*/ct-SVM multilabel classification models, the corresponding training set and the validation set were merged to train the new classifiers with the ct-SVM optimized parameters. Finally, an *internal* test set has been selected to validate our MODELS 1, 2, and 3, as described in the following paragraph.

**MODEL 1 (Threshold = 500 nM).** The binding affinity *K<sub>i</sub>* value 500 nM has been selected as threshold to label the hAR antagonists. The hAR subtypes distribution for the training, the validation, and the *internal* test data is reported in Table 4.

A preliminary model has been generated and validated by predicting the validation set. In the validation procedure, the best parameters have been selected (hA<sub>1</sub>R classifier, *C* = 4,

**Table 8.** Data Distribution in the *autoMEP/ct-SVM* MODEL 3 (Threshold = 100 nM)<sup>a</sup>

AR subtype	training set			validation set			internal test set		
	single-label	multilabel	no-label	single-label	multilabel	no-label	single-label	multilabel	no-label
hA <sub>1</sub> R	6	38		1	5		2	16	
hA <sub>2A</sub> R	36	54		8	13		12	26	
hA <sub>2B</sub> R	61	46		14	5		26	11	
hA <sub>3</sub> R	56	50		10	10		26	23	
total (514)	159	82	77	33	15	17	66	32	33

<sup>a</sup> Single-label antagonists occur only once in each class; multilabel antagonists belong to more than one class. Consequently, the sum of multilabel antagonists for all classes within a column is higher than the number of multilabel antagonists for the training set, the validation set, or the *internal* test set. No-label antagonists belong to none of the classes.

**Table 9.** *autoMEP/ct-SVM* MODEL 3: Performance Measures after the Prediction on the Validation Set and the *Internal* Test Set Are Summarized

model prediction	accuracy <sub>ML(α=1)</sub>	one-error <sup>a</sup>	coverage <sup>a</sup>	average precision <sup>a</sup>
validation set	0.68	0.17	1.77	0.89
test set	0.78	0.04	1.60	0.96

<sup>a</sup> Parameters have been calculated after discarding compounds that do not belong to any class in the validation and test set.

$\gamma = 2$ ; hA<sub>2A</sub>R classifier,  $C = 4$ ,  $\gamma = 0.5$ ; hA<sub>2B</sub>R classifier,  $C = 16$ ,  $\gamma = 0.5$ ; hA<sub>3</sub>R classifier,  $C = 16$ ,  $\gamma = 0.5$ ). After the training set and the validation set merged in a collection of 383 compounds, a new MODEL 1 has been carried out with the optimized parameters, and it has been applied to predict the *internal* test set. The statistical performances of MODEL 1, after the validation set and the *internal* test set prediction procedures, are summarized in Table 5.

The value of accuracy (0.86) on the test set and the measures of the ranking performance confirm that the MODEL 1 is statistically reliable. We have selected this model as a valuable filter of antagonists with hAR subtypes potency profile lower than 500 nM. MODEL 1 has been used in the final validation step, concerning the prediction of the *external* test set.

**MODEL 2 (Threshold = 250 nM).** Even in this case, we have selected a binding affinity  $K_i$  value (250 nM) as threshold to assign the hAR antagonists to the corresponding subtypes/classes. The relative hAR subtypes distribution for the training set, the validation, and the test set is reported in Table 6.

The first model has been built by using the training set. In the validation procedure, the best parameters have been selected (hA<sub>1</sub>R classifier,  $C = 4$ ,  $\gamma = 2$ ; hA<sub>2A</sub>R classifier,  $C = 8$ ,  $\gamma = 0.5$ ; hA<sub>2B</sub>R classifier,  $C = 4$ ,  $\gamma = 0.5$ ; hA<sub>3</sub>R classifier,  $C = 16$ ,  $\gamma = 1$ ). The same collection of 383 antagonists was utilized to generate the definitive MODEL 2 with the optimized parameters. It has then been applied for the prediction of the *internal* test set. The statistical performances of MODEL 2 in the prediction of the validation set and the *internal* test set are reported in Table 7.

The best value of accuracy (0.70) resulted in the validation step. The accuracy is highly affected by the low number of hA<sub>1</sub>R positive samples in both the training set and the validation set. The prediction performances on the *internal* test set and the measures of the ranking process highlight the statistical robustness of MODEL 2. Thus, it might correctly select hAR antagonists with positive subtypes

potency profile under 250 nM. Finally, MODEL 2 has been used to predict the *external* test set.

**MODEL 3 (Threshold = 100 nM).** In MODEL 3, the  $K_i$  value 100 nM has been used as threshold to evaluate the experimental binding affinity data and assign the corresponding labels to hAR antagonists, as summarized in the data distribution reported in Table 8.

In comparison with the previous data distributions, a higher number of single-label hAR antagonists is present in the training set and validation set (see Tables 4, 6, and 8). A preliminary model has been derived and validated by predicting the validation set to select the best parameters (hA<sub>1</sub>R classifier,  $C = 4$ ,  $\gamma = 0.5$ ; hA<sub>2A</sub>R classifier,  $C = 4$ ,  $\gamma = 0.5$ ; hA<sub>2B</sub>R classifier,  $C = 4$ ,  $\gamma = 0.5$ ; hA<sub>3</sub>R classifier,  $C = 4$ ,  $\gamma = 0.5$ ). After the training set and the validation set merged, the final MODEL 3 has been generated with the optimized parameters to predict the *internal* test set. The modeling performances, after the prediction of the validation set and the *internal* test set, are summarized in Table 9.

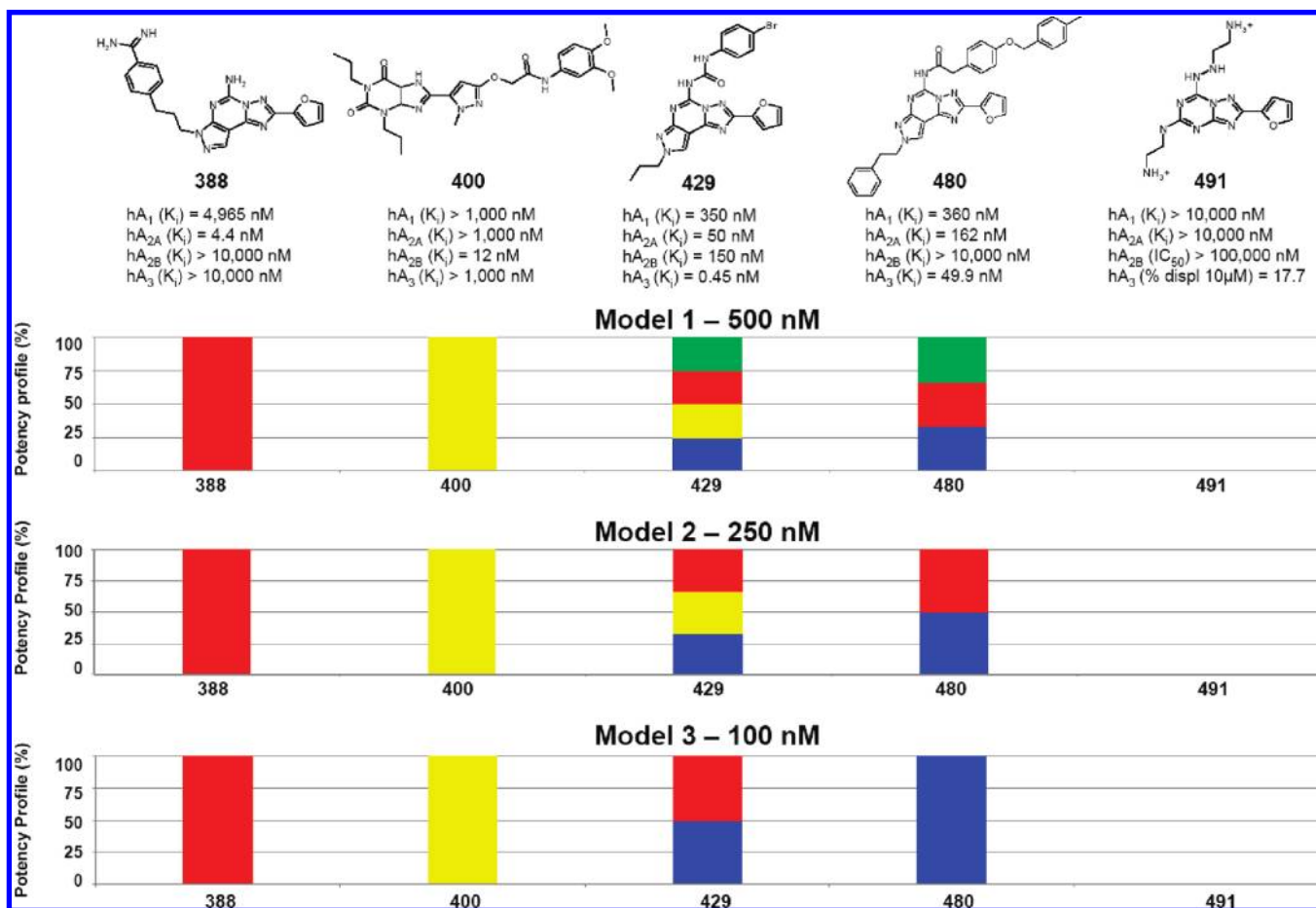
A value of accuracy of 0.78 on the test set and the measures of ranking performances confirms the statistical reliability of MODEL 3. Again, the limited information on hA<sub>1</sub>R antagonists is responsible for the low accuracy in the prediction of the validation set. However, the values for one-error, coverage, and average precision are similar in comparison with the previously reported MODELS 1 and 2. This model can be utilized to classify potential hAR antagonists with  $K_i$  value lower or higher than 100 nM. We have applied MODEL 3 for the prediction of the *external* test set.

**Internal Validation of the *autoMEP/ct-SVM* MODELS 1, 2, and 3.** As anticipated, our *autoMEP/ct-SVM* multilabel classification models have been further validated by predicting an *internal* test set to evaluate their *in series* applicability as “sieve system” for new xanthine and pyrazolo-triazolo-pyrimidine derivatives. In this validation process, each of the hA<sub>1</sub>R, hA<sub>2A</sub>R, hA<sub>2B</sub>R, and hA<sub>3</sub>R antagonists has been analyzed by MODELS 1, 2, and 3.

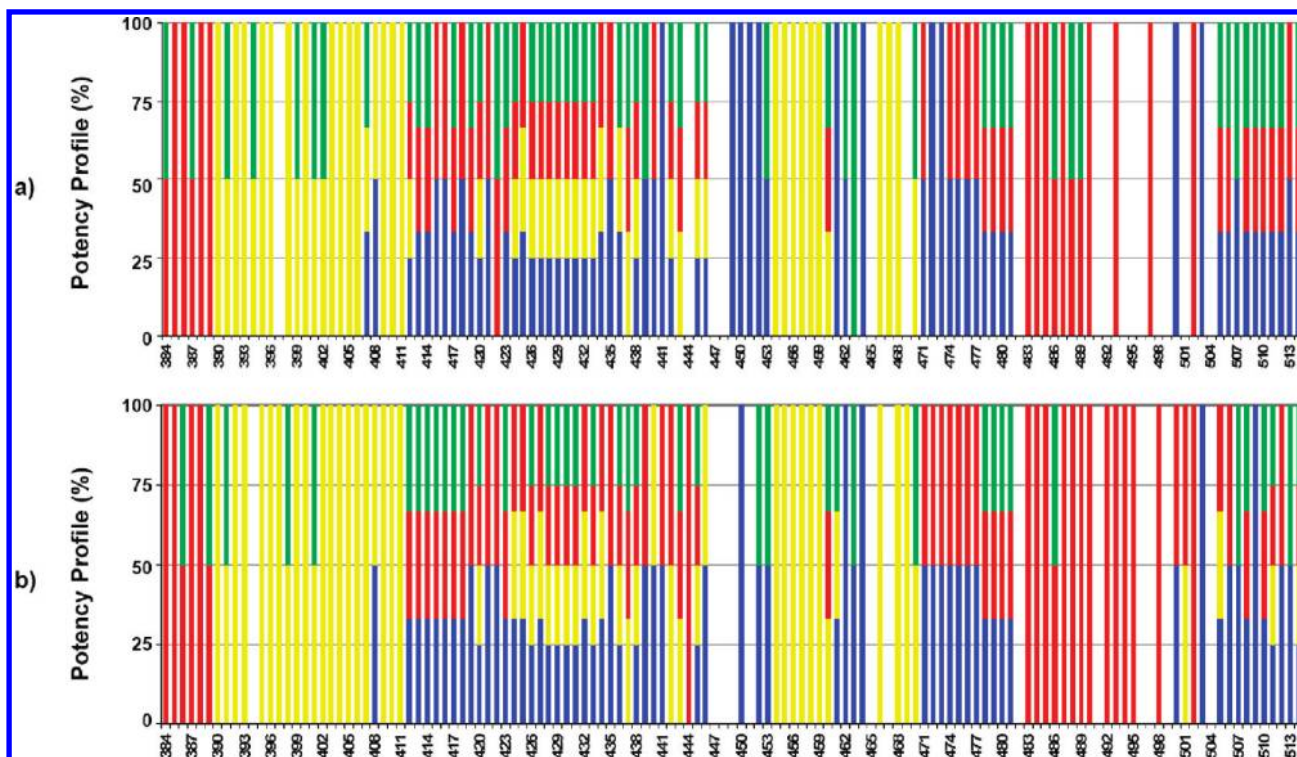
In Figure 5, the potency profiles of five structurally different and correctly predicted hAR antagonists have been considered as examples for the predictions interpretation.

In this potency profile representation, differently colored columns are used for each antagonist to show the predicted subtypes/labels by MODELS 1, 2, and 3. The predicted value “1” by the corresponding hAR binary classifier is represented with a color for each model to indicate the relative percentage (%) of potency profile. Thus, the prediction “1”/“0” refers to the corresponding hAR subtype  $K_i$  value, resulting to be lower/higher than the thresholds 500, 250, or 100 nM for the MODELS 1, 2, or 3, respectively. A predicted relative

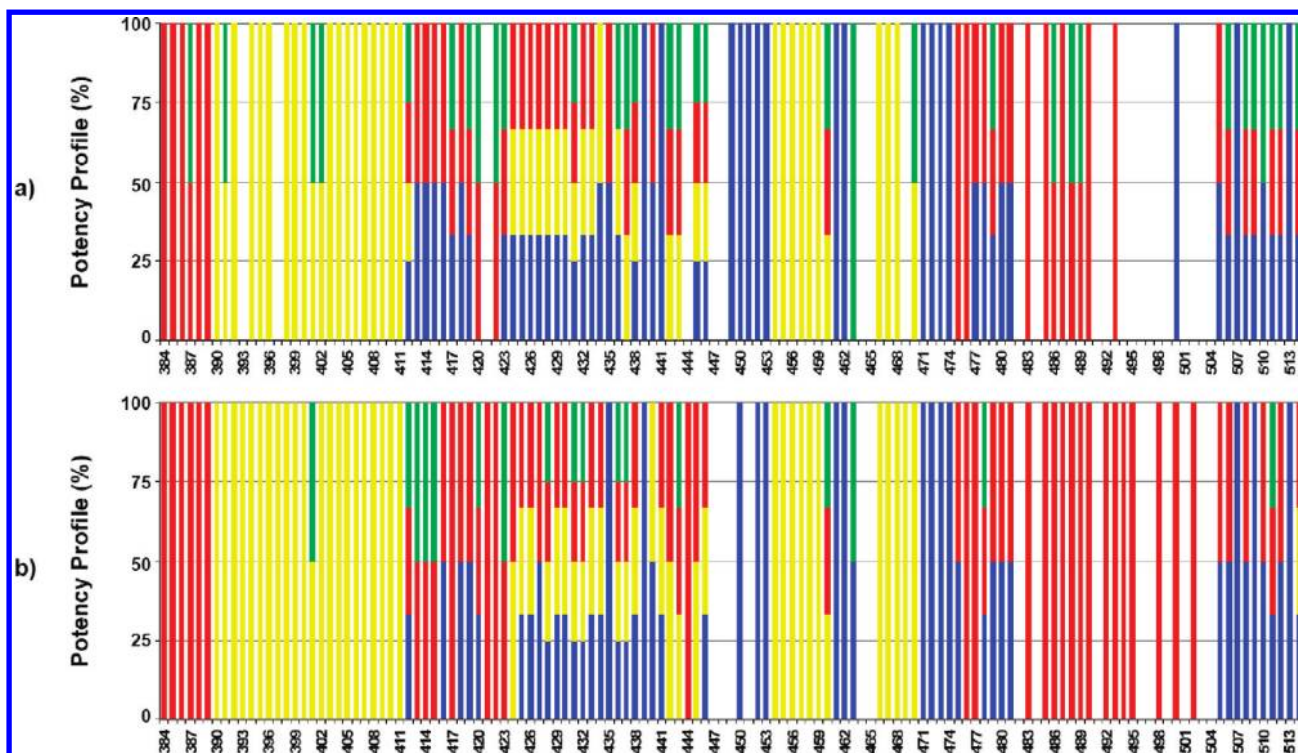




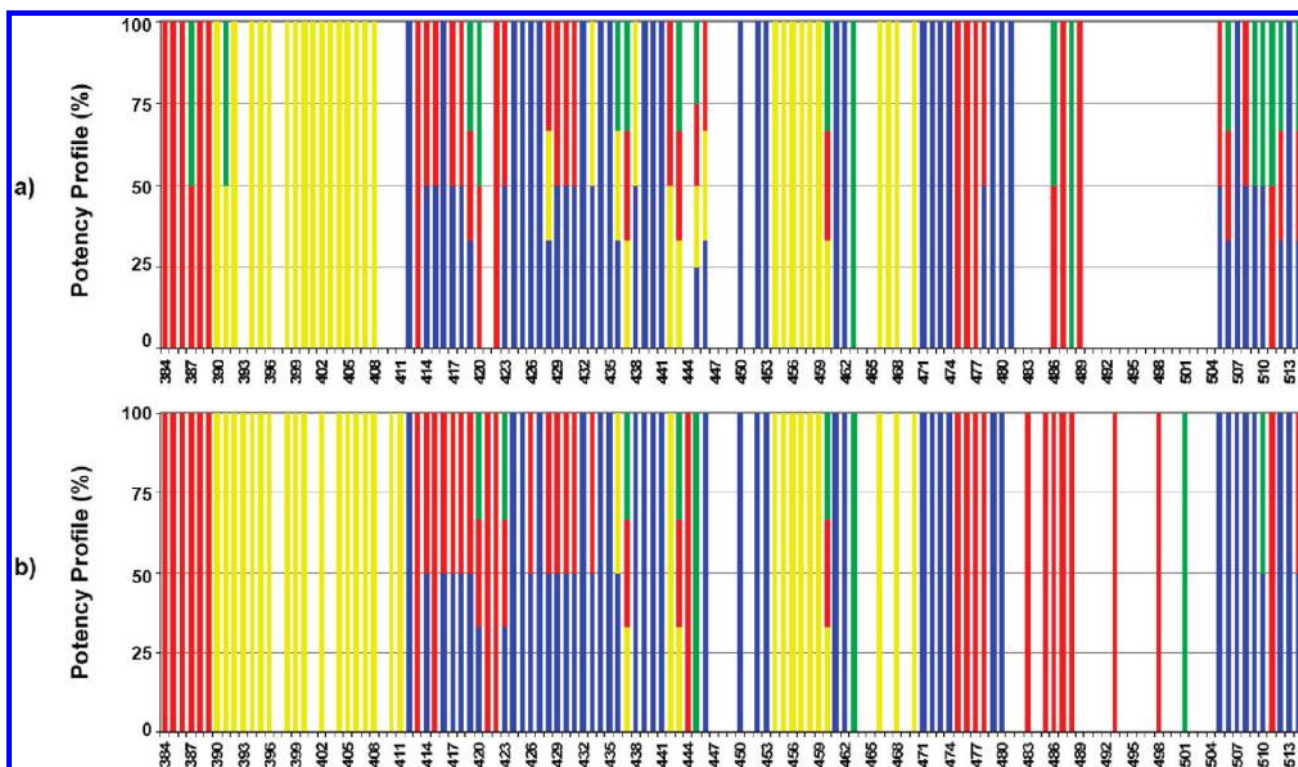
**Figure 5.** Flowchart of the correctly predicted profiles of five hAR antagonists by applying our *autoMEP/ct-SVM* multilabel classification models on the *internal* test set: the colored columns represent the relative percentage (%) of potency profiles corresponding to  $hA_1R$  (green),  $hA_{2A}R$  (red),  $hA_{2B}R$  (yellow), and  $hA_3R$  (blue) subtypes.



**Figure 6.** Graphical representation of the (a) experimental and (b) predicted by MODEL 1 (threshold = 500 nM) potency profiles for the *internal* test set. The colored columns show whether the  $K_i$  values to the relative human adenosine receptor (hAR) subtype are lower than 500 nM. The relative percentages (%) indicate positive potency profile for  $hA_1R$  (green),  $hA_{2A}R$  (red),  $hA_{2B}R$  (yellow), and  $hA_3R$  (blue) subtypes. If colored columns are missing, the  $K_i$  values are higher than 500 nM for all hAR subtypes.



**Figure 7.** Graphical representation of the (a) experimental and (b) predicted by MODEL 2 (threshold = 250 nM) potency profiles for the *internal* test set. The colored columns show whether the  $K_i$  values to the relative human adenosine receptor (hAR) subtype are lower than 250 nM. The relative percentages (%) indicate positive potency profile for hA<sub>1</sub>R (green), hA<sub>2A</sub>R (red), hA<sub>2B</sub>R (yellow), and hA<sub>3</sub>R (blue) subtypes. If colored columns are missing, the  $K_i$  values are higher than 250 nM for all hAR subtypes.



**Figure 8.** Graphical representation of the (a) experimental and (b) predicted by MODEL 3 (threshold = 100 nM) potency profiles for the *internal* test set. The colored columns show whether the  $K_i$  values to the relative human adenosine receptor (hAR) subtype are lower than 100 nM. The relative percentages (%) indicate positive potency profile for hA<sub>1</sub>R (green), hA<sub>2A</sub>R (red), hA<sub>2B</sub>R (yellow), and hA<sub>3</sub>R (blue) subtypes. If colored columns are missing, the  $K_i$  values are higher than 100 nM for all hAR subtypes.

percentage (%) of potency profile lower than 100% for at least one class means that the compound is more potent than the selected threshold  $K_i$  for multiple hAR subtypes. Moreover, the selectivity can be inferred by one-color columns

in all three models. In our first example, a N<sup>7</sup>-substituted pyrazolo-triazolo-pyrimidine analogue (compound **388**) resulted to have 100% hA<sub>2A</sub>R potency profile for all three models, with totally red columns (Figure 5). This corresponds

**Table 10.** *autoMEP/ct-SVM MODEL 1*: Statistical Parameters for Each Class after the Predictions on the *Internal* Test Set Are Reported

classes	TP	FP	TN	FN	recall	precision
hA <sub>1</sub> R	33	9	68	21	0.61	0.79
hA <sub>2A</sub> R	62	12	52	5	0.92	0.84
hA <sub>2B</sub> R	50	8	71	2	0.96	0.86
hA <sub>3</sub> R	60	2	65	4	0.94	0.97

**Table 11.** *autoMEP/ct-SVM MODEL 2*: Statistical Parameters for Each Class after the Predictions on the *Internal* Test Set Are Reported

classes	TP	FP	TN	FN	recall	precision
hA <sub>1</sub> R	11	6	93	21	0.34	0.65
hA <sub>2A</sub> R	55	12	61	3	0.95	0.82
hA <sub>2B</sub> R	48	6	75	2	0.96	0.89
hA <sub>3</sub> R	44	4	72	11	0.80	0.92

**Table 12.** *autoMEP/ct-SVM MODEL 3*: Statistical Parameters for Each Class after the Predictions on the *Internal* Test Set Are Reported

classes	TP	FP	TN	FN	recall	precision
hA <sub>1</sub> R	7	3	110	11	0.39	0.70
hA <sub>2A</sub> R	30	10	83	8	0.79	0.75
hA <sub>2B</sub> R	29	3	91	8	0.78	0.91
hA <sub>3</sub> R	45	1	81	4	0.92	0.98

to a  $K_i$  value lower than 100 nM only for the hA<sub>2A</sub>R subtype and a  $K_i$  value higher than 500 nM for the remaining hAR subtypes. Consequently, the compound **388** is hA<sub>2A</sub>R selective. In a similar case, the compound **400**, a xanthine derivative, is predicted to have a complete hA<sub>2B</sub>R potency profile by MODELS 1, 2, and 3. The corresponding experimental  $K_i$  values confirm hA<sub>2B</sub>R subtype selectivity. By observing the resulting columns, both compounds **429** and **480** show a mixed potency profile. They are N<sup>8</sup>-substituted pyrazolo-triazolo-pyrimidine analogues, and each *autoMEP/ct-SVM* model is able to inform about the experimental  $K_i$  values for all hAR subtypes. However, compound **429** is correctly predicted as potent hA<sub>2A</sub>R and hA<sub>3</sub>R antagonist, without any inference on hA<sub>3</sub>R selectivity. Finally, we have predicted a negative potency profile for all hAR subtypes in the last example (compound **491**), having a different chemical structure. In fact, no colored columns are present in the *autoMEP/ct-SVM* models prediction.

In our analysis, the *internal* test set has been sifted out by three quantitative filters, and we have compared the experimental and predicted relative percentages (%) of potency profile by MODELS 1, 2, and 3 for all hAR antagonists, as 1, 2, and 3 for all hAR antagonists, as illustrated in Figures 6–8, respectively.

As previously shown in Tables 5, 7, and 9, good values of accuracy were obtained for our *autoMEP/ct-SVM* multilabel classification models in the prediction of the collected *internal* test set (86%, 81%, and 78%, respectively). The satisfying prediction accuracies can be also inferred by the graphical comparison of the experimental (Figures 6a, 7a, and 8a) with the predicted (Figures 6b, 7b, and 8b) potency profiles for each model. In particular, the comparison of the experimental with the corresponding predicted graphical representations has shown high similarity in the color

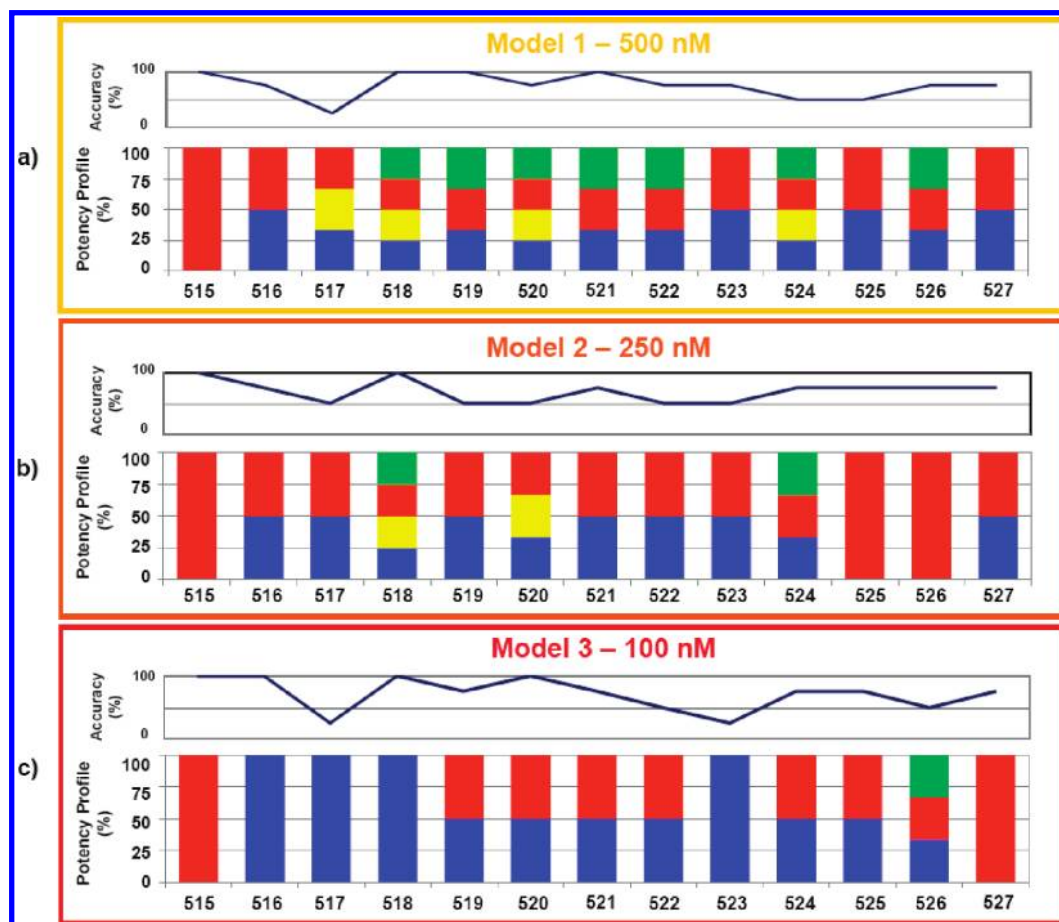
distribution. Moreover, the prediction results are very encouraging, as highlighted by the good values of the statistical base-class parameters recall, precision, and the number of TP and TN for hA<sub>2A</sub>R, hA<sub>2B</sub>R, and hA<sub>3</sub>R subtypes (Tables 10–12).

The analysis of the experimental and predicted hAR potency profiles of the *internal* test set by the above-mentioned models, with the exception of the hA<sub>1</sub>R subtype, again supports the quality of our filtering strategy. In the *internal* test set, 49 selective hAR antagonists have been considered. This subset comprises 1 hA<sub>1</sub>R selective antagonist (molecule **463**), 6 hA<sub>2A</sub>R selective antagonists (molecules **384–386**, **388–389**, **487**), 22 hA<sub>2B</sub>R selective antagonists (molecules **390**, **392**, **394–396**, **398**, **400**, **403–408**, **454–459**, **466–468**), and 20 hA<sub>3</sub>R selective antagonists (molecules **412**, **424–427**, **431–432**, **434**, **440–441**, **450**, **452**, **461**, **471–474**, **507**, **509**, **513**). After considering *in series* our *autoMEP/ct-SVM* models, 26 out of 49 selective hAR antagonists are perfectly predicted (molecules **385**, **388**, **390**, **392**, **395–396**, **400**, **404–406**, **408**, **425**, **431**, **450**, **454–459**, **466**, **468**, **471**, **474**, **487**, **507**), and we are able to infer hAR subtype selectivity of 10 compounds by analyzing their partially correct predictions (molecules **384**, **386**, **389**, **394**, **398**, **403**, **407**, **452**, **461**, and **513**). However, MODEL 3 has correctly assigned the potency profile of the other 10 out of 49 selective hAR antagonists (molecules **412**, **424**, **427**, **432**, **434**, **440–441**, **463**, **472–473**), as reported in Figures 6–8. Our approach is thus able to detect at least potent antagonists for the different hAR subtypes. Concerning the prediction of potent ( $K_i \leq 100$  nM) hAR antagonists, MODEL 3 has detected 7 out of 18 (39%) potent hA<sub>1</sub>R antagonists, 30 out of 38 (79%) potent hA<sub>2A</sub>R antagonists, 29 out of 37 (78%) potent hA<sub>2B</sub>R antagonists, and 45 out of 49 (96%) potent hA<sub>3</sub>R antagonists (see Table 12). Therefore, the independent application of each of our models can select with high accuracy hAR antagonists having binding affinity  $K_i$  value lower than 500, 250, or 100 nM, according to the aim of the filtering process.

**External Validation of the *autoMEP/ct-SVM* MODELS 1, 2, and 3.** In the optimization step of drug discovery, the principal application of the *autoMEP/ct-SVM* models described in the present work is the prediction of the complete hAR binding affinity profile and hAR subtypes selectivity of new potential antagonists. To evaluate the prediction capability of our approach, we have considered an additional test set (*external* test set), consisting of 13 novel N<sup>8</sup>-substituted pyrazolo-triazolo-pyrimidine analogues (**515–527**), which have been synthesized and pharmacologically characterized. The experimental human A<sub>1</sub>R, A<sub>2A</sub>R, A<sub>2B</sub>R, and A<sub>3</sub>R binding affinities are collected in Table 3. As illustrated in Figure 2, the *autoMEP* vectors of these new 13 hAR antagonists have been used as input matrix for the previously generated *autoMEP/ct-SVM* multilabel classification MODELS 1, 2, and 3. In most cases, our three models are able to assign the potency profile with at least the 75% of accuracy for each compound in the collected *external* test set (Figure 9).

The best prediction performance is achieved by MODEL 1, which represents the first sieve (see Figure 9a). By analyzing the statistical class-base parameters, good values of recall have been obtained for hA<sub>2A</sub>R and hA<sub>3</sub>R classifiers in MODEL 1 (1.00 and 1.00, respectively), MODEL 2 (1.00





**Figure 9.** AutoMEP/ct-SVM multilabel classification models: (a) *external* test set prediction by MODEL 1 – 500 nM, graphical representation indicating the percentage (%) of accuracy for each compound and the predicted relative percentages (%) of the potency profiles for hA<sub>1</sub>R (green), hA<sub>2A</sub>R (red), hA<sub>2B</sub>R (yellow), and hA<sub>3</sub>R (blue) subtypes; (b) *external* test set prediction by MODEL 2 – 250 nM, graphical representation indicating the percentage (%) of accuracy for each compound and the predicted relative percentages (%) of the potency profiles for hA<sub>1</sub>R (green), hA<sub>2A</sub>R (red), hA<sub>2B</sub>R (yellow), and hA<sub>3</sub>R (blue) subtypes; (c) *external* test set prediction by MODEL 3 – 100 nM, graphical representation indicating the percentage (%) of accuracy for each compound and the predicted relative percentages (%) of the potency profiles for hA<sub>1</sub>R (green), hA<sub>2A</sub>R (red), hA<sub>2B</sub>R (yellow), and hA<sub>3</sub>R (blue) subtypes.

and 1.00, respectively), and MODEL 3 (0.82 and 1.00, respectively). Moreover, hA<sub>2A</sub>R antagonists have been predicted also with high precision (0.92, 0.92, and 1.00 for MODELS 1, 2, and 3, respectively) (see the Supporting Information for details). It is noteworthy how most of the selective compounds in the test set are correctly classified. In more detail, compounds **515** and **518**, hA<sub>2A</sub>R and hA<sub>3</sub>R selective antagonists, respectively, are perfectly predicted, while compound **516** is erroneously recognized also as hA<sub>2A</sub>R antagonist by MODELS 1 and 2 (Table 3 and Figure 9).

These 13 derivatives are the result of various substitutions and homologies in two different positions to improve selectivity of the pyrazolo-triazolo-pyrimidine scaffold to hA<sub>2A</sub>R or hA<sub>3</sub>R subtypes. Considering their similar structure, our autoMEP/ct-SVM models are able at least to select almost all potent hA<sub>2A</sub>R antagonists (molecules **519–527**), without missing out any potent hA<sub>3</sub>R antagonist (molecule **520**) in the filtering procedure (see Figure 9c). Some exceptions, regarding the quality of potency profile predictions, are represented by hA<sub>1</sub>R predicted labels for compounds **517**, **522–524**, and **527**. Surprisingly, MODEL 3 has not misattributed any compound to hA<sub>2B</sub>R class. However, any additional information on hA<sub>1</sub>R selective antagonists in the training set might improve the predictivity for

the hA<sub>1</sub>R subtype (see Figure 9c). Concluding, the comparison of all experimental with the predicted hAR subtypes potency profiles by MODELS 1, 2, and 3 on the *external* test set has demonstrated the appreciable prediction performance of our autoMEP/ct-SVM “sieve system”.

## CONCLUSIONS AND PERSPECTIVES

In the present Article, a novel multilabel classification approach combining autoMEP molecular descriptors with support vector machine (autoMEP/ct-SVM) has been presented as a powerful tool for the prediction of hA<sub>1</sub>R, hA<sub>2A</sub>R, hA<sub>2B</sub>R, and hA<sub>3</sub>R subtypes potency profile and infer the potential selectivity of known xanthine and pyrazolo-triazolo-pyrimidine derivatives. Three statistically meaningful models have been generated from the same training set by using different binding affinity  $K_i$  values as thresholds for hAR classifiers, and, surprisingly, very positive results were achieved in the validation procedure. The application of our models *in series* has been able to filter and to assign with appreciable accuracy the hAR potency profile of 13 new pyrazolo-triazolo-pyrimidine analogues. In the search for selective hAR antagonists, we are continuously developing and analyzing new compounds. To further improve the predictivity of our dynamic autoMEP/ct-SVM strategy, we

aim at integrating new information on hAR antagonists in our data set, especially regarding the hA<sub>1</sub>R subtype.

**Abbreviations.** GPCRs, G protein-coupled receptors; AR, adenosine receptor; cAMP, cyclic adenosine monophosphate; h, human; SAR, structure–affinity relationship; r, rat; DPCPX, 8-cyclopentyl-1,3-dipropyl-xanthine; CGS 21680, 2-[4-(2-carboxyethyl)phenethyl]amino-5'-(N-ethylcarbamoyl)adenosine; I-AB-MECA, N<sup>6</sup>-(4-amino-3-iodobenzyl)-5'-(N-methylcarbamoyl)adenosine; NECA, 5'-(N-ethyl-carboxamido)adenosine; Cl-IB-MECA, 2-chloro-N<sup>6</sup>-(3-iodobenzyl)-5'-(N-methylcarbamoyl)adenosine; SAR, structure–affinity relationship; MEP, molecular electrostatic potential; SVM, support vector machine; molecular operating environment.

## ACKNOWLEDGMENT

The molecular modeling work coordinated by S.M. was carried out with financial support from the University of Padova, Italy, and the Italian Ministry for University and Research (MIUR), Rome, Italy. We thank the Molecular Networks GmbH (Erlangen, Germany; <http://www.molecular-networks.com>) for the assistance in using the ct-SVM classification method and Adriana modeling suite. S.M. is also very grateful to the Chemical Computing Group for the scientific and technical partnership.

**Supporting Information Available:** Training set (Table 1), validation set (Table 2), and internal test set (Table 3), Table 4 of elemental analyses of synthesized compounds, together with Table 5, Table 6, Table 7 reporting additional information on the validation process, and Table 8 of predictions on the external test set. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## REFERENCES AND NOTES

- Jacobson, K. A.; Gao, Z. G. Adenosine receptor as therapeutic targets. *Nat. Rev. Drug Discovery* **2006**, *5*, 247–264.
- Moro, S.; Gao, Z. G.; Jacobson, K. A.; Spalluto, G. Progress in pursuit of therapeutic adenosine receptor antagonists. *Med. Res. Rev.* **2006**, *26*, 131–159.
- Kadam, R. U.; Chavan, A. G.; Monga, V.; Kaur, N.; Jain, R.; Roy, N. Selectivity-based QSAR approach for screening and evaluation of TRH analogs for TRH-R1 and TRH-R2 receptors subtypes. *J. Mol. Graphics Modell.* **2008**, *27*, 309–320.
- Giorgi, I.; Leonardi, M.; Pietra, D.; Biagi, G.; Borghini, A.; Massarelli, I.; Ciampi, O.; Bianucci, A. M. Synthesis, biological assays and QSAR studies of N-(9-benzyl-2-phenyl-8-azapurin-6-yl)-amides as ligands for A<sub>1</sub> adenosine receptors. *Bioorg. Med. Chem.* **2009**, *17*, 1817–1830.
- Moro, S.; Bacilieri, M.; Deflorian, F.; Spalluto, G. G protein-coupled receptors as challenging druggable targets: insights from in silico studies. *New J. Chem.* **2006**, *30*, 301–308.
- Moro, S.; Deflorian, F.; Bacilieri, M.; Spalluto, G. Ligand-based homology modeling as attractive tool to inspect GPCR structural plasticity. *Curr. Pharm. Des.* **2006**, *12*, 2175–2185.
- Michielan, L.; Bolcato, C.; Stephanie, F.; Cacciari, B.; Bacilieri, M.; Klotz, K. N.; Kachler, S.; Pastorin, G.; Cardin, R.; Sperduti, A.; Moro, S. Combining selectivity and affinity predictions using an integrated Support Vector Machine (SVM) approach: an alternative tool to discriminate between the human adenosine A<sub>2A</sub> and A<sub>3</sub> receptor pyrazolo-triazolo-pyrimidine antagonists binding sites. *Bioorg. Med. Chem.* **2009**, *17*, 5259–5274.
- Fredholm, B. B.; Arslan, G.; Halldner, L.; Kull, B.; Schulte, G.; Wasserman, W. Structure and function of adenosine receptors and their genes. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **2000**, *362*, 364–374.
- Maemoto, T.; Tada, M.; Mihara, T.; Ueyama, N.; Matsuoka, H.; Harada, K.; Yamaji, T.; Shirakawa, K.; Kuroda, S.; Akahane, A.; Iwashita, A.; Matsuoka, N.; Mutoh, S. Pharmacological characterization of FR194921, a new potent, selective, and orally active antagonist for central adenosine A<sub>1</sub> receptors. *J. Pharmacol. Sci.* **2004**, *96*, 42–52.
- Ferre, S.; von Euler, G.; Johansson, B.; Fredholm, B. B.; Fuxe, K. Stimulation of high-affinity adenosine A<sub>2</sub> receptors decreases the affinity of dopamine D2 receptors in rat striatal membranes. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 7238–7241.
- Holgate, S. The identification of the adenosine A<sub>2B</sub> receptor as a novel therapeutic target in asthma. *Br. J. Pharmacol.* **2005**, *145*, 1009–1015.
- Xu, K.; Bastia, E.; Schwarzschild, M. Therapeutic potential of adenosine A<sub>2A</sub> receptor antagonists in Parkinson's disease. *Pharmacol. Ther.* **2005**, *105*, 267–310.
- Johnston, T. H.; Brochie, J. M. Drugs in development for Parkinson's disease: an update. *Curr. Opin. Invest. Drugs* **2006**, *7*, 25–32.
- Ribeiro, J. A.; Sebastiao, A. M.; de Mendonca, A. Adenosine receptors in the nervous system: pathophysiological implications. *Prog. Neurobiol.* **2002**, *68*, 377–392.
- Muller, C. E. Medicinal chemistry of adenosine A<sub>3</sub> receptor ligands. *Curr. Top. Med. Chem.* **2003**, *3*, 445–462.
- Jacobson, K. A. A<sub>3</sub> adenosine receptors. *Annual Reports in Medicinal Chemistry*; Elsevier: San Diego, CA, 2003.
- Okamura, T.; Kurogi, Y.; Hashimoto, K.; Sato, S.; Nishikawa, H.; Kiryu, K.; Nagao, Y. Structure–activity relationships of adenosine A<sub>3</sub> receptor ligands: new potential therapy for the treatment of glaucoma. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3775–3779.
- Baraldi, P. G.; Tabrizi, M. A.; Bovero, A.; Avitabile, B.; Preti, D.; Fruttarolo, F.; Romagnoli, R.; Varani, K.; Borea, P. A. Recent developments in the field of A<sub>2A</sub> and A<sub>3</sub> adenosine receptor antagonists. *Eur. J. Med. Chem.* **2003**, *38*, 367–382.
- Tabrizi, M. A.; Baraldi, P. G.; Preti, D.; Romagnoli, R.; Saponaro, G.; Baraldi, S.; Moorman, A. R.; Zaid, A. N.; Varani, K.; Borea, P. A. 1,3-Dipropyl-8-(1-phenylacetamide-1H-pyrazol-3-yl)-xanthine derivatives as highly potent and selective human A<sub>2B</sub> adenosine receptor antagonists. *Bioorg. Med. Chem.* **2008**, *16*, 2419–2430.
- Baraldi, P. G.; Cacciari, B.; Moro, S.; Spalluto, G.; Pastorin, G.; Da Ros, T.; Klotz, K. N.; Varani, K.; Gessi, S.; Borea, P. A. Synthesis, biological activity, and molecular modeling investigation of new pyrazolo[4,3-e]-1,2,4-triazolo[1,5-c]pyrimidine derivatives as human A<sub>3</sub> adenosine receptor antagonists. *J. Med. Chem.* **2002**, *45*, 770–780.
- Weyler, S.; Füller, F.; Diekmann, M.; Schumacher, B.; Hinz, S.; Klotz, K. N.; Müller, C. E. Improving potency, selectivity, and water solubility of adenosine A<sub>1</sub> receptor antagonists: xanthine modified at position 3 and related pyrimido[1,2,3-cd]purinediones. *Chem. Med. Chem.* **2006**, *1*, 891–902.
- Elzein, E.; Kalla, R.; Li, X.; Perry, T.; Parkhill, E.; Palle, V.; Varkhedkar, V.; Gimbel, A.; Zeng, D.; Lustig, D.; Leung, K.; Zablocki, J. Novel 1,3-dipropyl-8-(1-heteroaryl-methyl-1H-pyrazol-4-yl)-xanthine derivatives as high affinity and selective A<sub>2B</sub> adenosine receptor antagonists. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 302–306.
- Kalla, R. V.; Elzein, E.; Perry, T.; Li, X.; Gimbel, A.; Yang, M.; Zeng, D.; Zablocki, J. Selective, high affinity A<sub>2B</sub> adenosine receptor antagonists: N-1 monosubstituted 8-(pyrazol-4-yl)xanthines. *Bioorg. Med. Chem.* **2008**, *18*, 1397–1401.
- Baraldi, P. G.; Tabrizi, M. A.; Preti, D.; Bovero, A.; Romagnoli, R.; Fruttarolo, F.; Zaid, N. A.; Moorman, A. R.; Varani, K.; Gessi, S.; Merighi, S.; Borea, P. A. Design, synthesis, and biological evaluation of new 8-heterocyclic xanthine derivatives as highly potent and selective human A<sub>2B</sub> adenosine receptor antagonists. *J. Med. Chem.* **2004**, *47*, 1434–1477.
- Baraldi, P. G.; Preti, D.; Tabrizi, M. A.; Romagnoli, R.; Saponaro, G.; Baraldi, S.; Botta, M.; Bernardini, C.; Tafi, A.; Tuccinardi, T.; Martinelli, A.; Varani, K.; Borea, P. A. Structure–activity relationship studies of a new series of imidazo[2,1-f]purinones as potent and selective A<sub>3</sub> adenosine receptor antagonists. *Bioorg. Med. Chem.* **2008**, *16*, 10281–10294.
- Moro, S.; Bacilieri, M.; Ferrari, C.; Spalluto, G. Autocorrelation of molecular electrostatic potential surface properties combined with partial least squares analysis as alternative attractive tool to generate ligand-based 3D-QSARs. *Curr. Drug Discovery Technol.* **2005**, *2*, 13–21.
- Moro, S.; Bacilieri, M.; Cacciari, B.; Spalluto, G. Autocorrelation of molecular electrostatic potential surface properties combined with partial least squares analysis as new strategy for the prediction of the activity of human A<sub>3</sub> adenosine receptor antagonists. *J. Med. Chem.* **2005**, *48*, 5698–5704.
- Moro, S.; Bacilieri, M.; Cacciari, B.; Bolcato, C.; Cusan, C.; Pastorin, G.; Klotz, K. N.; Spalluto, G. The application of a 3D-QSAR (autoMEP/PLS) approach as an efficient pharmacodynamic-driven filtering method for small-sized virtual library: application to a lead optimization of a human A<sub>3</sub> adenosine receptor antagonist. *Bioorg. Med. Chem.* **2006**, *14*, 4923–4932.
- Bacilieri, M.; Kaseda, C.; Spalluto, G.; Moro, S. Response surface analysis as alternative 3D-QSAR tool: human A<sub>3</sub> adenosine receptor

- antagonists as a key study. *Lett. Drug Des. Discovery* **2007**, 4, 122–127.
- (30) Michielan, L.; Bacilieri, M.; Schiesaro, A.; Bolcato, C.; Pastorin, G.; Spalluto, G.; Cacciari, B.; Klotz, K. N.; Kaseda, C.; Moro, S. Linear and non-linear 3D-QSAR approaches in tandem with ligand-based homology modeling as computational strategy to depict the pyrazolo-triazolo-pyrimidine antagonists binding site of the human adenosine A<sub>2A</sub> receptor. *J. Comput. Inf. Model.* **2008**, 48, 350–363.
- (31) Burges, C. J. C. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Disc.* **1998**, 2, 121–167.
- (32) Cristianini, N.; Shawe-Taylor, J. *An Introduction to Support Vector Machines*; Cambridge University Press: New York, 2000.
- (33) Czerminski, R.; Yasri, A.; Hartsoogh, D. Use of support vector machine in pattern classification: application to QSAR studies. *Quant. Struct.-Act. Relat.* **2001**, 20, 227–240.
- (34) Warmuth, M. K.; Liao, J.; Ratsch, G.; Mathieson, M.; Putta, S.; Lemmen, C. Active learning with support vector machine in the drug discovery process. *J. Chem. Inf. Comput. Sci.* **2003**, 43, 667–673.
- (35) Jorissen, R. N.; Gilson, M. K. Virtual screening of molecular databases using a support vector machine. *J. Chem. Inf. Model.* **2005**, 45, 549–561.
- (36) Bruce, C. L.; Melville, J. L.; Pickett, S. D.; Hirst, J. D. Contemporary QSAR classifiers compared. *J. Chem. Inf. Model.* **2007**, 47, 219–227.
- (37) Boutell, M. R.; Luo, J.; Shen, X.; Brown, C. M. C. Learning multi-label scene classification. *Pattern Recognit.* **2004**, 37, 1757–1771.
- (38) Hristozov, D.; Gasteiger, J.; Da Costa, F. B. Multilabeled classification approach to find a plant source for terpenoids. *J. Chem. Inf. Model.* **2008**, 48, 56–67.
- (39) Baraldi, P. G.; Cacciari, B.; Romagnoli, R.; Spalluto, G.; Monopoli, A.; Ongini, E.; Varani, K.; Borea, P. A. 7-Substituted 5-amino-2-(2-furyl)pyrazolo[4,3-e]-1,2,4-triazolo[1,5-c]pyrimidines as A<sub>2A</sub> adenosine receptor antagonists: a study on the importance of modifications at the side chain on the activity and solubility. *J. Med. Chem.* **2002**, 45, 115–126.
- (40) Baraldi, P. G.; Cacciari, B.; Romagnoli, R.; Spalluto, G.; Varani, K.; Gessi, S.; Merighi, S.; Borea, P. A. Pyrazolo[4,3-e]-1,2,4-triazolo[1,5-c]pyrimidine derivatives: a new pharmacological tool for the characterization of the human A<sub>3</sub> adenosine receptor. *Drug. Dev. Res.* **2001**, 52, 406–415.
- (41) Baraldi, P. G.; Cacciari, B.; Romagnoli, R.; Spalluto, G.; Moro, S.; Klotz, K. N.; Leung, E.; Varani, K.; Gessi, S.; Merighi, S.; Borea, P. A. Pyrazolo[4,3-e]-1,2,4-triazolo[1,5-c]pyrimidine derivatives as highly potent and selective human A<sub>3</sub> adenosine receptor antagonists: influence of the chain at the N<sup>8</sup> pyrazole nitrogen. *J. Med. Chem.* **2000**, 43, 4768–4780.
- (42) Baraldi, P. G.; Cacciari, B.; Romagnoli, R.; Klotz, K. N.; Spalluto, G.; Varani, K.; Gessi, S.; Merighi, S.; Borea, P. A. Pyrazolo[4,3-e]-1,2,4-triazolo[1,5-c]pyrimidine derivatives as adenosine receptor ligands: a starting point for searching A<sub>2B</sub> adenosine receptor antagonists. *Drug Dev. Res.* **2001**, 53, 225–235.
- (43) Baraldi, P. G.; Fruttarolo, F.; Tabrizi, M. A.; Preti, D.; Romagnoli, R.; El-Kashef, H.; Moorman, A.; Varani, K.; Gessi, S.; Merighi, S.; Borea, P. A. Design, synthesis, and biological evaluation of C9- and C2-substituted pyrazolo[4,3-e]-1,2,4-triazolo[1,5-c]pyrimidines as new A<sub>2A</sub> and A<sub>3</sub> adenosine receptors antagonists. *J. Med. Chem.* **2003**, 46, 1229–1241.
- (44) *OpenMosix*; <http://www.openMosix.org>, 2004.
- (45) *ADRIANA.Code*, version 2.2; Molecular Networks GmbH: Erlangen, Germany, 2008.
- (46) R Development Core Team. *R: a language and environment for statistical computing, version 2.8.1*, 2008. URL: <http://www.r-project.org> (accessed February, 2009).
- (47) Dimitriadou, E.; Hornik, K.; Leisch, F.; Meyer, D.; Weingessel, A. e1071: Misc functions of the Department of Statistics (e1071); TU Wien, 2005.
- (48) Gasteiger, J.; Li, X.; Rudolph, C.; Sadowski, J.; Zupan, J. Representation of molecular electrostatic potentials by topological feature maps. *J. Am. Chem. Soc.* **1994**, 116, 4608–4620.
- (49) Gasteiger, J.; Marsili, M. Iterative partial equalization of orbital electronegativity - a rapid access to atomic charges. *Tetrahedron* **1980**, 36, 3219–3228.
- (50) Gasteiger, J.; Saller, H. Calculation of the charge distribution in conjugated systems by a quantification of the resonance concept. *Angew. Chem., Int. Ed. Engl.* **1985**, 24, 687–689.
- (51) Moreau, G.; Broto, P. Autocorrelation of molecular structures, application to SAR studies. *Nouv. J. Chim.* **1980**, 4, 757–764.
- (52) Moreau, G.; Broto, P. The autocorrelation of a topological structure: a new molecular descriptor. *Nouv. J. Chim.* **1980**, 4, 359–360.
- (53) Wagener, M.; Sadowski, J.; Gasteiger, J. Autocorrelation of molecular surface properties for modeling corticosteroid binding globulin and cytosolic Ah receptor activity by neural networks. *J. Am. Chem. Soc.* **1995**, 117, 7769–7775.
- (54) Bauknecht, H.; Zell, A.; Bayer, H.; Levi, P.; Wagener, M.; Sadowski, J.; Gasteiger, J. Locating biologically active compounds in medium-sized heterogeneous datasets by topological autocorrelation vectors: dopamine and benzodiazepine agonists. *J. Chem. Inf. Comput. Sci.* **1996**, 36, 1205–1213.
- (55) Vapnik, V. *The Nature of Statistical Learning Theory*; Springer: New York, 1995.
- (56) Vapnik, V. *Statistical Learning Theory*; Wiley: New York, 1998.
- (57) Smola, A. J.; Scholkopf, B. *Learning with Kernels. Support Vector Machines, Regularization, Optimization, and Beyond*; MIT Press: Cambridge, MA, 2002.
- (58) Pastorin, G.; Da Ros, T.; Spalluto, G.; Deflorian, F.; Moro, S.; Cacciari, B.; Baraldi, P. G.; Gessi, S.; Varani, K.; Borea, P. A. Pyrazolo[4,3-e]-1,2,4-triazolo[1,5-c]pyrimidine derivatives as adenosine receptor antagonists. Influence of the N5 substituent on the affinity at the human A<sub>3</sub> and A<sub>2B</sub> adenosine receptor subtypes: a molecular modeling investigation. *J. Med. Chem.* **2003**, 46, 4287–4296.
- (59) Karaneswsky, C.; Thomson, D.; Michellys, A.; Ruppar, P.; Chen, J. H. Preparation of benzoic acid phenylacetic derivatives as HNT-4<sup>o</sup> modulators. *PCT Int. Appl.* **2005**, WO2005009104200–2005009140716.
- (60) Klotz, K. N.; Hessling, J.; Hegler, J.; Owman, C.; Kull, B.; Fredholm, B. B.; Lohse, M. J. Comparative pharmacology of human adenosine receptor subtypes-Characterization of stably transfected receptors in CHO cells. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **1998**, 357, 1–9.
- (61) De Lean, A.; Hancock, A. A.; Lefkowitz, R. J. Validation and statistical analysis of a computer modeling method for quantitative analysis of radioligand binding data for mixtures of pharmacological receptor subtypes. *Mol. Pharmacol.* **1982**, 21, 5–16.
- (62) Klotz, K. N.; Cristalli, G.; Grifantini, M.; Vittori, S.; Lohse, M. J. Photoaffinity labeling of A<sub>1</sub> adenosine receptors. *J. Biol. Chem.* **1985**, 260, 14659–14664.
- (63) Cheng, Y. C.; Prusoff, H. R. Relationships between the inhibition constant (K<sub>i</sub>) and the concentration of inhibitor which causes 50% inhibition (IC<sub>50</sub>) of an enzymatic reaction. *Biochem. Pharmacol.* **1973**, 22, 3099–3108.
- (64) Ivanciuc, O. Application of support vector machines in chemistry. *Rev. Comput. Chem.* **2007**, 23, 291–400.

CI900311J