CHAPTER 5

Substructure-Based Virtual Screening for Adenosine A_{2A} Receptor Ligands

This chapter is based upon:

van der Horst, E.; van der Pijl, R. Mulder-Krieger, T.; Bender, A.; IJzerman, A. P. Substructure-Based Virtual Screening for Adenosine A_{2A} Receptor Ligands. *ChemMedChem* **2011**, *6*, 2302-2311.

5.1 Abstract

A virtual ligand-based screening approach was designed and evaluated for the discovery of new A_{2A} adenosine receptor (AR) ligands. For comparison and evaluation, the procedures in a recently published virtual screening study that used the A_{2A} AR X-ray crystal structure for the target-based discovery of new A_{2A} ligands were largely followed. Several screening models were constructed by deriving the distinguishing structural features from selected sets of A_{2A} AR antagonists, so-called frequent substructure mining. The best model in statistical terms was subsequently applied to large-scale virtual screening of a commercial vendor library. This resulted in the selection of 36 candidates for acquisition and testing. Of the selected candidates, eight compounds significantly inhibited radioligand binding at A_{2A} AR (more than 30%) at 10 µM, corresponding to a "hit rate" of 22%. This hit rate is quite comparable to that of the referenced target-based virtual screening study, while both approaches yield new, non-overlapping sets of ligands.

5.2 Introduction

G protein-coupled receptors (GPCRs) are a large family of cell surface receptors that share a common architecture including seven transmembrane helices. Members of this family respond to a variety of extracellular stimuli, such as ions, small molecules, peptides, proteins, and even photons.[1] GPCRs are involved in many physiological and pathophysiological processes. They are therefore extensively studied and profitable as therapeutic targets: more than 40% of commercially available drugs are thought to act through GPCRs.[2,3] However, due to the difficulties associated with membrane-protein crystallization, GPCR drug discovery has historically been limited by the scarcity of high-resolution X-ray crystal structures. Only recently, a small number of crystal structures of pharmacologically relevant GPCRs became available, which brings the total number of available structures to six at the time of writing.[4-9] One of these new structures is the structure of the human A_{2A} adenosine receptor (AR) in complex with the high-affinity, subtype-selective antagonist **1** (Figure 1, ZM241385 [10]).[7] The human A_{2A} AR is one of four AR subtypes (A_1, A_{2A}, A_{2B}, and A_3), all of which respond to the endogenous agonist adenosine, albeit at different concentrations.[11] The A_{2A} AR mediates suppression of inflammatory and immune responses,[12] suggesting possible applications for selective A_{2A} agonists for the control of disproportionate inflammatory responses.[13] Selective A_{2A} antagonists, on the other hand, may prove of value to block immunosuppressive pathways that hinder antitumor immunity.[14,15] The A_{2A} AR inhibits D_2 dopamine receptor function in the striatum. Preventing this inhibition using selective A_{2A} antagonists is actively explored as treatment for Parkinson's disease.[16] The hypothesized therapeutic effect of targeting A_{2A} AR is supported by epidemiological evidence linking the consumption of caffeine (**2**, Figure 1) with a lower risk of Parkinson's disease.[17] Caffeine is a nonselective AR receptor antagonist with micromolar affinity for the A_{2A} AR. The functional interaction of the A_{2A} AR with the abovementioned dopamine D_2 receptor, but also with glutamate receptors in the CNS further suggests the A_{2A} AR as potential target for the treatment of a range of other diseases too, including psychiatric disorders.[18]

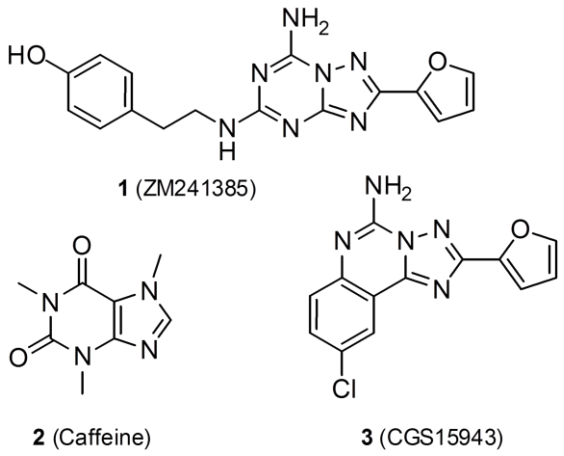


Figure 1. Chemical structures of known adenosine A_{2A} receptor antagonists.

With the recent publication of the A_{2A} AR X-ray crystal structure, AR drug discovery is not limited to ligand-based approaches or screening assays alone. The availability of the new A_{2A} AR structure has been a stimulus to the field, since it was successfully applied in two structure-based virtual screening studies.[19,20] These studies yielded high hit rates (41% and 35%); some of the active compounds had submicromolar affinities.

Although traditionally GPCR drug discovery did not have the privilege of available target structure information, a wealth of ligand information has been generated using ligand-based techniques and screening assays. With the ligand-based approaches, prior ligand knowledge is required, from which new ligands are derived. For the AR, almost all agonists are derivatives of the endogenous ligand adenosine, whereas AR antagonists are more diverse.[21] The AR antagonists can generally be divided into two classes according to their structure: the xanthine-like and the nonxanthine-like structures. Examples of the first class are xanthine derivatives such as caffeine (compound **2** in Figure 1) and theophylline. These have already been explored and optimized to great extent. Examples of the second class, the nonxanthine-like

structures, are compounds **1** and **3** (CGS15943 [22]) in Figure 1. These compounds share some structural resemblance: by aligning them using the furan substituent as an anchor, two rings of their ring system and the amine groups overlap. The (often unsubstituted) furan moiety is a common structural feature found in many A_{2A} antagonists.[23,24] It is typically linked to a triazole ring that is part of a ring-fused system of two or three heterocycles, with a single amine group attached. Describing molecules in terms of molecular parts, or fragments, such as rings and ring systems, substituents and functional groups, provides an abstraction for chemists to reason about molecules. This helps, for instance, for identification of a common core in a set of molecules, or for finding the substituents that contribute to bioactivity. It is also central to the planning of medicinal chemistry efforts, for instance for the design of derivatives around a common core to explore SAR, or to search for new ligands by exchanging molecular fragments between different structural classes.[25] Although these approaches may be relatively assured of success, truly new chemistry is only seldom being discovered this way; structural derivatives often resemble their parent structures. The similarity between known ligands that are analysed and the proposed candidates that result may be reduced by analyzing the structures in closer detail; instead of only focusing on predefined fragments, one might consider all possible structural patterns that are present in the molecules. This may result in the discovery of more unusual structural patterns that potentially offer greater diversity when applied to the discovery of new ligands.

In this study, we evaluated the application of a ligand-based virtual screening method for the discovery of new A_{2A} AR ligands. This method was designed to be complementary to the recent structure-based virtual screening studies, although it was exclusively based on ligand information only. Thus, also to evaluate our method, virtual screening was performed corresponding to the procedure of a recent structure-based virtual screening study for A_{2A} AR ligands by Katritch et al.[19] We constructed a computational screening model based on the structural features that distinguish A_{2A} AR antagonists from other (drug-like) compounds. This so-called frequent substructure

mining, i.e. the search for substructures that occur frequently in the A_{2A} AR ligand set but less so in the set with other drug-like compounds, was performed on selected libraries of A_{2A} AR antagonists of different activity ranges and using different chemical representations. Several screening models were constructed by varying the parameters of substructure generation and score calculation. These models were benchmarked and the best performing model was subsequently applied for large-scale screening on a commercial vendor library. Thirty-six hits from the virtual screening were acquired for testing. Of these, eight compounds inhibited binding of radioligand at the A_{2A} AR by more than 30% at 10 μM, corresponding to a "hit rate" of 22%.

5.3 Results

5.3.1 Identification of Discriminative Substructures for A_{2A} AR Antagonists

A large collection of human A_{2A} AR antagonists was retrieved from the ChEMBL database and grouped into three sets according to activity, i.e. the value of either pK_i, pIC50, or pEC50. The first set consisted of 892 low-affinity antagonists, with activity values between 5.0 and 8.0; the second set consisted of 255 high affinity antagonists that had activity values of 8.0 or higher; the third set consisted of the 1,147 antagonists of the two sets combined (the full activity range). For analysis of the antagonists, each antagonist set was compared against a background set of 10,000 drug-like molecules. Each pair of antagonist and background sets represented an individual training set from which a model was created. To analyze the structural features of the molecules, the molecular structures were first converted into a machine readable format, i.e. as molecular graphs. In addition to normal chemical representation, translation into one of three elaborate chemical representations was also explored as alternative representations when converting the source molecules to graphs. The translated source and background compounds were then subjected to frequent substructure mining. Frequent substructure mining is a data mining technique that finds all frequently occurring substructures that are present in a preset

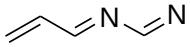
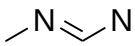
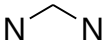
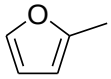
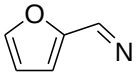
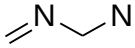
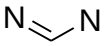
minimum number of molecules,[26,27] which in this study was set to 30% of the size of the set; a substructure is defined as any part of the molecule, ranging from a single atom to the complete structure. The number of generated substructures for each source set and chemical representation are provided in Table 1. The substructures generated from the three activity ranges in normal chemical representation, and the high-affinity source set in each of the three elaborate chemical representations are provided in Table SI1 to Table SI6 in Supporting Information. In general, smaller sets, such as the high-affinity antagonists set (255 molecules) result in significantly higher numbers of generated substructures compared to larger sets, such as the low-affinity (892 molecules) and combined antagonist (1,147 molecules) sets. With increasing set size, the chance for an individual substructure to occur more frequently than the set minimum decreases, resulting in finding fewer substructures. For infinitely large sets, only the most generic substructures would be found. In addition, the high mutual similarity between antagonists in the high-affinity set results in more substructures with frequencies above the support threshold.

Table 1. Number of generated substructures for each source set and chemical representation.

	Activity Range		
Representation[a]	**pK_i ≥ 8**	**pK_i ≥ 5**	**5 ≤ pK_i < 8**
Normal	**4,424**	**471**	**408**
Ar. bonds	**8,766**	**449**	**425**
Ar. atoms/bonds	**159,204**	**14,131**	**13,839**
Planar atoms/bonds	**160,363**	**13,879**	**12,922**

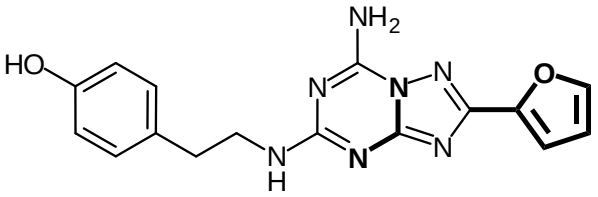
[a] Chemical representation of the molecules: “Normal” indicates common 2D structure drawing, “Ar. bonds”, “Ar. atoms/bonds”, and “Planar atoms/bonds” are three types of Elaborate Chemical Representations (ECR) that represent halogen atoms and heteroatoms with abstractions, and additional types for aromatic bonds, aromatic atom and bond type, or planar ring system atoms and bonds, respectively; see text for further information.

Table 2. Examples of discriminative substructures for high-affinity adenosine A_{2A} antagonists versus drug-like background compounds. For each substructure both absolute and relative occurrences are given. Note that the provided examples are all within the set of the 50 top ranking substructures. For further explanation, see text.

Nr	Substructure	A2A antagonists	Background compounds	Score contribution
a		**242 (94.9%)**	**326 (3.3%)**	**0.9164**
b		**252 (98.8%)**	**1,760 (17.6%)**	**0.8122**
c		**247 (96.9%)**	**1,955 (19.6%)**	**0.7731**
d		**209 (82.0%)**	**499 (5.0%)**	**0.7697**
e		**196 (76.9%)**	**63 (0.6%)**	**0.7623**
f		**201 (78.8%)**	**583 (5.8%)**	**0.7299**
g		**252 (98.8%)**	**2,758 (27.6%)**	**0.7124**

For the discussion of the results we will focus on the frequent substructure mining performed on the high-affinity A_{2A} AR antagonist set ($pK_i \geq 8$) represented in normal chemical representation. In total, 4,424 substructures were generated from this set (see Figure 1). A selection of these substructures is provided in Table 2. Note that the provided examples are all within the set of the 50 top ranking substructures (described below). All substructures in Table 2 are also present in compound **1** (note that substructures may overlap). For two of these substructures, **c** and **d** of Table 2, this is illustrated in Figure 2. This figure shows one example of how substructures are positioned in the molecules they originate from. Note that the methanediamine substructure, **c**, occurs three times in compound **1** (and also once in compound **3** and twice in compound **2**, Figure 1). For frequency calculations, however, it was counted

only once per molecule. Substructures **c** and **d** each represent one of two types of substructures that exist: substructures that are clear molecular fragments such as rings and functional groups (**d**) and substructures that have an unspecified shape (**c**).



1

Figure 2. The structure of compound 1 with two examples of discriminative substructures for A_{2A} antagonists highlighted in bold.

For each frequent substructure in the antagonist set, the occurrence in the background set was also determined. For instance, substructure **c** in Table 2 occurred in 247 of the 255 A_{2A} antagonists, or 96.9%, compared to 1,955 of the 10,000 (drug-like) background compounds, or 19.6%, while substructure d occurred in 82.0% and 5.0%, respectively. Since these substructures are frequently occurring in the A_{2A} antagonists and infrequently in background compounds they are signified as 'discriminative' for A_{2A} antagonists. This discriminative property is quantified by calculating the difference between the fraction of antagonists and the fraction of background compounds in which the substructures occur. This difference is referred to as 'score contribution' of a substructure and is used to rank the substructures. The top-ranked substructures, those with the highest score contribution, are the most 'discriminative' ones and were subsequently used for the screening. Note that although substructure **b** occurs more often in A_{2A} AR antagonists than **a** (252 vs. 242 times, respectively), **a** is ranked higher than **b**. This is because the frequency of **a** in the background set is considerably lower than that of **b**, resulting in a higher score contribution for **a**.

5.3.2 Score Calculation and Benchmark Screening

For the screening for A_{2A} antagonists, a score was calculated that was based on the presence of the top-ranked substructures in the compounds of the virtual library. Compounds with the highest score were considered more likely to be A_{2A} antagonists. Three sets of substructures were selected for screening. The first two sets consisted of the top 50 and the top 100 of the ranked substructures. The third set consisted of substructures ranked at positions 51 to 100 and served to investigate the impact of substructure ranking on screening. Three different methods of score calculation were tested: first, simple counting of substructures that match the chemical structure of a screening compound; second, summation of the score contributions of the matching substructures; third, multiplication of (the score contribution + 1) of all matching substructures. To determine the best method for finding A_{2A} AR ligands at the top of a ranked list, a small benchmark screening was performed on a set of 2000 drug-like decoy compounds spiked with 23 reference A_{2A} ligands This set was the same set as used for the benchmarking and validation of the referenced virtual screening study described by Katritch et al.,[19] and was provided by the first author of that study (private communication). Table 4 shows an example of score calculation based on the seven substructures from Table 2. Scores were calculated for one reference A_{2A} antagonist (compound **4**) and one decoy compound (compound **5**). The presence of substructures (rows) in one of the compounds (columns) is indicated by a mark. Score contributions are noted below each substructure. In the last row, the calculated scores for each of the aforementioned calculation methods are stated: counting, summation, and multiplication. The 'count' score, i.e. the number of substructures that match a compound, is higher for compound **4** than for compound **5** (six versus four, respectively). Ranking the compounds based on this score thus results in the reference A_{2A} antagonist (compound **4**) being ranked higher than the drug-like decoy (compound **5**). This holds also true for the other two scores.

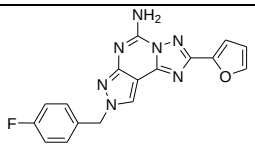
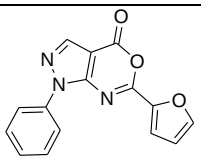
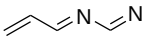
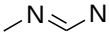
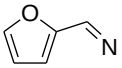
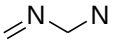
To optimize the screening protocol, a method was used to quantify the ability of the model to recognize A_{2A} AR antagonists, that is, to rank these more at the top. First,

compounds were ranked according to the calculated score; then, a subset of compounds was selected from the top down. Since the aim of ranking was to maximize the true positives, the top-set should be filled with reference A_{2A} AR ligands (true positives). The true positive rate (sensitivity) was plotted against the false positive rate (1 – specificity). This is known as a Receiver Operator Characteristic (ROC) plot. Since the goal was to maximize true positives and minimize false positives, the Area Under the Curve (AUC) of this ROC plot had to be maximized. The area under the curve in an ROC plot varies from 0.5 for a "random" model to 1.0 for a "perfect" model. The ROC scores for each model are provided in Table 3. These range from 0.98 for the best performing model to 0.92 for the worst. The best performing model consisted of the 100 top-ranked substructures derived from the high-affinity antagonist set represented in normal chemical representation, with 'multiplied' score calculation. The worst performing model consisted of the substructures derived from the high-affinity antagonist set in 'Planar' representation that were ranked at position 51 to 100, with either one of the three types of score calculation. Since virtual screening with 50 instead of 100 top-ranked substructures was 20% faster and performed equally well (similar ROC AUC of 0.98), this set was chosen for the large-scale virtual screening.

Table 3. ROC Scores for each substructure-based screening model, which expresses the performance of each model of identifying ligands from other drug-like compounds in the benchmark set. A score of 1.0 indicates a perfect model, while a score of 0.5 indicates a performance that is equal to random selection.

Representation / Score type	$pK_i \geq 8$			$pK_i \geq 5$			$5 \leq pK_i < 8$		
	1-100	1-50	51-100	1-100	1-50	51-100	1-100	1-50	51-100
Normal									
Counts	0.97997	0.97788	0.97278	0.96386	0.96250	0.95610	0.96302	0.94951	0.94840
Multiplied	0.98033	0.97809	0.97133	0.96415	0.96241	0.95599	0.96277	0.95039	0.95102
Summed	0.98004	0.97807	0.97133	0.96367	0.96224	0.95599	0.96253	0.95043	0.95100
Ar. bonds									
Counts	0.96512	0.96571	0.95442	0.97126	0.96671	0.97257	0.97292	0.96273	0.96953
Multiplied	0.96260	0.96548	0.95513	0.97041	0.96698	0.97416	0.97273	0.96295	0.96898
Summed	0.96253	0.96548	0.95513	0.96998	0.96698	0.97416	0.97253	0.96251	0.96898
Ar. atoms/ bonds									
Counts	0.92332	0.92317	0.91983	0.96622	0.94534	0.96609	0.96877	0.95235	0.97272
Multiplied	0.92321	0.92317	0.91983	0.96622	0.94534	0.96609	0.96870	0.95237	0.97317
Summed	0.92321	0.92317	0.91983	0.96622	0.94534	0.96609	0.96870	0.95237	0.97317
Planar atoms/bonds									
Counts	0.92220	0.92296	0.91778	0.96640	0.94515	0.96442	0.96941	0.95227	0.97195
Multiplied	0.92222	0.92296	0.91778	0.96640	0.94515	0.96442	0.96915	0.95227	0.97191
Summed	0.92222	0.92296	0.91778	0.96640	0.94515	0.96442	0.96915	0.95227	0.97191

Table 4. Example Score Calculation Based on a Selected Set of Substructures.

ID	Substructure[a]	 Compound **4**	 Compound **5**
a		0.9164	-
b		0.8122	-
c		0.7731	0.7731
d		0.7697	0.7697
e		0.7623	0.7623
f		-	0.7299
g		0.7124	-
		Count: 6	Count: 4
	Total Score[b]	Summed: 4.746	Summed: 3.035
		Multiplied: 32.886	Multiplied: 9.566

[a] List of substructures and their score contribution. [b] Calculated total score according to three different methods: "count" indicates counting of the number of matching substructures, "summed" indicates summation of the score contributions of these substructures, and "multiplied" indicates multiplication of (the score contribution + 1) of each substructure.

5.3.3 Substructure-based Virtual Screening

This 50 top-ranked model was subsequently used to screen a virtual library of commercially available compounds. For this, compounds from the ChemDiv Screening Compounds collection with molecular weight below 500 Dalton were used. This ChemDiv library represents an extensive collection of chemically diverse organic molecules. Each compound was given a score based on multiplication of each (score contribution + 1), for all the substructures that were present in the molecular structure of that compound. Compounds that had a substituted furan substituent were discarded. The rationale behind this filtering step was that many A_{2A} AR antagonists

contain a furan-2-yl group, and that substitutions of the 2-furanyl ring generally lead to decreased A_{2A} binding.[23,24] The high occurrence of this substructure is reflected in the list of substructures used for screening: the furan ring is present in three of the 50 top-ranked substructures (Table SI1 in Supporting Information), two of which are substructures d and e (Table 2). These substructures not only match the unsubstituted furan moiety but all substitution patterns of the furan ring, hence the extra filtering step. The remaining compounds were ranked according to score, and the top 1800 compounds with the highest score were considered as potential hits.

5.3.4 Selection of Candidate A_{2A} AR Ligands

The initial hits were clustered (Tanimoto distance, FCFP_4 fingerprints as available in Pipeline Pilot, see Experimental Section) into 41 clusters of sizes ranging from 1 to 503 molecules. Molecules with extreme logP values (calculated logP outside range -0.4 to 5.6) were disregarded. At least one compound was selected from each cluster, based on the values of the calculated solubility (logS), the calculated logP, and the score per heavy atom or ligand efficiency (LE). Thirty-six candidates were acquired for testing. For pragmatic reasons, compounds were ordered from one vendor only. This imposed a certain restriction on the present study, as our vendor, ChemDiv, was one of the four vendors eventually chosen by Katritch et al.[19] (and one of the six chosen by Carlsson et al.[20]) for compound acquisition.

5.3.5 Biological Evaluation in Radioligand Binding Studies

The 36 compounds that were selected from the virtual screening were first tested for binding in a radioligand displacement assay on HEK293 cell membranes expressing the human A_{2A} AR with [^{3}H]ZM241385 as the radioligand. This resulted in the identification of eight compounds that inhibited binding by ≥ 30% at 10 μM (as in the Carlsson study), corresponding to a "hit rate" of 22%. The dose-response curves were well behaved (with Hill slopes close to unity), as shown in Figure 3. The chemical structures of the eight compounds and their binding data for the A_{2A} adenosine receptor are

presented in Table 5, together with their affinities for the A_1 and A_3 AR, also determined in binding studies. The experimental results of the binding assays for all tested compounds are provided in Table SI8 of Supporting Information. Compound **7** had 770 nM affinity for A_{2A} AR with an 8-fold selectivity over A_1 and 1.5 fold selectivity over A_3 AR. Compound **6** had 1.24 µM affinity for the A_{2A} AR and was 2.6 fold selective over A_1. Compound **9** had an affinity of 8.72 µM for the A_{2A} AR and was subtype selective. Noteworthy are compound **12** and compound **14** (Table SI9 of Supporting Information). Compound **14** did not display significant affinity for the A_{2A} AR but was selective for the A_1 AR with 3.12 µM affinity. Compound **12** inhibited radioligand binding to the A_{2A} AR by 40% but had an affinity of 230 nM for the A_3 AR. None of the tested compounds displayed significant affinity for the A_{2B} AR. None of the ten compounds that were sampled from ranks 1800 to 3600 displayed any activity, which indicates that enrichment is found in the top-ranked structures.

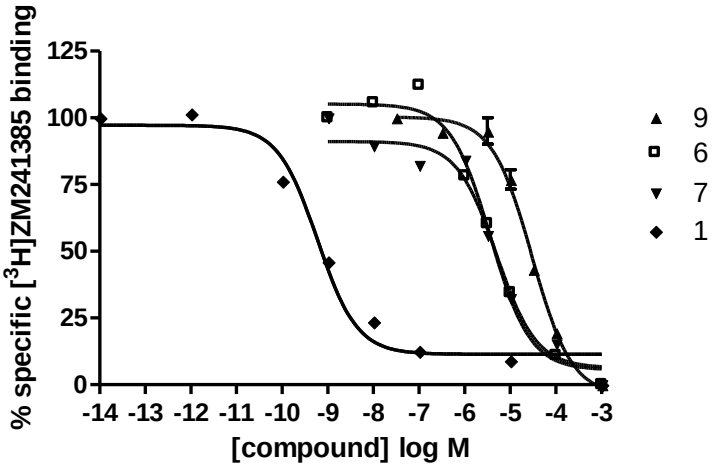
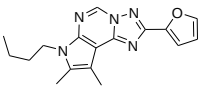
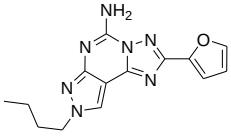
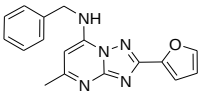
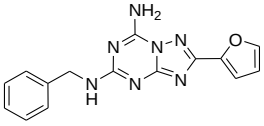
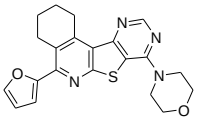
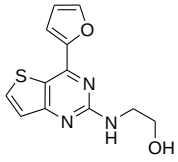
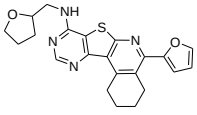
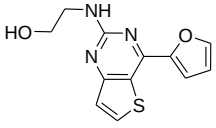
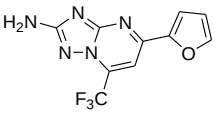
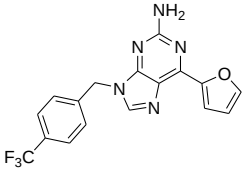
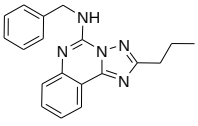
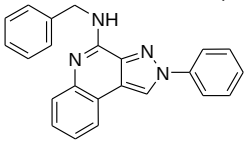


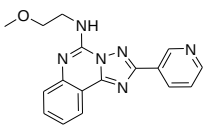
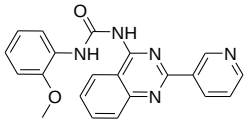
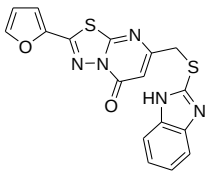
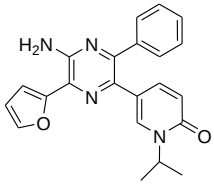
Figure 3. Dose-response curves of ligands **6**, **7**, and **9** for displacement of [^{3}H]ZM241385 from human A_{2A} AR. The dose-response curve of [^{3}H]ZM241385 (**1**) is included as a reference. Shown is a representative experiment performed in duplicate.

Table 5. Ligand Structures and Radioligand Binding Data for the Hits of the Substructure-based Screening.

ID (Rank)[a]	Ligand Structure	Head 3[b] A_1[c]	Head 3[b] A_{2A}[d]	Head 3[b] A_{2B}[e]	Head 3[b] A_3[f]	Most similar known AR ligand[g] (T_c)[h]
6 (10)		3.18 ± 0.36	1.24 ± 0.28	15%	40%	 (0.5)
7 (44)		5.97 ± 1.91	0.77 ± 0.12	0%	1.12 ± 0.18	 (0.63)
8 (475)		14%	39%	0%	0%	 (0.29)
9 (481)		33%	8.72 ± 1.90	0%	0%	 (0.34)
10 (518)		11%	32%	2%	0%	 (0.39)
11 (1594)		3.80 ± 0.57	42%	2%	5.64 ± 2.26	 (0.43)

(continued, see next page)

Table 5 continued

12 (1616)		3.09 ± 0.30	41%	13%	0.23 ± 0.07	 (0.39)
13 (1798)		32%	33%	0%	0%	 (0.27)

[a] Rank in the substructure-based virtual screening (ligands are ordered according to rank). [b] *K*i ± SEM (n = 3), or % displacement at 10 µM (n = 2). [c] Displacement or % displacement of specific [^{3}H]DPCPX binding to CHO cell membranes expressing human adenosine A_1 receptors. [d] Displacement or % displacement of specific [^{3}H]ZM241385 binding to HEK293 cell membranes expressing human adenosine A_{2A} receptors. [e] Displacement or % displacement of specific [^{3}H]PSB603 binding to CHO cell membranes expressing human adenosine A_{2B} receptors. [f] Displacement or % displacement of specific [^{3}H]PSB11 binding to HEK293 cells expressing human adenosine A_3 receptors. [g] Known adenosine receptor ligand with the highest chemical similarity (Tanimoto score) to the hit. [h] Tanimoto similarity coefficient based on ECFP4 fingerprints.

5.4 Discussion

The substructure-based screening method we developed proved useful to screen for new A_{2A} AR ligands. All models that were constructed performed well in retrospective screening of a small benchmark set. ROC scores were well above 0.90, indicating excellent performance. The suggested potential to find A_{2A} AR ligands was confirmed with the subsequent screening of a commercial vendor library: eight novel compounds were identified that inhibited radioligand binding (at 10 µM) by more than 30%, the best of which displayed an affinity of 770 nM for the A_{2A} AR.

The screening models that performed best according to the ROC score were all derived from the high-affinity ligands represented in normal chemical representation. The high accuracy of these models is somewhat expected considering the similarity between training set and test set: the reference ligands of the test set were high-affinity ligands

as well. In addition, the normal chemical representation offers the highest level of detail, which facilitates recognition of reference ligands. However, alternative (elaborate) chemical representations may also be beneficial for compound recall since these emphasize the more abstract binding features that are not recognized in normal chemical representation. Indeed, these models performed nearly as well as those produced with normal chemical representation and have the additional advantage of selecting more diverse chemistry as a result of abstractions.

The identified hits were filtered to be sufficiently different from known AR ligands according to the value of the Tanimoto similarity coefficient. Only one of the hits, compound **7** (Table 5), had a Tanimoto similarity score higher than 0.5 (i.e. 0.63), which is comparable to the chemical novelty found with structure-based screening. For instance, the screening studies of Carlsson et al. and Katritch et al. resulted in two out of seven, respectively four out of 13 ligands with a Tanimoto similarity score higher than 0.5 to known AR ligands.[19,20]

Interestingly, none of the tested compounds displayed binding at the A_{2B} AR. Consequently, selectivity of A_{2A} over A_{2B}, which is not considered trivial,[28] was reached. The same holds true for compounds with affinity for A_1, where selectivity over A_{2B} was (unintentionally) reached. These results are in line with the reputation of A_{2B} ARs as low-affinity receptors.[29]

A possible weakness of our approach lies in the screening method versus the method of candidate selection. These are not completely orthogonal which might negatively influence the outcome. The initial hits are identified based on a model derived from known A_{2A} AR antagonists and the subsequent candidates are selected based on dissimilarity (distance) to known AR ligands. Even though the high-affinity antagonists are only a subset of the total set of AR ligands, these sets are highly similar in terms of substructure distributions.

In terms of yield, our ligand-based method performed somewhat less than the two recent structure-based screening studies of Katritch et al. and Carlsson et al.[19,20] These

studies reported hit rates of 41% and 35%, of which the best binders had affinities of 0.06 μM and 0.2 μM, respectively. In addition, compared to ligand-based approaches, structure-based screening methods have the potential of exploring truly novel chemistry since these reflect the actual biological target. Ligand-based approaches on the other hand, may be prone to bias towards the ligands used for training. However, in practice, structure-based approaches may suffer from a similar ligand bias. First of all, the currently known GPCR crystal structures are co-crystalized with a ligand (except for two opsin structures,[30,31] which are not therapeutically relevant). Secondly, the parameters of a screening model are refined using ligand data. Examples are the inclusion and orientation of structured water molecules, orientation of amino acid residues, and the chosen scoring function. Interestingly, both aforementioned structure-based screening methods and the ligand-based screening method described here, resulted in the discovery of different ligands. This suggests that ligand-based methods and structure-based methods are complementary to each other and should be used side-by-side in the search for GPCR ligands. In fact, a virtual screening study may benefit from this complementarity by combining both methods in one virtual screening setup. For instance, the substructure-based part could be used for initial selection of compounds that would then be analyzed by the more computationally expensive structure-based virtual screening. Another example would be to apply the substructure-based screening to filter out compounds that contain substructures also found in a reference set, in order to screen compounds outside the chemical space of this reference set. A third example would be to combine the ranking obtained with each individual method to prioritize testing towards compounds that emerged as best in both methods.

5.5 Conclusion

In this study, we presented a virtual screening approach that was solely based on the structural features in existing ligands. Using a model constructed from these features,

new A_{2A} AR ligands were discovered from a commercial vendor library. From the 36 candidates that were selected for testing, eight compounds significantly inhibited A_{2A} AR radioligand binding, corresponding to a hit rate of 22%.

5.6 Experimental Section

5.6.1 Software

Except for chemical representation and frequent substructure mining, all data manipulation was performed in Pipeline Pilot 6.1.5.0 Student Edition (Scitegic). This includes the querying of molecular libraries, score and property calculation, and ranking. Translation of structures into different chemical representations was performed using a custom script; frequent substructure mining was performed using the program Gaston.[32]

5.6.2 Datasets and Chemical Representation

To assemble a set of human A2A AR antagonists, ligand structures and activity data were retrieved from the ChEMBL database,[33] selecting compounds with activity, i.e. either pKi, pIC50, or pEC50, of 5.0 or higher. Compounds that were explicitly annotated as antagonist were included, whereas compounds with reports of agonistic effects were removed. Subsequent manual inspection was performed to ensure further removal of any alleged agonists, for instance, compounds that were highly similar to adenosine. The antagonists were used to create three source datasets, based on the activity range of the antagonists. The first set consisted of 892 relatively low affinity antagonists (activity values ranging from 5.0 to 8.0), the second set consisted of 255 high affinity antagonists (activity values of 8.0 and higher), and the third set consisted of 1,147 antagonists of both sets combined. For analysis, each set was contrasted with a background of 'average' compounds (background set). The background set consisted of 10,000 randomly selected compounds from the drug-like subset of the ZINC

database (accessed: February 12, 2010), a collection of all available chemical compounds.[34] Together, the antagonist set and background compounds served as a single 'training set' for model creation.

Chemical structures were represented as graphs according to four representations: one 'normal' and three 'elaborate' chemical representations. These representations have been described in detail in the work of Kazius et al.[35] and chapter 3 (or reference 36). In short, 'normal' chemical representation translates chemical structures into graphs without modifications. The three 'elaborate' chemical representations differ as follows: the first uses a special bond type for aromatic bonds; the second has a special atom and bond type for aromatic atoms and bonds, and the third has special types for atoms and bonds in planar ring systems. In addition, all three elaborate representations used an abstraction for heteroatoms and halogen atoms. Labels attached to the abstract atoms specify the atom type and number of bonded hydrogens.

5.6.3 Generation of Frequent Substructures Occurring in Antagonists for the Human A_{2A} AR

Frequent substructure sets were generated using the frequent graph miner Gaston which finds all possible frequently occurring substructures in a set of molecules.[37] More details on the underlying algorithm are provided in chapters 2 and 3 (or references 36 and 38). For each substructure, the number of molecules the substructure occurred in was calculated. The difference between the relative occurrence (fraction) of a substructure in the antagonists set and the background set is the score contribution of that substructure. Substructures were ranked according to the score contribution in descending order. Three substructure sets were selected: the top 50 best substructures, the top 100 best substructures, and the substructures with rank 51 to 100.

In total, 36 substructure sets were generated corresponding to the combined features: three affinity classes times four chemical representations times three substructure

occurrence cut-offs. A number of these substructure sets are collected in Table SI1 to Table SI6 in Supporting Information.

5.6.4 Substructure-based Virtual Screening – Ranking of Compounds

To rank compounds in order of likelihood to be an A_{2A} receptor antagonist, a score was calculated based on the previously generated substructure set. The score for a compound was calculated as follows: for each substructure in the set, presence in the compound was determined. For the substructures that occurred in a compound, the score contribution was used to calculate the final score for that compound. Three different methods of score calculation were tested: summing the score contributions, multiplication of each (score contribution + 1), and counting the number of substructures that map. Compounds were ranked according to total score in descending order. Within this framework, the highest scoring compounds had the highest probability of being an adenosine A_{2A} receptor antagonist.

5.6.5 Small Scale Benchmark Screening

To find the best performing design and score calculation, the small-scale benchmark set from Katritch et al. was used.[19] This set consisted of 2,000 drug-like decoy molecules randomly selected from the ChemDiv drug-like discovery collection, spiked with 23 high affinity A_{2A} receptor antagonists (the chemical structures of the decoy molecules and the 23 antagonists are provided in Table SI9 of Supplemental Information). Receiver Operator Characteristic (ROC) curves were plotted and the Area Under the Curve (AUC) calculated with Pipeline Pilot 6.1.5.0 Student Edition. The substructure set and score calculation that resulted in the highest AUC, were selected for the large-scale virtual screening.

5.6.6 Virtual Screening of the ChemDiv Library and Candidate Selection

The complete ChemDiv vendor library was retrieved from ZINC and compounds with molecular weight above 500 were removed. This set was screened using the best

performing substructure set and score calculation as described in the previous paragraph for the small-scale screening benchmark. Compounds that contained a furan ring that was disubstituted, largely furanylfurans, were filtered out, as earlier SAR studies had shown that this is incompatible with high A_{2A} receptor affinity.[23,24] The set of 1,800 highest scoring compounds was taken as the initial hit set (Table SI7 of the Supporting Information). This number was chosen to be consistent with the selection of candidates by Katritch et al.[19] As a guideline for final compound selection, molecules were clustered using the 'Cluster Molecules' component from the Data Modeling Component Collection in Pipeline Pilot 6.1.5.0 Student Edition, based on FCFP_4 fingerprints and maximum Tanimoto distance to the cluster center set to 0.6. All other parameters of the 'Cluster Molecules'-component were kept at their default values. As known adenosine ligands are largely uncharged, charged molecules (at pH = 7.4) were filtered out, as well as molecules with calculated LogP values outside the range -0.4 to 5.6. From each cluster, at least one compound was selected based on predicted solubility, calculated LogP, and a measure similar to ligand efficiency in docking: the score divided by the number of heavy atoms. Compounds that had a Tanimoto similarity of 0.7 or more to known adenosine receptor ligands were discarded, to maximize the chance of identifying unprecedented hits. This set of 2,534 known adenosine receptor ligands was compiled from GLIDA, ChEMBL,[33] and the literature.[19,24,39,40] A final candidate set of limited size (36 compounds, see Table SI8 of Supplemental Information) was compiled and purchased for testing, based on chemical diversity and direct availability from ChemDiv Inc. (Chemical Diversity Research Institute, San Diego, CA, USA). To explore the top hits further, ten additional (diverse) compounds were selected that had a score equal or better than the lowest scoring reference compound but that were not part of the top 1,800 highest scoring compounds. These compounds were also ordered from ChemDiv Inc.

5.6.7 Biology

5.6.7.1 Binding Studies

[^{3}H]DPCPX and [^{3}H]ZM241385 (4-(2-[7-amino-2-(2-furyl)[1,2,4]triazolo[2,3-α][1,3,5]triazin-5-ylamino]ethyl)phenol) were purchased from ARC Inc. St Louis, USA. [^{3}H]PSB603 and [^{3}H]PSB11 were gifts from Prof. C.E. Müller (Bonn, Germany). CHO cells expressing the human adenosine A_1 receptor were provided by Dr. Andrea Townsend-Nicholson, University College of London, UK. HEK 293 cells stably expressing the human adenosine A_{2A} and A_3 receptor were gifts from Dr. Wang (Biogen) and Dr. K.-N. Klotz (University of Wűrzburg, Germany), respectively. CHO cells expressing the human A_{2B} receptor were provided by Dr. Steve Rees (GlaxoSmithKline, UK).

5.6.7.2 Adenosine A_1 Receptor

Affinity at the A_1 receptor was determined on membranes from CHO cells expressing the human receptors, using [^{3}H]DPCPX as the radioligand. Membranes containing 5 µg of protein were incubated in a total volume of 100 µL of 50 mM Tris•HCl (pH 7.4) and [^{3}H]DPCPX (final concentration 1.6 nM) for 1 h at 25 °C in a shaking water bath. Nonspecific binding was determined in the presence of 100 µM CPA. The incubation was terminated by filtration over pre-wetted Whatman GF/B filters under reduced pressure with a Brandel harvester. Filters were washed three times with ice-cold buffer and placed in scintillation vials. Emulsifier Safe (3.5 mL) was added, and after 2 h radioactivity was counted in a TriCarb 2900TR liquid scintillation counter.

5.6.7.3 Adenosine A_{2A} Receptor

At the A_{2A} receptor, affinity was determined on membranes from HEK 293 cells stably expressing this human receptor, using [^{3}H]ZM241385 as the radioligand. Membranes containing 40 µg of protein were incubated in a total volume of 100 µL of 50 mM Tris•HCl (pH 7.4) and [^{3}H]ZM241385 (final concentration 1.7 nM) for 2 h at 25 °C in a shaking water bath. Nonspecific binding was determined in the presence of 100 µM CGS. The incubation was terminated by filtration over pre-wetted Whatman GF/B filters under reduced pressure with a Brandel harvester. Filters were washed three

times with ice-cold buffer and placed in scintillation vials. Emulsifier Safe (3.5 mL) was added, and after 2 h radioactivity was counted in a TriCarb 2900TR liquid scintillation counter.

5.6.7.4 Adenosine A_{2B} Receptor

At the A_{2B} receptor, radioligand displacement was determined on membranes from CHO cells stably transfected with human A_{2B} receptor, using [^{3}H]PBS603 as the radioligand. Membranes containing 15 µg of protein were incubated in a total volume of 100 µL of 50 mM Tris•HCl (pH 7.4), 1U/mL ADA, 0.1 w/v % CHAPS (pH 8.2 at 5 °C), and [^{3}H]PBS603 (final concentration 1.0 nM) for 2 h at 25 °C in a shaking water bath. Nonspecific binding was determined in the presence of 100 µM NECA. The incubation was terminated by filtration over pre-wetted Whatman GF/C filters under reduced pressure with a Brandel harvester. Filters were washed three times with ice-cold 50 mM Tris•HCl, pH7.4 + 0.1% BSA buffer and placed in scintillation vials. Emulsifier Safe (3.5 mL) was added, and after 5 h radioactivity was counted in a TriCarb 2900TR liquid scintillation counter.

5.6.7.5 Adenosine A_3 Receptor

The affinity at the A_3 receptor was measured on membranes from HEK 293 cells stably expressing the human A_3 receptor, using [^{3}H]PSB11 as the radioligand. Membranes containing 25 µg of protein were incubated in a total volume of 100 µL of 50 mM Tris•HCl, 10 mM MgCl2, 1 mM EDTA, 0.01% CHAPS (pH 7.4), and [^{3}H]PSB11 (final concentration 4 nM) for 1 h at 37 °C in a shaking water bath. Nonspecific binding was determined in the presence of 100 µM R-PIA. The incubation was terminated by filtration over pre-wetted Whatman GF/B filters under reduced pressure with a Brandel harvester. Filters were washed three times with ice-cold buffer and placed in scintillation vials. Radioactivity was counted in a Wallac 1470 Wizard gamma counter.

5.6.7.6 Data Analysis

Ki values were calculated using a nonlinear regression curve-fitting program (GraphPad Prism 5.0, GraphPad Software Inc., San Diego, CA). Ki values of radioligands were 1.6,

1.7, 0.41, and 4.9 nM for [3H]DPCPX, [3H]ZM241385, [3H]PSB603, and [3H]PSB11, respectively.

5.7 Acknowledgements

The authors thank Vsevolod Katritch for providing the benchmark set, and Christa Mueller for providing [^{3}H]PSB603 and [^{3}H]PSB11. This study was performed within the framework of the Dutch Top Institute Pharma, project number: D1-105 (The GPCR Forum).

5.8 References

(1) R. Fredriksson, M. C. Lagerstrom, L.-G. Lundin, H. B. Schioth, *Mol. Pharmacol.* **2003**, *63,* 1256-1272.

(2) M. C. Lagerstrom, H. B. Schioth, *Nat. Rev. Drug Discov.* **2008**, *7,* 339-357.

(3) J. P. Overington, B. Al-Lazikani, A. L. Hopkins, *Nat. Rev. Drug Discov.* **2006**, *5,* 993-996.

(4) K. Palczewski, T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. L. Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto, M. Miyano, *Science* **2000**, *289,* 739-745.

(5) V. Cherezov, D. M. Rosenbaum, M. A. Hanson, S. G. F. Rasmussen, F. S. Thian, T. S. Kobilka, H.-J. Choi, P. Kuhn, W. I. Weis, B. K. Kobilka, R. C. Stevens, *Science* **2007**, *318,* 1258-1265.

(6) T. Warne, M. J. Serrano-Vega, J. G. Baker, R. Moukhametzianov, P. C. Edwards, R. Henderson, A. G. W. Leslie, C. G. Tate, G. F. X. Schertler, *Nature* **2008**, *454,* 486-491.

(7) V. Jaakola, M.T. Griffith, M. A. Hanson, V. Cherezov, E. Y. T. Chien, J. R. Lane, A. P. IJzerman, R. C. Stevens, *Science* **2008**, *322*, 1211-1217.

(8) B. Wu, E. Y. T. Chien, C. D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F. C. Bi, D. J. Hamel, P. Kuhn, T. M. Handel, V. Cherezov, R. C. Stevens, *Science* **2010**, *330*, 1066-1071.

(9) E. Y. T. Chien, W. Liu, Q. Zhao, V. Katritch, G. Won Han, M. A. Hanson, L. Shi, A. H. Newman, J. A. Javitch, V. Cherezov, R. C. Stevens, *Science* **2010**, *330,* 1091-1095.

(10) S. M. Poucher, J. R. Keddie, P. Singh, S. M. Stoggall, P. W. Caulkett, G. Jones, M. G. Coll, *Br. J. Pharmacol.* **1995**, *115*, 1096-1102.

(11) B. B. Fredholm, A. P. IJzerman, K. A. Jacobson, K.-N. Klotz, J. Linden, *Pharmacol. Rev.* **2001**, *53,* 527-552.

(12) T. M. Palmer, M. A. Trevethick, *Br. J. Pharmacol* **2008**, *153,* S27-S34.

(13) T. Eckle, M. Koeppen, H. K. Eltzschig, *Physiology (Bethesda, Md.)* **2009**, *24*, 298-306.

(14) A. Ohta, E. Gorelik, S. J. Prasad, F. Ronchese, D. Lukashev, M. K. K. Wong, X. Huang, S. Caldwell, K. Liu, P. Smith, J.-F. Chen, E. K. Jackson, S. Apasov, S. Abrams, M. Sitkovsky, *Proc. Natl. Acad. Sci.* **2006**, *103,* 13132 -13137.

(15) A. Ohta, M. Sitkovsky, *Curr. Opin. Pharmacol.* **2009**, *9,* 501-506.

(16) P. Jenner, A. Mori, R. Hauser, M. Morelli, B. B. Fredholm, J. F. Chen, *Parkinsonism Relat. Disord.* **2009**, *15,* 406-413.

(17) M. A. Hernán, B. Takkouche, F. Caamaño-Isorna, J. J. Gestal-Otero, *Ann. Neurol.* **2002**, *52,* 276-284.

(18) R. A. Cunha, S. Ferré, J.-M. Vaugeois, J.-F. Chen, *Curr. Pharm. Des.* **2008,** *14,* 1512-1524.

(19) V. Katritch, V.-P. Jaakola, J. R. Lane, J. Lin, A. P. IJzerman, M. Yeager, I. Kufareva, R. C. Stevens, R. Abagyan, *J. Med. Chem.* **2010**, *53*, 1799-1809.
(20) J. Carlsson, L. Yoo, Z.-G. Gao, J. J. Irwin, B. K. Shoichet, K. A. Jacobson, *J. Med. Chem.* **2010**, *53,* 3748-3755.
(21) K. A. Jacobson, Z.-G. Gao, *Nat. Rev. Drug Discov.* **2006**, *5,* 247-264.
(22) E. Ongini, S. Dionisotti, S. Gessi, E. Irenius, B. B. Fredholm, *Naunyn-Schmiedeberg's Arch. Pharmacol.* **1999**, *359,* 7-10.
(23) P. G. Baraldi, F. Fruttarolo, M. A. Tabrizi, D. Preti, R. Romagnoli, H. El-Kashef, A. Moorman, K. Varani, S. Gessi, S. Merighi, P. A. Borea, *J. Med. Chem.* **2003**, *46,* 1229-1241.
(24) M. Mantri, O. de Graaf, J. van Veldhoven, A. Göblyös, J. K. von Frijtag Drabbe Künzel, T. Mulder-Krieger, R. Link, H. de Vries, M. W. Beukers, J. Brussee, A. P. IJzerman, *J. Med. Chem.* **2008**, *51,* 4449-4455.
(25) J. P. D. van Veldhoven, L. C. W. Chang, J. K. von Frijtag Drabbe Künzel, T. Mulder-Krieger, R. Struensee-Link, M. W. Beukers, J. Brussee, A. P. IJzerman, *Bioorg. Med. Chem.* **2008**, *16,* 2741-52.
(26) S. Nijssen, J. N. Kok, *Proc. of the 2004 IEEE Int. Conf. Syst. Man Cybern.* **2004**, *vol. 5*, pp. 4571-4577.
(27) C. Borgelt, M. R. Berthold, D. E. Patterson in *Symbolic and Quantitative Approaches to Reasoning with Uncertainty; Lecture Notes in Computer Science*; Springer Berlin / Heidelberg, **2005**; Vol. 3571, pp. 1002-1013.
(28) J. Wei, S. Wang, S. Gao, X. Dai, Q. Gao, *J. Chem. Inf. Model.* **2007**, *47,* 613-625.
(29) M. W. Beukers, H. den Dulk, E. W. van Tilburg, J. Brouwer, A. P. IJzerman, *Mol. Pharmacol.* **2000**, *58,* 1349-1356.
(30) J. H. Park, P. Scheerer, K. P. Hofmann, H.-W. Choe, O. P. Ernst, *Nature* **2008**, *454,* 183-187.
(31) P. Scheerer, J. H. Park, P. W. Hildebrand, Y. J. Kim, N. Krauß, H.-W. Choe, K. P. Hofmann, O. P. Ernst, *Nature* **2008**, *455*, 497-502.
(32) S. Nijssen, "GASTON" to be found under http://www.liacs.nl/~snijssen/gaston/index.html, **2004**.
(33) P. de Matos, R. Alcantara, A. Dekker, M. Ennis, J. Hastings, K. Haug, I. Spiteri, S. Turner, C. Steinbeck, *Nucl. Acids Res.* **2010**, *38,* D249-254.
(34) J. J. Irwin, B. K. Shoichet, *J. Chem. Inf. Model.* **2005**, *45,* 177-182.
(35) J. Kazius, S. Nijssen, J. Kok, T. Bäck, A. P. IJzerman, *J. Chem. Inf. Model.* **2006**, *46,* 597-605.
(36) E. van der Horst, Y. Okuno, A. Bender, A. P. IJzerman, *J. Chem. Inf. Model.* **2009**, *49,* 348-360.
(37) S. Nijssen, J. N. Kok, *Proc. 10th ACM SIGKDD Int. Conf. Knowl. Discovery Data Min.* (Seattle, WA, USA) **2004**, pp. 647-652.
(38) E. van der Horst, A. P. IJzerman in *Fragment-Based Drug Discovery: a Practical Approach*, *1st ed.* (Eds.: E. R. Zartler, M. J. Shapiro), John Wiley and Sons, Ltd., Chichester, West-Sussex, United Kingdom, **2008**, pp. 199-222.
(39) L. C. W. Chang, R. F. Spanjersberg, J. K. von Frijtag Drabbe Künzel, T. Mulder-Krieger, J. Brussee, A. P. IJzerman, *J. Med. Chem.* **2006**, *49,* 2861-2867.

(40) L. C. W. Chang, J. K. von Frijtag Drabbe Künzel, T. Mulder-Krieger, J. Westerhout, T. Spangenberg, J. Brussee, A. P. IJzerman, *J. Med. Chem.* **2007**, *50,* 828-34.