

# Ligand-, structure- and pharmacophore-based molecular fingerprints: a case study on adenosine A<sub>1</sub>, A<sub>2A</sub>, A<sub>2B</sub>, and A<sub>3</sub> receptor antagonists

Francesco Sirci · Laura Goracci · David Rodríguez ·  
Jacqueline van Muijlwijk-Koezen · Hugo Gutiérrez-de-  
Terán · Raimund Mannhold

Received: 2 July 2012 / Accepted: 20 September 2012 / Published online: 12 October 2012  
© Springer Science+Business Media Dordrecht 2012

**Abstract** FLAP fingerprints are applied in the ligand-, structure- and pharmacophore-based mode in a case study on antagonists of all four adenosine receptor (AR) subtypes. Structurally diverse antagonist collections with respect to the different ARs were constructed by including binding data to human species only. FLAP models well discriminate “active” (=highly potent) from “inactive” (=weakly potent) AR antagonists, as indicated by enrichment curves, numbers of false positives, and AUC values. For all FLAP modes, model predictivity slightly decreases as follows: A<sub>2B</sub>R > A<sub>2A</sub>R > A<sub>3</sub>R > A<sub>1</sub>R antagonists. General performance of FLAP modes in this study is: ligand- > structure-

> pharmacophore- based mode. We also compared the FLAP performance with other common ligand- and structure-based fingerprints. Concerning the ligand-based mode, FLAP model performance is superior to ECFP4 and ROCS for all AR subtypes. Although focusing on the early first part of the A<sub>2A</sub>, A<sub>2B</sub> and A<sub>3</sub> enrichment curves, ECFP4 and ROCS still retain a satisfactory retrieval of actives. FLAP is also superior when comparing the structure-based mode with PLANTS and GOLD. In this study we applied for the first time the novel FLAPPharm tool for pharmacophore generation. Pharmacophore hypotheses, generated with this tool, convincingly match with formerly published data. Finally, we could demonstrate the capability of FLAP models to uncover selectivity aspects although single AR subtype models were not trained for this purpose.

**Electronic supplementary material** The online version of this article (doi:10.1007/s10822-012-9612-8) contains supplementary material, which is available to authorized users.

F. Sirci (✉) · L. Goracci  
Laboratory for Chemometrics and Molecular Modeling,  
Chemistry Department, University of Perugia, Via Elce  
di Sotto, 10, 06123 Perugia, Italy  
e-mail: francesco@chemiome.chm.unipg.it

D. Rodríguez · H. Gutiérrez-de-Terán  
Fundación Pública Galega de Medicina Xenómica–SERGAS,  
Complejo Hospitalario Universitario de Santiago, A Choupana,  
s/n, E-15706, Santiago de Compostela, Spain

J. van Muijlwijk-Koezen  
Leiden/Amsterdam Center for Drug Research (LACDR),  
Division of Medicinal Chemistry, Department of  
Pharmacochemistry, Faculty of Exact Sciences, VU University  
Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam,  
The Netherlands

R. Mannhold  
Department of Laser Medicine, Molecular Drug Research  
Group, Heinrich-Heine-Universität, Universitätsstraße 1,  
40225 Düsseldorf, Germany

**Keywords** Adenosine receptor subtypes · FLAP ·  
G protein coupled receptors (GPCRs) · GOLD · GRID ·  
PLANTS · ECFP4 · ROCS · Ligand-based virtual  
screening · Structure-based virtual screening ·  
Pharmacophore-based virtual screening

## Abbreviations

AR	Adenosine receptors
AUC	Area under the curve
ECFP4	Extended-connectivity fingerprint
EF	Enrichment factor
FLAP	Fingerprint of ligands and proteins
FN	False negative(s)
FP	False positive(s)
GOLD	Genetic optimization of ligand docking H-bonding, hydrogen bonding
LDA	Linear discriminant analysis
MIF	Molecular interaction fields
PLANTS	Protein-ligand ant colony optimization

QSAR	Quantitative structure–activity relationships
ROC	Receiver operating characteristic
ROCS	Rapid overlay of chemical structures

## Introduction

Endogenous adenosine affects numerous physiological and pathophysiological processes by stimulating four cell surface receptors, which are classified as subtypes A<sub>1</sub>R, A<sub>2A</sub>R, A<sub>2B</sub>R, and A<sub>3</sub>R. These are members of the class A of the superfamily of G-protein-coupled receptors [1]. The four adenosine receptors (ARs) have been pharmacologically characterized and cloned the A<sub>3</sub>R subtype being the last one in this regard [2, 3]. A<sub>1</sub>R and A<sub>3</sub>R cluster in one branch (human receptors present a sequence identity of 49 %) which preferentially couple to G<sub>i</sub> protein, thereby inhibiting adenylate cyclase and the production of cAMP. The second branch of ARs, involving A<sub>2A</sub>R and A<sub>2B</sub>R, share 59 % identity for the human receptors and couple to G<sub>s</sub> or G<sub>o</sub> protein, stimulating the production of cAMP. Furthermore, it was found that ARs can also couple to other second messenger systems, such as calcium or potassium channels (A<sub>1</sub>) and phospholipase C (A<sub>1</sub>R, A<sub>2B</sub>R, and A<sub>3</sub>R) [4].

The four ARs regulate various physiological processes, including: inhibition of lipolysis, neurotransmitter release and insulin/glucagon release, reduced heart rate, and sleep for the A<sub>1</sub>R; wakefulness, locomotion, and vasodilatation for the A<sub>2A</sub>R; vascular integrity regulated by A<sub>2B</sub>R and modulation of cerebral and cardiac ischemic damage, inflammation, normal and tumor cell growth by the A<sub>3</sub>R. An overview of the physiological regulator and distress role of adenosine is given by Fredholm [5].

Accordingly, AR antagonists have a therapeutic potential in the acute treatment of stroke and as antiasthmatic, antiallergic, cerebroprotective and anti-cancer agents [1, 6–8], as summarized in Table 1. Considering their great therapeutic potential, intense research efforts continue to be invested into the development of novel and selective AR

antagonists. Many ligand-, structure- and pharmacophoric-based studies have been reported in the literature for ARs modelling [9–20].

In this study, we developed ligand-, structure- and pharmacophore-based models for discriminating highly potent (“active”) from weakly potent (“inactive”) antagonists at all four adenosine receptor subtypes. For this purpose, the physicochemical property space was explored, with the aim to collect druggable and focused datasets for each AR subtype [21]. Only antagonists with reported affinity data (pK<sub>i</sub>) for human ARs were included in this study. Ligand- and structure-based approaches were applied, in order to build, validate and compare FLAP models [22, 23]. Performance of the ligand-based FLAP approach was compared with ROCS [24] and ECFP4 [25] approaches; performance of structure-based FLAP with PLANTS [26] and GOLD [27]. MIF analysis was applied to all AR subtype ligands in order to derive subtype specific pharmacophoric features. Finally, literature-known subtype selective antagonists were used for external prediction in each subtype model, in order to check the capability to retrieve selective compounds.

## Methods

### Dataset generation

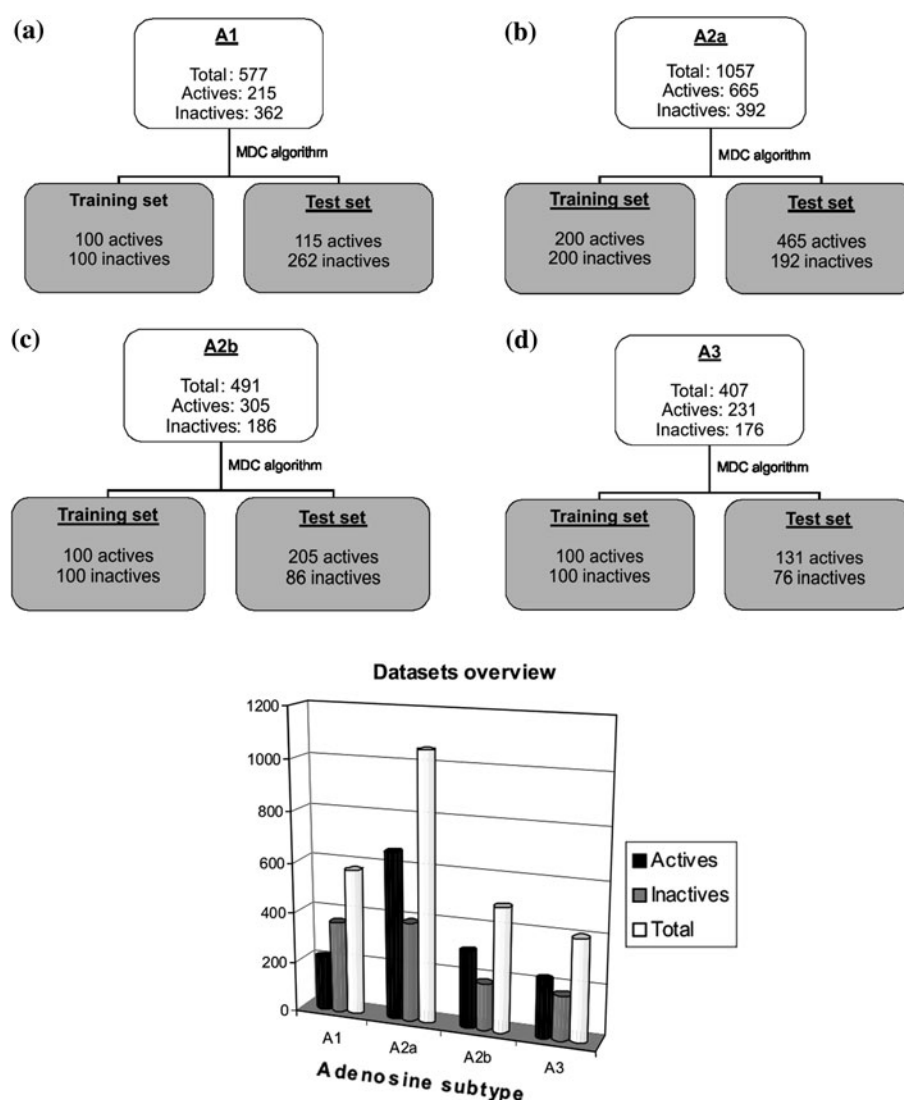
ChEMBL [28] is currently the most extensive public database of compound bioactivity data, including e.g. binding constants and functional pharmacology annotations. The ChEMBL search engine (<https://www.ebi.ac.uk/chembl>) has been used to retrieve AR antagonists, in order to build and validate in silico models for all four AR subtypes. We exclusively selected antagonists with reported affinity data (pK<sub>i</sub>) to human ARs (Fig. 1). Filtering for functional activity and target species was done using ChEMBL annotations taken from literature. Agonist structures and AR ligands with unknown affinity to human species were discarded. 3D-mol2 structures were generated from ChEMBL SMILES strings with CORINA v.3.46 [29]. Stereoisomers were automatically generated during the creation of FLAP database,

**Table 1** Therapeutic potential of AR antagonists

A <sub>1</sub> R antagonists	A <sub>2A</sub> R antagonists	A <sub>2B</sub> R antagonists	A <sub>3</sub> R antagonists
Asthma	Parkinson disease	Asthma	Asthma
Chronic heart failure	Huntington disease	Cardiovascular disorders	Cancer
Congestive heart failure	Cancer	Diarrhea	Glaucoma
COPD	Depression	Inflammation	Inflammation
Renal failure		Pain	Stroke

For a comprehensive overview, the reader is referred to Gessi et al. [6] as well as Müller and Jacobson [7]

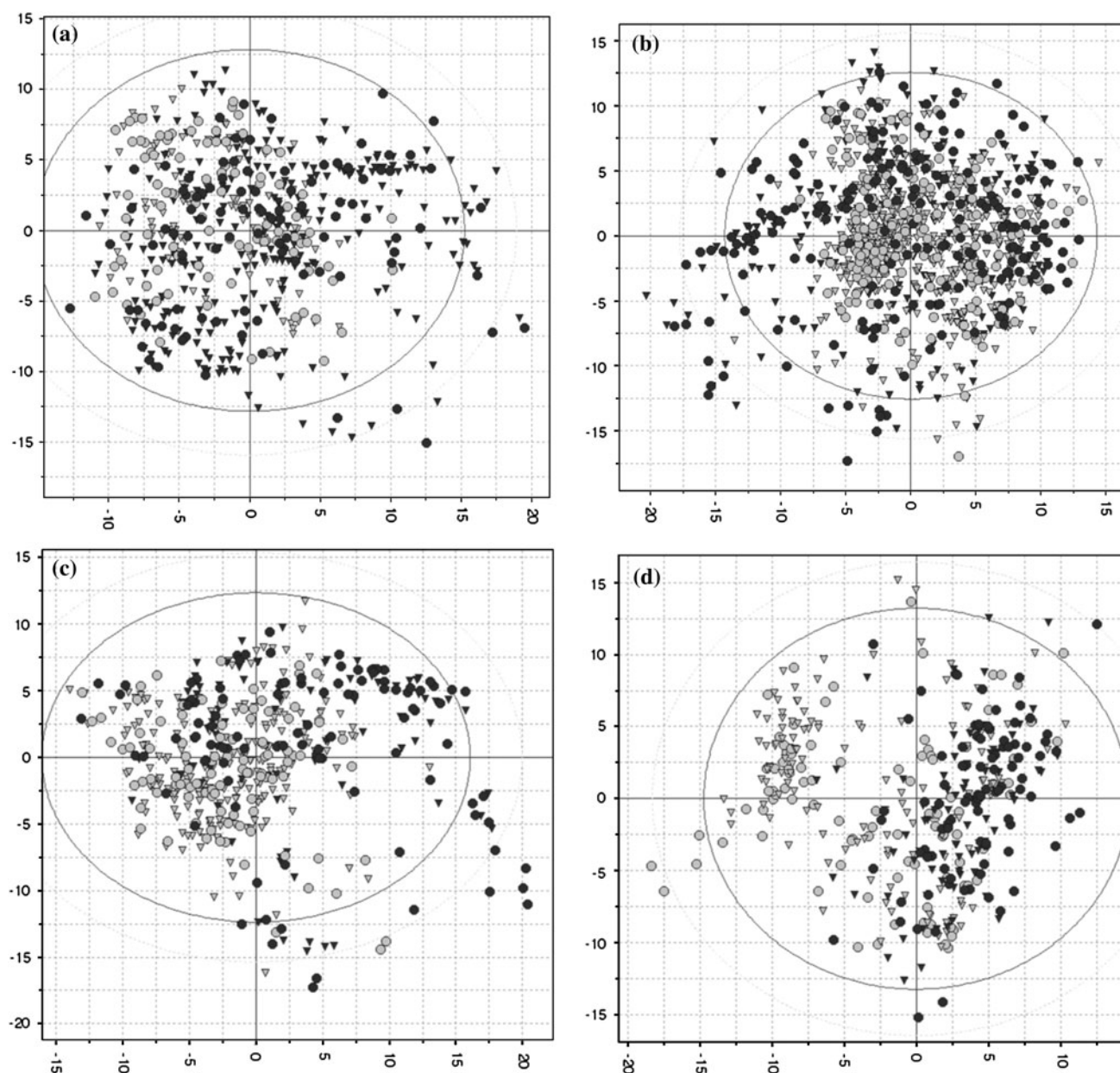
**Fig. 1** Top AR antagonist datasets used for FLAP modelling and for comparison with ROCS and ECFP4 performance; **a** A<sub>1</sub>R subtype collection; **b** A<sub>2A</sub>R subtype collection; **c** A<sub>2B</sub>R subtype collection; **d** A<sub>3</sub>R subtype collection. Bottom: Bar plot distribution of the different ARs collections splitted by the MDC algorithm



whereas they were retained when the bioactive stereoisomer is explicitly known from literature. Tautomers and protomers were generated using MoKa [30]; forms with <1 % predicted abundance were discarded.

For dataset generation, each AR subtype collection was initially characterized using VolSurf + descriptors [31, 32] and the multidimensional space generated was subsequently reduced applying a Principal Component Analysis (PCA). Selected compounds evenly cover the multivariate space (Fig. 2) and span an affinity range of roughly six log units with binding affinities ( $pK_i$ ) to human A<sub>1</sub>, A<sub>2A</sub> and A<sub>2B</sub> receptors ranging from 4.00 to 10.20. Compounds with  $pK_i$  values above 7.5 were defined as “actives” and compounds with  $pK_i$  values below 5.5 as “inactives”. Both groups of ligands formed our final datasets. Compounds with  $pK_i$  values from 7.5 to 5.5 were considered as borderline inactives and discarded from the original collection. Slightly different activity thresholds of  $pK_i > 7.0$  and  $pK_i < 5.0$  for actives and inactives, respectively, were

defined for constructing A<sub>3</sub> training and test sets. This was needed due to a lower amount of A<sub>3</sub> antagonists collected from ChEMBL. We built collections of structurally diverse compounds for the different ARs subtypes. In particular, training and test sets for each AR subtype were constructed by focussing on an even distribution of (1) chemical properties and (2) number of actives and inactives. For this purpose, each final dataset was split into training and a test set using the MDC (Most Descriptive Compounds) algorithm integrated in VolSurf+. The MDC algorithm involves a selection scheme that weighs the compounds according to their population density; for details see Hudson et al. [33]. This approach helped us to collect optimized training sets by selecting the most representative compounds according to their chemical properties; it also avoided an otherwise biased discrimination based only on ligand size and shape. Details of training and test set composition are given in Figs. 1 and 2 and in Supporting Table S1.



**Fig. 2** Score plot of Principal Component Analysis (PCA) calculated for all four AR subtypes. 128 pharmacokinetic descriptors were calculated using Volsurf+. The respective loading plots are available in Supporting Table S1. Active (in grey) and inactive molecules (in black) were divided in training set (circles) and test set (triangles).

Note that the use of the MDC algorithm ensures a homogeneous distribution of both the training and test sets in the PCA space. **a** A<sub>1</sub> PCA score plot; **b** A<sub>2A</sub> PCA score plot; **c** A<sub>2B</sub> PCA score plot; **d** A<sub>3</sub> PCA score plot

### FLAP (Fingerprint of ligands and proteins)

The program FLAP [22, 23] is an algorithm designed primarily for virtual screening, 3D QSAR, as well as pocket comparison via three-dimensional superposition. FLAP relies on the Molecular Interaction Fields (MIFs) generated through the GRID force-field [34] in order to evaluate the type, strength and direction of the interactions that a molecule can establish. The GRID probes H, DRY, O, and N1

are used for FLAP modeling. The hydrogen probe H is used to compute the shape of a small ligand or a cavity of a protein. The hydrophobic probe DRY finds places at which hydrophobic atoms on the surface of a target molecule will make favorable interactions with hydrophobic ligand atoms. The probe O represents the oxygen of a carbonyl group and describes hydrogen bond acceptor interactions with hydrogen bond donors from the target, and the probe N1 represents a nitrogen amide NH group and describes



hydrogen bond donor interactions with hydrogen bond acceptors from the target. It should be noted that the use of other probes (e.g., the hydrophobic C1 = probe instead of or combined with the aromatic/hydrophobic DRY probe) would give other four-point pharmacophores. The pragmatic choice for these four specific probes is however justified by the fact that they represent distinct ligand–protein interaction features, but keep the amount of variables (similarity scores for each probe and their combinations) manageable. Furthermore, the hydrogen-bond acceptor and hydrogen bond donor capacities of the target are described by the amide N1 probe (which are similar to N1=, NH=, N2, N2=, N3+, N2+ and O1 GRID probes) and carbonyl O probe (which are similar to O–, O:: and O = probes), respectively.

FLAP takes into account molecular flexibility, stereochemistry, protonation state, tautomerism, and all the other important factors in the molecular design of chemical compounds. FLAP can produce varied forms of different output types depending on the user requirements, and includes an optional training process which greatly facilitates the production of optimal models. Conceptually, the FLAP algorithm can be divided into three parts:

*In the first part*, all molecules of interest are described by the GRID force-field, and the resulting MIFs are reduced in complexity by extracting an adequate amount of hotspots which is proportional to the strength of interaction of the MIF considered. This mathematical description of the molecules is generated and saved for both the templates and for the ligand molecules.

*In the second part* of the algorithm a molecular template/s is/are chosen according to the best separation obtainable between actives and inactives. Then, the actual superposition of the ligand molecules onto the templates is performed. The superposition search process is a stochastic process based on quadruplets of points.

*The third part* concerns the evaluation of MIF volume superpositions, which is reported as similarity between the different structures on the basis of individual probes or the scores of probe combinations such as “Glob-P”; this is a global score produced by multiplying all individual Tanimoto similarity probe scores computed. In this sense, Glob-P represents a sort of “global similarity” of a single ligand screened over the template. Thus, for each candidate considered, similarity increases with the Glob-P value.

FLAP can either be employed in a ligand-based (the template is a small to medium-sized molecule), in a structure-based (the template is a large cavity of the protein molecule), or in a pharmacophore-based approach for virtual screening. In addition, FLAP represents a useful tool in Protein–Protein comparison. Successful applications have recently been published [23, 35–37]. Details about the

application pharmacophore (FLAPPharm) and structure-based frameworks are provided below.

The FLAPPharm tool inside the FLAP algorithm serves to generate pharmacophoric models. The method is described in detail somewhere else [38], and was extensively benchmarked [39] on a dataset of 81 X-ray derived alignment sets. As with all pharmacophore elucidation approaches, the assumption is that the input molecules are those active at a specific receptor. The first step of FLAPPharm consists of generating the appropriate conformers; for each ligand 1000s of conformers are generated, which are then filtered according to their energy, diversity, and pharmacophoric similarity to the other ligand conformers. Each conformer is then treated as a template, before a pruned tree-search is performed to align the other molecules. In this way, the method is template independent, since all ligand conformers are tested as a template. The alignment sets (models) that are generated are ranked and saved; the individual alignments and model scores come from a weighted sum of the hydrogen bond donor, hydrogen bond acceptor, hydrophobic, and shape similarities. From each model, a pharmacophoric pseudo-molecule is built. This ‘pharmacophore’ consists of common points where the ligands have common coincident atoms of the various types, and in place of the MIFs and pseudoMIFs are the PIFs (Pharmacophoric Interaction Fields) and pseudoPIFs. These PIFs are the mean average MIFs across the aligned ligands; fields regions common to all ligands will retain the same intensity, whereas regions common to 50 % of the ligands will be 50 % of the intensity and so on. In this way partial matching is inherently described by the fields without any user bias (for example the need to specify how many molecules must match a feature). The pharmacophoric pseudomolecule therefore retains the key features across each of the aligned ligands, and it can be used in the FLAP framework: for example as a template in virtual screening, or to dock it into a receptor to help validate or disprove the model. An additional advantage is that when screening, molecules’ similarities to the pharmacophore are provided, and therefore molecules can be ranked according to this continuous function as opposed to the typical pharmacophore constraint of whether a set of N features is matched or not.

FLAPPharm can provide a series of alignment models ranked by a global similarity score named S-Score. For each case, FLAPPharm tool was instructed to generate maximum five pharmacophoric hypotheses ranked by the S-Score. The speed accuracy was set to *medium* in order to minimize time speed calculation without decreasing alignment quality. In this case, FLAPPharm provides alignments by randomly sampling a lower amount of

quadruplets to overlap. For this study, the best S-Score pharmacophoric models were considered.

#### FLAPsite cavity definition

In the structure-based application of FLAP, the *FLAPsite* algorithm was applied for the identification of cavities in the 3D structures of the four ARs (see below for details on the generation of receptor models). The procedure starts by embedding the target protein into a grid with a spatial resolution of 1.0 Å. The detection routine through the GRID probe H identifies pocket points. The next step focuses on those grid points which are located within a distance of 4 Å from the closest protein atom, excluding protein surface points. For the remaining points the buriedness-index is calculated. This index gives an estimate of concavities and clefts in the protein surface.

To evaluate the embedding of a GRID probe, it is necessary to detect directions around the grid probe that are equally distributed. FLAP uses an octahedron to build an optimal distribution of vectors in the 3D-space. After a series of triangulations, it is possible to subdivide the eight faces of the octahedron, obtaining three additional vectors on each face. In this way we obtain a total of 30 vectors (6 primary vectors of octahedron plus 24 computed after triangulations) of length 10 Å and width 0.9 Å. The buriedness-index of a probe is increased by one whenever an atom of a protein is comprised within the search ray. GRID points with a buriedness-index lower than predetermined specific threshold are discarded. Remaining points are submitted to two basic morphological operations, erosion and dilation, which yields contrasting results. Erosion decreases object size and removes small anomalies by subtracting objects with a radius smaller than the structuring element. Dilation increases object size, filling in holes and broken areas, and connecting areas that are separated by spaces smaller than the size of the structuring element.

At this point primary pockets are detected. The following steps allow to eliminate too small pockets and to accentuate the degree of cavity curvature by using the hydrophobic probe DRY. When the physical geometry of two cavities is equal, the more hydrophobic one is prioritized (drugs usually bind hydrophobic sites).

#### ROCS shape-based similarity search

OpenEye ROCS (Rapid Overlay of Chemical Structures) [24] is a 3D-screening tool based on the concept that molecules with similar shape present an extensive volumetric overlay, whereas any volume mismatch contributes to shape dissimilarity. The method considers a set of ligands which are compared to one or more reference

structures. ROCS implements different measures of molecular similarity. OpenEye OMEGA v. 2.3.2 [40] was applied to generate consistent conformers to be overlapped by the ROCS method. In order to rank the screened compounds, we used the *ComboScore* scoring function in ROCS that includes the *ShapeTanimoto* score (which quantifies the steric overlap between two molecules), and the *ColorTanimoto* score (which represents the overlap of chemical functionalities calculated via the atomic charges). The *ShapeTanimoto* score is added to the *ColorTanimoto* score, resulting in the Tanimoto *ComboScore*. It can assume a value between 0 and 2.

#### Topological similarity search with ECFP4

Accelrys Scitegic Pipeline Pilot student version 2011 [41] was applied for the calculation of 2D Tanimoto based molecular similarity against the ChEMBL actives as references, using the topological ECFP4 fingerprint [25]. Extended connectivity fingerprints (ECFP) are circular fingerprints derived using a variant of the Morgan algorithm [42]. The procedure is very efficient and rapidly identifies topological properties of a chemical structure. In our study we used the Pipeline Pilot ECFP4 fingerprints to derive 2D Tanimoto similarities with template structures. The expressions “ClosestSimilarity” and “MaximumNumberOfClosest” were set for this study.

#### Receptor modelling and ligand docking

Homology model structures of each of the four human adenosine receptors were generated starting from the crystal structure of the inactive hA<sub>2A</sub>R, PDB code 3EML, [43] by means of homology modelling. The procedure was described recently in detail for the A<sub>2B</sub>R [44] and the A<sub>3</sub>R [20] models. Briefly, a family alignment was generated with ClustalX [45] and each template (hA<sub>2A</sub>R)–query (hA<sub>1</sub>R, hA<sub>2B</sub>R, hA<sub>3</sub>R) pairwise alignment was manually edited, if needed, and used as input for homology modelling with Modeller [46]. For each modelled subtype, 5 out of 15 initial models were selected on the basis of a double criteria: the stereochemical quality (reported by PROCHECK [47]) and the DOPEHR scoring function implemented in Modeller. These models were subjected to structure optimization and ranking according to the criteria in the Molprobit server [48] retaining a single model for each receptor. The loops of the selected structure were refined with the LoopModel routine [49] of Modeller, and the best model at this stage was selected following the same criteria stated above. Finally, the loop regions of each selected model were refined with partial energy minimizations with MacroModel [50]. These procedure yielded a 3D model for the hA<sub>1</sub>R, hA<sub>2B</sub>R and the hA<sub>3</sub>R receptors,

which together with the optimized crystal structure of A<sub>2A</sub>R (using the needed steps from the above protocol to model the missing loop regions) were used for the FLAP structure-based mode as described above, as well as for all docking studies depicted below.

Structure-based docking comparison was performed with the programs PLANTS (version 1.1) [26] and GOLD (version 5.0.1) [27]. PLANTS combines an ant colony optimization algorithm with an empirical scoring function for the prediction and scoring of binding poses in a protein structure [51]. For each compound, 10 poses were calculated and scored by the *Chemplp* scoring function at a speed setting of 1. The bindings pockets of ARs were defined by all residues within a 10 Å radius around the conserved asparagine in TM6 (Asn253<sup>6.55</sup>) [52], which is found to coordinate receptor binding to different ligands (PDB codes 3EML, 3QAK, 2YDO, 2YDV, 3PWH, 3REY and 3RFM [53]). All other options of PLANTS were left at their default setting.

GOLD docking software uses a genetic algorithm to obtain optimized docking poses. Docking was performed applying the GOLD fitness scoring function: Population size = 100; Selection pressure = 1.1; # Operations = 100,000; # islands = 3; niche size = 2; migration = 10; mutation = 95; crossover = 95. Same binding pocket detection parameters were set for GOLD docking.

## Results and discussion

The generation of optimized, i.e. highly diverse and evenly distributed, datasets strongly supports the derivation of valid computational models. In this study, we developed FLAP models for all AR subtypes. Such models have to be defined as “local”, since they were built by using focused datasets, according to the subtype considered. Their predictivity depends on the chemical diversity present in training and test sets used [54]. We focused our attention on the construction of structurally diverse collections for the four different AR subtypes. For this purpose, the ChEMBL compound collection was used for generating the datasets used in this study [28]. Accordingly, this is the first work on concomitant modelling of AR antagonists for all subtypes on the basis of maximally diverse and evenly distributed training and test sets.

In particular, training and test sets for each AR subtype were constructed by focussing onto these most relevant aspects: (1) the even distribution of chemical properties; (2) the even distribution of the number of actives and inactives to train FLAP models. For details see Figs. 1 and 2 “Methods” Section.

For ligand-, structure- and pharmacophore-based modelling, FLAP was instructed to generate all possible

stereoisomers for the dataset molecules, unless the bioactive form was reported in literature. During the modeling phase, FLAP automatically selects that tautomeric, protomeric, or stereo-isomeric form of a ligand that best fits the models. For each compound, FLAP was also instructed to generate at maximum 25 conformers with an RMSD distance of 0.3 Å, to avoid time consuming calculations during ligand overlap with the template structures and their respective MIFs.

The GRID probes H, DRY, O, and N1 were used for FLAP modeling of ligands and templates. From the options for template selection, available in FLAP, we applied Linear Discriminant Analysis (LDA) [55, 56]. In order to save computation time, we considered only the most active compounds ( $pK_i > 8.5$ ) out of the entire set as potential templates for the ligand-based mode. For the structure-based approach we considered as templates those cavities of each AR subtype that contain the key asparagine in TM6 (Asn<sup>6.55</sup>) involved in ligand binding in all receptor subtypes [8, 57]. Finally, pharmacophoric hypotheses were generated using literature-known selective antagonists to train the models [8].

### Ligand-based approach

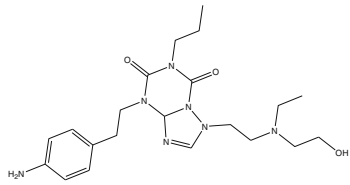
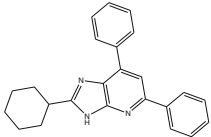
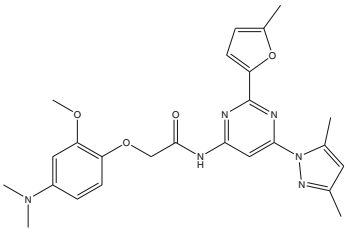
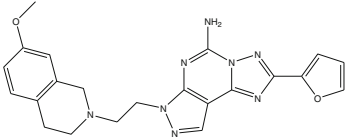
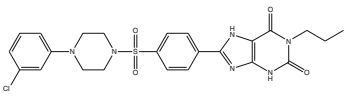
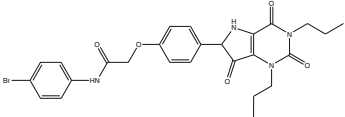
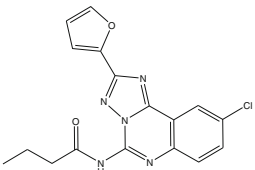
First step in the ligand-based FLAP mode is aligning AR dataset molecules onto representative templates, which are automatically selected by LDA in order to obtain the best discrimination between active and inactive antagonists. The choice to select two templates in this study considered the wide chemical diversity and size variation within the respective AR datasets and also represented the best compromise between speed of calculation and screening accuracy. Corresponding templates, their scaffolds and  $pK_i$  values for binding to human AR subtypes are given in Table 2.

Then, enrichment curves were used to compare the performance of the individual ligand-based FLAP models, demonstrating in general very good results in terms of retrieval and ranking of actives (Fig. 3, Table 3A).

For evaluating the number of false predictions, an LDA-regression (LDA-R) score was applied. This is a continuous score calculated by FLAP using the LDA tool, in order to estimate the possible classification of a ligand as active or inactive. The enrichment curves resulting from scoring of the training and test sets are reported in Fig. 3, ranked according to the LDA-R score. In addition, the figure shows enrichment curves (True positive rate vs. False positive rate) and the final AUC (Area Under the Curve). AUC is a measure of the successful discrimination between known actives and inactives by a given template.

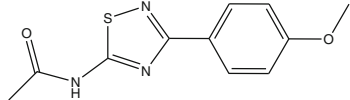
The A<sub>2A</sub> LDA model performed best among all ligand-based models, both for training and test set ranking. The

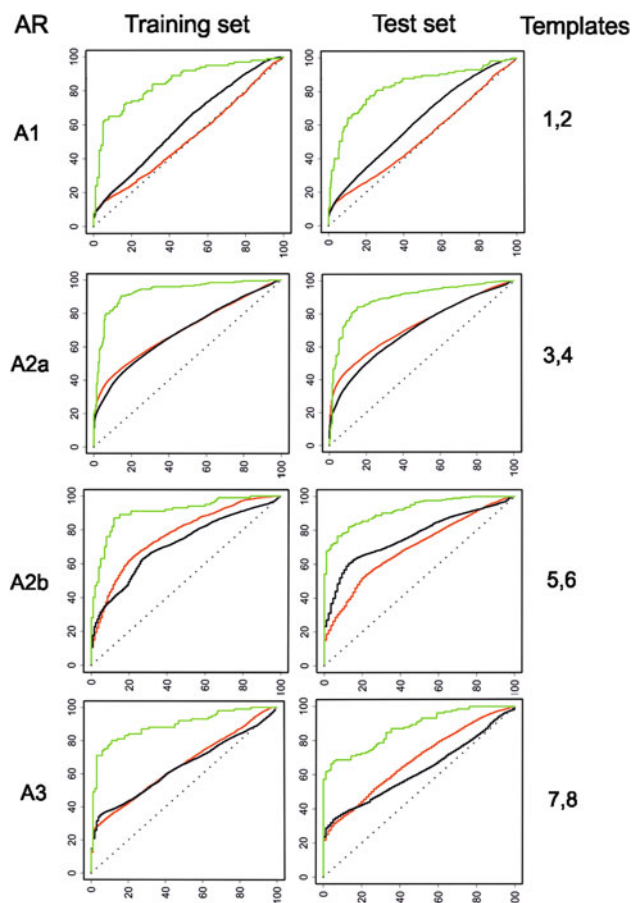
**Table 2** Templates used for ligand-based FLAP modeling

Template	2D Structure	Scaffold	pKi	Reference
<i>A<sub>1</sub>receptor</i>				
1 (ChEMBL591235)		Xanthine	9.24	[83]
2 (ChEMBL219160)		Amino-pyrimidine	9.05	[84]
<i>A<sub>2A</sub>receptor</i>				
3 (ChEMBL429125)		N-pyrimidinyl-2-phenoxyacetamide	9.70	[85]
4 (ChEMBL469674)		Pyrazolo-triazolo-pyrimidine	8.40	[75]
<i>A<sub>2B</sub>receptor</i>				
5 (ChEMBL483469)		Xanthine-8-yl-benzenesulfonamide	9.11	[86]
6 (ChEMBL456125)		1,3-dialkyl-deazaxanthine	9.00	[87]
<i>A<sub>3</sub>receptor</i>				
7 (ChEMBL319971)		Triazolo-quinazoline	7.84	[88]



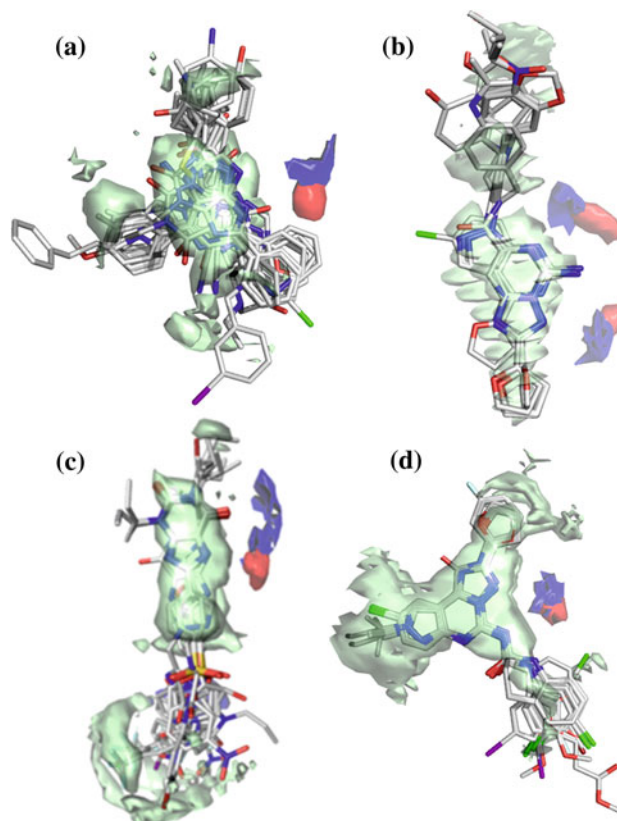
**Table 2** continued

Template	2D Structure	Scaffold	pKi	Reference
8 (040)		Thiadiazole	9.10	[89]

**Fig. 3** Enrichment curves of ligand-based modelling. False positive rate (x-axis) versus True positive rate (y-axis). Enrichment curves for FLAP models are given in *green*, for ECFP4 topological fingerprint models in *red* and for ROCS shape comparison models in *black*

low amount of false positives found in the top 20 % of the screened database and among the 100 top ranked compounds well demonstrates the LDA capability to efficiently discriminate actives from inactives of the training and the test sets by using only two templates. Similar good performance was found for the  $A_{2B}$  LDA model, especially, in the top 20 % of the screened database, while the amount of false positives for the training set in the 100 top ranked compounds increased.

The  $A_1$  and  $A_3$  models still maintain good results for the top 20 % of the screened database, whereas in the 100 top

**Fig. 4** Cumulative MIF analysis for the 20 top ranked actives on the respective templates; **a**  $A_1R$ , **b**  $A_{2A}R$ , **c**  $A_{2B}R$ , **d**  $A_3R$ . *Green*, *red* and *blue* fields respectively represent hydrophobic, donor and acceptor most relevant GRID interaction regions at  $-0.5$ ,  $-4.5$  and  $-5.5$  kcal/mol energy level

ranked the amount of false positive increased; however classification remained satisfactory. Encouragingly, the amount of false positives retrieved for the  $A_{2A}$ ,  $A_{2B}$  and  $A_3$  test sets is lower than for the respective training sets.

Beyond the derivation of enrichment curves, we applied cumulative MIF analysis in the ligand-based FLAP mode in order to identify the most relevant Molecular Interaction Fields (MIFs) shared by a set of superimposed molecules. This analysis was performed by aligning the 20 top ranked actives onto the optimally discriminative templates for each respective AR subtype (Fig. 4).

**Table 3** Number of FPs among top 20 % and 100 top-ranked compounds and AUC values

(A) Ligand-based approach					
A <sub>1</sub> R	Training set	Test set	A <sub>2A</sub> R	Training set	Test set
Top 20 %	3	17	Top 20 %	1	1
TOP-100	19	30	TOP-100	5	3
AUC	0.85	0.83	AUC	0.93	0.83
A <sub>2B</sub> R	Training set	Test set	A <sub>3</sub> R	Training set	Test set
Top 20 %	0	0	Top 20 %	1	0
TOP-100	13	1	TOP-100	18	10
AUC	0.91	0.93	AUC	0.89	0.87
(B) Structure-based approach					
A <sub>1</sub> R	Training set	Test set	A <sub>2A</sub> R	Training set	Test set
Top 20 %	8	2	Top 20 %	2	3
TOP-100	29	38	TOP-100	21	18
AUC	0.77	0.81	AUC	0.77	0.78
A <sub>2B</sub> R	Training set	Test set	A <sub>3</sub> R	Training set	Test set
Top 20 %	3	1	Top 20 %	2	0
TOP-100	22	11	TOP-100	25	17
AUC	0.82	0.82	AUC	0.79	0.76
(C) Pharmacophore-based approach					
A <sub>1</sub> R	Training set	Test set	A <sub>2A</sub> R	Training set	Test set
Top 20 %	2	34	Top 20 %	14	3
TOP-100	28	51	TOP-100	21	1
AUC	0.78	0.75	AUC	0.72	0.72
A <sub>2B</sub> R	Training set	Test set	A <sub>3</sub> R	Training set	Test set
Top 20 %	5	9	Top 20 %	3	3
TOP-100	28	3	TOP-100	31	22
AUC	0.81	0.80	AUC	0.76	0.69

Numbers of FPs and AUC values for training and test sets in AR subtype models, when using the ligand-based (**A**), structure-based (**B**), or pharmacophore-based mode (**C**)

In the case of the A<sub>1</sub>R, A<sub>2B</sub>R, and A<sub>3</sub>R datasets, cumulative MIF analysis revealed one donor and one acceptor region, located in the central core of the aligned ligands (Fig. 4). Such features are supposed to bind the conserved Asn<sup>6.55</sup>. This well agrees with other ligand-based hypotheses [8, 20, 58–62]. For the A<sub>2A</sub>R dataset, cumulative MIF analysis shows two acceptor and one donor regions for template 3 and two acceptor and two donor regions for the alignment onto template 4. The different regions detected by the two templates are due to different scaffolds of the overlapped ligands, which in turn generate different arrangements of MIFs. In fact, most of the A<sub>2A</sub>R antagonists used exhibit either a pyrazole-

triazole-pyrimidine or a pyrimidine scaffold (Fig. 4). In addition, MIFs well match with the interaction of ZM241385 with A<sub>2A</sub>R as revealed in the crystal structure [43].

FLAP screening results were compared with the results obtained using two different similarity search methods frequently applied in virtual screening: the 2D-based ECFP4 [25] method and the 3D-based ROCS method [24]. Such analyses were applied to each AR antagonist dataset used in the FLAP training phase. Starting from a similarity matrix, the resulting enrichment curve was calculated by averaging all single similarity values against each Ref. [63]. FLAP ranking was evaluated using the LDA-R

descriptor; this continuous score represents the probability of a molecule to be predicted as active (positive values) or inactive (negative values). Alternatively, the Tanimoto score served to assess the ECFP4 similarity ranking between reference and dataset structures. The *ComboScore* was employed within ROCS as metric for similarity (see “Methods” section for more details).

Figure 3 shows a superior performance of FLAP LDA models in discriminating active from inactive antagonists for all AR subtypes. However, especially focusing on the earlier part of the  $A_{2A}$ ,  $A_{2B}$  and  $A_3$  enrichment curves, ECFP4 and ROCS still retain a satisfactory retrieval of actives. Both ECFP4 and ROCS methods displayed very similar profiles of enrichment curves.

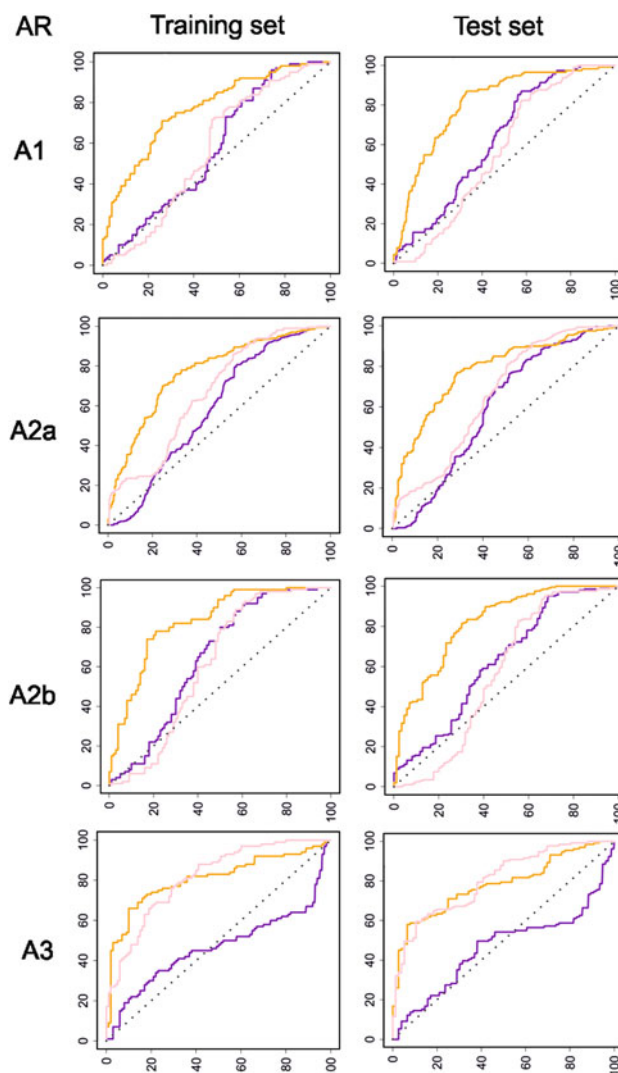
### Structure-based approach

In the structure-based mode, FLAP modelling considers protein cavities as template candidates using the *FLAPsite* tool (see “Methods” section). One of the crystal structure of  $A_{2A}R$  was used to represent this receptor, [43] while for the remaining members of the ARs homology-based models were created. *FLAPsite* detected 3, 3, 7, and 5 cavities for  $A_1$ ,  $A_{2A}$ ,  $A_{2B}$ , and  $A_3$  receptors, respectively. Afterwards we selected that cavity containing the family-conserved Asn<sup>6.55</sup>.

The selected cavity of each AR subtype was then used to screen the corresponding dataset of antagonists via the match between the MIFs of cavity and ligands. For each AR dataset, separate FLAP LDA models were built. Numbers of false positives are listed in Table 3B; enrichment plots are given in Fig. 5.

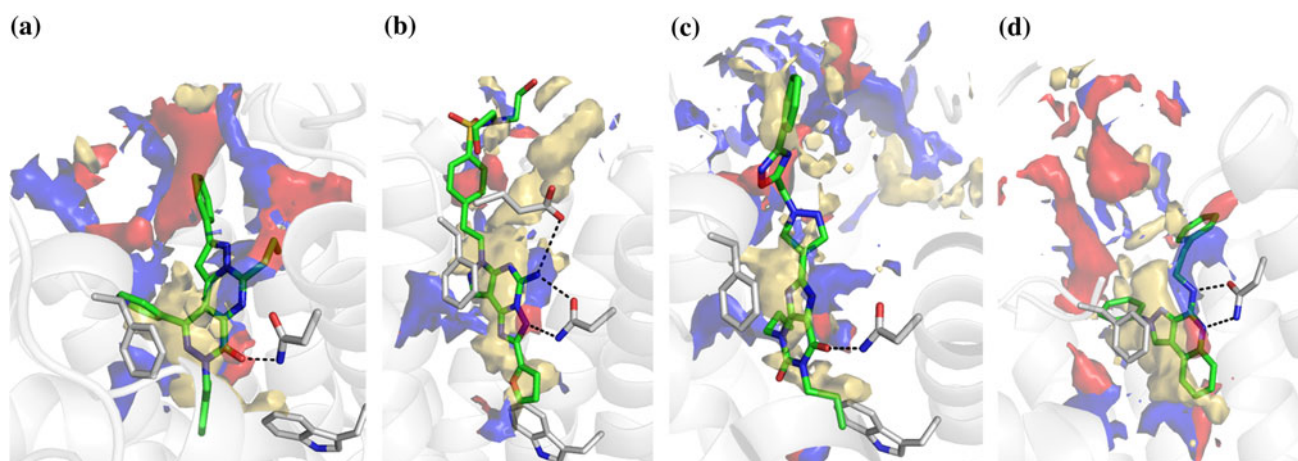
Examples of FLAP pose solutions for the highest LDA-R score are shown in Fig. 6. In the case of the  $A_{2A}R$ , validation of the FLAP model was performed in two different ways. In the first case, the pose of compound ZM241385 into the  $A_{2A}$  crystal structure was compared with the FLAP pose (Fig. S2). The FLAP-derived orientation closely resembles the crystallographic one, and represents a very good result because the FLAP procedure does not take into account any minimization process or energetic factors, but only overlapping in terms of quadruplets derived by the GRID MIFs. The RMSD value calculated for the crystallographic conformation and the FLAP pose is 1.83 Å. In both cases, the two H-bonds with Asn253<sup>6.55</sup> as well as the hydrophobic interactions with the many hydrophobic residues in pocket 1 were found. The second validation consisted in screening the  $A_{2A}$  training (200 actives; 200 inactives) and test sets (467 actives; 193 inactives). Virtual screening enrichment curves and results are shown in Table 3B and Fig. 5.

For all four AR subtype models, the number of FPs among the top 20 % of ranked compounds is remarkably



**Fig. 5** Enrichment curves of structure-based modelling. False positive rate (x-axis) versus True positive rate (y-axis). Enrichment curves for FLAP models are given in orange, for PLANTS “Chemplp” docking in purple, and for GOLD “GoldScore” docking in pink

low (close to zero), with the only exception in the  $A_1R$ -training set model, with 8 FPs. Among the 100 top-ranked compounds, the numbers of FPs (around 20) is also comparable among each receptor-model. Here again, the  $A_1R$  model exhibits a somewhat reduced performance with 29 FPs in the training and 38 FPs in the test sets. AUC values range close to 0.8 with slightly higher (better) values for the  $A_{2B}R$  model. Profiles of enrichment curves (Fig. 5) indicate best performance of the  $A_3R$  model and almost identical performance of the remaining three models. This is particularly true for the early first part of the enrichment ranking. As already observed for the ligand-based mode, the amount of false positive prediction for the  $A_{2A}R$ ,  $A_{2B}R$  and  $A_3$  test sets is surprisingly lower than for the respective training set (Table 3B).



**Fig. 6** Structure-based FLAP docking pose for the top ranked solution according to the LDA-R score. **a** CHEMBL1277925 pose into  $A_1R$  model, **b** CHEMBL22717 pose into  $A_{2A}R$  structure, **c** CHEMBL200779 pose into  $A_3R$ , **d** APQ029 into  $A_3R$  model. Gold

color, red and blue fields respectively represent hydrophobic, donor and acceptor GRID interaction regions at  $-0.5$ ,  $-4.5$  and  $-5.5$  kcal/mol energy level

The performance of structure-based FLAP models was compared with the outcome of two docking programs, PLANTS and GOLD; the corresponding enrichment curves are displayed in Fig. 5. The standard scoring function within each program was applied to rank the different datasets: *Chemplp* [51] for PLANTS and *GoldScore* for GOLD docking [27]. Notably, the FLAP LDA-R score shows a clear superior ranking as compared to PLANTS/*Chemplp* in all cases, while as compared to *GoldScore* this also holds for  $A_1R$ ,  $A_{2A}R$  and  $A_{2B}R$ . The only case when a docking algorithm retrieves a similar rate of actives is the GOLD program when applied to the  $A_3R$  subtype, both for the training and test sets.

Findings with the structure-based FLAP mode, reported above, nicely match with mutagenesis data found in literature. This holds for all four AR subtypes studied here.

$A_1R$  [10, 64]: Mutagenesis on Asn254<sup>6.55</sup> causes the complete loss of agonist and antagonist binding. The mutation Trp247<sup>6.48</sup> diminishes antagonist binding; accordingly FLAP models confirm the impact of Trp247<sup>6.48</sup> for receptor interaction.

$A_{2A}R$  [5, 10, 57, 64–66]: Alanine mutations of Asn253<sup>6.55</sup>, His250<sup>6.52</sup>, Glu169, Ile274<sup>7.39</sup>, Phe168<sup>EL2</sup> and Trp247<sup>6.48</sup> generate a complete loss of antagonist binding; all these residues generate MIFs which seem to be involved in most of the FLAP antagonist poses.

$A_{2B}R$  [10, 67, 68]: Alanine mutations of Asn254<sup>6.55</sup>, His251<sup>6.52</sup>, Trp247<sup>6.48</sup> and Phe187<sup>5.43</sup> abolish antagonist binding which well agrees with the MIFs found for the most part of FLAP poses.

$A_3R$  [10, 69]: Alanine mutation of Asn250<sup>6.55</sup> abolishes  $A_3R$  agonist and antagonist binding while mutation of His272<sup>7.43</sup> by Glu and Trp243<sup>6.48</sup> by either Ala or Phe

diminishes  $A_3R$  antagonist activity as corresponds to FLAP findings.

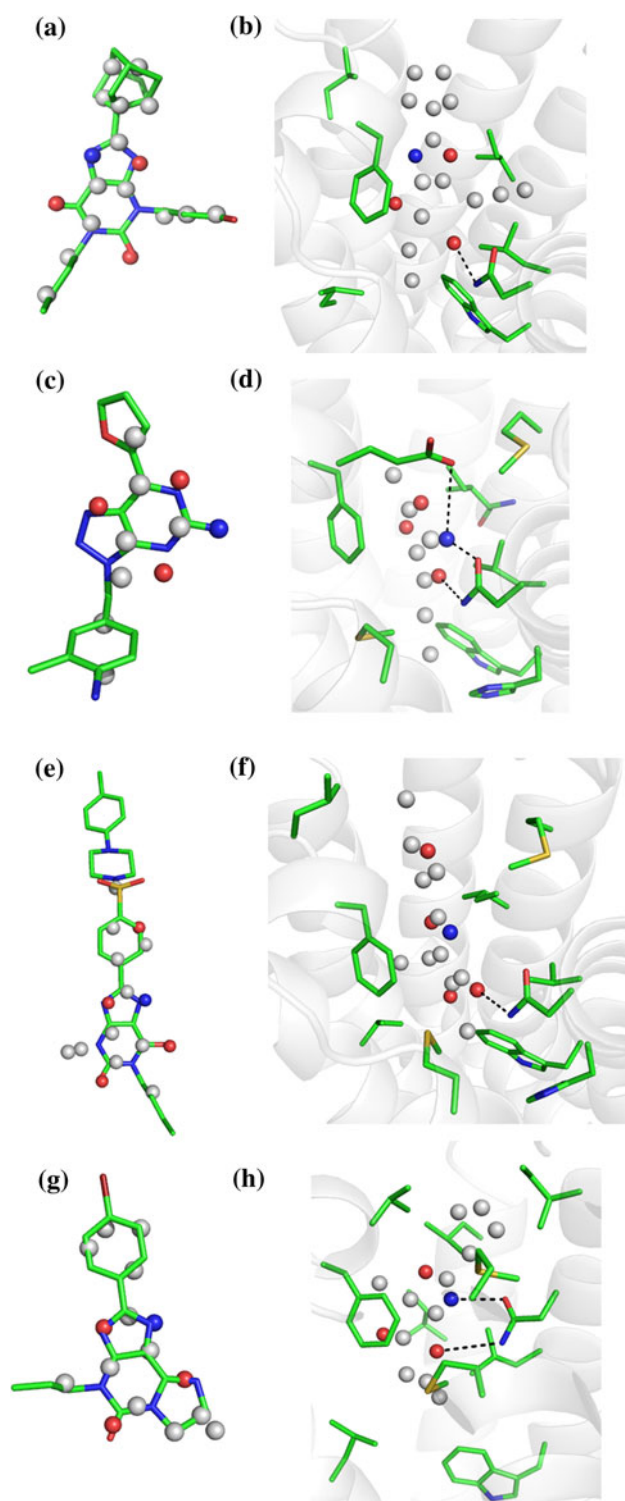
#### Pharmacophore-based approach

In this work, we applied and validated for the first time a novel procedure for generating pharmacophore hypotheses by alignment of a limited set of active molecules, called FLAPPharm (see “Methods” section). This tool uses such a set of actives to define their optimal 3D-superimposition and to derive a common 3D-pharmacophore. This entirely ligand-based method first generates 30 conformers for each selected molecule, within 20 kcal/mol of the minimum energy structure. GRID MIFs are calculated for the probes H, N1, DRY and O. Then all conformers of all selected actives are superimposed in order to find the best alignment model with the highest FLAP similarity score.

In this study, FLAP pharmacophore models were built for each AR subtype (Fig. 7). A collection of affinity data for AR antagonists with documented selectivity for one out of the four subtypes is given by Fredholm et al. [5]. Their chemical structures are listed in Fig. 8. Unfortunately, the amount of binding affinity data ( $IC_{50}$ ,  $K_i$ ) for selective ARs antagonists annotated in literature is quite low. In fact, the most part of the studies only report the % inhibition measurements at a certain concentration for all the ARs.

The Fig. 7 (panels a, c, e, g) shows the four different pharmacophoric models and the matching of the pharmacophoric points overlapped onto the most selective antagonist of the respective subtype. All the overlapped selective ligands perfectly fit almost all pharmacophoric points. FLAP can also dock the pharmacophore into the protein by the match of the pharmacophoric points and the GRID





**Fig. 7** FLAP ligand-based pharmacophoric models: **a** A<sub>1</sub>; **c** A<sub>2A</sub>; **e** A<sub>2B</sub>; **g** A<sub>3</sub>, coupled with the structure-based docking of the respective AR subtype pharmacophores: **b** A<sub>1</sub>; **d** A<sub>2A</sub>; **f** A<sub>2B</sub>; **h** A<sub>3</sub>. For clarity, in box **a**, **c**, **e** and **g** the best pharmacophoric hypotheses are shown together with the most selective antagonist of the respective AR subtypes

fields calculated for the cavity. The best poses of the pharmacophores into the respective AR models, according to their Glob-P similarity score, are reported in Fig. 7 (panels b, d, f, h). Each docked pharmacophore displays favourable interactions with the conserved Asn<sup>6.55</sup> residue that is essential for ligand binding.

The pharmacophoric model for the A<sub>1</sub>R was built by aligning eight selective A<sub>1</sub> antagonists on the basis of their GRID interaction features. One donor and three acceptor points were detected; they correspond to the xanthine scaffold frequently found in A<sub>1</sub> antagonists. One of the acceptor points binds the conserved residue Asn<sup>6.55</sup> present in the A<sub>1</sub>R binding site (Fig. 7b), as also confirmed by Dore et al. [70]. FLAPPharm alignment indicated three important hydrophobic regions around the central core supporting high affinity A<sub>1</sub>R binding.

Nine selective A<sub>2A</sub>R antagonists served to generate the respective pharmacophoric hypotheses. The best S-Score model, shown in Fig. 7d, comprises three acceptor and one donor regions together with an elongated distribution of hydrophobic points. One acceptor and one donor point are predicted to bind the conserved residue Asn<sup>6.55</sup> (Fig. 7d). Beyond binding to this residue, FLAPPharm docking predicts an extra interaction between Glu169 (EL2) and the donor point. This interaction well agrees with the crystallographic binding mode of ZM241385 in the receptor cavity [43].

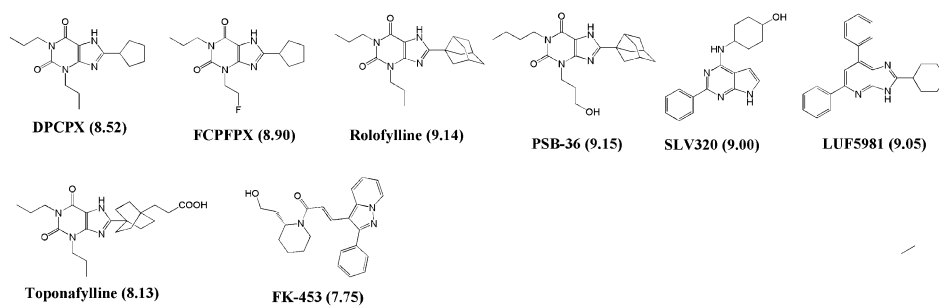
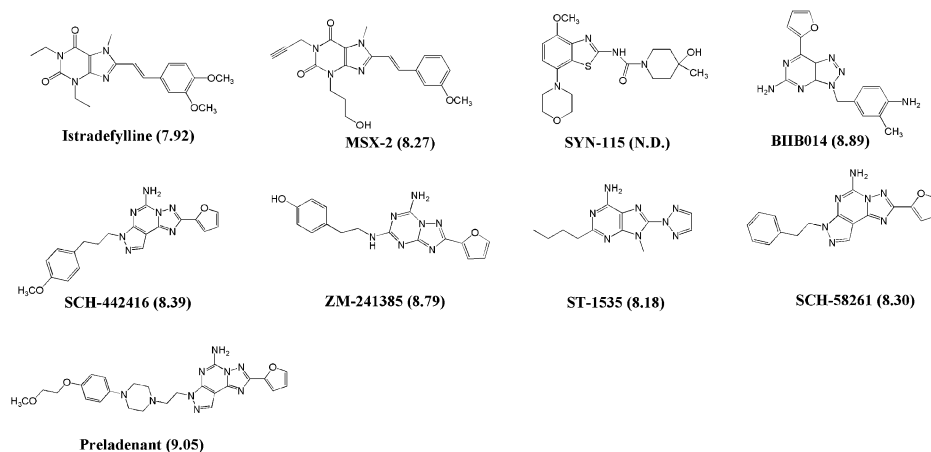
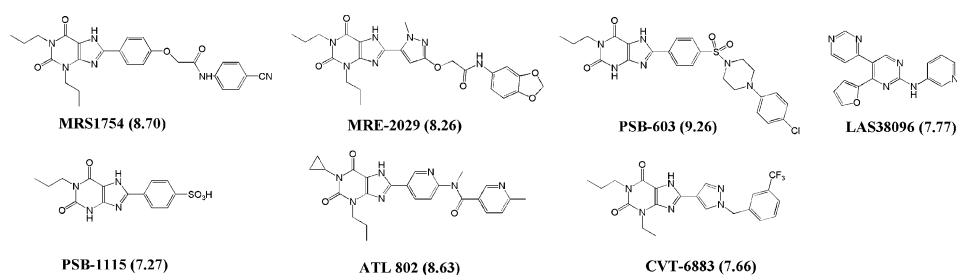
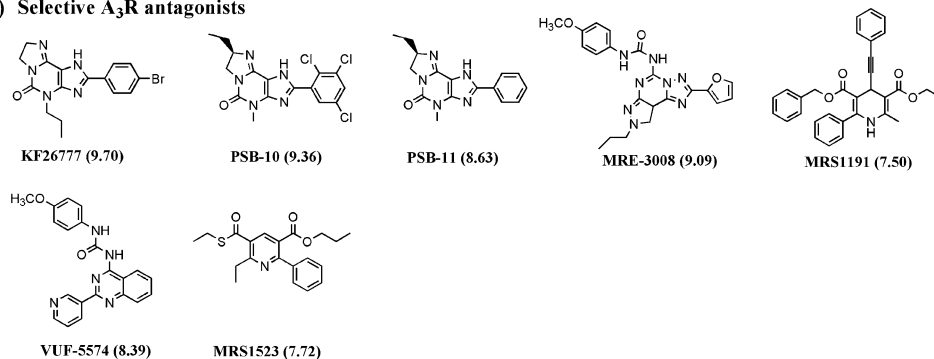
Seven selective antagonists were available for generating the A<sub>2B</sub>R pharmacophore. The best S-Score model displayed one donor and four acceptor points (Fig. 7e). Docking of the pharmacophore into the A<sub>2B</sub>R cavity showed an interaction of one acceptor with the conserved residue Asn254<sup>6.55</sup> (Fig. 7f). The geometrical arrangement of hydrophobic points of this model is similar to the A<sub>2A</sub>R pharmacophore.

Three acceptor and one donor point were deduced from FLAPPharm alignment using seven A<sub>3</sub>R selective antagonists. A triangular disposition of hydrophobic points was found. Docking into the A<sub>3</sub>R cavity displayed interaction of one donor and one acceptor with the conserved Asn<sup>6.55</sup> residue.

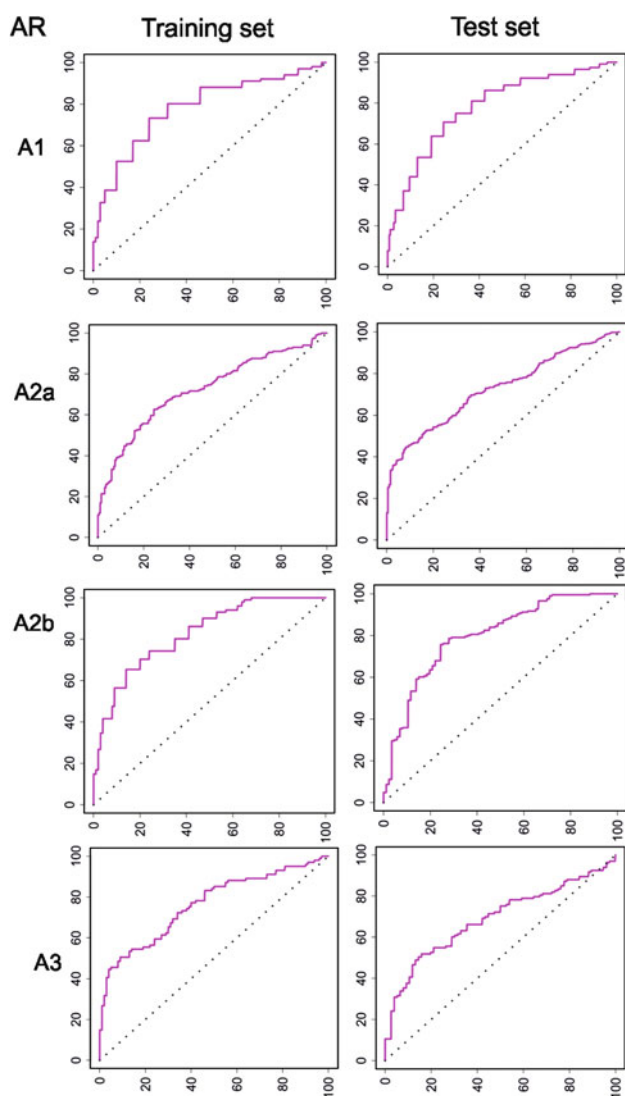
The similar geometrical arrangement of hydrophobic regions in A<sub>1</sub>R and A<sub>3</sub>R pharmacophores match their high sequence identity, thus suggesting a significant structural similarity in the binding cavity. The same conclusions can be drawn for the geometrical disposition of hydrophobic regions for A<sub>2A</sub>R and A<sub>2B</sub>R considering the similarity between the A<sub>2</sub> receptor subtypes [10].

It is worth mentioning that pharmacophoric points, derived here with the novel FLAPPharm tool, closely resemble previously reported hypotheses for the A<sub>1</sub>R



**(a) Selective A<sub>1</sub>R antagonists****(b) Selective A<sub>2A</sub>R antagonists****(c) Selective A<sub>2B</sub>R antagonists****(d) Selective A<sub>3</sub>R antagonists**

**Fig. 8** Subtype selective AR antagonists used to build FLAP pharmacophores [8]. Affinity pK<sub>i</sub> values are given in brackets



**Fig. 9** Enrichment curves of pharmacophore-based FLAP modelling. False positive rate (x-axis) versus True positive rate (y-axis)

[68, 69], A<sub>2A</sub>R [14, 17], A<sub>2B</sub>R [18], and A<sub>3</sub>R [15, 16, 19], thereby demonstrating the reliability of the new tool.

In order to check the impact and also to validate our pharmacophore hypotheses, we performed pharmacophore-based screening of the same training and test sets used above for ligand- and structure-based FLAP modeling. The pharmacophore models were able to correctly predict almost all active compounds of the respective training and test sets. The enrichment curves are shown in Fig. 9. All of the four enrichment curves show very good results in terms of the retrieval rate of actives (Table 3C).

In general, the amount of FPs among the top 20 % of ranked compounds is relatively low for all four AR subtypes; exceptions hold for the test set in pharmacophore-based FLAP modelling of A<sub>1</sub>R antagonists (34 FPs) and for the training set in FLAP modelling of the A<sub>2A</sub>R antagonists (14 FPs).

Among the 100 top-ranked compounds, the A<sub>2A</sub>R model showed the lowest amount of false positives both for the respective training (21 FPs) and test sets (1 FP). Predictions of FPs for A<sub>1</sub>R, A<sub>2B</sub>R and A<sub>3</sub>R training set compounds closely resemble with roughly 30. A very good performance was maintained for the classification of A<sub>2A</sub>R and A<sub>2B</sub>R test sets with 1 and 3 FPs, respectively. Satisfactory results were also obtained by the A<sub>3</sub>R pharmacophore-based FLAP modelling of A<sub>3</sub>R training set (31 FPs). The highest amount of wrongly predicted compounds was found for FLAP modelling of the A<sub>1</sub>R test set with 51 FPs.

#### Capability of FLAP in retrieving subtype selective antagonists

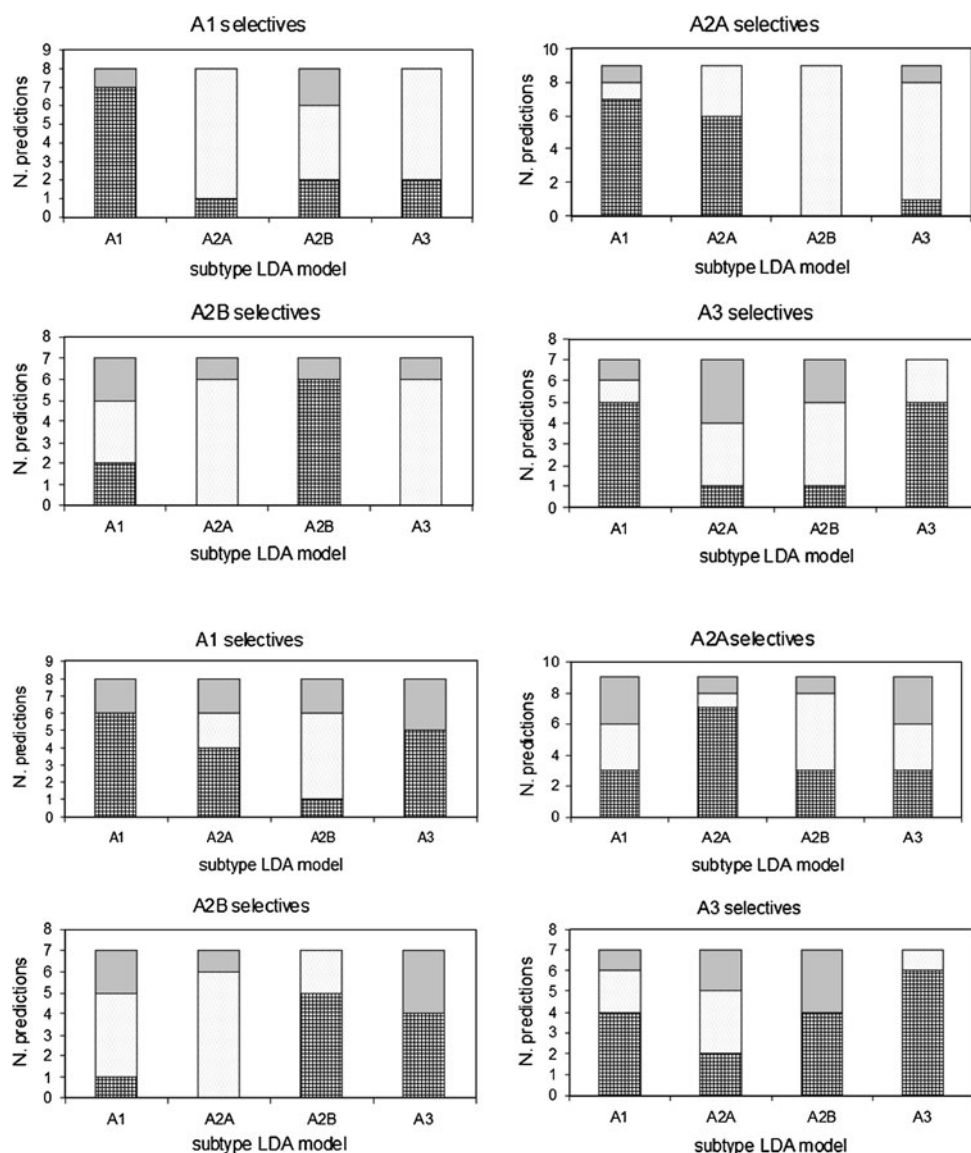
Selectivity aspects are crucial in medicinal chemistry and drug discovery. Drug discovery and design projects are focused on selective compounds for specific targets, in order to minimize side effects [11]. Optimizing subtype selectivity is particularly challenging since receptor subtypes often share high structural homology [10]. This implies similarity in cavity structure and in ligand–protein interactions common for the different subtypes.

Examples of subtype-selective scaffolds for ARs are reported in literature. 8-aryl or 8-cycloalkyl substituted xanthine and pyrazole-triazole-pyrimidines guide to selective A<sub>1</sub>R antagonists [71], whereas the xanthine scaffold substituted at N<sup>1</sup>, N<sup>3</sup> and C<sup>8</sup> positions lead to selective A<sub>2B</sub>R antagonists [72–74]. Substitutions at N<sup>7</sup> are essential to enhance A<sub>2A</sub>R selectivity [75], while N<sup>5</sup> and N<sup>8</sup> substitutions improve A<sub>3</sub>R selectivity [76].

The low number of AR antagonists tested on all four ARs, prevented us to generate ligand- and structure-based discriminative models based on selectivity aspects. Instead, we challenged our AR subtype models by testing if they are capable to retrieve their own selective antagonists; despite they were originally not trained on selectivity aspects.

For this purpose, we used the same compilation of subtype selective AR antagonists applied above to generate pharmacophoric hypotheses [8]. Few coincidences with AR test set molecules were found, whereas none of them belongs to the respective training sets. Thus, the FLAP LDA models were not biased by the presence of those selective antagonists used for such selectivity challenge.

For each compound, we generated 25 conformers with an RMSD value > 0.3 Å. These compounds served for external prediction by projection onto each previously developed ligand- and structure-based FLAP LDA model. We considered LDA-R scores in order to predict if a molecule belongs to the class of actives or inactives. According to the AR model considered, different tolerance thresholds were defined, in order to maximize the statistical significance of classifications.



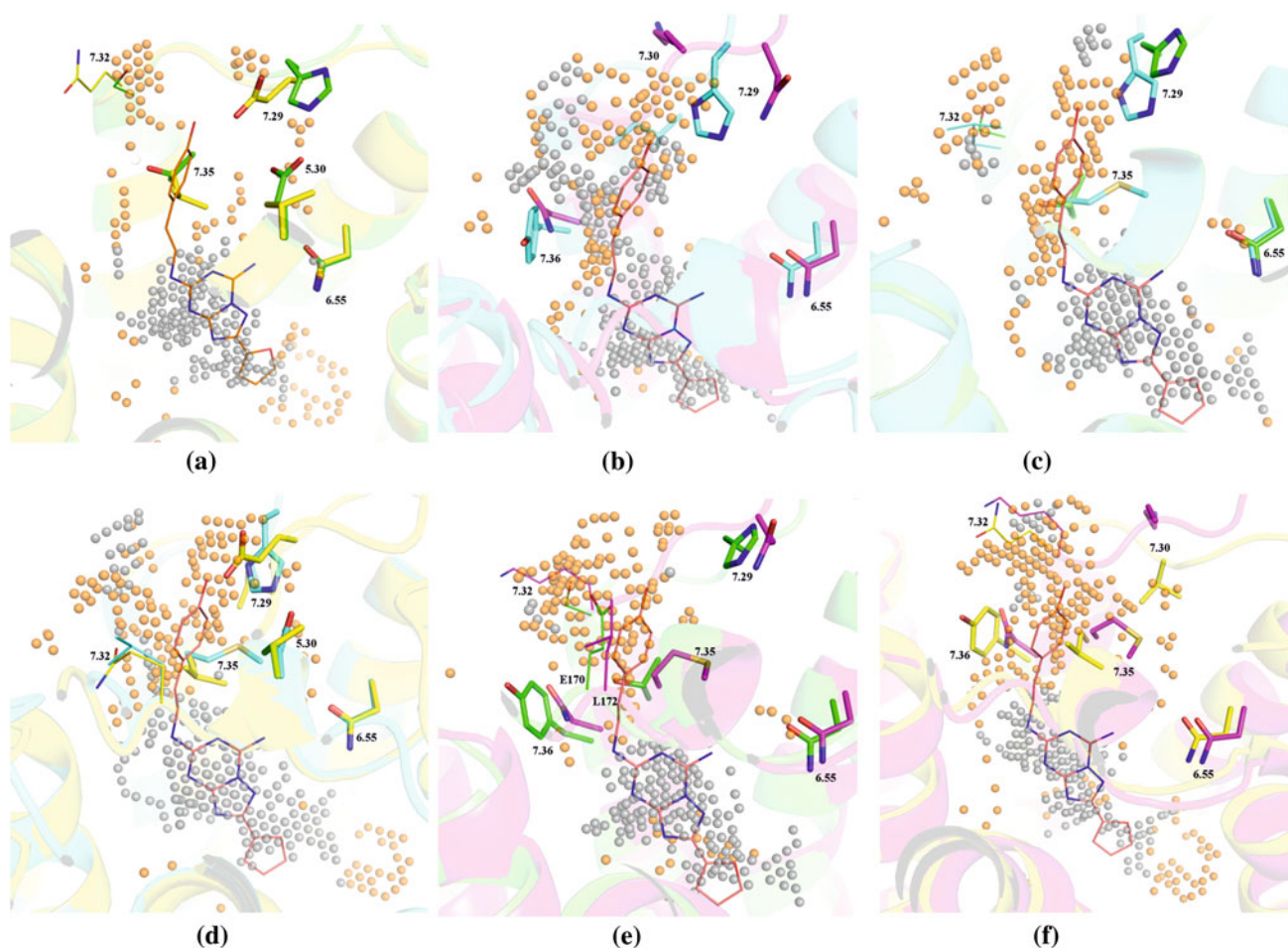
**Fig. 10** Selectivity validation for the respective AR subtypes *Top four bar plots* ligand-based LDA models, *Bottom four bar plots* structure-based LDA models. *Black grid symbol* indicates compounds predicted as selective according to the respective AR subtype model; *white colour with dots* indicates compounds predicted as not-selective to the respective AR subtype model; *grey colour* indicates uncertain prediction. The A<sub>1</sub> model well retrieves most A<sub>1</sub> selective test compounds (only one uncertain prediction out of 8 selective molecules); on the other hand, 4 out of 9 (44 %) A<sub>2A</sub> selective and 4 out of 7 (57 %) A<sub>3</sub> selective compounds are retrieved as well; only for A<sub>2B</sub> selective compounds (2 out of 7 (22 %)) retrieval is low. The A<sub>2A</sub> model efficiently retrieves its own A<sub>2A</sub> selective test compounds

(6 out of 9), perfectly rejects A<sub>2B</sub> selectives and fairly good rejects A<sub>1</sub> and A<sub>3</sub> selective compounds (picking 1 out of 8 (12 %) resp. 1 out of 7 (14 %)). Thus, the A<sub>2A</sub> LDA model might be useful in screening for A<sub>2A</sub> selective compounds. The A<sub>2B</sub> model nicely retrieves the A<sub>2B</sub> selectives (2 uncertain out of 7 (29 %)), perfectly rejects A<sub>2A</sub> selective compounds (3 uncertain out of 9 (33 %)) and rather well rejects A<sub>1</sub> (selecting only 1 out of 8 (12 %)) and A<sub>3</sub> selectives (1 uncertain prediction out of 7 (14 %)). The A<sub>3</sub> model fairly well retrieves the A<sub>3</sub> selectives (5 out of 7 (71 %)), and very well rejects A<sub>1</sub>, (2 wrong predictions out of 8 (25 %)), A<sub>2A</sub> (3 uncertain out of 9 (33 %)) and A<sub>2B</sub> selectives (5 uncertain out of 7 (71 %))

The LDA-R score described above served to check the capability of the respective subtype model to preferentially retrieve selective compounds, irrespective of the fact that these models were not trained in terms of selectivity aspects. Figure 10 shows the bar plot for the classification of the selective compounds as predicted by each FLAP LDA subtype model; black grid symbol indicates

prediction as selective (positive LDA-R values) whereas white colour with dots indicates prediction as not-selective (negative LDA-R values), grey colour defines an area of uncertainty in prediction.

Taken together, this analysis nicely indicates the potential capability of FLAP LDA models to uncover selectivity though subtype models were not specifically



**Fig. 11** Protein–Protein comparison of ARs. **a** A<sub>1</sub>R onto A<sub>3</sub>R; **b** A<sub>2A</sub>R onto A<sub>2B</sub>R; **c** A<sub>1</sub>R onto A<sub>2A</sub>R; **d** A<sub>2A</sub>R onto A<sub>3</sub>R; **e** A<sub>2B</sub>R onto A<sub>1</sub>; **f** A<sub>2B</sub>R onto A<sub>3</sub>R; with ZM241385 crystallographic pose (in wireframe blue). Orange circles indicate hydrophobic MIF

differences; gray circles indicate similarities. Sticks denote essential conserved residues for antagonist binding; thin wireframes denote marginal variable residues which might be involved in antagonist binding.

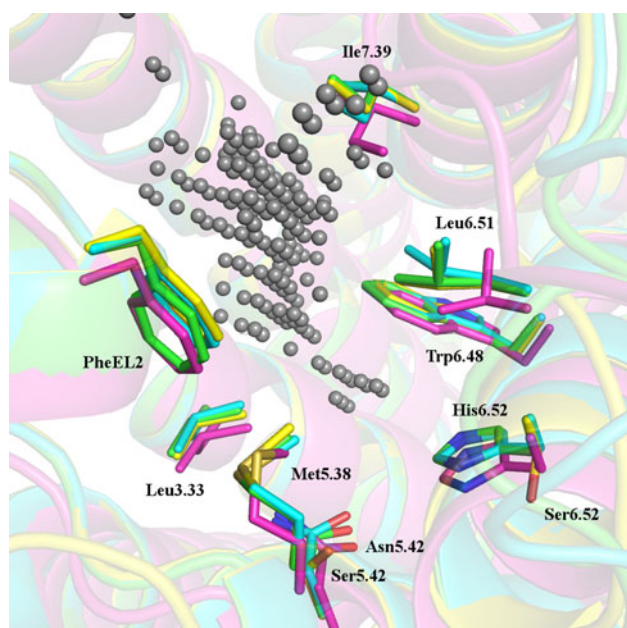
trained for revealing selectivity. Most ligand-based FLAP models displayed rather high sensitivity to retrieve their own selective antagonists. Only the A<sub>1</sub>R model showed low discriminant power cause it retrieved selective A<sub>2A</sub>, A<sub>2B</sub> and A<sub>3</sub> antagonists as well (Fig. 10a).

#### Selectivity aspects: a Protein–Protein comparison between receptor subtypes

In order to analyze aspects of AR antagonist selectivity, we performed Protein–Protein comparisons of all involved AR subtype pockets by comparing GRID MIF arrangements in the respective binding sites. Protein–Protein comparison has been applied in many studies in order to address selectivity [77]. We compared all six possible cavity pairs. Hydrophobicity seems to play a major role in ARs antagonist selectivity. Figure 11 shows the main differences found among the hydrophobic regions of the AR pockets. Grey spheres indicate regions in common between

overlapped cavities, while orange spheres highlight differences. In all cases, we found a shared hydrophobic region located at the central part of the overlapped cavities. Indeed, the respective Phe<sup>EL2</sup> residues seem to be mainly responsible for such hydrophobic interactions. Beyond it, some additional conserved residues are involved: Leu<sup>3.33</sup>, Ile<sup>7.39</sup>, Leu<sup>6.51</sup> (Val in the A<sub>2B</sub>R), Met<sup>5.38</sup> and Trp<sup>6.48</sup> (see Fig. 12 bottom). This common hydrophobic cavity might be important not only for antagonist binding, but also to maintain the receptor in the inactive conformation, in particular pertaining Trp<sup>6.48</sup> as reported in many studies [18, 69, 78–80]. It is noteworthy that the histidine residue located one turn of alpha helix above this tryptophan, His<sup>6.52</sup>, is substituted by the smaller Ser<sup>6.52</sup> in the particular case of A<sub>3</sub>R. This mutation is correlated with that of a proximal residue: the smaller Ser<sup>5.42</sup> in A<sub>3</sub>R substitutes the otherwise conserved Asn<sup>5.42</sup> in the other AR subtypes. These two serines form an extra cavity which is reported as responsible for specificity at the A<sub>3</sub>R receptor. Consequently,





**Fig. 12** ARs overlapping. Evidence of a common hydrophobic region generated by the highlighted conserved residues. This portion, especially Trp<sup>6.48</sup> residue is reported as responsible of antagonist binding and receptor activation. His<sup>6.52</sup> residue is not conserved in the A<sub>3</sub>R but is substituted by a Ser residue which is responsible of receptor specificity together with Ser<sup>5.42</sup> (see “Results and Discussion” section)

A<sub>3</sub>R antagonists are supposed to bind in this extra portion of the cavity by accommodating bulkier hydrophobic substituents which are linked to the typical hetero-aromatic central core [20, 81] (see orange spheres at the bottom right corner of Fig. 11a, d, f).

The A<sub>1</sub>R/A<sub>3</sub>R and the A<sub>2A</sub>R/A<sub>2B</sub>R overlap displayed minor differences in hydrophobic regions, confirming the phylogenetic similarity of the respective subtypes (Fig. 11a, b) [8]. One of the major differences between ARs subtypes involves position 7.35 which is characterized by smaller residues in the A<sub>1</sub>R (Thr)-A<sub>3</sub>R (Leu), compared to the more voluminous Met residues found in the A<sub>2A</sub>R-A<sub>2B</sub>R, and position 5.30 at the EL2 (Glu in A<sub>1</sub>, A<sub>2A</sub> and A<sub>2B</sub>) which varies only in the case of the A<sub>3</sub>R (Val). This glutamic acid is shown to form polar interactions with His<sup>7.29</sup> (A<sub>1</sub>R/A<sub>2A</sub>R) or Asn<sup>7.29</sup> (A<sub>2B</sub>R), favouring the closing of the binding site and offering additional hydrogen bond acceptor for the ligands, although some flexibility is expected according to crystal structures and MD simulations (Fig. 11a, d) [44, 70, 82]. A Valine in the equivalent position does lose this potential hydrogen bond interaction making the A<sub>3</sub>R cavity unique for accommodating bulkier hydrophobic substituents at the top of the site, as also experimentally demonstrated by Yaziji et al. [20]. Due to the less voluminous A<sub>1</sub>R/A<sub>3</sub>R 7.35 residues (Fig. 11a) generating less extended DRY field, hydrophobic differences become significant for the A<sub>2A</sub>R or A<sub>2B</sub>R overlapped

to the A<sub>1</sub>R or A<sub>3</sub>R (Fig. 11c, d, e, f). We also noticed other two important differences between the A<sub>2A</sub>R/A<sub>2B</sub>R DRY MIFs due to the nature of respectively His<sup>7.29</sup>/Asn<sup>7.29</sup> and Tyr<sup>7.36</sup>/Asn<sup>7.36</sup> (Fig. 11b). On one hand, the phenol ring of the co-crystallized ZM241385 antagonist well interacts with the A<sub>2A</sub>R Tyr<sup>7.36</sup> residue through  $\pi$ - $\pi$  stacking interaction, which is not possible to establish with the A<sub>2B</sub>R Asn residue. This might explain the experimentally increased preference for the A<sub>2A</sub>R ( $K_i$  A<sub>2A</sub> = 1.2 nM; IC<sub>50</sub> A<sub>2B</sub> = 42 nM [83]). However, it is well recognized that the aliphatic chain and the phenol ring in ZM241385 antagonist are very flexible within the A<sub>2B</sub>R binding site, according to the different crystal structures [43, 70]. On the other hand, the lower affinity for the A<sub>1</sub>R ( $K_i$  A<sub>1</sub> = 180 nM [83]) might be due to a smaller hydrophobic region, generated by the smaller and lower hydrophobicity of Thr<sup>7.35</sup> compared to the Met<sup>7.35</sup> of the A<sub>2A</sub>R or the A<sub>2B</sub>R (high density orange spheres; Fig. 11c, e). However, also shape differences of antagonists bounded were found important to enhance selectivity A<sub>1</sub>R/A<sub>2A</sub>R selectivity, as also reported in a recent study [84]. Small lipophilic substituents in 1,2,4-Triazine derivatives have been reported to enhance A<sub>2A</sub>R selectivity over A<sub>1</sub>R. This is given by a slight shape difference in the two cavities; A<sub>1</sub>R (Thr<sup>7.42</sup>) – A<sub>2A</sub>R (Ser<sup>7.42</sup>) and A<sub>1</sub>R (Val<sup>2.57</sup>) – A<sub>2A</sub>R (Ala<sup>2.57</sup>).

In addition, complete loss of affinity for the A<sub>3</sub>R (% inhibition A<sub>3</sub> = 5 % [83]) might be due to a combined effect of the less voluminous hydrophobicity of Leu<sup>7.35</sup> compared to Met<sup>7.35</sup>, as previously shown for the A<sub>1</sub>R, and the lack of hydrophobic interactions at the inner A<sub>3</sub>R cavity as discussed above (Fig. 11a, d, f).

## Conclusions

The performance of ligand-, structure- and pharmacophore-based molecular fingerprints was compared in this case study on antagonists of all four AR subtypes. A fundamental prerequisite for building valid computational models is the generation of highly diverse, evenly distributed datasets. For that purpose, we screened the comprehensive ChEMBL data collection to generate chemically highly diverse datasets exhibiting an even distribution of the number of actives and inactives to train the models for the four AR subtypes.

In the ligand-based FLAP mode, enrichment curves yielded following predictivity ranking: A<sub>2A</sub>R > A<sub>2B</sub>R > A<sub>3</sub>R > A<sub>1</sub>R antagonists. FLAP model performance was superior to ECFP4 and ROCS for all AR subtypes. However, when focusing on the early first part of enrichment curves, retrieval of actives, using ECFP4 or ROCS, remains satisfactory.



In the structure-based FLAP mode, the *FLAPsite* tool serves to consider AR cavities as putative templates. Out of the candidates that cavity is selected that contains the family-conserved Asn<sup>6.55</sup>. Ranking of model predictivity for the four subtypes is identical with the LB-approach. SB-FLAP models well match with mutagenesis studies thereby underlining FLAP model validity. Also in the structure-based version, FLAP clearly supersedes the performance of PLANTS and GOLD algorithms.

The novel FLAPPharm tool was used first time here for pharmacophore generation. It was first applied to a collection of AR antagonists from Fredholm [8] composed of structures exhibiting selectivity to one AR subtype over the remaining three. Using this set, pharmacophores were generated tightly resembling previous hypotheses which in turn demonstrates the reliability of the FLAPPharm tool. Then, we performed pharmacophore-based FLAP modeling on the same training and test sets as applied for the ligand- and structure-based FLAP modes. Also in this case, predictivity slightly decreases in the order  $A_{2A}R > A_{2B}R > A_{3R} > A_{1R}$  antagonists.

Finally, we demonstrated the capability of FLAP models to uncover selectivity aspects although single AR subtype models were not trained for revealing selectivity.

The above described robust and validated models obtained from optimized datasets will serve for prospective virtual screening for novel antagonists of all four AR subtypes in follow-up studies. In this work, the numbers of false positives for test sets were even lower than for training sets; this finding makes the models developed here particularly attractive for prediction of novel AR antagonists.

**Acknowledgments** We thank Molecular Discovery Ltd. for granting FLAP suite license. We also thank Lydia Siragusa for assisting us in the Protein–Protein study.

## References

1. Fredholm BB, Ijzerman AP, Jacobson KA, Klotz KN, Linden J (2001) *Pharmacol Rev* 53(4):527
2. Zhou QY, Li CY, Olah ME, Johnson RA, Stiles GL, Civelli O (1992) *Proc Nat Acad Sci USA* 89(16):7432
3. Salvatore CA, Jacobson MA, Taylor HE, Linden J, Johnson RG (1993) *Proc Nat Acad Sci USA* 90(21):10365
4. Ralevic V, Burnstock G (1998) *Pharmacol Rev* 50(3):413
5. Fredholm BB (2011) *Sleep Biol Rhythms* 9:24
6. Gessi S, Merighi S, Varani K, Borea PA (2011) *Advances in pharmacology* (San Diego, Calif) 61:41
7. Mueller CE, Jacobson KA (2011) *Biochimica Et Biophysica Acta-Biomembranes* 1808(5):1290
8. Fredholm BB, Ijzerman AP, Jacobson KA, Linden J, Mueller CE (2011) *Pharmacol Rev* 63(1):1
9. Langmead CJ, Andrews SP, Congreve M, Errey JC, Hurrell E, Marshall FH, Mason JS, Richardson CM, Robertson N, Zhukov A, Weir M (2012) *J Med Chem* 55(5):1904
10. Martinelli A, Tuccinardi T (2007) *Med Res Rev* 28(2):247
11. Michielan L, Stephanie F, Terfloeth L, Hristozov D, Cacciari B, Klotz K-N, Spalluto G, Gasteiger J, Moro S (2009) *J Chem Inf Model* 49(12):2820
12. Costanzi S, Tikhonova IG, Harden TK, Jacobson KA (2009) *J Comput Aided Mol Des* 23(11):747
13. Goddard WA III, Kim S-K, Li Y, Trzaskowski B, Griffith AR, Abrol R (2010) *J Struct Biol* 170(1):10
14. Wei J, Wang S, Gao S, Dai X, Gao Q (2007) *J Chem Inf Model* 47(2):613
15. Tafi A, Bernardini C, Botta M, Corelli F, Andreini M, Martinelli A, Ortore G, Baraldi PG, Fruttarolo F, Borea PA, Tuccinardi T (2006) *J Med Chem* 49(14):4085
16. Cheong SL, Federico S, Venkatesan G, Paira P, Shao Y-M, Spalluto G, Yap CW, Pastorin G (2011) *Bioorg Med Chem Lett* 21(10):2898
17. Xu Z, Cheng F, Da C, Liu G, Tang Y (2010) *J Mol Model* 16(12):1867
18. Cheng F, Xu Z, Liu G, Tang Y (2010) *Eur J Med Chem* 45(8):3459
19. Wei J, Li H, Qu W, Gao Q (2009) *Neurochem Int* 55(7):637
20. Yaziji V, Rodriguez D, Gutierrez-de-Teran H, Coelho A, Caamano O, Garcia-Mera X, Brea J, Isabel Loza M, Isabel Cadavid M, Sotelo E (2011) *J Med Chem* 54(2):457
21. Faller B, Ottaviani G, Ertl P, Berellini G, Collis A (2011) *Drug Discov Today* 16(21–22):976
22. Baroni M, Cruciani G, Sciabola S, Perruccio F, Mason JS (2007) *J Chem Inf Model* 47(2):279
23. Cross S, Baroni M, Carosati E, Benedetti P, Clementi S (2010) *J Chem Inf Model* 50(8):1442
24. Grant JA, Gallardo MA, Pickup BT (1996) *J Comput Chem* 17(14):1653
25. Rogers D, Hahn M (2010) *J Chem Inf Model* 50(5):742
26. Korb OS, Exner TE (2007) *Swarm Intell* 1:115
27. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) *J Mol Biol* 267:727
28. <https://www.ebi.ac.uk/chembl/db/>
29. Gasteiger J, Teckentrup A, Terfloeth L, Spycher S (2003) *J Phys Org Chem* 16(4):232
30. Milletti F, Storch L, Sforna G, Cruciani G (2007) *J Chem Inf Model* 47:2172
31. Cruciani C, Crivori P, Carrupt PA, Testa B (2000) *J Mol Struct Theochem* 503(1–2):17
32. Cruciani G, Pastor M, Guba W (2000) *Eur J Pharm Sci* 11:S29
33. Hudson BD, Hyde RM, Rahr E, Wood J (1996) *Quant Struct-Act Relat* 15(4):285
34. Carosati E, Sciabola S, Cruciani G (2004) *J Med Chem* 47(21):5114
35. Brincat JP, Carosati E, Sabatini S, Manfroni G, Fravolini A, Raygada JL, Pate D, Kaatz GW, Cruciani G (2011) *J Med Chem* 54(1):354
36. Carosati E, Mannhold R, Wahl P, Hansen JB, Fremming T, Zamora I, Cianchetta G, Baroni M (2007) *J Med Chem* 50(9):2117
37. Sciabola S, Stanton RV, Mills JE, Flocco MM, Baroni M, Cruciani G, Perruccio F, Mason JS (2010) *J Chem Inf Model* 50:155
38. Cross S, Baroni M, Goracci L, Cruciani G (2012) *J Chem Inf Model* 21:21
39. Cross S, Ortuso F, Baroni M, Costa G, Distinto S, Moraca F, Alcaro S, Cruciani G (2012) *J Chem Inf Model* 21:21
40. Bostrom J, Greenwood JR, Gottfries J (2003) *J Mol Graph Model* 21(5):449
41. <http://accelrys.com/products/pipeline-pilot/>
42. Morgan HL (1965) *J Chem Doc* 5(2):107
43. Jaakola V-P, Griffith MT, Hanson MA, Cherezov V, Chien EYT, Lane JR, Ijzerman AP, Stevens RC (2008) *Science* 322(5905):1211

44. Rodriguez D, Pineiro A, Gutierrez-de-Teran H (2011) *Biochemistry* 50(19):4194
45. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) *Nucleic Acids Res* 25(24):4876
46. Sali A, Blundell TL (1993) *J Mol Biol* 234(3):779
47. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) *J Appl Crystallog* 26:283
48. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, 3rd, Snoeyink J, Richardson JS, Richardson DC (2007) *Nucleic Acids Res* 35(Web Server issue):W375
49. Fiser A, Do RK, Sali A (2000) *Protein Sci* 9(9):1753
50. MacroModel, version 9.7, Schrödinger L, New York, NY, 2009
51. Korb O, Stutzle T, Exner TE (2009) *J Chem Inf Model* 49(1):84
52. Ballesteros JA, Weinstein H (1995) *Methods Neurosci* 25:366
53. Xu F, Stevens Raymond C (2011) *Structure* (London, England: 1993) 19(9):1204
54. Oprea TI, Gottfries J (2001) *J Comb Chem* 3(2):157
55. Fisher RA (1936) *Ann Eugen* 7:179
56. Mc Lachlan GJ (2004) *Wiley series in probability and statistics*
57. Jacobson KA, Linden J (2011) *Advances in pharmacology* 61 (XV–XVI)
58. Chang LCW, Spanjersberg RF, Kunzel JKVFD, Brussee J, Ijzerman AP (2006) *J Med Chem* 49(10):2861
59. Scheiff AB, Yerande SG, El-Tayeb A, Li W, Inamdar GS, Vasu KK, Sudarsanam V, Mueller CE (2010) *Bioorg Med Chem* 18(6):2195
60. Ferrarini PL, Betti L, Cavallini T, Giannaccini G, Lucacchini A, Manera C, Martinelli A, Ortore G, Saccomanni G, Tuccinardi T (2004) *J Med Chem* 47(12):3019
61. Novellino E, Cosimelli B, Ehlardo M, Greco G, Iadanza M, Lavecchia A, Rimoli MG, Sala A, Da Settimo A, Primofiore G, Da Settimo F, Taliani S, La Motta C, Klotz KN, Tuscano D, Trincavelli ML, Martini C (2005) *J Med Chem* 48(26):8253
62. Manetti F, Schenone S, Bondavalli F, Brullo C, Bruno O, Ranise A, Mosti L, Menozzi G, Fossa P, Trincavelli ML, Martini C, Martinelli A, Tintori C, Botta M (2005) *J Med Chem* 48(23):7172
63. Sirci F, Istyastono EP, Vischer HF, Kooistra AJ, Nijmeijer S, Kuijter M, Wijtmans M, Mannhold R, Leurs R, Esch IJPd, Graaf Cd (2012) *J Chem Inf Model* (submitted)
64. Kim SK, Gao ZG, Van Rompaey P, Gross AS, Chen A, Van Calenbergh S, Jacobson KA (2003) *J Med Chem* 46(23):4847
65. Kim JH, Wess J, Schoneberg T, Jacobson KA (1995) *J Biol Chem* 270:13987
66. Piirainen H, Ashok Y, Nanekar RT, Jaakola V-P (2011) *Biochimica Et Biophysica Acta-Biomembranes* 1808(5):1233
67. Ivanov AA, Baskin II, Palyulin VA, Piccagli L, Baraldi PG, Zefirov NS (2005) *J Med Chem* 48(22):6813
68. Ivanov AA, Barak D, Jacobson KA (2009) *J Med Chem* 52(10):3284
69. Gao ZG, Chen A, Barak D, Kim SK, Muller CE, Jacobson KA (2002) *J Biol Chem* 277(21):19056
70. Dore AS, Robertson N, Errey JC, Ng I, Hollenstein K, Tehan B, Hurrell E, Bennett K, Congreve M, Magnani F, Tate CG, Weir M, Marshall FH (2011) *Structure* 19(9):1283
71. Weyler S, Fuelle F, Diekmann M, Schumacher B, Hinz S, Klotz K-N, Mueller CE (2006) *ChemMedChem* 1(8):891
72. Elzein E, Rao KA, Li XF, Perry T, Parkhill E, Palle V, Varkhedkar V, Gimbel A, Zeng DW, Lustig D, Leung K, Zablocki J (2006) *Bioorg Med Chem Lett* 16(2):302
73. Kalla RV, Elzein E, Perry T, Li X, Gimbel A, Yang M, Zeng D, Zablocki J (2008) *Bioorg Med Chem Lett* 18(4):1397
74. Baraldi PG, Tabrizi MA, Preti D, Bovero A, Romagnoli R, Fruttarolo F, Zaid NA, Moorman AR, Varani K, Gessi S, Merighi S, Borea PA (2004) *J Med Chem* 47(6):1434
75. Baraldi PG, Tabrizi MA, Bovero A, Avitabile B, Preti D, Fruttarolo F, Romagnoli R, Varani K, Borea PA (2003) *Eur J Med Chem* 38(4):367
76. Baraldi PG, Cacciari B, Moro S, Spalluto G, Pastorin G, Da Ros T, Klotz KN, Varani K, Gessi S, Borea PA (2002) *J Med Chem* 45(4):770
77. Afzelius L, Raubacher F, Karlen A, Jorgensen FS, Andersson TB, Masimirembwa CM, Zamora I (2004) *Drug Metab Dispos* 32(11):1218
78. Hino T, Arakawa T, Iwanari H, Yurugi-Kobayashi T, Ikeda-Suno C, Nakada-Nakura Y, Kusano-Arai O, Weyand S, Shimamura T, Nomura N, Cameron AD, Kobayashi T, Hamakubo T, Iwata S, Murata T (2012) *Nature* 482(7384):237
79. Gao ZG, Kim SK, Biadatti T, Chen WZ, Lee K, Barak D, Kim SG, Johnson CR, Jacobson KA (2002) *J Med Chem* 45(20):4471
80. Jacobson KA, Gao ZG, Chen AS, Barak D, Kim SA, Lee K, Link A, Van Rompaey P, van Calenbergh S, Liang BT (2001) *J Med Chem* 44(24):4125
81. Katritch V, Kufareva I, Abagyan R (2011) *Neuropharmacology* 60(1):108
82. Xu F, Wu H, Katritch V, Han GW, Jacobson KA, Gao Z-G, Cherezov V, Stevens RC (2011) *Science* 332(6027):322
83. Lenzi O, Colotta V, Catarzi D, Varano F, Poli D, Filacchioni G, Varani K, Vincenzi F, Borea PA, Paoletta S, Morizzo E, Moro S (2009) *J Med Chem* 52(23):7640
84. Congreve M, Andrews SP, Doré AS, Hollenstein K, Hurrell E, Langmead CJ, Mason JS, Ng WI, Tehan B, Zhukov A, Weir M, Marshall FH (2012) *J Med Chem* 55:1898
85. Zhang X, Rueter JK, Chen Y, Moorjani M, Lanier MC, Lin E, Gross RS, Tellew JE, Williams JP, Lechner SM, Markison S, Joswig T, Malany S, Santos M, Castro-Palomino JC, Crespo MI, Prat M, Gual S, Diaz J-L, Saunders J, Slee DH (2008) *Bioorg Med Chem Lett* 18(6):1778
86. Borrmann T, Hinz S, Lertarelli DCG, Li W, Florin NC, Scheiff AB, Mueller CE (2009) *J Med Chem* 52(13):3994
87. Stefanachi A, Nicolotti O, Leonetti F, Cellamare S, Campagna F, Isabel Loza M, Manuel Brea J, Mazza F, Gavuzzo E, Carotti A (2008) *Bioorg Med Chem* 16(22):9780
88. Kim YC, Ji XD, Jacobson KA (1996) *J Med Chem* 39(21):4142
89. van Muijlwijk-Koezen JE (2001) *PhD thesis*