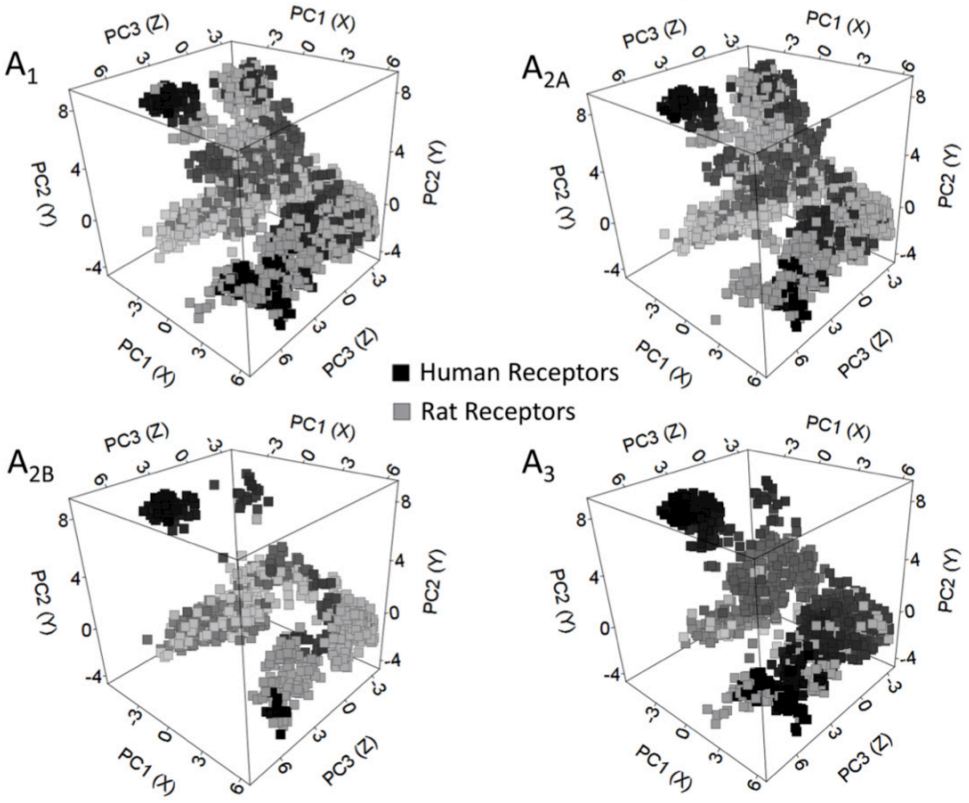
Déjà Vu - Réjà Vu : on knowledge-based approaches linking ligand and target information to bioactivity

Westen, G.J.P. van

Chapter 4

Identifying Novel Adenosine Receptor Ligands by Simultaneous Proteochemometric Modeling of Rat and Human Bioactivity Data



G.J.P. Van Westen, O.O. van den Hoven, R. van der Pijl, T. Mulder-Krieger, H. de Vries, J.K. Wegner, A.P. IJzerman, H.W.T. Van Vlijmen, and A. Bender; J. Med. Chem.; 2012: ***55*** *(16): 7010-7020.*

Contents

4.1 Abstract

The four subtypes of adenosine receptors form relevant drug targets in the treatment of e.g., diabetes and Parkinson's disease. In the present study we aimed at finding novel small molecule ligands for these receptors using virtual screening approaches based on proteochemometric (PCM) modeling. We combined bioactivity data from all human and rat receptors in order to widen available chemical space. After training and validating a proteochemometric model on this combined dataset (Q^2 of 0.73, RMSE of 0.61) we virtually screened a vendor database of 100,910 compounds. Of 54 compounds purchased, six novel high affinity adenosine receptor ligands were confirmed experimentally, one of which displayed an affinity of 7 nM on the human adenosine A1 receptor. We conclude that the combination of rat and human data performs better than human data only. Furthermore, we conclude that proteochemometric modeling is an efficient method to quickly screen for novel bioactive compounds.

4.2 Introduction

4.2.1 The Adenosine Receptors. G protein-coupled receptors (GPCRs) are membrane-bound proteins and targets for many hormones and neurotransmitters in the body. As such they are ideal drug targets with a large degree of inherent selectivity due to their tissue specific expression. The local hormone adenosine interacts with four different GPCRs, the adenosine A_1, A_{2A}, A_{2B} and A_3 receptors. These receptor subtypes are involved in many (patho)physiological processes, including diseases such as type 2 diabetes, heart arrhythmias, and Parkinson's disease.[1] In the current work we set ouy to identify novel small molecule ligands for these adenosine receptors using virtual screening approaches.

4.2.2 Proteochemometric Modeling. Different approaches exist to select potentially bioactive compounds using computational models. Conventionally, a structure-activity model can be created using known compounds.[2, 3] The obtained model can then be used to predict the modeled output variable for compounds that have not been experimentally tested, on the basis of the 'Molecular Similarity Principle' which states that similar compounds show similar activity.[4] However, in the case of the adenosine receptors we are dealing with *multiple similar* targets, rather than one target. Previously it has been shown that proteochemometric modeling (PCM),[5-8] is able to create robust predictive models for multiple similar targets.[9-11] As has been reviewed in detail before,[7] PCM takes both ligand- as well as target-similarity into account, and can thereby also benefit from the principle that '*similar targets bind similar ligands*'. Given the ability of PCM models to also consider ligands active on related receptors when predicting bioactivity against a particular receptor, this increases the likelihood of identifying both active compounds and novel active chemotypes. Hence, we chose to create a PCM model trained on the adenosine receptor subfamily, rather than to train individual bioactivity models. We hypothesized the PCM model to perform better than these individual models and hoped to find both compounds that are a selective ligand for a single receptor but also compounds that are globally active ligands active on the entire subfamily of human adenosine receptors.

4.2.3 Chemical space and target space. Chemical space can be characterized based on the similarity of the compounds that interact with adenosine receptors. It is this space that is exploited when a structure-activity model is created for one of the receptors, as chemicals predicted to be closely located to known ligands on a target are expected to be ligands of that target. Target space can be characterized by the similarity of the targets. It is this space that is exploited when a multiple sequence alignment shows that the A_{2A} and A_{2B} adenosine receptors are more similar than the A_{2A} and A_3 receptors (Supporting **Table S1**).[12] Since PCM uses both ligand and target to predict an output variable, it can thereby also consider the fact that some features have different effects on different targets to better fit the data.[13, 14] Therefore our hypothesis in the current work is that PCM models perform better in prediction of values for data points that were not originally in the training set when compared to conventional structure-activity models, a principle that has been shown before for different mutants of HIV reverse transcriptase.[10]

4.2.4 Inclusion of multiple species orthologs. Historically rat tissues have been the source of adenosine receptors for the testing of novel chemical entities, before the human receptors became available for in vitro testing.[12] As a result a large amount of historical data is available on the rat orthologs (identical receptors in other species) of the human adenosine receptors including affinity data of small molecules (5,397 data points in ChEMBL 2).[15] Recently it has been shown that in general, small molecule binding is conserved between human and rat orthologs. However, a species specific pharmacology is observed for the A_1 and A_{2A} receptor; relative to human receptors the average pKi for the A_1 receptor is -0.51 log units while 0.41 log units for the A_{2A} receptor.[16] Moreover, from the full sequence similarity it becomes apparent that rat and human orthologs show greater similarity (identity 84% ± 8) than human paralogs (similar receptors in the same species) among each other (48% ± 7) (Supporting **Table S1**).

As it has been previously shown that PCM can model paralog subfamilies,[17, 18] in this work we extend this approach by proposing to include *orthologs* in the training set, in order to capture the chemical space associated with these orthologs in the model as well, while at the same time considering target differences in the PCM model generation process. Through a combined virtual and experimental screening we hope to find both novel selective ligands (active on a single receptor) and novel global ligands (active on all human adenosine receptors) as both these ligand types are of interest to our research team.

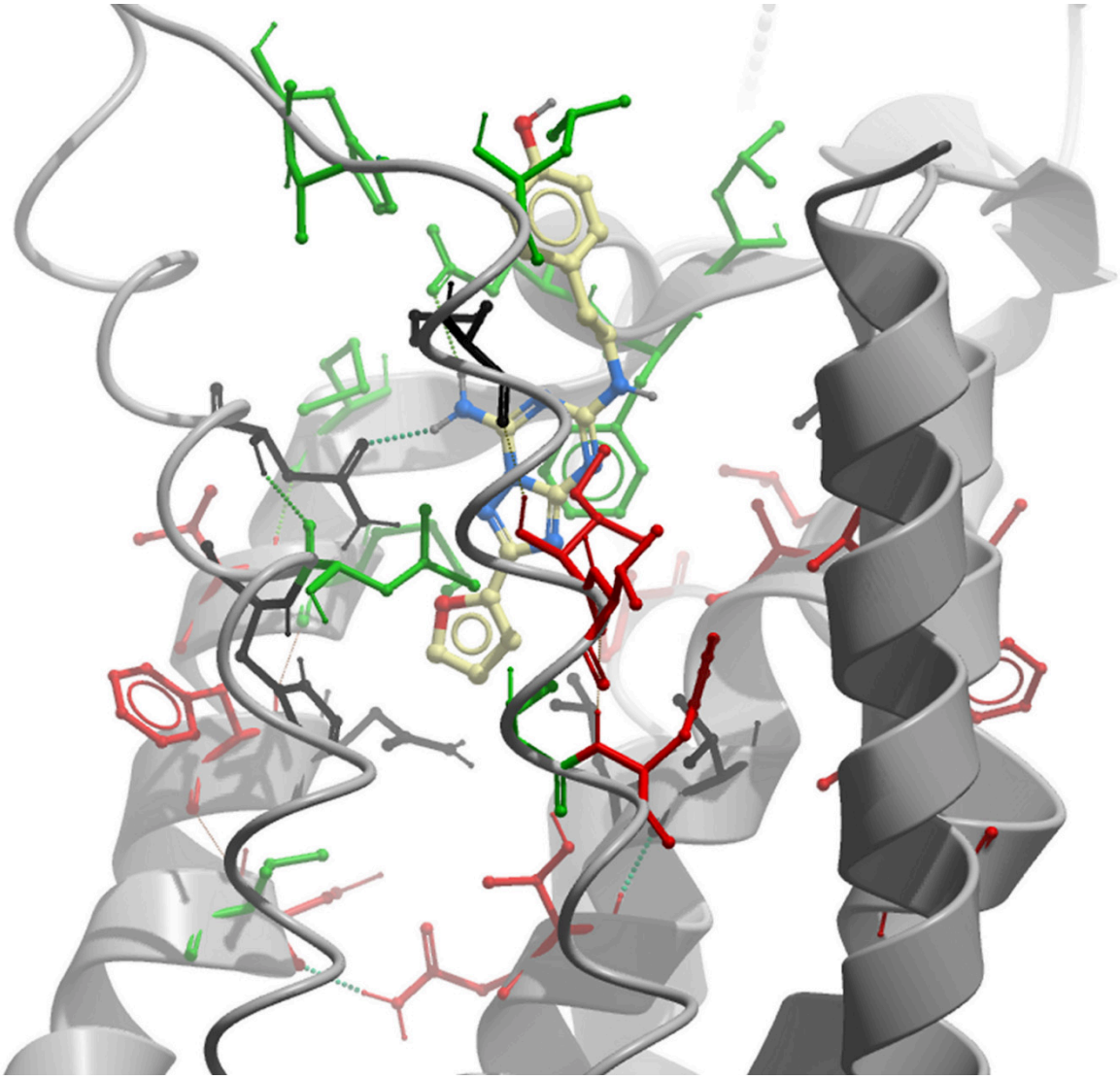


Figure 4.1: The binding site we selected to define the target similarity as visualized in PDB structure 3EML. The protein backbone is in gray and the co-crystallized ligand (ZM-241385) in ball and stick model. The green residues were obtained through selection of a 5 Å sphere around the co-crystallized ligand, red residues were obtained through TEA analysis (see **4.5.4**) and residues in gray/black occur in both analyses. Note that residues in both trans-membrane domains and extracellular loops were included.

4.3 Results and Discussion

4.3.1 Characterizing target space. Figure 4.1 shows the residues selected as binding site displayed in the adenosine A_{2A} receptor crystal structure containing antagonist ZM241385 (PDB code 3EML).[19] The figure displays the trans-membrane (TM) domains and extracellular loops (ELs) of the receptor. Individual amino acid side chains have been visualized in ball and stick model. The green residues were obtained through selection of a 5 Å sphere around the co-crystallized ligand, red residues were obtained through Two-Entropy Analysis (TEA) (see section **4.5.4**) and residues in gray/black occur in both analyses. **Figure 4.2** displays the results from a principal component analysis (PCA) of the ligand binding pocket in all receptors. The binding pocket was defined based on the same residues that were used to train the final model, representing *target space* (see section **4.5.4**). The PCA demonstrates that our binding site retains the pattern from full sequence similarity, in which receptor orthologs are more similar than paralogs. It should be noted however, that the difference between rat and human orthologs of the A_3 receptor is much larger than in any of the other three ortholog sets. This large difference is in agreement with the fact that compounds found to be active on the human A_3 receptor were much less active or even inactive on the rat A_3 receptor.[1] Therefore a full clustering of these two receptors based on the binding site would be contradictory to what we know from the chemical space of the two receptors, which is described below.

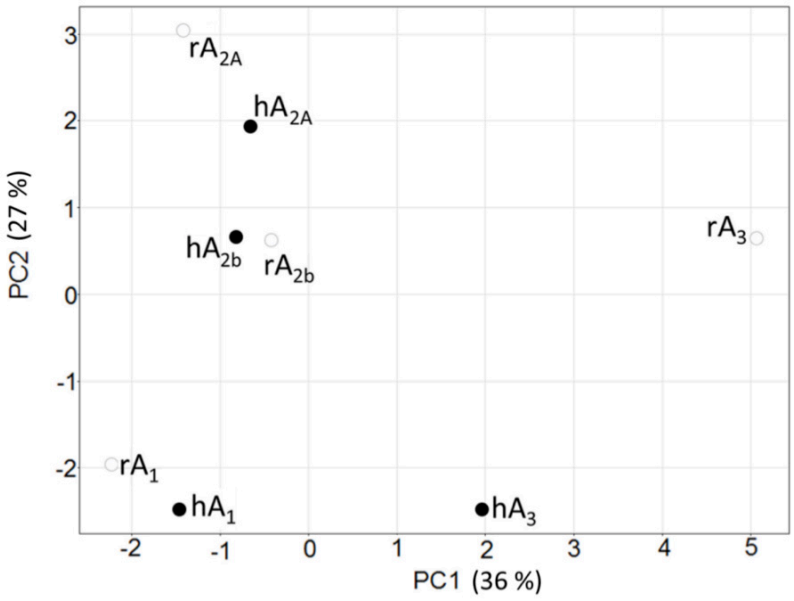


Figure 4.2: Principal component analysis of the similarity in target space. The adenosine receptor orthologs are more similar than their paralogs. human receptors are indicated with a black circle and rat receptors by a white circle. Both A_3 receptors are very different ('outliers'), while the A_{2A} and A_{2B} receptors cluster together. This observation is consistent with the fact that ligands active on the human A_3 receptor were often found to be inactive (or less active) on the rat A_3 receptor.

4.3.2 Characterizing chemical space. In addition to the analysis of target space, we also performed a PCA on the structures of all compounds (using the same descriptor as the final model) we had available in our data set, or *chemical space*, comprising a total of 10,999 data points. The results are displayed in **Figure 4.3** and data points are grouped by orthologs. In the same way as it could be observed in target space, a high similarity between orthologs is also visible in the chemical (ligand) space. Furthermore it becomes apparent that the chemical spaces for the A_1 and A_{2A} orthologs have been explored most extensively, while the chemical space for the A_{2B} orthologs has been sampled rather sparsely. Finally, the chemical space for the A_3 orthologs is dominated by compounds measured on the human ortholog, biasing in particular this dataset of active compounds. In fact, as mentioned earlier it has been hard to identify ligands (and in particular antagonists) that exhibit affinity for the rat A_3 receptor in previous work.[1]

The results from this PCA analysis show that there is significant (however in no case complete) overlap in the chemistry of the compounds that have been tested on the human as well as rat adenosine receptor subtypes (chemical space). Nevertheless, the number of identical compounds tested per ortholog pair is rather low (at about 5% of the total number of active compounds in our dataset, Supporting **Table S2**).

Likewise we compared the chemical space between the *paralogs* for both human and rat receptors (Supporting **Figure S1** and **S2**). The chemical space of annotated compounds for paralogs in the training set is very similar with the exception of the A_{2B} receptors. However, the points are colored according to their affinity on the receptors showing that the location of 'high affinity hotspots' differs between paralogs, while some hotspots are shared. This observation of high affinity hotspots' confirms that chemistry alone cannot explain the affinity differences between receptors, but also that selective compounds can be found within the training set, hence we expect our model to be able to predict selectivity.

The analysis of chemical space gave us confidence that bioactivity space between human and rat adenosine receptor orthologs is similar enough to allow us the use of PCM modeling approaches; still that it is also dissimilar enough to enable the discovery of novel bioactive ligands by considering bioactive space from both species in a single model.

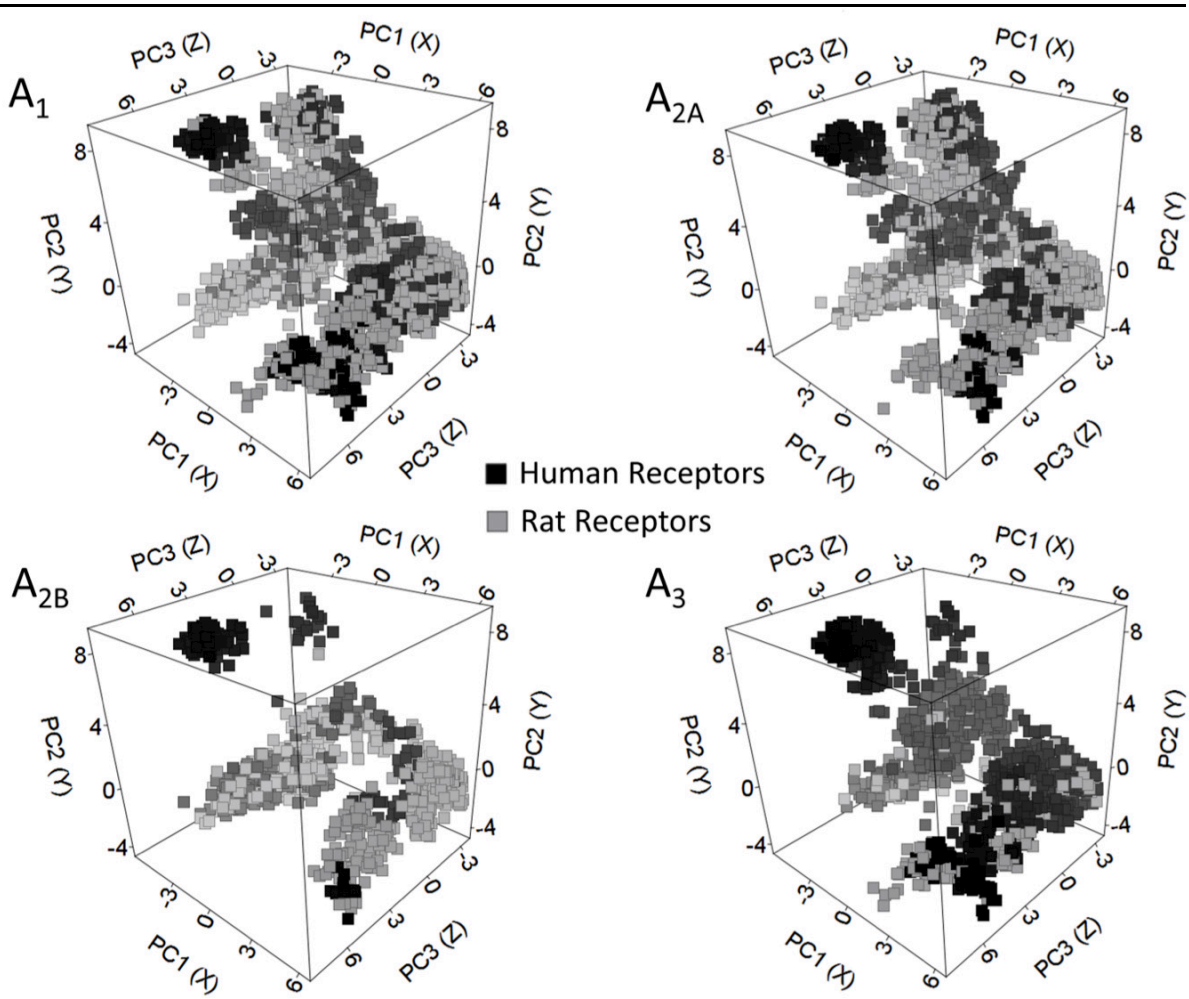


Figure 4.3: Principal component analysis of ligand chemical space. This PCA shows the large overlap in ligands that have been tested on ortholog pairs in the different species. The A_1 and A_{2A} receptors have the most densely populated chemical space, whereas A_{2B} has been explored the least. The space for the compounds tested on the A_3 receptors is dominated by compounds tested on the human ortholog. Note that the further along the x and z axes the point become lighter, black points fade to grey and grey points fade to white.

4.3.3 Target Descriptor. Firstly we identified the optimal selection method of the receptor binding site. (For a flow chart of the performed selections please see section **4.5**; here ECFP_4 fingerprints were employed as ligand descriptors, see section **4.5** section for further details.) In this part of the work we had a choice of 6 residue selection methods, two of which were structure based; the first one by selecting residues within a 5Å sphere around the co-crystallized ligand in PDB structure 3EML,[19] and the second one identical but using a 7 Å sphere. Furthermore, two selection methods were obtained utilizing TEA algorithms, selecting residues that were classified to be active in ligand recognition based on their evolutionary entropy.[20]

Here we used a conservative approach (TEA S), which selected a smaller number of residues, and a less restricted selection method (TEA L), which gave rise to a larger number of residues selected for model generation. Finally, we also evaluated two selection methods *combining* the a 5Å sphere with TEA S and one combining the 5Å sphere with TEA L. The best performing selection found was a combination of a 5 Å sphere around the co-crystallized ligand along with the small selection of TEA. This selection was named TEA S5 (Supporting **Figure S3** and Supporting **Table S3**).

During the optimization of our target descriptor by sampling different residue selection methods, we found that a larger selection is not always better. In fact, while the best performing binding site definition consisted of a combination of the two selection methods (crystal structure based and TEA based), it was in both cases the smallest residue selection within each method that performed optimally. Interestingly, we found that we needed to combine the crystal structure selection and TEA selection for optimal performance. In each individual selection method, both in the method based on the crystal structure alone and the method based solely on TEA, there was a pair of ortholog receptors that gave rise to an identical fingerprint. These were the A_{2A} receptors when selecting either a 5 Å or 7 Å sphere around the ligand in the crystal structure and the A_{2B} receptors when using the TEA based selection. However, since it was still possible to create predictive models in each individual case, it can be concluded that the activity space of these ortholog receptor pairs is highly similar (also see Supporting **Figure S1** and **S2**).

4.3.4 Ligand descriptor. Similar to the method used to identify the best descriptor binding site, we also identified the ligand descriptor giving rise to best modeling performance. Bender *et al.* in their analysis of descriptor space have shown that there is little difference between circular fingerprint performance and in this work similar results were obtained.[21] Our models identified the extended connectivity fingerprint using Sybyl atom typing (SCFP_4) to be the best-performing compound descriptor on this dataset (with an external validation RMSE of 0.70 log units and R_0^2 of 0.67) with three others close in performance. Those were FPFP_6 (RMSE 0.70 log units and R_0^2 0.68), EPFP_6 (RMSE 0.68 log units and R_0^2 0.69) and SPFP_6 (RMSE 0.69 log units and R_0^2 0.69) (see Supporting **Figure S4** and Supporting **Table S4**). Also in *predictive* power, *i.e.* the performance estimates in the cross validation compared to the external validation, the different fingerprints perform very similar but SCFP_4 better correlates to the external validation than in the others (see Supporting **Figure S4** and Supporting **Table S4**). We found this to be of high importance as we did not want to embark on a 'wet' experiment without having a fair estimate of model performance on unknown compounds.

4.3.5 Cross Validation. Finally, we sampled different cross-validation approaches by varying the amount of subdivisions in each cross-validation step. We observed that in the case of 5-fold cross validation the cross validation parameters are slightly worse compared with the external validation parameters, with a cross validated RMSE of 0.70 log units versus an RMSE for the external validation set of 0.68 log units. In addition the Q^2 is 0.69 in the cross validation and the R_0^2 is 0.71 in the external validation (Supporting **Figure S5**). When we increased the number of subdivisions, and hereby decreased the size of the fraction left out of the training during cross validation, this phenomenon was reversed. Hence the cross validated RMSE is slightly lower compared with RMSE in external validation (0.68 versus 0.70) and the Q^2 is slightly higher compared with R_0^2 (0.70 versus 0.69). This can indicate slight overtraining, as shown by Baumann,[22] which is the reason why we choose to implement 5-fold CV in the final model training procedure.

4.3.6 Final model training. The final model was trained on the full data set of eight receptors and 10,999 annotated data points. Given the preliminary results listed above, the model was built using SCFP_4 compound fingerprints and the TEA S5 residue selection. The training plot of the final model is shown in **Figure 4.4**, obtaining an R_0^2 of 0.95 and an RMSE of 0.26. The cross validated parameters, which constitutes a performance estimate, were a correlation coefficient (Q^2) of 0.73 and a prediction error (CV_RMSE) of 0.61 log units. This final model, created in Pipeline Pilot 8.5,[23] is provided in the Supporting Material. Furthermore, we also included in the Supporting Materials two tables showing the 25 substructures that have the largest positive (presence of these substructures leads to a higher pKi, Supporting **Table S5**) or largest negative effect on pKi (on average presence of these substructures leads to a lower pKi, Supporting **Table S6**). Finally, we have added the average effect on binding of the presence of the most occurring substructures. The top 100 most occurring substructures and their average effect on binding when present are given in Supporting **Table S7**.

The training plot of the final model, **Figure 4.4**, shows that especially compounds in the high pKi region (larger than 9) seem to be predicted more accurately. However, several compounds in this area, marked with a black circle, have been found to be underpredicted by a large margin. Upon identifying the outliers it was discovered these three points contained the same structure. Further literature studies showed that the outliers are all from Jacobson *et al.*[24] and that the original paper states that two of these compounds have been tested on the rat A_{2A} receptor, whereas ChEMBL 2 list them annotated to the human A_{2A} receptor. Moreover, the binding affinity values from this particular paper are much higher (pKi larger than 8.0) when compared with the affinity values these compounds and a large number of highly similar other compounds have on average in other papers (smaller than 7.0 and sometimes even smaller than 6.0). In Supporting **Table S8** are further details, i.e. the affinity of these particular compounds and a number of similar compounds in the training set. It would seem to be a reasonable explanation that these questionable values were experimental artifacts and a database annotation error. However, since this effect only occurs sporadically and only in the cases of these 8-cyclohexylcaffeine derivates we decided to keep the data points in the final model. It should be noted that our model was able to pick up these outlying experimental values. However; another consideration was that exhaustive checking of all 10,999 data points was practically infeasible.

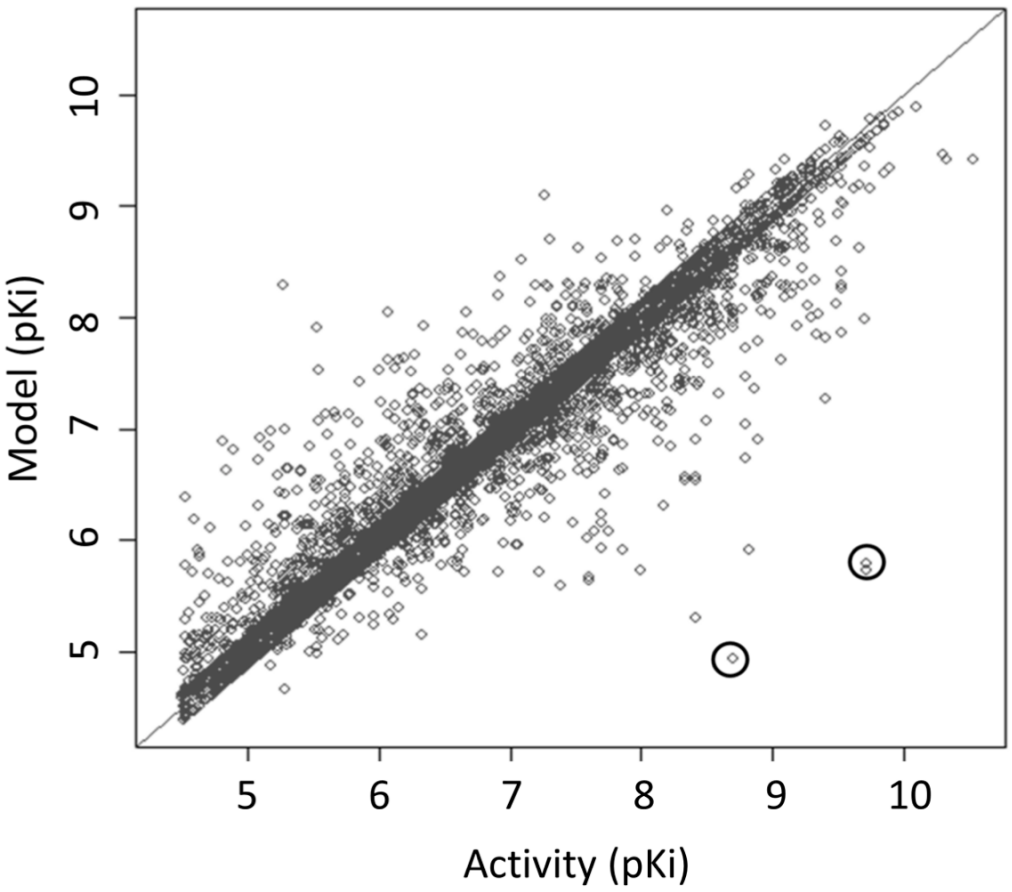


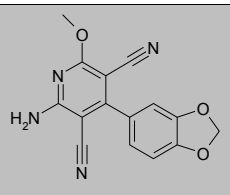
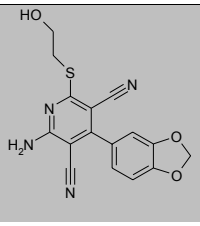
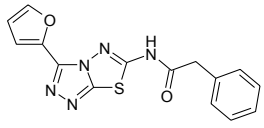
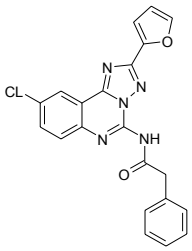
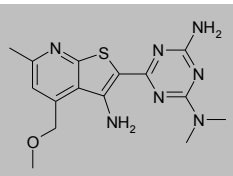
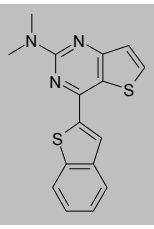
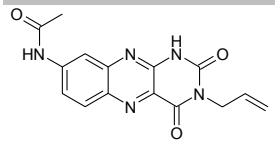
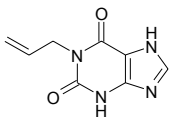
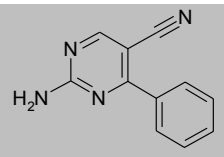
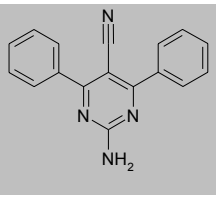
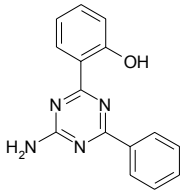
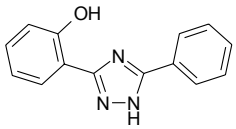
Figure 4.4: Cross validation plot of our final model correlating measured and predicted receptor affinities (pK_i values). The CV parameters were a Q^2 of 0.73 and a CV_RMSE of 0.61. The model fit had an R_0^2 of 0.95 and corresponding RMSE of 0.26. For an analysis of the outliers in the black circles see section **4.3.6**.

4.3.7 *In silico* model validation. Before applying our model in any virtual screening set-up, we performed several computational validation steps to ensure model predictivity and to prevent chance correlations from occurring. The learning curves (Supporting **Figure S6**) showed the maximal performance obtained was a prediction error of 0.62 log units (and corresponding R_0^2 of 0.71). In addition learning curves generated based on only the chemical space (Conventional structure-activity models rather than PCM models) showed that PCM is better able to model the ligand – target affinity than conventional single-target bioactivity models (Supporting **Figure S7** and Supporting **Figure S8**). The final model showed fair performance in external validation (Supporting **Figure S9**). It should be noted that our external validation consisted of compounds only tested on the human receptors. Interestingly the RMSE improved from 0.88 log units, when rat data was excluded from model training, to 0.82 log units, when these data were included (with the R_0^2 improving from 0.23 to 0.28).

Furthermore the model showed good performance in the decoy validation, since 33 of the 43 known actives were in the top 50 retrieved from 4,556 decoys (Supporting **Figure S10)**, with a runtime of 43 seconds. The highest predicted compound was LUF5957 with a predicted pKi of 9.02 (hA_1) and an experimentally determined pKi of 9.14 (hA_1). See Supporting **Table S9** for the structures of the four highest predicted decoys at rank 15, 21, 24 and 26. The 100-fold y-scrambled models plot shows a negative intersect with the Y-axis for both the R_0^2 and Q^2 regression lines as suggested to be characteristic for a predictive model by Eriksson *et al.* (Supporting **Figure S11**).[25]

The results from the different experiments show that the model appears to be statistically sound in nature. Several conclusions can be drawn from these results already. Firstly, it is difficult to train a model on public data gathered from a multitude of assays performed in different labs (also shown by Kramer *et al.*).[26] Secondly, our model is not based on chance correlations and has predictive power. Finally the pooling of data points from testing on rat receptors with data points from testing on human receptors has a positive effect on model performance. The RMSE improved from 0.88 to 0.82 upon inclusion of rat bioactivity data, likely by the inclusion of a much larger chemical space. Given the satisfying performance of our final model, we employed it in the next step to select novel potential adenosine receptor ligands from a chemical supplier, namely ChemDiv.

Table 4.1. Structures of the newly identified human adenosine receptor ligands.

Structure	Ki (μM, SEM and LE in parentheses) or % displacement at 10 μM				Most similar compound in training set	Similarity to training set (receptor)
	A_1	A_{2A}	A_{2B}	A_3		
 ***1***	**0.51 (±0.089, 0.39)**	**31 (±6.7, 0.28)**	32%	21%		0.65 (hA_1)
 ***2***	44%	**5.1 (±0.36, 0.32)**	-8%	3%		0.47 (rA_1)
 ***3***	35%	**1.6 (±0.31, 0.33)**	-5%	33%		0.30 (hA_1)
 ***4***	**3.2 (±0.13, 0.33)**	42%	14%	-3%		0.60 (rA_1)
 ***5***	**0.90 (±0.15, 0.55)**	**0.16 (±0.026, 0.62)**	-11%	13%		0.80 (hA_{2A})
 ***6***	**0.0072 (±0.0020, 0.56)**	**0.043 (±0.016, 0.51)**	**0.22 (±0.014, 0.46)**	**0.44 (±0.0073, 0.44)**		0.68 (hA_{2A})

Receptor affinity as determined in radioligand binding studies is shown as K_i value in μM or % displacement at 10 μM. Between parentheses the SEM in μM and the ligand efficiency (LE, see section **4.3.10**) in shown in kcal / mol per heavy atom. Also shown is the most similar compound in the training set (and the receptor it was annotated to) calculated as Tanimoto Similarity using the SCFP_4 fingerprint. Both entirely novel and atypical bioactive compounds have been identified (structures 3 and 4), as well as a fragment-like compound (structure 5) and a ligand with nanomolar activity (structure 6).

4.3.8 *In vitro* model validation. In our final 'wet' experimental validation we ordered 54 compounds that were indicated as active by our model on one or more of the adenosine receptors (see supporting SD file and Supporting **Table S10**). These 54 compounds were subsequently tested on all four human adenosine receptors (216 data points). Out of the total of 54 compounds tested six compounds were novel active compounds for the adenosine receptors (displacement larger 50% at a concentration of 10 µM; corresponding to a hit rate of 11%). Among the compounds were both selective ligands and highly active binders. For all six compounds active on either the human A_1 or human A_{2A} receptor in single-dose experiments full displacement curves were recorded, yielding K_i values. Furthermore, the pseudo Hill-coefficient was determined using variable slope regression in Graphpad Prism.[27] (The pseudo Hill-coefficients are listed in the supporting information along with all dose-response curves.)

Very diverse chemistry can be identified among the ligands found by our PCM model which are shown in **Table 4.1**. Two of the hits we found (compounds **1** and **2**) have a structure that resembles structures of known adenosine receptor ligands. However compound **3** and **4** have a structure that is not typical for compounds that are active on the adenosine receptors. Compound **5** shows a high affinity (0.90 µM on the human A_1 receptor and 0.30 µM on the human A_{2A} receptor), even though it is a very small fragment-like compound (MW 196). Finally, compound **6** even reached nanomolar affinities, even though no modifications or optimizations were performed on this compound. Note that the Tanimoto similarity to the training set based on the SCFP_4 fingerprint is as low as 0.30 in the case of compound **3** and reaching a maximum of 0.80 in the case of compound **5**. Furthermore, for two of the identified hits, the compound that is most similar in the training set has been annotated on the rat (A_1) receptor, further underlining the added value of the combination of human and rat orthologs. For additional details concerning the average and minimal similarity of the identified hits please see supporting **Table S11.** Shown in **Figure 4.5** are the curves used to determine the affinity of compound **6** on the human A_1 receptor. The full set of curves is contained in the supporting information.

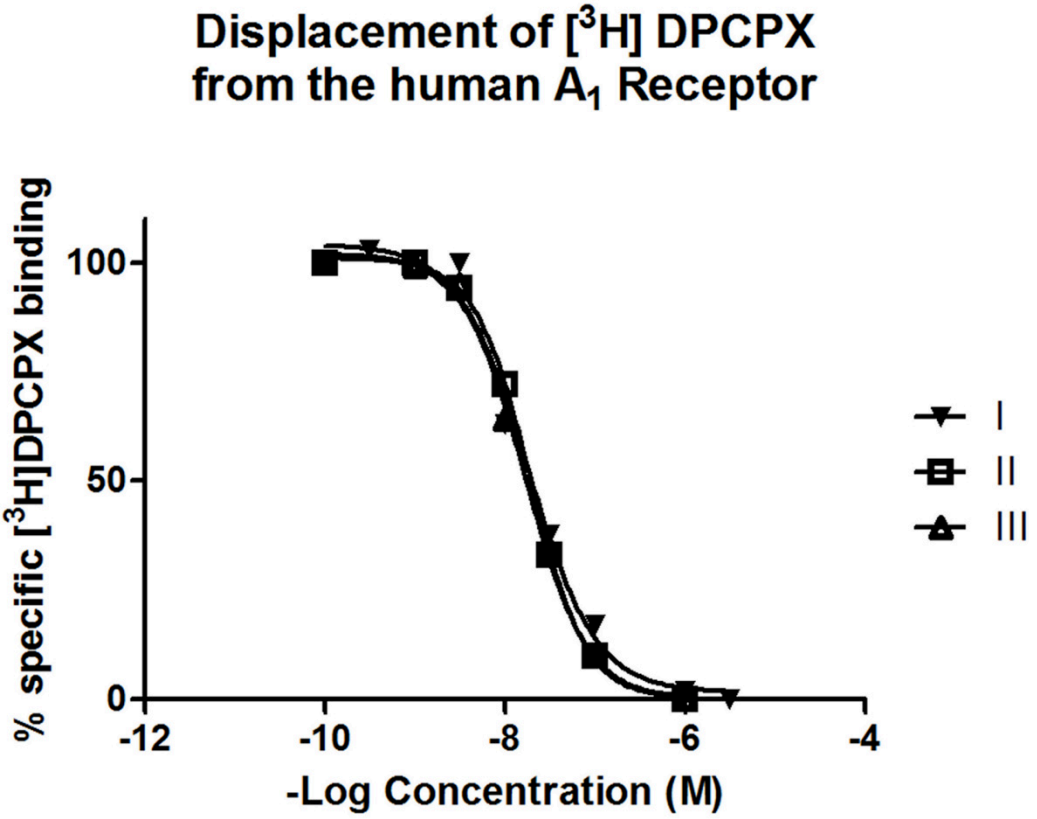


Figure 4.5: Typical dose response curve obtained during the in vitro model validation. Shown here are the dose response curves for compound **6** on the human A_1 receptor. The three curves performed in duplicate were obtained on different days. The pseudo Hill -coefficient was determined at -1.3 (± 0.1).

4.3.9 Implications on PCM performance. Since compounds **3** – **5** do not have a typical adenosine receptor template structure, we conclude that the PCM models obtained in this work are able to explore novel regions of bioactive chemical space. The ability of the model to find novel compounds is very likely the results of the larger chemical space covered in the training set in comparison with a conventional structure-activity model. (For a comparison of PCM and conventional structure-activity learning curves see Supporting **Figure S7** and Supporting **Figure S8**.) Together with the improved performance in the experimental validation, we show here the advantage PCM has due to its ability to characterize the full ligand – target interaction space.

However, while the PCM technique should in theory be able to predict bioactivity spectra, our experimental results indicate that our current model could not do so on the current data set. We were able to find active compounds and also selective compounds, but the compounds did not show selectivity as predicted by our model.

Furthermore, only one compound was found active on the human A_3 receptor, despite the fact that the model initially identified a much large number of compounds to have a pKi larger than 7.0 on the human A_3 receptor than on the other thee human receptors (see section **4.5.9** to **4.5.12** for details about compound selection). It is likely that this indicates that the model is not able to accurately model the bioactivity space for this receptor, in particular when we consider the large dissimilarity to the rat A_3 receptor.[1] The large dissimilarity combined with the low hit rate on the human A_3 receptor could indicate that the binding site definition is inaccurate. However it should be noted that this definition was based on only a single adenosine A_{2A} receptor crystal structure. As there are now more than a dozen GPCR crystal structures available, perhaps these can be used to better define the ligand binding site.

These two observations about the performance of the PCM model, the hit rate of 11% and the low performance for the human A_3 receptor, serve to illustrate that bioactivity models, like this model, are mainly a tool to *assist* in the process of medicinal chemistry. However this tool can be a very powerful tool as illustrated by the discovery of novel active compounds in the current work.

While this manuscript was completed, Langmead *et al.* published a structural virtual screening approach applied to the human adenosine A_{2A} receptor, identifying one out of 10 hits similar to compound **6,** subsequently optimized to be selective (for the structures see Supporting **Figure S12**).[28] It is interesting to see that we were able to identify a similar hit without the need for structural information.

4.3.10 Ligand Efficiency. Two of the identified novel ligands (**5** and **6**) showed a submicromolar affinity for (some of) the adenosine receptor subtypes. After calculating the ligand efficiency (LE),[29] we found that these two compounds both have a ligand efficiency higher than 0.5 kcal / mol per heavy atom on both the human A_1 and A_{2A} receptors. Furthermore, with the exception of compound **1** on the human A_{2A} receptor, all compounds have an LE higher than 0.30. Previously it has been shown that an LE of in the range of 0.30 – 0.40 constitutes a good value for lead optimization.[30-32] From the training set we also calculated the average LE (and standard deviation of this average) for ligands for each of the receptors, which was around 0.34 (supporting **Table S12**). These two compounds have a much higher LE, which renders them good starting points for the synthesis of a novel series now being pursued by our group.

4.3.11 PCM versus similarity searching. To place the performance of our PCM model in a broader context several similarity searching experiments were performed (supporting **Table S13**). To find 4 of the 6 hits, all compounds with a maximal (Tanimoto) similarity of 0.60 or higher should have been ordered, in ChemDiv this would have been 900 data points. However, the identification of all 6 hits would have required the purchase of all compounds with a maximal similarity of 0.30 or higher, a total of approximately 202,712 data points (on average approximately 50,000 compounds per human receptor). Moreover, the similarity searching was considerably slower than application of the PCM model. While the PCM model takes training time before it can be applied (3 hours on a Core i7 at 2.8 GHz with 16 GB of RAM), application afterwards is very quick, screening the full 791,162 compounds on all 4 receptors in 3 hours and 29 minutes (30 minutes for the filtered set; 100,910 compounds 15% of the total) using 6 threaded parallelization. Virtual screening using similarity searching on the same machine of all 791,162 compounds on the four human receptors took 71 hours and 29 minutes with a total time of 9 hours and 5 minutes for the filtered set (6 threaded calculation parallelization). The reason likely is that the PCM based approach requires a single calculation per data point (compound – receptor pair) whereas for the similarity searching, each compound requires the calculation of between 780 (hA_{2B}) and 1,661 (hA_3) Tanimoto similarities (indeed screening for the hA_{2B} receptor was considerably faster than the hA_3 receptor).

Likewise, we performed the decoy validation using a simple (Tanimoto) similarity based method (compounds ranked by maximal similarity to the training sets). Here we found that 40 of the known actives were in the top 50, with the highest predicted decoys at rank 35, 36, 38 and 43 (Supporting **Table S9**). However, while the PCM does not perform significantly better, the runtime for the similarity based approach was 6 minutes and 42 seconds (almost 10 times as long). In addition, similarity searching will not identify novel structures, which was the goal of this work.

These two similarity searching experiments demonstrate the added value of PCM as it displays a better performance and enrichment in prospective virtual screening combined with a significantly faster screening performance.

4.4 Conclusions

In this work we employed proteochemometric modeling (PCM) in order to identify novel human adenosine receptor ligands. By merging human and rat bioactivity data, we were able to identify six novel compounds that bind to members of the adenosine receptor family. One of these identified hits is very similar to a compound that was published recently (while we finalized this manuscript) using a structural rather than statistical approach. These novel ligands had an average Tanimoto similarity of 0.58 to the training set (ranging between 0.30 and 0.80). From the results we obtained we conclude that PCM is capable of capturing the full ligand – target space of a receptor subfamily rather than a single target. We showed that the addition of chemical and target information from orthologs can improve model quality when compared to creating a model based on a single species (prediction error decreased by 0.06 log units on this dataset). The ligand – targets spaces of human and rat adenosine receptors should not be regarded as separate entities and these spaces in fact overlap.

With the emergence and growth of large public databases such as ChEMBL,[15] PDB,[33] and Pubchem,[34] the PCM approach is likely to gain even further in momentum. In addition, the flexibility of this method may allow its application to other areas of drug discovery, such as receptor deorphanization or to different target families, such as the prediction of bioactivity profiles against kinases.

4.5 Experimental Section

4.5.1 Methods Overview. A flowchart of the modeling performed in this work is shown in **Figure 4.6**. The complete work can be divided in four major sections. Firstly, we created six different protein descriptors by varying the residue selection used to obtain them. Secondly, we created 16 different ligand fingerprints and also varied the maximal bond lengths in the substructures. Here we used a maximal diameter of either 4 or 6 bonds from a central atom. The third step consisted of finding the optimal combination of parameters for the training of the final model. These parameters included the different descriptors, method of cross validation and extended validation. The fourth and final step was the actual screening experiment where we combined both virtual screening and experimental validation. In the following section the individual procedures within each of these four blocks will be described starting with the descriptors for both the ligand and the receptors.

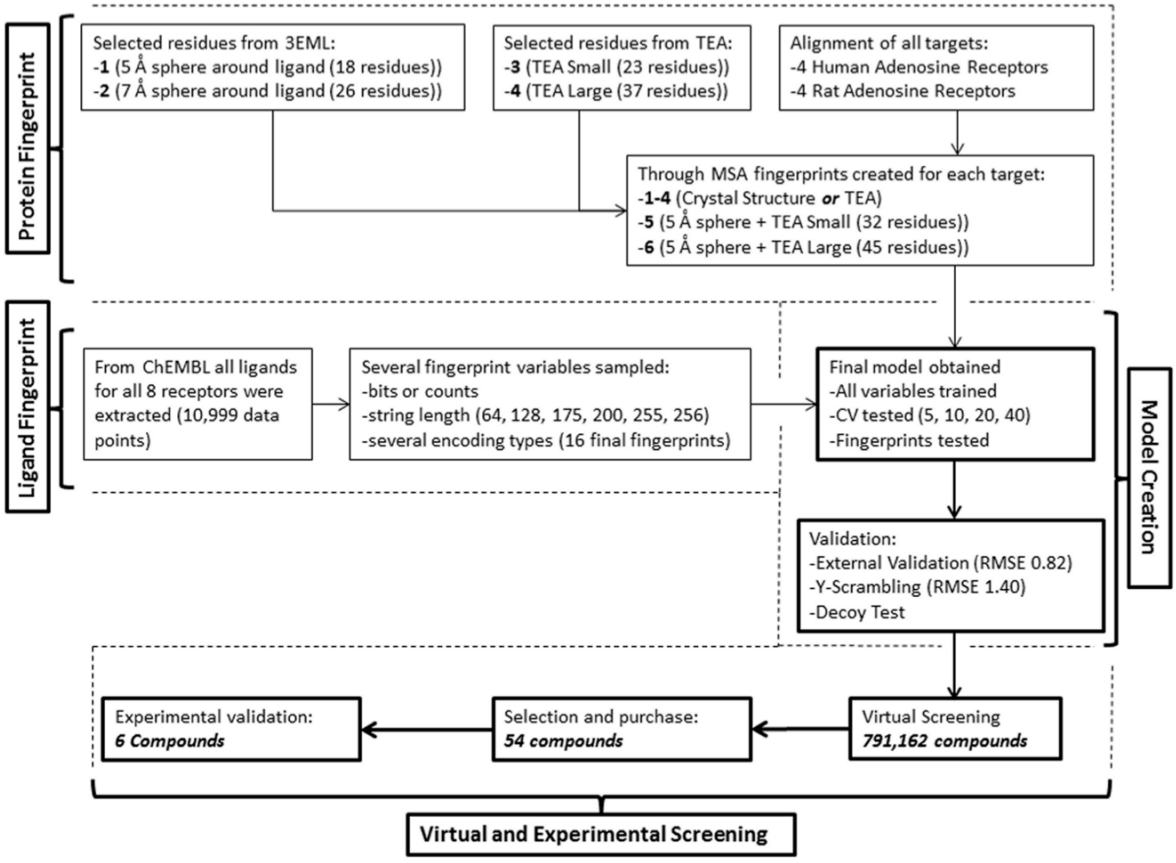


Figure 4.6: Flowchart of the work we performed.

4.5.2 Data set. The data set was obtained from ChEMBL 2.[15] From this database we selected all compounds that were tested on either human or rat adenosine receptors or both (Supporting Table **S2**). The selection was further narrowed to only include compounds for which a Ki value obtained from a radioligand binding assay was available. After selection the compounds were normalized and ionized at pH 7.4, they were assigned 2D coordinates and subsequently converted to fingerprints. All steps of this work were performed in Pipeline Pilot Student Edition version 6.1.5.[35]

The receptor sequences were obtained from Uniprot and aligned using ClustalW (Slow alignment, Gap Open 4, Gap Extend 4, available as supporting information).[36] This alignment was used to convert residues selected from the crystal structure to their ortholog and paralog counterparts. The residues selected by the TEA approach are provided in Ballesteros-Weinstein numbers and could be used directly.[37] After selection of the residues they were converted to a feature based protein fingerprint based on their single letter amino acid codes as we have done in previous work.[10]

4.5.3 Descriptor Benchmarking Approach. Before we could train our final predictive model, we sampled a multitude of parameters. We collected 6 selection methods to define our binding site residues (Supporting **Figure S3** and **Table S3**), 16 different types of circular fingerprints (Supporting **Figure S4** and **Table S4**), and 4 different folds of cross validation (CV) (Supporting **Figure S5**). From these options we wanted to select the optimal combination of variables. To identify the best combination, all models were built on 70% of the dataset (7,749 data points) and validated on the remaining 30% (3,250 data points). From the learning curves we already knew that this split was the optimal partition (Supporting **Figure S6**).

4.5.4 Protein descriptors. Sequences were encoded based on the binding site sequence in which each amino acid was represented as a single unique feature as was done in previous work.[10] However, these residues were selected in seven different ways and each selection was tested to find the best option to be used in the final model (Supporting **Figure S3**). The first two selection methods (1 and 2 in **Figure 4.6** were based on the crystal structure of the adenosine A_{2A} receptor bound to ZM241385 (PDB code 3EML).[19] Herein all residues were selected having any atom within either a 5 Å or a 7 Å sphere around the co-crystalized ligand. The third and fourth selection methods (3 and 4 in **Figure 4.6**) were based on a bioinformatics approach known as Two Entropy Analysis (TEA).[20] This method relies on quantifying the degree to which trans-membrane (TM) residues are conserved among Class-A GPCRs. Both the degree of conservation among GPCR subfamilies and the degree of conservation among the whole family are calculated. This calculation then serves as a basis to differentiate the function residues perform in individual GPCRs based on the difference between these degrees of conservation. Here we used a conservative approach (TEA S), which was small, and a less restricted selection method (TEA L), which was larger and included some of the residues from the 'mixed region' mentioned in the original publication.[20] The fifth method (TEA S5, 5 in **Figure 4.6**) was based on a combination of 1 one and 3, the sixth method (TEA L5, 6 in **Figure 4.6**) was based on a combination of method 1 and 4. Finally the seventh method consisted of simply using al TM residues. During this optimization, the ligand descriptor ECFP_4 was used (see **4.5.5** for further details).

The features describing the binding site were obtained by hashing an array of 58 physicochemical properties obtained from the AAindex database;[38] the indices employed can be found in Supporting **Table S14**. Finally, protein fingerprints were converted to an array of 175 features (Supporting **Figure S13**) which were then used in the modeling using Pipeline Pilot version 8.5.[23]

4.5.5 Compound descriptors. All descriptors were calculated in the academic version of PipelinePilot 6.1.5.[35] In the final model, ligands were described by Scitegic circular fingerprints (SCFP_4 type),[39, 40] which have previously been shown to capture a large amount of information with respect to compound bioactivity.[21] SCFP_4 descriptors provide individual substructures and treat these as a feature of a compound. We found them to perform the best and most consistent (Supporting **Figure S4**). These substructures have a maximal diameter of 4 bonds from a central atom. Finally ligand fingerprints were converted to an array of 175 features which were then used in the modeling (Supporting **Figure S13**).

4.5.6 Machine learning. Models were constructed in the academic version of Pipeline Pilot 6.1.5 using the R-statistics package. Support vector machines (SVM) as coded in the e1071 package were used for model creation.[41] Parameters gamma and cost were tuned over an exponential range and epsilon was set at 0.1. The optimal model was determined using cross validation before proceeding to experimental prospective validation of the model. The parameters used for validation were R_0^2, R^2, and RMSE.[42, 43]

4.5.7 *In silico* validation. We performed four different *in silico* validation experiments. Firstly a learning curve was generated, to spot possible discontinuous randomized splits data points, prevent overtraining and obtain an estimate of maximal performance that can be obtained on this dataset (Supporting **Figure S6**). Secondly, the obtained final model was subjected to external validation, applying the model to previously unseen compounds not part of our training set (**Supporting Figure S9**). Thirdly, the model was applied to a decoy set validation, to check performance in identifying unseen known actives from decoys (Supporting **Figure S10**).

This decoy set consisted of random selection of compounds from the ZINC database selected to resemble adenosine receptor ligands based on their physicochemical properties. These properties included molecular weight, number of hydrogen bond donors / acceptors, number of aromatic rings, calculated AlogP, average bond length, number of atoms, number of rotatable bonds, and formal charge (Supporting **Figure S14**). Finally y-scrambling was performed to estimate the possibility of chance correlations (Supporting **Figure S11**).

4.5.8 Virtual screening. Subsequent to our model validation we performed a virtual screening. We screened all compounds available in the ChemDiv database obtained *via* ZINC (accessed December 3rd 2009 consisting of 791,162 compounds) without any form of pre-filtering as we wanted a fair estimate of true model performance.[44] The main advantage of our statistical method is its throughput; typically one can screen the full ChemDiv database within 4 hours on a desktop machine (core i7 at 2.8 GHz with 16 GB RAM). In this case we screened 791,162 compounds on the four human Adenosine receptors (3,164,648 data points).

4.5.9 Selection Filters. On the resulting model output, several filters were applied: molecular solubility larger than -4 (solubility expressed as logS with S in mol/L),[45] AlogP between -0.4 and 5.6.[46] This reduced the number of data points from 3,164,648 to 403,640. The filtering could also very well have been applied before screening, but we wanted to see if our technique was capable of screening such a large number of compounds in a reasonable time. While more than 400,000 data points still represented a significant number the next step was binning and diversity clustering based on the chemistry of the compounds.

4.5.10 Binning. Subsequent to the classification ranking, compounds were binned in the following classes: predicted pKi between 5.0 and 6.0 (Bin 1; 180,419 data points), predicted pKi between 6.0 and 7.0 (Bin 2; 19,314 data points) and predicted pKi larger than 7.0 (Bin 3; 2,875 data points). The remainder was predicted to have a pKi smaller than 5.0 (201,032 data points) and were discarded together with those compounds not meeting the earlier physicochemical filters.

4.5.11 Clustering. Clustering was subsequently performed using the pipeline pilot component 'Cluster Molecules' on the individual bins with the aim of creating subsets containing different chemistry (Bin 1: 10 clusters for each receptor, Bin 2: 9 clusters for each receptor and Bin 3: 6 clusters for each receptor). The descriptor used was identical to the one we used to train the final models and the similarity coefficient used as the Tanimoto coefficient. For details about the obtained clusters see Supporting **Table S15** to **Table S18**. Bin 3 was to serve as a pool to select compounds for experimental validation. Note that Bin 3 is much larger in the case of the human adenosine A_3 receptor, this could indicate that the model is better able to find high affinity compounds for the human A_3 receptor. However, it is more likely that this indicates that the model is not able to accurately model the bioactivity space for this receptor, in particular when we consider the previously low hit rate in other studies on that receptor and the large dissimilarity to the rat A_3 receptor.

4.5.12 Final compound selection. Finally, compounds to be ordered were selected manually from Bin 3 for each receptor (predicted pKi larger than 7) with a focus on the selection of novel chemotypes. In total 54 compounds were selected. (Supporting **Table S10**, and supporting SD file).

These compounds were selected from different clusters (these clusters are indicated in bold in Supporting **Tables S15**-**S18** and are also present in the supporting SD file). The ignored clusters represented clusters that contained the remaining compounds (junk clusters) in the case of the human A_1, human A_{2A} and human A_3 receptor. In the case of human A_{2B} receptor the ignored cluster contained a number of compounds that were chemically very similar to compounds already selected for the hA_1 and hA_{2A} receptor (and were hence already going to be tested in the hA_{2B} receptor).

The compounds selected included both compounds that were predicted to be active on multiple receptors (like compound 5, predicted to be active on all 4 human receptors) and compounds predicted to be selective (like compound 3 predicted to be active on hA_{2A} and hA_{2B}). These compounds were subsequently ordered and tested *in vitro* on all receptors (216 data points), one compound (55 in Supporting **Table S10**) could not be ordered as it was unavailable from the supplier. 1H NMR and MS data are included in the Supporting Information for the found hits.

4.5.13 Ligand efficiency. Ligand efficiency (LE),[29, 31, 47] expressed in kcal / mol per heavy atom, was calculated according to eq. 1.

$$LE = \Delta G / N_{Non\text{-}hydrogen\ atoms} \quad (1)$$

To obtain ΔG we used eq. 2. ΔG was converted to kcal / mol.

$$\Delta G = -RT.\ln K_i \quad (2)$$

4.5.14 Similarity searching. All similarity searching experiments were performed in Pipeline Pilot version 8.5.[23] The Pipeline pilot similarity searching component was used and the search was done using SCFP_4 fingerprints. The component was optimized for speed rather than memory use and screening was done in parallel using 6 threads on a core i7 machine. For each receptor subtype, the subset of the training set regarding that subtype was used as reference compounds. For example, to identify similar compounds for the human A_1 receptor, we used the human A_1 receptor annotated compounds from ChEMBL 2.

4.5.15 Binding Studies. [3H]DPCPX and [3H]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. [3H]PSB603 and [3H]PSB11 were kind gifts from Prof C.E. Müller (Bonn, Germany). Chinese Hamster Ovary (CHO) cells expressing the human adenosine A_1 receptor were provided by Dr. Andrea Townsend-Nicholson, University College of London, UK. Human embryonic kidney (HEK) 293 cells stably expressing the human adenosine A_{2A} and human 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). Dose response curves for the found hits are included in the Supporting Information.

4.5.16 Human adenosine A_1 Receptor. Affinity at the A_1 receptor was determined on membranes from CHO cells expressing the human receptors, using [3H]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 [3H]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.

4.5.17 Human adenosine A_{2A} Receptor. At the A_{2A} receptor, affinity was determined on membranes from HEK 293 cells stably expressing this human receptor, using [3H]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 [3H]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 CGS21680. 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.

4.5.18 Human 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 [3H]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 [3H]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.

4.5.19 Human 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 [3H]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 [3H]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.

4.5.20 Data Analysis. K_i values were calculated using a nonlinear regression curve-fitting program (GraphPad Prism 5.0).[27] K_i values of radioligands were 1.6, 1.7, 0.41, and 4.9 nM for [^{3}H]DPCPX, [^{3}H]ZM241385, [^{3}H]PSB603, and [^{3}H]PSB11, respectively.

4.6 Supporting Information

Additional tables (Supporting **Tables S1 – S18**), figures (**Figures S1 – S14**), compound purity information (1H NMR, LC-MS and HR MS spectra), radioligand displacement curves for the found hits, the final model, an SD file with the ordered compounds, the multiple sequence alignment used to create the protein descriptor and a protocol to run the model are available online. These materials are available online at www.gjpvanwesten.nl.

4.7 Acknowledgments

GvW would like to thank Tibotec BVBA for funding. The authors would like to thank Jacobus van Veldhoven and Maris Vilums for help in analytical chemistry and Laura Heitman for the helpful discussions.

4.8 References

1. B.B. Fredholm, A.P. IJzerman, K.A. Jacobson*, et al.*; *International Union of Basic and Clinical Pharmacology. LXXXI. Nomenclature and Classification of Adenosine Receptors—An Update.* Pharmacol. Rev.; *2011*. **63** (1): 1-34.
2. R.M. Hyde and D.J. Livingstone; *Perspectives in QSAR: Computer chemistry and pattern recognition.* J. Comput.-Aided Mol. Des.; *1988*. **2**: 145-155.
3. K. Roy; *QSAR of Adenosine Receptor Antagonists II.* QSAR Comb. Sci.; *2003*. **22** (6): 614-621.
4. A. Bender and R.C. Glen; *Molecular similarity: a key technique in molecular informatics.* Org. Biomol. Chem.; *2004*. **2**: 3204-3218.
5. A. Kontijevskis, R. Petrovska, I. Mutule*, et al.*; *Proteochemometric analysis of small cyclic peptides' interaction with wild-type and chimeric melanocortin receptors.* Proteins: Struct., Funct., Bioinf.; *2007*. **69** (1): 83-96.
6. M. Lapinsh, P. Prusis, A. Gutcaits*, et al.*; *Development of proteo-chemometrics: a novel technology for the analysis of drug-receptor interactions.* Biochim. Biophys. Acta, Gen. Subj.; *2001*. **1525** (1-2): 180-190.
7. G.J.P. Van Westen, J.K. Wegner, A.P. IJzerman*, et al.*; *Proteochemometric Modeling as a Tool for Designing Selective Compounds and Extrapolating to Novel Targets.* Med. Chem. Commun.; *2011*. **2** (1): 16-30.
8. J. Wikberg, M. Lapinsh, and P. Prusis; *Proteochemometrics: A tool for modelling the molecular interaction space;* in *Chemogenomics in Drug Discovery - A Medicinal Chemistry Perspective*; H. Kubinyi and G. Müller; Editors. *2004*. p. 289 - 309.
9. M. Lapinsh, P. Prusis, S. Uhlen*, et al.*; *Improved approach for proteochemometrics modeling: application to organic compound - amine G protein-coupled receptor interactions.* Bioinformatics; *2005*. **21** (23): 4289-4296.
10. G.J.P. Van Westen, J.K. Wegner, P. Geluykens*, et al.*; *Which Compound to Select in Lead Optimization? Prospectively Validated Proteochemometric Models Guide Preclinical Development.* PLoS One; *2011*. **6** (11): e27518.
11. N. Weill and D. Rognan; *Development and Validation of a Novel Protein−Ligand Fingerprint To Mine Chemogenomic Space: Application to G Protein-Coupled Receptors and Their Ligands.* J. Chem. Inf. Model.; *2009*. **49** (4): 1049-1062.
12. B.B. Fredholm, A.P. IJzerman, K.A. Jacobson*, et al.*; *International Union of Pharmacology. XXV. Nomenclature and Classification of Adenosine Receptors.* Pharmacol. Rev.; *2001*. **53** (4): 527-552.

13. R. Guha and J.H. VanDrie; *Structure-Activity Landscape Index: Identifying and Quantifying Activity Cliffs.* J. Chem. Inf. Model.; *2008*. **48** (3): 646-658.

14. M.T. Sisay, L. Peltason, and J.r. Bajorath; *Structural Interpretation of Activity Cliffs Revealed by Systematic Analysis of Structure-Activity Relationships in Analog Series.* J. Chem. Inf. Model.; *2009*. **49** (10): 2179-2189.

15. A. Gaulton, L.J. Bellis, A.P. Bento*, et al.*; *ChEMBL: a large-scale bioactivity database for drug discovery.* Nucleic Acids Res.; *2011*. **40**: D1100 - D1107.

16. F.A. Kruger and J.P. Overington; *Global Analysis of Small Molecule Binding to Related Protein Targets.* PLoS Comput. Biol.; *2012*. **8** (1): e1002333.

17. J.E.S. Wikberg, F. Mutulis, I. Mutule*, et al.*; *Melanocortin receptors: Ligands and proteochemometrics modeling;* in *Melanocortin System*; D. Braaten; Editor *2003*: New York. p. 21-26.

18. M. Lapinsh, P. Prusis, T. Lundstedt*, et al.*; *Proteochemometrics modeling of the interaction of amine G-protein coupled receptors with a diverse set of ligands.* Mol. Pharmacol.; *2002*. **61** (6): 1465-1475.

19. V.P. Jaakola, M.T. Griffith, M.A. Hanson*, et al.*; *The 2.6 Angstrom Crystal Structure of a Human A2A Adenosine Receptor Bound to an Antagonist.* Science; *2008*. **322** (5905): 1211-1217.

20. K. Ye, E.W.M. Lameijer, M.W. Beukers*, et al.*; *A two-entropies analysis to identify functional positions in the transmembrane region of class A G protein-coupled receptors.* Proteins: Struct., Funct., Bioinf.; *2006*. **63** (4): 1018-1030.

21. A. Bender, J.L. Jenkins, J. Scheiber*, et al.*; *How Similar Are Similarity Searching Methods? A Principal Component Analysis of Molecular Descriptor Space.* J. Chem. Inf. Model.; *2009*. **49** (1): 108-119.

22. K. Baumann; *Cross-validation is dead. Long live crossvalidation! Model validation based on resampling.* J. Cheminf.; *2010*. **2** (Suppl 1): O5.

23. Accelrys Software Inc *Pipeline Pilot Professional Edition* Scitegic Version 8.5

24. K.A. Jacobson, P.J.M. Van Galen, and M. Williams; *Adenosine receptors: pharmacology, structure-activity relationships, and therapeutic potential.* J. Med. Chem.; *1992*. **35** (3): 407-422.

25. L. Eriksson, J. Jaworska, A.P. Worth*, et al.*; *Methods for reliability and uncertainty assessment and for applicability evaluations of classification-and regression-based QSARs.* Environ. Health Perspect.; *2003*. **111** (10): 1361-1375.

26. J.R. Lane, C. Klein Herenbrink, G.J.P. Van Westen*, et al.*; *A Novel Nonribose Agonist, LUF5834, Engages Residues That Are Distinct from Those of Adenosine-Like Ligands to Activate the Adenosine A2a Receptor.* Mol. Pharmacol.; *2012*. **81** (3): 475-487.

27. GraphPad Software Inc *GraphPad Prism* 5.0

28. C.J. Langmead, S.P. Andrews, M. Congreve*, et al.*; *Identification of Novel Adenosine A2A Receptor Antagonists by Virtual Screening.* J. Med. Chem.; *2012*. **55**: 1904 - 1909.

29. P.R. Andrews, D.J. Craik, and J.L. Martin; *Functional group contributions to drug-receptor interactions.* J. Med. Chem.; *1984*. **27** (12): 1648-1657.

30. A.L. Hopkins, C.R. Groom, and A. Alex; *Ligand efficiency: a useful metric for lead selection.* Drug Discov. Today; *2004*. **9** (10): 430-431.

31. D. Tanaka, Y. Tsuda, T. Shiyama*, et al.*; *A Practical Use of Ligand Efficiency Indices Out of the Fragment-Based Approach: Ligand Efficiency-Guided Lead Identification of Soluble Epoxide Hydrolase Inhibitors.* J. Med. Chem.; *2010*. **54** (3): 851-857.

32. C. Abad-Zapatero and J.T. Metz; *Ligand efficiency indices as guideposts for drug discovery.* Drug Discov. Today; *2005*. **10** (7): 464-469.

33. H.M. Berman, J. Westbrook, Z. Feng*, et al.*; *The Protein Data Bank* Nucleic Acids Res.; *2000*. **28**: 235-242.

34. E.E. Bolton, Y. Wang, P.A. Thiessen*, et al.*; *PubChem: Integrated Platform of Small Molecules and Biological Activities;* in *Annual Reports in Computational Chemistry*; A.W. Ralph and C.S. David; Editors. *2008*; Elsevier. p. 217-241.

35. Accelrys Software Inc *Pipeline Pilot Student Edition* Scitegic Version 6.1.5

36. E. Jain, A. Bairoch, S. Duvaud*, et al.*; *Infrastructure for the life sciences: design and implementation of the UniProt website.* BMC Bioinformatics; *2009*. **10** (1): 136-155.

37. J.A. Ballesteros and H. Weinstein; *Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors;* in *Methods in Neurosciences*; C.S. Stuart; Editor *1995*; Academic Press. p. 366-428.

38. S. Kawashima, H. Ogata, and M. Kanehisa; *AAindex: Amino Acid Index Database.* Nucleic Acids Res.; *1999*. **27** (1): 368-369.

39. R.C. Glen, A. Bender, C.H. Arnby*, et al.*; *Circular fingerprints: Flexible molecular descriptors with applications from physical chemistry to ADME.* IDrugs; *2006*. **9** (3): 199 - 204.

40. D. Rogers and M. Hahn; *Extended-Connectivity Fingerprints.* J. Chem. Inf. Model.; *2010*. **50** (5): 742-754.

41. E. Dimitriadou, K. Hornik, F. Leisch*, et al. Misc Functions of the Department of Statistics (e1071)* TU Wien 2006 1.5-15

42. A. Tropsha, P. Gramatica, and Vijay K. Gombar; *The Importance of Being Earnest: Validation is the Absolute Essential for Successful Application and Interpretation of QSPR Models.* QSAR Comb. Sci.; *2003*. **22** (1): 69-77.

43. A. Tropsha; *Predictive Quantitative Structure-Activity Relationships Modeling;* in *Handbook of Chemoinformatics Algorithms*; J. Faulon and A. Bender; Editors. *2010*.

44. J.J. Irwin and B.K. Shoichet; *ZINC − A Free Database of Commercially Available Compounds for Virtual Screening.* J. Chem. Inf. Model.; *2005*. **45** (1): 177-182.

45. I.V. Tetko, V.Y. Tanchuk, T.N. Kasheva*, et al.*; *Estimation of Aqueous Solubility of Chemical Compounds Using E-State Indices.* J. Chem. Inf. Comput. Sci.; *2001*. **41** (6): 1488-1493.

46. A.K. Ghose, V.N. Viswanadhan, and J.J. Wendoloski; *Prediction of Hydrophobic (Lipophilic) Properties of Small Organic Molecules Using Fragmental Methods: An Analysis of ALOGP and CLOGP Methods.* J. Phys. Chem.; *1998*. **102** (21): 3762-3772.

47. I.D. Kuntz, K. Chen, K.A. Sharp*, et al.*; *The maximal affinity of ligands.* Proc. Natl. Acad. Sci. U. S. A.; *1999*. **96** (18): 9997-10002.