

Comparative Analysis of Putative Agonist-Binding Modes in the Human A₁ Adenosine Receptor

Hugo Gutiérrez-de-Terán,^[a] Manuel Pastor,^[a] Nuria B. Centeno,^[a] Johan Åqvist,^[b] and Ferran Sanz^{*[a]}

A recent study reported a model of the human A₁ adenosine receptor and its agonist binding site, proposing two putative binding modes in the same binding site for the natural agonist, adenosine. The present work investigates the flexibility of this binding site by exhaustive exploration with the natural agonist and with three other adenosine derivatives: N6-cyclopentyladenosine (CPA), 2-chloro-N6-cyclopentyladenosine (CCPA), and 5'-N-ethylcarboxamidoadenosine (NECA). Our aim was to find a common binding mode for agonists that would explain the role

in the binding process of the different substitutions allowed at the 2, N6, and 5' positions of adenosine. This problem was addressed through docking simulations, molecular dynamics studies, and estimations of the ligand-binding free energy with both the AUTODOCK scoring function and the linear interaction energy (LIE) approach. The results point to a single receptor-binding position that explains the effects of the different chemical modifications on the adenosine derivatives considered here.

Introduction

The nucleoside adenosine exerts its extracellular effects through coupling with transmembrane adenosine receptors (ARs), a family of G protein-coupled receptors (GPCRs). In the last few decades much effort has been made in pursuit of pharmacological understanding of the four adenosine receptor subtypes: named A₁, A_{2a}, A_{2b}, and A₃. Among them, the A₁ subtype is of particular interest, both because it is widely studied^[1] and because of the physiological functions in which it is involved, making it attractive as a potential drug target.^[2,3] In this respect, A₁AR agonists can be of therapeutic interest in different biological systems: in the CNS, this receptor is located on cortex, hippocampus, cerebellum, and thalamus, where its activation leads to neuroprotective effects through the blockade of neurotransmitter release.^[4] In the heart, A₁AR agonists cause negative chronotropic, dromotropic, and inotropic effects, which makes them attractive as potential cardioprotective and anti-infarct agents.^[5]

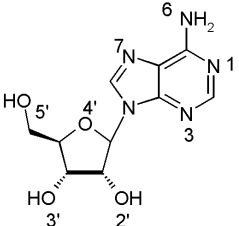
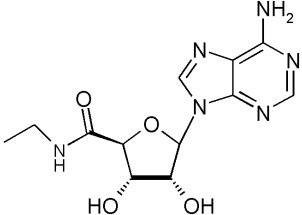
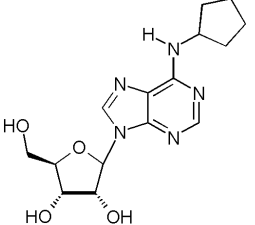
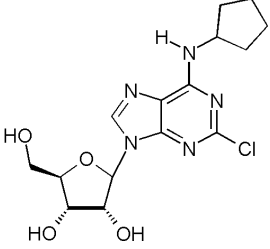
Selective and potent agonists for the A₁ receptor have been developed, providing tools for better understanding of its pharmacology. Agonists for ARs are chemically related to the natural agonist adenosine (see Table 1). The ribose moiety appears to be essential for both affinity^[2] and agonistic profile^[6] of adenosine derivatives. Single modifications on either the 2'- or the 3'-hydroxyl group lead to molecules that show a partial agonist profile,^[7,8] while removal of both substituents at the same time confers antagonistic properties.^[9] Modifications of the ribose moiety maintaining a full agonist profile are limited to 5'-uronamido derivatives, of which NECA (5'-N-ethylcarboxamidoadenosine) is a potent non-A₁-selective example.^[10] The introduction of small aromatic or cycloalkyl substituents at the N6 position leads to more potent and selective A₁ derivatives, such as N6-cyclopentyladenosine (CPA).^[2] Another substitution

position is the 2-position, where a chlorine atom in conjunction with N6 substitution yielded CCPA (2-chloro-N6-cyclopentyladenosine),^[11] the most potent agonist known for the A₁AR.^[12] Thus, molecules such as NECA, CPA, and CCPA have been used for years to study this receptor subtype (see refs. [6,13], and references therein). In order to investigate the location of the agonist binding site and the role of several amino acids known to influence ligand binding, we have recently developed a model for the human A₁ adenosine receptor^[14] (hA₁AR) based on the X-ray crystal structure of bovine rhodopsin.^[15] Using this model and the GROUP module in the GRID software package,^[16] we have exhaustively explored the putative ribose-binding site. A single solution was found, located between helices 1, 2, and 7 and composed of residues Ser1.46, Asp2.50, His7.43, and Ser7.46. The high polarity of this pocket is able to establish enough hydrogen bond (H-bond) interactions to balance the ribose desolvation energy, and also showed close similarity to other experimentally determined ribose-binding sites. This ribose-binding pocket was a starting anchoring point for the search for the binding position for the natural agonist, adenosine (ADO), carried out by docking and molecular dynamics (MD) simulations, and concluded that two adenosine-binding modes were possible, both involving heli-

[a] H. Gutiérrez-de-Terán, Dr. M. Pastor, Dr. N. B. Centeno, Prof. F. Sanz
Research Group on Biomedical Informatics (GRIB)
Institut Municipal d'Investigació Mèdica, Universitat Pompeu Fabra
Carrer Dr. Aiguader 80, 08003 Barcelona (Spain)
Fax: (+34) 93-224-0875
E-mail: fsanz@imim.es

[b] Prof. J. Åqvist
Department of Cell and Molecular Biology, Uppsala University, BMC
Box 596, 751 24 Uppsala (Sweden)

Table 1. A₁ adenosine receptor agonists studied.

Compound ^[a]	Name	A ₁ AR K _i [nM]	ΔG ^[d] [kcal mol ⁻¹]
	ADO	73 ^[b]	-9.8
	NECA	13.6 ^[c]	-10.8
	CPA	2.25 ^[c]	-11.9
	CCPA	0.83 ^[c]	-12.5

[a] Standard numbering of adenosine derivatives is shown for adenosine. [b] Inhibition of adenylate cyclase in rat A₁AR.^[10] [c] Values in human A₁AR extracted from ref. [12]. [d] ΔG calculated by use of K_i = e^{ΔG/RT} at 25 °C.

ces 1, 2, 3, 6, and 7 and almost the same residues. One binding mode had the adenosine amino group facing the extracellular side, interacting with Thr7.42, while in the other complex obtained the amino group was linked with another threonine (Thr3.36) located in helix 3, therefore facing towards the central core of the helical bundle. The adenine moiety was in each case stabilized by hydrophobic interactions with Val3.32, a conserved residue in ARs that corresponds to the aspartate that binds the basic nitrogen of the ligands in aminergic receptors.^[17] Both binding modes may explain mutagenesis data concerning residues such as Thr3.36,^[18] or His7.43,^[19,20] and consequently no criteria for determining the preferred binding mode were found in the previous study. This work reports a deeper and comparative analysis of such binding modes by considering the docking of other relevant A₁ agonists. This analysis is intended to explore structure–activity relationships in adenosine derivatives, identifying the role of the different chemical substitutions and their relative counterparts in the re-

ceptor. Moreover, the obtained docking model should be predictive for the differences in potency of the ligands studied. To make progress towards this goal, we made use of different computational simulation methods such as docking, molecular dynamics, and binding free energy calculations.

Computational Methods

Human A₁ adenosine receptor model: The previously reported^[14] human A₁ adenosine receptor model (hA₁AR) was used in these calculations. This model incorporates only the seven-helix bundle, discarding the loop regions. Briefly, each helix was separately built and geometrically optimized, followed by packing of the seven helices with use of the experimentally measured structure of bovine rhodopsin^[15] as a template. Finally, the model was energy-optimized and refined by molecular dynamics techniques. The generalized numbering scheme proposed by Ballesteros and Weinstein^[21] has been used to name the residues through the text. This nomenclature allows easier identification of the residues within the seven-helix bundle, as well as simpler association of equivalent residues of different GPCRs or of the same receptor in different species.

Modeling of the ligands: The four agonists studied—ADO, NECA, CPA, and CCPA—are depicted in Table 1, together with their experimentally determined binding data. The 3D structures of adenosine derivatives were built by modification of the crystallographic coordinates of adenosine (refcode ADENOS10 in the Cambridge structural database^[22]) with appropriate substituents by use of the Builder module in InsightII.^[23] Each molecule was then energy-minimized with the semiempirical AM1 Hamiltonian as implemented in MOPAC93.^[24] Final geometries and partial charges obtained in this step were used as input for AUTODOCK computations.

In order to identify necessary force-field parameters for the molecular dynamics and binding free energy calculations, we performed single-point ab initio calculations with the aid of the GAUSSIAN98 suite of programs,^[25] with use of the Hartree–Fock method and the 6-31G* basis set. The restrained electrostatic potential fitting procedure^[26] was applied on the calculated electrostatic potentials in order to obtain partial atomic charges compatible with the Amber Parm94^[27] force-field.

Docking studies: A docking exploration with AUTODOCK3.0^[28] was performed for each agonist examined. This program allows full flexibility in the ligands, while keeping the geometry of the receptor frozen. The exploration of docking positions was carried out through 100 runs of the Lamarckian genetic algorithm (LGA), with use of AUTODOCK default parameters. The resulting docking positions were clustered according to an RMS criterion of 1 Å. The force-field implemented in the program was modified by use of a united-atom model for both protein and ligand, and 12–10 hydrogen bond parameters for all the O...H and N...H interactions, whenever the involved nitrogen atoms have a lone pair. We restrained the docking exploration to the region containing the residues experimentally known to influence agonist binding, by defining a 20³ Å³ grid centered on Nδ of His7.43.

Molecular dynamics and binding free energy predictions: In addition to the empirical binding energies provided by AUTODOCK, free energies of binding were calculated by the linear interaction energy (LIE) method.^[29,30] This method is based on thermal conformational sampling of the ligand, both in the free state (i.e., solvated in water) and bound to the solvated protein. The estimated

energy of binding is calculated as a linear combination of the differences in the average ligand-surroundings interactions (where "surroundings" relates to protein and water when the bound state is considered, or just water if considering the free state). Interaction energies are split into electrostatic and van der Waals terms, and weighted by different factors:

$$\Delta G = \alpha \left(\langle V_{l-s}^{\text{vdW}} \rangle_{\text{p}} - \langle V_{l-s}^{\text{vdW}} \rangle_{\text{w}} \right) + \beta \left(\langle V_{l-s}^{\text{el}} \rangle_{\text{p}} - \langle V_{l-s}^{\text{el}} \rangle_{\text{w}} \right) + \gamma \quad (1)$$

Here the $\langle \rangle$ terms denote thermal averages sampled during molecular dynamics of the electrostatic (el) and van der Waals (vdW) energies for the ligand atoms in the protein (p) and water (w) environments. The coefficients in the above equation are taken as $\beta = 0.5$, when the linear electrostatic response approximation is valid,^[31] and $\alpha = 0.181$ for the nonpolar interactions. In our case, given that all ligands are neutral molecules with two or more hydroxyl groups (which has been shown to cause deviations from exact linear response^[32]), the parameter β was set at the previously determined value of 0.33.^[30] The constant parameter $\gamma \neq 0$ can be adjusted by least-squares optimization in order to improve the calculated absolute binding energies quantitatively with respect to the experimental ones. Such a constant term appears to be dependent on the nature of binding site studied, and this issue is discussed elsewhere.^[33] Here, we are mainly concerned with relative binding energies, and also have only a few compounds to work with, so a detailed optimization of this parameter is not warranted.

The program Q^[34] was used for the MD simulations and analysis, with employment of the Amber parm94 force-field^[27] implemented in the program. In both the "bound" and the "free" states, the ligand was solvated with a randomly oriented TIP3P^[35] water sphere of 18 Å radius, in which water molecules were confined to the sphere through a radial restraining potential.^[34,36,37] In the "bound" simulation, ionic groups of the protein within this sphere that are not too close to the boundary (Asp2.50 and Glu1.39) were modeled as charged. The resulting net charge for the sphere of simulation was -2 , since there were no Arg or Lys residues that could be ionized to counterbalance negative charges. Furthermore, neither of the histidines inside the simulation sphere (His7.43 and His6.51) are likely to be positively charged according to experimental data.^[38]

The solvent was relaxed by use of a strong thermal bath coupling ($\tau = 1$ fs), a temperature of 0 K, and restraint of solute atoms with a force constant that decreased from 50 to 25, 10, and 5 kcal mol⁻¹ Å⁻² in blocks of 0.8 ps each. This step, which is very similar to a steepest-descent energy minimization, was followed by a heating phase consisting of blocks of 6 ps at 50 K, 150 K, and 300 K, while the bath coupling was relaxed ($\tau = 5$ fs) and the 5 kcal mol⁻¹ Å⁻² force constant was maintained on all the protein and ligand atoms. In some cases, a set of five simulated annealing (SA) cycles was performed at this point, consisting of a 4 ps smooth ($\tau = 50$ fs) warming up until 900 K followed by a fast cooling down similar to energy minimization ($\tau = 5$ fs), in order to allow the system to explore conformational changes requiring crossing of high-energy barriers from the initial docking position. A harmonic constraint of 10 kcal mol⁻¹ Å⁻² was applied on the C α protein atoms in this step to keep the trace of the protein fixed while the ligand and protein side chains were relaxed.

Productive MD simulations for the data collection were performed at 300 K (coupling to the temperature bath was set to $\tau = 10$ fs) with a time step of 1.5 fs. No cut-off was applied to the ligand interactions, while a 10 Å cut-off was applied to other non-bonded interactions, together with use of the local reaction field ap-

proach^[39] for long-range electrostatic interactions beyond that cut-off. A 10 kcal mol⁻¹ Å⁻² force constant was imposed on the atoms in the boundary zone between 16.5 and 18 Å from the sphere center and, due to the absence of loops and the membrane, a force constant of 5 kcal mol⁻¹ Å⁻² was maintained for the alpha carbons inside the sphere. (Note that, due to the size of the protein, no water penetrates into the membrane region.) The rest of the protein inside that sphere and all ligand atoms were completely free to move, while the protein atoms outside the sphere of simulation were fixed by a high (100 kcal mol⁻¹ Å⁻²) harmonic constraint. Energy data were collected at this stage for 300–1500 ps starting from a moment at which the system could be considered equilibrated. The length of the sampling stage was such that the energy data could be split into two halves with equal average energies.

In the "free" simulation, the solvent was first relaxed by means of 0.2 ps MD through a strong bath coupling ($\tau = 0.2$ fs) followed by 15 ps (relaxing bath coupling to $\tau = 10$ fs) at temperature of 300 K and application of a harmonic constraint (10 kcal mol⁻¹ Å⁻²) on all ligand atoms. For the data collection, the central atom of the ligand was geometrically restrained to the center of the grid by use of a 50 kcal mol⁻¹ Å⁻² force constant, to ensure a homogenous solvation. Time step and bath coupling were set to the same values as in the "bound" state.

Analysis of the MD results was performed with the tools provided in the Q software package plus additional visual inspection with VMD.^[40]

Results and Discussion

The starting point for this study was the previously reported hA₁AR adenosine binding site that, given the relative symmetry of H-bond acceptor groups in adenosine derivatives,^[41] was able to accommodate two potential agonist-binding modes.^[14] In order to deal with this ambiguity, we picked three classic and potent full A₁AR agonists (NECA, CPA, and CCPA) for which experimental affinity constants had been determined for the human receptor, as well as the natural ligand, adenosine (see Table 1), and studied their binding to our hA₁AR model. This was done in a two-step protocol as follows. Firstly, an automatic docking exploration provided ideas about the most relevant binding positions for each compound. Then, each relevant position was further analyzed by MD and an estimation of the free energy of binding by the linear interaction energy (LIE) approach was performed.

Docking studies

An exhaustive docking exploration of the binding site with each of the four ligands considered was performed with the aid of AUTODOCK software.^[28] This program has been previously used for docking of flexible ligands into models of GPCRs,^[42–44] and its usefulness has been further tested by us in the experimental case of bovine rhodopsin.^[14]

Results of docking explorations are summarized in Table 2. For each of the ribose-unmodified ligands (ADO, CPA, and CCPA), two binding positions appear, in agreement with the re-

Table 2. Summary of the docking results for the four A₁AR agonists, showing the energy of binding predicted by the scoring function and the ranking of each solution among those provided by AUTODOCK.

Compound	Position A		Position B	
	ΔG	Rank	ΔG	Rank
ADO	-10.6	1 ^[a]	-10.4	3 ^[a]
NECA	-	-	-11.1	1
CPA	-10.7	6	-13.5	1
CCPA	-10.8	6	-12.7	1

[a] Docking solutions 1–3 for ADO are, in practice, energetically equivalent.

sults reported in our previous article. Both positions show the adenosine scaffold located in the cavity between helices 2, 3, and 7 and involve essentially the same residues, but playing different roles in agonist binding in each case. These docking positions are referred to from now on as “position A” and “position B” (see Figure 1). It should be pointed out that other docking positions were also found for each ligand, but that none of them showed consistency among the ligand series and/or agreement with mutagenesis data, thus comprising low-quality solutions not selected for further investigation. In position A, the docking positions of ADO and CPA show the 5'-hydroxyl of adenosine interacting with the carboxylic acid of Asp2.50, whereas the OH3' and OH2' groups are surrounded by Ser1.46, His7.43, and Ser7.46, and the N6 amino group points towards Thr3.36. In this orientation, the hydrophobic pocket that accommodates the N6 hydrophobic substitution (N6-cyclopentyl in the cases of CPA and CCPA) is composed of Leu3.33, Thr3.36, Leu6.51, and Thr7.42 (involving the methyl groups of the side chains in the case of threonines). The CCPA ligand is predicted to dock slightly closer to the extracellular part of the receptor, thus missing interactions with helices 1 and 2. The substituent at position 2 (chlorine in CCPA) points towards the extracellular part of the receptor in the region between helices 1 and 7. According to the AUTODOCK scoring function, this docking solution is isoenergetic for the three ribose-unmodified ligands, thus not explaining the experimental increase in affinity found when 2 and N6 substituents are present. Furthermore, this binding mode was not found for NECA, probably due to steric hindrance between the 5' substituent and Asp2.50. With the other orientation (position B), AUTODOCK provided this solution as the best ranked binding position for all the ligands except for ADO (for which the scoring energies for positions A and B are very similar, though). In this case, Asp2.50 interacts with the ribose group through the 2' and 3' hydroxyl groups, also surrounded by Ser1.46 and Ser7.46, and the 5' hydroxyl points towards the backbone oxygen of His7.43. The amino group is close to residue Thr7.42, thus pointing upwards in this orientation to the extracellular side of the receptor. The hydrophobic pocket needed for the N6 substitution is made up of Leu6.51, Ile7.39, Thr7.42, and to a lesser extent by Leu3.33 and Thr7.35. These hydrophobic interactions, related to the structural difference between CPA and ADO, imply an increase in the free energy of binding, both found experimentally and estimated by the

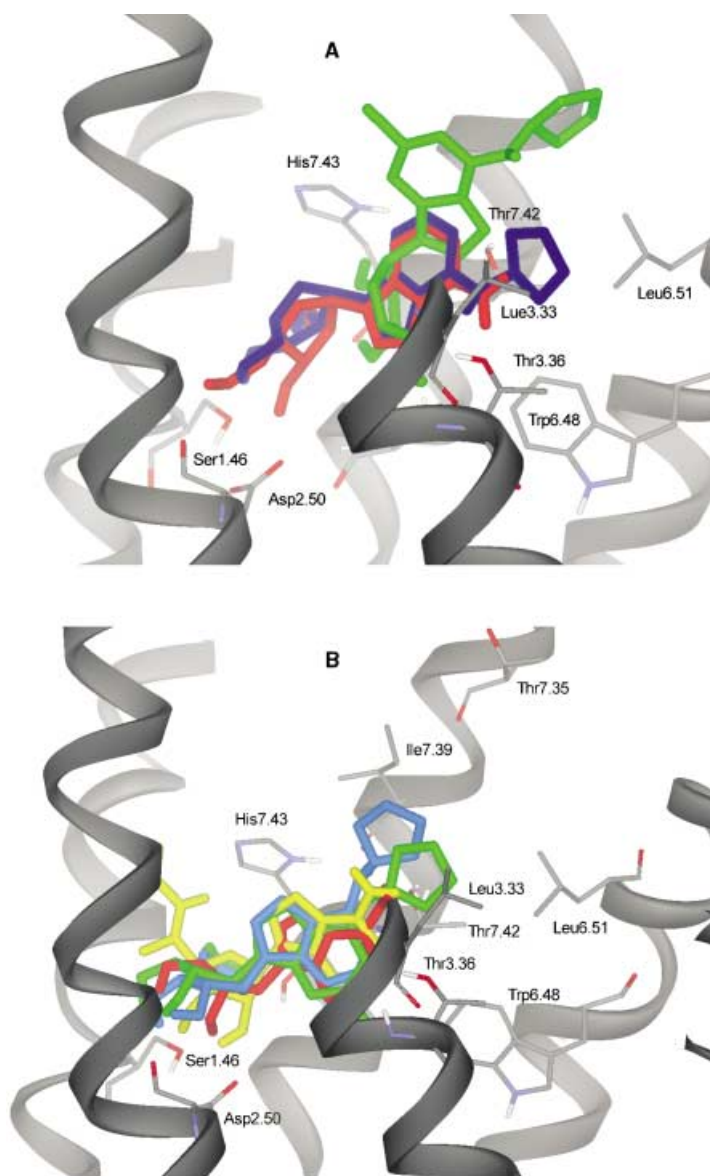


Figure 1. Superposition of agonists considered in the docking positions A and B obtained by AUTODOCK. Ligands: red (ADO), blue (CPA), green (CCPA), and yellow (NECA). Note that this last ligand was not found in position A (see text for details).

AUTODOCK scoring function (Table 2). The 2-Cl substituent in CCPA is accommodated between residues Thr3.36 and Trp6.48, facing in this case towards the cytosolic side of the receptor. However, this chemical modification is not translated into an increase in the estimated receptor affinity with respect to CPA as expected from experimental data; on the contrary, a small decrease is predicted. With regard to the O5' position of adenines, when a carboxamide group is present (5'-N-ethylcarboxamide in NECA), the NH points towards His7.43, whilst the ethyl chain is located between helices 1 and 2. The estimated free energy of binding of NECA lies between the values calculated for ADO and CPA, in accordance with experimental data.

These docking results point to position B as the one that best explains the experimental results, for several reasons.

Firstly, the ranking within the series obtained in this position with the scoring function estimation of the free energy of binding is fairly close to the experimentally ascertained one. Additionally, the estimated affinity is greater in position B than in position A in all cases (with the exception of equal energetic values for ADO). Finally, position A was not automatically found for NECA and the ligand CCPA did not fit well with respect to both ADO and CPA. The correct ranking of compounds in the docking position B is disrupted by the CPA–CCPA pair, the ranking of which is inverted. These two ligands present the greatest similarity in the data set, both in structure and in experimentally determined ΔG ($\Delta\Delta G=0.6$ kcal mol⁻¹), and the prediction of such a small difference may be beyond the scope of the AUTODOCK scoring function, especially when the only structural difference lies in a chlorine atom, which could be deficiently parametrized.

Molecular dynamics and binding free energy predictions

In attempted prediction of binding positions for a certain ligand on a protein, an automatic docking exploration provides good starting points for further detailed analyses of the stability of ligand–protein complexes and the interactions involved in binding. Such analyses are preferably done by molecular dynamics (MD) simulations, thus allowing the effects of the receptor flexibility and the explicit water solvation on the complex of interest to be taken into account. With the LIE approach, it is possible to estimate ligand-binding free energies from MD simulation averages of the non-bonded intermolecular ligand potential energies. In our case, these estimated binding free energies might not be expected to reproduce the absolute values for each complex quantitatively for several reasons. Firstly, it has been shown earlier that hydrophobic binding sites require a negative value of the constant γ in Equation (1), which in the case of a very hydrophobic site such as that of retinol binding protein can reach a value of about -7 kcal mol⁻¹.^[33] Secondly, the fact that we are working here with a homology model that does not incorporate loop regions and the membrane environment of the hA₁AR structure may have some effect on the absolute ΔG scale of the calculated energies. Nevertheless, the influence of these features is expected to be the same through the whole data set, since the ligands considered are full agonists and should consequently share a common binding site. Thus, the differences in binding free energy and the ranking predicted for the series of agonists should in our opinion be reliable, as has been found in previous studies,^[33,45] and therefore be informative about the relevance of each binding mode proposed. In this context, it should be noted that the LIE method without reparameterization has been successfully used to predict relative binding

free energies within the 1 kcal mol⁻¹ range.^[33,45] However, methods using conformational sampling are always associated with convergence errors and it is essential to provide estimates of these in order to judge whether predicted binding energy differences are statistically significant.

The free energy of binding estimation obtained with the AUTODOCK scoring function for the first docking position studied (position A) was not very encouraging (see Table 2). In fact, the influence of the different chemical substitutions in the adenosine scaffold is not reflected in any improvement in the affinity of the compounds. Because of the lack of good superposition of the four agonists after AUTODOCK computations (which locate CCPA slightly closer to the extracellular side than the other ligands and do not predict binding for NECA) the MD simulations of CCPA and NECA started from the geometry found in the CPA and ADO simulations, after mutation of the corresponding ligand structures. The LIE predictions of ligand affinities do not predict the ranking of binding free energies better than AUTODOCK in this binding position A (see Table 3).

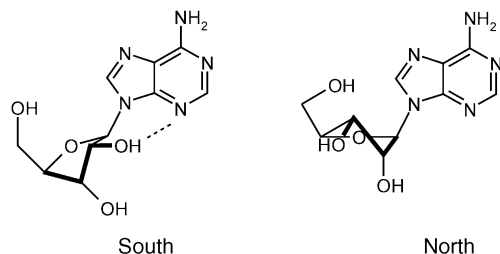
Table 3. Average ligand–surroundings interaction energies [kcal mol⁻¹] for each compound in water and in the two binding positions together with the calculated relative binding free energies.

Ligand	Water		Position A			Position B		$\Delta\Delta G_{bind}^{calcd[a]}$
	$\langle V_{l-s}^{el} \rangle_w$	$\langle V_{l-s}^{vdw} \rangle_w$	$\langle V_{l-s}^{el} \rangle_p$	$\langle V_{l-s}^{vdw} \rangle_p$	$\Delta\Delta G_{bind}^{calcd[a]}$	$\langle V_{l-s}^{el} \rangle_p$	$\langle V_{l-s}^{vdw} \rangle_p$	
ADO	-69.4	-16.5	-61.7	-33.0	-0.5 ± 0.4	-63.3	-34.8	-1.3 ± 0.6
NECA	-76.1	-20.2	-49.6	-41.9	4.8 ± 2.1	-77.1	-42.7	-4.4 ± 1.4
CPA	-65.0	-25.7	-46.1	-50.7	2.1 ± 0.5	-65.1	-48.4	-4.2 ± 0.4
CCPA	-58.3	-26.5	-48.8	-50.7	-1.3 ± 0.9	-65.1	-47.5	-6.3 ± 0.5

[a] Binding free energies calculated from Equation (1) without any optimization of the constant γ .

In fact, only the LIE estimation of CCPA binding free energy showed this compound with a slight improvement with respect to the natural agonist (ADO), while both CPA and NECA were predicted to bind more poorly than ADO (positive values of $\Delta\Delta G$, in disagreement with experimentally ascertained data). A careful analysis of the protein–ligand interactions reveals a common pattern of interactions between the ribose and the protein. The hydrogen bond requirements of the ribose moiety in the protein binding site are not well satisfied, and only a single strong interaction was achieved, between OH5' and Asp2.50. The other two hydroxyl groups (at 2' and 3' positions) display H-bonds only with low occurrence with surrounding side chains (His7.43 and Ser1.40 and Asp2.50 respectively), the first hydroxyl being most involved in intramolecular interactions. This binding profile is different in the ribose-modified ligand NECA. The MD simulation of this last ligand started from the corresponding mutation of adenosine to NECA, in such a way that the 5'-N-ethylcarboxamido substitution was located between helices 2 (Asp2.50), 3 (Leu3.35), and 6 (Phe6.44). The 2' hydroxyl group of NECA shows a stable H-bond with His7.43, and the adenine core overlaps well with that of the ribose-unmodified ligands. However, these new contacts are not able to explain the experimental increase in affinity due to the 5'-N-alkylcarboxamido modification of aden-

osine. Taking the high polarity of the ribose into account, it would be expected that the ligand–receptor complex should be able to generate enough H-bond interactions to balance the ribose desolvation energy. In contrast, the electrostatic energy term is more favorable (more negative values for $\langle V_{1-5}^{\text{el}} \rangle$) in water than in the binding position A for the four agonists studied (Table 3). Additionally, all the ligands adopt a south conformation (Scheme 1) in this binding position, with an in-



Scheme 1. South and north conformations of adenosine. The dashed line indicates the intramolecular hydrogen bond between OH2' and N3 present in the south conformation.

tramolecular hydrogen bond between OH2'...N3. This feature is in disagreement with the results of conformational analysis of ribose-modified adenosine derivatives,^[46] which show a preference for the north-*anti* conformation. The adenine core makes favorable hydrophobic interactions with Val3.32 in all four cases, and an H-bond between N3 and His7.43 is achieved intermittently. The behavior of the N6 position is stable along the MD simulation, with a single H-bond with Thr3.36 and, in the case of CPA and CCPA, favorable hydrophobic interactions with the previously described hydrophobic pocket are attained (residues Thr7.42, Leu6.51, Leu3.33, and Thr3.36). Nevertheless, these additional hydrophobic interactions do not contribute to any enhancement of the estimation of the free energy of binding of CPA, while in the case of CCPA the combination of the N6 and 2-Cl substitutions leads to a slight increase in predicted affinity with respect to the natural agonist ($\Delta\Delta G = 0.8 \text{ kcal mol}^{-1}$).

As was also the case with the AUTODOCK scoring function, the LIE predictions of the free energy of binding are much better for the other binding position considered (position B). The potency profile of the four ligands studied is now close to the observed order of potency (see Table 3), although the low value predicted for the binding of CPA results in an inverted order of the NECA–CPA pair. Analysis of the MD simulations in more detail suggests an explanation for these results and allows an interpretation of how the ranking of compounds is related to the chemical modifications on the adenosine scaffold. First of all, there is a better arrangement of the ribose in its polar binding site than in the binding position A, it adopting the north-*anti* conformation in the three ribose-unmodified ligands in this case (Figure 2A, C, and D). Stronger electrostatic interactions are achieved: both the hydroxyl groups at the 2' and 3' ribose positions show stable H-bonds with Asp2.50,

whereas the 3' one also interacts with Ser7.46, whilst OH5' displays an alternating H-bond with Ser1.40 and His7.43. In the case of NECA (Figure 2B), the chemical modification on 5' induces changes in this interaction profile: the 5'-amide makes additional H-bonds with the side chains of Asp2.50 and Ser1.46, while the side chain of His7.43 changes its conformation slightly due to a stable H-bond with O4'. The ethyl chain of the *N*-ethylcarboxamido group is located in the groove between helices 1 and 2, surrounded by hydrophobic residues (Ala1.43, Val1.47, Gly2.54, and Ile2.58). This pocket is well conserved between A₁ and A_{2A} receptors (the only change is Val1.47Ile in A_{2A}), which is in agreement with the non-selective profile of this ligand. Additionally, this pocket shows steric hindrance for large substituents, a fact that could explain the low tolerance for such modifications in the 5'-carboxamido derivatives.^[47] The different arrangement of this ligand in "position B" results in stronger binding for NECA than for the related 5' unsubstituted agonist, the natural ligand adenosine, although in this case the ligand NECA adopts the south conformation.

In all cases the adenine core is stabilized by hydrophobic interactions with the hydrophobic side chains of Val3.32 and Trp6.48. Available mutagenesis data for A₁ and A_{2A} ARs reveal that a hydrophobic amino acid is necessary at position 3.32.^[18,48] With regard to Trp6.48, which is highly conserved in GPCRs, experiments on A₃AR have shown its critical role in receptor activation,^[49] which is consistent with results obtained on other GPCRs.^[50,51] Additional interactions of the adenine core involve one H-bond between the nitrogen in position 1 of the adenine ring and the hydroxyl group of Thr3.36, bridged through a water molecule, whereas the amino group in N6 is located close to Thr7.42. These interactions are maintained in the four complexes during the MD simulations, although slight differences in the strength of a given contact are observed among ligands. With regard to the N6 substitution on the adenosine scaffold, the cyclopentyl group of CPA and CCPA appears to be surrounded by a hydrophobic pocket formed by Leu6.51, Ile7.39, Thr7.42, and in a minor way Leu3.33 and Thr7.35 (Figure 2C and D), a pocket that contributes to an increase in the estimated affinity of both ligands with respect to adenosine, in agreement with experimental data (see Table 3). Several pieces of experimental data point towards these residues in the ligand-binding process. Mutation of Leu3.33 to alanine results in a substantial reduction in binding not only of the agonist, but also of N-O840, an adenine-like antagonist bearing a cyclopentyl group in the same position as CPA.^[18] It is also worth noting that CPA exhibits a more than 10 000-fold greater affinity for hA₁ than for hA_{2B},^[12,52] a related receptor in which Thr7.42 is replaced by a serine and Leu3.33 is replaced by a valine. We presume that this affinity difference could be related to an optimal complementarity, and consequently improved hydrophobic interactions, between the N6-substituent and the corresponding receptor pocket of the A₁ receptor. With regard to the nature of the residue 7.35, it has also been implicated in species differences in the binding of N6-substituted adenine analogues.^[53]

The calculated LIE binding free energies predict the CPA complex in binding position B to be more than 2 kcal mol⁻¹

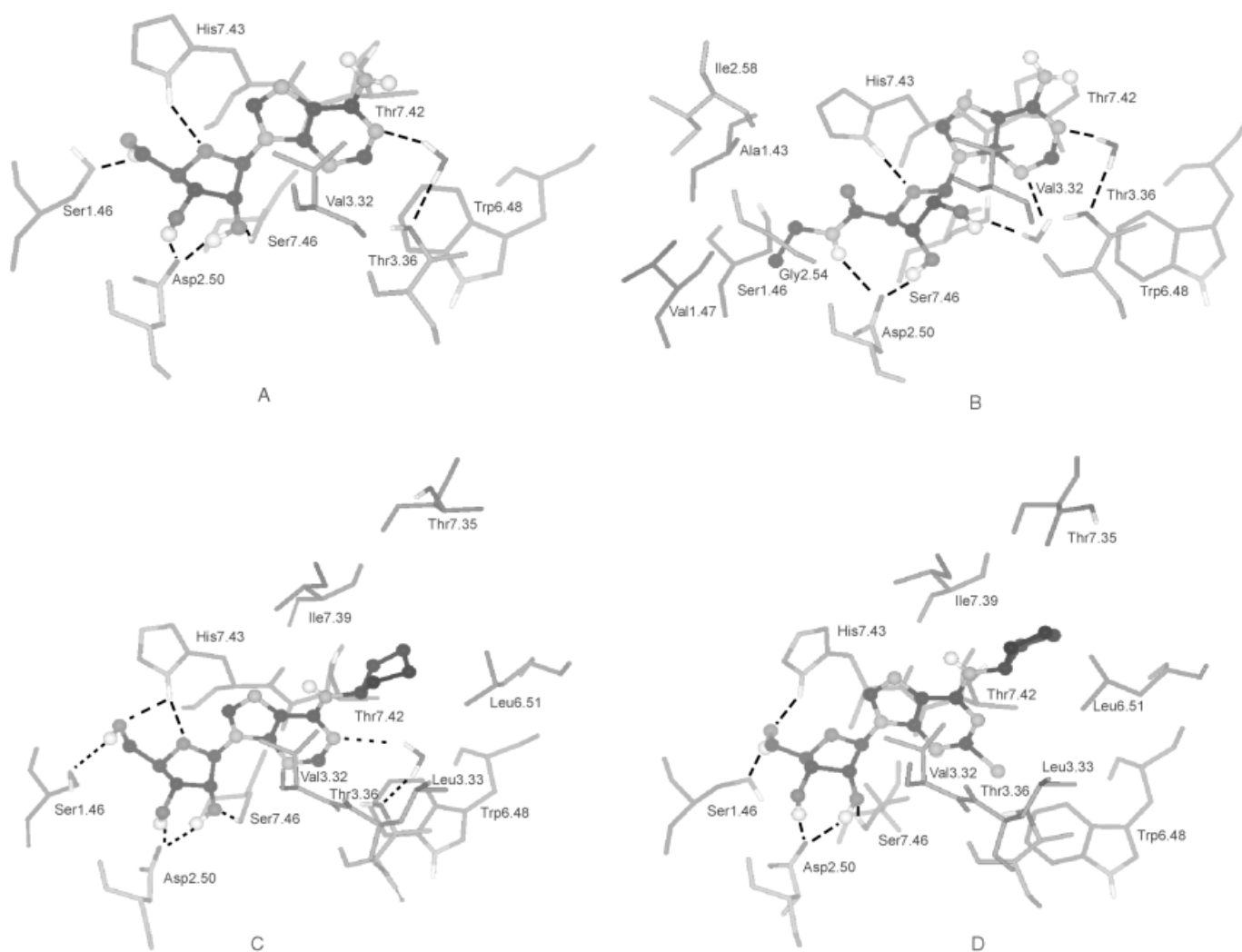


Figure 2. Snapshots from MD simulations of ADO (A), NECA (B), CPA (C), and CCPA (D) in binding position B. Dashed lines indicate hydrogen bonds frequently detected in the MD trajectory.

more stable than the corresponding one with ADO (see Table 2). The interactions between the shared adenosine scaffold (ribose and adenine) and the receptor are the same in both molecules and consequently the *N*6-cyclopentyl group appears to be responsible for the enhancement of binding affinity for CPA in relation to ADO.

If we look at the influence of the 2-chloro substitution in binding position B, the ligand bearing such a chemical modification (CCPA) is predicted to bind more strongly than the corresponding ligand without the chlorine, CPA ($\Delta\Delta G = 2.1 \text{ kcal mol}^{-1}$). This difference is in qualitative agreement with the experimentally found difference of $0.6 \text{ kcal mol}^{-1}$ reported by Klotz et al.,^[12] but clearly overestimated here (chlorine charges resulting from the RESP procedure may, however, apparently be more uncertain than those for other atoms^[54]). The chlorine atom of CCPA is favoring the anchoring of the ligand into the binding site, making weakly favorable electrostatic interactions with Thr3.36, while the interactions of the common scaffold of both ligands are conserved. On the other hand, the

main difference between these two ligands comes from the behavior in water of CCPA, which alternates frequently between the south conformation and the north conformation, while CPA does not explore the south conformation in water. This conformational equilibrium results in a ligand electrostatic interaction energy ($\langle V_{l-s}^{\text{el}} \rangle_w$) that is less negative for CCPA, with a concomitant increase in its predicted affinity. NECA would bind more strongly to the hA_1AR than to ADO ($\Delta\Delta G_{\text{bind}} = -3.1 \text{ kcal mol}^{-1}$) and more weakly than CCPA ($1.9 \text{ kcal mol}^{-1}$), again in qualitative agreement with the experimentally determined differences. On the other hand, our results predict essentially the same binding energy for NECA as for CPA. Thus, only for this pair of ligands is the experimental ranking not reproduced, since there is an experimentally measured difference of $1.1 \text{ kcal mol}^{-1}$ in favor of CPA. It should be noted here that the predicted binding energy of NECA is associated with larger convergence errors than the other compounds, which makes the comparison for this particular pair of molecules more uncertain.

Conclusion

The previously reported hA₁AR putative agonist binding site has been investigated by docking and binding free energy predictions of representative agonists, in an attempt to reproduce the known structure–activity relationships of A₁AR agonists^[6] in our molecular model. In the first part of this study, an exhaustive docking exploration of the receptor model with the four agonists considered was performed, with the aid of AUTODOCK. This was aimed at covering all available possibilities for agonist–receptor docking and yielded the result that no other solution apart from those denoted “A” and “B” should be considered for further investigation. With regard to these two docking positions, the empirical scoring function implemented in the docking software pointed to position B as the more predictive of experimentally measured affinity data. In addition, the superposition of the four ligands on the predicted position B was much better than on the alternative position A, which was only found for three of them.

Analysis of both docking alternatives by the MD/LIE approach had a double objective, the first being to provide a detailed description of the ligand–receptor interactions in a more realistic, dynamic environment, with the aim of better identification of the residues involved in each considered binding mode and the effects of the considered chemical modifications of adenosine. MD simulations allow relaxation of the receptor–ligand complexes, and therefore the resulting pattern of specific interactions may be different from the initial static picture of a docked complex. In fact, some complexes needed an equilibration period after which the average energies were stabilized, reflecting that small conformational changes can lead to increased ligand–receptor affinities. The second objective was to assess the docking results through estimation of the ligand-binding free energy obtained by the LIE method. This kind of assessment of docking results by the LIE approach was found useful in the investigation of the binding of blockers to K⁺ channels by Luzhkov and Åqvist.^[55] In our case, the LIE method predicts more stable complexes for the binding position B in all cases, with binding energy differences between position A and B that stretch from -0.8 to -9.2 kcal mol⁻¹ (extreme values for ADO and NECA, respectively). The other issue that arises is whether the affinity variations between compounds are correctly predicted in position B and the structural reasons for these differences. A comparison of the estimated (both by the AUTODOCK scoring function and by LIE approaches) and experimentally ascertained free energies of binding for position B is depicted in Figure 3. Here, the absolute values derived from the LIE approach have been adjusted to experimental values by use of a least-squares-derived constant parameter $\gamma = -7.3$ kcal mol⁻¹, that is found to be identical to that obtained earlier for retinol-binding protein.^[33] Although the goal of this work is not the prediction of absolute binding free energies, comparison of the differential values obtained is facilitated by the inclusion of this parameter in the LIE equation. As discussed above, the interpretation of the γ parameter is related to binding site hydrophobicity and possibly to neglected features in our homology-derived model: that is, absence of

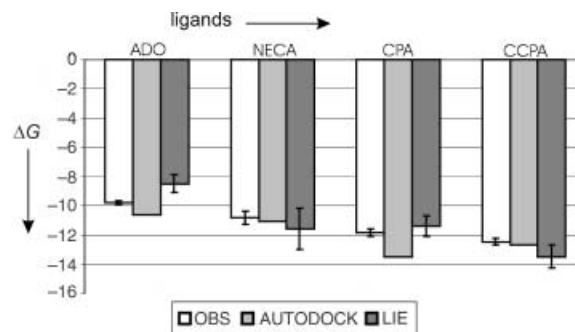


Figure 3. Comparison of experimentally measured (OBS) and estimated (AUTODOCK and LIE) free energies of binding for binding position B. Standard errors of experimental energies are derived from values in the original references.

loop regions and membrane environment, that should affect all agonists in the same way. The origin of a negative γ for hydrophobic binding sites has been addressed elsewhere^[45] and can be understood from extrapolation of hydrophobic solvation energies to zero solute size in polar and nonpolar environments. For a discussion about the relevant thermodynamic cycles describing the microscopic binding process, see also ref. [56]. Figure 3 shows that the ranking within the series is quite reasonable by both AUTODOCK and LIE approaches, even if the ligand CPA is overestimated by the scoring function and underestimated by the LIE approach. Moreover, the role of the chemical substitutions that enhance the affinity with respect to the unmodified adenosine can be explained in terms of ligand–receptor interactions. A hydrophobic pocket is identified between helices 3, 6, and 7 for bulky N6 substitutions characteristic of A₁-selective agonists; the 2-chloro substitution, present in the most potent agonist of the series, is well accommodated between helices 3 and 6; finally, 5'-substituents on 5'-N-alkylcarboxamidoadenosines are located in a binding cavity with limited available space between helices 1, 2, and 7, thus not suitable for large N-alkyl modifications. Additionally, the proposed model for agonists binding to the hA₁AR is explanatory for some mutagenesis data that affect all the compounds in the series, such as those related to Val3.32,^[48] Thr3.36,^[18] and His7.43.^[19,20] Our binding mode is also consistent with the possible implication of Trp6.48 in receptor activation,^[49] although this biological process can not be simulated with the current methodology. Recently, a model of agonist–receptor interactions for the A_{2A} subtype has been published.^[57] In that work, the proposed binding site partially overlaps that postulated in this study, although differences are also observed. In addition to the fact that the two studies have been carried out in related but different receptors (hA_{2A}AR vs. hA₁AR), it should also be pointed out that different strategies underlie the two studies. The main aim of the study of Kim and co-workers^[57] was to generate a model consistent with the experimental evidence. With this purpose the authors introduced manual interventions into the modeling process. Conversely, the present study uses automatic protocols free of human intervention to obtain ligand–protein models. Experimentally derived evidence is then used a posteriori to choose among different solutions and to assess the quality of the

models. If account is taken of the complexity of the interpretation of mutagenesis data^[58] the two modeling approaches are, in our opinion, complementary and useful for dealing with prediction of ligand-binding positions in these systems. Besides the relevance of these modeling results for guiding the design of more potent agonists of the A₁AR, it should be pointed out that the strategy described, consisting of obtaining docking positions automatically and assessing them by checking their consistency for a series of ligands, could be useful in other similar challenging situations.

Acknowledgements

The authors thank Almirall Prodesfarma S.A. and Dr. Mabel Loza for their information and support. They are also grateful to Dr. Jordi Villà-Freixa for helpful discussions and careful reading of the manuscript. H.G.T. acknowledges CIRIT of the Generalitat de Catalunya for his Ph.D. grant.

Keywords: adenosine receptors · binding free energy · density functional calculations · docking exploration · homology modeling · molecular dynamics

- [1] C. E. Müller, *Farmaco* **2001**, *56*, 77–80.
- [2] C. Müller, B. Stein, *Curr. Pharm. Design.* **1996**, *2*, 501–530.
- [3] S. A. Poulsen, R. J. Quinn, *Bioorg. Med. Chem.* **1998**, *6*, 619–641.
- [4] H. L. Haas, O. Selbach, *Naunyn Schmiedeberg's Arch. Pharmacol.* **2000**, *362*, 375–381.
- [5] J. M. Downey, G. S. Liu, J. D. Thornton, *Cardiovasc. Res.* **1993**, *27*, 3–8.
- [6] C. E. Müller, *Curr. Med. Chem.* **2000**, *7*, 1269–1288.
- [7] S. Vittori, A. Lorenzen, C. Stannek, S. Costanzi, R. Volpini, I. J. AP, J. K. Kunzel, G. Cristalli, *J. Med. Chem.* **2000**, *43*, 250–260.
- [8] E. M. van der Wenden, J. K. von Frijtag Drabbe Kunzel, R. A. A. Mathot, M. Danhof, A. P. IJzerman, W. Soudijn, *J. Med. Chem.* **1995**, *38*, 4000–4006.
- [9] M. J. Lohse, K. N. Klotz, E. Diekmann, K. Friedrich, U. Schwabe, *Eur. J. Pharmacol.* **1988**, *156*, 157–160.
- [10] J. W. Daly, W. L. Padgett, *Biochem. Pharmacol.* **1992**, *43*, 1089–1093.
- [11] M. J. Lohse, K. N. Klotz, U. Schwabe, G. Cristalli, S. Vittori, M. Grifantini, *Naunyn Schmiedeberg's Arch. Pharmacol.* **1988**, *337*, 687–689.
- [12] K. N. Klotz, J. Hessling, J. Hegler, C. Owman, B. Kull, B. B. Fredholm, M. J. Lohse, *Naunyn Schmiedeberg's Arch. Pharmacol.* **1998**, *357*, 1–9.
- [13] K. N. Klotz, *Naunyn Schmiedeberg's Arch. Pharmacol.* **2000**, *362*, 382–391.
- [14] H. Gutiérrez-de-Terán, N. B. Centeno, M. Pastor, F. Sanz, *Proteins*, **2004**, *54*, 705–715.
- [15] K. Palczewski, T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. Le Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto, M. Miyano, *Science* **2000**, *289*, 739–745.
- [16] P. J. Goodford, *J. Med. Chem.* **1985**, *28*, 849–857.
- [17] K. Kristiansen, W. K. Kroeze, D. L. Willins, E. I. Gelber, J. E. Savage, R. A. Glennon, B. L. Roth, *J. Pharmacol. Exp. Ther.* **2000**, *293*, 735–746.
- [18] S. A. Rivkees, H. Barbhuiya, A. P. IJzerman, *J. Biol. Chem.* **1999**, *274*, 3617–3621.
- [19] K. Klotz, M. Lohse, U. Schwabe, *J. Biol. Chem.* **1988**, *263*, 17522–17526.
- [20] M. E. Olah, H. Ren, J. Ostrowski, K. A. Jacobson, G. L. Stiles, *J. Biol. Chem.* **1992**, *267*, 10764–10770.
- [21] J. A. Ballesteros, H. Weinstein in *Methods in Neurosciences* (Eds.: P. M. Conn, S. C. Sealfon), Academic Press, San Diego, **1994**, pp. 366–428.
- [22] F. H. Allen, O. Kennard, *Chem. Des. Autom. News.* **1993**, *8*, 31–37.
- [23] Insight II, Molecular simulations Inc., San Diego, **2000**.
- [24] J. J. Stewart, *J. Comput. Aided. Mol. Des.* **1990**, *4*, 1–105.
- [25] Gaussian 98 (Revision A.7), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, J. A. Pople, Gaussian, Inc., Pittsburgh, PA, **1998**.
- [26] C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, *97*, 10269–10280.
- [27] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
- [28] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J. Comp. Chem.* **1998**, *19*, 1639–1662.
- [29] J. Åqvist, C. Medina, J. E. Samuelsson, *Protein. Eng.* **1994**, *7*, 385–391.
- [30] T. Hansson, J. Marelus, J. Åqvist, *J. Comput. Aided. Mol. Des.* **1998**, *12*, 27–35.
- [31] F. S. Lee, Z. T. Chu, M. B. Bolger, A. Warshel, *Prot. Eng.* **1992**, *5*, 215–228.
- [32] J. Åqvist, T. Hansson, *J. Phys. Chem.* **1996**, *100*, 9512–9521.
- [33] J. Åqvist, J. Marelus, *Comb. Chem. High Throughput. Screen.* **2001**, *4*, 613–626.
- [34] J. Marelus, K. Kolmodin, I. Feierberg, J. Åqvist, *J. Mol. Graph. Modelling* **1999**, *16*, 213–225.
- [35] W. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.
- [36] J. Essex, W. Jorgensen, *J. Comp. Chem.* **1995**, *16*, 951–972.
- [37] G. King, A. Warshel, *J. Chem. Phys.* **1989**, *91*, 3647–3661.
- [38] G. Allende, V. Casado, J. Mallol, R. Franco, C. Lluis, E. I. Canela, *J. Neurochem.* **1993**, *60*, 1525–1533.
- [39] F. S. Lee, A. Warshel, *J. Chem. Phys.* **1992**, *97*, 3100–3107.
- [40] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph. Model.* **1996**, *14*, 33–38.
- [41] E. van der Wenden, S. L. Price, R. P. Apaya, A. P. IJzerman, W. Soudijn, *J. Comp. Aid. Mol. Des.* **1995**, *9*, 44–54.
- [42] M. Rashid, P. Manivet, H. Nishio, J. Pratuangdejkul, M. Rajab, M. Ishiguro, J. M. Launay, T. Nagatomo, *Life Sci.* **2003**, *73*, 193–207.
- [43] M. Mahmoudian, *J. Mol. Graph. Model.* **1997**, *15*, 149–153.
- [44] A. Gieldon, R. Kazmierkiewicz, R. Slusarz, J. Ciarkowski, *J. Comput. Aided. Mol. Des.* **2001**, *15*, 1085–1104.
- [45] M. Almlöf, B. O. Brandsdal, J. Åqvist, *J. Comput. Chem.* **2004**, *25*, in press..
- [46] P. Franchetti, L. Cappellacci, S. Marchetti, L. Trincavelli, C. Martini, M. R. Mazzoni, A. Lucacchini, M. Grifantini, *J. Med. Chem.* **1998**, *41*, 1708–1715.
- [47] M. de Zwart, A. Kourounakis, H. Kooijman, A. L. Spek, R. Link, J. K. von Frijtag Drabbe Kunzel, A. P. IJzerman, *J. Med. Chem.* **1999**, *42*, 1384–1392.
- [48] Q. Jiang, B. X. Lee, M. Glashofer, A. M. van Rhee, K. A. Jacobson, *J. Med. Chem.* **1997**, *40*, 2588–2595.
- [49] Z.-G. Gao, A. Chen, D. Barak, S.-K. Kim, C. E. Muller, K. A. Jacobson, *J. Biol. Chem.* **2002**, *277*, 19056–19063.
- [50] J. A. Javitch, J. A. Ballesteros, H. Weinstein, J. Chen, *Biochemistry* **1998**, *37*, 998–1006.
- [51] J. Marie, E. Richard, D. Pruneau, J. L. Paquet, C. Siatka, R. Languier, C. Ponce, P. Vassault, T. Groblewski, B. Maigret, J. C. Bonnafous, *J. Biol. Chem.* **2001**, *276*, 41100–41111.
- [52] J. Linden, T. Thai, H. Figler, X. Jin, A. S. Robeva, *Mol. Pharmacol.* **1999**, *56*, 705–713.
- [53] A. L. Tucker, A. S. Robeva, H. E. Taylor, D. Holeton, M. Bockner, K. R. Lynch, J. Linden, *J. Biol. Chem.* **1994**, *269*, 27900–27906.
- [54] M. Almlöf, July 2003, personal communication.
- [55] V. B. Luzhkov, J. Åqvist, *FEBS Lett.* **2001**, *495*, 191–196.
- [56] Y. Sham, Z. T. Chu, H. Tao, A. Warshel, *Proteins* **2000**, *39*, 393–407.
- [57] S.-K. Kim, Z.-G. Gao, P. V. Rompaey, A. S. Gross, A. Chen, S. V. Calenbergh, K. A. Jacobson, *J. Med. Chem.* **2003**, *46*, 4847–4859.
- [58] D. Colquhoun, *Br. J. Pharmacol.* **1998**, *125*, 924–947.

Received: November 11, 2003