Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling 26 (2008) 874–883
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Prediction of the binding mode of biarylpropylsulfonamide allosteric AMPA receptor modulators based on docking, GRID molecular interaction fields and 3D-QSAR analysis

Kasper Harpsøe, Tommy Liljefors, Thomas Balle *

Department of Medicinal Chemistry, Faculty of Pharmaceutical Sciences, University of Copenhagen, Universitetsparken 2, DK-2100 Copenhagen, Denmark

Received 21 March 2007; received in revised form 7 June 2007; accepted 8 June 2007
Available online 14 June 2007

Abstract

A novel approach of combining flexible molecular docking, GRID molecular interaction fields, analysis of ligand–protein hydrogen bond interactions, conformational energy penalties and 3D-QSAR analysis was used to propose a binding mode in the dimer interface of the iGluR2 receptor for the biarylpropylsulfonamide class of positive allosteric AMPA modulators. Possible binding poses were generated by flexible molecular docking. GRID molecular interaction fields of the binding site, ligand–protein hydrogen bonding interactions and conformational energy penalties were used to select the most likely binding mode. The selected binding poses were subjected to a 3D-QSAR analysis using previously published activity data. The resulting model (2 LVs, $R^2 = 0.89$, $q^2 = 0.61$) predicted the activities of the compounds in the test set with a standard deviation on error of prediction of 0.17. The proposed binding mode was validated by interpretation of the PLS-coefficient regions from the 3D-QSAR analysis in terms of interactions between the receptor and the modulators.

Keywords: Biarylpropylsulfonamides; AMPA receptor; Allosteric modulators; Binding mode; Docking; Glide; GRID; 3D-QSAR; GOLPE

1. Introduction

The α-amino-3-hydroxy-5-methyl-4-isoxazolepropanoic acid (AMPA) receptors belong to the class of ionotropic glutamate receptors (iGluRs). The AMPA receptors are widespread in the mammalian nervous system and play a central role in, e.g. learning and memory. Enhancement of these vital functions by positive allosteric modulators potentiating the agonist response has been suggested to be beneficial for treatment of cognitive disorders and several compounds have entered preclinical and clinical trials [1]. The AMPA receptors are comprised of iGluR1-4 subunits and are known to be tetrameric complexes [2] organized as dimers of dimers [3,4]. Each monomer is comprised of an extracellular portion containing a large N-terminal domain and a ligand-binding domain (LBD) comprised of two sub-domains designated S1 and S2, a transmembrane portion that forms part of the ion channel [5] and a long intracellular C-terminal domain. Binding of an agonist induces a conformational change in the LBD of the dimer. The dimer interface remains intact and the strain due to the conformational change is transferred to the ion channel and causes it to open [6]. In the process called desensitization the dimer interface then rearranges which decouples ligand binding from ion channel opening. This results in ion channel closure while the agonist is still bound [7] in contrast to deactivation where the ion channel closes as a result of unbinding of the agonist [8].

Crystal structures of the LBD of iGluR2 in complex with an agonist and the positive allosteric AMPA modulators **1** (cyclothiazide), **2** (aniracetam) and **3** (CX614) (Chart 1) are currently available [7,9]. Compounds **1**–**3** inhibit desensitization and slow deactivation to a variable degree and thus potentiate the AMPA receptor response [9–11] by binding in the dimer interface [7,9]. Two molecules of **1** bind in two symmetrical binding sites [7] while **2** and **3** bind in the center of the dimer interface overlapping both binding sites for **1** [9] (Fig. 1).

The four AMPA receptor subunits are expressed as two different splice forms called flip and flop, which differ in the S2 sub-domain of the LBD by 9–11 residues [12]. A single residue

* Corresponding author. Tel.: +45 35 30 60 00; fax: +45 35 30 60 40.
E-mail address: tb@farma.ku.dk (T. Balle).

1093-3263/$ – see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2007.06.002

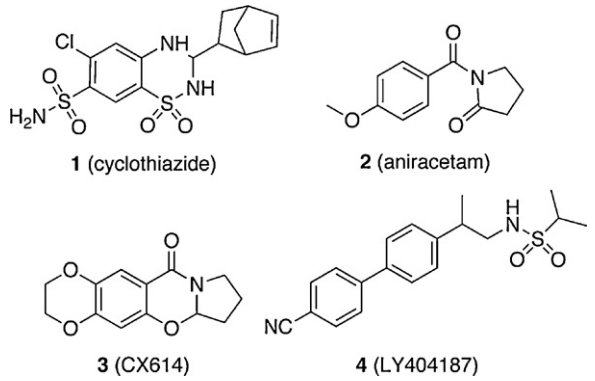


Chart 1. Chemical structure of four positive allosteric AMPA modulators.

(residue 754 in iGluR2) located in the dimer interface, which is a serine in the flip and an asparagine in the flop splice form is responsible for a marked preference of **1** for the flip splice form and for a weak preference of **2** for the flop splice form [10]. Compound **3** also displays a weak preference for the flop splice form [11]. Another positive allosteric AMPA receptor modulator, **4** (LY404187, Chart 1) belongs to a recently reported series of biarylpropylsulfonamides including one of the most potent AMPA receptor potentiators reported to date [13]. Like compound **1**, compound **4** inhibits desensitization of AMPA receptor responses [13] and displays a preference for the flip splice form [14] but the effect on deactivation has so far not been reported. Since the same residue is responsible for the flip/flop preferences of **1**, **2** and **4** [10,14], it is likely that the binding site of **4** overlaps with those of **1** and **2**, which is also suggested by Morrow et al. [15]. However, there are currently no crystal structure of **4** bound to an AMPA receptor and since a molecular superimposition of **4** and any of the compounds **1**–**3** is not obvious, it is of high interest to identify the binding mode of **4** in the AMPA receptor dimer interface. The determination of the binding mode for **4** would enable structure-based ligand design within the series of biarylpropylsulfonamides and aid in the elucidation of the effect on desensitization and deactivation.

In the following we propose a binding mode for **4** and a series of derivatives of **4** [16] (Table 1) in the dimer interface of iGluR2. The proposed binding mode is based on flexible molecular docking using Glide [17]. In a critical assessment of 10 docking programs and 37 scoring functions against eight proteins of seven

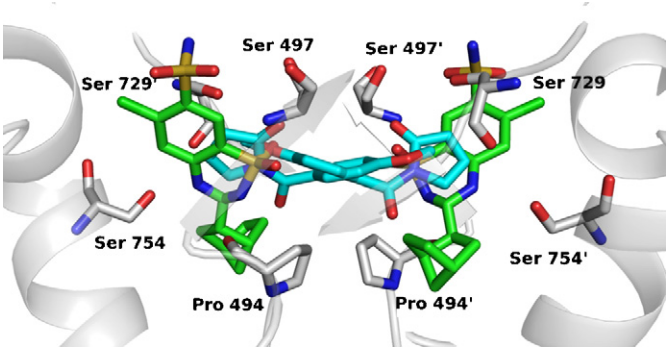


Fig. 1. Binding mode of compounds **1** (green carbons) and **2** (cyan carbons – two overlapping orientations are present in the structure, each with an occupancy of 0.5) in the dimer interface of iGluR2.

protein types, Warren et al. [18] conclude that docking programs are usually able to generate multiple ligand poses that include binding modes similar to the crystallographically observed binding mode. However, scoring functions are less successful at identifying the observed binding pose in the set of docked poses. For this reason, the docked poses were carefully analyzed using molecular interaction fields calculated using GRID [19] and ligand–protein hydrogen bond interactions including the possible influence of strongly bound water molecules as identified by GRID. In addition, conformational energy penalties [20] were taken into account. Finally, the docked poses for the series of biarylpropylsulfonamides were subjected to a 3D-QSAR analysis using GRID/GOLPE [21]. If the binding poses obtained are valid, the calculated positive and negative PLS-coefficient regions should be possible to interpret in terms of interactions in the binding pocket displayed by the docked compounds [22]. This novel approach of combining several methods to analyze docked poses should give a better validated binding mode compared to that obtained by using a general purpose scoring function.

2. Methodology

2.1. Alignments

The sequence of the iGluR2$_{flop}$ LBD from the species *Rattus Norvegicus* was retrieved from the Protein Data Bank [23] using the aniracetam-iGluR2 crystal structure [9]. Since the first two amino acids in the sequence are cloning artifacts, they were deleted before alignment. The incorporated GT-linker was maintained in the sequence to mark its position. The sequence of iGluR4$_{flip}$ (primary accession number P19493) was retrieved from the UniProt Knowledgebase [24]. The full length sequence was aligned to the sequence of the iGluR2$_{flop}$ LBD using T-Coffee, version 2.11 [25] with default settings and the none-matching sequence was discarded (see Supplementary data).

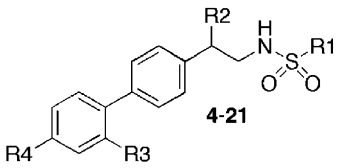
2.2. Molecule preparation

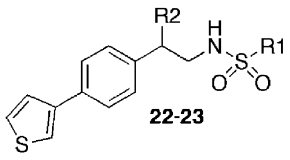
The aniracetam-iGluR2 crystal structure [9] was imported into Maestro [26], Asn 754 of both chains A and C was mutated to a serine, water molecules were deleted, the atom types and bond orders of the agonist were corrected and hydrogen atoms were added. All modulator molecules were built in Maestro and energy minimized using MacroModel [27] with the MMFFs force field and the TNCG minimization method. The calculations were performed without a cut-off distance for non-bonded interactions and an aqueous solvation model.

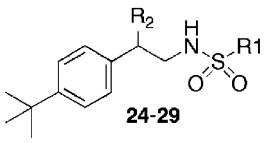
2.3. Characterization of the binding site

The binding cavity was characterized by calculations of molecular interaction fields using the program GRID [28]. Hydrogen atoms were deleted and the atom types of HETATM records in the pdb file were corrected according to the grub.dat file included in GRID. The box dimensions were set to X_{max}/X_{min}, 26/12; Y_{max}/Y_{min}, 85/71; Z_{max}/Z_{min}, 37/17 and grid

Table 1
Structures, pharmacological data and EC_{50} values predicted by 3D-QSAR for the compounds studied







Compounds[a]	R1	R2	R3	R4	EC_{50}[b]	Calculated EC_{50}[c]
4	$CH(CH_3)_2$	CH_3	H	CN	0.29	0.16[d]
5	$CH(CH_3)_2$	CH_3	H	NH_2	0.13	0.18
6	$CH(CH_3)_2$	CH_3	H	CHO	0.25	0.12
7	$CH(CH_3)_2$	CH_3	H	CH_3	0.27	0.17
8	$CH(CH_3)_2$	CH_3	H	H	1.0	1.6
9	$CH(CH_3)_2$	CH_3	H	CO_2H	1.4	NI[e]
10	$CH(CH_3)_2$	CH_3	H	Cl	>3	NI[f]
11	$CH(CH_3)_2$	CH_3	H	CF_3	>3	NI[f]
12	$CH(CH_3)_2$	H	F	H	7.2	5.7[d]
13	CH_3	CH_3	F	H	19.6	15.9[d]
14	CF_3	CH_3	F	H	32.9	20.4
15	CH_2CH_3	CH_3	F	H	5.4	7.5
16	$(CH_2)_2CH_3$	CH_3	F	H	23.8	31.7
17	$CH(CH_3)_2$	CH_3	F	H	4.4	5.3
18	$N(CH_3)_2$	CH_3	F	H	4.0	5.8
19	Phenyl	CH_3	F	H	62.5	51.6
20	Benzyl	CH_3	F	H	>100	NI[e,f]
21	$(CH_2)_3CH_3$	CH_3	F	H	>100	NI[e,f]
22	CH_3	CH_3			4.5	4.1
23	$CH(CH_3)_2$	CH_3			0.66	1.2
24	$CH(CH_3)_2$	H			12.8	5.6
25	$CH(CH_3)_2$	CH_3			1.2	2.8
26	$CH(CH_3)_2$	CH_2CH_3			2.0	3.8
27	$CH(CH_3)_2$	$(CH_2)_2CH_3$			27.9	7.6
28	$CH(CH_3)_2$	Benzyl			>100	NI[e,f]
29	$CH(CH_3)_2$	Phenylethyl			>100	NI[f]

[a] All chiral compounds were tested as racemic mixtures.
[b] EC_{50} values in μM have been published previously by Ornstein et al. [16] and are based on potentiation of glutamate-induced currents on $GluR4_{flip}$ receptors relative to the effect of 100 μM **1**.
[c] EC_{50} values in μM as calculated from the 3D-QSAR model.
[d] External validation set.
[e] NI, not included in the 3D-QSAR as a good docking pose was not obtained.
[f] NI, not included in the 3D-QSAR model as the activity was not accurately determined.

spacing was set to 0.33 Å (NPLA = 3). All HETATM records were included in the calculations (NETA = 100) and the methyl (C_3) and water (OH_2) probes were used. When GRID was utilized to position water molecules, the grid spacing was set to 0.2 (NPLA = 5). The utility programs in GRID, Minim and Filmap, were employed to locate minima using a threshold of −2 kcal/mol and position the water molecules using interpolation. Water molecules not mediating hydrogen bonds between the protein and the modulator were deleted.

2.4. Docking

Dockings and refinements were performed using Glide [29]. Before docking, the position of polar hydrogen atoms in serines, etc., were optimized by running the first step of refinement in the protein preparation utility of Glide. A grid was prepared with the center defined by residues Pro 494 and Ser 497. All three midpoint box diameters were set to 14 Å and the length of ligands were set to 10 Å. The preliminary docking with **5** employed a scaling of 0.7 for all ligand atoms and the maximal output was set to 10 poses per ligand. As the protein structure was obtained with **2** bound and a mutation has been introduced in the binding site the interactions between the modulator and the protein may be non-optimal. In order to obtain more reliable poses for the biarylpropylsulfonamide modulators in following dockings the protein with the selected pose of **5** present in the binding site was refined using all five refinement steps of the Glide protein preparation utility (RMSD = 0.45 Å measured on all heavy atoms – structure is included in Supplementary data). Sequential dockings only differed from the preliminary by scaling the ligand atoms by 0.9 instead of 0.7.

2.5. Conformational analysis

The conformational analysis of **5** was performed in MacroModel with the force field, minimization method and

non-bonded interaction cut-off mentioned above. A Monte Carlo conformational search was performed with an aqueous solvation model, the automatic setup procedure was used to define torsion angles to sample and the energy window was set to 6 kcal/mol (25 kJ/mol). The gas phase energy for the obtained global energy minimum conformation was calculated. The binding conformation obtained from docking was minimized in the gas phase using flat-bottomed Cartesian constraints with a force constant of 120 kcal/mol Å^2 (500 kJ/mol Å^2) allowing atoms to move 0.3 Å. The conformational energy penalty of binding was calculated as the difference between the gas phase energy of the binding conformation and the gas phase energy of the global minimum conformation in water.

2.6. 3D-QSAR GRID calculations

The interaction energies were calculated using GRID with a grid size of (X = 21 Å, Y = 20 Å, Z = 26 Å). Energies were calculated with a grid spacing of 1 Å (NPLA = 1) for the water (OH_2) and methyl (C_3) probes. The dielectric constant for the solvent around the target molecule (DWAT) was set to 4.0 to mimic the interior of the protein. Maximum cut-off energy (EMAX) was set to 5 kcal/mol.

2.7. GOLPE analysis

Sixteen compounds (**5**–**8**, **14**–**19**, **22**–**27**) were included in the training set and 3 compounds (**4**, **12**, **13**) were used for external validation. In the final model, all 19 compounds were included in the training set. The partial least-square (PLS) models were calculated using GOLPE [21,30]. The dependent variables were imported as pEC_{50}-values. GOLPE automatically rejects X-variables having a total sum of square less than 10^{-7}. Further pre-treatment was performed as follows: absolute critical X-values below |0.10| were deleted. Minimum S.D. cut-off was set to 0.10 for all probes. Second and third level X-variables were deleted. The initial PLS-model was build using these data. Groups of variables were generated using smart region definition (SRD) [31] with the default number of seeds. Variables within a distance of 1 Å to the seeds were grouped. Neighboring groups within a distance of 2 Å containing the same information were collapsed. The effect of the individual groups of variables was evaluated using the fractional factorial design (FFD) procedure including 20% dummy variables. A fold-over design was used by reversing the signs of the columns in the original design matrix. Only groups having a positive influence on the model larger than the average dummy variable were included in the final model. Cross-validation was performed using the ''leave five random groups out'' procedure; twenty reduced models were built and evaluated. Each time 20% of the molecules were randomly chosen and left out of the model. The reduced models were then used to predict the affinity of the molecules left out. Finally, the obtained model was used to predict the affinities of an external validation set of three compounds.

3. Results and discussion

3.1. Protein selection

The EC_{50} values listed in Table 1 represents the potentiation of the response mediated by 100 μM L-glutamate in HEK-293 cells expressing the $iGluR4_{flip}$ subunit relative to that of 100 μM **1** [16]. A crystal structure of the LBD of the $iGluR4_{flip}$ subunit is currently unavailable. However, the LBD of $iGluR2_{flop}$, for which several crystal structures are available [6,7,9], is closely related to that of $iGluR4_{flip}$. The sequence identity between the two LBDs is 89%, and the only differing amino acid located in the binding site of **1** is the flip/flop residue, Ser/Asn 754. The high degree of similarity of iGluR2 and iGluR4 in the area corresponding to the binding pocket is also reflected in the similar EC_{50} values reported for **4**; 0.15 and 0.21 μM for $iGluR2_{flip}$ and $iGluR4_{flip}$, respectively [32]. Thus, under the assumption that the binding site of the biarylpropylsulfonamides in Table 1 overlaps with the binding site of **1** as discussed in the introduction, the crystal structures of the LBD of $iGluR2_{flop}$ should be suitable for the investigation of the binding mode of he compounds studied. Compared to other crystal structures of the AMPA receptor LBD, those co-crystallized with **2** and **3** display a different conformation of Ser 497 and 729 in the binding site of the modulators. As the biarylpropylsulfonamides are assumed to bind in the same area [15], the structure of $iGluR2_{flop}$ co-crystallized with **2** was selected for this work and Asn 754 was replaced by a serine to mimic the flip splice form.

3.2. Docking and evaluation of docked poses for compound 5

To aid in the evaluation of docked poses, the protein was characterized by calculations of molecular interaction fields using the methyl (C_3) and water (OH_2) probes in the program GRID. Compound **5** (Table 1), which is the most active of the compounds studied, was selected for the initial identification of the binding mode as it can be expected to have the best fit to the receptor. It contains a chiral center at the carbon carrying the R2 substituent (Table 1) with unreported stereochemistry and the sulfonamide hydrogen and nitrogen lone pair exchange positions due to inversion about the sulfonamide nitrogen. Thus, four possible configurations were built and docked into the receptor structure using Glide. Glide has previously been evaluated on compound **1** resulting in an RMSD of 0.19 Å on all heavy atoms compared to the crystal structure pose [33]. A similar evaluation was performed on compounds **2** and **3** resulting in RMDS values of 0.61 and 0.44 Å, respectively (data not shown).

The fit of the docked poses to the molecular interaction fields of the receptor, the hydrogen bonding pattern and the conformational energy of the poses were considered in the selection of the most likely binding mode. The selected output poses of **5** were ranked as number 1 and number 4 for the (*R*)- for the (*S*)-enantiomers, respectively. Both poses display similar binding modes with the biphenyl moiety located near

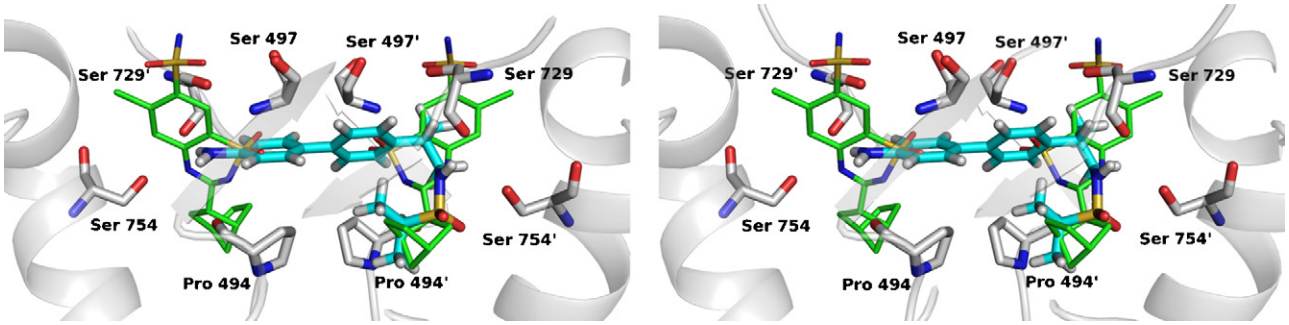


Fig. 2. Predicted binding mode (stereo view) of compound **5** (cyan carbons) and the crystal structure-binding mode of CTZ (green carbons) in the dimer interface of iGluR2.

the center of the dimer interface and the isopropylsulfonamide protruding into the binding cavity of the norbornenyl group of **1**. Fig. 2 shows that the selected binding pose of (*S*)-**5** overlaps with the binding sites of **1**. The isoenergy contour calculated by using the GRID methyl probe at −2 kcal/mol, representing favorable van der Waals interactions with the protein, almost fully encloses the hydrophobic parts of the molecule (Fig. 3). This indicates a good fit to the protein cavity. At the same time, the hydrophobic parts of the molecule does not overlap with favorable sites for water molecules indicating insignificant loss of binding affinity due to desolvation.

The hydrogen bond interactions with the protein of the selected poses of (*S*)- and (*R*)-**5** differ significantly. The sulfonamide nitrogen atom hydrogen bonds to the backbone carbonyl of Ser 729 (Fig. 4a) and Pro 494 (Fig. 4b) for (*S*)- and (*R*)-**5**, respectively. The hydroxy group of Ser 754 is within hydrogen bonding distance of one of the sulfonamide oxygen atoms for both enantiomers (Fig. 4). Assuming a similar binding mode for **4**, this hydrogen bond can explain its reported flip preference [14]. Within 3 Å of the second sulfonamide oxygen of (*S*)-**5**, GRID calculations using the water probe identify the location of a water molecule with an interaction energy to the protein below −12 kcal/mol (Fig. 4a). This strongly bound water molecule, also present in the aniracetam-iGluR2 [9] and other iGluR2 crystal structures [6], may form a hydrogen bond to (*S*)-**5**, whereas the isopropyl group of (*R*)-**5** is located 2 Å from this position indicating a steric clash with the water molecule. Furthermore, the second sulfonamide oxygen of (*R*)-**5** does not lie within hydrogen bonding distance of any donor groups of the protein or calculated positions of tightly bound water molecules (Fig. 4b). This indicates a loss of binding affinity due to uncompensated energy cost of desolvation. These observations qualitatively predict (*S*)-**5** to be the most active enantiomer. This is furthermore supported by comparison of the conformational energy penalties of the bound conformations of the two enantiomers. For (*R*)-**5**, 1.7 kcal/mol is required to adopt the suggested binding mode whereas the corresponding energy for (*S*)-**5** is −1.3 kcal/mol, suggesting an energy gain, when calculated as described by Boström et al. [20] The identified strongly bound water molecules described above was considered important for the binding conformation of the modulators. Thus, a water molecule at the position indicated by GRID was inserted and treated as part of the protein in subsequent dockings.

The (*S*)-enantiomers of compounds **4**–**29** in Table 1 were docked to the protein, refined in the presence of (*S*)-**5** as described in Section 2. The highest-ranking pose of each modulator with a binding mode similar to the one described above for (*S*)-**5** was selected for the following work. With respect to compounds **9**, **20** and **28** no such pose was identified. The reason for this will be discussed below.

3.3. 3D-QSAR analysis

The docked poses for 19 compounds with well-defined activities and reasonable binding poses were used directly as

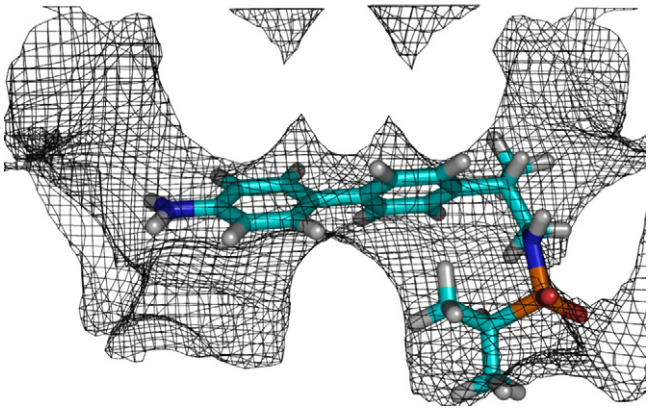


Fig. 3. The docked pose of the (*S*)-enantiomer of **5** in the methyl probe energy contour at −2 kcal/mol.

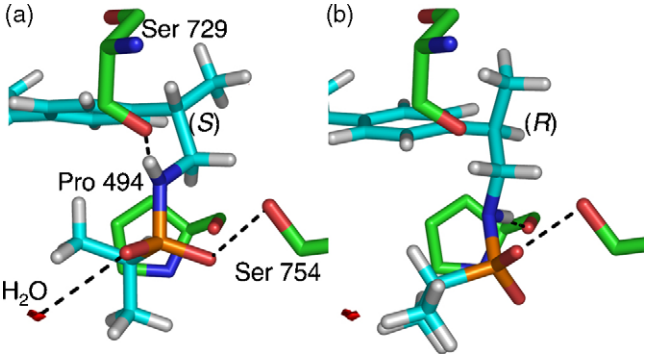


Fig. 4. Hydrogen bonds (dashed lines) between compound **5** and the receptor and the position of a calculated tightly bound water molecule (red surface marked OH_2). (a) The (*S*)-enantiomer. (b) The (*R*)-enantiomer.

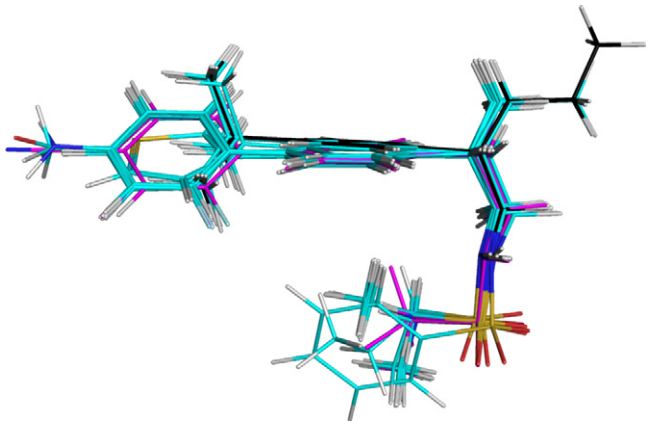


Fig. 5. Superimposition of all selected docking poses. Training set (cyan), validation set (magenta). Compound **27** of the training set is shown in black.

input for the 3D-QSAR. The data set was divided into two sets; a training set of 16 compounds (**5**–**8**, **14**–**19**, **22**–**27**) and a validation set of 3 compounds (**4**, **12** and **13**). The external validation set was selected based on structural diversity and variation within the activity range. A superimposition of the selected poses obtained from docking is shown in Fig. 5. Descriptors (molecular interaction energies) were calculated using the water (OH_2) probe and the methyl (C_3) probe in GRID and the following PLS analysis was performed using GOLPE.

The statistical properties of the resulting models are summarized in Table 2. A large number of the interaction energies calculated by GRID do not contribute to explain the observed differences in activity and can therefore be deleted since they only add noise to the PLS-model. Following the variable pre-treatment protocol described in Section 2.7, the number of *X*-variables were reduced from 21800 to 2327 without affecting the predictivity of the model (as judged by q^2). Further variable selection was performed using the smart region definition [31] in combination with the fractional factorial design procedures as implemented in GOLPE. The SRD procedure group variables close in space and collapses groups of variables containing virtually the same information to regions. These regions are then evaluated using the FFD procedure. In the FFD procedure a large number of models are calculated with some regions left out in order to evaluate the importance of the individual regions. Only regions contributing in a positive manner to the predictivity of the model are included in the final model. The FFD procedure reduced the number of variables to 1178 with a significant improvement of the predictivity of the model as determined by q^2 (from $q^2 = 0.35$ to $q^2 = 0.61$, Table 2)

Table 2
Properties of 3D-QSAR models

	Number of compounds	Number of variables	LVs	R^2	q^2	SDEP[a]
After pre-treatment	16	2327	2	0.87	0.35	0.20
After FFD selection	16	1178	2	0.89	0.61	0.17
After FFD selection	19	1212	2	0.93	0.68	N.A.

N.A.: Not applicable. LVs: Latent variables.
[a] Standard deviation on error of prediction (SDEP) for external validation set (log(10) units), SDEP = $(\sum(Y_{Exp} - Y_{Calc})^2/N)$.

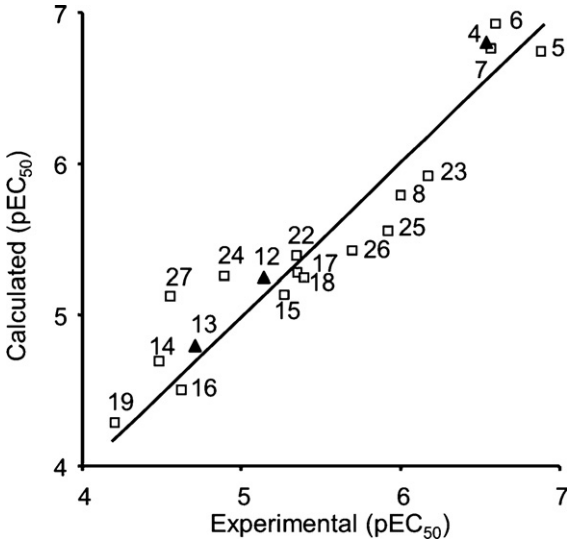


Fig. 6. Experimental and calculated activities (pEC_{50}). Training set (□), validation set (▲).

for a model with only two latent variables (LVs). The model was further tested using an external validation set (i.e. a set of compounds not included in the training set). The activities of these compounds were predicted with a standard deviation on error of prediction (SDEP) of 0.17. The experimental and calculated activities for the compounds in the training set and in the external validation set are listed in Table 1 and shown graphically in Fig. 6. The calculated activities lie within 0.4 log(10) units of the experimental values for all compounds except for compound **27**. Compound **27** is unique in the sense that the R2 *n*-propyl substituent reaches into an area that is not covered by the other compounds included in the 3D-QSAR and it can therefore not be expected to be well described by the model. Compound **27** is shown in black in Fig. 5.

It would have been obvious to use the 3D-QSAR model to predict the EC_{50} values for the compounds with inaccurately determined activities (i.e. compounds **10**, **11**, **20**, **21**, **28** and **29**). However, for compounds **20**, **21** and **28** no reasonable binding pose was obtained and compounds **10**, **11** and **29** represents "one of a kind" in relation to the compounds in the training set. The R4 substituents of compounds **10** and **11** are chemically different from the remaining R4 substituents and the R2 substituent of **29** extends beyond the borders of the 3D-QSAR model.

In order to use all available information in the following interpretation of the contour maps of the PLS-coefficients, a final model comprising the compounds from both the training and test sets ($r^2 = 0.93$, $q^2 = 0.68$) was built following the procedure described above. The contour maps of the PLS-coefficients for a model with 2 LV's are shown for the methyl probe in Fig. 7a and for the water probe in Fig. 7b. In both figures compound **5**, which is the most active compound (Table 1), is shown to illustrate the size of the regions.

The contour map of the PLS-coefficients for the methyl probe show three distinct areas of positive coefficients marked A, B and C in Fig. 7a. These areas are located close to the substituents R1, R2 and R4, respectively. An unfavorable interaction (positive interaction energy) between the molecule

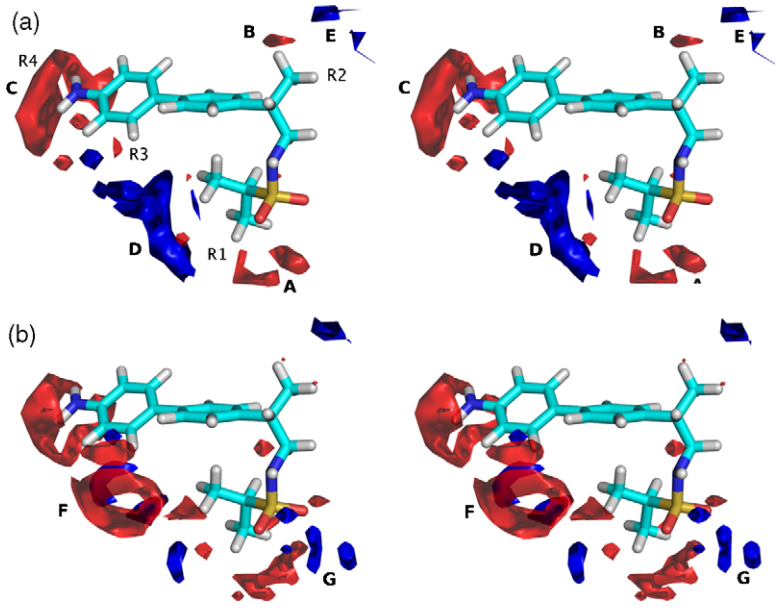


Fig. 7. Contour maps for the PLS-coefficients (2 LV's) at the +0.0018 and −0.0017 level (stereo view). (a) C_3 probe; (b) OH_2 probe. Red areas (transparent) indicate positive coefficients; blue areas indicate negative coefficients. The (*S*)-enantiomer of **5** is shown to illustrate the size of the regions.

and the methyl probe in these areas is predicted to lead to increased activity. Distinct areas with negative coefficients corresponding to the R1/R3 and the R2 substituents are marked D and E, respectively. An unfavorable interaction between a molecule and the probe in these areas is predicted to lead to decreased activity.

The contribution of the water probe to the model is ambiguous, since this probe apart from being both hydrogen bond donor and acceptor also contains a steric component. Consequently, the areas of interest for the water probe are the regions in space surrounding the molecules where the plot deviates from the corresponding plot of the methyl probe. These regions are marked F and G in Fig. 7b. Region F is an area with positive coefficients situated as a shell around the R3 substituent. Favorable electrostatic interactions (negative interaction energy) between the molecule and the probe in this area are predicted to lead to decreased activity. The region marked G in Fig. 7b is constituted of several distinct areas having negative coefficients, situated in close proximity to the oxygens of the sulfonamide moiety. A favorable electrostatic interaction (hydrogen bond) between the molecule and the probe in this area is predicted to lead to increased affinity.

3.4. Comparison of PLS-coefficient contour maps and receptor interactions for the docked compounds

If the proposed binding mode and the docking-based alignment of the compounds are valid, the positive and negative regions A-G obtained by the 3D-QSAR analysis and displayed in Fig. 7 should, as mentioned in the Introduction, represent interactions in the binding cavity.

The data set in Table 1 includes three different compound classes containing a phenyl/substituted phenyl (**4**–**21**), a thiophenyl (**22**, **23**) or a *tert*-bytyl group (**24**–**29**), respectively. The contour maps in Fig. 7 only identify variations of the activity due to structural changes in areas corresponding to the R1–R4 substituents. The 3D-QSAR results and the observed similar activities of compounds **8**, **23** and **25** are consistent with a close overlap of the core structure of the docking output poses (Fig. 5) and an equally good fit to the GRID-energy contour of the methyl probe at −2 kcal/mol (Fig. 3).

In the following, the PLS-coefficient regions A-G in Fig. 7 will be interpreted in terms of interactions to the protein shown by the substituents R1–R4 of the docked compounds.

3.4.1. R1 substituents

The 3D-QSAR analysis indicates favorable interactions between an isopropyl group and the receptor (region A in Fig. 7) whereas larger substituents cause steric repulsions (right part of region D). The GRID-energy contour of the methyl probe representing positions of energetically favorable van der Waals interactions of the binding cavity shows a pocket corresponding to the location of the R1 substituent in the docked compounds (Fig. 3). The pocket is complementary to an isopropyl substituent. Smaller substituents as methyl (**13**), ethyl (**15**) and trifluromethyl (**14**) do not utilize this pocket completely which results, as observed (Table 1), in decreased activities compared to that of isopropyl (**17**). Larger substituents do not fit the pocket, which will result in a steric clash with the receptor consistent with the low activities of compounds **16** and **19**–**21**. This lack of fit explains why reasonable binding poses of **20** and **21** having the largest R1 substituents in the series of compounds in Table 1 could not be obtained by docking.

3.4.2. R2 substituents

3D-QSAR predicts that small R2 alkyl substituents are favorable (region B in Fig. 7) and substituents larger than an ethyl group are unfavorable (region E). This rank order of activities is demonstrated by compounds **24**–**27** (Table 1). In

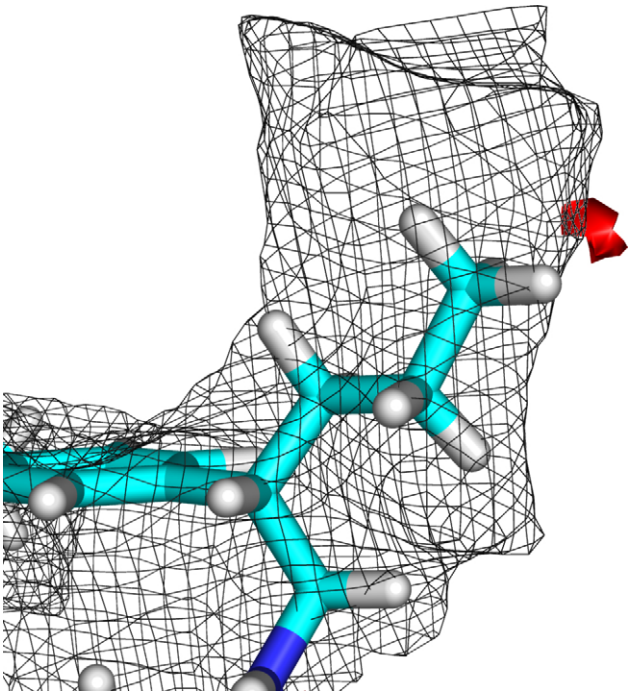


Fig. 8. The docked pose of the (*S*)-enantiomer of **27** superimposed on the interaction energy contour of the methyl probe (grey grid, −2 kcal/mol) and the water probe (red surface, −10 kcal/mol).

contrast, the GRID-energy contour of a methyl probe at −2 kcal/mol (Fig. 3) shows a large receptor cavity that clearly can accommodate the R2 propyl substituent of **27** (Fig. 8). This discrepancy may be explained by the presence of a water molecule that needs to be displaced when compounds containing large R2 substituents bind to the protein. GRID calculations using the water probe confirm the presence of a water molecule with an interaction energy of −10 kcal/mol in close proximity to the R2 propyl substituent of docked **27** as shown in Fig. 8. A water molecule is also found in this position in the aniracetam-iGluR2 [9] and other iGluR2 crystal structures [6,7]. Displacement of this tightly bound water molecule by a hydrophobic group will significantly reduce the binding affinity in accordance with the 3D-QSAR. The effect in the activity is more dramatic with larger hydrophobic R2 substituents, which will completely replace the water molecule (compounds **28** and **29**, Table 1). Thus, by taking the water molecule into account, region E in PLS-coefficient contour maps in Fig. 7 can fully be interpreted in terms of interactions between the docked ligands and the protein.

3.4.3. R3 substituents

The only non-hydrogen R3 substituent in the studied series of compounds is a fluorine atom. The 3D-QSAR analysis shows unfavorable steric (Fig. 7a, left part of region D) as well as unfavorable electrostatic interactions (Fig. 7b, region F). Inspection of the interactions between docked ligands with R3 = F and the protein reveals a short distance (2.24 Å) between the fluoro substituent and a hydrogen atom of Gly 731 indicating steric repulsion. In addition, the electronegative fluoro substituent is located in the vicinity of the backbone carbonyl oxygen atoms of Met 496 of Lys 730 (3.34 and 3.58 Å, respectively) indicating electrostatic repulsions. However, as the effect of R3 substitution is based on limited variation the results should be interpreted with care and the 4-fold decrease in activity (**8** versus **17**) could also at least partly be explained by difference in desolvation energy between the phenyl and the fluorophenyl group.

3.4.4. R4 substituents

Despite the difference in size and electrostatic properties of a cyano (**4**), an amino (**5**) and a methyl (**7**) substituent, the 3D-QSAR contour maps (Fig. 7) show that the effect of R4 substitution is mainly steric (favorable van der Waals interactions). Strong hydrogen bonds between the hydrophilic R4 substituents and the receptor are thus precluded. This indicates that the binding site should contain a pocket that can accommodate hydrophobic substituents as well as hydrogen bond donor and acceptor groups of different sizes. In agreement with this, Fig. 9 shows that the R4 substituent is positioned in a large pocket of the protein and GRID calculations (water probe) with a docked modulator in the binding site, exemplified by compound **5** in Fig. 9, show that a network of loosely bound water molecules solvates the R4 substituent. The water network may adapt to the properties of the R4 substituent explaining the similar activities of compounds **4**–**7**.

The chloro and trifluoromethyl derivatives, **10** and **11** (Table 1) are more than 20-fold less active than **5** and a possible explanation can be found in the electrostatic properties of the surrounding receptor atoms in analogy to the R3 substituent. The partially negatively charged backbone carbonyl oxygen atoms of Pro 494, Phe 495 and Ser 729 create an unfavorable environment for aromatic trifluoromethyl and chloro substituents.

As mentioned above, a reasonable docked pose for compound **9** could not be obtained despite its high activity (Table 1). The compound was docked with the R4 carboxyl substituent in neutral as well as in anionic form. The lack of explicit water molecules in the large cavity surrounding the R4 substituent (Fig. 9) results in strong interactions between the R4 substituent

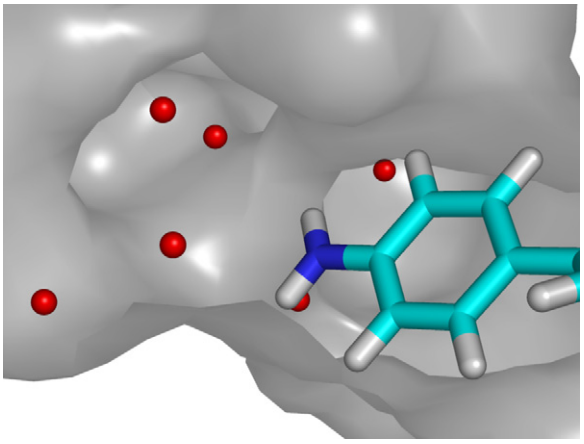


Fig. 9. The docked pose of the (*S*)-enantiomer of **5** and calculated water molecules in the binding site of the protein represented as the surface of nearby residues.

and hydrogen bond acceptors and donors in the large cavity, which drastically distorts the binding poses obtained.

4. Conclusions

A binding mode in the dimer interface of the iGluR2/4 receptor for allosteric AMPA modulators of the biarylpropylsulfonamides class of compounds have been proposed based on docking, analysis of hydrogen bonding patterns, calculation of GRID molecular interaction fields and calculation of conformational energy penalties. A predictive 3D-QSAR model was obtained based on the docking output poses. Interpretation of PLS-coefficient regions in terms of interactions with the protein shown by the substituents R1–R4 of the docked compounds validates the deduced binding mode of the compounds.

The proposed binding site of the modulators overlaps with those of other allosteric AMPA modulators known from crystal structures. A direct interaction is observed between the modulators and Ser 754, which has previously been shown to affect their potentiating effect. This interaction may explain the flip preference of **4** (LY404187).

In this study, a combination of several modeling approaches was utilized to investigate the binding mode of the biarylpropylsulfonamide modulators. The results from the individual methods are consistent, serving as a mutual internal validation, and supplement each other.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2007.06.002.

References

[1] M.D. Black, Therapeutic potential of positive AMPA modulators and their relationship to AMPA receptor subunits. A review of preclinical data, Psychopharmacology 179 (2005) 154–163.
[2] S. Matsuda, Y. Kamiya, M. Yuzaki, Roles of the N-terminal domain on the function and quaternary structure of the ionotropic glutamate receptor, J. Biol. Chem. 280 (2005) 20021–20029.
[3] W. Tichelaar, M. Safferling, K. Keinanen, H. Stark, D.R. Madden, The three-dimensional structure of an ionotropic glutamate receptor reveals a dimer-of-dimers assembly, J. Mol. Biol. 344 (2004) 435–442.
[4] M. Mansour, N. Nagarajan, R.B. Nehring, J.D. Clements, C. Rosenmund, Heteromeric AMPA receptors assemble with a preferred subunit stoichiometry and spatial arrangement, Neuron 32 (2001) 841–853.
[5] E. Gouaux, Structure and function of AMPA receptors, J. Physiol. - London 554 (2004) 249–253.
[6] N. Armstrong, E. Gouaux, Mechanisms for activation and antagonism of an AMPA-Sensitive glutamate receptor: crystal structures of the GluR2 ligand binding core, Neuron 28 (2000) 165–181.
[7] Y. Sun, R. Olson, M. Horning, N. Armstrong, M. Mayer, E. Gouaux, Mechanism of glutamate receptor desensitization, Nature 417 (2002) 245–253.
[8] A. Arai, G. Lynch, The waveform of synaptic transmission at hippocampal synapses is not determined by AMPA receptor desensitization, Brain Res. 799 (1998) 230–234.
[9] R.S. Jin, S. Clark, A.M. Weeks, J.T. Dudman, E. Gouaux, K.M. Partin, Mechanism of positive allosteric modulators acting on AMPA receptors, J. Neurosci. 25 (2005) 9027–9036.
[10] K.M. Partin, M.W. Fleck, M.L. Mayer, AMPA receptor flip/flop mutants affecting deactivation, desensitization, and modulation by cyclothiazide, aniracetam, and thiocyanate, J. Neurosci. 16 (1996) 6634–6647.
[11] A.C. Arai, M. Kessler, G. Rogers, G. Lynch, Effects of the potent ampakine CX614 on hippocampal and recombinant AMPA receptors: Interactions with cyclothiazide and GYKI 52466, Mol. Pharmacol. 58 (2000) 802–813.
[12] B. Sommer, K. Keinanen, T.A. Verdoorn, W. Wisden, N. Burnashev, A. Herb, M. Kohler, T. Takagi, B. Sakmann, P.H. Seeburg, Flip and flop—a Cell-Specific Functional Switch in Glutamate-Operated Channels of the Cns, Science 249 (1990) 1580–1585.
[13] P.J. Baumbarger, M. Muhlhauser, J. Zhai, C.R. Yang, E.S. Nisenbaum, Positive modulation of α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) receptors in prefrontal cortical pyramidal neurons by a novel allosteric potentiator, J. Pharmacol. Exp. Ther. 298 (2001) 86–102.
[14] J.C. Quirk, E.S. Nisenbaum, Multiple molecular determinants for allosteric modulation of alternatively spliced AMPA receptors, J. Neurosci. 23 (2003) 10953–10962.
[15] J.A. Morrow, J.K. Maclean, C. Jamieson, Recent advances in positive allosteric modulators of the AMPA receptor, Curr. Opin. Drug Discov. Devel. 9 (2006) 571–579.
[16] P.L. Ornstein, D.M. Zimmerman, M.B. Arnold, T.J. Bleisch, B. Cantrell, R. Simon, H. Zarrinmayeh, S.R. Baker, M. Gates, J.P. Tizzano, D. Bleakman, A. Mandelzys, K.R. Jarvie, K. Ho, M. Deverill, R.K. Kamboj, Biarylpropylsulfonamides as novel, potent potentiators of 2-amino-3-(5-methyl-3-hydroxyisoxazol-4-yl)-propanoic acid (AMPA) receptors, J. Med. Chem. 43 (2000) 4354–4358.
[17] R.A. Friesner, J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, M.P. Repasky, E.H. Knoll, M. Shelley, J.K. Perry, D.E. Shaw, P. Francis, P.S. Shenkin, Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, J. Med. Chem. 47 (2004) 1739–1749.
[18] G.L. Warren, C.W. Andrews, A.M. Capelli, B. Clarke, J. LaLonde, M.H. Lambert, M. Lindvall, N. Nevins, S.F. Semus, S. Senger, G. Tedesco, I.D. Wall, J.M. Woolven, C.E. Peishoff, M.S. Head, A critical assessment of docking programs and scoring functions, J. Med. Chem. 49 (2006) 5912–5931.
[19] P.J. Goodford, A computational-procedure for determining energetically favorable binding-sites on biologically important macromolecules, J. Med. Chem. 28 (1985) 849–857.
[20] J. Boström, P.O. Norrby, T. Liljefors, Conformational energy penalties of protein-bound ligands, J. Comput. Aid. Mol. Des. 12 (1998) 383–396.
[21] M. Baroni, G. Costantino, G. Cruciani, D. Riganelli, R. Valigi, S. Clementi, Generating optimal linear PLS estimations (GOLPE)—an advanced chemometric tool for handling 3D-QSAR problems, Quant. Struct. -Act. Relat. 12 (1993) 9–20.
[22] B. Bordas, I. Belai, A. Lopata, Z. Szanto, Interpretation of scoring functions using 3D molecular fields. Mapping the diacyl-hydrazine-binding pocket of an insect ecdysone receptor, J. Chem. Inf. Model. 47 (2007) 176–185.
[23] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The protein data bank, Nucleic Acids Res. 28 (2000) 235–242.
[24] A. Bairoch, R. Apweiler, C.H. Wu, W.C. Barker, B. Boeckmann, S. Ferro, E. Gasteiger, H.Z. Huang, R. Lopez, M. Magrane, M.J. Martin, D.A. Natale, C. O'Donovan, N. Redaschi, L.S.L. Yeh, The universal protein resource (UniProt), Nucleic Acids Res. 33 (2005) D154–D159.
[25] C. Notredame, D.G. Higgins, J. Heringa, T-Coffee: a novel method for fast and accurate multiple sequence alignment, J. Mol. Biol. 302 (2000) 205–217.
[26] Maestro, version 7.5.106, 2006, Schrödinger Inc., Portland, OR.
[27] Macromodel, version 9.1.106, 2006, Schrödinger Inc., Portland, OR.
[28] GRID, version 22a, 2005, Molecular Discovery Ltd., Pinner, Middlesex.
[29] Glide, version 4.0.108, 2006, Schrödinger Inc., Portland, OR.
[30] GOLPE, version 4.5.12, 2002, Multivariate Infometric Analysis S.r.l., Perugia, Italy.

[31] M. Pastor, G. Cruciani, S. Clementi, Smart region definition: a new way to improve the predictive ability and interpretability of three-dimensional quantitative structure-activity relationships, J. Med. Chem. 40 (1997) 1455–1464.

[32] P. Miu, K.R. Jarvie, V. Radhakrishnan, M.R. Gates, A. Ogden, P.L. Ornstein, H. Zarrinmayeh, K. Ho, D. Peters, J. Grabell, A. Gupta, D.M. Zimmerman, D. Bleakman, Novel AMPA receptor potentiators LY392098 and LY404187: effects on recombinant human AMPA receptors in vitro, Neuropharmacology 40 (2001) 976–983.

[33] K. Harpsøe, T. Varming, A.H. Gouliaev, D. Peters, T. Liljefors, Identification of a putative binding site for 5-alkyl-benzothiadiazides in the AMPA receptor dimer interface, J. Mol. Graph. Model. 26 (2007) 213–225.