# **Supporting Information**

# New insights into the GABA<sub>A</sub> receptor structure and orthosteric ligand binding: Receptor modeling guided by experimental data

Tommy Sander, Bente Frølund, Anne Techau Bruun, Ivaylo Ivanov, J. Andrew McCammon, and Thomas Balle\*.

\* Corresponding author. Email: tb@farma.ku.dk.

# **METHODS**

# Model building

The homology modeling process was conducted in three steps as outlined below: 1) generation of initial models, 2) sampling of the  $\beta$ 5-L5' loop and F-loop, and 3) building the refined model based on the two previous steps. The program Modeller  $9v7^1$  was used in all runs and was setup to perform a thorough (*slow*) refinement. When building the entire  $\alpha_1\beta_2\gamma_2$  pentamer, i.e. steps 1 and 3, Modeller's *automodel* class was used. Symmetry restraints were imposed on the refinement process to ensure (nearly) identical pairs of  $\beta_2\alpha_1$  interfaces, defined as:  $\beta_2$  residues D24-P34, V53-D56, D95-F105, A135-R141, E153-I164 and K196-R207, and  $\alpha_1$  residues D9-D19, L22, D43-P51, T60-R66, V82-M89, K92, N110-L117, L127-T133 and T171-Y190. For loop sampling, the *loopmodel* class in Modeller was used.

- 1) Four hundred initial models were generated based on the alignment shown in main article Figure 2, so that only certain templates were used in certain regions, as indicated in the figure. Residue specific restraints were imposed as outlined in Table S.I.-1. Two models (no. 53 and 226, referred to as models 1a and 1b, respectively) were selected for further use. Using the amino acid rotameric library implemented in the molecular modeling program Maestro v. 9.0,<sup>2</sup> residue R66 in all  $\alpha_1$  chains was put in a conformation so that its guanidinium head group was inside the putative binding pocket, in accordance with reports in the literature.<sup>3-5</sup> A resulting steric repulsion with  $\alpha_1$  F45 was relieved by putting this residue in another energetically favorable conformation, likewise from Maestro's rotameric library, with a  $\chi_1$  torsion angle of ~180° (identical to what is seen for the similar Y23 residue in the GLIC structure aligning with this position in GABA<sub>A</sub>R).
- 2) In the second step, three rounds of loop sampling of the initial models were performed, namely, a) of the F-loop of the  $\alpha_1$  subunit, b) of the  $\beta$ 5-L5' loop segment in the  $\alpha_1$  subunit, and c) of the  $\beta$ 5-L5' loop segment in the  $\gamma_2$  subunit. In each run, 500 loop models were generated.

- a. Sampling of the F-loop (residues  $\alpha_1$  E169-L187) was performed on the  $\beta_2$ - $\alpha_1$  dimer (chains A-B) of the initial model 1a, imposing residue specific distance restraints as specified in Table S.I.-1. The  $\alpha_1$  subunit of the best model chosen for further use (no. 417) is referred to as model 2a.
- b. Sampling of the  $\alpha_1 \beta$ 5-L5' loop (residues H109-K116) was performed on the  $\beta_2$ - $\alpha_1$ - $\beta_2$  trimer (chains A-B-C) of the initial model 1b. The  $\alpha_1$  subunit of the resulting best model (no. 381) is referred to as model 2b.
- c. The loop conformation of model 2b was inserted in the  $\alpha_1$  and  $\beta_2$  subunits of the initial model 1b, and in  $\beta_2$  chains the loop residues were manually mutated to those found in the  $\beta_2$  sequence. The  $\alpha_1$ - $\gamma_2$ - $\beta_2$  trimer (chains D-E-A) from this structure was now used in a similar round of sampling of the  $\gamma_2 \beta_5$ -L5' loop (residues A119-M130). Residue specific restraints were imposed on the  $\gamma_2$  loop sampling as outlined in Table S.I.-1. The  $\gamma_2$  subunit of the model selected as the best from this sampling (no. 358) is referred to as model 2c.
- 3) In the final homology modeling step, 400 models of the GABA<sub>A</sub> receptor EC domain were built using the selected models from steps 1 and 2 as templates. The primary templates were model 1a for  $\beta_2$ , and model 1b for  $\alpha_1$  and  $\gamma_2$  subunits because the best geometry and ProSA z-scores were obtained for the respective model-subunit combinations. For a few short regions the roles were switched so that model 1a was used for  $\alpha$ , and model 1b for  $\beta_2$ , because here the backbone conformation was flipped for some residues relative to the original structural templates. Model 2a acted as template for the Floop in all five subunits, model 2b as template for the  $\beta$ 5-L5' loop in  $\alpha_1$  and  $\beta_2$  chains, and model 2c as template for the  $\beta$ 5-L5' loop in  $\gamma_2$ . Specific details of which templates were used in which regions can be found in Fig. S.I.-1. Of the generated models, no. 193 was selected as the overall best (see Table S.I.-2) and is referred to as the refined model (see selection criteria in the main article). Backbone conformations resulting in Ramachandran plot violations (three residues in loop regions, none near the orthosteric or BZD binding sites) were manually adjusted to the proper configuration using the "rotate peptide plane" tool in Maestro. The model was treated according to the Protein Preparation procedure<sup>6</sup> using exhaustive sampling of H-bond networks and otherwise standard settings. This procedure adds hydrogen atoms, connects disulfides, probes the optimal flip orientation and tautomeric state of glutamine, asparagine and histidine residues, optimizes H-bond networks, and performs a geometry optimization to a maximum RMSD of 0.3Å. The Nand C-termini were modeled in their protonated ammonium and deprotonated carboxylate states, respectively, and all basic and acidic residues were in their default protonated states (neutral His, protonated Arg and Lys, and deprotonated Asp and Glu).

# **Molecular Dynamics (MD)**

An MD simulation of the refined GABA<sub>A</sub>R model after the above mentioned geometry optimization was set up and carried out in explicit solvent consisting of TIP3P type<sup>7</sup> water molecules. Initially, the program GRID<sup>8,9</sup> was used to place water molecules in favorable areas throughout the structure, using the water probe (OH2) and a grid spacing of 0.5 Å (*NPLA=2*) to calculate interaction energies between protein and water, and minima below -5 kcal/mol were occupied. A total of 1855 water molecules were added and subsequently geometry optimized around the model with all protein atoms kept frozen in space. This complex was centered in a cubic box with side lenghts of 114.5 Å, corresponding to a 15 Å buffer distance between the protein and each side of the box, which was filled with an additional 40,214 waters as well as with 113 Na<sup>+</sup> and 112 Cl<sup>-</sup> ions to neutralize the system and afford a ~0.15M solution resembling physiological conditions. The system was setup with periodic boundary conditions and comprised in total 42,258 molecules, or 143,490 atoms.

The 12-step MD protocol comprising two rounds of energy minimization, nine equilibration steps and a production run was as follows:

- Minimization with 50 kcal/mol/Å<sup>2</sup> position restraints on protein atoms. 10 steps of steepest descent followed by conjugate gradient until convergence at 50 kcal/mol/Å.
- 2) Minimization like above but without position restraints. Convergence at 5 kcal/mol/Å.
- 12 ps at 10 K in the NVT ensemble, time steps of 1:1:3 fs (bonded:near:far interactions), protein heavy atoms restrained with 50 kcal/mol/Å<sup>2</sup>.
- 4) 24 ps at 10 K, protein heavy atoms restraints kept.
- 5) 60 ps at 300 K, protein heavy atoms restraints kept.
- 6) 120 ps, protein heavy atoms restraints lowered to 5 kcal/mol/ $Å^2$ .
- 7) 120 ps, protein Ca's restrained with 5 kcal/mol/Å<sup>2</sup>, side chain heavy atoms with 1 kcal/mol/Å<sup>2</sup>.
- 8) 120 ps, Cα retraints kept, side chains released.
- 9) 120 ps, C $\alpha$  retraints lowered to 3 kcal/mol/Å<sup>2</sup>.
- 10) 120 ps, C $\alpha$  retraints lowered to 1 kcal/mol/Å<sup>2</sup>.
- 11) 9.6 ns, C $\alpha$  retraints lowered to 0.25 kcal/mol/Å<sup>2</sup>.
- 12) 48 ns production run.

As is the default in the used MD program (Desmond), pressure and temperature during the equilibration steps 3-11 were coupled with the Berendsen methods<sup>10</sup> (only temperature in step 3 where the default constant-NVT ensemble was used). During the production run, the Nosé-

Hoover thermostat<sup>11</sup> and Martyna-Tobias-Klein barostat<sup>12</sup> were used. Unless specifically stated above, default Desmond settings were used: OPLS 2005 force field, constant-NPT ensemble at 1 atm pressure and 300 K temperature, 2:2:6 fs integration steps (bonded:near:far interactions, where "far" is > 9Å), constrained covalent bonds between heavy atoms and hydrogens using the SHAKE algorithm,<sup>13</sup> and smooth particle mesh Ewald electrostatics<sup>14</sup> beyond the short range Coulombic interactions cutoff at 9Å. The use of shorter integration time steps in step 3 is standard procedure in Desmond in the initial heating step.

Similar MD simulations were performed for the *Ac*-AChBP apo structure (PDB: 2byn) and the EC domain of the ELIC structure (PDB: 2vl0, chains A-E). Non-protein molecules and moieties were deleted from the *Ac*-AChBP structure, as were the N-terminal FLAG epitopes so that each chain sequence started with 1-HSQ. For the ELIC structure, the transmembrane part starting at residues 200-PSY was deleted from each chain in the original PDB structure. The two pentamers were treated like the GABA<sub>A</sub>R model as described above (protein preparation, GRID solvation etc.). The MD protocol outlined above was followed with the exceptions that, 1) the production runs were terminated after 30.0 ns for *Ac*-AChBP and 35.6 ns for ELIC, and 2) for *Ac*-AChBP, equibration steps 9-11 were only run for 60 ps, 60 ps, and 120 ps, respectively.

# DISCUSSION

#### Alignment considerations and alterations

A proper sequence alignment is perhaps the single most critical step in homology modeling.<sup>15</sup> In some regions of the GABA<sub>A</sub>R sequences there is little consensus in the literature as to how they should align with the other Cys-loop receptors.<sup>16-23</sup> With sequence identities between the GABA<sub>A</sub>R and templates falling in the low range of 13-18% (calculated with BioEdit<sup>24</sup> using the alignment in Figure 2 of the main text), special attention was required in this phase. We settled on an iterative protocol where a thoroughly verified structural alignment of the templates was expanded by adding all human nAChR sequences because of their close relation to the mouse and *Torpedo* nAChR structures, and finally adding the human GABA<sub>A</sub>R sequences. Motifs that are conserved among the different Cysloop receptor subfamilies were thus better identified. However, we still found that manual alterations were needed because such motifs were not always properly caught by the automated procedure (in our case the ClustalX program), and because certain regions with poor sequence similarity were inappropriately aligned.

As stated in the Methods section of the main manuscript, manual adjustments were made in four regions, namely, 1) in and after the N-terminal  $\alpha$ -helix, 2) in the L5- $\beta$ 5' segment, 3) in loop F, and 4) in loop C. The rationale behind this is described in the following (a detailed comparison of the sequence alignments before and after the alterations is given in Fig. S.I.-2).

At the N-terminal part highly conserved motifs corresponding to the GABA<sub>A</sub>R  $\alpha_1$  sequence <sup>17</sup><u>ILDRLLDGYDNRLRP</u> were misaligned by ClustalX. This was due to the presence of sequences in the alignment with insertions in this region (*Bt*-AChBP, *Torpedo*  $\delta$  subunit, and human nAChR  $\beta_2$ ,  $\beta_4$  and  $\delta$  subunits). Manually altering the alignment not only reestablished these motifs at the sequence level, but also facilitated that the highly conserved Tyr residue corresponding to  $\alpha_1$  Y25 was positioned in the model so that it interacts with the likewise highly conserved  $\alpha_1$  D71 at the end of the  $\beta_2$  strand, which would otherwise pack in hydrophobic surroundings. The same interaction is observed in the mouse nAChR  $\alpha_1$ structure, and hence we are confident that the altered alignment is correct.

In the L5- $\beta$ 5' segment, between the absolutely conserved residues P96 and G124 (GABA<sub>A</sub>R  $\alpha_1$  numbering), two-four extra residues are found in all GABA<sub>A</sub>R subunits compared to the other Cys-loop sequences. We kept the insertions before the  $\beta$ 5' strand in order to align  $\alpha_1$  R119 (conserved among all GABA<sub>A</sub>R subunits) with *Ls*-AChBP R104, hence facilitating the salt bridge formation described above for e.g. the  $\alpha_1$  R119 $\cdots\beta_2$  D163 ion pair. The actual position of the insertion was chosen arbitrarily and probably had minimal impact on the final model due to the subsequent loop sampling performed on this segment.

The long loop F between  $\alpha_1$  W170 and  $\alpha_1$  Y190 is particularly challenging as it has virtually no conserved regions among the Cys-loop family members, and in addition the structural templates have markedly different conformations of this loop. Therefore, we used the cysteine accessibility data reported by Newell & Czajkowski<sup>25</sup> to guide the alignment so that the initial models were in reasonable accordance with the observations of that study. Like for the  $\beta 5$ - $\beta 5$ ' segment this has probably only had minor impacts on the final structure because of the subsequent loop sampling. As our primary focus was on the orthosteric binding site, no attempts were made to optimize loop F in a similar way specifically for the  $\beta_2$  and  $\gamma_2$  subunits. We therefore chose to let their sequences follow that of the  $\alpha_1$  as closely as possible while making sure that hydrophobic packing of hydrophilic/charged residues did not occur.

Loop C in GABA<sub>A</sub>R subunits is two to three residues shorter than in the relevant templates (i.e. the AChBPs, mouse  $\alpha_1$ , and *Torpedo*  $\alpha$  subunit) and hence presents an important challenge in the alignment process. Main fix points were  $\beta_2$  F200 and Y205 which align with highly conserved aromatic residues. The *Ac*-AChBP segment <sup>188</sup>YSCCPEPY was aligned with  $\beta_2$  <sup>200</sup>FS--TGSY, rather than <sup>200</sup>FSTGS--Y as in the initial ClustalX alignment. This afforded a C-loop with good backbone geometry (i.e. in favored/allowed areas in the

Ramachandran plot) and with  $\beta_2$  T202 facing the binding site. This is in apparent accordance with experimental data demonstrating that in this position a residue with a hydroxyl group (Thr or Ser) is pivotal for activating the receptor with GABA or muscimol.<sup>26,27</sup>

It should be noted that the vast amount of experimental data that have been published for the GABA<sub>A</sub>R receptor is not always easy to interpret, may sometimes appear contradictory, and hence should be interpreted with caution. For instance, in one of the studies referred to above for loop C,<sup>27</sup> residues  $\beta_2$  F200-T202 were found to be solvent inaccessible, and the authors further concluded that they do not line the binding site. This is not compatible with our model. However, regardless of which alignment (if any) is correct for the above mentioned loop C stretch, the  $\beta_2$  F200 residue does with high probability align with a tyrosine conserved among AChBP and nAChR sequences (and also ELIC) that clearly lines the binding site and directly engages in ligand binding. Similar uncertainties likely apply to the other parts of the model where such data were taken into consideration, e.g. in  $\alpha_1$  loop F. However for this region, in contrast to loop C, we have neither a reliable structural basis from the templates nor any contradictory results to question the biochemical data of Newell & Czajkowski,<sup>25</sup> so here priority was given to optimize adherence to their results.

In general we believe that the above summarized manual adjustments to the automatically generated sequence alignment are reasonable and have resulted in a more realistic  $GABA_AR$  model.

Atom 1 <sup>a</sup>	Atom 2	Form <sup>b</sup>	Distance (std.dev.)
Pentar	mer generation re	straints (ste	ps 1 and 3)
$\alpha_1$ D54 $C\gamma^{c}$	α <sub>1</sub> R220 Cζ	G	4.2 Å (0.1 Å)
$\alpha_1 \operatorname{R119} C \zeta^d$	β <sub>2</sub> D163 <i>C</i> γ	G	4.2 Å (0.1 Å)
β <sub>2</sub> E153 <i>Cδ</i>	β <sub>2</sub> K196 Nζ	G	3.3 Å (0.1 Å)
$\alpha_1 \text{ R73 } C \zeta^d$	α <sub>1</sub> L118 <i>Cβ</i>	L.b.	18 Å (1 Å)
β <sub>2</sub> R86 <i>C</i> ζ	β <sub>2</sub> L118 <i>Cβ</i>	L.b.	9 Å (0.5 Å)
$\beta_2 \ Y97 \ C\zeta^d$	β <sub>2</sub> L99 <i>Cβ</i>	U.b.	4.2 Å (0.1 Å)
$\beta_2$ Y97 $C\delta 2^{d}$	β <sub>2</sub> E155 <i>C</i> α	U.b.	5.0 Å (0.1 Å)
$\beta_2$ D95 $O\delta l^{d}$	$\beta_2$ S156 $O\gamma$	U.b.	2.8 Å (0.1 Å)
$\beta_2D95\;\textit{O\delta2}^{d}$	β <sub>2</sub> Y157 N	U.b.	3.0 Å (0.1 Å)
$\beta_2D95\;\textit{O\delta2}^{d}$	$\beta_2 G158 N$	U.b.	3.0 Å (0.1 Å)
	F-loop sampli	ng restraint	S
$\alpha_1$ W170 C $\zeta$ 3 <sup>d,e</sup>	α <sub>1</sub> I44 <i>Cβ</i>	U.b.	5.0 (0.1 Å)
α <sub>1</sub> Α175 <i>Cβ</i>	α <sub>1</sub> I44 <i>Cβ</i>	U.b.	5.5 Å (0.3 Å)
α <sub>1</sub> S177 <i>Cβ</i>	$\beta_2$ Y157 Ca	L.b.	21 Å (0.3 Å)
α <sub>1</sub> V178 <i>Cβ</i>	$\beta_2$ Y157 Ca	U.b.	17.5 Å (0.3 Å)
α <sub>1</sub> V178 <i>Cβ</i>	α <sub>1</sub> V197 <i>Cβ</i>	L.b.	11 Å (0.3 Å)
α <sub>1</sub> V179 <i>Cβ</i>	α <sub>1</sub> I44 <i>Cβ</i>	U.b.	9.0 Å (0.3 Å)
α <sub>1</sub> V179 <i>Cβ</i>	$\alpha_1$ V46 <i>C</i> $\beta$	L.b.	5.5 Å (0.3 Å)
α <sub>1</sub> V180 <i>Cβ</i>	$\beta_2$ Y157 Ca	U.b.	20 Å (0.3 Å)
α <sub>1</sub> V180 <i>Cβ</i>	$\beta_2 R207 C \alpha$	U.b.	20 Å (0.3 Å)
α <sub>1</sub> V180 <i>Cβ</i>	$\alpha_1$ V46 <i>C</i> $\beta$	L.b.	9.0 Å (0.3 Å)
α <sub>1</sub> V180 <i>Cβ</i>	α <sub>1</sub> L192 <i>Cβ</i>	L.b.	8.0 Å (0.3 Å)
α <sub>1</sub> Α181 <i>Cβ</i>	β <sub>2</sub> L99 <i>Cβ</i>	L.b.	17 Å (0.3 Å)
α <sub>1</sub> D183 <i>C</i> γ	$\beta_2 R207 C \alpha$	U.b.	15.5 Å (0.3 Å)
	$\beta$ 5- $\beta$ 5' loop samp	ling (γ <sub>2</sub> sub	unit)
γ <sub>2</sub> W123 <i>Cζ3</i>	γ <sub>2</sub> L143 <i>Cδ1</i>	U.b.	5.0 Å (0.1 Å)
$\gamma_2$ I124 Ca	γ <sub>2</sub> R144 <i>Cβ</i>	U.b.	7.0 Å (0.1 Å)
γ <sub>2</sub> M130 <i>Cβ</i>	γ <sub>2</sub> T142 <i>Cβ</i>	U.b.	4.5 Å (0.1 Å)

Table S.I.-1. Atom specific distance restraints imposed in model building.

<sup>a</sup> Atom specification format: subunit - residue - PDB atom name.

<sup>b</sup> The Modeller distance restraint type enforced. G, Gaussian; U.b., Upper bound; L.b., Lower bound. <sup>c</sup> Also set for the corresponding residues/atoms in  $\beta_2$  subunits. <sup>d</sup> Also set for the corresponding residues/atoms in the other subunits.

<sup>e</sup> Also imposed as restraint in the final model building step

Table S.I.-2. Modeller *molpdf* and DOPE scores, ProSA z-scores, and OPLS 2001 energies of all homology models top ranked in either Modeller score or within top 50 of both. Highlighted rows indicate models selected for further use from each run.

Model	Modelle	er scores		(by chai	ProSA	z-score eighted a	verage)		OPLS 2001 energy
no.	<i>molpdf</i> <sup>a</sup>	DOPE	$A / \beta_2$	$\mathbf{B}$ / $\alpha_1$	$C$ / $\beta_2$	$D / \alpha_1$	$E$ / $\gamma_2$	Avg <sup>b</sup>	(kcal/mol) <sup>c</sup>
			Initi	al model	generatio	on			
53 (1a)	29715	-112206	-3.80	-4.92	-3.85	-4.43	-4.43	-4.29	478
79	29909	-111935	-3.72	-4.92	-3.72	-5.05	-4.52	-4.39	559
99	29809	-112522	-3.67	-5.15	-3.84	-5.06	-4.34	-4.41	483
163	29850	-112693	-3.41	-4.88	-3.90	-4.91	-4.46	-4.31	478
205	29490	-110644	-3.82	-4.39	-3.92	-4.31	-4.74	-4.24	543
226 (1b)	29697	-112143	-3.66	-5.17	-3.76	-5.05	-4.68	-4.47	470
262	30245	-113466	-3.80	-5.17	-3.82	-5.06	-4.60	-4.49	494
367	29621	-111940	-3.87	-5.04	-3.77	-4.97	-4.48	-4.43	501
398	29899	-111993	-3.77	-4.79	-3.79	-4.54	-4.62	-4.30	502
			F-loop	sampling	g (α <sub>1</sub> subu	nit) <sup>d</sup>			
24	68.7	-1799	-	-5.32	-	-	-	-	203
33	69.3	-1866	-	-5.38	-	-	-	-	188
38	77.1	-1878	-	-5.38	-	-	-	-	188
70	70.3	-1866	-	-5.29	-	-	-	-	190
74	60.4	-1874	-	-5.20	-	-	-	-	189
98	78.5	-1803	-	-5.14	-	-	-	-	189
195	48.5	-1849	-	-5.33	-	-	-	-	192
207	59.2	-1902	-	-5.17	-	-	-	-	186
219	80.7	-1811	-	-5.18	-	-	-	-	192
225	61.2	-1834	-	-5.36	-	-	-	-	196
250	82.9	-1825	-	-5.13	-	-	-	-	193
271	50.8	-1929	-	-5.12	-	-	-	-	191
275	70.7	-1908	-	-5.01	-	-	-	-	193
276	58.8	-1900	-	-5.22	-	-	-	-	187
313	77.7	-1816	-	-5.24	-	-	-	-	189
345	59.2	-1781	-	-4.98	-	-	-	-	192
351	64.5	-1882	-	-5.28	-	-	-	-	191
407	12.6	-1793	-	-5.17	-	-	-	-	187
417 (2a)	75.6	-1966	-	-5.41	-	-	-	-	185
436	68.5	-1916	-	-5.32	-	-	-	-	188

Model	Modelle	er scores		(by chai	ProSA	z-score eighted a	verage)		OPLS 2001 energy
no.	molpdf <sup>a</sup>	DOPE	A / $\beta_2$	$\mathbf{B}$ / $\alpha_1$	$C$ / $\beta_2$	$\mathbf{D}$ / $\boldsymbol{\alpha}_1$	$E$ / $\gamma_2$	Avg <sup>b</sup>	(kcal/mol) <sup>c</sup>
469	76.1	-1810	-	-5.34	-	-	-	-	193
471	320.0	-1978	-	-5.27	-	-	-	-	239
		/	B5-L5'lo	op sampl	ing (α <sub>1</sub> su	bunit) <sup>e</sup>			
9	22.4	-1121	-	-5.11	-	-	-	-	295
10	12.4	-1250	-	-5.07	-	-	-	-	297
64	25.3	-1155	-	-4.90	-	-	-	-	293
96	21.2	-1152	-	-5.08	-	-	-	-	297
97	$1095.0^{\rm f}$	-1305	-	-5.01	-	-	-	-	302
108	8.4	-1190	-	-5.09	-	-	-	-	299
115	16.2	-1185	-	-5.08	-	-	-	-	299
117	20.7	-1237	-	-5.13	-	-	-	-	293
122	19.8	-1206	-	-5.08	-	-	-	-	297
183	23.6	-1166	-	-4.87	-	-	-	-	297
317	23.7	-1180	-	-5.10	-	-	-	-	294
318	23.2	-1195	-	-5.13	-	-	-	-	294
355	13.6	-1249	-	-5.05	-	-	-	-	299
364	22.5	-1187	-	-5.03	-	-	-	-	296
381 (2b)	26.9	-1202	-	-5.13	-	-	-	-	291
414	17.2	-1184	-	-5.10	-	-	-	-	293
465	22.0	-1190	-	-5.13	-	-	-	-	298
			в5-L5'lo	op sampl	ing (y <sub>2</sub> su	bunit) <sup>g</sup>			
6	150.9	-2161	-	-	-	-	-4.82	-	305
49	23.7	-1984	-	-	-	-	-4.61	-	289
146	47.1	-2122	-	-	-	-	-4.81	-	301
165	27.5	-2116	-	-	-	-	-4.86	-	296
194	35.1	-2073	-	-	-	-	-4.72	-	284
321	33.1	-2068	-	-	-	-	-4.82	-	300
358 (2c)	32.6	-2143	-	-	-	-	-4.79	-	282
447	41.4	-2105	-	-	-	-	-4.90	-	290
499	44.9	-2089	-	-	-	-	-4.87	-	288
			Refin	ed model	l generati	on			
99	5611	-112296	-3.75	-5.32	-4.11	-5.75	-4.47	-4.68	491
100	5709	-112046	-4.01	-5.29	-4.08	-5.73	-4.58	-4.74	467

Model	Modelle	er scores		(by chai	ProSA in and wo	z-score eighted a	verage)		OPLS 2001 energy
no.	<i>molpdf</i> <sup>a</sup>	DOPE	A / $\beta_2$	$\mathbf{B}$ / $\alpha_1$	$C$ / $\beta_2$	$D / \alpha_1$	$E$ / $\gamma_2$	Avg <sup>b</sup>	(kcal/mol) <sup>c</sup>
112	5762	-112740	-4.25	-5.32	-4.09	-5.65	-4.78	-4.82	420
193	5651	-112350	-4.37	-5.30	-4.07	-5.83	-4.51	-4.82	413 <sup>h</sup>
354	5721	-112340	-3.95	-5.13	-4.03	-5.59	-4.44	-4.63	430
358	5542	-111242	-4.00	-4.93	-4.13	-5.47	-4.52	-4.61	488
359	5729	-112032	-3.82	-5.22	-4.05	-5.72	-4.50	-4.67	436
386	5637	-112102	-3.71	-5.29	-4.09	-5.78	-4.64	-4.71	455
396	5667	-112213	-4.18	-4.96	-4.28	-5.58	-4.45	-4.69	429

<sup>a</sup> Note that the *molpdf* scores are incomparable between runs.

<sup>b</sup> Average z-score per chain weighted after number of residues in each chain ( $\alpha_1$ , 213;  $\beta_2$  and  $\gamma_2$ , 211).

<sup>c</sup> After being subjected to the Protein Preparation geometry optimization.<sup>6</sup>

<sup>d</sup> The OPLS 2001 energy of the input dimer was 211 kcal/mol.

<sup>e</sup> The OPLS 2001 energy of the input trimer was 307 kcal/mol.

<sup>f</sup> The high *molpdf* for model 97 is due to the  $\alpha_1$  H109 imidazole ring having a severely distorted configuration with crossing covalent bonds. This was fixed prior to running pprep.

<sup>g</sup> The OPLS 2001 energy of the input trimer was 291 kcal/mol.

<sup>h</sup> OPLS 2001 energy before manually adjusting the three peptide bonds causing Ramachandran plot violations (see text). After adjustments the energy was 401 kcal/mol as reported in the main article. (ProSA z-scores were identical before and after adjustments)

Compound <sup>a</sup>	R <sub>1</sub>	<b>R</b> <sub>2</sub>	$K_i (\mu M)$	Ref <sup>b</sup>	Gscore
	R <sub>1</sub>	OH N			
	HN R <sub>2</sub>	Õ			
1 (4-PIOL)	Н	Н	9.1	29	-7.2
· · · ·	Н	methyl	37	29	-7.0
	methyl	Н	10	29	-7.2
	ethyl	Н	6.3	29	-7.1
	propyl	Н	6.6	29	-6.8
	butyl	Н	7.7	29	-7.5
	hexyl	Н	4.5	29	-6.5
	octyl	Н	1.8	29	-5.2
	cyclohexyl	Н	4.9	29	-7.1
2	phenyl	Н	0.22	30	-7.4
3	benzyl	Н	3.8	29	-7.2
	2-phenylethyl	Н	5.0	29	-6.9
	3-phenylpropyl	Н	1.1	29	-7.5
	diphenylmethyl	Н	0.96	29	-6.9
	2,2-diphenylethyl	Н	0.36	29	-6.8
4	3,3-diphenylpropyl	Н	0.068	29	-7.8
	4,4-diphenylbutyl	Н	0.70	29	-8.4
5	3-biphenyl	Н	0.010	31	-8.3
	4-biphenylmethyl	Н	0.4	29	-7.9
	1-naphthyl	Н	0.82	30	-7.3
	2-naphthyl	Н	0.036	30	-7.0
	1-naphthylmethyl	Н	0.10	29	-7.6
	1-naphthylethyl	Н	1.7	29	-7.6
6	2-naphthylmethyl	Н	0.049	29	-7.9
7	1-phenyl-2-naphthylmethyl	Н	0.021	30	-8.5
	1-fluoro-2-naphthylmethyl	Н	0.019	30	-7.9
	1-chloro-2-naphthylmethyl	Н	0.016	30	-7.8
	1-cyano-2-naphthylmethyl	Н	0.028	30	-8.0
	1-methylthio-2-naphthylmethyl	Н	0.028	30	-8.0
	1-phenylthio-2-naphthylmethyl	Н	0.250	30	-8.0
8	1-bromo-2-naphthylmethyl	Н	0.011	30	-7.5
9	5-bromo-2-naphthylmethyl	Н	0.080	30	-7.8

Table S.I.-3. Structure, binding affinity and docking score (Glide Gscore<sup>28</sup>) for the highest ranked docking pose of all 52 4-PIOL and 4-PHP derived ligands docked to the model binding site.

<b>Compound</b> <sup>a</sup>	<b>R</b> <sub>1</sub>	<b>R</b> <sub>2</sub>	$K_i(\mu M)$	Ref <sup>b</sup>	Gscore
10	7-bromo-2-naphthylmethyl	Н	0.109	30	-7.8
11	8-bromo-2-naphthylmethyl	Н	0.045	30	-7.7
	2-naphthylethyl	Н	0.49	29	-7.9
	9-anthracylmethyl	Н	5.9	29	-7.4
	R	OH			
	$\bigwedge$				
	HŃ	$R_2$			
<b>12</b> (4-PHP)	Н	Н	10	32	-7.4
	phenyl	Н	0.022	32	-7.1
13	3-biphenyl	Н	0.0028	32	-7.8
14	2-naphthylmethyl	Н	0.033	32	-7.7
15	1-bromo-2-naphthylmethyl	Н	0.0095	32	-7.3
	Н	methyl	5.0	32	-7.5
	Н	phenyl	0.27	32	-7.1
	Н	2-tolyl	0.67	32	-7.2
	Н	3-tolyl	0.24	32	-7.6
	Н	4-tolyl	0.32	32	-7.1
	Н	benzyl	0.36	32	-7.8
16	Н	3-biphenyl	0.030	32	-8.5
17	Н	4-biphenyl	0.42	32	-7.0
18	Н	2-naphthylmethyl	0.0030	32	-8.9
	2-naphthylmethyl	phenyl	1.5	32	-7.8
19	phenyl	2-naphthylmethyl	0.022	32	-9.5

 $^{a}$  Compound number for those mentioned in the main article.  $^{b}$  Literature reference for the given  $K_{i}$  value.

Fig. S.I.-1. Alignments of the initial models 1a-b and 2a-c to each chain sequence in the refined model, highlighting (in blue) which templates were used in which regions in step 3 of the homology modeling protocol. Model 1a and 1b sequences correspond to the relevant chain of the refined model. Models 2a and 2b are only  $\alpha_1$  sequences; model 2c is only  $\gamma_2$ .

model ra	9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL	KFKGP	79
model-1b	9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL	KFKGP	79
model-2a	9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL	KFKGP	79
model-2b	9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL	KFKGP	79
refined- $\alpha$ 1	d 9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL	KFKGP	79
model-1a	80 MTVLRLNNLMASKIWTPDTFFHNGKKSVAHNMTMPNKLLRITEDGTLLYTMRLTVRAECPMHLEDF	PMDAH	150
model-1b	80 MTVLRLNNLMASKIWTPDTFFHNGKKSVAHNMTMPNKLLRITEDGTLLYTMRLTVRAECPMHLEDF	PMDAH	150
model-2a	80 MTVLRLNNLMASKIWTPDTFFHNGKKSVAHNMTMPNKLLRITEDGTLLYTMRLTVRAECPMHLEDF	PMDAH	150
model-2b	80 MTVLRLNNLMASKIWTPDTFFHNG <mark>KKSVAHNMTMPNKLL</mark> RITEDGTLLYTMRLTVRAECPMHLEDF	PMDAH	150
refined- $\alpha$ 1	a 80 MTVLRLNNLMASKIWTPDTFFHNGKKSVAHNMTMPNKLLRITEDGTLLYTMRLTVRAECPMHLEDF	PMDAH	150
model-1a	a 151   ACPLKFGSYAYTRAEVVYEWTREPARSVVVAEDGSRLNQYDLLGQTVDSGIVQSSTGEYVVMTTHF	HLKRK	221
model-1b	511 ACPLKFGSYAYTRAEVVYEWTREPARSVVVAEDGSRLNQYDLLGQTVDSGIVQSSTGEYVVMTTHF	HLKRK	221
model-2a	3 151 ACPLKFGSYAYTRAEVVYEWTREPARSVVVAEDGSRLNQYDLLGQTVDSGIVQSSTGEYVVMTTHF	HLKRK	221
model-2b	511 ACPLKFGSYAYTRAEVVYEWTREPARSVVVAEDGSRLNQYDLLGQTVDSGIVQSSTGEYVVMTTHF	HLKRK	221
refined- $\alpha$ 1	d 151 ACPLKFGSYAYTRAEVVYEWTREPARSVVVAEDGSRLNQYDLLGQTVDSGIVQSSTGEYVVMTTHF	HLKRK	221
model-1a	7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL	SYNVI	77
model-1a model-1b	7   SNMSLVKETVDRLLKGYD I RLRPDFGGPPVAVGMN I D I AS I DMVSEVNMDYTLTMYFQQAWRDKRL 7   SNMSLVKETVDRLLKGYD I RLRPDFGGPPVAVGMN I D I AS I DMVSEVNMDYTLTMYFQQAWRDKRL	SYNV I SYNV I	77 77
model-1a model-1b model-2a	7   SNMSLVKETVDRLLKGYD I RLRPDFGGPPVAVGMN I D I AS I DMVSEVNMDYTLTMYFQQAWRDKRL 7   SNMSLVKETVDRLLKGYD I RLRPDFGGPPVAVGMN I D I AS I DMVSEVNMDYTLTMYFQQAWRDKRL 9   DNTTVFTR I LDRLLDGYDNRLRPGLGERVTEVKTD I FVTSFGPVSDHDMEYT I DVFFRQSWKDERL	SYNVI SYNVI KFKGP	77 77 79
model-1a model-1b model-2a model-2b	<ul> <li>7 SNMSLVKETVDRLLKGYD I RLRPDFGGPPVAVGMN I D I AS I DMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYD I RLRPDFGGPPVAVGMN I D I AS I DMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>9 DNTTVFTR I L DRLLDGYDNRLRPGLGERVTEVKTD I FVTSFGPVSDHDMEYT I DVFFRQSWKDERL</li> <li>9 DNTTVFTR I L DRLLDGYDNRLRPGLGERVTEVKTD I FVTSFGPVSDHDMEYT I DVFFRQSWKDERL</li> </ul>	SYNVI SYNVI KFKGP KFKGP	77 77 79 79
model-1a model-1b model-2a model-2b refined-β2	<ul> <li>7 SNMSLVKETVDRLLKGYD I RLRPDFGGPPVAVGMN I D I AS I DMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYD I RLRPDFGGPPVAVGMN I D I AS I DMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>9 DNTTVFTR I LDRLLDGYDNRLRPGLGERVTEVKTD I FVTSFGPVSDHDMEYT I DVFFRQSWKDERL</li> <li>9 DNTTVFTR I LDRLLDGYDNRLRPGLGERVTEVKTD I FVTSFGPVSDHDMEYT I DVFFRQSWKDERL</li> <li>7 SNMSLVKETVDRLLKGYD I RLRPDFGGPPVAVGMN I D I AS I DMVSEVNMDYTLTMYFQQAWRDKRL</li> </ul>	SYNVI SYNVI KFKGP KFKGP SYNVI	77 77 79 79 77
model-1a model-1b model-2a model-2b refined-β2	<ul> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> </ul>	SYNVI SYNVI KFKGP KFKGP SYNVI	77 77 79 79 77
model-1a model-1b model-2a model-2b refined-β2 model-1a	<ul> <li>7 SNMSLVKETVDRLLKGYD I RLRPDFGGPPVAVGMN I D I AS I DMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYD I RLRPDFGGPPVAVGMN I D I AS I DMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>9 DNTTVFTR I L DRLLDGYDNRLRPGLGERVTEVKTD I FVTSFGPVSDHDMEYT I DVFFRQSWKDERL</li> <li>9 DNTTVFTR I L DRLLDGYDNRLRPGLGERVTEVKTD I FVTSFGPVSDHDMEYT I DVFFRQSWKDERL</li> <li>7 SNMSLVKETVDRLLKGYD I RLRPDFGGPPVAVGMN I D I AS I DMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>78 PLNLTLDNRVADQLWVPDTYFLNDKKSFVHGVTVKNRM I RLHPDGTVLYGLR I TTTAACMMDLRRY</li> </ul>	SYNVI SYNVI KFKGP KFKGP SYNVI	77 77 79 79 77
model-1a model-1b model-2a model-2b refined-β2 model-1a model-1b	<ul> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>78 PLNLTLDNRVADQLWVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>78 PLNLTLDNRVADQLWVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> </ul>	SYNVI SYNVI KFKGP KFKGP SYNVI PL DEQ PL DEQ	77 77 79 77 148 148
model-1a model-1b model-2a model-2b refined-β2 model-1a model-1b model-2a	<ul> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>9 DNTTVFTRILDRLLGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>78 PLNLTLDNRVADQLWVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>78 PLNLTLDNRVADQLWVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>80 MTVLRLNNLMASKIWTPDTFFHNGKKSVAHNMTMPNKLLRITEDGTLLYTMRLTVRAECPMHLEDF</li> </ul>	SYNVI SYNVI KFKGP KFKGP SYNVI PL DEQ PL DEQ PMDAH	77 77 79 77 148 148
model-1a model-1b model-2a model-2b refined-β2 model-1a model-1b model-2a model-2b	<ul> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>78 PLNLTLDNRVADQLWVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>78 PLNLTLDNRVADQLWVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>80 MTVLRLNNLMASKIWTPDTFFHNGKKSVAHNMTMPNKLLRITEDGTLLYTMRLTVRAECPMHLEDF</li> <li>80 MTVLRLNNLMASKIWTPDTFFHNGKKSVAHNMTMPNKLLRITEDGTLLYTMRLTVRAECPMHLEDF</li> </ul>	SYNVI SYNVI KFKGP KFKGP SYNVI PLDEQ PLDEQ PMDAH PMDAH	77 79 79 77 148 148 150 150
model-1a model-2b model-2b refined-β2 model-1a model-1b model-2b refined-β2	<ul> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>9 DNTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>78 PLNLTLDNRVADQLWVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>80 MTVLRLNNLMASKIWTPDTFFHNGKKSVAHNMTMPNKLLRITEDGTLLYTMRLTVRAECPMHLEDF</li> <li>27 8 PLNLTLDNRVADQLWVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> </ul>	SYNVI SYNVI KFKGP KFKGP SYNVI PLDEQ PMDAH PMDAH PLDEQ	77 79 79 77 148 150 150 148
model-1a model-2b model-2b refined-β2 model-1a model-1b model-2a refined-β2	<ul> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>9 DNTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>78 PLNLTLDNRVADQLWVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>80 MTVLRLNNLMASKIWTPDTFFHNGKKSVAHNMTMPNKLLRITEDGTLLYTMRLTVRAECPMHLEDF</li> <li>278 PLNLTLDNRVADQLWVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> </ul>	SYNVI SYNVI KFKGP SYNVI PLDEQ PLDEQ PMDAH PLDEQ	77 79 79 77 148 150 150 148
model-1a model-2b model-2b refined-β2 model-1a model-1b model-2a model-2b refined-β2 model-1a	<ul> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>9 DNTTVFTRILDRLLGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>78 PLNLTLDNRVADQLWVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>80 MTVLRLNNLMASKIWTPDTFFHNGKKSVAHNMTMPNKLLRITEDGTLLYTMRLTVRAECPMHLEDF</li> <li>80 MTVLRLNNLMASKIWTPDTFFHNGKKSVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>81 49 NCTLEIESYGYTTDDIEFYWRGDDNAVTGVTKIELPQFSIVDYKLITKKVVFSTGSYPRLSLSF</li> </ul>	SYNVI SYNVI KFKGP KFKGP SYNVI PLDEQ PLDEQ PMDAH PLDEQ KLKRN	77 77 79 79 77 148 150 150 148 217
model-1a model-2a model-2a refined-β2 model-1a model-1a model-2a refined-β2 model-1a model-1a	<ul> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>9 DNTTVFTRILDRLLGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>78 PLNLTLDNRVADQLWVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>80 MTVLRLNNLMASKIWTPDTFFHNGKKSVAHNMTMPNKLLRITEDGTLLYTMRLTVRAECPMHLEDF</li> <li>81 MTVLRLNNLMASKIWTPDTFFHNGKKSVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>82 MTVLRLNNLMASKIWTPDTFFHNGKKSVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>84 MTVLRLNNLMASKIWTPDTFFHNGKKSVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>85 MTVLRLNNLMASKIWTPDTFFHNGKKSVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>86 MTVLRLNNLMASKIWTPDTFFHNGKKSVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>98 MTVLRLNNLMASKIWTPDTFFHNGKKSVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>99 NCTLEIESYGYTTDDIEFYWRGDDNAVTGVTKIELPQFSIVDYKLITKKVVFSTGSYPRLSLSF</li> </ul>	SYNVI SYNVI KFKGP KFKGP SYNVI PLDEQ PMDAH PLDEQ KLKRN KLKRN	77 79 79 77 148 150 150 148 217 217
model-1a model-1b model-2a model-2b refined-β2 model-1b model-2a model-2b refined-β2 model-1a model-1b model-2a	<ul> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>9 DNTTVFTRILDRLLGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIFFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>80 MTVLRLNNLMASKIWTPDTFFHNGKKSVAHNMTMPNKLLRITEDGTLLYTMRLTVRAECPMHLEDF</li> <li>80 MTVLRLNNLMASKIWTPDTFFHNGKKSVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>81 NCTLEIESYGYTTDDIEFYWRGDDNAVTGVTKIELPQFSIVDYKLITKKVVFSTGSYPRLSLSF</li> <li>9149 NCTLEIESYGYTTDDIEFYWRGDDNAVTGVTKIELPQFSIVDYKLITKKVVFSTGSYPRLSLSF</li> <li>9151 ACPLKFGSYAYTRAEVVYEWTREPARSVVVAEDGSRLNQYDLLGQTVDSGIVQSSTGEYVVMTTHF</li> </ul>	SYNVI SYNVI KFKGP KFKGP SYNVI PLDEQ PMDAH PLDEQ KLKRN KLKRN KLKRN HLKRK	77 77 79 77 148 150 150 148 217 217 221
model-1a model-1b model-2a model-2b refined-β2 model-1a model-2a model-2a model-1a model-1a model-2a	<ul> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>78 PLNLTLDNRVADQLWVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>80 MTVLRLNNLMASKIWTPDTFFHNGKKSVAHNMTMPNKLLRITEDGTLLYTMRLTVRAECPMHLEDF</li> <li>80 MTVLRLNNLMASKIWTPDTFFHNGKKSVAHNMTMPNKLLRITEDGTLLYTMRLTVRAECPMHLEDF</li> <li>8149 NCTLEIESYGYTTDDIEFYWRGDDNAVTGVTKIELPQFSIVDYKLITKKVVFSTGSYPRLSLSF</li> <li>149 NCTLEIESYGYTTDDIEFYWRGDDNAVTGVTKIELPQFSIVDYKLITKKVVFSTGSYPRLSLSF</li> <li>151 ACPLKFGSYAYTRAEVVYEWTREPARSVVVAEDGSRLNQYDLLGQTVDSGIVQSSTGEYVVMTTHF</li> </ul>	SYNVI SYNVI KFKGP KFKGP SYNVI PLDEQ PLDEQ PMDAH PNDAH PLDEQ KLKRN KLKRN KLKRN HLKRK	77 79 79 77 148 150 150 148 217 217 221 221
model-1a model-2b model-2a model-2b refined-β2 model-1a model-2a model-2b refined-β2 model-2a model-2b refined-β2	<ul> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>9 DNTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTDIFVTSFGPVSDHDMEYTIDVFFRQSWKDERL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>7 SNMSLVKETVDRLLKGYDIRLRPDFGGPPVAVGMNIDIASIDMVSEVNMDYTLTMYFQQAWRDKRL</li> <li>78 PLNLTLDNRVADQLWVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLRRY</li> <li>80 MTVLRLNNLMASKIWTPDTFFHNGKKSVAHNMTMPNKLLRITEDGTLLYTMRLTVRAECPMHLEDF</li> <li>81 MVLRLNNLMASKIWTPDTFFHNGKKSVAHNMTMPNKLLRITEDGTLLYTMRLTVRAECPMHLEDF</li> <li>8149 NCTLEIESYGYTTDDIEFYWRGDDNAVTGVTKIELPQFSIVDYKLITKKVVFSTGSYPRLSLSF</li> <li>151 ACPLKFGSYAYTRAEVYEWTREPARSVVAEDGSRLNQYDLLGQTVDSGIVQSSTGEYVVMTTHF</li> <li>149 NCTLEIESYGYTTDDIEFYWRGDDNAVTGVTKIELPQFSIVDYKLITKKVVFSTGSYPRLSLSF</li> </ul>	SYNVI SYNVI KFKGP SYNVI PLDEQ PLDEQ PMDAH PMDAH PLDEQ KLKRN KLKRN HLKRK KLKRN	77 79 79 77 148 150 150 148 217 221 221 217 221

model-1a	22	22 VPEGDVTVILNNLLEGYDNKLRPDIGVKPTLIHTD	MYVNSIGPVNAINMEYTIDIFFAQTWYDRRLKFNST	92
model-1b	22	22 VPEGDVTVILNNLLEGYDNKLRPDIGVKPTLIHTD	MYVNSIGPVNAINMEYTIDIFFAQTWYDRRLKFNST	92
model-2a	9	9 DNTTVFTRILDRLLDGYDNRLRPGLGERVTEVKTD	IFVTSFGPVSDHDMEYTIDVFFRQSWKDERLKFKGP	79
model-2c	22	22 VPEGDVTVILNNLLEGYDNKLRPDIGVKPTLIHTD	MYVNSIGPVNAINMEYTIDIFFAQTWYDRRLKFNST	92
refined-γ2	22	22 VPEGDVTVILNNLLEGYDNKLRPDIGVKPTLIHTD	MYVNSIGPVNAINMEYTIDIFFAQTWYDRRLKFNST	92
model-1a	93	93   I K V L R L N S N M V G K I W I P D T F F R N S K K A D A H W I T T P	NRMLRIWNDGRVLYTLRLTIDAECQLQLHNFPMDEH	163
model-1b	93	93 IKVLRLNSNMVGKIWIPDTFFRNSKKADAHWITTP	NRML R I WNDG R V L Y T L R L T I DA E CQ L Q L H N F P M D E H	163
model-2a	80	80 MTVLRLNNLMASKIWTPDTFFHNGKKSVAHNMTMP	NKLLRITEDGTLLYTMRLTVRAECPMHLEDFPMDAH	150
model-2c	93	93 I K V L R L N S N M V G K I W I P D T F F R N S K K A D A H W I T T P	NRML RIWNDGRVLYTLRLTIDAECQLQLHNFPMDEH	163
refined-γ2	93	93 I KVL RL NSNMVGK I WI PDTFFRNSKKADAHWI TTP	NRMLRIWNDGRVLYTLRLTIDAECQLQLHNFPMDEH	163
model-1a	164	64   SCPLEFSSYGYPREEIVYQWKRSSVE - VGDTRSW -	RLYQFSFVGLRNTTEVVKTTSGDYVVMSVYFDLSRR	232
model-1b	164	64 SCPLEFSSYGYPREEIVYQWKRSSVE-VGDTRSW-	RL Y <mark>QFSFVGLRNTTEVVKTTSGDYVVMSVYFDLSRR</mark>	232
model-2a	151	51 ACPLKFGSYAYTRAEVVYEWTREPARSVVVAEDGS	RL NQYDL LGQT VDSG I VQSSTGEYVVMTTHFHL KRK	221
model-2c	164	64 SCPLEFSSYGYPREEIVYQWKRSSVE-VGDTRSW-	RLYQFSFVGLRNTTEVVKTTSGDYVVMSVYFDLSRR	232
refined-γ2	164	64 SCPLEFSSYGYPREEIVYQWKRSSVE - VGDTRSW -	RLYQFSFVGLRNTTEVVKTTSGDYVVMSVYFDLSRR	232

Fig. S.I.-2. Comparison of the ClustalX alignments before and after manual alterations of the GABA<sub>A</sub>R sequences relative to the structural template alignment. Image is generated with JalView.<sup>33</sup> Columns are colored by conservation at 17% threshold. A) Original (left) and altered (right) alignment in the N-terminal  $\alpha$ -helix region. B) Original (left) and altered (right) alignment in the L5-L5' region. C) Original (top) and altered (bottom) alignment in loops F and C.

# A)

Ac-AChBP	1 HSQANLMRLKSDLF NRS PMYPGP - TK DDPL	29	1 HSQANLMRLKSDLF NRS PMYPGP - TK DDPL	29
Ls-AChBP	1 DRADILYNIR QTS R <mark>P</mark> DVI <mark>P</mark> - TQR - DR <mark>P</mark> V	26	1 DRADILYNIR QTS RPDVI <mark>P</mark> - TQR - DR <mark>PV</mark>	26
Bt-AChBP	1 Q I RWTLL NQ I TGESDV I P - L SN - NT PL	25	1 Q I RWTLL NQ I TGESDV I P - L SN - NT PL	25
ELIC	1	2	1	2
GLIC	1 V <mark>SPP</mark> PIA - DE <mark>P</mark> L	12	1 VSP <mark>P</mark> PPIA - DE <mark>PL</mark>	12
M. ACh α1	1 SEHETRLEAKLF EDY SSVVRP - VEDHREIV	29	1 SEHETRLEAKLF EDY SSVVRP - VEDHREIV	29
T. ACh α	1 SEHETRL VANLL ENY NKV I RP - VEHHTHFV	29	1 SEHETRL VANL L ENY NKV I RP - VEHHTHF V	29
T. ACh β	1 SVMEDTLLSVLF ENY N <mark>P</mark> KV <mark>RP</mark> - SQTV <mark>G</mark> DKV	29	1 SVMEDTLLSVLF ENY NPKVRP - SQTVGDKV	29
T. ACh γ	1 NEE <mark>G</mark> RLIEKLL <mark>G</mark> DY DKRI <mark>KP</mark> - AKTLDHVI	28	1 NEEGRLIEKLL GDY DKRIKP - AKTLDHVI	28
T. ACh δ	1 VNEEERLINDLLIVNKY NKHVRP - VKHNNEVV	31	1 VNEEERLINDLLIVNKY NKHVRP - VKHNNEVV	31
GABA α1	3 LQDELKDNTTVFTRILDRLLD <mark>G</mark> YDN <mark>R</mark> LR <mark>PG</mark> LGERVT	38	9 DNTTVFTRILDRLL DGY DNRLRP - GLG - ERVT	38
GABA β2	1 - SVNDPSNMSLVKETVDRLLKGYDIRLRPDFGGPPV	35	6 SNMSLVKETVDRLL KGY DIRLRP - DFG - GPPV	35
GABA γ2	16 WVLTPKVPEGDVTVILNNLLEGYDNKLRPDIGVKPT	51	22 VPEGDVTVILNNLL EGY DNKLRP - DIG - VKPT	51

### B)

Ac-AChBP86WTPDITAYS-STRPVQVLSPQIAVVTHDG11386WTPDITAYSSTRPVQVLSPQIAVVTHDG113Ls-AChBP81WVPDLAAYN-AISKPEVLTPQLARVVSDG10881WVPDLAAYNAISKPEVLTPQLARVVSDG108Bt-AChBP81WTPDLSFYN-AIAAPELLSADRVVVSDG10881WTPDLSFYNAIAAPELLSADRVVVSKDG108Bt-AChBP81WTPDLSFYN-AIAAPELLSADRVVVSDG10881WTPDLSFYNAIAAPELLSADRVVVSKDG108BLC62WVPALEFIN-VVGSPDT-ONKKIMLFPDG8862WVPALEFINVVGSPDT-GNKRUMLFPDG88GLIC68WIPEIRFVN-VENARDA-DVVDISVSPDG9468WIPEIRFVNVENARDA-DVVDISVSPDG94M.AChα186WRPDVVLYN-NADGDFAIV-KFTKVLLDYTG11486WPDUVLYNNADGDFAIV-KTLLDYTG114T.AChα86WPDUVLYN-NADGDFAIV-KFTKVLLDYTG11486WPDUVLYNNADGDFAIV-HMTKLLDYTG114T.AChβ86WIPDIVLMN-NNDGSFEIT-LHVNVVQHTG11385WLPDVVLEN-NVDGGFEVA-YYANVLVYNDG113T.AChβ88WIPDIVLQN-NNDGYNVA-YFCNVLVRPNG11688WIPDIVLQNNNDGYNVA-YFCNVLVRPNG116GABAα194WTPDTFFHNGKKSVAHNMTMPNKLLRITEDG12494WTPDTFFHNGKKSVAHNMTMPNKLLRITEDG121GABAβ291WVPDTFFRNSKKADAHWITTPNRMLRIWNDG137107WIPDTFFRNSKKADAHWITTPNRMLRIWNDG137												
Ls-AChBP81WV PDLAAYN - A I S - · K PEVLTPQLARVVS DG10881WV PDLAAYN A I S - · K PEVLTPQ - · · L ARVVS DG108BFAChBP81WT PDLSFYN - A I A - · A PELLSADRVVVSK DG10881WT PDLSFYN A I A - · A PELLSAD - · · RVVSK DG108ELIC62WV PALEFIN · VVG - · SPDT · GNKRLMLFPDG8862WV PALEFIN · VVG - · · SPDT · GNKRLMLFPDG88GLIC68WI PE I RFVN · VEN - · ARDA · DVVD I SVSPDG9468WI PE I RFVN VEN - · ARDA · DVV - · D I SVSPDG94M. ACh α186WP PD VLYN · NADGDFAIV · KFTKVLLDYTG11486WR PD VVLYN NADGDFAIV · KTT · · · K VLLDYTG114T. ACh α86WI PD I VLM · NADGDFAIV · HMTKLLLDYTG11486WP PD VVLYN NADGDFAIV · HMT · · · KLLDYTG114T. ACh β86WI PD I VLM · NADGGFEVA · YYANVLVYNDG11385WL PD VVLEN · VVL · · · NVLVHDG113T. ACh β88WI PD I VLM · NNDGSFEIT · LHVVVLVHDG11388WI PD I VLM NNDGGFEVA · YYA · · · NVLVYNDG113T. ACh β88WI PD I VLM · NNDGQYNVA · YFC · · · VLVRPNG11688WI PD I VLM NNDGQYNVA · YFC · · · · VLVRPNG116GABA α194WT PD TFFHNGKKSVAHNMTMPNKLLRITEDG12494WT PD TFFHNGKK · · SVAHNMTMPNKLLRITEDG121GABA β291WV PD TFFRNSKKADAHWITTPNRMLRIWNDG137107WI PD TFFRNSKK · · ADAHWITTPNRMLRIWNDG137	Ac-AChBP	86	WT P D I T A	AYS	<u>;</u> - STR <mark>P</mark> VQV	LSPQIAVVTHDG	113	86 W	T P D I T	AYS	STR PVQVLSPQ IAVVTHDG	113
BEAChBP81WTPDLSFYNAIAAPELLSADRVVVSKDG10881WTPDLSFYNAIAAPELLSADRVVJSKDG108ELIC62WVPALEFIN-VCGSPDT-GNKRLMLFPDG8862WVPALEFINVVGSPDT-GNKRLMLFPDG88GLIC68WIPEIRFVN-VENARDA-DVVDISVSPDG9468WIPEIRFVN-VENARDA-DVVDISVSPDG94M.ACha186WRPDVVLYN-NADGDFAIV-KFTKVLLDYTG11486WRPDVVLYNNADGDFAIV-KFTKVLLDYTG114T.ACha86WLPDLVLYN-NADGDFAIV-HMTKLLLDYTG11486WQPDIVLMNNDGSFEIT-LHVNVLVQHTG114T.ACha86WLPDVVLEN-NNDGSFEIT-LHVNVLVQHTG11486WQPDIVLMNNDGSFEIT-LHVNVLVQHTG114T.ACha86WLPDVVLEN-NNDGSFEIT-LHVNVLVQHTG11385WLPDVVLEN-VNDGGFEVA-YYANVLVYNDG113T.ACha88WIPDIVLQN-NNDGQYNVA-YFCNVLVRPNG11688WIPDIVLQNNNDGQYNVA-YFCNVLVRPNG116GABA $\alpha$ 194WTPDTFFHNGKKSVAHNMTMPNKLLRITEDG12494WTPDTFFHNGKKSVAHNMTMPNKLLRITEG121GABA $\beta$ 2107WIPDTFFRNSKKADAHWITTPNRMLRIWNDG137107WIPDTFFRNSKK-ADAHWITTPNRMLRIWNDG137	Ls-AChBP	81	WVPDLAA	A Y N	- A I S KPEV	L T PQL A R V V S DG	108	81 W	VV P D L A	A Y <mark>N</mark>	A I S K <mark>P</mark> EVLT <mark>P</mark> Q LARVVSDG	108
ELIC62WV PALEFIN-VVGSPDT-GNKRLMLFPDG8862WV PALEFINVVGSPDT-GNKRLMLFPDG88GLC68WIPEIRFVN-VENARDA-DVVDISVSPDG9468WIPEIRFVNVENARDA-DVVDISVSPDG94M.ACh $\alpha$ 186WR PDVLYN-NADGDFAIV-KFTKVLLDYTG11486WR PDVLYNNADGDFAIV-KFTKVLLDYTG114T.ACh $\alpha$ 86WIPDIVLWN-NADGDFAIV-HMTKLLLDYTG11486WR PDVLYNNADGDFAIV-HMTKULLDYTG114T.ACh $\beta$ 88WOPDIVLWN-NDGSFEIT-LHVNVLVQHTG11486WQ PDIVLWNNDGGFEIV-HMTKULLDYTG114T.ACh $\beta$ 85WIPDIVLQN-NNDGGFEVA-YYANVLVYNDG11385WIPDIVLQN-NNDGGYVA-YFCNVLVRNG113T.ACh $\beta$ 88WIPDIVLQN-NNDGQYNVA-YFCNVLVRNG11688WIPDIVLQN-NNDGQYNA-YFCNVLVRNG116GABA $\alpha$ 194WTPDTFFHNGKKSVAHNMTMPNKLLRITEDG12194WTPDTFFHNGKKSVAHNMTMPNKLLRITEG121GABA $\beta$ 291WVPDTYFLNDKKSFVHGVTVKNRMIRLHPDG137107WIPDTFFRNSKKADAHWITTPNRMLRIWNDG137	Bt-AChBP	81	WTPDLSF	Y N	- A I A APEL	LSADRVVVSKDG	108	81 W	VT PDL SI	FYN	AIAAPELLSADRVVVSKDG	108
GLIC         68         WIPEIRFVN-VENARDA-DVVDISVSPDG         94         68         WIPEIRFVNVENARDA-DVVDISVSPDG         94           M. ACh α1         80         WRPDVVLYN-NADGDFAIV-KFTKVLLDYTG         114         86         WRPDVVLYNNADGDFAIV-KFT         114         86         WRPDVVLYNNADGDFAIV-KT         114	ELIC	62	WV <mark>P</mark> ALEF	= I N	- VV <mark>G</mark> SPDT	- GNKRLML F P DG	88	62 W	VV <mark>P</mark> ALEI	FIN	VV <mark>G</mark> SPDT - <mark>G</mark> NK RLML FPDG	88
M. ACh α186WRPDVVLYN-NADGDFAIV-KFTKVLLDYTG11486WRPDVVLYNNADGDFAIV-KFTKVLLDYTG114T. ACh α86WLPDLVLYN-NADGDFAIV-HMTKLLDYTG11486WLPDLVLYNNADGDFAIV-HMTKLLDYTG114T. ACh β86WQPDIVLMN-NNDGSFEIT-LHVVVLVQHTG11486WQPDIVLMNNDGSFEIT-LHVNVLVQHTG114T. ACh β85WLPDVVLEN-NVDGQFEVA-YYANVLVYNDG11385WLPDVVLEN-NVDGQFEVA-YYANVLVYNDG113T. ACh δ88WIPDIVLQN-NNDGQYNVA-YFCNVLVRPNG11688WIPDIVLQNNNDGQYNVA-YFCNVLVRPNG116GABA α194WTPDTFFHNGKKS-SVAHNMTMPNKLLRITEDG12412494WTPDTFFHNGKKSVAHNMTMPNKLLRITEG121GABA β291WVPDTYFLNDKKSFVHGVTVKNRMIRLHPDG137107WIPDTFFRNSKK-ADAHWITTPNRMLRIWNDG137	GLIC	68	WI <mark>P</mark> EIRF	= V N	- VEN ARDA	- DVVDISVSPDG	94	68 <mark>W</mark>	I <mark>P</mark> E I RI	F V <mark>N</mark>	VEN ARDA - DVV DISVSPDG	94
T. ACh $\alpha$ 86WL PDL VL YN - NADG DFAIV - HMTKLLLDYTG11486WL PDL VL YN NADG DFAIV - HMT KLLLDYTG114T. ACh $\beta$ 86WL PD VL WN - NNDG SFEIT - LHVNVL VQHTG11486WQ PDI VL MNNDC SFEIT - LHV NVL VQHTG114T. ACh $\beta$ 85WL PD VVL EN NVDG GFEVA - YYANVL VYNDG11385WL PD VVL EN NVDG GFEVA - YYANVL YNDG113T. ACh $\delta$ 88WI PD I VL QN - NNDG QYNVA - YFC NVL VR PNG11688WI PD I VL QN NNDG QYNVA - YFC NVL VR PNG116GABA $\alpha$ 194WT PD TFFHNGKKS VAHNMT MPNKLL RI TEDG12494WT PD TFFHNGKK SVAHNMT MPNKL RI TEDG121GABA $\beta$ 291WV PD TYFL NDKKSFVHGVT VKNRMI RL HPDG12191WV PD TYFL NDKK SFVHGVT VKNRMI RL HPDG127GABA $\gamma$ 2107WI PD TFFFNSKKA DA HWI TT PNRM RI WNDG137107WI PD TFFRNSKK - ADA HWI TT PNRM RI WNDG137	M. ACh α1	86	WR <mark>P</mark> DVVI	- Y N	- NADGDFAIV	- KF T K V L L D Y T G	114	86 W	VR <mark>P</mark> DVVI	LY <mark>N</mark>	NADGDFAIV-KFTKVLLDYT <mark>G</mark>	114
T. ACh β86 WQ PDIVLMN - NNDGSFEIT - LHVNVLVQHTG11486 WQ PDIVLMNNNDGSFEIT - LHV NVLVQHTG114T. ACh γ85 WL PDVVLEN - NVDGQFEVA - YYANVLVYNDG11385 WL PDVVLEN NVDGGFEVA - YYA NVLVYNDG113T. ACh δ88 WI PDIVLQN - NNDGQYNVA - YFCNVLVRPNG11688 WI PDIVLQNNNDGQYNVA - YFC NVLVRPNG116GABA α194 WT PDTFFHNGKKSVAHNMTMPNKLLRITEDG12494 WT PDTFFHNGKK - SVAHNMTMPNKLLRITEDG124GABA β291 WV PDTYFLNDKKSFVHGVTVKNRMIRLHPDG12191 WV PDTFFHNGKK - SFVHGVTVKNRMIRLHPDG121GABA β2107 WI PDTFFRNSKKADAHWITTPNRMLRIWNDG137107 WI PDTFFRNSKK - ADAHWITTPNRMLRIWNDG137	T. ACh $\alpha$	86	WL <mark>P</mark> DL VI	- Y N	- NADGDFAIV	- HMTKLLLDYT <mark>G</mark>	114	86 W	VL <mark>P</mark> DL VI	LYN	NADGDFAIV-HMTKLLLDYT <mark>G</mark>	114
Τ. ΑCh γ         85 WL PDVVLEN - NVDGQFEVA - YYANVLVYNDG         113         85 WL PDVVLEN NVDGQFEVA - YYA NVLVYNDG         113           Τ. ACh δ         88 WI PDI VLQN - NNDGQYNVA - YFCNVLVRPNG         116         88 WI PDI VLQNNNDGQYNVA - YFCNVLVRPNG         116           GABA α1         94 WTPDTFFHNGKKSVAHNMTMPNKLLRITEDG         124         94 WTPDTFFHNGKK SVAHNMTMPNKLLRITEDG         124           GABA β2         91 WVPDTYFLNDKKSFVHGVTVKNRMIRLHPDG         121         91 WVPDTYFLNDKK SVHOVTVKNRMIRLHPDG         121           GABA β2         91 WVPDTFFFNSKKADAHWITTPNRMLRIWNDG         137         107 WIPDTFFRNSKK - ADAHWITTPNRMLRIWNDG         137	T. ACh β	86	WQ <mark>P</mark> DIVL	MN	- NNDGSFEIT	- L H V N V L V Q H T <mark>G</mark>	114	86 W	VQ P D I V I	LMN	NNDGSFEIT-LHVNVLVQHT <mark>G</mark>	114
Τ. ACh δ         88 WIPDIVLQN-NNDGQYNVA-YFCNVLVRPNG         116         88 WIPDIVLQNNNDGQYNVA-YFCNVLVRPNG         116           GABA α1         94 WTPDTFFHNGKKS-SVAHNMTMPNKLLRITEDG         124         94 WTPDTFFHNGKKSVAHNMTMPNKLLRITEDG         124           GABA β2         91 WVPDTYFLNDKKSFVHGVTVKNRMIRLHPDG         121         91 WVPDTYFLNDKK-SFVHGVTVKNRMIRLHPDG         121           GABA γ2         107 WIPDTFFFNGKK-SFVHGVTVKNRMIRLHPDG         121         107 WIPDTFFRNGKKSFVHGVTVKNRMIRLHPDG         121	T. ACh γ	85	WL <mark>P</mark> D V V I	EN	- NVDGQFEVA	- YYANVL VYNDG	113	85 W	VL <mark>P</mark> DVVI	LEN	NVDGQFEVA - YYA NVLVYNDG	113
GABA α194WTPDTFFHNGKKSVAHNMTMPNKLLRITEDG12494WTPDTFFHNGKKSVAHNMTMPNKLLRITEDG124GABA β291WVPDTYFLNDKKSFVHGVTVKNRMIRLHPDG12191WVPDTYFLNDKKSFVHGVTVKNRMIRLHPDG121GABA γ2107WIPDTFFRNSKKADAHWITTPNRMLRIWNDG137107WIPDTFFRNSKKADAHWITTPNRMLRIWNDG137	T. ACh δ	88	WIPDIVL	QN	- NNDGQYNVA	- YFCNVLVRPNG	116	88 <mark>W</mark>	VI <mark>P</mark> DIVI	LQ <mark>N</mark>	NNDGQYNVA - YFC NVL VRPNG	116
GABAβ2 91 WVPDTYFLNDKKSFVHGVTVKNRMIRLHPDG 121 91 WVPDTYFLNDKKSFVHGVTVKNRMIRLHPDG 121 GABAγ2 107 WIPDTFFRNSKKADAHWITTPNRMLRIWNDG 137 107 WIPDTFFRNSKKADAHWITTPNRMLRIWNDG 137	GABA α1	94	WT <mark>P</mark> D T F F	= H N	GKKSVAHNMT	MPNKLLRITEDG	124	94 W	VT <mark>P</mark> D T F I	FHN	GKK SVAHNMTMPNKLLRITEDG	124
GABA γ2 107 WIPDTFFRNSKKADAHWITTPNRMLRIWNDG 137 107 WIPDTFFRNSKKADAHWITTPNRMLRIWNDG 137	GABA β2	91	WVPDTYF	ELN	DKKSFVHGVT	VKNRMIRLHPDG	121	91 W	V <mark>P D T</mark> Y I	FLN	DKK SFVHGVTVKNRMIRL HPDG	121
	GABA γ2	107	WIPDTFF	RN	SKKADAHWIT	T P N R M L R I W N D G	137 1	07 W	I <mark>P</mark> DTFI	FR <mark>N</mark>	SKK ADAHWITTPNRMLRIWNDG	137

# C)

Ac-AChBP	154 IDLK-TDTDQ-V-DLSSYYASSKYEILSATQTRQVQHYSCCPEPY	IDV 198
Ls-AChBP	149 ISVD-PTTEN-SDDSEYFSQYSRFEILDVTQKKNSVTYSCCPEAY	EDV 194
Bt-AChBP	149 FALI-TGEE-GV-VNIAEYFDSPKFDLLSATQSLNRKKYSCCENMY	DDI 194
ELIC	128 QQL RF SD I - QVY - TEN I DNEE I DEWWI RKASTHISDIRYDHLSSVQPNQNEF	SRI 180
GLIC	136 IVLA-VDLEK-V-GK-NDDVFLTGWDIESFTAVVKPANFALEDRLE	SKL 180
M. ACh α1	156 VAIN-PESDQ-P-DLSNFMESGEWVIKEARGWKHWVFYSCCPTTPY	DI 201
T. ACh α	156 VSIS-PESDR-P-DLSTFMESGEWVMKDYRGWKHWVYYTCCPDTPY	DI 201
T. ACh β	156 VILQ-HALDA-M-I-NQDAFTENGQWSIEHKPSRKNWRSDDPSY	EDV 198
T. ACh γ	155 VNLQ-LSAEEGI-DPEDFTENGEWTIRHRPAKKNYNWQLTKDDIDF	QEI 201
T. ACh δ	158 ISMD-LI-I-DPEAFTENGEWEIIHKPAKKNIYGDKFPNGTNY	200 V D C
GABA α1	166 VVYEWTREPAR SVVVAEDGS RL NQYDLLGQTVDSG I VQSSTGE Y	/V <mark>M</mark> 212
GABA β2	163 IEFYWRGDDNAVTGVTKIELPQFSIVDYKLITKKVVFSTGSY	RL 207
GABA γ2	179 IVYQWKRSSVEVGDTRSWRLYQFSFVGLRNTTEVVKTTSGDY	/V <mark>M</mark> 223
		109
Ac-AChBP	154 IDLK-TDTD-Q-VDLSSYYASSKYEILSATQTRQVQHYSCCPEPY	I DV 198
Ac-AChBP Ls-AChBP	154 IDLK-TDTD-Q-VDLSSYYASSKYEILSATQTRQVQHYSCCPEPY 149 ISVD-PTTE-N-S-DDSEYFSQYSRFEILDVTQKKNSVTYSCCPEAY	I DV 198 E DV 194
Ac-AChBP Ls-AChBP Bt-AChBP	154 IDLK-TDTD-Q-VDLSSYYASSKYEILSATQTRQVQHYSCCPEPY 149 ISVD-PTTE-N-S-DDSEYFSQYSRFEILDVTQKKNSVTYSCCPEAY 149 FALI-TGEEGV-V-NIAEYFDSPKFDLLSATQSLNRKKYSCCENMY 128 QQLRESD	I DV 198 E DV 194 D D I 194 S R I 180
Ac-AChBP Ls-AChBP Bt-AChBP ELIC	154 IDLK-TDTD-Q-VDLSSYYASSKYEILSATQTRQVQHYSCCPEPY 149 ISVD-PTTE-N-S-DDSEYFSQYSRFEILDVTQKKNSVTYSCCPEAY 149 FALI-TGEEGV-V-NIAEYFDSPKFDLLSATQSLNRKKYSCCENMY 128 QQLRF-SDI-QVY-TENIDNEEIDEWWIRKASTHISDIRYDHLSSVQPNQNEF 136 LVLA-VDIE-K-VGK-NDDVEITQNDIESETAVVKPANEALED	I DV 198 E DV 194 D D I 194 S R I 180 S K I 180
Ac-AChBP Ls-AChBP Bt-AChBP ELIC GLIC	154       IDLK-TDTD-Q-V-DLSSYYASSKYEILSATQTRQVQHYSCCPEPY         149       ISVD-PTTE-N-S-DDSEYFSQYSRFEILDVTQKKNSVTYSCCPEAY         149       FALI-TGEEGV-V-NIAEYFDSPKFDLLSATQSLNRKKYSCCENMY         128       QQLRF-SDI-QYY-TENIDNEEIDEWWIRKASTHISDIRYDHLSSVQPNQNEF         136       VAL-VDLE-K-V-GK-NDDVFLTGWDIESFTAVVKPANFALEDRLE         156       VAL-VDES-K-V-GLSEWWIRKASTHISDIRYDHLSSCCP	I DV 198 E DV 194 D D I 194 S R I 180 S K L 180
Ac-AChBP Ls-AChBP Bt-AChBP ELIC GLIC Μ. ACh α1	<ul> <li>154 IDLK-TDTD-Q-V-DLSSYYASSKYEILSATQTRQVQHYSCCPEPY</li> <li>149 ISVD-PTTE-N-S-DDSEYFSQYSRFEILDVTQKKNSVTYSCCPEAY</li> <li>149 FALI-TGEEGV-V-NIAEYFDSPKFDLLSATQSLNRKKYSCCENMY</li> <li>128 QQLRF-SDI-QVY-TENIDNEEIDEWVIRKASTHISDIRYDHLSSVQPNQNEF</li> <li>136 IVLA-VDLE-K-V-GK-NDDVFLTGWDIESFTAVVKPANFALEDRLE</li> <li>156 VAIN-PESD-Q-P-DLSNFMESGEWVIKEARGWKHWVYSCCPTPY</li> <li>156 VAIN-PESD-Q-P-DLSTEMESGEWVIKEARGWKHWVYSCCPTPY</li> </ul>	I D V 198 E D V 194 D I 194 S R I 180 S K L 180 D I 201
Ac-AChBP Ls-AChBP Bt-AChBP ELIC GLIC M. ACh α1 T. ACh α T. ACh β	154       IDLK-TDTD-Q-V-DLSSYYASSKYEILSATQTRQVQHYSCCPEPY         149       ISVD-PTTE-N-S-DDSEYFSQYSRFEILDVTQKKNSVTYSCCPEAY         149       FALI-TGEEGV-V-NIAEYFDSPKFDLLSATQSLNRKKYSCCENMY         128       QQLRF-SDI-QYY-TENIDNEEIDEWWIRKASTHISDIRYDHLSSVQPNQNEF         136       IVLA-VDLE-K-V-GK-NDDVFLTGWDIESFTAVVKPANFALEDRLEE         156       VAIN-PESD-Q-P-DLSNFMESGEWVIKEARGWKHWVFYSCCPTTPY         156       VSIS-PESD-R-P-DLSTFMESGEWVIKENKESSKDPRGWINNYYTCCPDTPY         156       VSIS-PESD-R-P-DLSTFMESGEWVIKESRKWWSSICHWVYTCCPDTPY	I D V 198 E D V 194 D D I 194 S R I 180 S K L 180 D I 201 D I 201
Ac-AChBP Ls-AChBP Bt-AChBP ELIC GLIC M. ACh α1 T. ACh α T. ACh α T. ACh α	154       IDLK-TDTD-Q-V-DLSSYYASSKYEILSATQTRQVQHYSCCPEPY         149       ISVD-PTTE-N-S-DDSEYFSQYSRFEILDVTQKKNSVTYSCCPEAY         149       FALI-TGEEGV-V-NIAEYFDSPKFDLLSATQSLNRKKYSCCENMY         128       QQLRF-SDI-QYY-TENIDNEEIDEWWIRKASTHISDIRYDHLSSVQPNQNEF         136       IVLA-VDLE-K-V-GK-NDDVFLTGWDIESFTAVVKPANFALEDRLE         156       VAIN-PESD-Q-P-DLSNFMESGEWVIKARGWKHWVYYSCCPTTPY         156       VSIS-PESD-R-P-DLSTFMESGEWVIKDYRGWKHWVYYTCCPDTPY         156       VILQ-HALDAM-I-NQDAFTENGWSIEHKPSRKWNWOITKDDTPY	I D V 198 E D V 194 D I 194 S R I 180 S K L 180 D I 201 D I 201 E D V 198
Ac-AChBP Ls-AChBP Bt-AChBP ELIC GLIC M. ACh α1 T. ACh α T. ACh β T. ACh β	154       IDLK-TDTD-Q-V-DLSSYYASSKYEILSATQTRQVQHYSCCPEPY         149       ISVD-PTTE-N-S-DDSEYFSQYSRFEILDVTQKKNSVTYSCCPEAY         149       FALI-TGEEGV-V-NIAEYFDSPKFDLLSATQSLNRKKYSCCENMY         128       QQLRF-SDI-QVY-TENIDNEEIDEWWIRKASTHISDIRYDHLSSVQPNQNEF         136       VLA-VDLE-K-V-GK-NDDVFLTGWDIESFTAVVKPANFALEDRLE         156       VAIN-PESD-Q-P-DLSTFMESGEWVIKEARGWKHWVFYSCCPDTPY         156       VILQ-HALDAM-INQDAFTENGEWSIEHKPSRKNWRSDDPSY         158       SND-I	I D V 198 E D V 194 D I 194 S R I 180 S K L 180 D I 201 D I 201 E D V 198 Q E I 201
Ac-AChBP Ls-AChBP Bt-AChBP ELIC GLIC M. ACh α1 T. ACh α T. ACh α T. ACh γ T. ACh β	154       IDLK - TDTD - Q - V - DLSSYY ASSKYE ILSATQTRQVQHYSCCP EPY         149       ISVD - PTTE - N - S - DDSEYFS QYSRFE ILDVTQKKNSVTYSCCP EAY         149       FALI - TGEEGV - V - NIAEYFD SPKFDL LSATQSLNRKKYSCCE NMY         128       QQLRF - SD I - QYY - TENIDNEE IDEWWI RKASTHJSD IRYDHLSSVQPNQNEF         136       IVLA - VDLE - K - V - GK - NDD VFLTGWI ESFTAVVKPANFALED RLEE         156       VAIN - PESD - Q - P - DLSNFM ESGEWVI KEARGWKHWVFYSCCP TTPY         156       VSIS - PESD - R - P - DLSTFM ESGEWVI KEARGWKHWVFYSCCP TTPY         156       VSIS - PESD - R - P - DLSTFM ESGEWVI KEARGWKHWVFYSCCP TTPY         156       VSIS - PESD - R - P - DLSTFM ESGEWVI KEARGWKHWVFYSCCP TPY         156       VSIS - PESD - R - P - DLSTFM ESGEWVI KEARGWKHWVFYCCP DTPY         156       VSIS - PESD - R - P - DLSTFM ESGEWVI KEARGWKHWVFYCCP DTPY         156       VSIS - PESD - R - P - DLSTFM ESGEWVI KDYRGWKHWVYTCCP DTPY         157       VNLQ - HALDAM - I - NQDAFT ENGEWT I RHRPAKKNYNQLTKD DIDFO         158       VSILQ - LSAE - EGI DPEAFT ENGEWT I HKPAKKNYNQLTKD DIDFO         158       ISMD - L I - I - DPEAFT ENGEWE I I HKPAKKNYSONG I KDS GNY         156       VVAEDGS - ENN WYN I OVYSGE I YOSS	I DV 198 E DV 194 D I 194 S R I 180 S K L 180 D I 201 D I 201 E DV 198 Q E I 201 Q DV 200
Ac-AChBP Ls-AChBP ELACHBP ELIC GLIC M. ACh α1 T. ACh α T. ACh α GABA α1 CABA β2	154       IDLK-TDTD-Q-V-DLSSYYASSKYEILSATQTRQVQHYSCCPEPY         149       ISVD-PTTE-N-S-DDSEYFSQYSRFEILDVTQKKNSVTYSCCPEAY         149       FALI-TGEEGV-V-NIAEYFDSPKFDLLSATQSLNRKKYSCCENMY         128       QQLRF-SDI-QYY-TENIDNEEIDEWWIRKASTHISDIRYDHLSSVQPNQNEF         136       IVLA-VDLE-K-V-GK-NDDVFLTGWDIESFTAVVKPANFALEDRLE         156       VAIN-PESD-Q-P-DLSNFMESGEWVIKEARGWKHWVYSCCPTTPY         156       VSIS-PESD-R-P-DLSTFMESGEWVIKDYRGWHWVYTCCPDTPY         155       VILQ-HALDAM-I-NQDAFTENGEWTIEHKPSRKNWSNDDISFY         156       VILQ-LSAE-EGI-DPEDFTENGEWTIHKPAKKNYNQLTKDDIYY         158       ISMD-LI-I-DPEAFTENGEWEIIHKPAKKNYNGLTKDDIYY         158       ISMD-LI-DPEAFTENGEWEIIHKPAKKNYGDKFPNGTNY         164       VYE-WTRE-P-ARSVVVAEDGSENGEWEIIHKPAKKNYSG-T	I D V 198 E D V 194 D I 194 S R I 180 S K I 180 D I 201 E D V 198 Q E I 201 Q D V 200 V M 212 P R 207
Ac-AChBP Ls-AChBP Bt-AChBP ELIC GLIC GLIC M. ACh α1 T. ACh α T. ACh α T. ACh β T. ACh δ GABA α1 GABA β2	154       IDLK-TDTD-Q-V-DLSSYYASSKYEILSATQTRQVQHYSCCPEPY         149       ISVD-PTTE-N-S-DDSEYFSQYSRFEILDVTQKKNSVTYSCCPEAY         149       FALI-TGEEGV-V-NIAEYFDSPKFDLLSATQSLNRKKYSCCENMY         128       QQLRF-SDI-QYY-TENIDNEIDEWWIRKASTHISDIRYDHLSSVQPNQNEF         136       IVLA-VDLE-K-V-GK-NDDVFLTGWDIESFTAVVKPANFALEDRLE         156       VAIN-PESD-Q-P-DLSNFMESGEWVIKEARGWKHWVYYSCCPTTPY         156       VSIS-PESD-R-P-DLSTFMESGEWVIKDYRGWKHWVYYTCCPDTPY         156       VILQ-HALDAM-I-NQDAFTENGGWSIEHKPSRKNWRSD	I DV         198           DV         194           DI         194           S RI         180           S KL         180           DI         201           DI         201           DI         201           DI         201           QU         198           QU         198           QU         198           QU         198           QU         198           QU         198           QU         200           VVM         212           PRL         207           VVM         212



Fig. S.I.-3: Ramachandran plot for the refined model. Triangles represent Gly residues. Generated with Procheck.<sup>34</sup>

Fig. S.I.-4: ProSA-web<sup>35</sup> analysis of the local model quality, showing the improvements from the initial (black) to the refined model (red). This can also be thought of as the individual residue contributions to the z-score. In general, regions above zero in the shown plots indicate potentially problematic regions.  $\beta_2$  subunit comparison (chain A) to the initial model 1a is shown to the left, and  $\alpha_1$  subunit comparison (chain B) to initial model 1b is shown to the right. A z-score ("knowledge-based energy") is calculated for each residue, however to improve readability the plot is smoothed using two different window sizes (10 and 40). With window size 40, the average score is calculated for each 40-residue interval (*i*,*i*+39) and assigned to the central residue (*i*+19) of that interval.





Fig. S.I.-5. Chain and residue specific Cα RMSF for the production MD simulation.

# **Reference List**

- 1. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 1993;234:779-815.
- 2. Maestro. New York, NY: Schrödinger, LLC; 2009.
- Boileau AJ, Evers AR, Davis AF, Czajkowski C. Mapping the agonist binding site of the GABA<sub>A</sub> receptor: Evidence for a β-strand. J Neurosci 1999;19:4847-4854.
- Holden JH, Czajkowski C. Different residues in the GABA<sub>A</sub> receptor α<sub>1</sub>T60-α<sub>1</sub>K70 region mediate GABA and SR-95531 actions. J Biol Chem 2002;277:18785-18792.
- Jansen M, Rabe H, Strehless A, Dieler S, Debus F, Dannhardt G, Akabas MH, Luddens H. Synthesis of GABA<sub>A</sub> receptor agonists and evaluation of their α-subunit selectivity and orientation in the GABA binding site. J Med Chem 2008;51:4430-4448.
- Schrödinger Suite 2009 Protein Preparation Wizard; Epik version 2.0, Schrödinger, LLC, New York, NY, 2009; Impact version 5.5, Schrödinger, LLC, New York, NY, 2009; Prime version 2.1, Schrödinger, LLC, New York, NY, 2009.
- Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926-935.
- Goodford PJ. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. J Med Chem 1985;28:849-857.
- 9. GRID. Pinner, Middlesex: Molecular Discovery, Ltd.; 2005.
- 10. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684-3690.
- 11. Nosé S. A molecular dynamics method for simulations in the canonical ensemble. Mol Phys 1984;52:255-268.
- 12. Martyna GJ, Tobias DJ, Klein ML. Constant pressure molecular dynamics algorithms. J Chem Phys 1994;101:4177-4189.
- 13. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 1977;23:327-341.
- 14. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. J Chem Phys 1995;103:8577-8593.
- 15. Kryshtafovych A, Venclovas C, Fidelis K, Moult J. Progress over the first decade of CASP experiments. Proteins 2005;61 Suppl 7:225-236.
- Campagna-Slater V, Weaver DF. Molecular modelling of the GABA<sub>A</sub> ion channel protein. J Mol Graph Model 2007;25:721-730.

- 17. Harrison NJ, Lummis SC. Molecular modeling of the GABA<sub>C</sub> receptor ligand-binding domain. J Mol Model 2006;12:317-324.
- Ernst M, Brauchart D, Boresch S, Sieghart W. Comparative modeling of GABA<sub>A</sub> receptors: Limits, insights, future developments. Neuroscience 2003;119:933-943.
- 19. Ernst M, Bruckner S, Boresch S, Sieghart W. Comparative models of GABA<sub>A</sub> receptor extracellular and transmembrane domains: Important insights in pharmacology and function. Mol Pharmacol 2005;68:1291-1300.
- Adamian L, Gussin HA, Tseng YY, Muni NJ, Feng F, Qian H, Pepperberg DR, Liang J. Structural model of ρ1 GABA<sub>C</sub> receptor based on evolutionary analysis: Testing of predicted protein-protein interactions involved in receptor assembly and function. Protein Sci 2009;18:2371-2383.
- 21. Melis C, Lummis SC, Molteni C. Molecular dynamics simulations of GABA binding to the GABA<sub>C</sub> receptor: The role of Arg104. Biophys J 2008;95:4115-4123.
- 22. Mokrab Y, Bavro VN, Mizuguchi K, Todorov NP, Martin IL, Dunn SM, Chan SL, Chau PL. Exploring ligand recognition and ion flow in comparative models of the human GABA type A receptor. J Mol Graph Model 2007;26:760-774.
- 23. Law RJ, Lightstone FC. Modeling neuronal nicotinic and GABA receptors: Important interface salt-links and protein dynamics. Biophys J 2009;97:1586-1594.
- 24. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 1999;41:95-98.
- Newell JG, Czajkowski C. The GABA<sub>A</sub> Receptor α<sub>1</sub> Subunit Pro174-Asp191 Segment Is Involved in GABA Binding and Channel Gating. J Biol Chem 2003;278:13166-13172.
- 26. Amin J, Weiss DS.  $GABA_A$  receptor needs two homologous domains of the  $\beta$ -subunit for activation by GABA but not by pentobarbital. Nature 1993;366:565-569.
- 27. Wagner DA, Czajkowski C. Structure and dynamics of the GABA binding pocket: A narrowing cleft that constricts during activation. J Neurosci 2001;21:67-74.
- Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 2004;47:1739-1749.
- Frølund B, Jørgensen AT, Tagmose L, Stensbøl TB, Vestergaard HT, Engblom C, Kristiansen U, Sanchez C, Krogsgaard-Larsen P, Liljefors T. Novel class of potent 4-arylalkyl substituted 3-isoxazolol GABA<sub>A</sub> antagonists: Synthesis, pharmacology, and molecular modeling. J Med Chem 2002;45:2454-2468.
- Frølund B, Jensen LS, Guandalini L, Canillo C, Vestergaard HT, Kristiansen U, Nielsen B, Stensbøl TB, Madsen C, Krogsgaard-Larsen P, Liljefors T. Potent 4aryl- or 4-arylalkyl-substituted 3-isoxazolol GABA<sub>A</sub> antagonists: Synthesis, pharmacology, and molecular modeling. J Med Chem 2005;48:427-439.

- Frølund B, Jensen LS, Storustovu SI, Stensbøl TB, Ebert B, Kehler J, Krogsgaard-Larsen P, Liljefors T. 4-aryl-5-(4-piperidyl)-3-isoxazolol GABA<sub>A</sub> antagonists: Synthesis, pharmacology, and structure-activity relationships. J Med Chem 2007;50:1988-1992.
- Møller HA, Sander T, Kristensen JL, Nielsen B, Krall J, Bergmann ML, Christiansen B, Balle T, Jensen AA, Frølund B. Novel 4-(piperidin-4-yl)-1hydroxypyrazoles as γ-aminobutyric acid<sub>A</sub> receptor ligands: Synthesis, pharmacology, and structure-activity relationships. J Med Chem 2010;53:3417-3421.
- Clamp M, Cuff J, Searle SM, Barton GJ. The Jalview Java alignment editor. Bioinformatics 2004;20:426-427.
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 1993;26:283-291.
- Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 2007;35:W407-W410.