Supplementary Material

Comments on 5-HT_{2A} model and sequence alignment published by other groups

(1) *Comparing our model with the* ***Chambers and Nichols***[36] *model on 5-HT_{2A}*

(i) The model of 5-HT_{2A} by Chambers and Nichols was constructed using bovine rhodopsin as a template. We have used the template β_2-AR to build our model, which is better in two ways (a) sequence identity of 5-HT_{2A} with β_2-AR is much better than with Bovine Rhodopsin, (b) both β_2-AR and 5-HT_{2A} belong to amine GPCR, which is a sub-group of Class A GPCR.

(ii) Moreover, in **Chambers and Nichols** model, the network of polar interaction ("ionic lock") was preserved between Arg residue (R135 in bovine rhodopsin) from conserved E/DRY motif of TM3 helix and Glu residue (E247 in bovine rhodopsin) of TM6, which bridges these two transmembrane helices, stabilizing the inactive-state conformation. It is also mentioned that the same interaction is preserved in all class A GPCRs. But the analogous polar interactions are broken in all the ligand-activated GPCR crystal structures, β_2-AR, β_1-AR and $A_{2A,}$ that were published very recently. Please see reference number "61" in the main article. Since we have used β_2-AR as a template, our model also retains the lack of network of polar interaction ("ionic lock").

(2) *Comparison of our model with* ***Evers and co-workers'*** *model*[37]

Evers and co-workers also generated the protein models by applying ligand-supported homology modelling using bovine rhodopsin as a template. The authors claims that the comparison of the different methods in retrieving known antagonists from the virtual libraries shows that the ligand-based screening techniques outperform the molecular docking approach when sufficient ligand information is used for the generation of models. However, he agrees that the present study cannot provide an exhaustive comparison of all currently available virtual screening methods. Also, they suggest that protein models can be used as a structural basis for the generation of relevant binding poses and ligand alignments, which would be useful for the subsequent generation of 3D-QSAR models. In another article published by the same group, which came on the same issue of the same journal (Evers, A and Klabunde, T. 2005, 48:1088-1097), the authors claim that structure-based homology models may be used as a structural basis for GPCR lead finding and compound optimization. When considering the ligand-supported homology model used in that study, one would like to be cautious that the model could be inaccurate with the docking poses due to the distant homology to bovine Rhodopsin (template). It would be enthusiastic to see the similar kind of virtual screening on 5-HT_{2A} model built using better template such as adrenergic receptors.

Since 3D-coordinates and the information on model quality through PROCHECK or other methods were not available for the model constructed either by **Chambers and Nichols** or by **Evers and co-workers**, we are unable to perform detailed structure comparisons of our model with theirs.

(3) *Comparison of our model with **Bruno and co-workers'** model*[35]

The paper by Bruno and co-workers' model on 5-HT_{2A} was published subsequent to the analysis described and reported in our manuscript.

(i) Brunos' model was built by using β_2-AR as a template, which was considered as a more suitable template than the bovine Rhodopsin as we have done. The alignment constructed by us to model the template is consistent with the alignment constructed by Bruno and co-workers except that the alignment at TM5 was eight residues longer than their alignment at TM5.

(ii) Our model is consistent with their model in preserving the disulphide bridge between C148 on TM3 and C227 on ECL2.

(iii) The model coordinates could not be obtain for a rigorous comparison. It is not clear if their model includes N-terminal and C-terminal loop regions, but, it seems that the model was built with ICL3 region (viewed through their results part of root mean square fluctuation on residues. Compared to their model, our model was built with the loop regions which includes N-terminal, C- terminal and ICL3 region, which further raise a question of stability of their incomplete monomeric structure that went through molecular dynamics and derived docking studies.

(iv) Though, the PROCHECK results were included in Bruno and coworkers' article, we are unable to find the statistical numbers for residues falling under allowed regions to compare with our PROCHECK results. But, comparing the results graphically, our result shows that fewer residues of our model fall outside the allowed regions of Ramachandran plot despite the presence of residues at N-terminal and ICL3 regions in our model (Supplementary Table 1).

(4) *Comparing our 5-HT_{2A} sequence alignment with **Johnson and co-workers'** alignment*[41]

Our sequence comparison of human 5-HT_{2A} with monkey, pig and rat 5-HT_{2A} was consistent with the alignment by Johnson and co-workers. Analysis of our alignment result shows that the most variable regions are found in the N- and C-terminal regions (Supplementary Figures 3, 4 and 5) and very few variable residues found at TM regions.

Legend to Supplementary Figure

Supplementary Figure SF1: **Human amine GPCR sequence alignment**. The human amine GPCR sequence alignment (42 sequences) includes sequences from dopamine receptors, histamine receptors, adrenergic receptors, angiotensin type-II receptors, trace amine associated receptors and 5-HT receptors and were referred as human amine GPCR sequences. This alignment shows the conserved residues and motifs, which are described inside result part under the same heading.

Supplementary Figure SF2: **Human 5-HT class GPCR alignment.** There are 12 sequences from 5-HT_1, 5-HT_2, 5-HT_4, 5-ht_5, 5-HT_6 and 5-HT_7 receptors and their subtypes of humans were refer to human 5-HT receptor sequences.

Supplementary Figure SF3: **Sequence Alignment of 5-HT2a orthologous receptors.** Orthologous of 5-HT_{2A} receptor sequences (14 sequences) were aligned to identify the amino acid exchanges

Supplementary Figure SF4: **Comparison of human 5-HT_{2A} sequence against rat 5-HT_{2A}.** A close comparison of sequences between human and rat (*Rattus norvegieus*) 5-HT_{2A} receptor was performed, to identify residues which are specific to the corresponding species and are reported in Figure 3 in the main article.

Supplementary Figure SF5: **Comparison of human 5-HT_{2A} sequence against rodents 5-HT_{2A}.** The 5-HT_{2A} sequences of rodents such as *Rattus norvegieus*, *Mus musculus*, *Cricetulus griesius*, *Chinese hamster* and *Mesocricetus auratus* were compared against human 5-HT_{2A}.

Supplementary Figure SF6: **Comparison of human 5-HT_{2A} sequence against human 5-HT_{2B} and 5-HT_{2C} sequences.** Comparison of amino acid exchanges across the subclasses, 5-HT_{2A}, 5-HT_{2B} and 5-HT_{2C}.

Supplementary Figure SF7: **PROCHECK results on built model.** Fully allowed regions are marked in yellow and partially allowed regions in red. A vast majority of the points are within allowed and partially allowed regions. The analyzed results were tabulated in Supplementary Table 1. Results of applying the validation program, PROCHECK, on the final model of 5-HT_{2A} (A) before loop (N-terminal, ICL3 and C-terminal regions) building, and (B) after loop building.

Supplementary Figure SF8: **Verify3d results on built model.** The built model was validated through verify3d server and results were considered tabulated in Supplementary Table 1.

Supplementary Figure SF9: Analysis of correlation of the binding energy with experimental affinity values derived from the literature[25, 40-43]. Similar experimental data are not available for dopamine.

Supplementary Figure SF10: **Comparison of side chain orientation of conserved residues on each TM helices.** The orientation of conserved residues on each TM helices of the model were compared with the available crystal structures (Bovine Rhodopsin, Turkey β_1-AR and Human Adenosin A_{2A} receptor) by superposing them, which excludes the template crystal structure. (A) Asn in TM1, Asp in TM2, Arg in TM3 and

Trp in TM4 were compared for their side chain orientation and represented by sticks. (B) The conserved Pro residue on TM5, 6 and 7 was compared for their side chain orientation and represented by sticks

Supplementary Figure SF11: **Representation of docked poses for each ligand.** Ligands are coloured in cyan and the residue which makes H-bond with the ligand coloured in pink. The secondary structural regions are marked, from where the residue comes to make the H-bond. H-bonds are marked in dotted yellow line.

Supplementary Figure SF12a: **Ligplot representation**. Representation of interaction is depicted for the ligands such as Serotonin, Dopamine, DOI and LSD. H-bond, hydrophobic and hydrophilic residues are ploted and key for the same is given at the end of the figure.

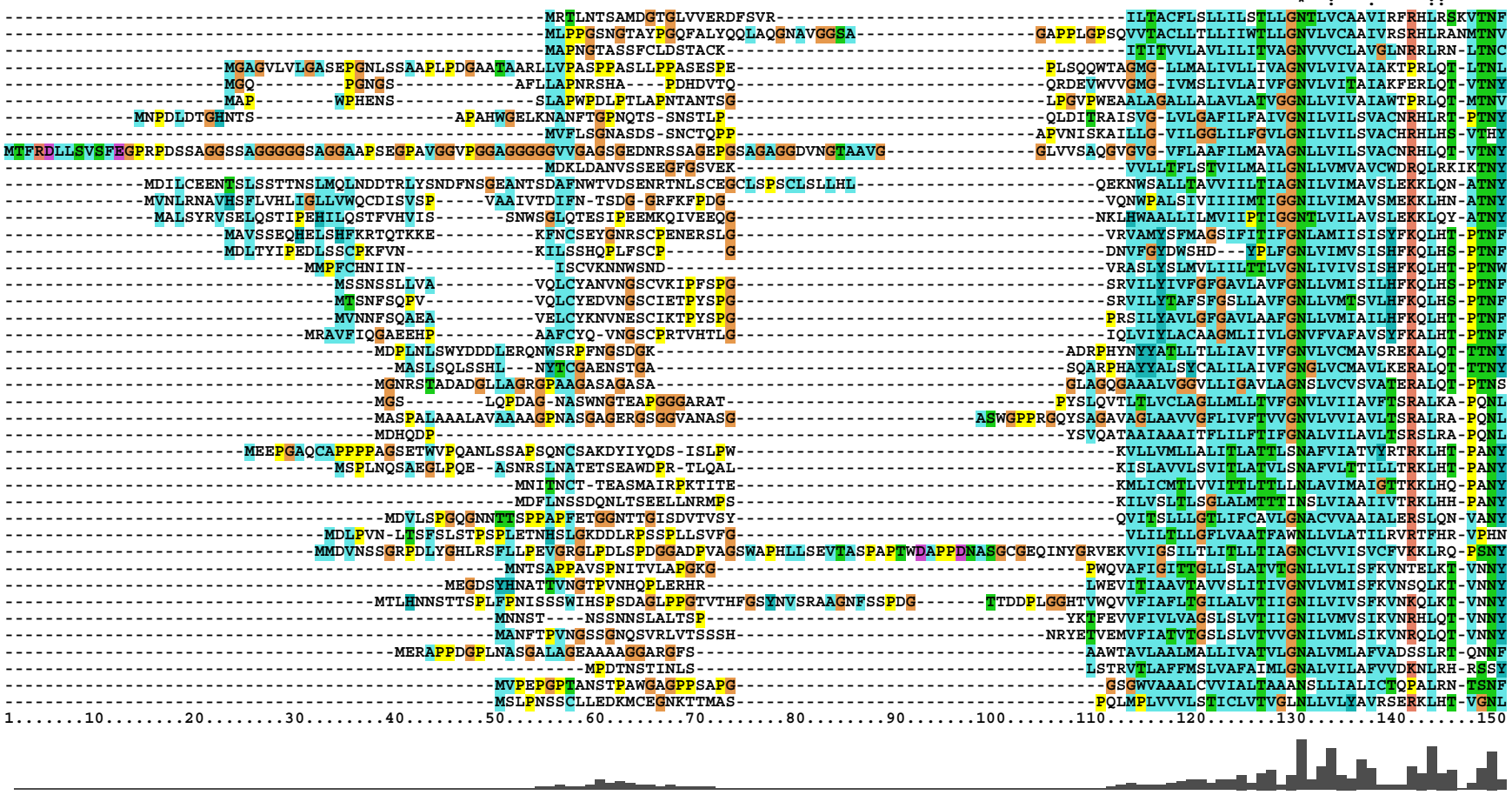
Supplementary Figure SF12b: **Ligplot representation.** Representation of interaction is depicted for the ligands such as Ketanserin, Haloperidol, Clozapine and Risperidone. H-bond, hydrophobic and hydrophilic residues are ploted and key for the same is given at the end of the figure.

Legend to Supplementary Table

Supplementary Table ST1: **Residue length for each TM helix in the built model and their sequence identity with template β2-AR (PDB id 2RH1)**. The regions are also marked for their structural validity through Ramachandran plot and verify3d.

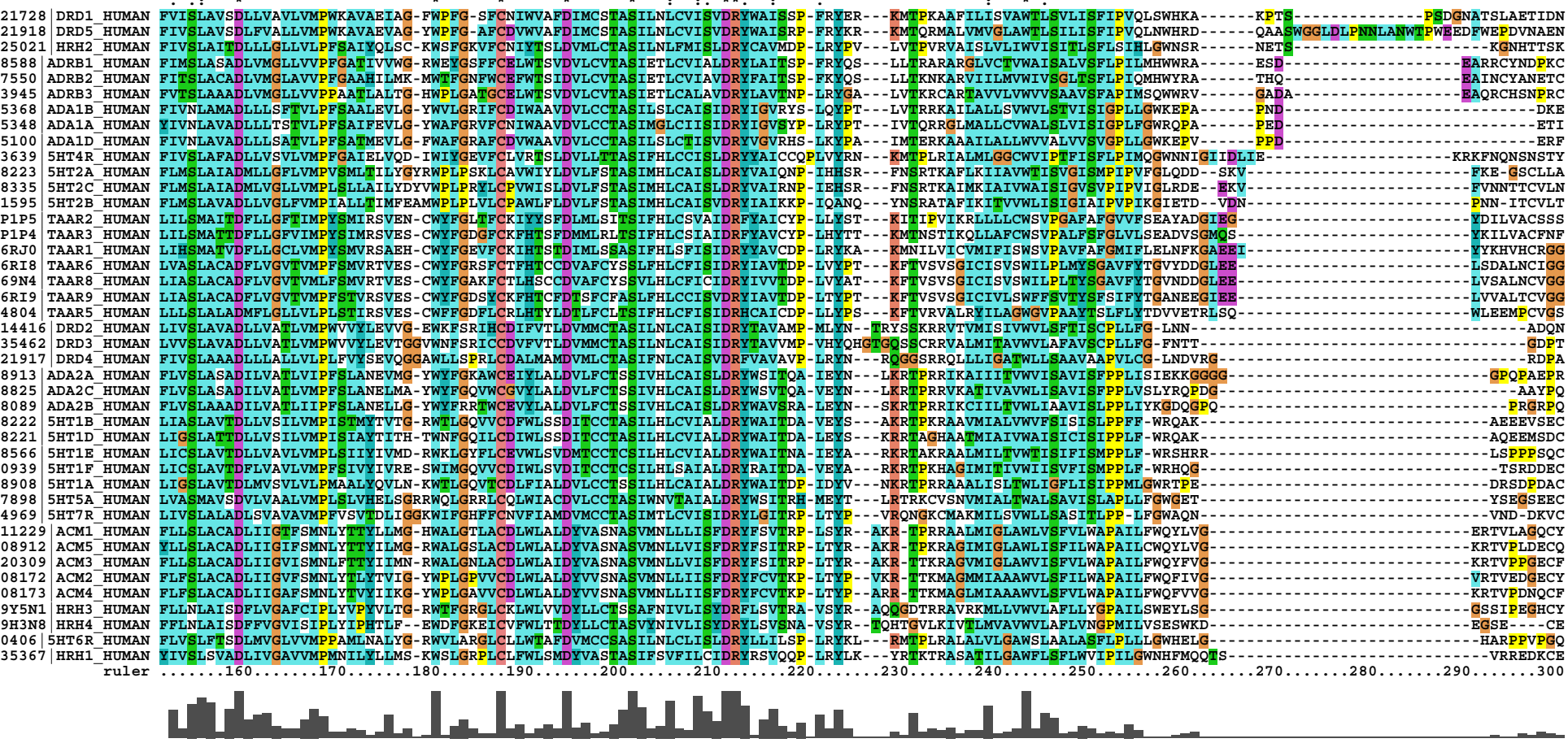
Supplementary Figure 1

Human amine GPCR alignment



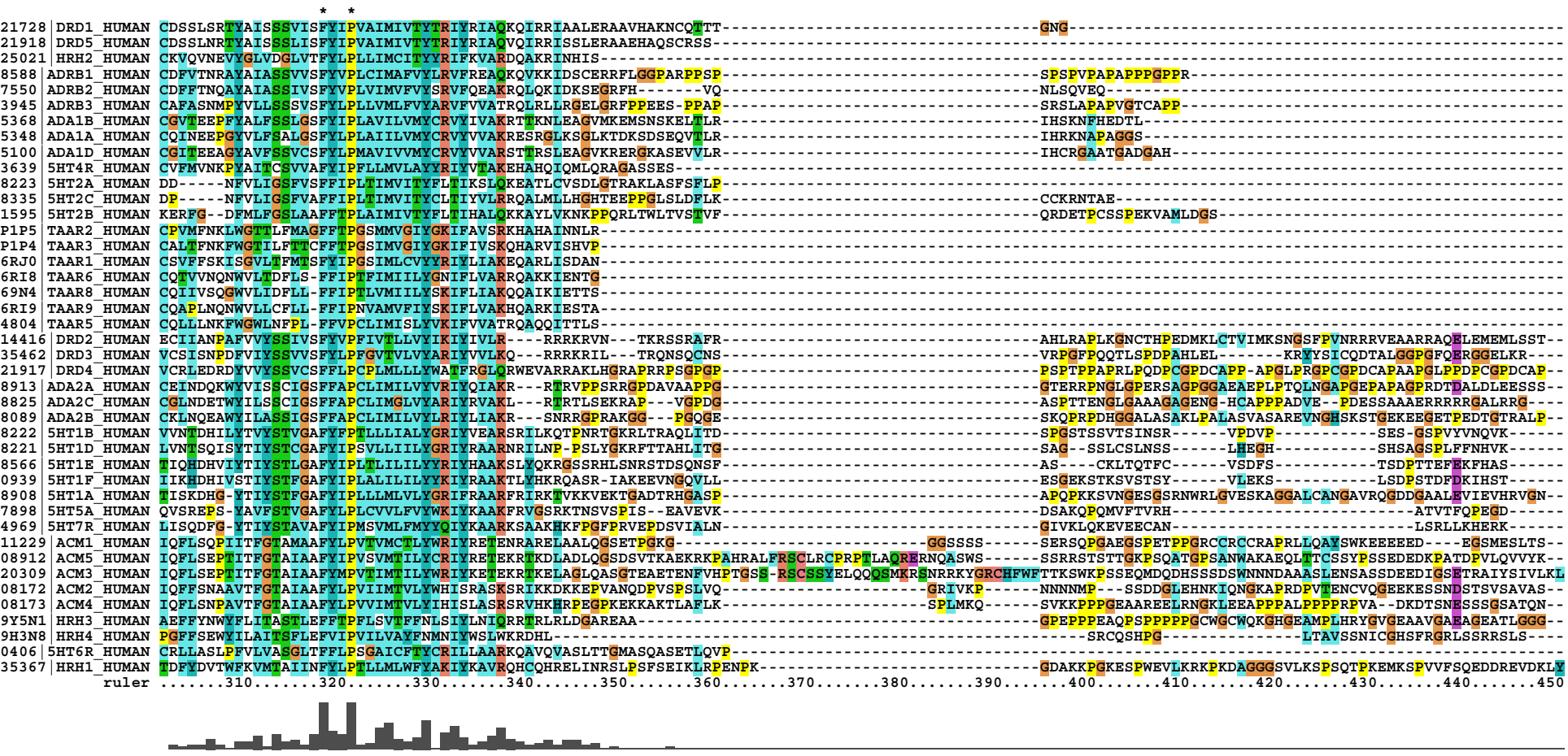
Supplementary Figure 1

Human amine GPCR alignment



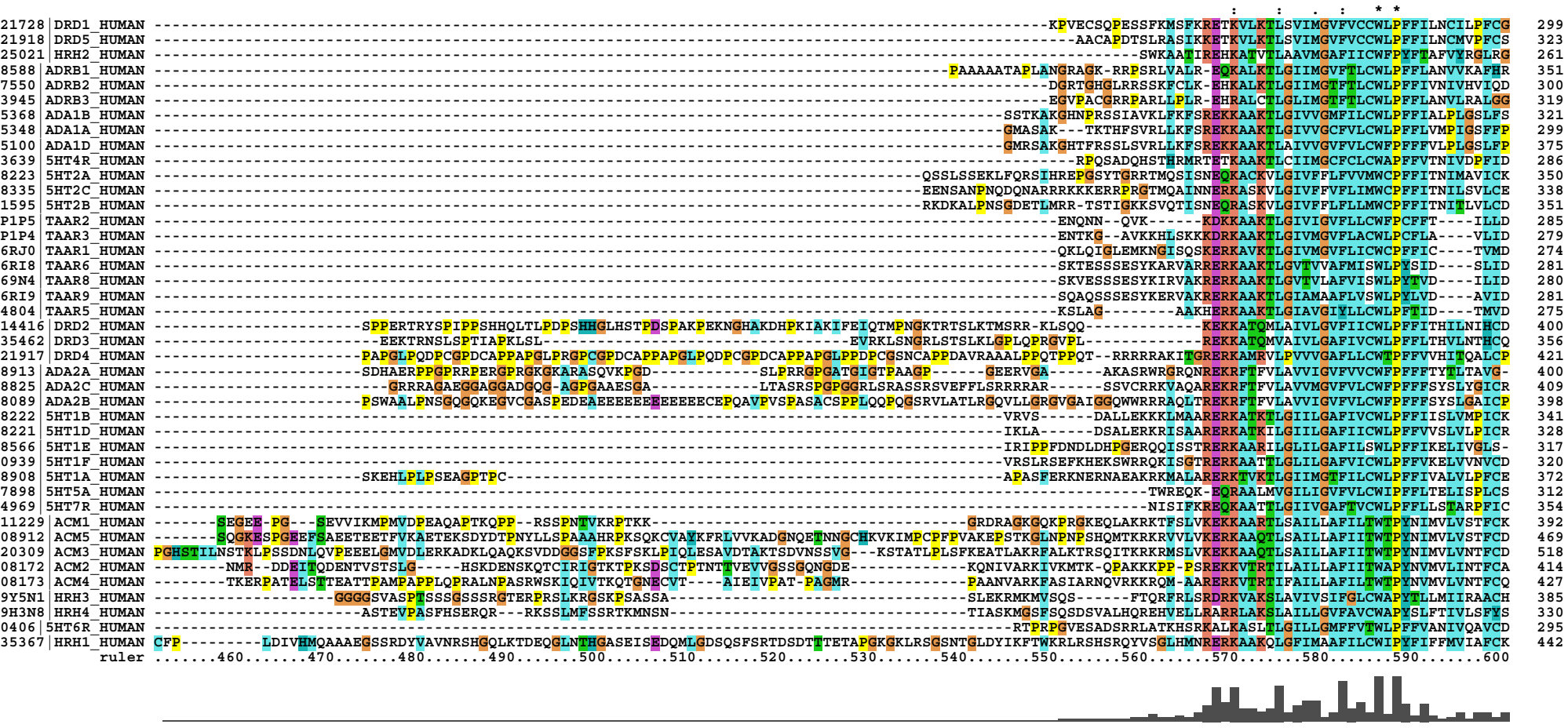
Supplementary Figure 1

Human amine GPCR alignment



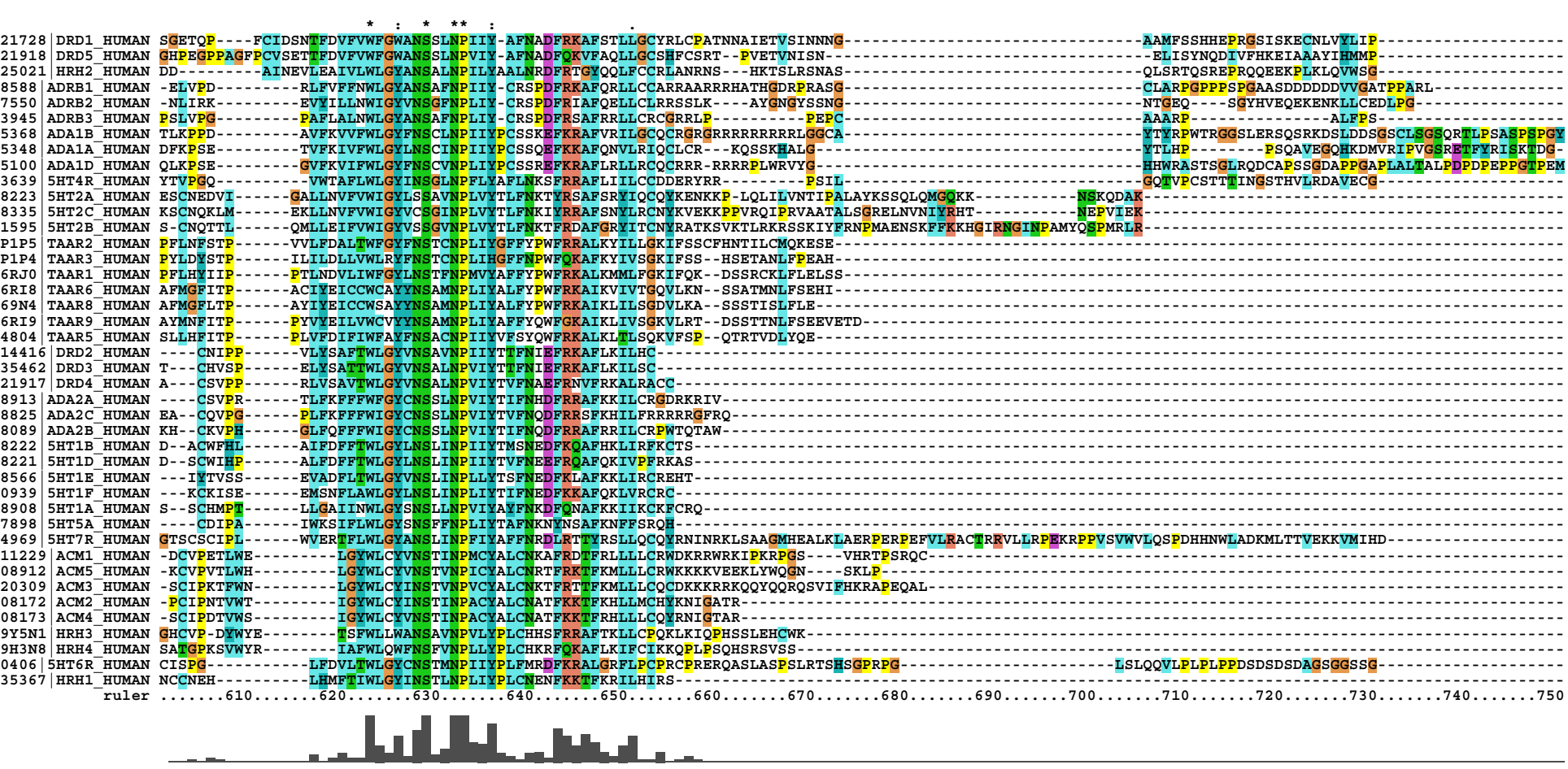
Supplementary Figure 1

Human amine GPCR alignment



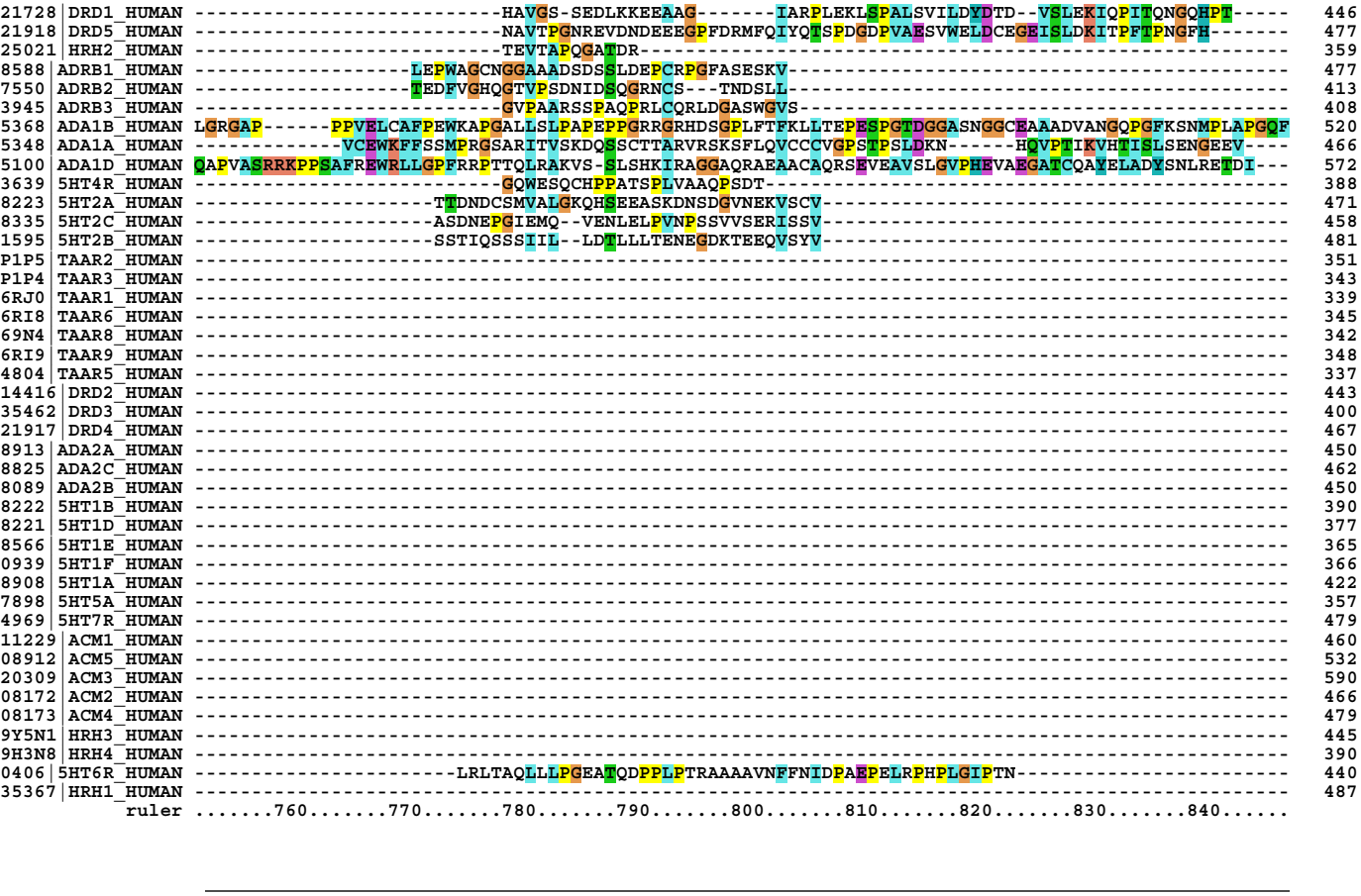
Supplementary Figure 1

Human amine GPCR alignment



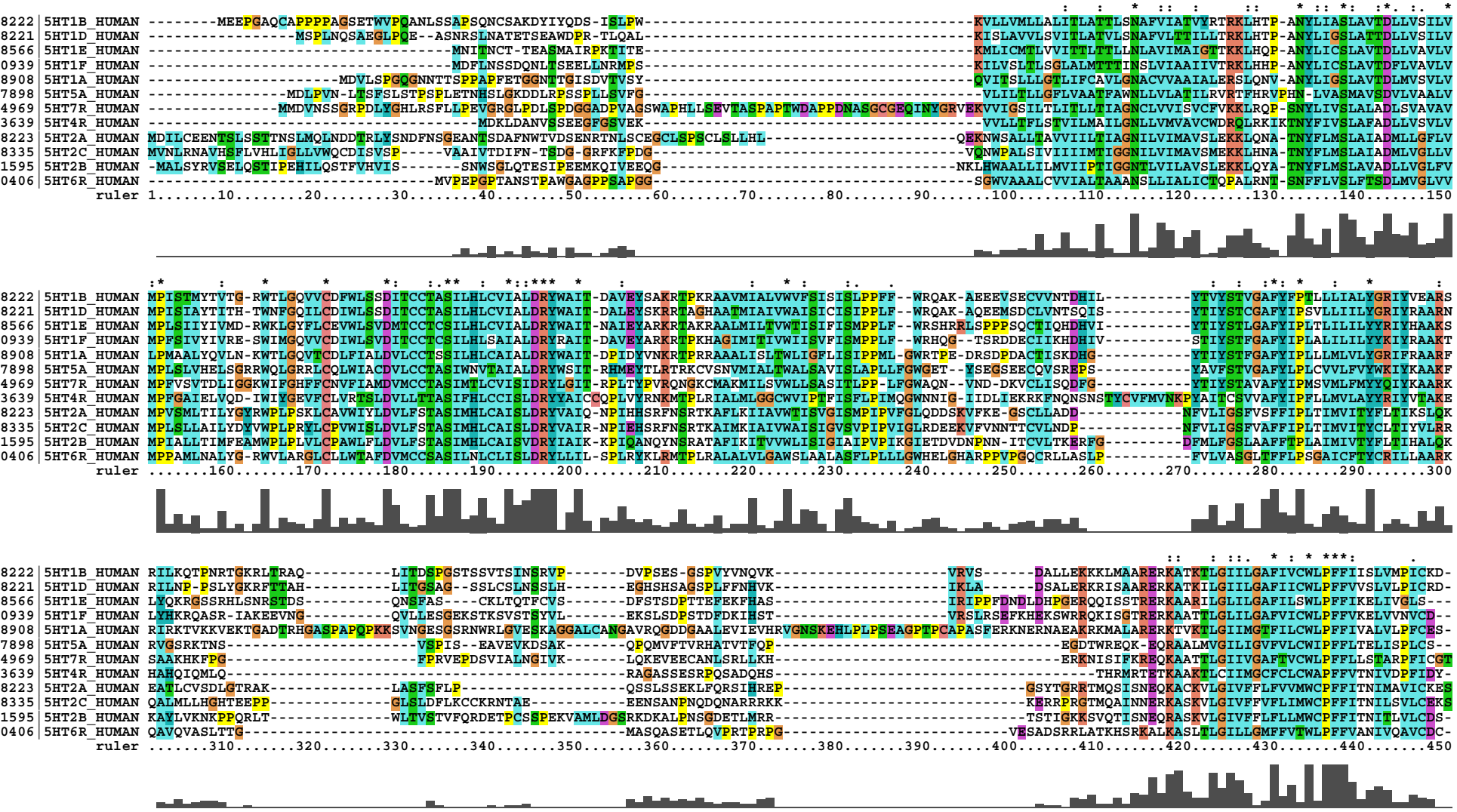
Supplementary Figure 1

Human amine GPCR alignment



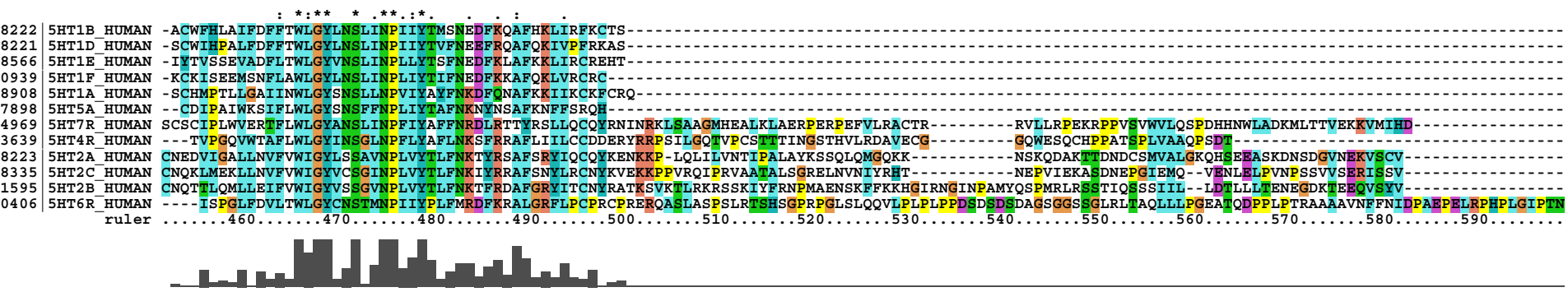
Supplementary Figure 2

Human 5-HT class GPCR alignment



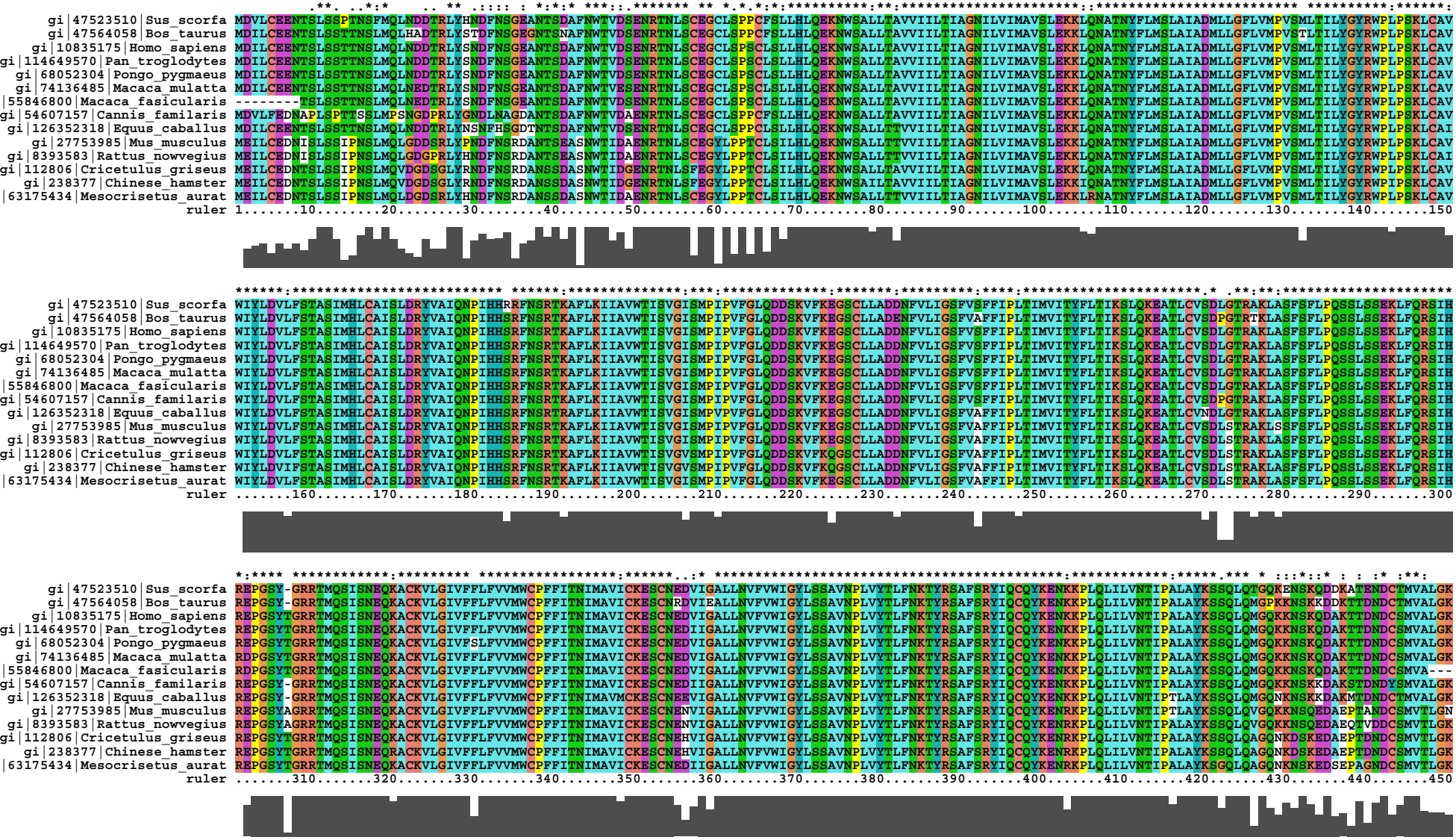
Supplementary Figure 2

Human 5-HT class GPCR alignment



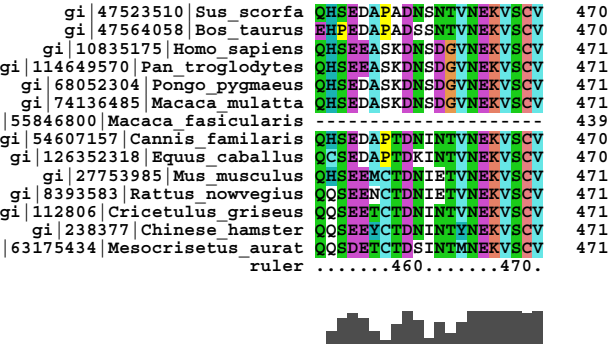
Supplementary Figure 3

Alignment of 5-HT2A orthologus



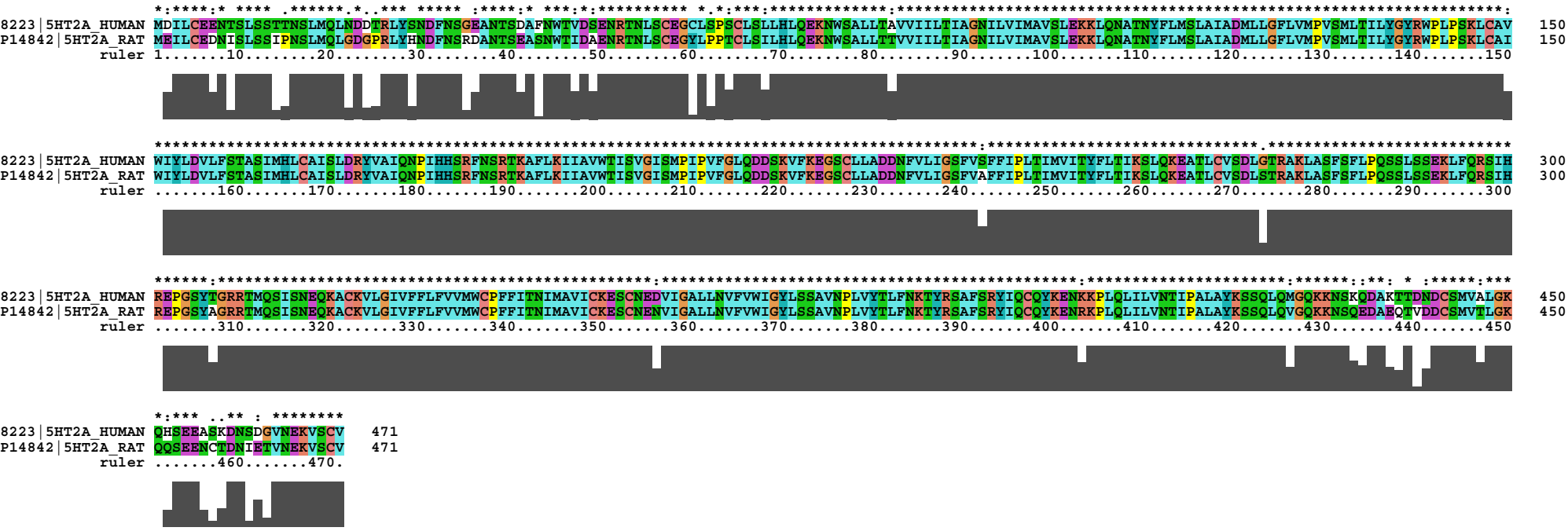
Supplementary Figure 3

Alignment of 5-HT2A orthologus



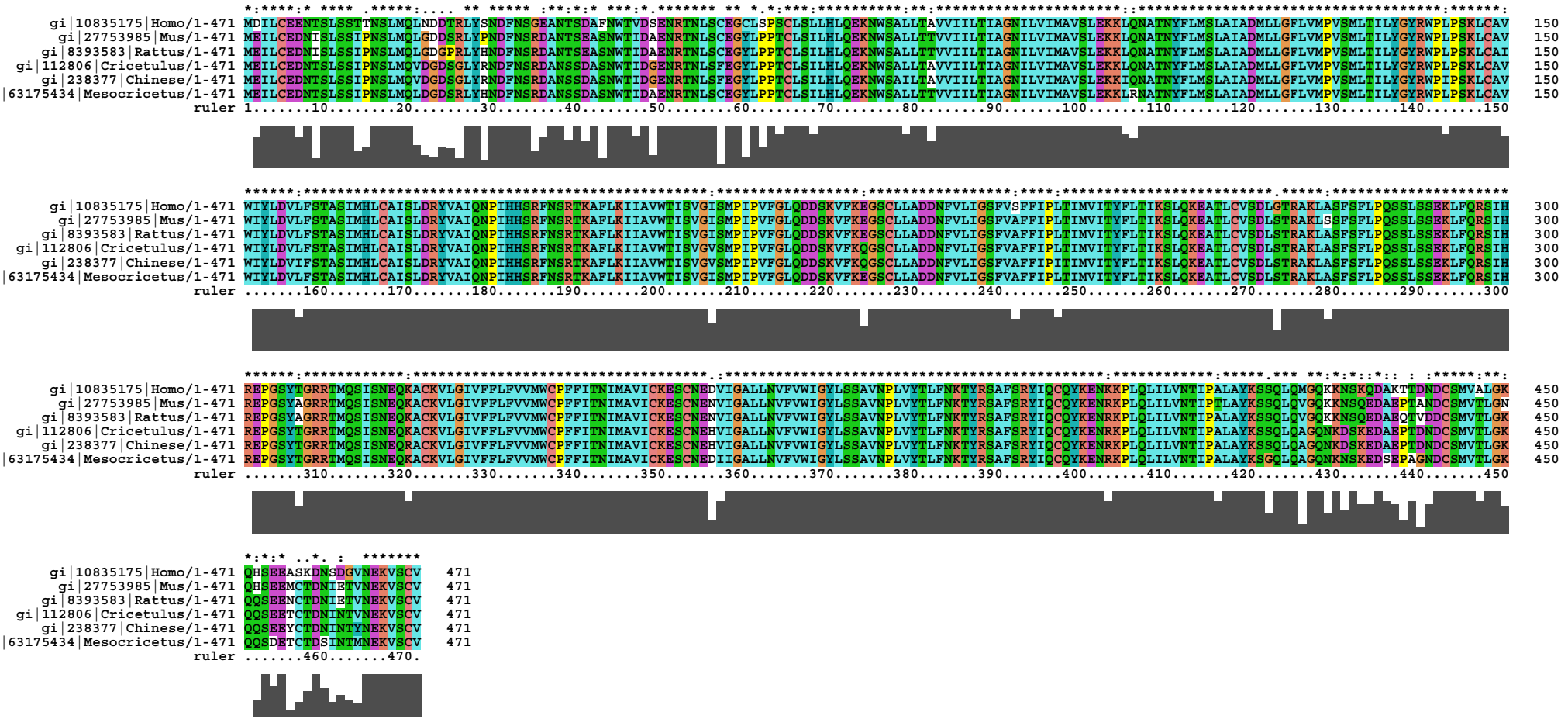
Supplementary Figure 4

Sequence comparison of human 5-HT2A with rat 5-HT2A



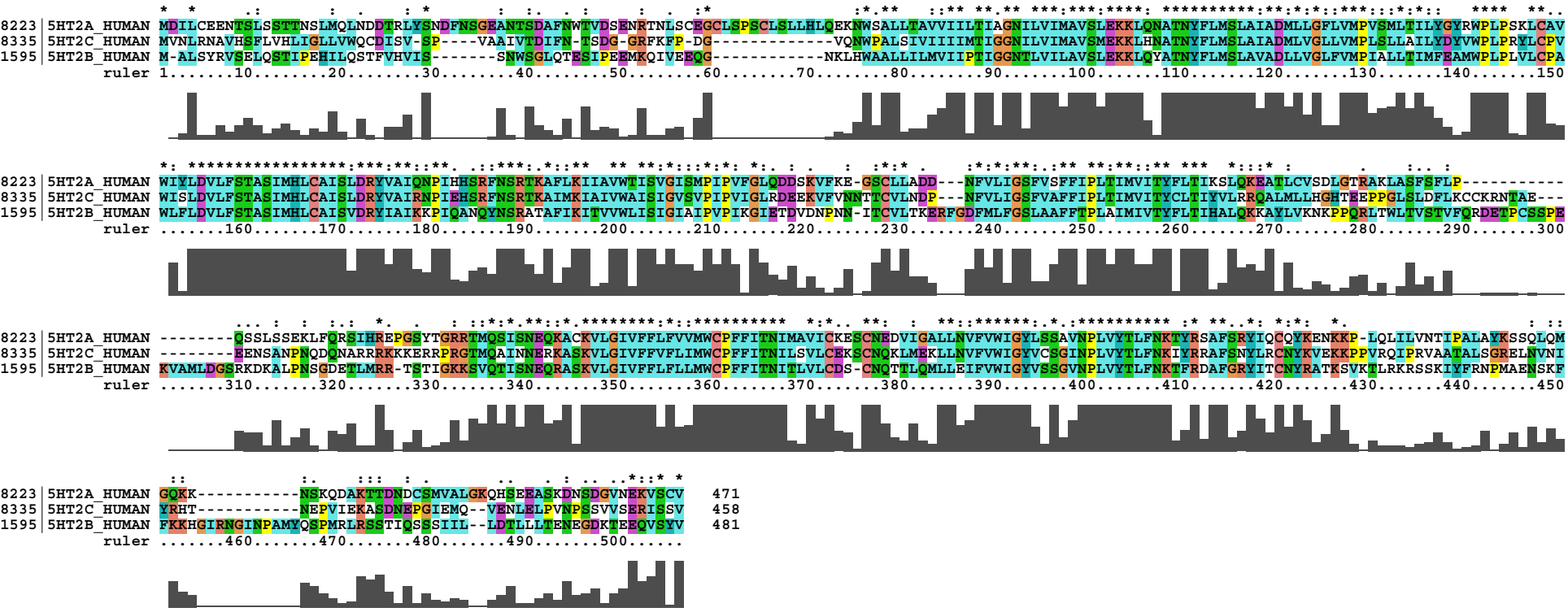
Supplementary Figure 5

Sequence comparison of human 5-HT2A with rodents' 5-HT2A Page 1 of 1



Supplementary Figure 6

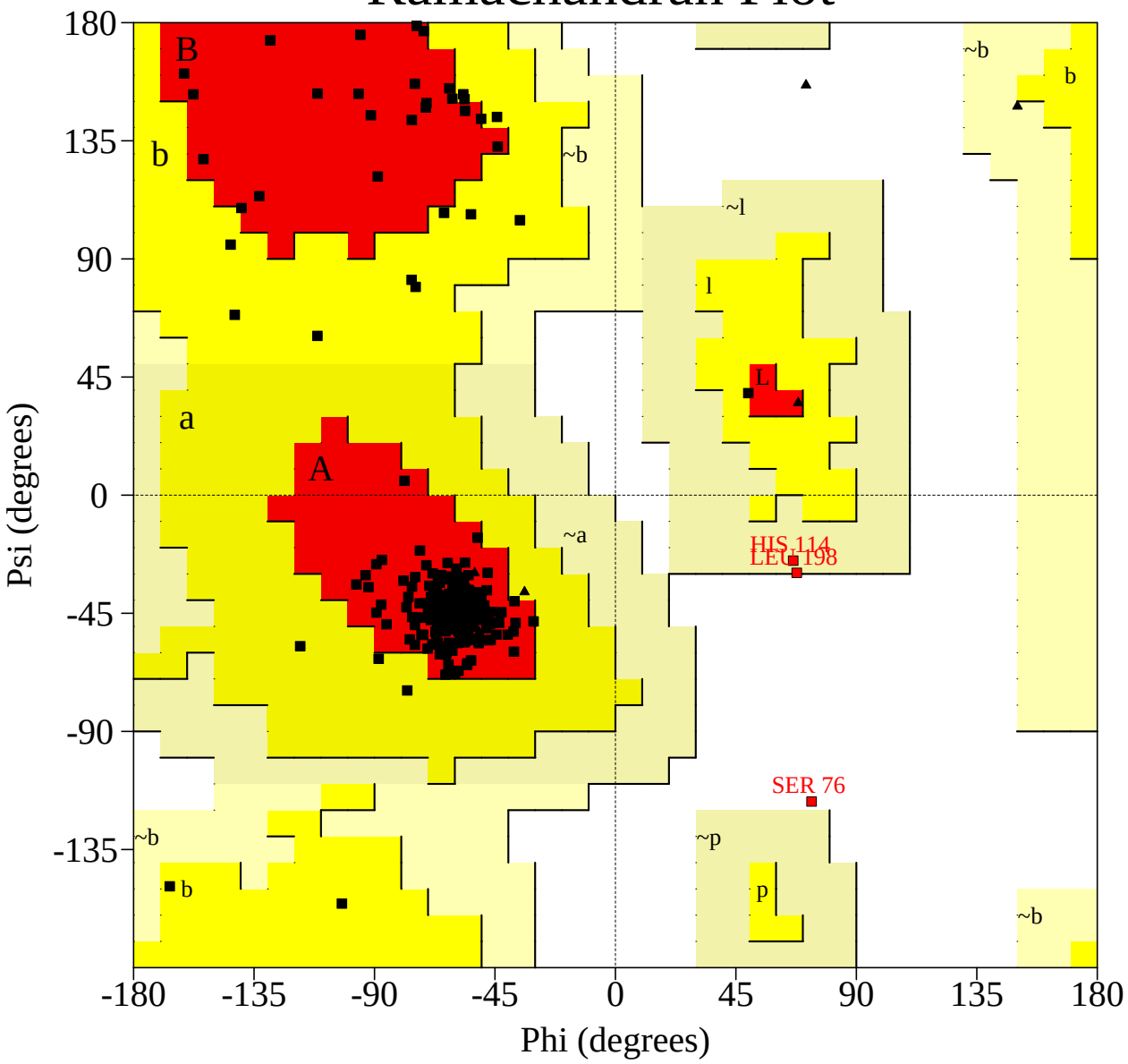
Human 5-HT2 subclass GPCR alignment



PROCHECK

Supplementary Figure 7a

Ramachandran Plot



Plot statistics

Residues in most favoured regions [A,B,L]	242	92.4%
Residues in additional allowed regions [a,b,l,p]	17	6.5%
Residues in generously allowed regions [~a,~b,~l,~p]	2	0.8%
Residues in disallowed regions	1	0.4%
	----	------
Number of non-glycine and non-proline residues	262	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	10	
Number of proline residues	9	

Total number of residues	283	

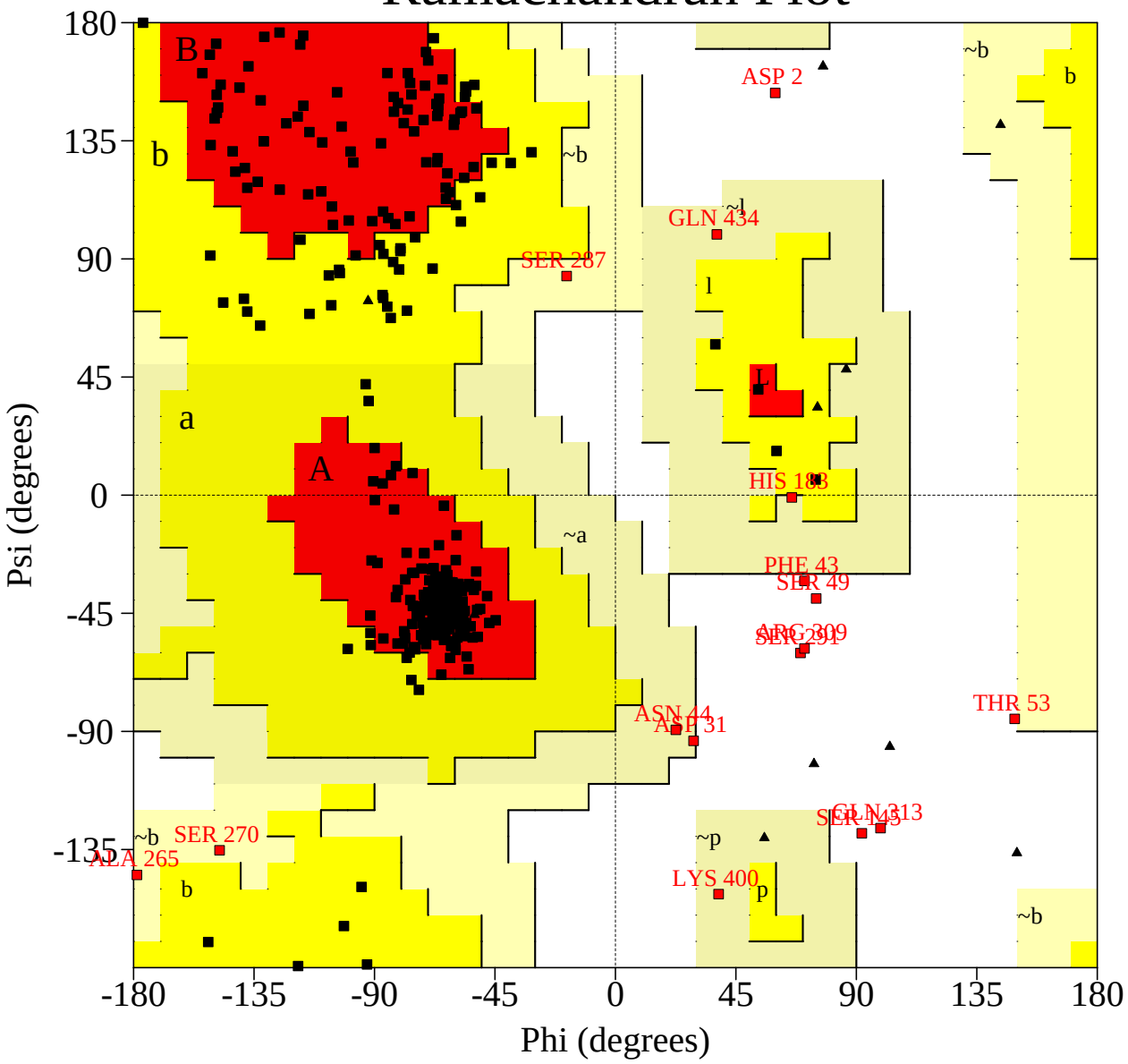
Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

new-entry_01.ps

PROCHECK

Supplementary Figure 7b

Ramachandran Plot



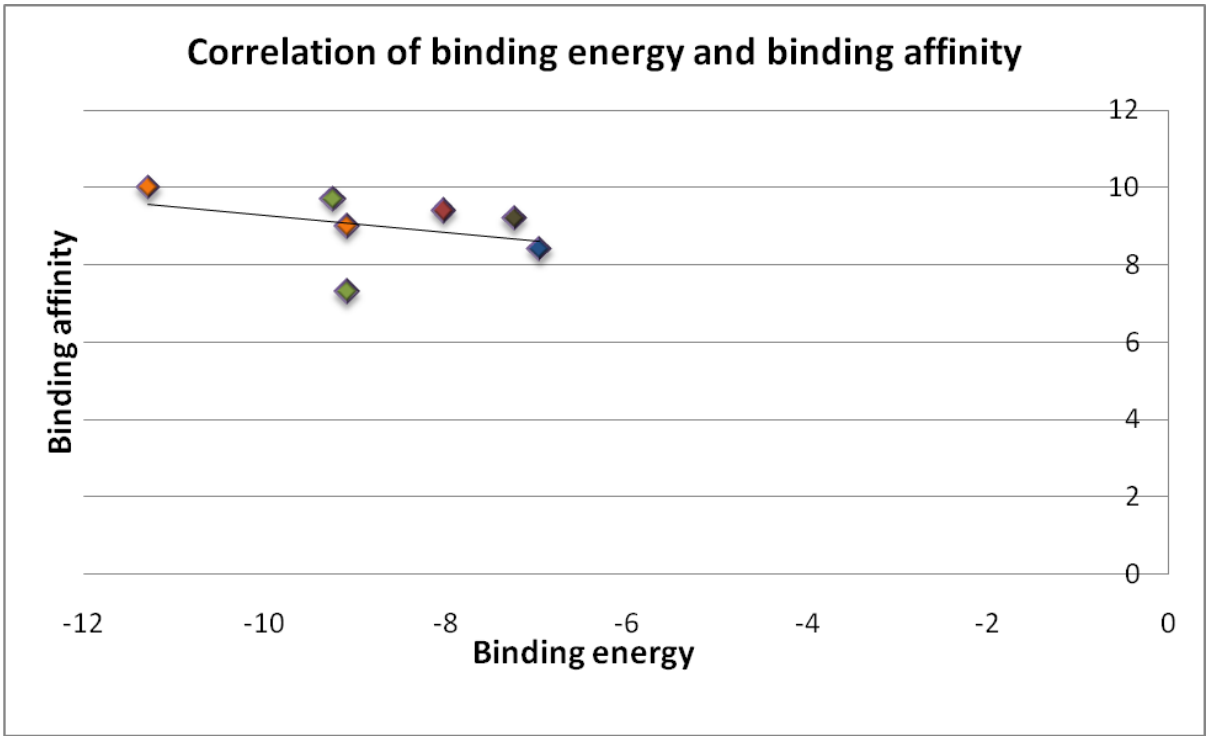
Plot statistics

Residues in most favoured regions [A,B,L]	370	84.7%
Residues in additional allowed regions [a,b,l,p]	51	11.7%
Residues in generously allowed regions [~a,~b,~l,~p]	8	1.8%
Residues in disallowed regions	8	1.8%
	----	------
Number of non-glycine and non-proline residues	437	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	18	
Number of proline residues	14	

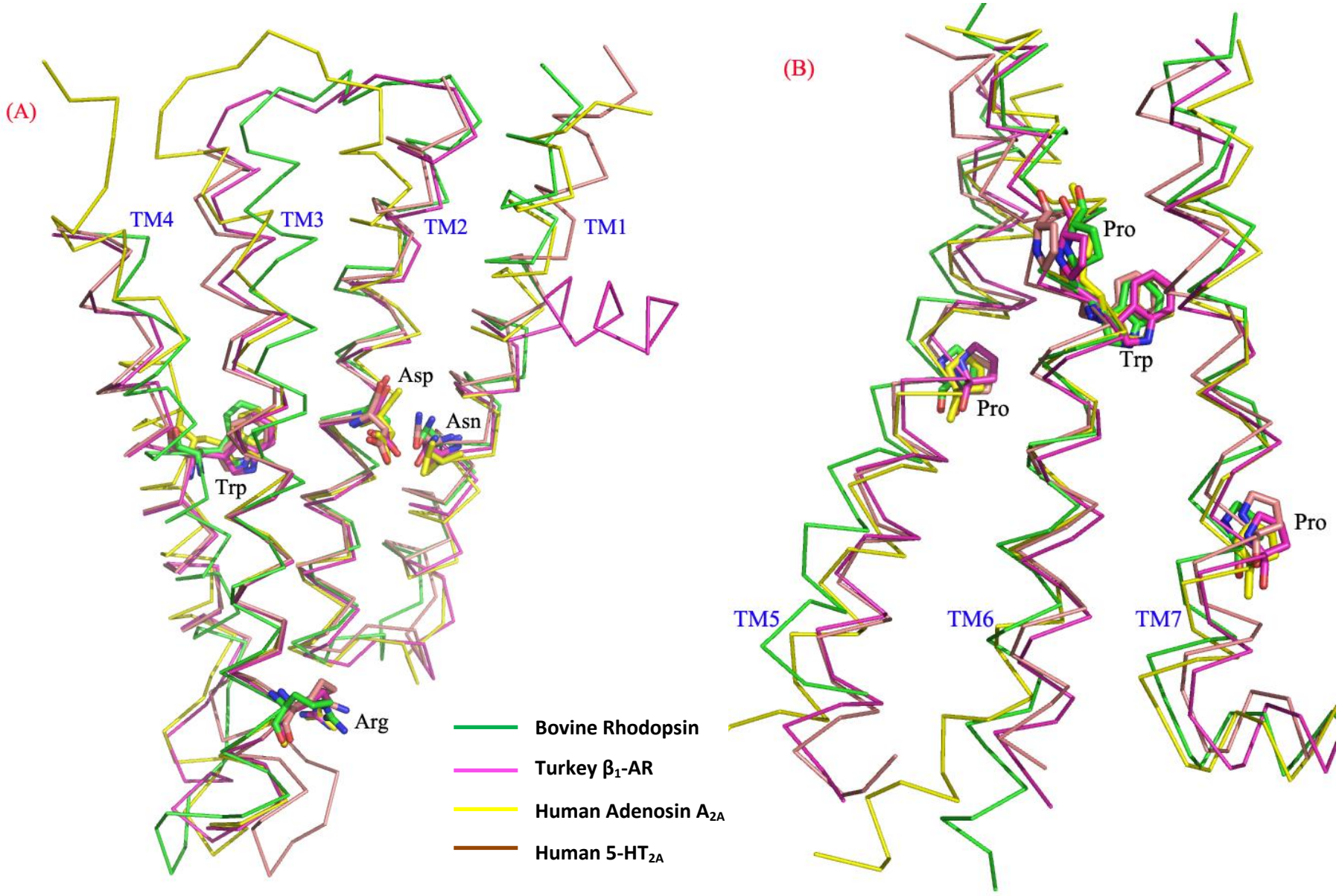
Total number of residues	471	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

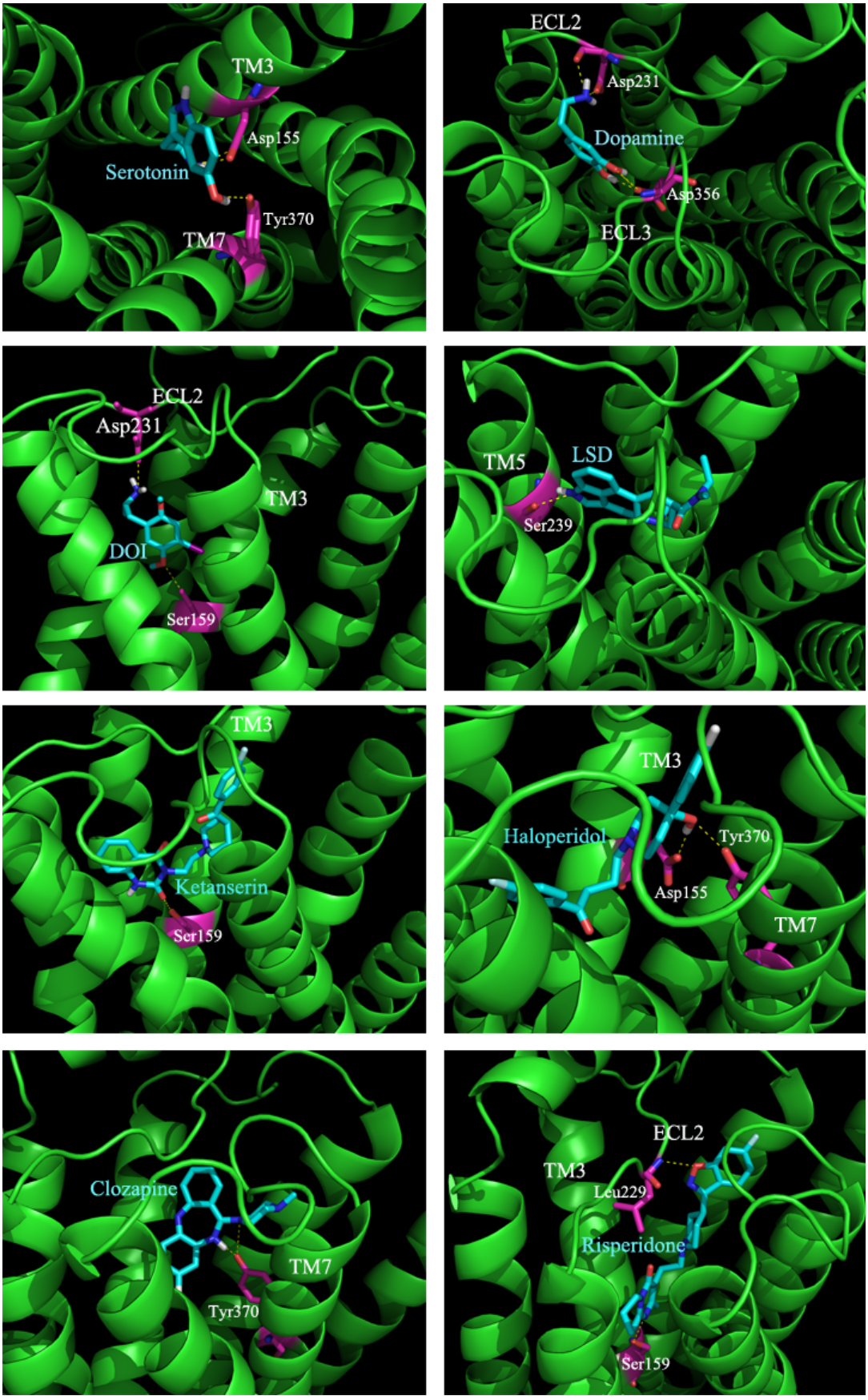
Supplementary figure 9



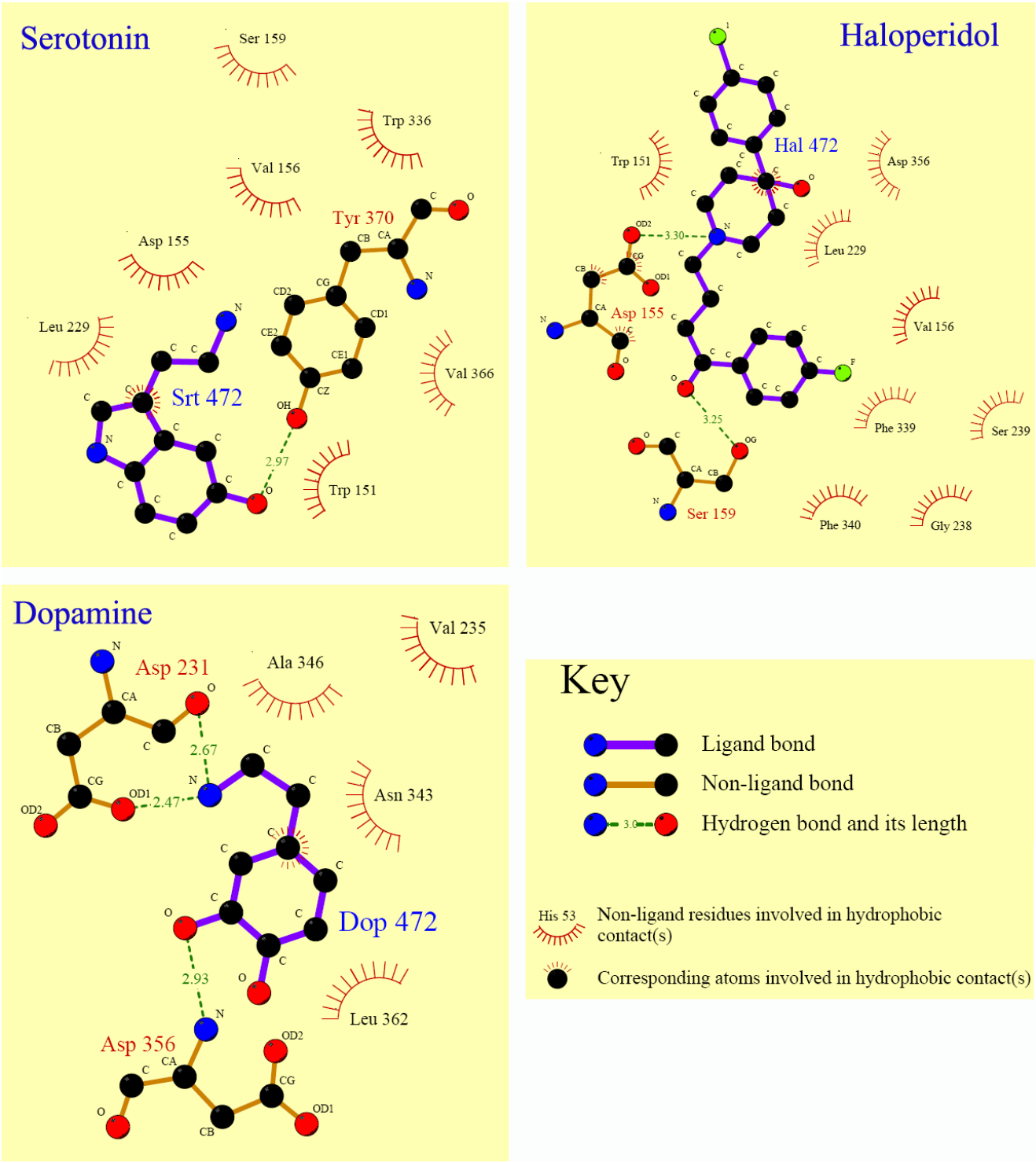
Supplementary Figure 10



Supplementary Figure 11



Supplementary Figure 12



Supplementary Table 1

Structural Region		Residue range	Sequence identity with β_2-AR (2RH1)	Results of Ramachandran plot	Results of Verify3D
Transmembrane regions	TM 1	70-100	23%	Allowed region	+ve region
	TM2	108-137	47%	-do-	-do-
	TM3	144-178	49%	-do-	-do-
	TM4	189-212 (213)	33%	-do-	-do-
	TM5	(233)234-264	32%	-do-	-do-
	TM6	318-348	55%	-do-	-do-
	TM7	359-383	50%	-do-	-do-
	TM8/Small horizontal helix (next to TM7)	385-395 (384^{th} residue is in between 7^{th} and 8^{th} TM region)	27%	-do-	-do-
Intracellular and extracellular loops	N-terminal	1-69	-	Few residues in generously allowed and disallowed regions	Half of residues in –ve region
	ICL1	101-107	29%	Allowed region	+ve region
	ECL1	138-143	20%	-do-	+ve region
	ICL2	179-188	10%	-do-	Negative region
	ECL2	213 -233 (214-232)	19%	-do-	+ve region
	ICL3	265-317	-	Few residues in generously allowed and disallowed regions	Few residues in –ve region
	ECL3	349-359	20%	Allowed region	+ve region
	C-terminal	396-471	-	-do-	+ve region