Supporting Information

Comparative Study of Allosteric GPCR Binding Sites and Their Ligandability Potential

Sonja Peter[1,2], Lydia Siragusa[3,4], Morgan Thomas[1,6], Tommaso Palomba[3], Simon Cross[3], Noel M. O'Boyle[1], Dávid Bajusz[5], György G. Ferenczy[5], György M. Keserű[5], Giovanni Bottegoni[2,7], Brian Bender[1], Ijen Chen[1], Chris De Graaf[1]**

[1]Nxera Pharma UK, Computational Chemistry, Steinmetz Building, Granta Park, Cambridge CB21 6DG, UK. *ijen.chen@nxera.life; *chrisdgrf@gmail.com

[2]Department of Biomolecular Sciences, University of Urbino Carlo Bo, Piazza Rinascimento 6, 61029 Urbino, Italy

[3]Molecular Discovery Ltd., Kinetic Business Centre, Theobald Street, Elstree, Borehamwood, WD6 4PJ Hertfordshire, United Kingdom

[4]Molecular Horizon srl, via Montelino 30, 06084 Bettona (PG), Italy

[5]Medicinal Chemistry Research Group and Drug Innovation Centre, HUN-REN Research Centre for Natural Sciences, Magyar tudósok krt. 2, 1117 Budapest, Hungary

[6]Yusuf Hamied Department of Chemistry, University of Cambridge, CB2 1EW Cambridge, United Kingdom

[7]Institute of Clinical Sciences, University of Birmingham, Edgbaston, B15 2TT Birmingham, United Kingdom

Table S1. Mechanisms of action observed by GPCR modulators (modified from Congreve et al.[4], Fasciani et al.[53] and Grundmann et al.[54])

Orthosteric Agonist	When endogenous or synthetic agonists bind to the primary binding site, they cause a change in receptor conformation, which stabilizes an active state of the receptor.
Positive Allosteric Modulator (PAM)	PAMs bind to an allosteric site on the receptor separate from the endogenous binding pocket, increasing the affinity and/or efficacy of an orthosteric ligand resulting in enhanced receptor activation when an orthosteric ligand is present.
Allosteric Agonist	The binding of an allosteric agonist alone is sufficient to activate the receptor. These compounds also enhance the affinity and/or efficacy of an orthosteric ligand.
Orthosteric Antagonist	When orthosteric antagonists bind to a receptor, it prevents the activation of that receptor.
Negative Allosteric Modulator (NAM)	NAMs bind to an allosteric site on the receptor separate from the endogenous binding pocket, reducing the affinity and/or efficacy of an orthosteric ligand resulting in a decrease in receptor activation when an orthosteric ligand is present.
Allosteric Antagonist	The binding of an allosteric antagonist alone is sufficient to prevent receptor activation. These compounds also decrease the affinity and/or efficacy of an orthosteric ligand.
Biased Allosteric Modulator (BAM)	A biased allosteric modulator is a substance that selectively influences one signaling pathway over others by binding to an allosteric site on a protein, resulting in a biased or specific cellular response.

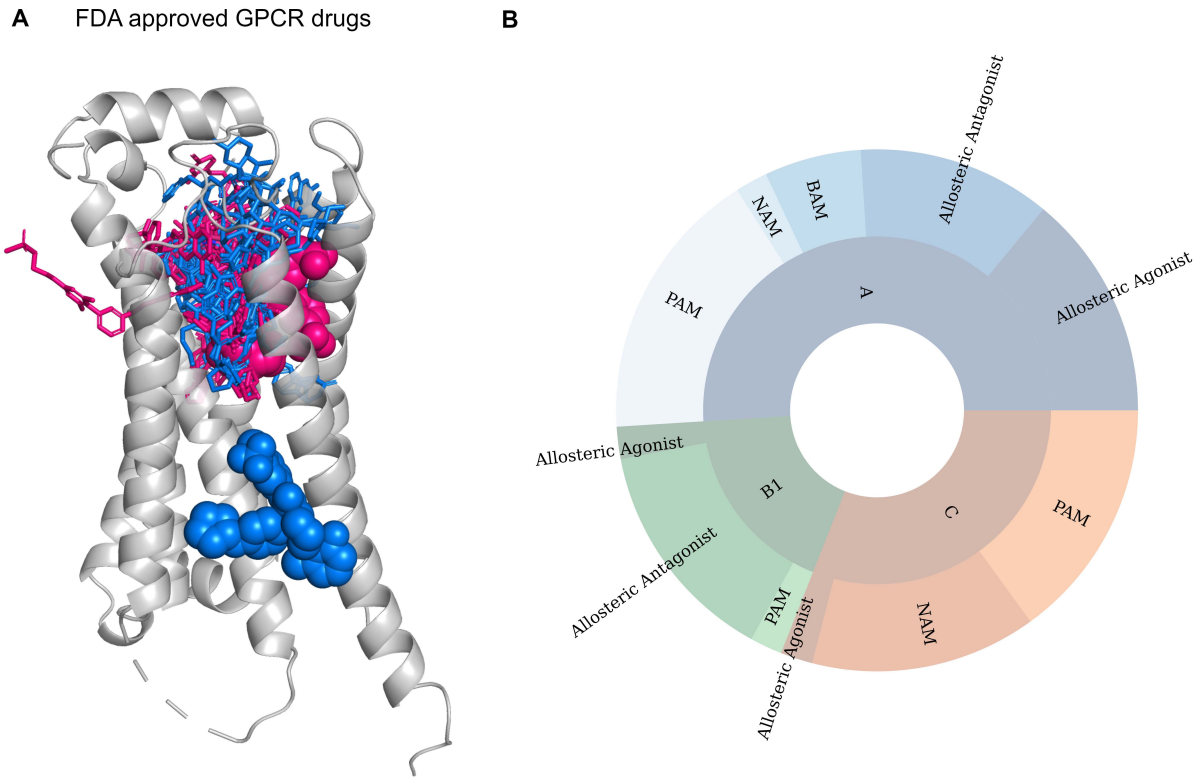


Figure S1. Overview of public structural information on allosteric small molecule bound GPCRs (December 2023). (A) Overlay of all structurally resolved orthosteric (sticks) and allosteric (spheres) FDA approved drugs (N=217). Modulators that activate GPCR activity are pink and those that inhibit GPCR activity are blue. (B) Nested pie chart depicts the allosteric ligand modalities covered by Class A, B1, and C publicly available allosteric PDB dataset. Specifically, ligand-bound structures with Allosteric Agonist (A: 14, B1, 2, C: 2), Allosteric Antagonists (A: 12, B1: 14), Positive Allosteric Modulators (PAM. A: 17, B1: 2, C: 15), Negative Allosteric Modulators (NAM, A: 2, C: 14), and Biased Allosteric Modulators (BAM, A: 6) are known. Refer to Table S1 for the explanation of binding modes.

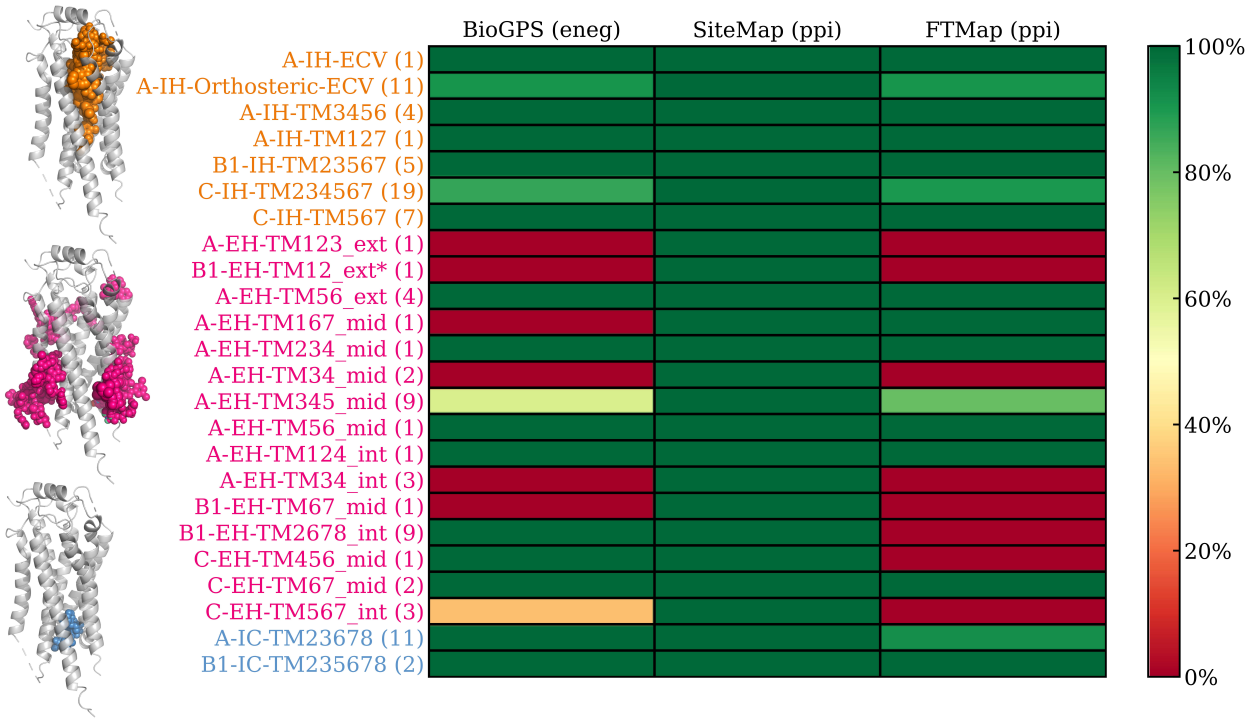


Figure S2. Binding site detection of known allosteric inhibitors for BioGPS, SiteMap, and FTMap. BioGPS was run using energetic probes (eneg) to detect binding sites (2.2b). SiteMap and FTMap were run with the flag that allows for the detection of shallow binding pockets (ppi). The site detection was also counted if only a subpocket was detected. The binding sites are differentiated by their location within the receptor: intrahelical binders are shown in orange spheres, extrahelical modulators are depicted as pink spheres, and intracellular ligands are illustrated as blue spheres. The number following the binding site annotation indicates the number of structures with a known allosteric ligand. The total number is 101, instead of 100 as PDB: 8JD5 has two allosteric ligands binding at C-IH-TM234567 and C-EH-TM67_mid, respectively. Additionally, the image shows the overlay of all the binding sites to the sphingosine 1-phosphate receptor 1 (S1PR1) in a grey ribbon (PDB: 7EO4). All intrahelical allosteric ligands bind at a site distinct from the endogenous ligand. ACM2, ACM4, GPR52, and MRGX1 bind in the extracellular vestibule. The pocket-first approach assigns them to A-IH-Orthosteric-ECV as the pocket obtained by the binding site detection tool extends into the orthosteric site. *Similarly, the allosteric ligand of GLP1R binds between TM1 and TM2; the pocket obtained by the binding site detection tools extends into the orthosteric site.

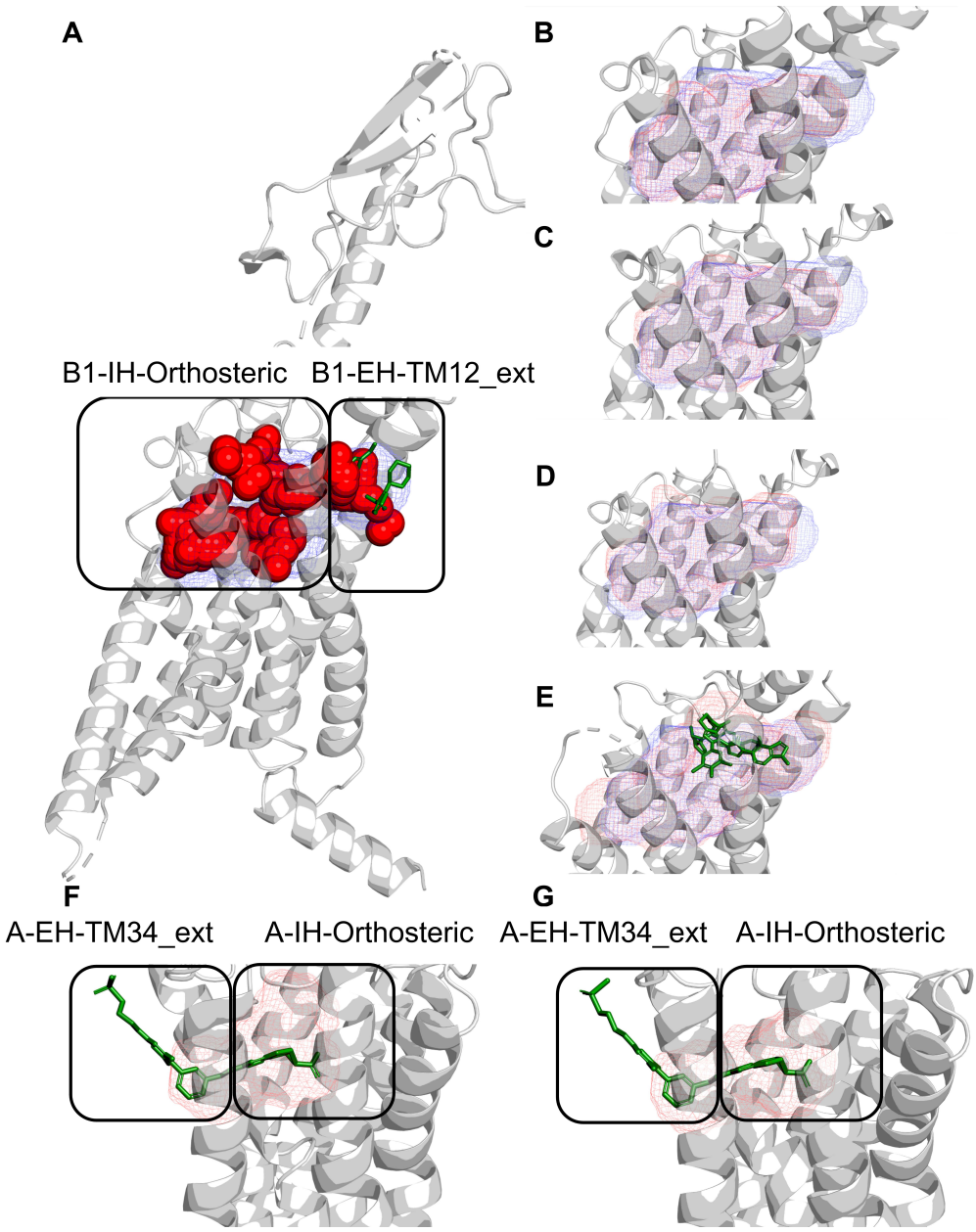


Figure S3. Annotation for GLP1R PAM and FFAR1 orthosteric agonist. (A) The predicted pocket for GLP1R (PDB: 6VCB) by SiteMap is represented as red SitePoints, while BioGPS is depicted as a blue mesh. GLP1R is shown as a gray cartoon with PAM LSN3160440 as green licorice. Both binding site detection tools predict a site that overlaps with the PAM and extends into the B1-IH-Orthosteric site. The ligand-based annotation would classify PAM LSN3160440 as B1-EH-TM12_ext, while the overall pocket would be annotated as B1-IH-Orthosteric. In the case of the PAM binding site of GLP1R, the annotation of B1-EH-TM12_ext* was chosen throughout the manuscript to reflect the allosteric ligand position. Similarly, the pocket detected in (B) active GPCRdb AlphaFold model[20] of (C-D) GLP1R (PDB: 6X18, 7DUQ), and (E) GLP1R (PDB: 6X19) that are shown as red mesh were annotated B1-IH-TM12_ext*. For the ligand property analysis, the orthosteric agonist CHU-128 (green licorice) is annotated as B1-IH-Orthosteric. (F, G) The predicted pockets (red mesh) for the orthosteric agonist MK-8666, depicted as green licorice, bound to FFAR1 (gray cartoon) for PDB: 5TZY and PDB: 5TZR, respectively. The pocket property analysis excludes the pocket volume since the MK-8666 occupies A-IH-Orthosteric and A-EH-TM34_ext. For the ligand property analysis, MK-8666 is annotated as A-IH-Orthosteric.

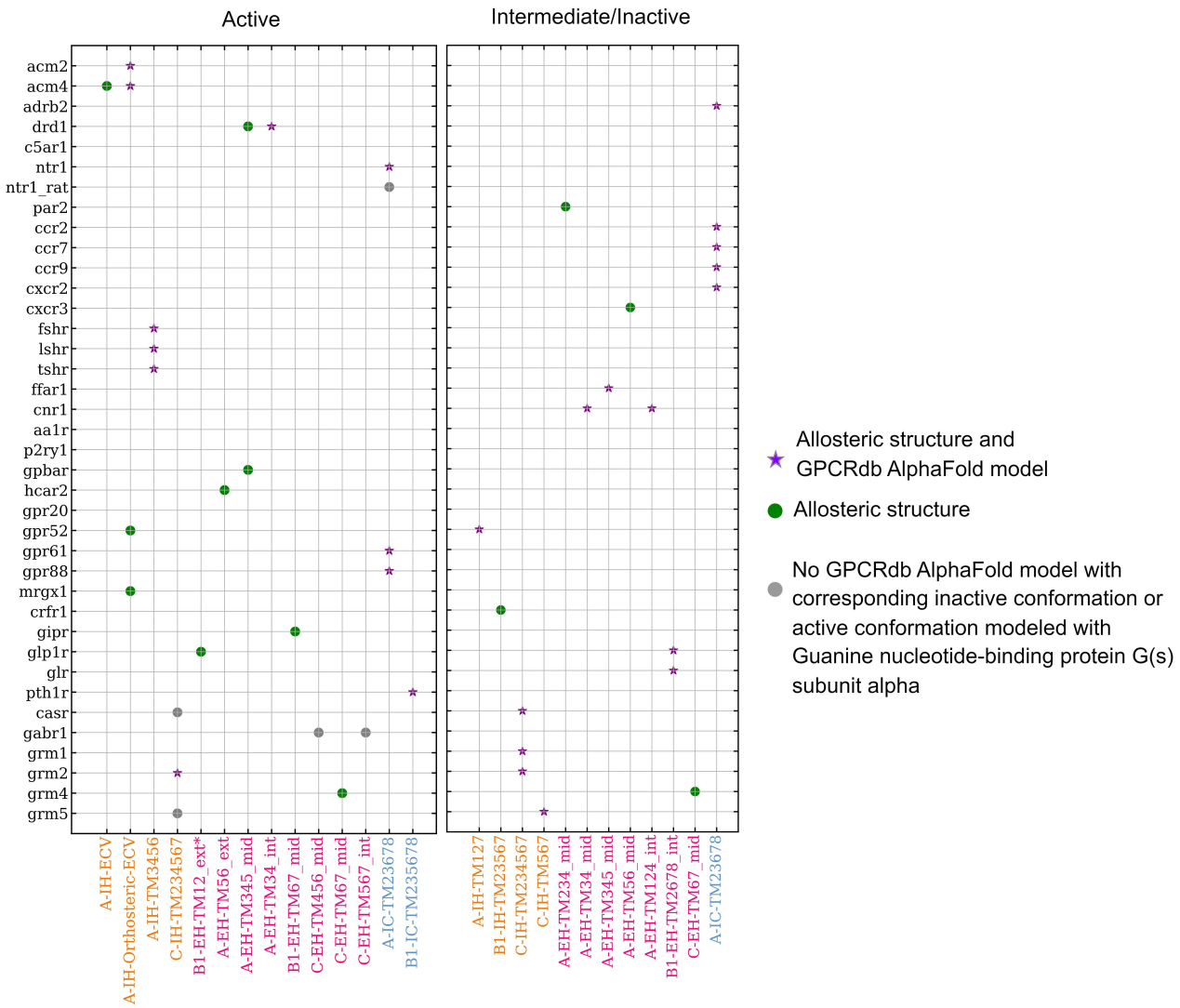


Figure S4. The detection of binding sites in G-protein-coupled receptors (GPCRs) depends on their structural conformation and the allosteric site. If a binding site is detected only in the Allosteric structure, it is represented by a green circle. If it is also found in the GPCRdb AlphaFold models[20], it is represented by a by a purple star. If no GPCRdb AlphaFold model[20] with the corresponding inactive conformation or active conformation modeled with a G(s) is available, the binding site is shown as grey circle.

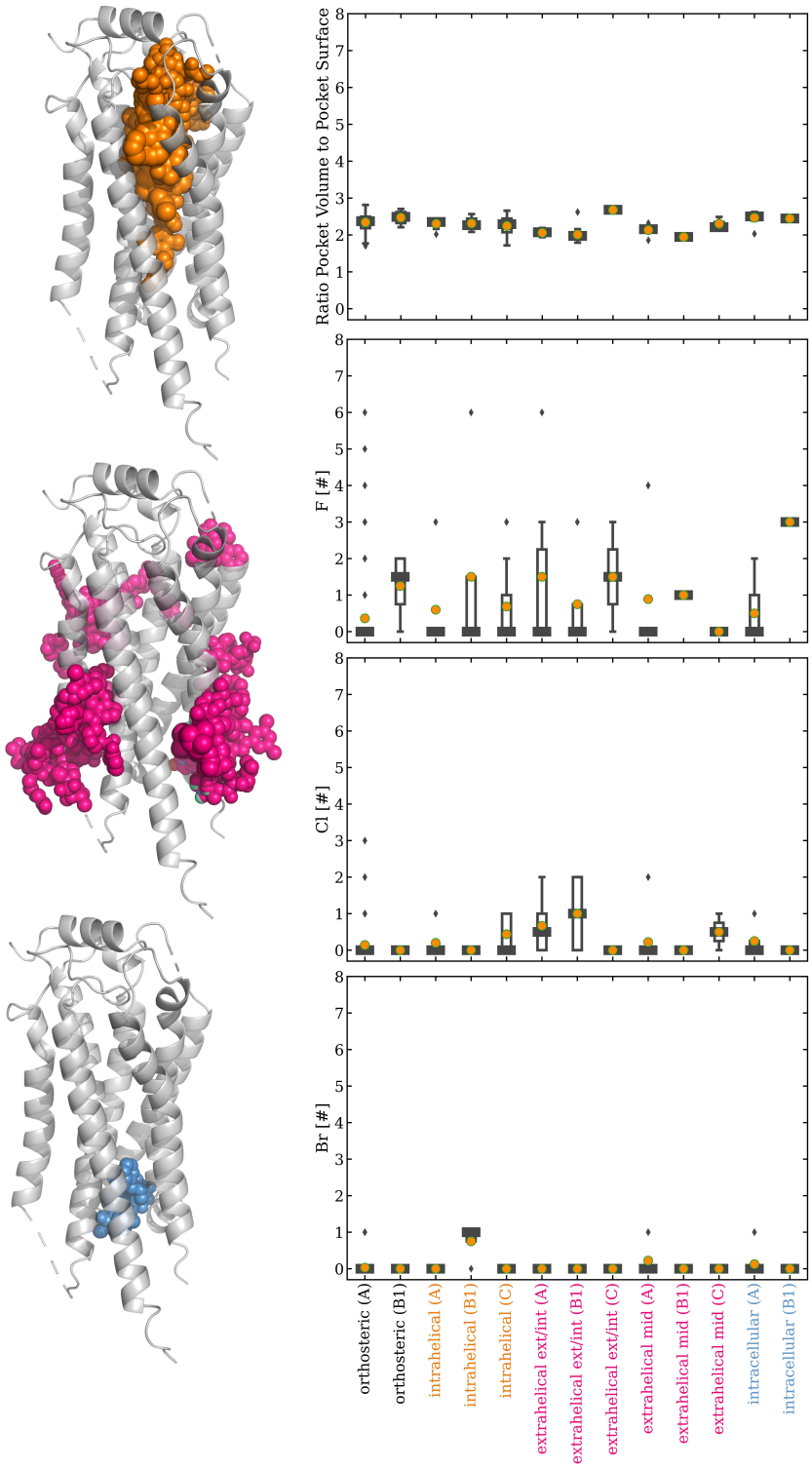


Figure S5. Overview of the ratios of pocket volume to pocket surface and binding site distributions of halogens (F, Cl, and Br) for Class A, B1, and C. For halogen counts, duplicate ligands were removed. A breakdown of the number of ligands for each binding site location is available in Supplementary Data Table S8. For a more detailed Class-specific breakdown of descriptor values, please refer to Supplementary Data Table S6 and Data Table S7.

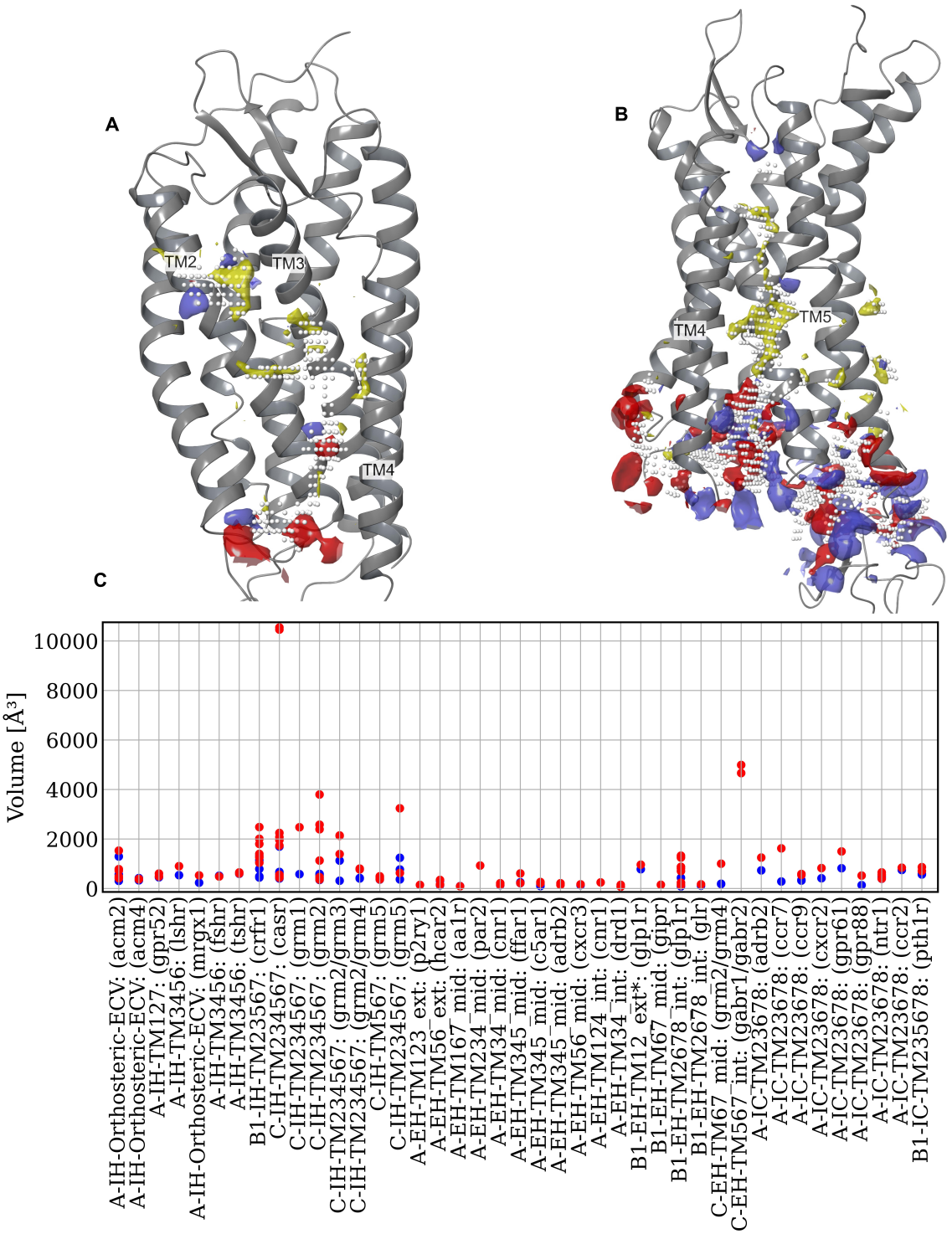


Figure S6. Description of two binding pockets identified by SiteMap (ppi) with default settings allowing shallow binding site detection and comparison of the pocket volume for the same structure and binding site between SiteMap and SiteMap (ppi). White spheres represent the pockets, and the purinergic receptor P2Y (P2Y1R, PDB: 4XNV) is shown as a grey ribbon. (A) The first binding site, A-EH-TM123_ext, extends to the interior A-EH-TM124_int site. (B) The second binding site, A-EH-TM345_mid pocket extends into the intracellular site. (C) Pocket volume determined by SiteMap in blue and SiteMap (ppi) in red of known allosteric binding sites. Especially for pockets in Class C dimer, namely C-IH-TM234567 (CASR) and C-EH-TM567_int (GABR1/2), the pocket volume can be over 4000 Å³.

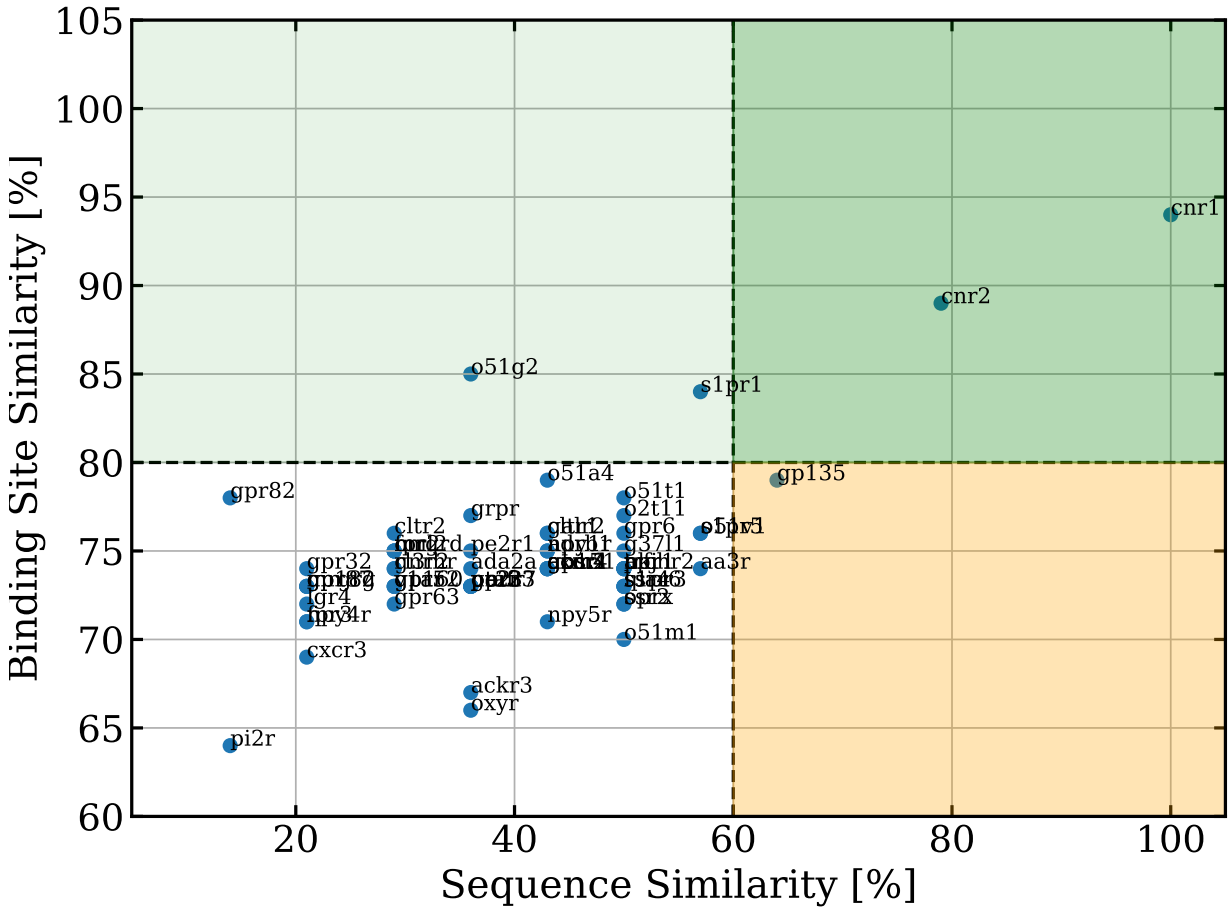


Figure S7. Binding Site versus Sequence similarity for A-EH-TM124_int demonstrating that sequence and binding site similarity correlation. A threshold of 80% was chosen to distinguish between similar binding site and a threshold of 6% was chosen to distinguish between similar sequence. Proteins IDs that are in the white quadrant have neither high binding site similarity, nor do they have high sequence similarity. Protein IDs that are in the green quadrant have both high sequence and binding site similarity. Even though O51G2 has a high binding site similarity, it has a low sequence similarity. For GP135, S1PR5, and AA3R the sequence similarity is high, but the binding site similarity compared to S1PR1 is low.

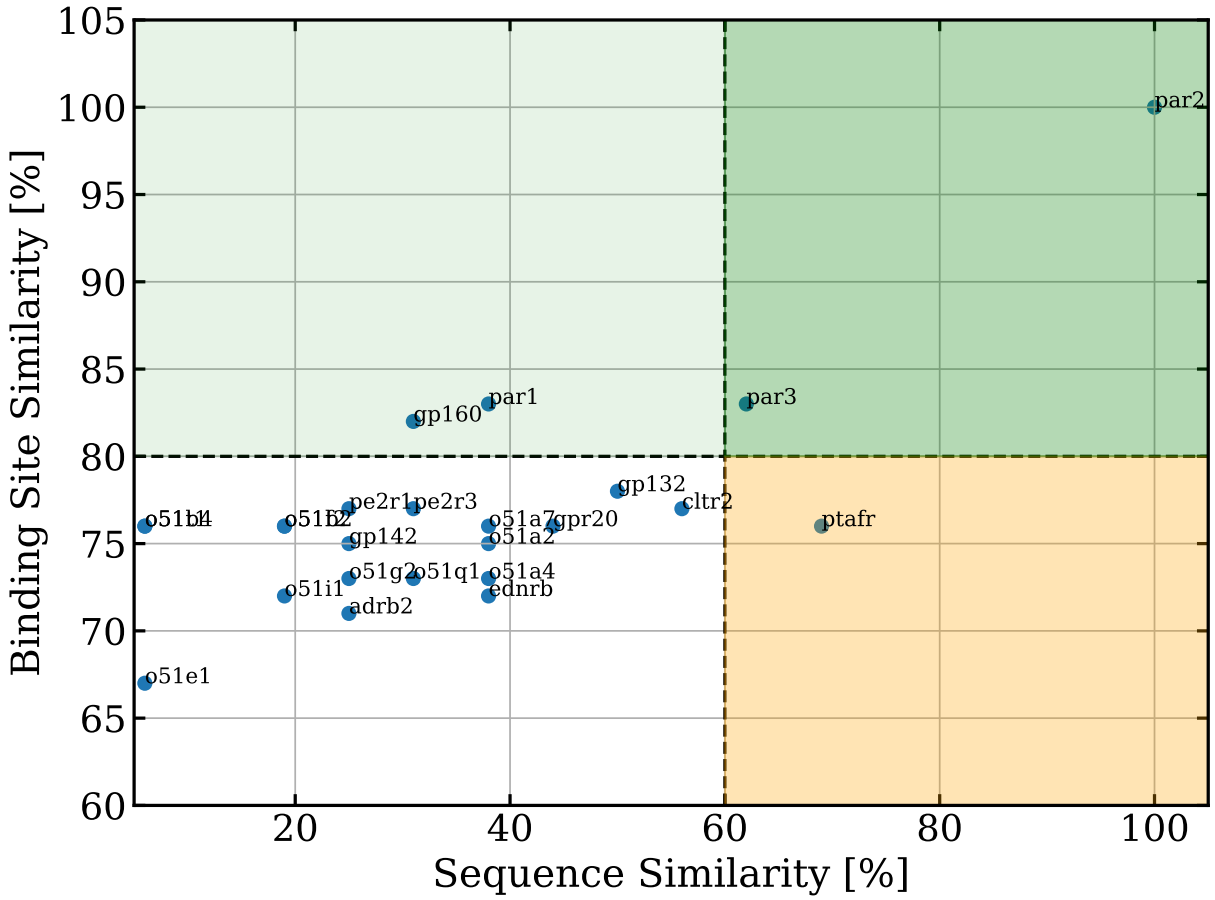
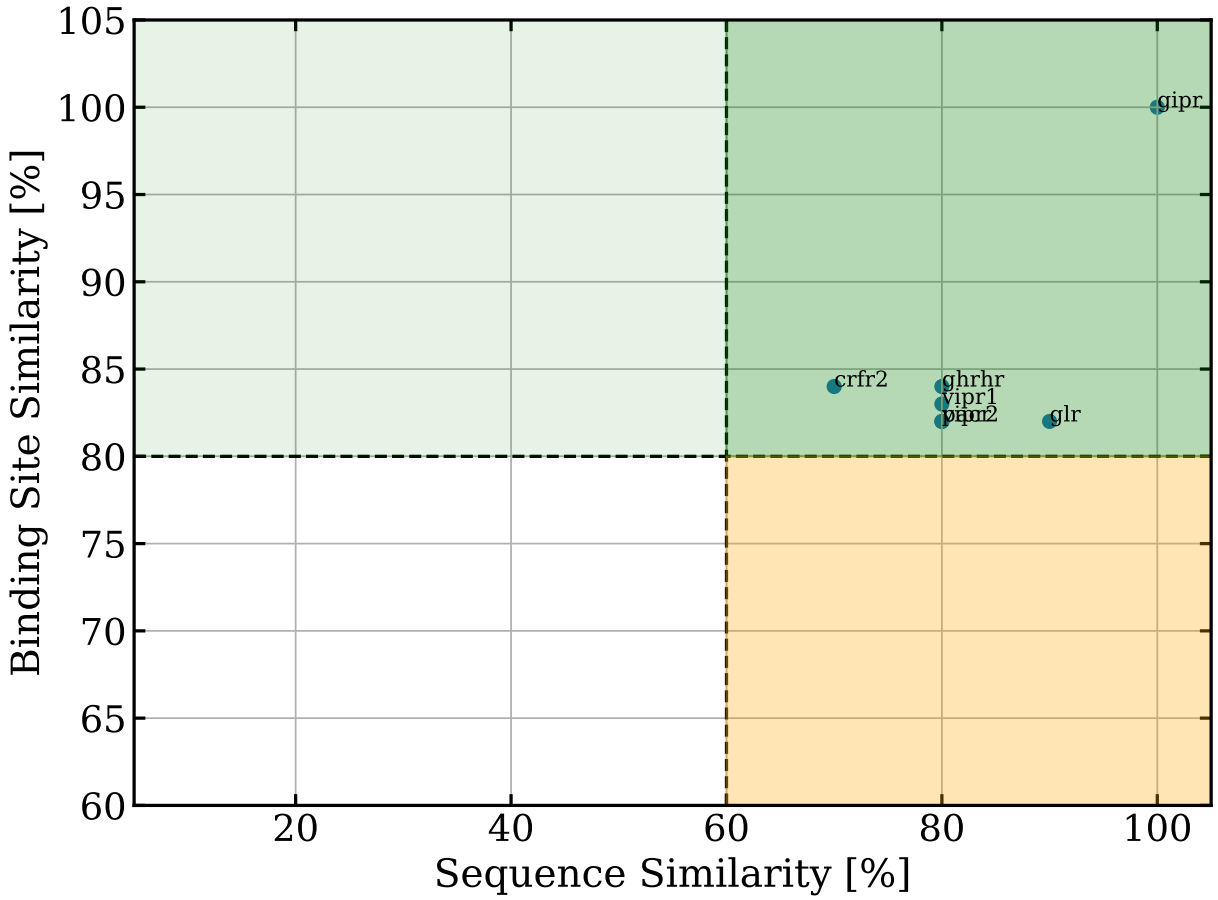


Figure S8. Binding Site versus Sequence similarity for A-EH-TM234_mid demonstrating that sequence and binding site similarity correlation. A threshold of 80% was chosen to distinguish between similar binding site and a threshold of 6% was chosen to distinguish between similar sequence. Proteins IDs that are in the white quadrant have neither high binding site similarity, nor do they have high sequence similarity. Protein IDs that are in the green quadrant have both high sequence and binding site similarity. Even though PAR1 and GP160 have a high binding site similarity, they have a low sequence similarity (colored by light green). For PTAFR the sequence similarity is high, but the binding site similarity compared to PAR3 is low.



100

101 **Figure S9.** Binding Site versus Sequence similarity for B1-EH-TM67_mid demonstrating that sequence
102 and binding site similarity correlation. A threshold of 80% was chosen to distinguish between similar
103 binding site and a threshold of 6% was chosen to distinguish between similar sequence. All protein IDs that
104 had a pocket at B1-EH-TM67_mid had a high binding site and sequence similarity.

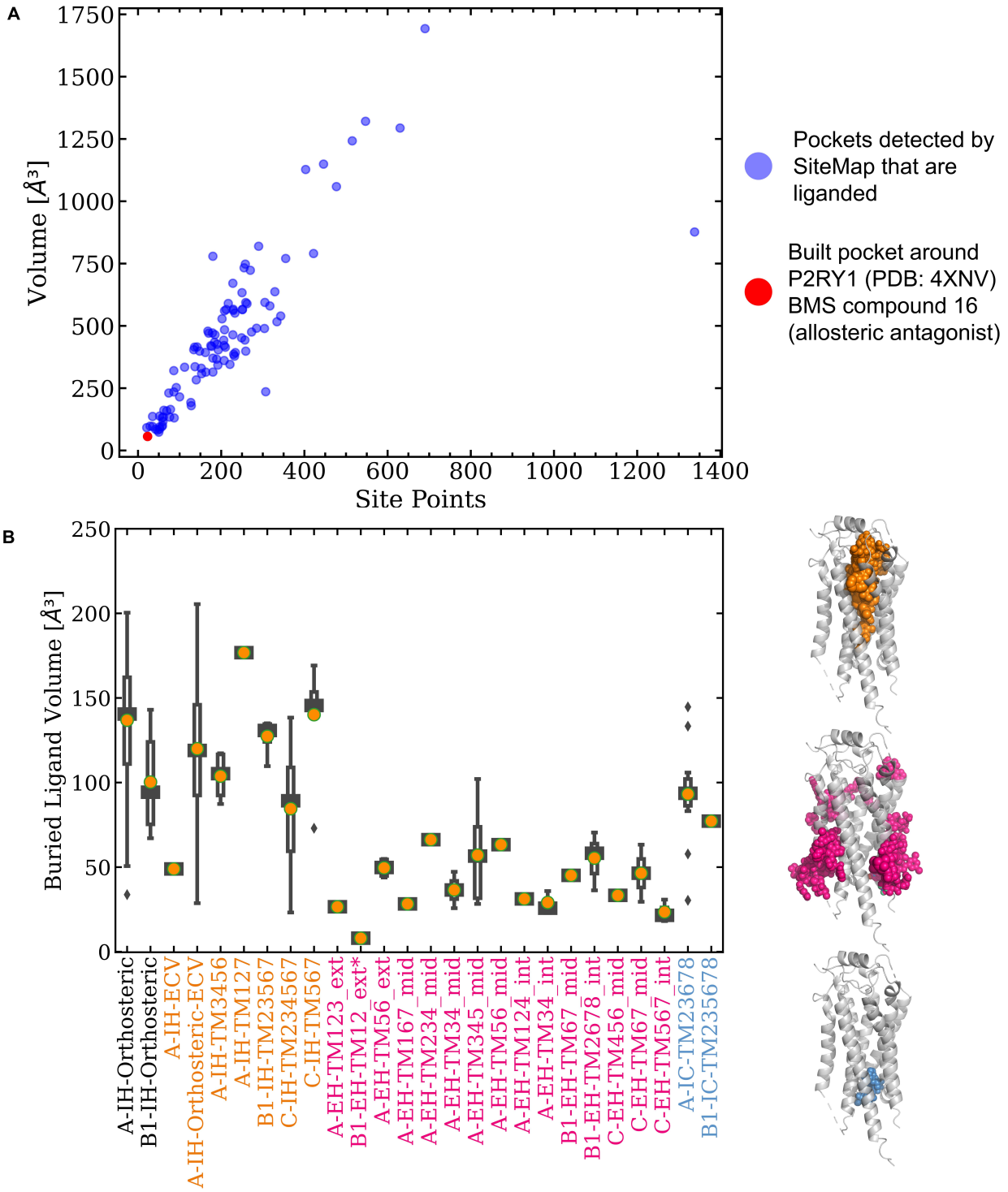


Figure S10. Small and shallow pocket of P2RY1 binding at A-EH-TM123_ext. (A) Volume versus Site Points of P2RY1 pocket build around the allosteric ligand (PDB: 4XNV) in red) and other liganded SiteMap pockets (in blue) demonstrating the shallowness of the P2RY1 pocket. (B) Buried volumes of pockets build around ligands binding to the orthosteric, intrahelical (IH), extrahelical (EH), and intracellular site (IC). The buried volume of A-EH-TM123_ext, where the NAM of P2RY1 binds is with A-EH-TM34_mid and A-EH-TM167_mid with ~25 Å³.

Table S2. Overview of each allosteric ligand, indicating which other GPCR families (as per protein ID) share similar ligands based on a threshold of 0.6 or 0.75 ligand similarity.

Protein ID	Ligand PDB code	0.60	0.75
aa1r	XTD	['ada1a', 'acm2', 'aa1r']	['ada1a', 'acm2', 'aa1r']
acm2	2CU	['grm3', 'gp146', 'ts1r2', 'ada2b', 'fzd5', 'apj', 'cckar', 'lt4r2', 't2r43', 'gpr55', 'pkr1', 'p2ry4', 'gp157', 'gpr35', 'acm4', 'gpr3', '5ht1e', 'pe2r1', 'agrl4', 't2r39', 'grm8', 'mrgre', 'cltr1', '5ht6r', 'rxfp2', 'agrg5', 'lgr4', 't2r10', 'hrh1', 'nmur2', 'gpr84', 'nk1r', 'c5ar2', 'cnr2', 'ada1a', 'npff2', 's1pr5', 'ssr5', 'ur2r', 'adrb1', 'ednra', 'npbw2', 'gpr17', 'ta2r7', 'npy1r', 'mrgx1', 'ffar3', 'sucr1', 'gasr', 'mas1l', 'aa2br', 'ta2r1', 'p2y10', 'tshr', 'cxcr6', 'ssr3', 's1pr3', 'c3ar', 'lpar6', '5ht5a', 'pth1r', 'ccr6', 'agrg3', 't2r16', 'taar1', 'ackr1', 'gp162', 'drd1', 'fzd8', 'o51e1', 'mchr1', 'agrl2', 'ccrl2', 'acm2', 'mtr1b', 'p2ry2', 'pi2r', 'grpr', 't2r45', 'fpr1', 'ntr2', 'par1', 'grm5', 'casr', 't2r38', 'bkrb1', 'gpr45', 'mrgrd', 'agtr1', 'gp156', '5ht1d', 'acm5', 'gpr34', 'gpc5d', 'ccr10', 't2r42', 'fzd4', 'ada2c', 'ts1r3', 'grm2', 'oprk', 'gpr25', 'gpbar', 'ffar2', 'ghrhr', 'gp174', 'npy6r', 's1pr4', 'ssr4', '5ht7r', 'ghsr', 'cxcr1', 'agrf5', 'gpr85', 'ccr1', 'glp1r', 'v2r', 'v1br', 't2r60', 'lpar1', 'taar6', 'lgr5', 'gabr2', 'aa1r', '5ht2a', 'cml2', 'galr1', 'ta2r', 'lshr', 'gpr83', 'ccr7', 'gp119', 'kissr', 'gpr88', 'ox1r', 'ssr2', 's1pr2', 'agra2', 'gpr61', 'ffar4', 'rl3r2', 'gpc6a', 'p2y11', 'oprm', 'pd2r2', 'fzd2', 'gp141', 'grm4', 'gpr52', 'p2ry8', 'agrb1', 'xcr1', 'gpr39', 'opsg', '5ht1b', 'gpr4', 'acm3', 'hcar1', 'gpr32', 'gp150', 'crfr1', 'vipr2', 'fzd9', 'prlhr', 'bai1', 'sctr', 'agrl3', 'qrfpr', 'cxcr2', 'gpr15', 'ccr9', 'adrb3', 'pacr', 'ffar1', 'mrgx3', 'opn3', 'p2y14', 't2r19', 'ta2r5', 'taar5', 'lgr6', 'gnrhr', 'drd5', 'brs3', 'lpar2', 'gpr75', 'hrh3', 'ccr2', 'trfr', 'gpr1', 'gpr37', 'agtr2', 'mrgrg', 't2r50', 'v1ar', 'glp2r', 'bkrb2', 'ogr1', 't2r30', 'gpr26', '5ht4r', 'fzd7', 'npy5r', 'rai3', 'grm1', 'nk3r', 't2r41', 'fshr', 'gp135', 'fzd10', 'oxer1', 'or1g1', 'fpr3', 'opsd', 'gipr', 'agrb2', 'fzd1', 'ox2r', 'gp142', 'grm7', 'par3', 'gpr20', 'vipr1', 'or1a1', 'gp153', 'crfr2', 'ptafr', '5ht1a', 'pd2r', 'gpr31', 'npsr1', 'hcar2', 'mc5r', 'acthr', 'ccr4', 'gpr18', 'lpar4', 'etbr2', 'drd3', 'cml1', 'galr2', 'ackr3', 't2r14', 'gabr1', 'taar3', 'ta2r8', '5ht2b', 'agrg1', 'oprx', 'p2y12', 'opn5', 'taar8', 'ta2r3', 'agra1', 'gpr62', 'rl3r1', 'gp171', 'gpr78', 's1pr1', 'ssr1', 'mc3r', 'cxcr4', 'opsr', 'gp182', 'lpar3', 'calrl', 'c5ar1', 'cnr1', 'gper1', 'hrh2', 'nmur1', 'ccr3', 'gpr87', 't2r13', 'ackr4', 'rxfp1', 'drd4', 'mrgx2', 'ta2r4', 'g37l1', 'ccr8', 'ednrb', 'npbw1', 'cxcr3', 'mshr', 'npff1', 'ada1b', 'gp176', 'adrb2', 'glr', 'mtr1l', 't2r40', 'pkr2', 'npy2r', 'lt4r1', 'gpr27', 't2r31', 'par4', 'psyr', 'ts1r1', 'fzd6', 'ada2a', 'pf2r', 'nk2r', 'cltr2', 'mrgrf', 'pe2r2', 't2r20', '5ht1f', 'npy4r', 'p2ry1', 'mtr1a', 'gp152', 'hcar3', 'gp139', 'gpr6', 'o51e2', 'oprd', 'mchr2', 'pe2r4', 'oxyr', 'agrl1', 'pth2r', 'gp148', 'gp143', 'grm6', 'oxgr1', 'ntr1', 'aa3r', 'calcr', 'mtlr', 'gpr21', 'par2', 'fpr2', 'gp132', 'aa2ar', 'agrb3', 't2r46', 'smo', 'ada1d', 'gp183', 'agrf1', 'gpr12', 'cxcr5', 'gp101', 'p2y13', 'opn4', 'taar9', 'mc4r', 'mrgx4', 'gpr63', 'ackr2', 'mas', 'galr3', 'gp161', 'drd2', 'cx3c1', '5ht2c', 'ta2r9', 'nmbr', 'taar2', 'gpr19', 'ccr5', 'hrh4', 'lpar5']	[]

acm4	IUE	['grm3', 'gp146', 'ts1r2', 'ada2b', 'fzd5', 'apj', 'cckar', 'lt4r2', 't2r43', 'gpr55', 'pkr1', 'p2ry4', 'gp157', 'gpr35', 'acm4', 'gpr3', '5ht1e', 'pe2r1', 'agrl4', 't2r39', 'grm8', 'mrgre', 'cltr1', '5ht6r', 'rxfp2', 'agrg5', 'lgr4', 't2r10', 'hrh1', 'nmur2', 'gpr84', 'nk1r', 'c5ar2', 'cnr2', 'ada1a', 'npff2', 's1pr5', 'ssr5', 'ur2r', 'adrb1', 'ednra', 'npbw2', 'gpr17', 'ta2r7', 'npy1r', 'mrgx1', 'ffar3', 'sucr1', 'gasr', 'mas1l', 'aa2br', 'ta2r1', 'p2y10', 'tshr', 'cxcr6', 'ssr3', 's1pr3', 'c3ar', 'lpar6', '5ht5a', 'pth1r', 'ccr6', 'agrg3', 't2r16', 'taar1', 'ackr1', 'gp162', 'drd1', 'fzd8', 'o51e1', 'mchr1', 'agrl2', 'ccrl2', 'acm2', 'mtr1b', 'p2ry2', 'pi2r', 'grpr', 't2r45', 'fpr1', 'ntr2', 'par1', 'grm5', 'casr', 't2r38', 'bkrb1', 'gpr45', 'mrgrd', 'agtr1', 'gp156', '5ht1d', 'acm5', 'gpr34', 'gpc5d', 'ccr10', 't2r42', 'fzd4', 'ada2c', 'ts1r3', 'grm2', 'oprk', 'gpr25', 'gpbar', 'ffar2', 'ghrhr', 'gp174', 'npy6r', 's1pr4', 'ssr4', '5ht7r', 'ghsr', 'cxcr1', 'agrf5', 'gpr85', 'ccr1', 'glp1r', 'v2r', 'v1br', 't2r60', 'lpar1', 'taar6', 'lgr5', 'gabr2', 'aa1r', '5ht2a', 'cml2', 'galr1', 'ta2r', 'lshr', 'gpr83', 'ccr7', 'gp119', 'kissr', 'gpr88', 'ox1r', 'ssr2', 's1pr2', 'agra2', 'gpr61', 'ffar4', 'rl3r2', 'gpc6a', 'p2y11', 'oprm', 'pd2r2', 'fzd2', 'gp141', 'grm4', 'gpr52', 'p2ry8', 'agrb1', 'xcr1', 'gpr39', 'opsg', '5ht1b', 'gpr4', 'acm3', 'hcar1', 'gpr32', 'gp150', 'crfr1', 'vipr2', 'fzd9', 'prlhr', 'bai1', 'sctr', 'agrl3', 'qrfpr', 'cxcr2', 'gpr15', 'ccr9', 'adrb3', 'pacr', 'ffar1', 'mrgx3', 'opn3', 'p2y14', 't2r19', 'ta2r5', 'taar5', 'lgr6', 'gnrhr', 'drd5', 'brs3', 'lpar2', 'gpr75', 'hrh3', 'ccr2', 'trfr', 'gpr1', 'gpr37', 'agtr2', 'mrgrg', 't2r50', 'v1ar', 'glp2r', 'bkrb2', 'ogr1', 't2r30', 'gpr26', '5ht4r', 'fzd7', 'npy5r', 'rai3', 'grm1', 'nk3r', 't2r41', 'fshr', 'gp135', 'fzd10', 'oxer1', 'or1g1', 'fpr3', 'opsd', 'gipr', 'agrb2', 'fzd1', 'ox2r', 'gp142', 'grm7', 'par3', 'gpr20', 'vipr1', 'or1a1', 'gp153', 'crfr2', 'ptafr', '5ht1a', 'pd2r', 'gpr31', 'npsr1', 'hcar2', 'mc5r', 'acthr', 'ccr4', 'gpr18', 'lpar4', 'etbr2', 'drd3', 'cml1', 'galr2', 'ackr3', 't2r14', 'gabr1', 'taar3', 'ta2r8', '5ht2b', 'agrg1', 'oprx', 'p2y12', 'opn5', 'taar8', 'ta2r3', 'agra1', 'gpr62', 'rl3r1', 'gp171', 'gpr78', 's1pr1', 'ssr1', 'mc3r', 'cxcr4', 'opsr', 'gp182', 'lpar3', 'calrl', 'c5ar1', 'cnr1', 'gper1', 'hrh2', 'nmur1', 'ccr3', 'gpr87', 't2r13', 'ackr4', 'rxfp1', 'drd4', 'mrgx2', 'ta2r4', 'g37l1', 'ccr8', 'ednrb', 'npbw1', 'cxcr3', 'mshr', 'npff1', 'ada1b', 'gp176', 'adrb2', 'glr', 'mtr1l', 't2r40', 'pkr2', 'npy2r', 'lt4r1', 'gpr27', 't2r31', 'par4', 'psyr', 'ts1r1', 'fzd6', 'ada2a', 'pf2r', 'nk2r', 'cltr2', 'mrgrf', 'pe2r2', 't2r20', '5ht1f', 'npy4r', 'p2ry1', 'mtr1a', 'gp152', 'hcar3', 'gp139', 'gpr6', 'o51e2', 'oprd', 'mchr2', 'pe2r4', 'oxyr', 'agrl1', 'pth2r', 'gp148', 'gp143', 'grm6', 'oxgr1', 'ntr1', 'aa3r', 'calcr', 'mtlr', 'gpr21', 'par2', 'fpr2', 'gp132', 'aa2ar', 'agrb3', 't2r46', 'smo', 'ada1d', 'gp183', 'agrf1', 'gpr12', 'cxcr5', 'gp101', 'p2y13', 'opn4', 'taar9', 'mc4r', 'mrgx4', 'gpr63', 'ackr2', 'mas', 'galr3', 'gp161', 'drd2', 'cx3c1', '5ht2c', 'ta2r9', 'nmbr', 'taar2', 'gpr19', 'ccr5', 'hrh4', 'lpar5']	['acm4', 'acm2']
acm4	IUI	['acm4', 'acm2', 'acm5', 'ghsr', 'acm3']	['acm4']
acm4	XNO	['acm4', 'acm2', 'acm5', 'acm3']	['acm4', 'acm2', 'acm5', 'acm3']
adrb2	8VS	['adrb1', 'v2r', 'adrb2']	['adrb1', 'v2r', 'adrb2']
adrb2	KBY	['adrb1', 'v2r', 'adrb2']	['adrb1', 'v2r', 'adrb2']

adrb2	M3J	['taar1', 'adrb2']	[]
c5ar1	9P2	['c5ar1']	['c5ar1']
c5ar1	EFD	['c5ar1']	['c5ar1']
casr	H43	['casr']	['casr']
casr	YP1	['casr', 'adrb2']	['casr']
casr	9IG	['casr']	['casr']
casr	YP4	['casr']	['casr']
ccr2	VT5	['ccr1', 'ccr2']	['ccr1', 'ccr2']
ccr7	JLW	['cxcr1', 'ccr7', 'cxcr2']	['cxcr1', 'ccr7', 'cxcr2']
ccr9	79K	['ccr9', 'ccr2']	['ccr9']
cnr1	9GL	['cnr2', 'cnr1']	['cnr2', 'cnr1']
cnr1	7IC	['npsr1', 'cnr1']	['cnr1']
crfr1	1Q5	['crfr1']	['crfr1']
cxcr2	EBX	['cxcr1', 'cxcr2', 'cxcr3']	[]
cxcr3	43I	['cxcr1', 'cxcr3']	['cxcr3']
drd1	G4C	['drd1', '5ht2b']	['drd1']
ffar1	6XQ	['ffar1']	[]
ffar1	7OS	['ffar1']	['ffar1']
fshr	O6F	['fshr']	[]
gabr1	QDA	[]	[]
gabr1	FN0	['cckar', 'gabr2', 'gabr1']	['cckar', 'gabr2', 'gabr1']
gipr	41Y	['gipr']	['gipr']
glp1r	97Y	['glr']	[]
glp1r	QW7	[]	[]
glp1r	HNO	['glp1r']	['glp1r']
glr	5MV	['glp1r', 'vipr2', 'gipr', 'glr']	['glp1r', 'vipr2', 'gipr', 'glr']
glr	97V	[]	[]
gpbar	FX0	['gpbar']	['gpbar']
gpr52	EN6	['gpr52']	['gpr52']
gpr88	J5F	['oprk', 'gpr88']	['gpr88']
grm1	FM9	['grm5', '5ht2a', 'grm1']	['grm5', '5ht2a', 'grm1']
grm2	HZR	['grm2']	['grm2']
grm2	J9R	['grm2']	['grm2']
grm2	J9U	[]	[]
grm2	ZQY	[]	[]
grm4	BQI	['grm4']	['grm4']
grm4	BK0	['grm5', 'grm4']	[]
grm5	2U8	['grm5']	['grm5']
grm5	51D	['grm5']	[]

grm5	51E	['grm5']	['grm5']
grm5	D7W	['grm5']	['grm5']
grm5	D8B	['grm5', 'grm4', 'grm1']	['grm5']
grm5	4YI	['grm5', 'grm4']	[]
grm5	YKU	['grm5']	['grm5']
hcar2	IX8	['hcar2']	['hcar2']
lshr	55Z	['tshr', 'lshr', 'gnrhr', 'fshr']	['lshr', 'fshr']
mrgx1	U39	['hrh1', 'mrgx1']	['hrh1', 'mrgx1']
ntr1	SRW	['ntr1']	['ntr1']
p2ry1	BUR	['p2ry2', '5ht2a', 'p2y14', 'p2ry1', 'drd2']	['p2ry2', 'p2y14', 'p2ry1']
par2	8UN	[]	[]
par2	Antibody	['crfr1']	[]
tshr	HOI	['tshr', 'oprd']	['tshr']

114

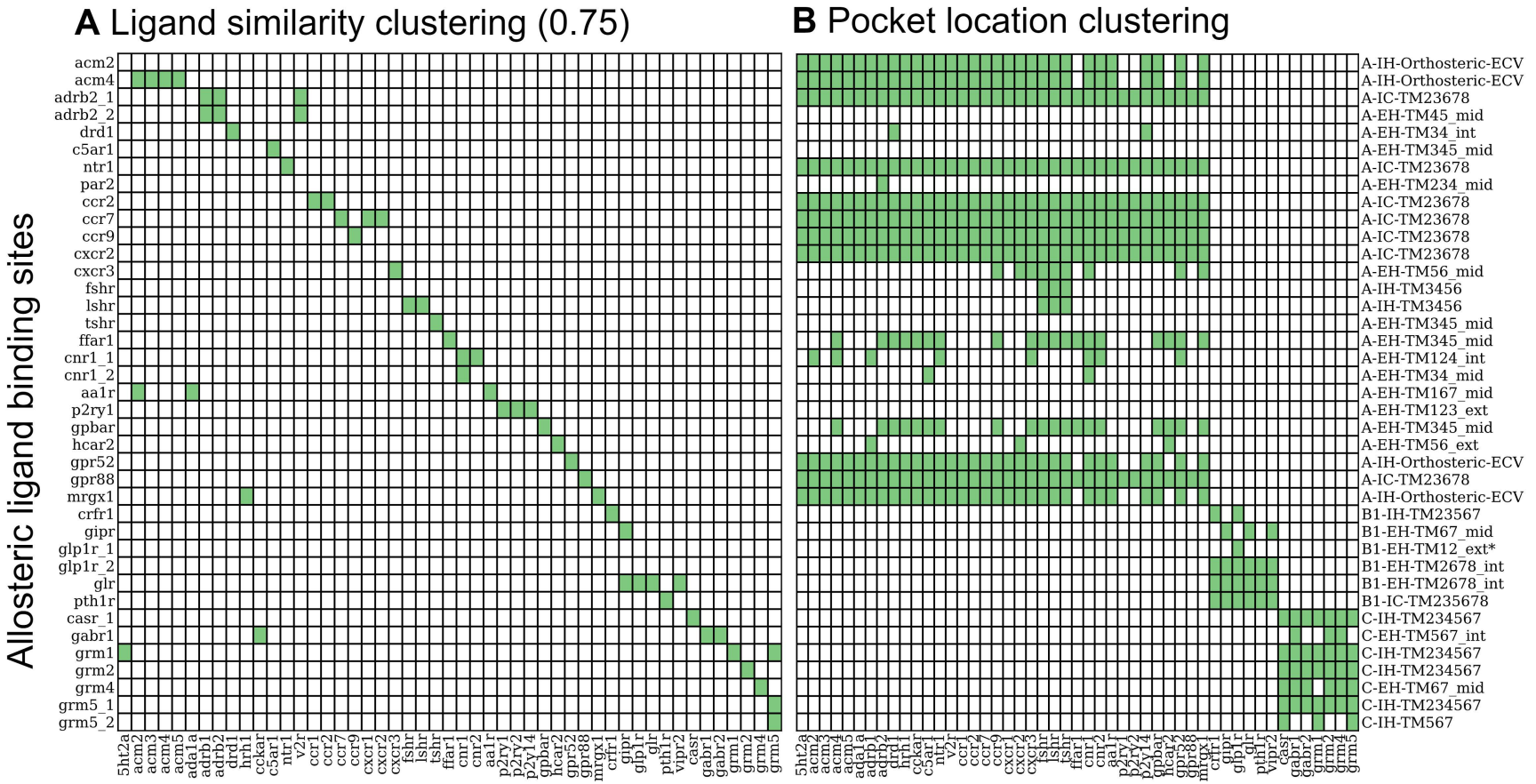


Figure S11. Ligandability analysis using ligand-based and pocket-based approaches (y-axis) and GPCR families that have at least one similar bioactive compound to one of the allosteric small molecule references (x-axis). (A) The heatmap indicates the identification of similar bioactive compounds (pX ≥ 5 for each GPCR listed on the X-axis) to the bound allosteric ligand on the Y-axis shown by the bound GPCR name (similarity ≥ 0.75). Green: found; white: no similar bioactive compounds other than to itself. Please refer to Table S2 for detailed table regarding similar compounds to allosteric ligands found for a threshold of a Tanimoto score 0.60 and 0.75. (B) The heatmap shows if a pocket is detected (in green) for the GPCRs either with known bound allosteric ligands or having similar active compounds as the allosteric GPCR ligand. Sites that were not detected by BioGPS (geo) are colored in white.

References

(1) Hauser, A. S.; Attwood, M. M.; Rask-Andersen, M.; Schiöth, H. B.; Gloriam, D. E. Trends in GPCR Drug Discovery: New Agents, Targets and Indications. *Nat Rev Drug Discov* **2017**, *16* (12), 829–842. https://doi.org/10.1038/nrd.2017.178.

(2) Congreve, M.; Graaf, C. de; Swain, N. A.; Tate, C. G. Impact of GPCR Structures on Drug Discovery. *Cell* **2020**, *181* (1), 81–91. https://doi.org/10.1016/j.cell.2020.03.003.

(3) Graaf, C. de; Song, G.; Cao, C.; Zhao, Q.; Wang, M.-W.; Wu, B.; Stevens, R. C. Extending the Structural View of Class B GPCRs. *Trends Biochem Sci* **2017**, *42* (12), 946–960. https://doi.org/10.1016/j.tibs.2017.10.003.

(4) Congreve, M.; Oswald, C.; Marshall, F. H. Applying Structure-Based Drug Design Approaches to Allosteric Modulators of GPCRs. *Trends Pharmacol Sci* **2017**, *38* (9), 837–847. https://doi.org/10.1016/j.tips.2017.05.010.

(5) Hedderich, J. B.; Persechino, M.; Becker, K.; Heydenreich, F. M.; Gutermuth, T.; Bouvier, M.; Bünemann, M.; Kolb, P. The Pocketome of G-Protein-Coupled Receptors Reveals Previously Untargeted Allosteric Sites. *Nat Commun* **2022**, *13* (1), 2567. https://doi.org/10.1038/s41467-022-29609-6.

(6) Kooistra, A. J.; Munk, C.; Hauser, A. S.; Gloriam, D. E. An Online GPCR Structure Analysis Platform. *Nat. Struct. Mol. Biol.* **2021**, *28* (11), 875–878. https://doi.org/10.1038/s41594-021-00675-6.

(7) Kooistra, A. J.; Mordalski, S.; Pándy-Szekeres, G.; Esguerra, M.; Mamyrbekov, A.; Munk, C.; Keserű, G. M.; Gloriam, D. E. GPCRdb in 2021: Integrating GPCR Sequence, Structure and Function. *Nucleic Acids Res.* **2020**, *49* (D1), gkaa1080-. https://doi.org/10.1093/nar/gkaa1080.

(8) Desaphy, J.; Bret, G.; Rognan, D.; Kellenberger, E. Sc-PDB: A 3D-Database of Ligandable Binding Sites—10 Years On. *Nucleic Acids Res.* **2015**, *43* (D1), D399–D404. https://doi.org/10.1093/nar/gku928.

(9) Stank, A.; Kokh, D. B.; Fuller, J. C.; Wade, R. C. Protein Binding Pocket Dynamics. *Acc. Chem. Res.* **2016**, *49* (5), 809–815. https://doi.org/10.1021/acs.accounts.5b00516.

(10) Ballante, F.; Kooistra, A. J.; Kampen, S.; Graaf, C. de; Carlsson, J. Structure-Based Virtual Screening for Ligands of G Protein–Coupled Receptors: What Can Molecular Docking Do for You? *Pharmacol. Rev.* **2021**, *73* (4), 527–565. https://doi.org/10.1124/pharmrev.120.000246.

(11) Wakefield, A. E.; Bajusz, D.; Kozakov, D.; Keserű, G. M.; Vajda, S. Conservation of Allosteric Ligand Binding Sites in G-Protein Coupled Receptors. *J Chem Inf Model* **2022**, *62* (20), 4937–4954. https://doi.org/10.1021/acs.jcim.2c00209.

(12) Brenke, R.; Kozakov, D.; Chuang, G.-Y.; Beglov, D.; Hall, D.; Landon, M. R.; Mattos, C.; Vajda, S. Fragment-Based Identification of Druggable 'Hot Spots' of Proteins Using Fourier Domain Correlation Techniques. *Bioinformatics* **2009**, *25* (5), 621–627. https://doi.org/10.1093/bioinformatics/btp036.

(13) Baroni, M.; Cruciani, G.; Sciabola, S.; Perruccio, F.; Mason, J. S. A Common Reference Framework for Analyzing/Comparing Proteins and Ligands. Fingerprints for Ligands And Proteins (FLAP): Theory and Application. *J Chem Inf Model* **2007**, *47* (2), 279–294. https://doi.org/10.1021/ci600253e.

(14) Cross, S.; Baroni, M.; Carosati, E.; Benedetti, P.; Clementi, S. FLAP: GRID Molecular Interaction Fields in Virtual Screening. Validation Using the DUD Data Set. *J Chem Inf Model* **2010**, *50* (8), 1442–1450. https://doi.org/10.1021/ci100221g.

(15) Siragusa, L.; Cross, S.; Baroni, M.; Goracci, L.; Cruciani, G. BioGPS: Navigating Biological Space to Predict Polypharmacology, Off-targeting, and Selectivity. *Proteins Struct Funct Bioinform* **2015**, *83* (3), 517–532. https://doi.org/10.1002/prot.24753.

(16) Siragusa, L.; Spyrakis, F.; Goracci, L.; Cross, S.; Cruciani, G. BioGPS: The Music for the Chemo- and Bioinformatics Walzer. *Mol. Inform.* **2014**, *33* (6-7), 446–453. https://doi.org/10.1002/minf.201400028.

(17) Vass, M.; Kooistra, A. J.; Yang, D.; Stevens, R. C.; Wang, M.-W.; Graaf, C. de. Chemical Diversity in the G Protein-Coupled Receptor Superfamily. *Trends Pharmacol Sci* **2018**, *39* (5), 494–512. https://doi.org/10.1016/j.tips.2018.02.004.

(18) Berman, H. M.; Bhat, T. N.; Bourne, P. E.; Feng, Z.; Gilliland, G.; Weissig, H.; Westbrook, J. The Protein Data Bank and the Challenge of Structural Genomics. *Nat Struct Biol* **2000**, *7* (Suppl 11), 957–959. https://doi.org/10.1038/80734.

(19) Pándy-Szekeres, G.; Munk, C.; Tsonkov, T. M.; Mordalski, S.; Harpsøe, K.; Hauser, A. S.; Bojarski, A. J.; Gloriam, D. E. GPCRdb in 2018: Adding GPCR Structure Models and Ligands. *Nucleic Acids Res.* **2017**, *46* (Database issue), gkx1109-. https://doi.org/10.1093/nar/gkx1109.

(20) Pándy-Szekeres, G.; Caroli, J.; Mamyrbekov, A.; Kermani, A. A.; Keserű, G. M.; Kooistra, A. J.; Gloriam, D. E. GPCRdb in 2023: State-Specific Structure Models Using AlphaFold2 and New Ligand Resources. *Nucleic Acids Res.* **2022**, *51* (D1), D395–D402. https://doi.org/10.1093/nar/gkac1013.

(21) Halgren, T. A. Identifying and Characterizing Binding Sites and Assessing Druggability. *J Chem Inf Model* **2009**, *49* (2), 377–389. https://doi.org/10.1021/ci800324m.

(22) Tubert-Brohman, I.; Sherman, W.; Repasky, M.; Beuming, T. Improved Docking of Polypeptides with Glide. *J Chem Inf Model* **2013**, *53* (7), 1689–1699. https://doi.org/10.1021/ci400128m.

(23) Mason, J. S.; Bortolato, A.; Weiss, D. R.; Deflorian, F.; Tehan, B.; Marshall, F. H. High End GPCR Design: Crafted Ligand Design and Druggability Analysis Using Protein Structure, Lipophilic Hotspots and Explicit Water Networks. *Silico Pharmacol.* **2013**, *1* (1), 23. https://doi.org/10.1186/2193-9616-1-23.

(24) Bortolato, A.; Tehan, B. G.; Bodnarchuk, M. S.; Essex, J. W.; Mason, J. S. Water Network Perturbation in Ligand Binding: Adenosine A2A Antagonists as a Case Study. *J. Chem. Inf. Model.* **2013**, *53* (7), 1700–1713. https://doi.org/10.1021/ci4001458.

(25) Lomize, M. A.; Pogozheva, I. D.; Joo, H.; Mosberg, H. I.; Lomize, A. L. OPM Database and PPM Web Server: Resources for Positioning of Proteins in Membranes. *Nucleic Acids Res.* **2012**, *40* (D1), D370–D376. https://doi.org/10.1093/nar/gkr703.

(26) Lomize, M. A.; Lomize, A. L.; Pogozheva, I. D.; Mosberg, H. I. OPM: Orientations of Proteins in Membranes Database. *Bioinformatics* **2006**, *22* (5), 623–625. https://doi.org/10.1093/bioinformatics/btk023.

(27) Harding, S. D.; Armstrong, J. F.; Faccenda, E.; Southan, C.; Alexander, S. P. H.; Davenport, A. P.; Spedding, M.; Davies, J. A. The IUPHAR/BPS Guide to PHARMACOLOGY in 2024. *Nucleic Acids Res.* **2023**, gkad944. https://doi.org/10.1093/nar/gkad944.

(28) Davies, M.; Nowotka, M.; Papadatos, G.; Dedman, N.; Gaulton, A.; Atkinson, F.; Bellis, L.; Overington, J. P. ChEMBL Web Services: Streamlining Access to Drug Discovery Data and Utilities. *Nucleic Acids Res.* **2015**, *43* (W1), W612–W620. https://doi.org/10.1093/nar/gkv352.

(29) Zdrazil, B.; Felix, E.; Hunter, F.; Manners, E. J.; Blackshaw, J.; Corbett, S.; Veij, M. de; Ioannidis, H.; Lopez, D. M.; Mosquera, J. F.; Magarinos, M. P.; Bosc, N.; Arcila, R.; Kizilören, T.; Gaulton, A.; Bento, A. P.; Adasme, M. F.; Monecke, P.; Landrum, G. A.; Leach, A. R. The ChEMBL Database in 2023: A Drug Discovery Platform Spanning Multiple Bioactivity Data Types and Time Periods. *Nucleic Acids Res.* **2023**, gkad1004. https://doi.org/10.1093/nar/gkad1004.

(30) Landrum. *RDKit: Open-Source Cheminformatics*; 2022.

(31) Møller, T. C.; Moreno-Delgado, D.; Pin, J.-P.; Kniazeff, J. Class C G Protein-Coupled Receptors: Reviving Old Couples with New Partners. *Biophys. Rep.* **2017**, *3* (4–6), 57–63. https://doi.org/10.1007/s41048-017-0036-9.

(32) Li, X. X.; Lee, J. D.; Massey, N. L.; Guan, C.; Robertson, A. A. B.; Clark, R. J.; Woodruff, T. M. Pharmacological Characterisation of Small Molecule C5aR1 Inhibitors in Human Cells Reveals Biased Activities for Signalling and Function. *Biochem. Pharmacol.* **2020**, *180*, 114156. https://doi.org/10.1016/j.bcp.2020.114156.

(33) Leach, K.; Wen, A.; Cook, A. E.; Sexton, P. M.; Conigrave, A. D.; Christopoulos, A. Impact of Clinically Relevant Mutations on the Pharmacoregulation and Signaling Bias of the

Calcium-Sensing Receptor by Positive and Negative Allosteric Modulators. *Endocrinology* **2013**, *154* (3), 1105–1116. https://doi.org/10.1210/en.2012-1887.

(34) Nemeth, E. F.; Heaton, W. H.; Miller, M.; Fox, J.; Balandrin, M. F.; Wagenen, B. C. V.; Colloton, M.; Karbon, W.; Scherrer, J.; Shatzen, E.; Rishton, G.; Scully, S.; Qi, M.; Harris, R.; Lacey, D.; Martin, D. Pharmacodynamics of the Type II Calcimimetic Compound Cinacalcet HCl. *J. Pharmacol. Exp. Ther.* **2004**, *308* (2), 627–635. https://doi.org/10.1124/jpet.103.057273.

(35) Davey, A. E.; Leach, K.; Valant, C.; Conigrave, A. D.; Sexton, P. M.; Christopoulos, A. Positive and Negative Allosteric Modulators Promote Biased Signaling at the Calcium-Sensing Receptor. *Endocrinology* **2012**, *153* (3), 1232–1241. https://doi.org/10.1210/en.2011-1426.

(36) Leach, K.; Gregory, K. J.; Kufareva, I.; Khajehali, E.; Cook, A. E.; Abagyan, R.; Conigrave, A. D.; Sexton, P. M.; Christopoulos, A. Towards a Structural Understanding of Allosteric Drugs at the Human Calcium-Sensing Receptor. *Cell Res.* **2016**, *26* (5), 574–592. https://doi.org/10.1038/cr.2016.36.

(37) Zhuang, Y.; Krumm, B.; Zhang, H.; Zhou, X. E.; Wang, Y.; Huang, X.-P.; Liu, Y.; Cheng, X.; Jiang, Y.; Jiang, H.; Zhang, C.; Yi, W.; Roth, B. L.; Zhang, Y.; Xu, H. E. Mechanism of Dopamine Binding and Allosteric Modulation of the Human D1 Dopamine Receptor. *Cell Res.* **2021**, *31* (5), 593–596. https://doi.org/10.1038/s41422-021-00482-0.

(38) Müller, K.; Faeh, C.; Diederich, F. Fluorine in Pharmaceuticals: Looking Beyond Intuition. *Science* **2007**, *317* (5846), 1881–1886. https://doi.org/10.1126/science.1131943.

(39) Meanwell, N. A. Fluorine and Fluorinated Motifs in the Design and Application of Bioisosteres for Drug Design. *J. Med. Chem.* **2018**, *61* (14), 5822–5880. https://doi.org/10.1021/acs.jmedchem.7b01788.

(40) Price, M. R.; Baillie, G. L.; Thomas, A.; Stevenson, L. A.; Easson, M.; Goodwin, R.; McLean, A.; McIntosh, L.; Goodwin, G.; Walker, G.; Westwood, P.; Marrs, J.; Thomson, F.; Cowley, P.; Christopoulos, A.; Pertwee, R. G.; Ross, R. A. Allosteric Modulation of the Cannabinoid CB1 Receptor. *Mol. Pharmacol.* **2005**, *68* (5), 1484–1495. https://doi.org/10.1124/mol.105.016162.

(41) Shao, Z.; Yan, W.; Chapman, K.; Ramesh, K.; Ferrell, A. J.; Yin, J.; Wang, X.; Xu, Q.; Rosenbaum, D. M. Structure of an Allosteric Modulator Bound to the CB1 Cannabinoid Receptor. *Nat. Chem. Biol.* **2019**, *15* (12), 1199–1205. https://doi.org/10.1038/s41589-019-0387-2.

(42) Cheng, R. K. Y.; Fiez-Vandal, C.; Schlenker, O.; Edman, K.; Aggeler, B.; Brown, D. G.; Brown, G. A.; Cooke, R. M.; Dumelin, C. E.; Doré, A. S.; Geschwindner, S.; Grebner, C.; Hermansson, N.-O.; Jazayeri, A.; Johansson, P.; Leong, L.; Prihandoko, R.; Rappas, M.; Soutter, H.; Snijder, A.; Sundström, L.; Tehan, B.; Thornton, P.; Troast, D.; Wiggin, G.; Zhukov, A.; Marshall, F. H.; Dekker, N. Structural Insight into Allosteric Modulation of Protease-Activated Receptor 2. *Nature* **2017**, *545* (7652), 112–115. https://doi.org/10.1038/nature22309.

(43) Jiao, H.; Pang, B.; Liu, A.; Chen, Q.; Pan, Q.; Wang, X.; Xu, Y.; Chiang, Y.-C.; Ren, R.; Hu, H. Structural Insights into the Activation and Inhibition of CXC Chemokine Receptor 3. *Nat. Struct. Mol. Biol.* **2024**, 1–11. https://doi.org/10.1038/s41594-023-01175-5.

(44) Wang, X.; Wang, M.; Xu, T.; Feng, Y.; Shao, Q.; Han, S.; Chu, X.; Xu, Y.; Lin, S.; Zhao, Q.; Wu, B. Structural Insights into Dimerization and Activation of the MGlu2–MGlu3 and MGlu2–MGlu4 Heterodimers. *Cell Res.* **2023**, *33* (10), 762–774. https://doi.org/10.1038/s41422-023-00830-2.

(45) Wakefield, A. E.; Mason, J. S.; Vajda, S.; Keserű, G. M. Analysis of Tractable Allosteric Sites in G Protein-Coupled Receptors. *Sci. Rep.* **2019**, *9* (1), 6180. https://doi.org/10.1038/s41598-019-42618-8.

(46) Ding, T.; Karlov, D. S.; Pino-Angeles, A.; Tikhonova, I. G. Intermolecular Interactions in G Protein-Coupled Receptor Allosteric Sites at the Membrane Interface from Molecular Dynamics Simulations and Quantum Chemical Calculations. *J. Chem. Inf. Model.* **2022**, *62* (19), 4736–4747. https://doi.org/10.1021/acs.jcim.2c00788.

(47) Shah; Lind, C.; Pascali, F. D.; Penn, R. B.; Jr, A. D. M.; Deshpande, and D. A. In Silico Identification of a B2-Adrenoceptor Allosteric Site That Selectively Augments Canonical B2AR-Gs Signaling and Function. *Proceedings of the National Academy of Sciences* **2022**, *119*, e2214024119.

(48) Inan, T.; Flinko, R.; Lewis, G. K.; MacKerell, A. D.; Kurkcuoglu, O. Identifying and Assessing Putative Allosteric Sites and Modulators for CXCR4 Predicted through Network Modeling and Site Identification by Ligand Competitive Saturation. *J. Phys. Chem. B* **2024**, *128* (21), 5157–5174. https://doi.org/10.1021/acs.jpcb.4c00925.

(49) Liu, H.; Kim, H. R.; Deepak, R. N. V. K.; Wang, L.; Chung, K. Y.; Fan, H.; Wei, Z.; Zhang, C. Orthosteric and Allosteric Action of the C5a Receptor Antagonists. *Nat. Struct. Mol. Biol.* **2018**, *25* (6), 472–481. https://doi.org/10.1038/s41594-018-0067-z.

(50) Mao, C.; Shen, C.; Li, C.; Shen, D.-D.; Xu, C.; Zhang, S.; Zhou, R.; Shen, Q.; Chen, L.-N.; Jiang, Z.; Liu, J.; Zhang, Y. Cryo-EM Structures of Inactive and Active GABAB Receptor. *Cell Res.* **2020**, *30* (7), 564–573. https://doi.org/10.1038/s41422-020-0350-5.

(51) Persechino, M.; Hedderich, J. B.; Kolb, P.; Hilger, D. Allosteric Modulation of GPCRs: From Structural Insights to in Silico Drug Discovery. *Pharmacol. Ther.* **2022**, *237*, 108242. https://doi.org/10.1016/j.pharmthera.2022.108242.

(52) Burggraaff, L.; Veen, A. van; Lam, C. C.; Vlijmen, H. W. T. van; IJzerman, A. P.; Westen, G. J. P. van. Annotation of Allosteric Compounds to Enhance Bioactivity Modeling for Class A GPCRs. *J. Chem. Inf. Model.* **2020**, *60* (10), 4664–4672. https://doi.org/10.1021/acs.jcim.0c00695.

(53) Fasciani, I.; Petragnano, F.; Aloisi, G.; Marampon, F.; Carli, M.; Scarselli, M.; Maggio, R.; Rossi, M. Allosteric Modulators of G Protein-Coupled Dopamine and Serotonin Receptors: A New Class of Atypical Antipsychotics. *Pharmaceuticals* **2020**, *13* (11), 388. https://doi.org/10.3390/ph13110388.

(54) Grundmann, M.; Bender, E.; Schamberger, J.; Eitner, F. Pharmacology of Free Fatty Acid Receptors and Their Allosteric Modulators. *Int. J. Mol. Sci.* **2021**, *22* (4), 1763. https://doi.org/10.3390/ijms22041763.