https://doi.org/10.1038/s41586-019-1141-3

Structural basis of ligand recognition at the human MT_1 melatonin receptor

Benjamin Stauch[1,2,16], Linda C. Johansson[1,2,16], John D. McCorvy[4,15], Nilkanth Patel[1,3], Gye Won Han[1,2], Xi-Ping Huang[4,5], Cornelius Gati[6,7], Alexander Batyuk[8], Samuel T. Slocum[4], Andrii Ishchenko[1,2], Wolfgang Brehm[9], Thomas A. White[9], Nairie Michaelian[1,2], Caleb Madsen[10], Lan Zhu[11], Thomas D. Grant[12], Jessica M. Grandner[1,3], Anna Shiriaeva[1,2], Reid H. J. Olsen[4], Alexandra R. Tribo[4], Saïd Yous[13], Raymond C. Stevens[1,2,3], Uwe Weierstall[10,11], Vsevolod Katritch[1,2,3], Bryan L. Roth[4,5,14]*, Wei Liu[11]* & Vadim Cherezov[1,2,3]*

Melatonin (*N*-acetyl-5-methoxytryptamine) is a neurohormone that maintains circadian rhythms[1] by synchronization to environmental cues and is involved in diverse physiological processes[2] such as the regulation of blood pressure and core body temperature, oncogenesis, and immune function[3]. Melatonin is formed in the pineal gland in a light-regulated manner[4] by enzymatic conversion from 5-hydroxytryptamine (5-HT or serotonin), and modulates sleep and wakefulness[5] by activating two high-affinity G-protein-coupled receptors, type 1A (MT_1) and type 1B (MT_2)[3,6]. Shift work, travel, and ubiquitous artificial lighting can disrupt natural circadian rhythms; as a result, sleep disorders affect a substantial population in modern society and pose a considerable economic burden[7]. Over-the-counter melatonin is widely used to alleviate jet lag and as a safer alternative to benzodiazepines and other sleeping aids[8,9], and is one of the most popular supplements in the United States[10]. Here, we present high-resolution room-temperature X-ray free electron laser (XFEL) structures of MT_1 in complex with four agonists: the insomnia drug ramelteon[11], two melatonin analogues, and the mixed melatonin–serotonin antidepressant agomelatine[12,13]. The structure of MT_2 is described in an accompanying paper[14]. Although the MT_1 and 5-HT receptors have similar endogenous ligands, and agomelatine acts on both receptors, the receptors differ markedly in the structure and composition of their ligand pockets; in MT_1, access to the ligand pocket is tightly sealed from solvent by extracellular loop 2, leaving only a narrow channel between transmembrane helices IV and V that connects it to the lipid bilayer. The binding site is extremely compact, and ligands interact with MT_1 mainly by strong aromatic stacking with Phe179 and auxiliary hydrogen bonds with Asn162 and Gln181. Our structures provide an unexpected example of atypical ligand entry for a non-lipid receptor, lay the molecular foundation of ligand recognition by melatonin receptors, and will facilitate the design of future tool compounds and therapeutic agents, while their comparison to 5-HT receptors yields insights into the evolution and polypharmacology of G-protein-coupled receptors.

To investigate the basis of ligand recognition by MT_1, we crystallized the receptor using an intracellular loop 3 fusion with *Pyrococcus abyssi* glycogen synthase (PGS)[15]. To increase receptor thermostability, we introduced nine point mutations (D73$^{2.50}$N, L95^{ECL1}F, G104$^{3.29}$A, F116$^{3.41}$W, N124$^{3.49}$D, C127$^{3.52}$L, W251$^{6.48}$F, A292$^{7.50}$P and N299$^{8.47}$D; superscripts represent Ballesteros–Weinstein nomenclature[16], see Methods), which were essential for crystallization. In combination, these modifications result in an approximately 40-fold reduction in melatonin affinity (approximately 14-fold in the presence of a physiological concentration of sodium; Extended Data Table 1), primarily caused by the allosteric sodium site[17] mutation D73$^{2.50}$N, which is known to affect coupling to G proteins[18], and N299$^{8.47}$D, which is located at the G protein interface (Extended Data Tables 1 and 2).

MT_1 was crystallized in lipidic cubic phase (LCP; Extended Data Fig. 1), and serial femtosecond crystallography (SFX) was used to obtain structures of MT_1 bound to the insomnia drug ramelteon (2.8 Å resolution), the high-affinity agonist 2-phenylmelatonin (2-PMT; 2.9 Å), 2-iodomelatonin, and agomelatine (both 3.2 Å). Data collection and refinement statistics for all MT_1 structures are provided in Extended Data Table 3. The overall receptor conformation in all four structures is nearly identical (Cα root mean squared deviation (r.m.s.d.) < 0.3 Å); therefore, we focus our analysis below on the highest-resolution MT_1–ramelteon structure unless otherwise noted.

The structure of MT_1 displays the canonical GPCR topology of a heptahelical transmembrane bundle (7TM) with an extracellular N terminus, three extracellular loops (ECLs), three intracellular loops (ICLs), and a short amphipathic helix VIII oriented parallel to the membrane (Fig. 1). On the basis of structural alignment with the active[19] and inactive states[20] of the β_2-adrenoreceptor (β_2AR), we concluded that the 7TM of MT_1 is captured in an inactive conformation. Similar inactive states have been observed for agonist-bound complexes without intracellular binding partners of some GPCRs that have naturally weak coupling between the extracellular and intracellular sides (for example, β_2AR), or those in which this coupling was disrupted by mutations[21,22].

ECL2 forms a short β-hairpin anchored to helix III by the conserved disulfide bridge between C100$^{3.25}$ and C177^{ECL2} and a number of polar interactions with helices III, VI and VII, tightly sealing off the extracellular entrance to the orthosteric binding site (Fig. 1), which is among the smallest of any receptor solved to date (710 Å^3; Extended Data Table 4).

Potential ligand access to the orthosteric binding site is provided by a channel between helices IV and V that opens towards the lipid bilayer (Fig. 1b, d). The channel is located about 7–11 Å below the hydrophobic membrane boundary, is about 10 Å long from the entrance to the centre of the ligand binding site, has a minimum diameter of 3 Å (about 4–5 Å otherwise), and is lined with a number of hydrophobic residues and H195$^{5.46}$, which is highly conserved in melatonin receptors.

[1]Bridge Institute,USC Michelson Center for Convergent Biosciences, University of Southern California, Los Angeles, CA, USA. [2]Department of Chemistry, University of Southern California, Los Angeles, CA, USA. [3]Department of Biological Sciences, University of Southern California, Los Angeles, CA, USA. [4]Department of Pharmacology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [5]National Institute of Mental Health Psychoactive Drug Screening Program, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [6]SLAC National Accelerator Laboratory, Bioscience Division, Menlo Park, CA, USA. [7]Stanford University, Department of Structural Biology, Stanford, CA, USA. [8]Linac Coherent Light Source, SLAC National Accelerator Laboratory, Menlo Park, CA, USA. [9]Center for Free-Electron Laser Science, DESY, Hamburg, Germany. [10]Department of Physics, Arizona State University, Tempe, AZ, USA. [11]School of Molecular Sciences and Biodesign Center for Applied Structural Discovery, Biodesign Institute, Arizona State University, Tempe, AZ, USA. [12]Hauptman-Woodward Institute, Department of Structural Biology, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, Buffalo, NY, USA. [13]Univ Lille, Inserm, CHU Lille, UMR-S 1172 - JPArc - Centre de Recherche Jean-Pierre AUBERT Neurosciences et Cancer, Lille, France. [14]Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [15]Present address: Department of Cell Biology, Neurobiology and Anatomy, Medical College of Wisconsin, Milwaukee, WI, USA. [16]These authors contributed equally: Benjamin Stauch, Linda C. Johansson. *e-mail: bryan_roth@med.unc.edu; w.liu@asu.edu; cherezov@usc.edu

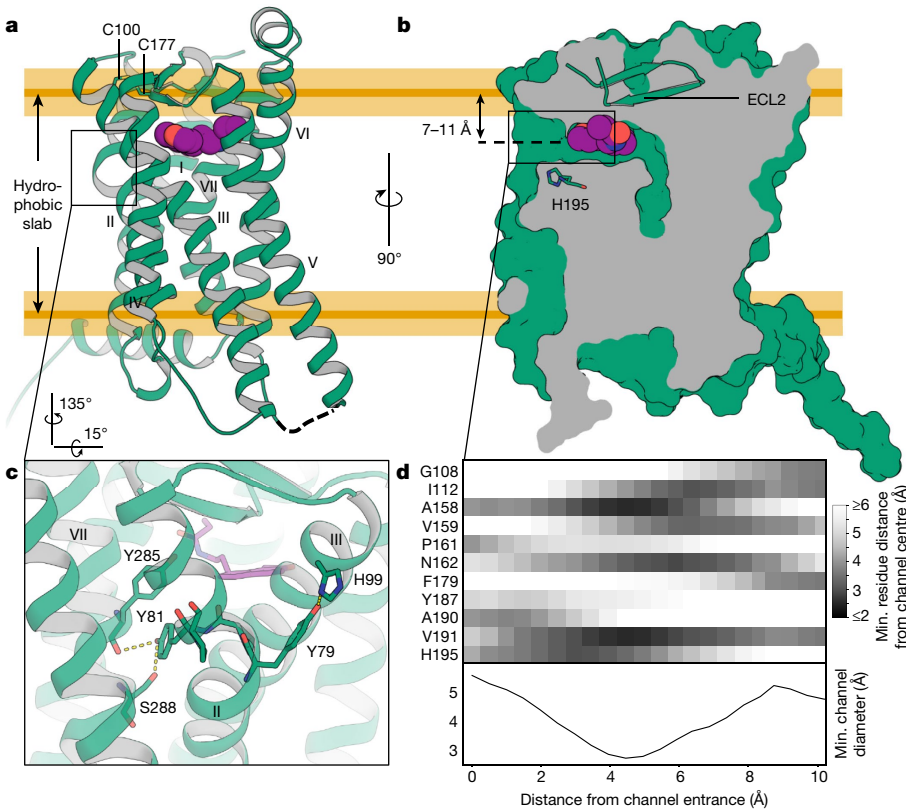


Fig. 1 | Structural features of MT_1. a, Overall architecture of MT_1 (green; disulfide, yellow; helices, labelled with Roman numerals) bound to ramelteon (purple). ICL3 was replaced by a fusion protein and is shown as a dashed line. Approximate boundaries of the hydrophobic slab corresponding to the lipid tails are derived from molecular dynamics simulation and indicated by orange lines (yellow shaded areas represent s.d.). ECL2 closes off the binding site to the extracellular space. **b**, Section through the receptor illustrating the lateral ligand access channel. **c**, Details of the YPYP motif in helix II that forms a bulge in proximity to the ligand. Residues are shown as green sticks, and hydrogen bonds as dashed lines. **d**, Proximity matrix of pore-lining residues, and minimum diameter profile across the length of the channel, calculated using spherical probes.

A part of helix II that participates in shaping the ligand binding pocket contains a distinct YPYP motif (residues $Y79^{2.56}$ to $P82^{2.59}$), which introduces a bulge and kink to helix II and places the ring planes of its tyrosine residues in an antiparallel orientation (Fig. 1c). $Y79^{2.56}$ interacts with $H99^{3.24}$, while $Y81^{2.58}$ points towards the receptor core and interacts with the backbone of $Y285^{7.43}$ and the side chain of $S288^{7.46}$. This YPYP motif is unique to melatonin receptors and the closely related orphan receptor GPR50. It is not found in any other membrane protein structures or predicted transmembrane segments of the human proteome (see Methods). Although $P82^{2.59}$ is conserved across class A receptors (74% of receptors), an equivalent of $P80^{2.57}$ is not present in any other human GPCR[22]. Mutation of any residue in the YPYP motif lowers thermostability by 6–10 °C and impairs receptor function (Extended Data Tables 5, 6), suggesting that this motif is an important structural element of melatonin receptors that contributes to receptor stability and activation.

Melatonergic ligands display a remarkably limited number of chemotypes. Indeed, over 80% of high-affinity melatonin ligands (from 391 compounds in ChEMBL24[23] (https://www.ebi.ac.uk/chembl/) with MT_1 affinity of less than 100 nM), most of which are agonists, possess a core that is either a (hetero)indene (mostly indole, benzofuran, benzoxazole) or a naphthalene bioisostere. We crystallized MT_1 with ligands from both of these major chemotypes, and together they revealed the conserved receptor interaction points and pharmacophore of melatonergic ligands.

Ramelteon engages in three anchoring interactions, including aromatic stacking of its heterocyclic core with $F179^{ECL2}$, and hydrogen bonds between its alkylamide tail and $Q181^{ECL2}$ and between its dihydrofuran moiety and $N162^{4.60}$ (Fig. 2a). The methoxy and alkylamide groups of 2-PMT, 2-iodomelatonin, and agomelatine (Fig. 2b) interact with $N162^{4.60}$ and $Q181^{ECL2}$ in a similar manner to ramelteon. The ligand core forms a close contact with $G108^{3.33}$, suggesting that a larger residue in this position would clash with the ligand. Indeed, the $G108^{3.33}$A mutation practically abolishes ligand binding (Extended Data Table 5). The binding site contains a hydrophobic sub-pocket, which accommodates the phenyl and iodine substituents of the co-crystallized melatonin analogues (Fig. 2c).

Mutations of the ligand coordinating residues $F179^{ECL2}$A and $Q181^{ECL2}$A are detrimental to receptor stability and activation (Extended Data Tables 5, 6), while $Q181^{ECL2}$E is functionally better tolerated. The $N162^{4.60}$A mutation inactivates the receptor (Extended Data Table 6) but does not affect ligand affinity or receptor stability (Extended Data Tables 1, 5), reflecting the importance of the methoxy or equivalent substituent for ligand efficacy[24]. These findings are consistent with molecular dynamics simulations of the receptor–ligand complexes (Extended Data Fig. 2), wherein equilibrium interactions between ligands and two anchor residues, $F179^{ECL2}$ and $Q181^{ECL2}$, are very stable and are stronger than the interaction with $N162^{4.60}$. Together, these findings explain the tight steric fit and lipophilic interactions between ligands and receptor, mainly mediated by ECL2, as primary determinants of affinity. The importance of ECL2 for ligand binding in melatonin receptors is further highlighted by the recent demonstration that transplanting ECL2 of MT_1 to GPR50, a paralogue melatonin-related receptor that does not bind melatonin[25], is sufficient to restore ligand-dependent activation in the resulting chimaera[26]. Analysis of the otherwise remarkably conserved binding site shows that the main difference between MT_1 or MT_2 and GPR50 is in ECL2 (Fig. 2d).

Notably, although residue $N255^{6.52}$ does not interact with ligands in our structures (Fig. 2a, c), and mutation $N255^{6.52}$A does not affect ligand affinity (Extended Data Table 1), this mutation reduces receptor stability and signalling (Extended Data Tables 5, 6), suggesting that this

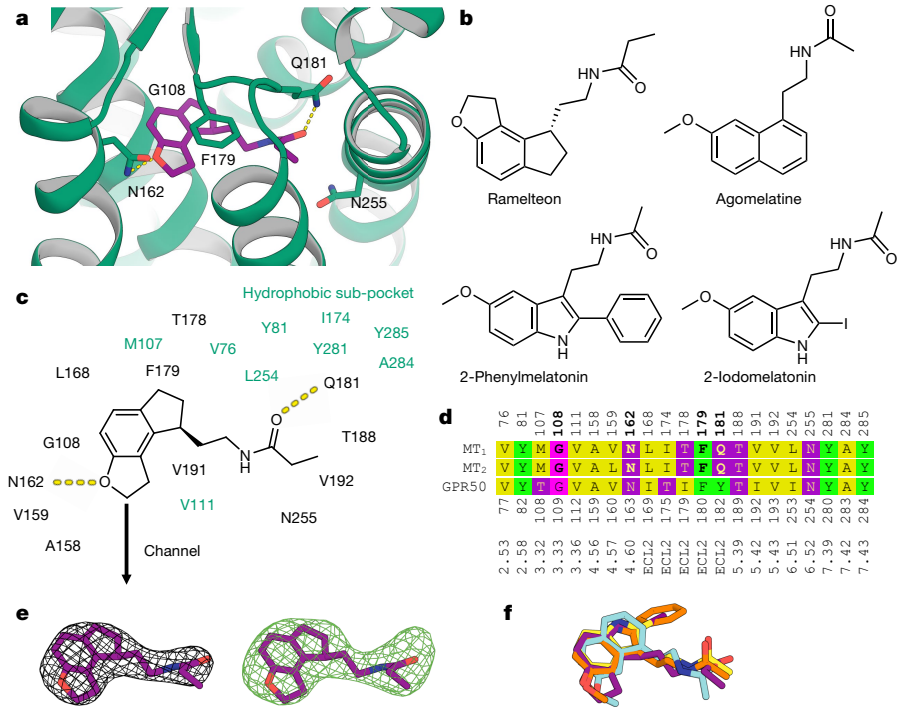


Fig. 2 | Ligand recognition at MT_1. **a**, Ramelteon (purple) forms specific interactions with side chains of $N162^{4.60}$ and $Q181^{ECL2}$ of MT_1 (green), and stacks with $F179^{ECL2}$ (all side chains shown as green sticks). **b**, Chemical structures of melatonergic ligands that were co-crystallized with MT_1 in this work. **c**, Binding site composition and interactions of ramelteon as a representative of all four complexes. The hydrophobic sub-pocket that accommodates substituents at the 2 position of indole-like ligands is shown in green. **d**, Structure-guided sequence alignment of the binding site residues of MT_1 (residue numbering above), MT_2, and GPR50 (residue numbering below). Ligand-interacting residues in MT_1 and MT_2 are highlighted in bold, and Ballesteros–Weinstein residue numbering is provided for reference. **e**, $2mF_o - DF_c$ electron density map (grey mesh) in the binding site of the MT_1–ramelteon complex, contoured at 1.0σ, and simulated annealing $mF_o - DF_c$ omit map (green mesh), contoured at 3.0σ. Electron density maps for other ligands are shown in Extended Data Fig. 1. **f**, Overlay of experimental ligand conformations of ramelteon (purple), 2-PMT (orange), 2-iodomelatonin (yellow), and agomelatine (cyan) after receptor superimposition. The conformations of receptor side chains in the binding site are very similar between the complexes and are omitted for clarity.

residue might be involved in receptor activation. Small ligands such as melatonin or agomelatine are affected more strongly by these effects than bulkier ligands (Extended Data Fig. 2e, Extended Data Table 6). In all of our MT_1 structures we observe a small blob of electron density (about 6σ in $mF_o - DF_c$ maps) close to $N255^{6.52}$ (Extended Data Fig. 1e). We tentatively attribute this density to propan-2-ol, an essential additive in the crystallization of MT_1. This additive was not needed to crystallize MT_2, in which mutation of $N^{6.52}$ does not affect signalling[14]. Under physiological conditions, $N255^{6.52}$ is likely to remain hydrated rather than interacting directly with the ligand, as also suggested by our molecular dynamics simulations (Extended Data Fig. 2d).

The existence of the lateral channel is supported by studies of the bitopic ligand CTL 01-05-B-A05, a symmetrical agomelatine dimer linked by an ethoxyethane spacer (see Methods). This ligand was docked into the MT_1 binding site, indicating that it can be accommodated only if its aliphatic linker region protrudes through the channel, and, therefore, one of its monomers must pass through the channel upon binding (Fig. 3). We chose a residue in the vicinity of the channel entrance and designed a point mutation ($A190^{5.41}F$) that was predicted to keep the channel open but would interfere with the placement of peripheral moieties of the bitopic ligand. Signalling for monotopic ligands was only slightly affected, whereas the bitopic ligand lost about one order of magnitude in potency (Extended Data Table 7). On the other hand, mutation $A158^{4.56}M$, which was designed to block the channel, caused loss of functional activity for all tested agonists, consistent with the hypothesis that the ligand entry in MT_1 is largely defined by the lateral channel.

Melatonin and serotonin are ancient molecules with physiolog-ical roles that predate the evolution of nervous systems, multicellu-larity and their appropriation as a hormone and a neurotransmitter,

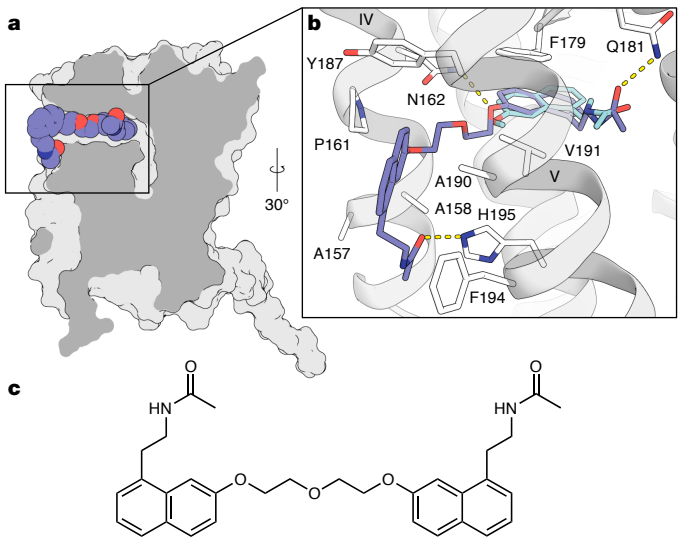


Fig. 3 | Docking model of bitopic ligand. **a**, Section through MT_1 (grey) showing the best docked pose of the bitopic ligand CTL 01-05-B-A05 (spheres with slate blue carbons). **b**, Details of receptor–ligand interactions. Dashed lines represent hydrogen bonds between ligand moieties and receptor residues $N162^{4.60}$, $Q181^{ECL2}$, and $H195^{5.46}$. **c**, Chemical structure of CTL 01-05-B-A05. The ligand protrudes from the lateral channel between helices IV and V. Its core shows minor displacement compared to the experimentally determined conformation of agomelatine (cyan sticks), and it forms favourable interactions with several residues (white sticks) in the periphery of the channel.

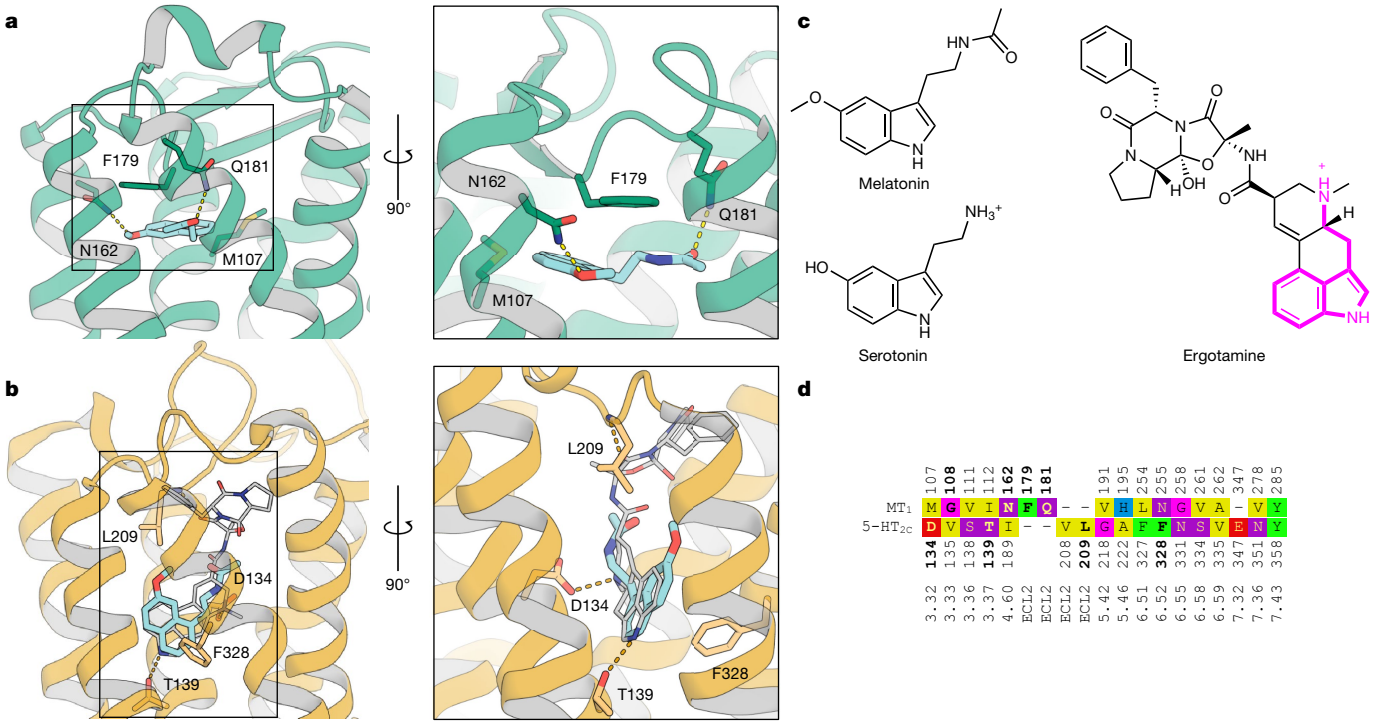


Fig. 4 | Comparison between MT_1 and 5-HT_{2C}. a, Crystal structure of agomelatine (cyan sticks, $pK_i(MT_1) = 8.8$) bound to MT_1 (green cartoon). Important residues in the binding site are shown as sticks with green carbons. **b**, Crystal structure of ergotamine (grey sticks) bound to 5-HT_{2C} (orange cartoon), and docking model of agomelatine (cyan sticks, $pK_i(5\text{-}HT_{2C})$[13] = 6.2). **c**, Chemical structures of melatonin, serotonin, and ergotamine with the substructure shared between serotonin and ergotamine highlighted (magenta). **d**, Structure-guided sequence alignment in the binding sites of MT_1 and 5-HT_{2C} receptors with reference sequence numbering (above/below), and Ballesteros–Weinstein numbering (bottom). Residues that participate in tight ligand interactions in the respective structure are bold.

respectively[27,28]. Their primordial functions, which result from their chemical properties, are still evident by their occurrence in plants and their antioxidant abilities[27,28]. The structural similarity and overlapping biosynthetic pathways of melatonin and serotonin contrast with their very different pharmacological roles and physicochemical parameters. Melatonin is soluble in both water and lipids and can traverse membranes by passive diffusion, whereas serotonin, owing to its charge, has to be secreted and actively transported[29]. Strikingly, 98% of high-affinity melatonin receptor ligands (<100 nM in the ChEMBL database) are neutral at physiological pH, whereas, depending on subtype, 80–99% of 5-HT receptor ligands are charged (see Methods).

Melatonin and 5-HT receptors share a low level of sequence identity (about 20–25%) and even lower binding site identity (1 out of 16 residues; Fig. 4d). A key question is how these receptor families evolved to bind structurally similar endogenous molecules while maintaining orthogonal functions. Very few ligands bind both melatonin and 5-HT receptors, most notably the psychedelic alkaloid 5-methoxy-*N*,*N*-dimethyltryptamine (5-MeO-DMT)[30] and the antidepressant agomelatine[12,13], albeit with reduced potency.

The docking model of agomelatine in the 5-HT_{2C} receptor shows a similar pose to the ergotamine core[31], but has a perpendicular orientation to agomelatine in the MT_1 structure and involves completely different residues (Fig. 4). This orthogonal binding mode and the low sequence conservation between melatonin and 5-HT receptors are testament to disparate evolution driven by their distinct cognate ligands. Their receptors are co-expressed in a number of tissues[32], where promiscuous binding of endogenous molecules has to remain the exception. The two structurally similar ligands can be reliably discriminated by (i) structurally different binding pockets; in particular, the lack in melatonin receptors of an equivalent residue to negatively charged $D^{3.32}$, which serves as primary anchor for all aminergic receptors; and (ii) the use of different physical ligand properties (restricting access of charged ligands to the MT_1 binding site through the membrane channel).

In conclusion, the binding site of melatonin in the MT_1 receptor is completely different from that of the related metabolite serotonin in 5-HT receptors, and the MT_1 receptor contains an unexpected channel that provides ligands lateral access to the receptor binding site from within the lipid bilayer. This channel represents a novel allosteric site that can be targeted by rational structure-based drug design. The atypical entry mechanism could impose constraints on ligand dimensions and physicochemical properties and can be exploited in the future development of synthetic melatonergic agents to address the need for safer sleeping aids without the potential for abuse, and for antidepressants that benefit from melatonin–5-HT polypharmacology by design.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-019-1141-3.

Received: 28 July 2018; Accepted: 25 March 2019;
Published online 24 April 2019.

1. Pévet, P. Melatonin receptors as therapeutic targets in the suprachiasmatic nucleus. *Expert Opin. Ther. Targets* **20**, 1209–1218 (2016).
2. Hardeland, R., Pandi-Perumal, S. R. & Cardinali, D. P. Melatonin. *Int. J. Biochem. Cell Biol.* **38**, 313–316 (2006).
3. Dubocovich, M. L. & Markowska, M. Functional MT1 and MT2 melatonin receptors in mammals. *Endocrine* **27**, 101–110 (2005).
4. Ganguly, S., Coon, S. L. & Klein, D. C. Control of melatonin synthesis in the mammalian pineal gland: the critical role of serotonin acetylation. *Cell Tissue Res.* **309**, 127–137 (2002).
5. Tosini, G., Owino, S., Guillaume, J. L. & Jockers, R. Understanding melatonin receptor pharmacology: latest insights from mouse models, and their relevance to human disease. *BioEssays* **36**, 778–787 (2014).
6. Dubocovich, M. L. et al. International Union of Basic and Clinical Pharmacology. LXXV. Nomenclature, classification, and pharmacology of G protein-coupled melatonin receptors. *Pharmacol. Rev.* **62**, 343–380 (2010).
7. Stoller, M. K. Economic effects of insomnia. *Clin. Ther.* **16**, 873–897, discussion 854 (1994).
8. Jockers, R. et al. Update on melatonin receptors: IUPHAR Review 20. *Br. J. Pharmacol.* **173**, 2702–2725 (2016).

9. Zlotos, D. P. Recent progress in the development of agonists and antagonists for melatonin receptors. *Curr. Med. Chem.* **19**, 3532–3549 (2012).
10. Clarke, T. C., Black, L. I., Stussman, B. J., Barnes, P. M. & Nahin, R. L. Trends in the use of complementary health approaches among adults: United States, 2002–2012. *Natl. Health Stat. Report* **79**, 1–16 (2015).
11. Owen, R. T. Ramelteon: profile of a new sleep-promoting medication. *Drugs Today (Barc)* **42**, 255–263 (2006).
12. Millan, M. J. et al. The novel melatonin agonist agomelatine (S20098) is an antagonist at 5-hydroxytryptamine2C receptors, blockade of which enhances the activity of frontocortical dopaminergic and adrenergic pathways. *J. Pharmacol. Exp. Ther.* **306**, 954–964 (2003).
13. Guardiola-Lemaitre, B. et al. Agomelatine: mechanism of action and pharmacological profile in relation to antidepressant properties. *Br. J. Pharmacol.* **171**, 3604–3619 (2014).
14. Johansson, L. C. et al. XFEL structures of the human MT2 melatonin receptor reveal the basis of subtype selectivity. *Nature* https://doi.org/10.1038/s41586-019-1144-0 (2019).
15. Yin, J., Mobarec, J. C., Kolb, P. & Rosenbaum, D. M. Crystal structure of the human OX2 orexin receptor bound to the insomnia drug suvorexant. *Nature* **519**, 247–250 (2015).
16. Ballesteros, J. A. & Weinstein, H. in *Methods in Neurosciences* Vol. 25 (ed. Sealfon, S. C.) 366–428 (Academic, 1995).
17. Katritch, V. et al. Allosteric sodium in class A GPCR signaling. *Trends Biochem. Sci.* **39**, 233–244 (2014).
18. White, K. L. et al. Structural connection between activation microswitch and allosteric sodium site in GPCR signaling. *Structure* **26**, 259–269.e5 (2018).
19. Rasmussen, S. G. et al. Crystal structure of the β2 adrenergic receptor–G_s protein complex. *Nature* **477**, 549–555 (2011).
20. Cherezov, V. et al. High-resolution crystal structure of an engineered human β2-adrenergic G protein-coupled receptor. *Science* **318**, 1258–1265 (2007).
21. Stauch, B. & Cherezov, V. Serial femtosecond crystallography of G protein-coupled receptors. *Annu. Rev. Biophys.* **47**, 377–397 (2018).
22. Isberg, V. et al. GPCRdb: an information system for G protein-coupled receptors. *Nucleic Acids Res.* **44** (D1), D356–D364 (2016).
23. Bento, A. P. et al. The ChEMBL bioactivity database: an update. *Nucleic Acids Res.* **42**, D1083–D1090 (2014).
24. Dubocovich, M. L. Luzindole (N-0774): a novel melatonin receptor antagonist. *J. Pharmacol. Exp. Ther.* **246**, 902–910 (1988).
25. Reppert, S. M., Weaver, D. R., Ebisawa, T., Mahle, C. D. & Kolakowski, L. F. Jr Cloning of a melatonin-related receptor from human pituitary. *FEBS Lett.* **386**, 219–224 (1996).
26. Clement, N. et al. Importance of the second extracellular loop for melatonin MT_1 receptor function and absence of melatonin binding in GPR50. *Br. J. Pharmacol.* **175**, 3281–3297 (2018).
27. Azmitia, E. C. Serotonin and brain: evolution, neuroplasticity, and homeostasis. *Int. Rev. Neurobiol.* **77**, 31–56 (2007).
28. Tan, D. X. et al. Melatonin: a hormone, a tissue factor, an autocoid, a paracoid, and an antioxidant vitamin. *J. Pineal Res.* **34**, 75–78 (2003).
29. Yu, H., Dickson, E. J., Jung, S. R., Koh, D. S. & Hille, B. High membrane permeability for melatonin. *J. Gen. Physiol.* **147**, 63–76 (2016).
30. de la Fuente Revenga, M. et al. Novel *N*-acetyl bioisosteres of melatonin: melatonergic receptor pharmacology, physicochemical studies, and phenotypic assessment of their neurogenic potential. *J. Med. Chem.* **58**, 4998–5014 (2015).
31. Peng, Y. et al. 5-HT_{2C} receptor structures reveal the structural basis of GPCR polypharmacology. *Cell* **172**, 719–730.e14 (2018).
32. Pasqualetti, M. et al. Distribution and cellular localization of the serotonin type 2C receptor messenger RNA in human brain. *Neuroscience* **92**, 601–611 (1999).

Acknowledgements We thank M. Chu, C. Hanson, K. Villers, J. Velasquez, and H. Shaye for technical support, and D.R. Mende for useful discussion of sequence analysis. This research was supported by the National Institutes of Health (NIH) grants R35 GM127086 (V.C.), R21 DA042298 (W.L.), R01 GM124152 (W.L.), R01 MH112205 (B.L.R.), and U24DK116195 (B.L.R.), the NIMH Psychoactive Drug Screening Program contract (B.L.R.), F31-NS093917 (R.H.J.O.), the National Science Foundation (NSF) BioXFEL Science and Technology Center 1231306 (B.S., W.L., U.W., T.D.G., V.C.), EMBO ALTF 677-2014 (B.S.), HFSP long-term fellowship LT000046/2014-L (L.C.J.), and a postdoctoral fellowship from the Swedish Research Council (L.C.J.). C.G. thanks the SLAC National Accelerator Laboratory and the Department of Energy for financial support through the Panofsky fellowship. T.A.W. and W.B. acknowledge financial support from the Helmholtz Association via Programme-Oriented Funding. Parts of this research were carried out at the LCLS, a National User Facility operated by Stanford University on behalf of the US Department of Energy and supported by the US Department of Energy Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515.

Reviewer information *Nature* thanks Christian Siebold, Ieva Sutkeviciute, Jean-Pierre Vilardaga and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions B.S., L.C.J., W.L. and V.C. conceived the project, analysed data and wrote the manuscript with contributions from all authors. B.S. and L.C.J. cloned and characterized the receptor, generated all constructs, crystallized the receptor, prepared all crystal samples and figures, solved and refined the structures, and assisted in generating mutant constructs for binding and functional analyses. B.S. designed thermostabilizing point mutations and performed sequence analysis. J.D.M., X.-P.H. and S.T.S. performed radioligand binding and functional experiments, assisted in generating mutant and wild-type constructs used for binding and functional analyses, and analysed all binding and functional data. A.I., N.M., A.S., L.Z. and W.L. assisted in XFEL sample preparation. G.W.H. performed structure refinement and quality control. B.S., L.C.J., A.B., L.Z., W.L. and V.C. collected XFEL data. C.G., W.B., T.A.W. and T.D.G. processed XFEL data and solved the indexing ambiguity. C.M. and U.W. operated the LCP injector during XFEL data collection. N.P. performed molecular docking and molecular dynamics calculations and assisted in preparing figures. J.M.G. assisted in docking calculations. V.K. supervised molecular docking and molecular dynamics calculations. R.H.J.O. assisted with molecular biology and generating mutant constructs. A.R.T. assisted with generating mutant constructs and functional experiments. S.Y. synthesized the bitopic compound, analysed data and edited the paper. R.C.S. contributed to study design and selection of chemical compounds for receptor stabilization and functional characterization, supervised protein expression and edited the paper. B.L.R. supervised pharmacological experiments and edited the paper. W.L. supervised the LCP crystallization and optimization experiments. V.C. coordinated and supervised the whole project.

Competing interests The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41586-019-1141-3.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1141-3.

Correspondence and requests for materials should be addressed to B.L.R., W.L. and V.C.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Design of thermostabilizing point mutations. Wild-type human MT_1 showed extremely low expression, yield, and stability for pursuing structural studies. To predict stabilizing point mutations, we searched the MT_1 reference sequence (UniProt[33] identifier P48039) against the database of non-redundant (nr) protein sequences using the *blastp* algorithm[34,35], excluding 'low quality' and exactly identical sequences. Standard parameters were used, and alignments were scored using the BLOSUM62 matrix. The 100 closest identified homologues were aligned using COBALT[36], yielding the sequence conservation profile P_{MT1} for receptor species orthologues. Next, a conservation profile Q_{ClassA} for all transmembrane residues of human class A GPCRs (that is, paralogous sequences) was generated using the structure-guided sequence alignment of GPCRdb[22,37]. Corresponding residues in P_{MT1} and Q_{ClassA} were mapped using the generic residue numbering of GPCRdb[37] for the sequence of human MT_1, which is contained in both alignments. For each of these residues, the dissimilarity of the two conservation profiles was quantified using the Kullback–Leibler divergence[38] $D_{KL}(P||Q)$ with $P(i)$ and $Q(i)$ being the relative frequencies of each amino acid in profiles P_{MT1} and Q_{ClassA}, respectively, for that particular residue. Residues were ranked from most to least diverging, and the top twenty most divergent residues were mutated and tested. Of all mutations identified, $A^{7.50}P$, $N^{3.49}D$, and $C^{3.52}L$ had the most beneficial effect on aSEC (analytical size-exclusion chromatography) and thermostability, and facilitated crystallization.

Further mutations that improved expression, yield, monodispersity, thermostability, or crystallizability, were $D^{2.50}N$[18,39], $G^{3.29}A$ (the most beneficial mutation out of an exhaustive Gly-to-Ala scan[40]), $F^{3.41}W$[41], $W^{6.48}F$ of the proposed rotamer toggle switch[42], $N^{8.47}D$, and $L95^{ECL1}F$, which restores a WxFG motif commonly found in class A GPCRs[43].

Crystallization of MT_1. The codon-optimized nucleotide sequence of human MT_1 was synthesized by GenScript and subcloned into a modified pFastBac1 baculovirus expression vector (Invitrogen) containing an N-terminal haemagglutinin (HA) signal sequence, Flag tag, 10× His tag, and PreScission protease (PSP) cleavage site. Site-directed mutagenesis was carried out using oligonucleotides (IDT) with internal mismatches and AccuPrime Pfx polymerase (Thermo Fisher Scientific) and verified by Sanger sequencing (Genewiz). After truncating 11 N-terminal (including both glycosylation sites of the receptor) and 25 C-terminal amino acids, we replaced intracellular receptor amino acid residues 219–227 with the 196-amino-acid catalytic domain of *Pyrococcus abyssi* glycogen synthase (PGS, UniProt Q9V2J8)[15]. Several beneficial point mutations were identified (see above), nine of which were used in the final crystallized construct (MT_1-CC).

The resulting MT_1-CC chimaera was expressed in *Spodoptera frugiperda* (Sf9, purchased from ATCC, CRL-1711, authenticated by supplier using morphology and growth characteristics, certified mycoplasma-free) insect cells using the Bac-to-bac baculovirus expression system (Invitrogen). Cells with a density of (2–3) × 10^6 cells ml^{-1} were infected with baculovirus at 27 °C at a multiplicity of infection of 5, harvested by centrifugation 48 h after infection and stored at −80 °C until use. The membrane fraction was isolated from 3 l of biomass using repeated Dounce homogenization and ultracentrifugation in hypotonic (twice, 10 mM HEPES pH 7.5, 10 mM $MgCl_2$, 20 mM KCl) and hypertonic (three times, 10 mM HEPES pH 7.5, 10 mM $MgCl_2$, 20 mM KCl, 1 M NaCl) buffer. After membrane isolation, the ligand (ramelteon, Apex Biosciences; agomelatine, Sigma-Aldrich; 2-phenylmelatonin, Tocris; or 2-iodomelatonin, Tocris) concentration was 100 μM in all buffers.

Washed membranes were incubated in hypotonic buffer in the presence of 2 mg ml^{-1} iodoacetamide (Sigma-Aldrich) for 45 min, and receptor was subsequently extracted from membranes in a volume of 200 ml by addition of 2× solubilization buffer (100 mM HEPES pH 7.5, 300 mM NaCl, 2% (wt/vol) n-dodecyl-β-D-maltopyranoside (DDM, Anatrace) and 0.4% (wt/vol) cholesterol hemisuccinate (CHS, Sigma-Aldrich)) for 3 h. After overnight incubation with 1 ml of Talon (immobilized metal affinity chromatography, IMAC) resin (Clontech) in the presence of 150 mM NaCl and 20 mM imidazole, the sample was washed on a gravity column (Bio-Rad) with 12 column volumes (cv) of wash buffer 1 (50 mM HEPES pH 7.5, 800 mM NaCl, 25 mM imidazole pH 7.5, 10 mM $MgCl_2$, 10% (vol/vol) glycerol, 0.05%/0.01% (wt/vol) DDM/CHS) followed by 6 cv of wash buffer 2 (50 mM HEPES pH 7.5, 150 mM NaCl, 50 mM imidazole, 10% (vol/vol) glycerol, 0.025%/0.005% (wt/vol) DDM/CHS). The sample was eluted in 3.75 cv of elution buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 250 mM imidazole pH 7.5, 10% (vol/vol) glycerol, 0.025%/0.005% (wt/vol) DDM/CHS) after discarding the first 0.6 cv of elution flow-through, and subsequently concentrated to a volume of about 400 μl using an Amicon centrifugal filter with 100 kDa molecular weight cutoff (Millipore). The concentrated sample and 30 IU of His-tagged PSP (GenScript) were concomitantly passed over a PD MiniTrap G-25 desalting column (GE Healthcare) to remove imidazole and adjust detergent concentration to 0.05%/0.01% (wt/vol) DDM/CHS. After overnight incubation, cleaved tags and protease were removed by reverse IMAC, and the receptor was concentrated to ~20–30 mg ml^{-1} as above.

The receptor was reconstituted into LCP by mixing two volumes of purified receptor with three volumes of molten monoolein/cholesterol (9:1 wt/wt) using coupled gas-tight 100-μl syringes (Hamilton) as previously described[44]. For initial crystallization trials, protein-laden LCP boli (40 nl) were dispensed and overlaid with 800 nl precipitant in 96-well glass sandwich plates (Marienfeld) using the NT8-LCP system (Formulatrix). Initial crystals of MT_1 bound to the high-affinity agonists ramelteon and 2-PMT were 30–50 μm in size in the maximum dimension and could not be optimized to diffract better than 5 Å at a microfocus synchrotron beamline (Extended Data Fig. 1a, b). We then shifted our focus to obtaining small, high-density crystals grown in syringes[45]. To prepare samples for XFEL data collection, strings of 5 μl protein-laden LCP were injected into syringes containing 50 μl precipitant solution (60–100 mM potassium phosphate monobasic, 32–35% (vol/vol) PEG 400, 100 mM HEPES pH 7.0, 1 mM ligand, 2.5% (vol/vol) DMSO, 1.5% (vol/vol) propan-2-ol), and incubated for 3–10 days at 17 °C. After expulsion of excess precipitant, equivalent syringes were pooled together, and consolidated samples were reconstituted by addition of 7.9 monoacylglycerol (MAG) lipid[46].

Crystallographic data collection, structure solution and refinement. Data were collected at the CXI instrument of the Linac Coherent Light Source (LCLS)[47] at photon energies of 9.52 and 9.83 keV, pulse durations of 30 and 43 fs, pulse repetition rate of 120 Hz, approximately 10^{11} photons per pulse (5–11% transmission), and beam size 1.5 × 1.5 μm^2. MT_1-CC micro-crystals (~5–10 μm maximum dimension, Extended Data Fig. 1c) were delivered at room temperature in a vacuum chamber to the intersection with the XFEL beam using an LCP jet viscous medium injector[46] with a 50-μm nozzle at a flow rate of 0.3–0.36 μl min^{-1}, and diffraction images were collected using the CSPAD detector[48]. Crystal hits were identified using Cheetah[49] (more than 20 Bragg peaks of minimum 3 pixels in size with a signal-to-noise ratio better than 4 using an intensity threshold of 200 detector intensity units). Patterns were indexed, integrated, and merged with CrystFEL (v0.6.2)[50], using MOSFLM[51], DirAx[52], and XDS[53] for indexing with tightened tolerances of 1% on reciprocal unit cell dimensions and 0.5° on reciprocal unit cell angles. The total number of collected images/hits/indexed images are as follows: 517,530/43,306/42,679 (MT_1-CC–ramelteon), 726,497/119,563/99,897 (MT_1-CC–2-PMT), 429,006/21,776/21,038 (MT_1-CC–2-iodomelatonin), 466,602/45,820/43,423 (MT_1-CC–agomelatine). Integration radii of 3, 5, and 6 pixels with per pattern resolution cut-offs 1.2 nm^{-1} above the conservative resolution estimates for each crystal were used (push-res option). The default values were used for all other options. On the basis of apparent metric symmetry, data were first processed in a cubic space group; however, no molecular replacement solution could be found. A tetragonal lattice was then considered with the length of the unique axis *c* very close to that of the other two axes. This led to an indexing ambiguity where each crystal could be indexed in three different ways related by a permutation of the axes ($hkl \rightarrow klh \rightarrow lhk$). The correct assignment from the three options must be made for each crystal. Although the ambigator program implemented in CrystFEL can only distinguish between two possible indexing assignments at a time[54], in this case the three-fold ambiguity could be resolved by applying the algorithm iteratively at least twice using the same axis permutation operator $hkl \rightarrow klh$, which generates all three indexing options cyclically.

The structure of MT_1-CC–2-PMT was then solved by molecular replacement (MR) in space group $P4\,2_1\,2$ using the backbone of the helical bundle of the human OX_2 orexin receptor[15] (PDB ID: 4S0V) as a search model, followed by a separate search against the PGS coordinates from the same structure using Phaser 2.1[55]. One copy of each molecule in the asymmetric unit was readily identified with TFZ scores (*Z*-score of the translation function) of 9.8 for both partial and final solution. The packing in our MT_1 crystals was found to be predominantly mediated by PGS, with the receptor crystallized as a monomer in the asymmetric unit, forming a layered structure with receptors from adjacent layers coordinated head to head (Extended Data Fig. 1f).

All refinements were performed using Refmac5[56] and Buster v.2.10.2[57] followed by manual examination and rebuilding of the refined coordinates in the program Coot[58] using both $2mF_o - DF_c$ and $mF_o - DF_c$ maps, as well as omit maps. Ligand restraints were generated using Prodrg[59]. The remaining three structures were solved by MR using the pruned protein coordinates of MT_1-CC–2-PMT as a search model, and repeating the refinement procedure described above. The Ramachandran plot determined by MolProbity[60] indicates that with the exception of Y79 of the YPYP motif all residues are in favoured/allowed regions: 97.1/2.7% of residues (MT_1-CC–ramelteon), 96/3.8% of residues in MT_1-CC–2-PMT, 96.4/3.4% of residues in MT_1-CC–2-iodomelatonin, 95.4/4.4% of residues in MT_1-CC–agomelatine. The final data collection and refinement statistics are shown in Extended Data Table 3. Simulated annealing omit maps ($mF_o - DF_c$) were calculated using Phenix[61].

Thermostability studies. To assess the stability of the solubilized receptor, 1–3 μg purified protein was incubated in a volume of 100 μl at 4 °C for 20 min in the presence of 1.5 μM 7-diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin (CPM, Sigma-Aldrich)[62], added as a stock solution in DMSO (1% final concentration, vol/vol), 50 mM HEPES pH 7.5, 150 mM NaCl, 10% (vol/vol) glycerol, and 0.05%/0.01% (wt/vol) DDM/CHS. After incubation of the sample at room temperature for 5 min, thermal unfolding of the receptor was induced and monitored using a Rotor-Gene Q instrument (QIAGEN) between 25 °C and 95 °C (+2 °C min^{-1}) at wavelengths of 365 nm (excitation) and 460 nm (emission), and gain settings of −2 to −1, and melting temperatures were extracted from the first derivative of the melting curve.

Synthesis of bitopic compound. *N*-{2-[7-(2-{2-[8-(2-Acetylaminoethyl)-naphthalen-2-yloxy]-ethoxy}-ethoxy)-naphthalen-1-yl]-ethyl}acetamide (compound **2**, CTL 01-05-B-A05) was obtained in one step by condensation of the described[63] *N*-[2-(7-hydroxynaphth-1-yl)ethyl]acetamide (**1**) with bis(2-bromoethyl)ether in the presence of potassium carbonate in acetonitrile[64]. The synthesis scheme, experimental section, and NMR spectra are shown in Supplementary Figs. 7, 8.

Radioligand binding assays. Binding assays were performed using HEK293T cells (purchased from ATCC, CRL-11268, authenticated by supplier using morphology, growth characteristics and STR profiling, certified mycoplasma-free) transfected with wild-type or mutants of MT_1. All binding assays were performed in standard binding buffer (50 mM Tris, 10 mM $MgCl_2$, 0.1 mM EDTA, 0.1% BSA, 0.01% ascorbic acid, pH 7.4) using [^{3}H]melatonin (PerkinElmer, specific activity = 77.4–84.7 Ci/mmol) as the radioligand. For competitive binding, assays were performed with various concentrations of cold unlabelled ligands (100 fM to 10 μM), [^{3}H]melatonin (0.2–1.7 nM), and resuspended membranes in a total volume of 125 μl. Competition assays were incubated for 4 h at 37 °C in a humidified incubator until harvesting. For constructs with reduced binding affinity for [^{3}H]melatonin, homologous competition binding assays were carried out at two [^{3}H]melatonin concentrations (about 1–2 and 5–10 nM) to estimate K_d values. For all assays, non-specific activity was defined as the addition of 5 μM 2-PMT. For determining the effect of NaCl on melatonin binding affinity, 147 mM NaCl was included in the binding buffer. Plates were harvested using vacuum filtration onto 0.3% polyethyleneimine pre-soaked 96-well Filtermat A (PerkinElmer) and washed three times with 250 μl per well of cold wash buffer (50 mM Tris, pH 7.4). Filters were dried and scintillation cocktail (Meltilex, PerkinElmer) was melted and allowed to cool to room temperature. Afterwards, filter plates were placed in cassettes and read using a Wallac TriLux Microbeta counter (PerkinElmer). Data were analysed either using 'One-site-homologous' (to yield K_d) or 'One-site-Fit K_i' (to yield K_i) using GraphPad Prism 7.0.

MT_1 $G_{i/o}$-mediated cAMP inhibition assay. MT_1 $G_{i/o}$-mediated cAMP inhibition assays were performed in HEK293T cells (ATCC CRL-11268, mycoplasma-free) co-expressing the cAMP biosensor GloSensor-22F (Promega) and human MT_1. Transfected cells were seeded (10–15,000 cells per 40 μl per well) into poly-L-lysine-coated white 384 clear-bottom tissue culture plates in DMEM containing 1% dialysed FBS. On the next day, ligands were diluted from 10 μM to 100 fmol in HBSS (Hank's balanced salt solution), 20 mM HEPES, 0.1% BSA, 0.01% ascorbic acid, pH 7.4. Medium was removed on 384-well plates and 20 μl drug buffer (HBSS, 20 mM HEPES, pH 7.4) was added per well and allowed to equilibrate for at least 15 min at room temperature. To start the assay, cells were treated with 5 μl per well of 5× concentrated ligands in HBSS, 20 mM HEPES, 0.1% BSA, 0.01% ascorbic acid, pH 7.4 using a FLIPR (Molecular Devices). After 15 min, cAMP accumulation was initiated by addition of 10 μl per well of 0.3 μM isoproterenol (final concentration) in GloSensor reagent. Luminescence per well per second (LCPS) was read on a Wallac TriLux Microbeta counter (PerkinElmer). Data were normalized to maximum cAMP inhibition by melatonin (100%) and basal cAMP accumulation by isoproterenol (0%), and analysed using the sigmoidal dose–response function in GraphPad Prism 7.0.

Transmembrane segment sequence analysis. To analyse amino acid frequencies and distributions in transmembrane segments, two data sets were constructed: one data set comprising experimentally determined, representative structures of membrane proteins (experimental-TM); and one data set comprising the subset of the human proteome with transmembrane regions predicted by a computational method (human-TM).

The experimental-TM data set was constructed as follows. A list of experimental polypeptide X-ray structures was retrieved from the OPM database[65], which provides quaternary assemblies and membrane boundaries for all experimental membrane protein structures from the Protein Data Bank[66]. Because OPM also contains non-helical proteins (for example, β-barrel folds) that would skew our analysis, only entries containing at least one helical membrane-spanning element were retained. To this end, the secondary structure of each polypeptide chain was computed using DSSP[67] v3.0, and for each annotated TM segment that was determined to be at least 80% helical, leading and trailing non-helical residues were removed. After removing one more residue from either side so as to minimize contamination by flanking regions, segments longer than 17 amino acids were retained. Next, for duplicate polypeptide chains only those with the highest resolution were retained. The unique chains were clustered hierarchically using single linkage clustering as implemented in the treecluster routine of Biopython[68]. When extracting amino acid sequences, only residues with a residue numbering between 1 and 999 were retained to eliminate expression tags and soluble GPCR fusion partners. Distances were calculated from global sequence alignments for each pair of sequences, using the pairwise2 routine of Biopython with gap open and extension penalties of −10 and −0.5 and the BLOSUM62 substitution matrix[69]. Gaps at alignment ends were not penalized, and sequence identity (in %) was calculated over all experimentally observed residues and converted to distance as 1 − identity/100%. The 'tree' resulting from the clustering was cut and clades collapsed to clusters with representative polypeptide chains chosen from each cluster such that (i) cluster representatives had the minimum average distance to all other entries in the same cluster, and (ii) cluster representatives were at most 95% identical to each other. The cluster representatives then formed a balanced data set of 2,965 TM segments in 741 polypeptide chains and 546 PDB entries with reduced sequence bias.

To construct the human-TM data set, 20,245 reference sequences representing the human proteome were downloaded from UniProt[33] and submitted to the TMHMM Server v2.0[70] to computationally predict TM segments. This analysis identified 5,225 proteins with at least one TM segment of no less than 18 amino acids, 2,286 of which were single-pass proteins, and 18,959 such TM segments overall.

The observed amino acid frequencies in the TM segments of the two data sets were compared to their expected frequencies, which were calculated by shuffling the primary sequence of each protein in the data set, and counting the occurrence of a given amino acid in a fraction of the length of that randomized sequence that corresponds to the experimentally determined or computationally predicted TM content (see above). Statistics were accumulated per data set by shuffling each protein sequence 10,000 times, with the average count for a given amino acid approximating its expected value. Despite amino acid counts being discrete, they were found to be sufficiently well represented by a normal distribution, as manually inspected by quantile–quantile plots. Significance was estimated by comparing the observed count with the sample mean of the distribution of 10,000 randomized counts and integrating over the two more extreme tails of this distribution. Similarly, patterns of consecutive amino acids were analysed in both data sets following the outline of an earlier study[71]. Observed counts for patterns in TM segments were compared to expected counts given the distribution of individual amino acids in TM segments; that is, instead of randomizing the entire protein sequence as above, only amino acids within TM segments were shuffled, and instances of a given pattern were counted in this randomized set of (independent) TM segments. Statistics were accumulated and *P* values estimated as above.

We analysed the two data sets for the occurrence of the YPYP motif and, given the significantly lower abundance of proline residues in TM segments (observed fraction 0.51–0.66 of their expected number in human proteome and PDB, respectively, $P < 0.0001$), found this motif occurring in melatonin receptors to be statistically significant ($P < 0.025$), and not to exist in receptors other than the melatonin receptors.

Cheminformatics. Ligand data sets for a given receptor, identified by its UniProt[33] accession, were retrieved from ChEMBL24[23], and limited to data points with available binding ('B') data. Affinities were converted to pK_i, and for ligands with multiple reported affinity values, pK_i values were averaged. Substructure searches and matching were performed with rdkit (http://www.rdkit.org), and ligand protonation states estimated with OpenBabel[72] at pH 7.4.

Channel diameter determination. Channel dimensions were obtained using the CAVER analyst v2.0[73]. After addition of hydrogens to the coordinates of the MT_1–ramelteon crystal structure using PyMOL[74], a tunnel starting point was placed manually close to the perceived channel entrance. Using default program parameters, channel dimensions were extracted and trimmed to the segment between the channel entrance and ligand centroid. The residue proximity matrix was calculated for pore-lining residues within 5 Å of the respective probe position, with distances calculated as minima to any residue atom including hydrogens.

Molecular docking. MT_1 ligands obtained from the ChEMBL database[23] were docked into the ramelteon-bound crystal structure using energy-based docking in ICM-Pro v3.8-6[75]. Receptor structures were optimized using ICM docking protocols. Ligands selected for docking were converted from chemical structures to optimize their three-dimensional geometry and charges according to the MMFF force field[76]. Docking grid maps were generated using the receptor model, and ligands were docked using the biased probability Monte Carlo (BPMC) sampling and optimization method with a conformational sampling thoroughness of 50, starting with random initial ligand conformations. At least five independent runs were performed for the docking of the ligand set, and consistent dock poses were selected for further interaction analysis. The docking protocol did not use

distance restraints or any other a priori derived information to guide ligand–receptor interactions. The same procedure was applied to dock agomelatine to the 5-HT_{2C} structure[31] (PDB ID: 6BQG).

Molecular dynamics simulation. The experimental structure of MT_1 was pre-processed to assign protonation states and to model missing side chains using ICMFF energy-based optimization protocols available in the ICM-Pro molecular modelling package[75] (v3.8-6). The missing ICL3 coordinates were modelled using loop modelling and regularization protocols in ICM-Pro. The pre-processed structure was subjected to molecular dynamics simulation as previously described[77], using input files generated by the CHARMM-GUI web server[78]. For initial membrane embedding, the MT_1 models were superimposed to the 5-HT_{2B} (PDB ID: 4NC3) receptor structure[79] retrieved from the OPM database[65]. The structure was simulated in a periodic box with dimensions (x, y, z) of 75.5 Å, 75.5 Å, 105.4 Å, containing lipids (129 copies of the POPC lipid), 10,284 water molecules, 27 sodium ions, and 41 chloride ions. Parameters for melatonin and 2-PMT were obtained using CGenFF[80], available with the CHARMM-GUI server. After the initial energy minimization, the system was equilibrated for 10 ns, followed by production runs of up to 1 μs.

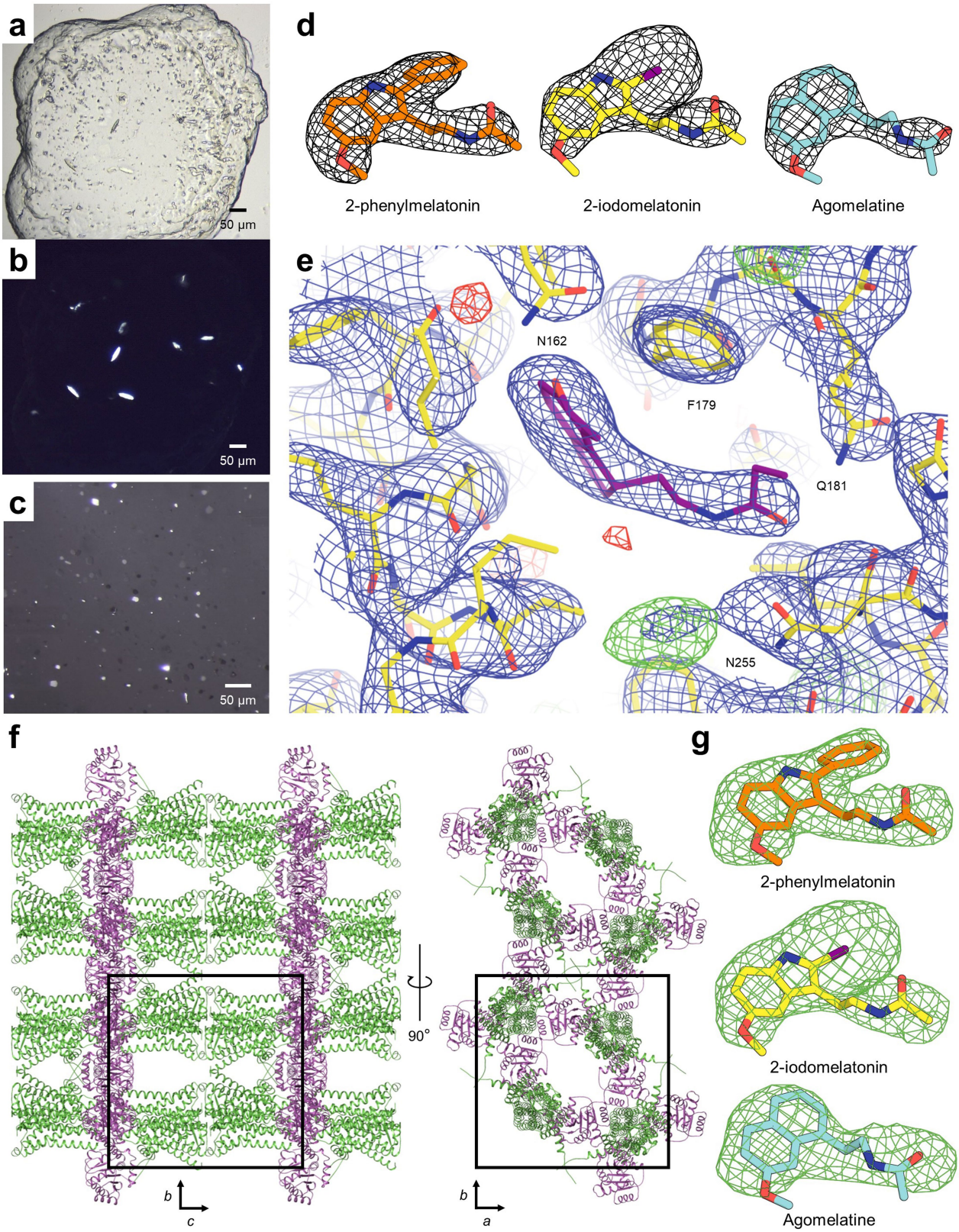
Membrane boundaries were derived as follows: for a set of 1,000 snapshots evenly spaced in time across the 1-μs trajectory, coordinates were superimposed to the crystal structure reference to eliminate receptor motion, the membrane being oriented along the x–y plane. Then, the positions of the carbonyl oxygen atoms of POPC molecules furthest from the lipid head groups were used to represent the boundaries of the hydrophobic slab. For each snapshot, this yielded a set of 129 oxygen atom coordinates, 64 (65) of which represent the upper (lower) leaf of the bilayer. Next, two planes were fit to these coordinates, representing the upper and lower membrane boundaries, respectively, using a least-squares method, minimizing the vertical (z-coordinate) distance from the points to the plane. The parameters of the resulting pair of planes were used to determine statistics over the course of the simulation, such as bilayer thickness and buriedness of the lateral ligand access channel below the extracellular membrane leaf, the position of which was defined as the mid-point between the smallest and largest z-coordinate of the side chains of channel-lining receptor residues $Y187^{5.38}$ and $H195^{5.46}$, respectively. This channel coordinate was found to fluctuate very little over the course of the simulation (s.d. = 0.4 Å). The channel entrance was found to lie 6.5 to 10.9 Å below the upper boundary of the hydrophobic slab (values from quartiles of the distance distribution; mean 8.7 ± 3.4 Å s.d., median 8.8 Å), and the average thickness of the hydrophobic slab was 30.0 ± 1.7 Å.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

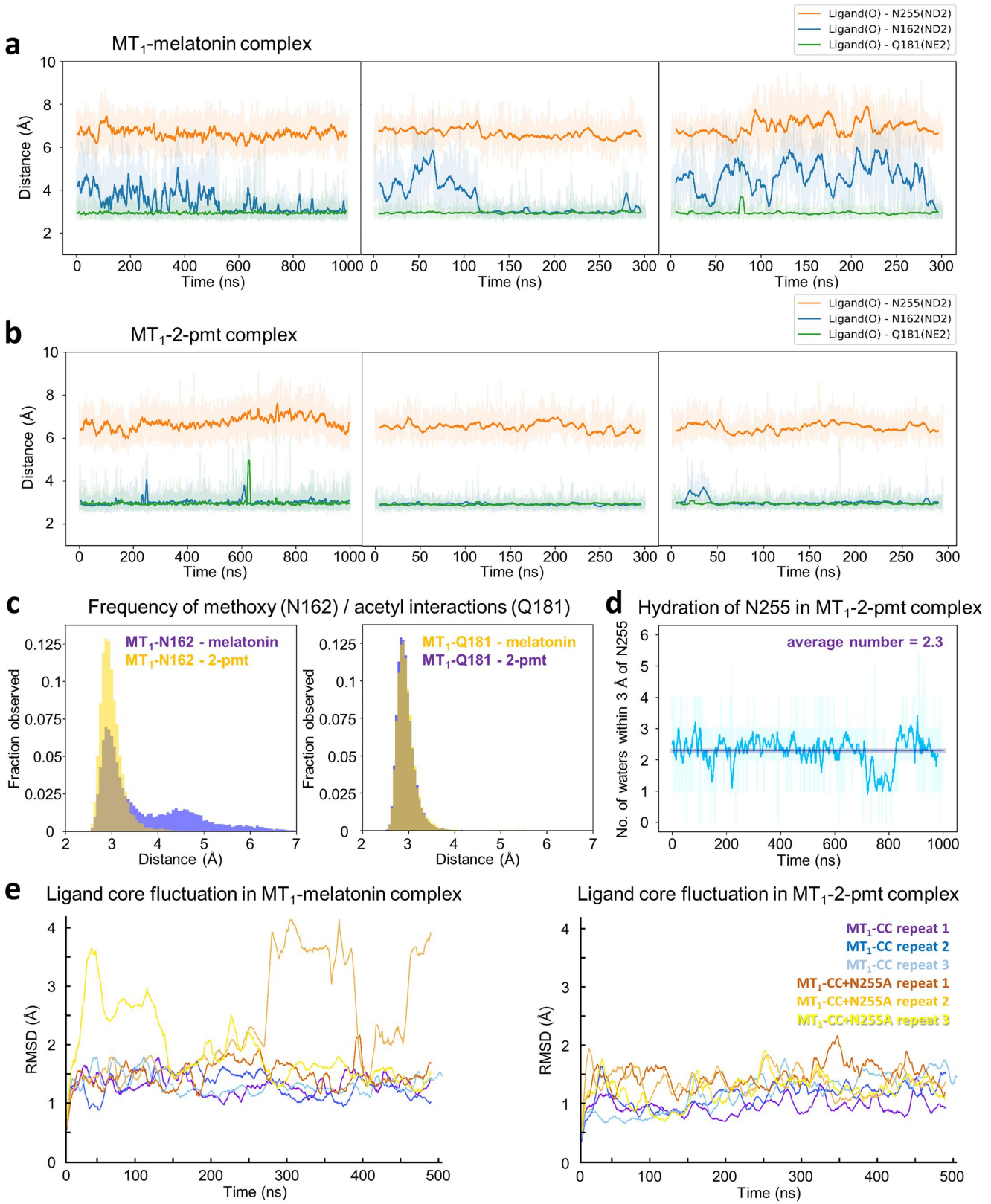
Coordinates and structure factors were deposited in the Protein Data Bank (PDB) under the following accession codes: 6ME2 (MT_1-CC–ramelteon), 6ME3 (MT_1-CC–2-PMT), 6ME4 (MT_1-CC–2-iodomelatonin), and 6ME5 (MT_1-CC–agomelatine).

33. The UniProt Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **45**, D158–D169 (2017).
34. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
35. Johnson, M. et al. NCBI BLAST: a better web interface. *Nucleic Acids Res.* **36**, W5–W9 (2008).
36. Papadopoulos, J. S. & Agarwala, R. COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinformatics* **23**, 1073–1079 (2007).
37. Isberg, V. et al. Generic GPCR residue numbers — aligning topology maps while minding the gaps. *Trends Pharmacol. Sci.* **36**, 22–31 (2015).
38. Kullback, S. & Leibler, R. A. On information and sufficiency. *Ann. Math. Stat.* **22**, 79–86 (1951).
39. White, J. F. et al. Structure of the agonist-bound neurotensin receptor. *Nature* **490**, 508–513 (2012).
40. Shibata, Y. et al. Thermostabilization of the neurotensin receptor NTS1. *J. Mol. Biol.* **390**, 262–277 (2009).
41. Roth, C. B., Hanson, M. A. & Stevens, R. C. Stabilization of the human β2-adrenergic receptor TM4–TM3–TM5 helix interface by mutagenesis of $Glu122^{3.41}$, a critical residue in GPCR structure. *J. Mol. Biol.* **376**, 1305–1319 (2008).
42. Bhattacharya, S., Hall, S. E. & Vaidehi, N. Agonist-induced conformational changes in bovine rhodopsin: insight into activation of G-protein-coupled receptors. *J. Mol. Biol.* **382**, 539–555 (2008).
43. Klco, J. M., Nikiforovich, G. V. & Baranski, T. J. Genetic analysis of the first and third extracellular loops of the C5a receptor reveals an essential WXFG motif in the first loop. *J. Biol. Chem.* **281**, 12010–12019 (2006).
44. Caffrey, M. & Cherezov, V. Crystallizing membrane proteins using lipidic mesophases. *Nat. Protocols* **4**, 706–731 (2009).
45. Liu, W., Ishchenko, A. & Cherezov, V. Preparation of microcrystals in lipidic cubic phase for serial femtosecond crystallography. *Nat. Protocols* **9**, 2123–2134 (2014).
46. Weierstall, U. et al. Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nat. Commun.* **5**, 3309 (2014).
47. Boutet, S. & Williams, G. J. The Coherent X-ray Imaging (CXI) instrument at the Linac Coherent Light Source (LCLS). *New J. Phys.* **12**, 035024 (2010).
48. Hart, P. et al. The Cornell-SLAC Pixel Array Detector at LCLS. *SLAC-PUB* https://ieeexplore.ieee.org/abstract/document/6551166 (2012).
49. Barty, A. et al. *Cheetah*: software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. *J. Appl. Crystallogr.* **47**, 1118–1131 (2014).
50. White, T. A. et al. CrystFEL: a software suite for snapshot serial crystallography. *J. Appl. Crystallogr.* **45**, 335–341 (2012).
51. Leslie, A. G. The integration of macromolecular diffraction data. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 48–57 (2006).
52. Duisenberg, A. Indexing in single-crystal diffractometry with an obstinate list of reflections. *J. Appl. Crystallogr.* **25**, 92–96 (1992).
53. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
54. White, T. A. et al. Recent developments in *CrystFEL*. *J. Appl. Crystallogr.* **49**, 680–689 (2016).
55. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
56. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).
57. *BUSTER v. 2.10.2* (Global Phasing Ltd., Cambridge, 2017).
58. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
59. Schüttelkopf, A. W. & van Aalten, D. M. PRODRG: a tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 1355–1363 (2004).
60. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).
61. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
62. Alexandrov, A. I., Mileni, M., Chien, E. Y., Hanson, M. A. & Stevens, R. C. Microscale fluorescent thermal stability assay for membrane proteins. *Structure* **16**, 351–359 (2008).
63. Depreux, P. et al. Synthesis and structure-activity relationships of novel naphthalenic and bioisosteric related amidic derivatives as melatonin receptor ligands. *J. Med. Chem.* **37**, 3231–3239 (1994).
64. Yous, S. et al. Novel naphthalenic ligands with high affinity for the melatonin receptor. *J. Med. Chem.* **35**, 1484–1486 (1992).
65. 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.* **40**, D370–D376 (2012).
66. Berman, H. M. et al. The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
67. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637 (1983).
68. Cock, P. J. et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422–1423 (2009).
69. Henikoff, S. & Henikoff, J. G. Performance evaluation of amino acid substitution matrices. *Proteins* **17**, 49–61 (1993).
70. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580 (2001).
71. Senes, A., Gerstein, M. & Engelman, D. M. Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with β-branched residues at neighboring positions. *J. Mol. Biol.* **296**, 921–936 (2000).
72. O'Boyle, N. M. et al. Open Babel: An open chemical toolbox. *J. Cheminform.* **3**, 33 (2011).
73. Kozlikova, B. et al. CAVER Analyst 1.0: graphic tool for interactive visualization and analysis of tunnels and channels in protein structures. *Bioinformatics* **30**, 2684–2685 (2014).
74. Schrödinger, L. L. C. The PyMOL Molecular Graphics System, Version 1.8 (2015).
75. Abagyan, R., Totrov, M. & Kuznetsov, D. ICM—A new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem.* **15**, 488–506 (1994).
76. Halgren, T. A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **17**, 490–519 (1996).
77. Zhang, H. et al. Structural basis for selectivity and diversity in angiotensin II receptors. *Nature* **544**, 327–332 (2017).
78. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).
79. Liu, W. et al. Serial femtosecond crystallography of G protein-coupled receptors. *Science* **342**, 1521–1524 (2013).
80. Vanommeslaeghe, K. et al. CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690 (2010).



Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Crystals, ligand electron density maps, and packing of MT_1. a, **b**, Bright field (**a**) and cross-polarized (**b**) images of representative MT_1–2-PMT crystals, optimized for synchrotron data collection (representing three independent crystallization setups). **c**, Cross-polarized image of representative MT_1–ramelteon crystals used for XFEL data collection (representing two independent crystallization setups). **d**, $2mF_o - DF_c$ ligand electron density maps of MT_1 co-crystallized with 2-PMT (orange), 2-iodomelatonin (yellow), and agomelatine (cyan), contoured at 1.0σ (grey mesh). **e**, $2mF_o - DF_c$ (blue, contoured at 1.0σ) and $mF_o - DF_c$ (green/red, $\pm 3.5\sigma$) electron density maps of MT_1–ramelteon (ligand purple, protein yellow) illustrating the small, unassigned electron density close to $N255^{6.52}$ that is tentatively attributed to the essential additive 2-propan-ol. The distance from this electron density to the closest ligand atom is approximately 4.8 Å. **f**, Packing of MT_1–PGS crystallized in the $P4\,2_1\,2$ space group. The receptor is shown in green and the PGS fusion protein is shown in purple. **g**, Simulated annealing $mF_o - DF_c$ omit maps (green mesh) of 2-PMT (orange sticks), 2-iodomelatonin (yellow), and agomelatine (cyan), contoured at 3.0σ.



Extended Data Fig. 2 | Molecular dynamics simulations. a, **b**, Distance plots for interactions between residues in MT_1 ($N162^{4.60}$, atom type ND2 (N^{δ}); $Q181^{ECL2}$, atom NE2 (N^{ε}); $N255^{6.52}$, atom ND2), and the closest oxygen atoms of the methoxy and acetyl groups, respectively, in the ligands melatonin (**a**) and 2-PMT (**b**) from three independent simulation runs. **c**, Distance histograms for interactions of methoxy with $N162^{4.60}$ (left), and $Q181^{ECL2}$ with the ligand acetyl tail (right), in melatonin and 2-PMT complexes. **d**, Hydration of residue $N255^{6.52}$ over the course of a 1-μs simulation of the MT_1–2-PMT complex from three independent simulations. **e**, Stability of ligand binding in simulations of MT_1 complexes. Time dependence of r.m.s.d. for non-hydrogen atoms of melatonin shown for MT_1–melatonin complex (left) and MT_1–2-PMT complex (right). Three independent simulations of crystal construct (purple, blue, light blue) and crystal construct with $N255^{6.52}$A mutation (orange, light orange, yellow) are shown, spanning 1.5 μs of cumulative time per system. Sampling rate was 10 frames per ns, and solid lines represent moving average values from 50 frames in all cases.

Extended Data Table 1 | MT_1 radioligand affinity

	melatonin	2-pmt	ramelteon	agomelatine	2-iodomelatonin
	K_d, nM [n] (pK_d ± s.e.m.)	K_i, nM [n] (pK_i ± s.e.m.)	K_i, nM [n] (pK_i ± s.e.m.)	K_i, nM [n] (pK_i ± s.e.m.)	K_i, nM [n] (pK_i ± s.e.m.)
WT	0.85 [11] (9.07±0.13)	0.42 [3] (9.46±0.18)	0.43 [4] (9.45±0.18)	1.67 [3] (8.79±0.08)	0.12 [3] (9.92±0.07)
WT + NaCl	2.46 [6] (8.61±0.11)	ND	ND	ND	ND
MT_1-CC (Sf9)	36.0 [7] (7.44±0.07)	7.41 [6] (8.13±0.09)	18.2 [3] (7.74±0.15)	29.5 [3] (7.53±0.01)	15.5 [3] (7.81±0.06)
MT_1-CC (Sf9) + NaCl	34.3 [5] (7.46±0.07)	ND	ND	ND	ND
MT_1-9mut	44.7 [7] (7.35±0.05)	9.70 [3] (8.01±0.24)	2.97 [3] (8.53±0.02)	14.9 [3] (7.83±0.04)	ND
MT_1-9mut + NaCl	32.4 [5] (7.49±0.03)	ND	ND	ND	ND
MT_1-PGS	1.62 [3] (8.79±0.05)	0.16 [4] (9.79±0.03)	ND	ND	ND
D73N (2.50)	195.0 [3] (6.71±0.06)	5.45 [3] (8.26±0.20)	ND	ND	ND
L95F (ECL1)	1.42 [3] (8.85±0.01)	0.24 [3] (9.63±0.10)	ND	ND	ND
G104A (3.29)	0.86 [3] (9.07±0.07)	0.12 [3] (9.91±0.08)	ND	ND	ND
F116W (3.41)	2.42 [3] (8.62±0.05)	0.37 [3] (9.43±0.06)	ND	ND	ND
N124D (3.49)	0.76 [3] (9.12±0.04)	0.09 [3] (10.03±0.12)	ND	ND	ND
C127L (3.52)	0.10 [3] (9.98±0.05)	0.04 [3] (10.40±0.03)	ND	ND	ND
W251F (6.48)	2.88 [3] (8.54±0.09)	0.06 [3] (10.23±0.17)	ND	ND	ND
A292P (7.50)	4.90 [3] (8.31±0.00)	1.27 [3] (8.90±0.07)	ND	ND	ND
N299D (8.47)	8.38 [3] (8.08±0.24)	3.29 [3] (8.48±0.21)	ND	ND	ND
H99A (3.24)	2.82 [3] (8.55±0.07)	0.28 [3] (9.55±0.05)	ND	ND	ND
N162A (4.60)	1.45 [3] (8.85±0.06)	0.20 [3] (9.71±0.06)	0.10 [3] (10.00±0.08)	0.74 [3] (9.14±0.07)	ND
F179A (ECL2)	*no specific binding up to 7 nM of radioligand*				
Q181A (ECL2)	1.80 [3] (8.85±0.23)	0.44 [3] (9.42±0.16)	0.36 [3] (9.58±0.26)	1.19 [3] (8.97±0.10)	ND
Q181E (ECL2)	1.85 [3] (8.73±0.03)	0.24 [3] (9.63±0.04)	0.28 [3] (9.57±0.10)	2.80 [3] (8.58±0.11)	ND
N255A (6.52)	2.04 [3] (8.79±0.25)	0.29 [3] (9.55±0.06)	0.43 [3] (9.39±0.11)	2.42 [3] (8.67±0.16)	ND
H195A (5.46)	*low expression / no specific binding up to 7 nM of radioligand*				
M107A (3.32)	0.65 [3] (9.23±0.14)	0.054 [3] (10.29±0.10)	0.08 [3] (10.15±0.13)	0.14 [3] (9.85±0.04)	ND

Affinity was determined using radioligand competition binding and [^{3}H]melatonin to yield K_d or K_i affinity estimates (number of independent experiments in square brackets) for wild-type (WT) and mutant receptors expressed in HEK293T cells. MT_1-CC, crystal construct (including PGS fusion) expressed in Sf9 cells; MT_1-9mut, construct combining all crystallogenic mutations, expressed in HEK293T cells. MT_1–PGS, wild-type construct with PGS fusion. ND, not determined. Binding isotherms are shown in Supplementary Fig. 1. To determine the effect of NaCl, binding assays were performed in the presence of 147 mM NaCl (binding isotherms in Supplementary Fig. 2).

Extended Data Table 2 | Functional data ($G_{i/o}$ Glosensor) for crystallogenic mutants

	melatonin			2-pmt		
	EC_{50}, nM [n] (pEC_{50})	**% E_{MAX} (% E_{MAX}*)**	**Δlog(E_{MAX} / EC_{50})**	**EC_{50}, nM [n] (pEC_{50})**	**% E_{MAX} (% E_{MAX}*)**	**Δlog(E_{MAX} / EC_{50})**
WT	0.0045 [23] (11.34±0.07)	100 (100)	0	0.0047 [16] (11.32±0.13)	100 (107±3)	0
MT_1-PGS	0.50 [12] (9.30±0.14)	42±2 (100)	-2.42	0.22 [12] (9.65±0.13)	39±2 (100±8)	-2.09
D73N (2.50)	11.78 [10] (7.93±0.07)	106±4 (100)	-3.39	0.5860 [10] (9.23±0.19)	110±3 (111±3)	-2.05
L95F (ECL1)	0.0042 [10] (11.37±0.23)	89±7 (100)	-0.02	0.0033 [6] (11.48±0.31)	92±5 (110±6)	+0.12
G104A (3.29)	0.0058 [7] (11.24±0.14)	100±4 (100)	-0.11	0.0046 [6] (11.34±0.19)	100±3 (107±3)	+0.02
F116W (3.41)	0.0069 [8] (11.16±0.13)	108±5 (100)	-0.15	0.0031 [7] (11.51±0.13)	103±3 (102±3)	+0.20
N124D (3.49)	0.0405 [7] (10.39±0.12)	104±9 (100)	-0.93	0.0068 [6] (11.17±0.14)	110±3 (117±3)	-0.11
C127L (3.52)	0.0111 [7] (10.95±0.19)	105±11 (100)	-0.37	0.0044 [6] (11.36±0.12)	115±5 (116±5)	+0.09
W251F (6.48)	0.0603 [7] (10.22±0.11)	99±7 (100)	-1.13	0.0041 [6] (11.38±0.11)	95±5 (104±5)	+0.04
A292P (7.50)	0.0187 [7] (10.73±0.07)	114±6 (100)	-0.56	0.0030 [6] (11.52±0.16)	111±5 (105±5)	+0.24
N299D (8.47)	0.0644 [7] (10.19±0.12)	107±4 (100)	-1.12	0.0299 [6] (10.53±0.20)	98±5 (100±5)	-0.81
MT_1-9mut	*no activity*					

Data were acquired with wild-type MT_1 and crystallogenic mutants by using GloSensor to measure $G_{i/o}$-mediated cAMP inhibition via isoproterenol stimulation. Data represent means of EC_{50} (number of independent experiments in square brackets), represented as mean pEC_{50} ± s.e.m. as well as E_{MAX} shown as mean $\%E_{MAX}$ ± s.e.m. $\%E$max is relative to wild-type receptor in columns, and $\%E$max* is relative to melatonin in rows. Mutant effects were calculated by the change in relative activity, or log(E_{MAX}/EC_{50}), subtracting wild-type from mutant activity. Dose–response curves are shown in Supplementary Fig. 3.

Extended Data Table 3 | MT_1 crystallographic data collection and refinement statistics

	MT_1-CC-ramelteon[a]	MT_1-CC-2-pmt[b]	MT_1-CC-2-iodomelatonin[c]	MT_1-CC-agomelatine[d]
Data collection				
Space group	P4 2_1 2	P4 2_1 2	P4 2_1 2	P4 2_1 2
Cell dimensions				
a, *b*, *c* (Å)	122.3, 122.3, 122.8	122.4, 122.4,122.8	122.6, 122.6, 123.3	122.2, 122.2,122.8
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	29.7-2.80 (2.92-2.80)	30.6-2.90 (3.04-2.90)	28.9-3.20 (3.42-3.20)	28.9-3.20 (3.42-3.20)
R_{split}	0.100 (3.56)	0.098 (2.96)	0.150 (2.10)	0.120 (1.67)
$I/\sigma I$	7.9 (0.33)	9.4 (0.39)	6.5 (0.57)	8.7 (0.70)
$CC_{1/2}$ (%)	99.93 (56.0)	99.95 (54.5)	99.87 (60.4)	99.85 (54.5)
Completeness (%)	100 (100)	100 (100)	100 (100)	100 (100)
Redundancy	1,003 (187.4)	2,323 (489.3)	595 (85.8)	1,131 (151.6)
Refinement				
Resolution (Å)	29.67-2.80	30.00-2.90	28.91-3.20	28.87-3.20
No. reflections	23,262	21,300	15,957	15,853
R_{work}/R_{free}	0.204/0.230	0.203/0.229	0.209/0.249	0.232/0.257
No. atoms				
Protein	3,738	3,753	3,702	3,668
Ligand	19	23	18	18
Lipid and other	40	19	16	9
B-factors (Å^2)				
Receptor	114.4	113.4	133.0	117.9
PGS	104.1	99.7	124.5	109.7
Ligand	101.9	94.3	126.4	105.3
Lipids and other	107.8	106.9	118.5	110.0
R.m.s. deviations				
Bond lengths (Å)	0.008	0.009	0.009	0.009
Bond angles (°)	0.94	0.98	0.95	0.93

Number of crystals used for structure determination: [a]46,679, [b]99,897, [c]21,038, and [d]42,423. Values in parentheses are for highest-resolution shell.

Extended Data Table 4 | Volumes of enclosed binding sites of class A GPCRs

receptor	PDB ID	ligand	Volume, Å^3
MT_1	6ME2	ramelteon	710
MT_1	6ME3	2-pmt	710
MT_1	6ME4	2-iodomelatonin	714
MT_1	6ME5	agomelatine	702
CB_1	5XRA	CHEMBL1683648	928
CB_1	5U09	taranabant	1043
S1P1	3V2W	sphingosine mimetic	936
FFA1	5TZR	MK-8666	378*
LPA1	4Z35	ONO-9910539	904
rhodopsin (active)	2X72	all-*trans* retinal	894
rhodopsin (inactive)	1U19	11-*cis* retinal	598

*Large part of ligand is outside binding pocket.

Extended Data Table 5 | Thermostability data

Mutant	apo T_m, °C	mlt T_m, °C	2-pmt T_m, °C	rmt, T_m, °C	mlt ΔT_m, °C	2-pmt ΔT_m, °C	rmt ΔT_m, °C
MT_1-CC	59.4±0.5	67.3±0.3	75.6±0.1	73.9±0.2	+7.9	+16.2	+14.5
N162A (4.60)	59.9±0.8	67.1±0.8	ND	73.3±0.8	+7.2	ND	+13.4
F179A (ECL2)	55.1±0.6	55.2±0.5	ND	59.9±0.1	+0.1	ND	+4.8
Q181A (ECL2)	60.8±0.4	65.0±0.2	72.6±0.1	74.0±0.4	+4.2	+11.8	+13.2
N255A (6.52)	59.6±0.6	60.3±0.4	66.8±0.3	64.6±0.4	+0.7	+7.2	+5.0
Q181A/N255A	61.7±0.2	62.1±0.4	ND	64.9±0.4	+0.4	ND	+3.2
M107A (3.32)	55.3±1.1	62.8±0.8	ND	69.4±1.0	+7.5	ND	+14.1
M107D (3.32)	ND	61.7±0.3	ND	67.9±0.6	ND	ND	ND
G108A (3.33)	60.8±0.5	61.0±0.6	ND	65.0±0.1	+0.2	ND	+4.2
F194A (5.45)	57.9±0.2	65.6±0.1	ND	73.3±0.3	+7.7	ND	+15.4
H195A (5.46)	56.7±1.5	66.2±0.2	ND	71.8±0.2	+9.5	ND	+15.1
F196A (5.47)	52.9±0.1	58.2±0.3	ND	66.5±0.1	+5.3	ND	+13.6
Y79A (2.56)	ND	55.7±0.5	ND	63.3±0.7	ND	ND	ND
Y79F (2.56)	ND	61.0±0.1	ND	68.6±0.2	ND	ND	ND
P80A (2.57)	ND	59.8±0.8	ND	67.2±0.8	ND	ND	ND
Y81A (2.58)	ND	60.3±0.1	ND	67.4±0.5	ND	ND	ND
Y81F (2.58)	55.8±0.1	65.3±0.1	ND	72.5±0.1	+9.6	ND	+16.8
P82A (2.59)	ND	57.8±0.2	ND	65.1±0.4	ND	ND	ND
H99L (3.24)	ND	59.7±0.1	ND	67.9±0.2	ND	ND	ND

Melting temperature (T_m, mean ± s.d. for $n = 3$ independent experiments) for crystallized construct (MT_1-CC) and several of its mutants purified in the absence (apo) and presence (100 μM) of ligand (mlt, melatonin; 2-PMT, 2-phenylmelatonin; rmt, ramelteon). ND, not determined. Melting curves are shown in Supplementary Fig. 4.

Extended Data Table 6 | Functional data (G$_{i/o}$ GloSensor) for mutants of the YPYP motif and the ligand binding site

	melatonin			2-pmt			ramelteon			agomelatine		
	EC$_{50}$, nM [n] (pEC$_{50}$)	**%E$_{MAX}$ (%E$_{MAX}$*)**	**Δlog (E$_{MAX}$/ EC$_{50}$)**	**EC$_{50}$, nM [n] (pEC$_{50}$)**	**%E$_{MAX}$ (%E$_{MAX}$*)**	**Δlog (E$_{MAX}$/ EC$_{50}$)**	**EC$_{50}$, nM [n] (pEC$_{50}$)**	**%E$_{MAX}$ (%E$_{MAX}$*)**	**Δlog (E$_{MAX}$/ EC$_{50}$)**	**EC$_{50}$, nM [n] (pEC$_{50}$)**	**%E$_{MAX}$ (%E$_{MAX}$*)**	**Δlog (E$_{MAX}$/ EC$_{50}$)**
WT	0.0045 [23] (11.34±0.07)	100 (100)	0	0.0047 [16] (11.32±0.13)	100 (107±3)	0	0.0035 [12] (11.45±0.15)	100 (98±3)	0	0.0084 [10] (11.07±0.11)	100 (107±5)	---
Y79A (2.56)	0.2380 [4] (9.62±0.14)	121±17 (100)	-1.64	0.0166 [4] (10.78±0.10)	81±2 (71±1)	-0.64	0.0339 [4] (10.47±0.09)	88±2 (71±1)	-1.04	2.3988 [4] (8.62±0.09)	81±2 (71±1)	-2.55
Y79F (2.56)	0.0537 [4] (10.27±0.08)	98±2 (100)	-1.08	0.0052 [4] (11.28±0.07)	92±2 (100±2)	-0.08	0.0060 [4] (11.22±0.07)	100±2 (100±2)	-0.23	0.3802 [4] (9.42±0.07)	92±2 (100±2)	-1.69
P80A (2.57)	0.3020 [4] (9.52±0.10)	79±2 (100)	-1.92	0.0316 [4] (10.50±0.09)	74±2 (100±2)	-0.95	0.0251 [4] (10.60±0.08)	81±2 (100±2)	-0.95	0.3715 [4] (9.43±0.10)	74±2 (100±2)	-1.77
Y81A (2.58)	0.0347 [4] (10.46±0.17)	107±2 (100)	-0.85	0.0012 [4] (11.91±0.06)	100±2 (100±2)	+0.59	0.0048 [4] (11.32±0.06)	109±2 (100±2)	-0.10	0.0891 [4] (10.05±0.07)	100±2 (100±2)	-1.02
P82A (2.59)	0.1514 [4] (9.82±0.06)	57±2 (100)	-1.77	0.0269 [4] (10.57±0.10)	53±2 (100±3)	-1.03	0.0575 [4] (10.24±0.10)	58±2 (100±3)	-1.45	0.2042 [4] (9.69±0.12)	53±2 (100±3)	-1.66
H99A (3.24)	0.0034 [4] (11.47±0.08)	91±2 (100)	+0.09	0.0014 [4] (11.85±0.09)	86±2 (100±2)	+0.46	0.0011 [4] (11.96±0.11)	93±2 (100±2)	+0.48	0.0040 [4] (11.40±0.10)	85±2 (100±2)	+0.26
N162A (4.60)							*no activity*					
F179A (ECL2)	5957 [5] (5.23±0.07)	124±2 (100)	-6.02	748 [5] (6.13±0.05)	118±0 (96±3)	-5.12	501 [5] (6.30±0.08)	125±2 (89±4)	-5.06	1159 [5] (5.94±0.11)	118±2 (97±3)	-5.06
Q181A (ECL2)							*no activity*					
Q181E (ECL2)	1.2882 [5] (8.89±0.07)	131±2 (100)	-2.34	0.0461 [5] (10.34±0.09)	116±3 (95±3)	-0.92	0.4831 [5] (9.32±0.08)	123±4 (92±3)	-2.05	5.3951 [5] (8.27±0.06)	118±3 (96±3)	-2.74
H195A (5.46)							*low expression*					
N255A (6.52)	0.2366 [5] (9.63±0.16)	134±3 (100)	-1.59	0.0152 [5] (10.82±0.19)	118±2 (94±1)	-0.44	0.0396 [5] (10.40±0.18)	116±4 (85±3)	-0.99	1.8707 [5] (8.73±0.06)	111±5 (88±4)	-2.30

Data were acquired with wild-type MT$_1$ and mutants by using GloSensor to measure G$_{i/o}$-mediated cAMP inhibition via isoproterenol stimulation. Data represent means of EC$_{50}$ (number of independent experiments in square brackets), shown as mean pEC$_{50}$ ± s.e.m. as well as E_{MAX}, shown as mean %E_{MAX} ± s.e.m. %E_{MAX} is relative to wild-type receptor in columns, and %E_{MAX}* is relative to melatonin in rows. Mutant effects were calculated by the change in relative activity, or log(E_{MAX}/EC$_{50}$) subtracting wild-type from mutant activity. ND, not determined. Dose–response curves are shown in Supplementary Fig. 5.

Extended Data Table 7 | Functional data ($G_{i/o}$ GloSensor) for mutants of the lateral channel

	melatonin			ramelteon			bitopic ligand CTL 01-05-B-A05		
	EC_{50}, nM [n] (pEC_{50})	**$\%E_{MAX}$ ($\%E_{MAX}$*)**	**Δlog (E_{MAX} / EC_{50})**	**EC_{50}, nM [n] (pEC_{50})**	**$\%E_{MAX}$ ($\%E_{MAX}$*)**	**Δlog (E_{MAX} / EC_{50})**	**EC_{50}, nM [n] (pEC_{50})**	**$\%E_{MAX}$ ($\%E_{MAX}$*)**	**Δlog (E_{MAX} / EC_{50})**
WT	0.0045 [23] (11.34±0.07)	100 (100)	0	0.0035 [12] (11.45±0.15)	100 (98±2)	0	0.69 [4] (9.16±0.09)	100 (124±3)	0
A158M (4.56)			*no activity*				ND		
A190F (5.41)	0.0141 [9] (10.85±0.20)	110±4 (100)	-0.45	0.0084 [6] (11.07±0.43)	119±7 (106±6)	-0.31	5.75 [4] (8.24±0.13)	107±3 (120±3)	-0.89

Data were acquired with wild-type MT_1 and mutants by using GloSensor to measure $G_{i/o}$-mediated cAMP inhibition via isoproterenol stimulation. Data represent means of EC_{50} (number of independent experiments in square brackets), represented as mean pEC_{50} ± s.e.m. as well as E_{MAX}, shown as mean $\%E_{MAX}$ ± s.e.m. $\%E_{MAX}$ is relative to wild-type receptor in columns, and $\%E_{MAX}$* is relative to melatonin activity in rows. Mutant effects were calculated by the change in relative activity, or $\log(E_{MAX}/EC_{50})$ subtracting wild-type from mutant activity. ND, not determined. Dose–response curves are shown in Supplementary Fig. 6.