The Serotonin G-Protein Coupled Receptor 5HT2A: Molecular and Multiple Sequence Analysis

Jorge E. Chalit^{1,*}, Emily J. Ragan^{1,**}.

1. Department of Chemistry & Biochemistry, Metropolitan State University of Denver, Denver, CO, 80217-3362 USA. *e-mail: jorgechalitresearch@gmail.com, **e-mail: eragan@msudenver.edu

ABSTRACT: The serotonin receptor is an integral part of human biochemistry, especially in regard to neuropsychiatric illness, but is also found across animal species. There is a need for a deeper molecular understanding of the serotonin receptor in regard to novel treatments in psychiatry. The present review and subsequent analysis provide detailed insight regarding the function, composition, and molecular interactions of the 5HT2A receptor and its associated G-protein complex. A multiple sequence analysis of the 5HT2A and G α q proteins among four vertebrate species was performed to elucidate protein conservation across taxa. The sequence alignment demonstrated regions of high conservation across the entire length of the G α q protein sequences and in specific regions of the 5HT2A receptor, including five alpha helices that interact with the G α q protein. Chimera was used to examine the location and molecular level interactions of key 5HT2A and G α q residues in the cryoEM structure of 5HT2A with mini-G α q and beta/gamma (PDB 6WHA). This analysis provided additional insight into interactions around residues known to be important for function and allowed the identification of additional residues predicted to lead to anomalous function if mutated.

1. Introduction

Serotonin, formally called 5-hydroxytryptamine (5-HT), is colloquially termed "the happy molecule," due to its relation to mood.¹ The neurotransmitter molecule undergoes a complex signaling pathway through several G-protein coupled receptors (GPCRs), of which the 5-HT receptor family includes seven: 5-HT1 through 5-HT7.² Each receptor binds serotonin, which subsequently triggers a signaling cascade throughout the cell. Each GPCR is composed of a transmembrane domain made of seven alpha helix proteins that interact with the G-protein, which consists of three subunits: alpha, beta, and gamma (Figure 1). There are several different G-protein alpha subunits; the 5HT2A receptor in particular interacts with the Gaq isotype. In the serotonin pathway, the G-protein complex is activated by the binding of the signaling hormone, serotonin (Figure 1). A conformational change in structure induces binding of GTP in place of GDP in the alpha subunit and its separation from the adjacent beta and gamma units.³ Once activated, the single Gaq unit couples to phospholipase C, which catalyzes the formation of inositol triphosphate (IP3) and diacylglycerol (DAG) from phospholipid phosphatidylinositol-4,5-bisphosphate (PIP2).^{3,4} IP3 stimulates the release of Ca²⁺, while DAG stimulates protein kinase C (PKC). Both PKC and Ca²⁺ aid with cellular neurotransmitter release.³ Without regulation, the signaling molecules will continue to produce their downstream effects, so the cell contains inherent elements that turn signaling off. One of these regulatory elements is arrestin, which binds the phosphorylated cytoplasmic tail of the receptor protein, thereby preventing further interaction with the G-protein. Other notable mechanisms of down-regulation include GTP hydrolysis into GDP in the Gaq subunit, dephosphorylation of IP3, deacylation of DAG, and the removal of Ca²⁺ from the cytoplasm.³



Figure 1: The GPCR signaling pathway during serotonin signaling. **(1)** Inactive receptor and G protein complex prior to signaling. **(2)** Serotonin binds the receptor and initiates the signaling cascade. **(3)** Gaq binds GTP and activates PLC. **(4)** The Gaq-PLC complex hydrolyzes PIP2 into DAG and IP3; IP3 stimulates Ca²⁺ release. **(5)** Activated PKC and Ca²⁺ stimulate cellular effects, including neurotransmitter release. **(6)** The complex returns to the inactive conformation.

Dysfunctions in these serotonin-triggered GPCR signaling pathways are theorized to lead to mood-related disorders.¹ The various serotonin receptors modulate behaviors such as anxiety, aggression, risk assessment, cooperation, and depressive behavior, as well as other functions like wakefulness, food appetence, memory, and motor function.⁵ There are several receptors linked to mood regulation within the serotonin family, which include 5HT1A, 5HT1B, 5HT2A, 5HT2B, and 5HT2C.^{1,6} The

current literature agrees that 5HT2A is one of the receptors that plays a modulatory role in neuropsychiatric illnesses.⁷ Standard treatments for affective disorders include selective serotonin reuptake inhibitor (SSRI) therapy, though SSRIs do not target the receptor directly. SSRIs elicit their therapeutic effects by limiting the reuptake of serotonin at the synapse and thereby increasing the serotonin available to interact with the receptor.⁶ However, clinical efficacy of SSRIs can vary between patients,⁸ and some studies show that a sizeable portion of individuals still don't achieve remission despite repeated treatments.9 Newer studies are focusing on psychedelic treatments for mental illness, including the use of compounds like LSD and psilocybin.^{10,11} Most psychedelic compounds, including LSD and psilocybin, rely on the 5HT2A receptor for their effects.^{12,13} Å recently published set of structures (Figure 2) illustrated the amino acid composition and interactions of the activated 5HT2A receptor bound to the exogenous ligand 25CN-NBOH, a psychedelic compound analog, and endogenous serotonin.¹⁴

The 5HTR2A gene is most highly expressed in the brain's frontal cortex and anterior cingulate cortex.¹⁵ The anterior cingulate cortex is known as the emotional center of the brain, so the receptors in that brain region are key components of mood regulation, albeit alongside other contributory regions in the limbic system.¹⁶ It's worth noting that the receptor also shows relatively high expression in the aortic, coronary, and tibial arteries, which suggests that serotonin is also involved in cardio-vascular physiology. Interestingly, recent science shows that the 5HT2A receptor indeed may play a role in peripheral diseases such as hypertension and atherosclerosis, necessitating an integrative approach to medical intervention for serotonin-based ailments.¹⁷ Indeed, there must be more to 5HT2A than just a role in mood, and investigating 5HT2A in model organisms could be helpful in further exploring its function and role.

Despite much 5HT2A research centering on human mental illness, the serotonin molecule is known to modulate both vertebrate and invertebrate nervous systems.⁵ 5HT2A receptors diverged approximately 450-500 million years ago and are found across gnathostomata, a diverse group of vertebrates that include fish, amphibians, reptiles, and mammals.^{18,19} Likewise, the Gaq protein is part of both human and animal physiology through its presence in vertebrates and invertebrates.^{5,20} Given that protein sequence determines structure and function, proteins that have similar sequences are likely to elicit the same function even in dissimilar species. In 2007, Raote, Bhattacharya, and Panicker performed a sequence analysis for the 5HT2A protein between humans, rats, and mice.⁷ They elucidated extensive similarities between the species, with few mutations in the protein sequence, most of which would have little to no effect on the receptor's function. However, comparing mammalian species is more likely to reveal similarities in sequence identity given the relatedness of the species. Therefore, it may be more useful to compare human 5HT2A and Gaq with more distant orthologs.

A comparison of multiple vertebrate 5HT2A receptors and $G\alpha q$ sequences will be helpful in assessing the conservation of these sequences, including where they interact with each other, across taxa. In general, sequence components with higher conservation are likely to be more important for structure and function of the protein. This paper compares 5HT2A and $G\alpha q$ protein sequences from four different vertebrate species, including a fish, in order to identify amino acids conserved over evolutionary time. This study integrates findings from this cross-



Figure 2: The 5HT2A receptor bound to the G α q-protein complex (PDB 6WHA). The 5HT2A protein is highlighted in cyan, the G α q protein in yellow, the G β protein in green, and the G γ protein in orange. Agonist 25CN-NBOH (gray) is bound to 5HT2A at the active site.

species analysis with the recent structure of the human 5HT2A receptor and its interaction with the human $G\alpha q$ protein.¹⁴

2. Materials & Methods

2.1. BLAST & Alignments

The FASTA-format protein sequences from human 5HT2A were entered into Basic Local Alignment Search Tool (BLAST) at NCBI to perform a thorough search of homologous protein species.²¹ The reference proteins database was employed with all other parameters on their respective default settings; the species Homo sapiens (humans) was excluded from the organism search set. The BLAST search revealed numerous species for comparison, of which three evolutionarily diverse vertebrates were selected: *Pelodiscus sinensis* (Chinese softshell turtles), Corvus moneduloides (New Caledonian crows), and Carassius auratus (goldfish). The species were selected based on their wide range (54%-90%) in percent identity match with the H. sapiens sequence in order to analyze protein sequence conservation broadly. The Gaq protein sequence underwent an identical BLAST search, and the same species were selected for comparison. The FASTA-format sequences for the homologs were then entered into Clustal Omega and MView to perform a sequence alignment comparison.^{22,23} For Clustal Omega, the default settings were used except for the "order" parameter was set to "input." The results from Clustal Omega were then copied directly into MView and all default parameters applied.

2.2 3D Structure Analysis

When researchers solve the structure of a protein using X-ray crystallography, nuclear magnetic resonance, or cryogenic electron microscopy (cryoEM), they place the structure coordinates in the protein data bank (PDB) so other researchers can use the structure and perform investigations and visualizations of the atomic level structure. In order to perform a thorough investigation of the 5HT2A receptor's biochemical structure, an analysis of the structure of the 5HT2A and mini-G α q protein coupled during *in vitro* binding conditions (PDB file 6WHA) was conducted using the UCSF Chimera Extensible Molecular

Modeling System.^{14,24} The mini-G α q is an engineered G α q that is shorter than, but reliably models, the wild-type G α q.^{25,26} The three-dimensional viewing model allowed for visualization of key amino acid residues related to protein function. The amino acids were analyzed in terms of hydrogen bonding and hydrophobic interactions with their respective counterparts based on distances calculated in Chimera, and a fundamental understanding of chemical interactions. Information about the structural and functional role of regions of the human 5HT2A receptor and G protein, including a mini-G α q, was obtained from a recent cryo-EM paper and subsequently related to the other vertebrate sequences in a multiple sequence alignment.¹⁴ Regions with clear significance in the human structure were a priority in the analysis.

3. Results & Discussion

The multiple sequence alignment shown in Figure 3 compares the 5HT2A receptor amino acid sequence between humans, Chinese softshell turtles, New Caledonian crows, and goldfish. The areas of lowest interspecies conservation are in the first ~100 amino acids of the protein and the last ~40 residues in the C terminus, although the very last seven residues are well conserved. Kim et. al determined several key residues in the 5HT2A protein that are important in ligand activation of the Gprotein, including 1181, L325, and N384.¹⁴ These three residues, all located in the cytosolic end that interacts directly with the Gaq-protein, are completely conserved across the species.

	cov	pid	1		30
1 Human5HT2A	100.0%	-	-		
2 Turtle5HT2A		75.9%		MYTKGYNFLFFHCTSAPVSDIAMDSLC-DEENPTVNSLI-OLNHERFYDNVFMPGEVNTSHLFNLTMDS	
3 Crow5HT2A	98.78	75.78			
4 Goldfish5HT2A	95.8%	54.28		MNLYANISKPVPKSIFTSSNLMPFDPDTWTPSPDIMLGNVVRNSSFGCRSWADT-EASLFPNVSISSTSSEK	
	cov	pid	81	. 1 N107	160
1 Human5HT2A	100.0%	100.0%		- <mark>Enrenls<mark>c</mark>eg<mark>els</mark>ps<mark>c-</mark>lsllhl<u>geknms</u>alltavviiltitagnilvimavslekk<mark>l()n</mark> entflmslaindmllgplv</mark>	
2 Turtle5HT2A	99.2%	75.9%		-ENLTNLS <mark>C</mark> DGRVTPP <mark>C</mark> HS <u>ElfHL</u> SEKNMPALLTVIVITLTIAGNILVIMAVSLEKKL(MNTNYFLMSLAINDMLLGFLV	
3 Crow5HT2A	98.7%			-DULTNISCESNMESICCPSONNMPALLTMIVIIITIAONILVIMAVELEKKI(NATNIPLMELAIADHLLOPLV	
4 Goldfish5HT2A	95.8%	54.2%		TEPLSHORCHERETEREELVRANHALLILVVVIIAVNL PRICH PRICH PRICH PRICH PRICH	
				1181 R185	
1 Human5HT2A	cov	pid 100.0%	161	2	240
2 Turtle5HT2A		75.9%			
3 Crow5HT2A	98.78			NOVEMENT AND A DETERMINED OF A STATE OF TAKEN A THEORY IN THE STATE AND	
4 Goldfish5HT2A		54.28		MPV SMLTILIGYTM PLPT <mark>KLCAV</mark> HIYLDVLPSTA SIMBLCAIS JDRY AIRNI I EBSRIMIRTKAP <mark>A</mark> KI IAVWTI SVGI S MPV SMLTILIGYTM PLP <mark>RKLCA</mark> I WIYLDVLPSTA SIMBLCAIS IDRY AIRNI I EBSRIMIRTKAPAKI IAVWTI SVGI S MPV SHVTIVGYSM PLPASL <mark>CP</mark> P WIYLDVLPSTA SIMBL <mark>C</mark> AIS IDRY VAIRNI I TEVRINIRTKAPAKI TAVWTI S <mark>A</mark> GI S	
	cov	pid	241	L261	320
1 Human5HT2A		100.0%		NPIPVFGLQDDS <mark>KVFK</mark> <u>BC</u> SCLLADDNFVLIGSFVSFFIPLTINVITTFLTIKSL <mark>NK</mark> BAT <mark>CVS</mark> DLGT <mark>RAK</mark> LAS	
2 Turtle5HT2A	99.2%			NPVPVEGLQDDSKVFXKCS <mark>CLLADZS</mark> PVLICSPVAFFIPLTINVVTFLTIKSLVKSAT.CINDLCKKTAFAS NPIPVEGLQDDSKVFXKDIFKKDI <mark>CLLAD</mark> DNFVLVCSPVAFFIPLTINVVTFLTIKSLVKSAN.CVNDVGPKTAFAS	
3 Crow5HT2A 4 Goldfish5HT2A	98.78	75.78		MPIPY GLODDSKY Y KOIFKKDICLLADSKYVLYG SFYRFIPIT MYVTYLTIK LYKER CYNDYGPKTYRS MPIPYLGLRDETKYFKDCSCLITDNSFYLIGSFYRFYDIIIMYVTYFLIISALYSEAI.CLDCLYPRYMSTGF	
outerstates	,,,,,,,				
				L325	
	COV	-	321	-FSFLPOSSUSS-EK-DFORSTHREPGSYTGRRINGSUSNEOKACKUDIVFPUEVWCPFFITNIMAVIC	100
1 Human5HT2A 2 Turtle5HT2A		100.0%		-FSFLPQSSLSS-EK-LFQRSIEREPGSYTGRRENQSISNEQKACKULJIVFF.FVVNNCPFFITNIMAVIC -FSFFPQSSLSS-EK-LFQRSANRDLRNSGRRENQSISNEQKASKULJIVFF.FVVNNCPFFITNVMAVIC	
3 Crow5HT2A		75.78		-PSFLPQSSISSEK-LFQRSINRDAGTSGRRINGSISNEQRASKVLJIVFF.FVVHHCPFFIINVHAVIC	
4 Goldfish5HT2A		54.28		TIN PL PGPS: SPSEKKLFLR3SLSREVGAESGVVMPPFGENNVGSISNEORASKVL IVVFF FVVMCPFFITNVLAVVC	
	cov	pid	401	N384	180
1 Human5HT2A		100.0%		RESONEDVIGALLAVFVMIGY.SSAVAPLVYT.FMKTYRSAFIRYIQCOYKENKK PLOLILVATIPALAYKSSOLONGO	
2 Turtle5HT2A	99.28			NK-SONERVIGULIN FUNIGELSSAUN PLUTELEN KEERSAFIIRET OCREKK PLOLILUNTIPALAK SSOLOTAO	
3 Crow5HT2A	98.78	75.78		XE-S <mark>C</mark> KEEVIGGLLNVFVNIGYLSSAVNPLVYTLFNKTYRSAFIRY I Q <mark>C</mark> RYKEEKKFFQLILVNTIPALAYDSSQLQLAQ	
4 Goldfish5HT2A	95.8%	54.2%		EPATCDADVMNSLENVEVNVCXLSSAVNPLVXTLINCTYRSAFARTARCCFB8BKKPLQLILVNTIPPLATOSTELPLTG	
	cov	nid	481	A447 H452	
1 Human5HT2A		100.0%		KKNSKQDAKTTONDCS KVAL-GKCHSEEASKDNSDGVNEKVSCV	
2 Turtle5HT2A		75.9%		MASLAKEA MMAKDISMVAI-GIEPLOSISKGSISPVSELVSCV	
3 Crow5HT2A		75.78		M.SL.KEA MMAKDYST TI-GTERLDGTSKSSIGPGNEK SGV	
4 Goldfish5HT2A	95.8%	54.2%		SIGNGD SLFLPNKKELLSKSSKNESKSEL	

Figure 3: MView alignment of 5HT2A amino acid sequences for human, Chinese softshell turtles, New Caledonian Crows, and goldfish. "Cov" indicates coverage - the percent that a species matches the human protein in terms of amino acid sequence length. "Pid" indicates the percent identity between individual amino acids. Dots are 10 amino acids apart from one another; each line contains 80 amino acids, including dashes for gaps in the alignment. The magenta rectangles denote important amino acids related to function that were analyzed in this study using human sequence numbering. The blue rectangles indicate the area of low conservation in the receptor protein. The green rectangles outline the five regions in the alpha helices that interact with the Gαq protein. The red rectangle outlines the transmembrane region that is completely conserved. See Table 1 for an explanation of the amino acid color schemes.

Table 1. MSA Amino Acid Color Scheme

BRIGHT GREEN	Hydrophobic
DARK GREEN	Large hydrophobic
DARK BLUE	Negative charge
LIGHT BLUE	Small alcohol
YELLOW	Cysteine
PURPLE	Polar
RED	Positive charge

This table shows the appropriate color coding for the amino acid residues in each multiple sequence alignment. Note that in the alignments, there are residues that are not colored due to their mismatch with the primary human sequence.

Kim et al. showed that the variations I181A or I181E led to weaker activation by the ligands 25CN-NBOH, LSD, or serotonin.¹⁴ The mutations L325A and N384A reduced receptor-Gaq coupling.¹⁴ Given the weaker overall 5HT2A receptor interaction with Gaq when I181, L325, and N384 were mutated in the Kim et. al study,¹⁴ along with their conservation across taxa, it's reasonable to conclude that they're important for 5HT2A receptor function.

The 5HT2A alignment also shows high conservation in other regions that are important for binding of serotonin and other serotonergic molecules like LSD and 25CN-NBOH. Kim et al. identified residues N107 and R185 as having a crucial role in Gaq signaling. Both of those amino acids are also completely conserved across the analyzed species. There is a notable region of high conservation between T253 and Q262, with the only difference being in goldfish at positions K259 and S260. Within that region, L261 is of particular interest as it resides in a hydrophobic pocket that Kim et al. described as critical for binding Gaq (Figure 4).¹⁴ The high amount of conservation at critical amino acids between the aforementioned species indicates that the 5HT2A receptor is well-conserved protein that has retained key aspects of structure and function over ~400 million years of evolutionary time. Analyzing residue composition using multiple sequence analyses provides a useful tool for



Figure 4: Key hydrophobic region between 5HT2A (cyan) and Gaq (yellow). All four highlighted amino acid side chains are completely conserved residues across the four species investigated (L261 and A321 of 5HT2A, L236 and L240 of the mini-Gaq/L349 and L353 in the full Gaq sequence).

predicting other potentially critical amino acids in the 5HT2A protein.

Another interesting region of complete conservation across all four species in the 5HT2A protein is in an alpha helix from W151 to Y174 (Figure 5A). This region is not discussed in the Kim et al. paper, but their cryo-EM structure allowed for visualization of its 3D structure in the human protein and analysis of the biochemical interactions involved. Figure 5A was generated using Chimera and PDB 6WHA to highlight the intramembranous region that is completely conserved between species. The helix from W151 to Q178 is buried in the middle of five other membrane spanning helices. Residue S159 forms two hydrogen bonds with the ligand 25CN-NBOH (Figure 5B). This well-conserved helix ends in an intracellular loop with residues N179 through N187, which interact with G α q, suggesting that the helix is important for linking ligand binding with G α q activation.

The percent identity and coverage between sequences were higher when analyzing the G α q-protein than the 5HT2A receptor protein, suggesting even higher G α q-protein conservation across taxa. Figure 6 shows the multiple sequence alignment for G α q proteins in humans, Chinese softshell turtles, New Caledonian crows, and goldfish. There is greater than 90% overall identity from goldfish to humans in the G α q (Figure 6), whereas the 5-HT2A protein only shows 54.2% overall identity between goldfish and humans (Figure 3). The G α qs in all species contain 359 amino acids, whereas the 5HT2A proteins differ in sequence length (471 amino acids in humans and crows, 490 in turtles, 491 in goldfish).

 $G\alpha q$ amino acid Q350 is completely conserved between all four species and corresponds to the mini-G αq Q237 that lost its signaling capacity when experimentally mutated to alanine.¹⁴ The G αq protein interacts with the 5HT2A protein along its alpha unit amino acids K345-V359 (corresponding to K233-V247 in the cryo-EM structure), which are all completely conserved in the species investigated. It's worth noting that the five alpha helical regions from the 5HT2A protein involved in the interaction with the G αq unit are highly conserved between the four species but show more variation than seen in the G αq protein (Figure 3). Higher variation in residue composition likely means fewer functional constraints.



Figure 5: A) The region highlighted in purple corresponds to the highly conserved transmembrane helix in the 5HT2A protein sequence. Residues W151 (top) and Y174 (bottom) are also shown. Residue W151 is closer to the extracellular side. **B)** This figure provides a focused view of the highly conserved transmembrane helix at the S159 position, which is a residue that is also highly conserved. S159 forms two hydrogen bonds with the ligand 25CN-NBOH.

1 Human 2 Turtle 3 Crow 4 Goldfish	cov pid 1
1 Human 2 Turtle 3 Crow 4 Goldfish	cov pid 81 . 1 . . . 160 100.0% 100.0% QNIFTAMQANIRANDZLKIPYKYEHNKAHAQLVREVDVEKVSAFENPYVDAIKSUNDPGIQECYDRREVQLSDSTKYY . <t< td=""></t<>
1 Human 2 Turtle 3 Crow 4 Goldfish	cov pid 161 2 240 100.0% 100.0% INDLDRVADPAYLPTQQDVLRVRVPTTGILEYPFDLQSVIFRMVDVGCQRSERRKHIECFENVTSIMFLVALSETDQVLV 100.0% 98.9% INDLDRVADPAYLPTQQDVLRVRVPTTGILEYPFDLQSVIFRMVDVGCQRSERRKHIECFENVTSIMFLVALSETDQVLV 100.0% 97.8% INDLDRVADPAYLPTQQDVLRVRVPTTGILEYPFDLQSVIFRMVDVGCQRSERRKHIECFENVTSIMFLVALSETDQVLV 100.0% 97.8% INDLDRVADPAYLPTQQDVLRVRVPTTGILEYPFDLQSVIFRMVDVGCQRSERRKHIECFENVTSIMFLVALSETDQVLV 100.0% 92.2% INSLORTAGESVVTQQDVLRVRVPTTGILEYPFDLQSVIFRMVDVGGQRSERRKHIECFENVTSIMFLVALSETDQVLV
1 Human 2 Turtle 3 Crow 4 Goldfish	cov pid 241 : .
1 Human 2 Turtle 3 Crow 4 Goldfish	cov pid 321 i i 359 100.0% 100.0% DKI IYSHFTCATDIENTRFYFAN KDTIIQ NIKEYNLV 100.0% 98.9% DKI IYSHFTCATDIENTRFYFAN KDTIIQ NIKEYNLV 100.0% 97.8% DKI IYSHFTCATDIENTRFYFAN KDTIIQ NIKEYNLV 100.0% 92.2% EXIIYSHFTCATDIENTRFYFAN KDTIIQ IKEYNLV

Figure 6: MView alignment for $G\alpha q$ amino sequences in human, Chinese softshell turtles, New Caledonian Crows, and goldfish. "Cov" indicates coverage - the percent that a species matches the human protein in terms of amino acid sequence length. "Pid" indicates the percent identity between individual amino acids. Dots are 10 amino acids apart from one another; each line contains 80 amino acids. The magenta rectangles denote the crucial glutamine that is involved in a critical interaction with the 5HT2A receptor protein and the N352 that is not conserved compared with goldfish. The blue rectangle outlines the general $G\alpha q$ domain that interacts with the 5HT2A receptor. The green rectangle outlines the region of the $G\alpha q$ that interacts with the G β domain. See Table 1 for an explanation of the color schemes.

One region in $G\alpha q$ that has variation is human, turtle and crow residue N352 (corresponding to N239 in the mini-G α q), which is a threonine in goldfish. Figure 7 shows that N239 of the mini-Gaq is in proximity to H183 of the 5HT2A protein, though not within range to contribute a meaningful electrostatic interaction. A threonine at the same position would likely exhibit similar biochemical behavior, so the difference between goldfish and the other species is predicted to be minimal at this location. Goldfish are not typically used as model organisms, but a similar species, zebrafish, are. The zebrafish 5HT2A receptor shares 92% sequence identity with the goldfish receptor. Zebrafish may be a viable candidate for future studies focused on the 5HT2A receptor. For example, there is current research using zebrafish as a model for human 5HT2A-mediated diseases such as rhabdomyolysis.²⁷ A cursory analysis of zebrafish compared with humans demonstrates 55.7% sequence identity (Supplementary Figure 1). Given the findings of the present study, identifying regions of high conservation between zebrafish and humans may point to viable candidates for functional studies.

The conservation of 5HT2A and G α q across multiple species underscores their involvement in essential biological processes. It's worth noting that a mutation in the G α q would affect more than just mood dysregulation. As previously mentioned, new studies show concomitant somatic illness with psychiatric disturbances.¹⁷ The G α q is also involved in the epinephrine and glucagon signaling pathways, albeit via slightly different mechanisms. Epinephrine acts on adrenergic receptors throughout the body, inducing vascular smooth muscle contraction, pupil dilation, heart rate elevation, and other physiologic effects.²⁸ Glucagon is involved in hepatic gluconeogenesis and fat metabolism, among others.²⁹ Given the importance of those two hormones in the body, and their mediation by GPCRs, it's reasonable to propose that mutations in the G α q would lead to other negative somatic effects.



Figure 7: N239 (left) of the mini-G α q (yellow) protein is shown next to H183 (right) of the 5HT2A (cyan) protein. The measured distance of 4.207 Å between the sidechains is outside of the range for hydrogen-bonding. Altering the structure via conformational change could place the amino acids in hydrogen-bonding proximity.

Beyond its interaction with multiple receptors, another likely reason for high sequence conservation in G α q is its interaction with the adjacent G β subunit. The amino acids closest to the Nterminus on the mini-G α q interact with its β domain; the alphahelix produced by residues V5-R32 of the mini-G α q subunit (corresponding to L11-R38 of the G α q human protein) interacts with the beta-sheet formed by residues K78-L95 of the G β unit (Figure 8A). Specifically, D20 and R15 of the mini-G α q (corresponding to E26 and I21 of the human G α q) form hydrogen bonds with K89 and the backbone at V90 of the G β domain, respectively (Figure 8B). The regions in the G α q protein interacting with the G β subunit also have high interspecies conservation, as illustrated in Figure 6. Given that the G α q unit interacts with both the 5HT2A and the G β domain, it makes sense that the protein is well-conserved across homologous taxa.

Regions of low conservation may not play an essential role in function. Harvey et al. analyzed human genetic variations in the 5HT2A protein sequence and their respective differences in responding to antipsychotics and endogenous serotonin signaling.³⁰ The researchers found that the mutations T25N, A447V, and H452Y in the 5HT2A receptor did not significantly alter serotonin's signaling potential. T25 is near the N-terminus of



Figure 8: A) mini-G α q (yellow) alpha helix is shown with hydrogen bonds interacting with the beta-sheet of the G β subunit (green). **B)** The two key interactions between the beta strand and the alpha-helix are the hydrogen bonds between mini-G α q D20/G β K89 (top blue lines between strand and helix) and mini-G α q R15/G β V90 backbone (bottom blue lines between strand and helix).

the 5HT2A sequence (Figure 3), in a very poorly conserved region across the four species investigated. Likewise, A447 and H452 are in a poorly conserved region near the C-terminus (Figure 3). While in this example mutational evidence preceded the analysis, it highlights the potential benefit of using multiple sequence analyses to determine regions that likely need to be conserved for function or that lack conservation and may be less critical for function.

As demonstrated here, regions that are highly conserved between distant homologs tend to be associated with regions that are important in function. Although not definitive, these predictions can help identify amino acids in which missense mutations would be deleterious. Future molecular and genetic studies would benefit from first inspecting a protein of interest using multiple sequence analysis to perform a better-guided structural investigation. A diverse range of organisms, such as goldfish, turtles, and crows here, provided useful insight into homologous 5HT2A and Gaq proteins, including identification of a highly conserved helix in 5HT2A that may link ligand binding with a change in interaction with the Gaq.

Acknowledgements

Molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311. This paper evolved out of an MSU Denver Biochemistry II course bioinformatics project supported by NSF IOS 1656407 to Dr. Emily Ragan. Thank you to the two anonymous reviewers as well as Dr. Megan Filbin for providing helpful feedback during the review process.

References

(1) Ruhé, H. G.; Mason, N. S.; Schene, A. H. Mood Is Indirectly Related to Serotonin, Norepinephrine and Dopamine Levels in Humans: A Meta-Analysis of Monoamine Depletion Studies. *Mol. Psychiatry* **2007**, *l* 2(4),331–359. https://doi.org/10.1038/sj.mp.4001949.

(2) Westkaemper, R.B. Structure and Function Reveal Insights in the Pharmacology of 5-HT Receptor Subtypes. In *The Serotonin Receptors: From Molecular Pharmacology to Human Therapeutics*; Roth, B. L., Ed.; Humana Press, 2006. https://doi.org/10.1007/978-1-59745-080-5.

(3) Lounsbury, K. *Signal Transduction and Second Messengers*, 1st ed.; Elsevier Inc., 2009. https://doi.org/10.1016/B978-0-12-369521-5.00006-3.

(4) Hoyer, D.; Clarke, D. E.; Fozard, J. R.; Hartig, P. R.; Martin, G. R.; Mylecharane, E. J.; Saxena, P. R.; Humphrey, P. P. A. VII. International Union of Pharmacology Classification of Receptors for 5-Hydroxytryptamine (Serotonin). *Pharmacol. Rev.* **1994**, *46* (2), 157–203.

(5) Bacqué-cazenave, J.; Bharatiya, R.; Barrière, G.; Delbecque, J. P.; Bouguiyoud, N.; Di Giovanni, G.; Cattaert, D.; De Deurwaerdère, P. Serotonin in Animal Cognition and Behavior. *Int. J. Mol.* Sci. **2020**, 21 (5), 1–23. https://doi.org/10.3390/ijms21051649.

(6) Fakhoury, M. Revisiting the Serotonin Hypothesis: Implications for Major Depressive Disorders. *Mol. Neurobiol.* **2016**, *53* (5), 2778–2786. https://doi.org/10.1007/s12035-015-9152-z.

(7) Raote, I.; Bhattacharya, A.; Panicker, M. M. Serotonin 2A (5-HT2A) Receptor Function: Ligand-Dependent Mechanisms and Pathways. In *Serotonin Receptors in Neurobiology*; Chattopadh-yay, A., Ed.; CRC Press/Taylor & Francis: Boca Raton, FL, 2007.

(8) Peciña, M.; Karp, J. F.; Mathew, S.; Todtenkopf, M. S.; Ehrich, E. W.; Zubieta, J. K. Endogenous Opioid System Dysregulation in Depression: Implications for New Therapeutic Approaches. *Mol. Psychiatry* **2019**, *24* (4), 576–587. https://doi.org/10.1038/s41380-018-0117-2.

(9) Rush, J. A.; Trivedi, M. H.; Wisniewski, S. R.; Nierenberg, A. A.; Stewart, J. W.; Warden, D.; Niederehe, G.; Thase, M. E.; Lavori, P. W.; Lebowitz, B. D.; McGrath, P. J.; Rosenbaum, J. F.; Sackeim, H. A.; Kupfer, D. J.; Luther, J.; Maurizio Fava, M. STAR-D (2006; AjPsych) Tiered Approach for Depression. *Am J Psychiatry* **2006**, *163* (11), 1905–1917.

(10) Fuentes, J.J; Fonseca, F.; Elices, M.; Farré, M.; Torrens, M. Therapeutics Use of LSD in Psychiatry: A Systematic Review of Randomized-Controlled Clinical Trial. *Front. Psychiatry* **2020**, *10* (January), 1-14. https://doi.org/10.3389/fpsyt.2019.00943.

(11) Davis, A. K.; Barrett, F. S.; May, D. G.; Cosimano, M. P.; Sepeda, N. D.; Johnson, M. W.; Finan, P. H.; Griffiths, R. R. Effects of Psilocybin-Assisted Therapy on Major Depressive Disorder: A Randomized Clinical Trial. *JAMA Psychiatry* **2020**, 1–9. https://doi.org/10.1001/jamapsychiatry.2020.3285.

(12) Kometer, M.; Schmidt, A.; Jäncke, L.; Vollenweider, F. X. Activation of Serotonin 2A Receptors Underlies the Psilocybin-Induced Effects on α Oscillations, N170 Visual-Evoked Potentials,

and Visual Hallucinations. J. Neurosci. **2013**, 33 (25), 10544–10551. https://doi.org/10.1523/JNEUROSCI.3007-12.2013.

(13) López-Giménez, J. F.; González-Maseo, J. Hallucinogens and Serotonin 5-HT2A Receptor-Mediated Signaling Pathways. Curr. Top. Behav. Neurosci. **2018**, No. 36, 45–73. https://doi.org/10.1007/7854.

(14) Kim, K.; Che, T.; Panova, O.; DiBerto, J. F.; Lyu, J.; Krumm, B. E.; Wacker, D.; Robertson, M. J.; Seven, A. B.; Nichols, D. E.; Shoichet, B. K.; Skiniotis, G.; Roth, B. L. Structure of a Hallucinogen-Activated Gq-Coupled 5-HT2A Serotonin Receptor. Cell **2020**, 182 (6), 1574-1588.e1-e7. https://doi.org/10.1016/j.cell.2020.08.024.

(15) Aguet, F.; Barbeira, A.; Bonazzola, R.; Brown, A.; Castel, S.; Jo, B.; Kasela, S.; Kim-Hellmuth, S.; Liang, Y.; Oliva, M.; Parsana, P.; Flynn, E.; Frésard, L.; Gamazon, E.; Hamel, A.; He, Y.; Hormozdiari, F.; Mohammadi, P.; Muñoz-Aguirre, M.; Lappalainen, T. The GTEx Consortium Atlas of Genetic Regulatory Effects across Human Tissues; 2019. https://doi.org/10.1101/787903

(16) Sapolsky, R. M. Feeling Someone's Pain, Understanding Someone's Pain, Alleviating Someone's Pain. In Behave: The Biology of Humans at Our Best and Worst. Penguin Press: New York, 2017; 528-531.

(17) Nichols, C. D. Serotonin 5-HT2A Receptor Function as a Contributing Factor to Both Neuropsychiatric and Cardiovascular Diseases. Cardiovasc. Psychiatry Neurol. **2009**, 2009, 1–8. https://doi.org/10.1155/2009/475108.

(18) Peroutka, S.J.; Howell, T.A. The Molecular Evolution of G Protein-Coupled Receptors: Focus on 5-Hydroxytryptamine Receptors. Neuropharmacology **1994**, 33 (3–4), 319–324. https://doi.org/10.1016/0028-3908(94)90060-4.

(19) Moutkine, I.; Collins, E. L.; Béchade, C.; Maroteaux, L. Evolutionary Considerations on 5-HT2 Receptors. Pharmacological Research **2019**, 140, 14–20. https://doi.org/10.1016/j.phrs.2018.09.014.

(20) Strathmann, M.; Simon, M. I. G Protein Diversity: A Distinct Class of Alpha Subunits Is Present in Vertebrates and Invertebrates. Proc Natl Acad Sci U S A **1990**, 87 (23), 9113–9117. https://doi.org/10.1073/pnas.87.23.9113.

(21) Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. (1990) "Basic local alignment search tool." J. Mol. Biol. 215:403-410.

(22) Madeira, F.; Park, Y. M.; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A. R. N.; Potter, S. C.; Finn, R. D.; Lopez, R. The EMBL-EBI Search and Sequence Analysis Tools APIs in 2019. Nucleic Acids Res. **2019**, 47 (W1), W636–W641. https://doi.org/10.1093/nar/gkz268.

(23) Brown N.P.; Leroy C.; Sander C. (1998) MView: a web-compatible database search or multiple alignment viewer. Bioinformatics 14(4):380-1. Pu-Med: 9632837 DOI: 10.1093/bioinformatics/14.4.380

(24) UCSF Chimera--a visualization system for exploratory research and analysis. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. J Comput Chem. 2004 Oct;25(13):1605-12. (25) Nehmé, R.; Carpenter, B.; Singhal, A.; Strege, A.; Edwards, P. C.; White, C. F.; Du, H.; Grisshammer, R.; Tate, C. G. Mini-G Proteins: Novel Tools for Studying GPCRs in Their Active Conformation The First Mini-G Protein Developed Was Mini-G s. Here We Extend the Family of Mini-G Proteins to Include Mini-G Olf, Mini-G I1, Mini-G O1 and the Chimeras Mini-G s/q and Mini-G. PLoS One **2017**, 1–26.

(26) Wan, Q.; Okashah, N.; Inoue, A.; Nehme, R.; Carpenter, B.; Tate, C. G.; Lambert, N. A. Mini G Protein Probes for Active G Protein–Coupled Receptors (GPCRs) in Live Cells. J. Biol. Chem. **2018**, 293 (19), 7466–7473. https://doi.org/10.1074/jbc.RA118.001975.

(27) Kawahara, G.; Maeda, H.; Kikura-Hanajiri, R.; Yoshida, K. ichi; Hayashi, Y. K. The Psychoactive Drug 25B-NBOMe Recapitulates Rhabdomyolysis in Zebrafish Larvae. Forensic Toxicol. **2017**, 35 (2), 369–375. https://doi.org/10.1007/s11419-017-0366-9.

(28) Dalal R.; Grujic D. Epinephrine, 2020. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK482160/ (Accessed 01/01/2021).

(29) Habegger, K. M.; Heppner, K. M.; Geary, N.; Bartness, T. J.; DiMarchi, R.; Tschöp, M. H. The Metabolic Actions of Glucagon Revisited. Nat. Rev. Endocrinol. **2010**, 6 (12), 689–697. https://doi.org/10.1038/nrendo.2010.187.

(30) Harvey, L.; Reid, R. E.; Ma, C.; Knight, P. J. K.; Pfeifer, T. A.; Grigliatti, T. A. Human Genetic Variations in the 5HT2A Receptor: A Single Nucleotide Polymorphism Identified with Altered Response to Clozapine. Pharmacogenetics **2003**, 13 (2), 107–118. https://doi.org/10.1097/00008571-200302000-00007.



Supplemental Figure 1: MView alignment for 5HT2A protein of zebrafish and humans. "Cov" indicates coverage - the percent that a species matches the human protein in terms of amino acid sequence length. "Pid" indicates the percent identity between individual amino acids. See Table 1 for an explanation of the color schemes. Dots are 10 amino acids apart from one another; each line contains 80 amino acids, including dashes for gaps in the alignment.