The organizing unit of GPCR activation-switch architecture: a phylogenetic-scale census

Huazhang Shen[1]

[1] Independent researcher

Corresponding author: Huazhang Shen (hzshen@gmail.com) · ORCID 0009-0006-8208-3111

Abstract

The intracellular "ionic lock" between Arg3.50 of the D(E)RY motif and an acidic residue at position 6.30 is a textbook feature of class A G-protein-coupled receptor (GPCR) activation. Across an 11,609-sequence cross-species survey, the complete classic lock is present in only about a quarter of receptors — consistent with prior reports that ~34% of class A receptors carry both partners (Kling et al. 2016) and that the 6.30 acidic residue is conserved in only 25-31% of sequences (Hauser et al. 2021), both of which already flag the mechanism as receptor-dependent. We classify Arg3.50-protection mechanisms genome-wide into six categories and, using a cross-validated multivariate model, find that ligand family adds no predictive power for mechanism beyond receptor gene identity — the same mechanism recurs across aminergic, peptide, and nucleotide families. We then show that the three components most implicated in the activation switch — the 6.30 lock, Tyr5.58, and the sodium-pocket residues — coevolve as a coupled network (Pagel pairwise tests, all $p < 10^{-17}$; a joint eight-state corHMM model favoring full coupling over independence, ΔAICc = 439 on the full gene tree). This signal persists at reduced effect size (ΔAICc = 220; pairwise $p < 10^{-7}$) when restricted to true multi-species orthologous groups excluding within-species paralogs, and is robust to leave-one-clade-out removal. Finally, precise heavy-atom distance measurements on experimentally determined structures show that only two of five candidate receptors (PAR2, AGTR2) have a genuine salt-bridge-range intra-TM6 acidic residue near Arg6.30, both in engineered rather than wild-type constructs — a candidate proximity configuration, not a demonstrated stabilizing interaction. Together these results argue that activation-switch architecture is organized not by ligand identity but by gene-and-lineage-specific, phylogenetically coupled local solutions.

Keywords

G protein-coupled receptor; ionic lock; Arg3.50; DRY motif; phylogenetic comparative methods; coevolution; corHMM; molecular evolution; activation switch; orthogroup

Introduction

The activation of a class A GPCR is classically described in terms of a small set of conserved microswitches, chief among them the “ionic lock”: a salt bridge tethering the Arg of the D(E)RY motif at the cytoplasmic end of TM3 to an acidic residue at position 6.30, at the intracellular tip of TM6. Breaking this contact is a hallmark of the inactive-to-active transition, and the lock features prominently in textbook cartoons of receptor activation.

Yet the lock is, in a strict sense, the exception rather than the rule, and this is not a new observation. Systematic residue-level surveys spanning nearly two decades have repeatedly found that the complete contact is present in only a minority of class A receptors: an early structural study of CCR5 noted that roughly 30% of class A receptors, including the entire chemokine subfamily, carry a positively charged residue at 6.30 rather than the canonical acidic partner (Springael et al. 2007); site-directed mutagenesis of the histamine H4 receptor later showed its high constitutive activity is explained by a substitute Asp5.69-Arg6.31 salt bridge rather than a missing lock being simply absent (Schneider et al. 2010); Kling et al. (2016) report that just 34% of class A GPCRs carry both an arginine at 3.50 and a glutamate at 6.30, and explicitly note that “the exact function of Arg3.50 is likely to be receptor-dependent and must be characterized independently for every GPCR”; and Hauser et al. (2021), in a superfamily-wide alignment, find the 6.30 acidic residue conserved in only 25-31% of sequences — far below the 65-86% conservation of the reference position 4.49. The rarity of the classic lock, and the expectation that its replacement is receptor-specific, are therefore already established.

What neither of these studies does is (a) systematically classify the alternative Arg3.50-protection strategies into discrete structural categories across the full class A sequence space, (b) test directly — with a predictive model rather than aggregate entropy statistics — whether these categories track ligand family or a finer unit of receptor identity, or (c) test whether the individual structural components implicated in the activation switch (the 6.30 lock, Tyr5.58, the sodium-pocket residues) coevolve as a coupled system once the phylogenetic non-independence of related sequences is corrected for. We address these three gaps directly. The field’s working taxonomy classifies receptors by ligand and implicitly assumes that mechanism tracks ligand; here we test that assumption with a cross-validated multivariate model at phylogenetic scale, and find that it does not hold — ligand family adds no predictive power for mechanism beyond gene identity, and the organizing unit of activation-switch architecture is instead a coevolving network of switch residues that tracks receptor lineage independently of what the receptor binds.

We build the argument in four steps. First, a genome-wide census of Arg3.50-protection mechanisms establishes both the rarity of the classic lock and, via a direct predictive test, the fact that the alternatives do not map onto ligand family. Second, an orthogroup-resolved re-analysis confirms that the organizing unit descends below the ligand label to individual genes and lineages, and reports honestly where this re-analysis weakens the original signal. Third, phylogenetic comparative modeling shows that the switch components coevolve as a coupled network, a result we further test for robustness to paralog non-independence by restricting to true multi-species orthologous groups. Fourth, structural analysis of solved structures, using precise heavy-atom distance measurements, documents where a candidate intra-TM6 configuration near Arg6.30 does and does not reach genuine salt-bridge range.

Results

1. The classic ionic lock is a minority mechanism, and its alternatives do not map to ligand family

We classified the Arg3.50-protection mechanism of each of 11,609 cross-species class A sequences into six categories on the basis of the residues occupying the DRY motif and the 6.30 position and its structural neighborhood (Methods). The classic cross-helical lock — a charged residue at 6.30 forming an intracellular salt bridge with the DRY arginine — accounts for only **25.6%** of the library. The remainder distribute across alternative arrangements, the two largest being a neutral 6.30 with a nearby local acidic residue (**42.4%**) and a same-charge 6.30 with an intra-helical rescue partner (**23.0%**) (Figure 1).

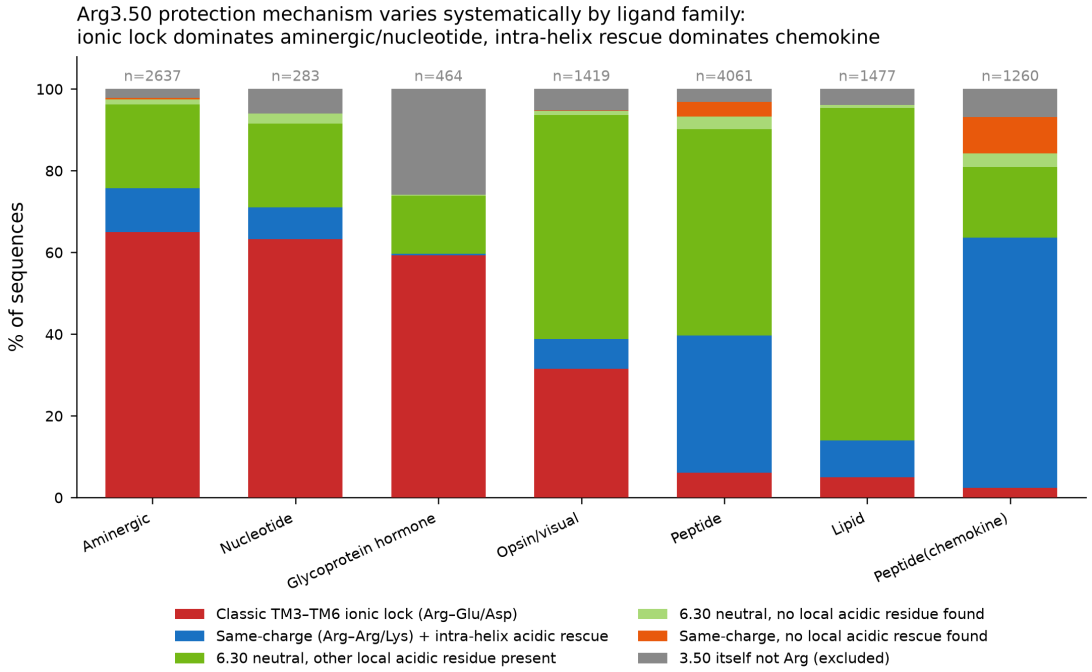

Figure 1. Mechanism-by-family. Six-category Arg3.50-protection mechanism distribution, split by ligand family; establishes the missing-lock puzzle and the many-to-many structure.

This 25.6% figure is directly comparable in construction — a *complete-combination* proportion, not a single-residue conservation rate — to Kling et al.'s 34% co-occurrence of Arg3.50 and Glu6.30. The two are drawn from different samples (the GMOS database versus our 11,609-sequence cross-species set) and so are not identical, but they agree in direction and magnitude, and both establish that the intact lock is present in only a minority of receptors.

The central observation of this section is that these mechanisms do not partition by ligand. The relationship between ligand family and Arg3.50-protection mechanism is many-to-many: a single mechanism recurs across aminergic, glycoprotein-hormone, and nucleotide families, and conversely a single family contains multiple mechanisms (Figure 2). The peptide-receptor family shows the highest mechanism entropy in the library (1.60 bits), yet when the same sequences are resolved to the level of individual genes, 58.3% of genes are

≥90% consistent across species. In other words, the apparent mechanistic diversity of a ligand family is an aggregation artifact: the unit at which mechanism is actually determined lies below the family, at the gene and lineage.

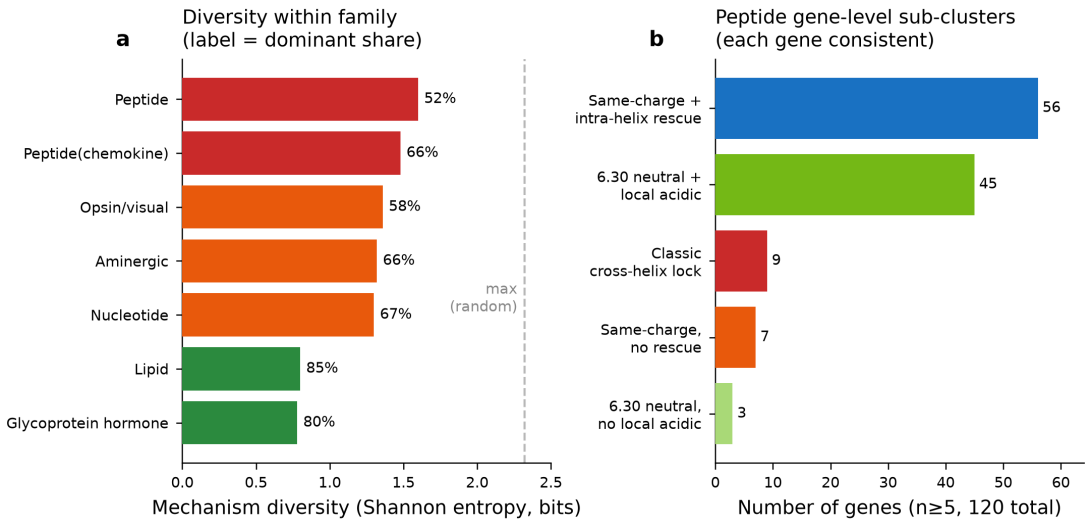

Figure 2. Mechanism-diversity-within-family. Per-family mechanism entropy (panel a) and gene-level consistency (panel b).

We tested this directly (new in this revision) with a cross-validated multivariate classification model on n=11,059 sequences (excluding 550 with a non-Arg 3.50): 5-fold stratified cross-validation comparing ligand_family-only, gene(best_ref)-only, and combined feature sets as predictors of the 5-category mechanism label. Ligand family alone achieves cross-validated log-loss = 0.944 (pseudo-R^2 = 0.231 relative to the majority-class baseline); gene identity alone achieves 0.896 (pseudo-R^2 = 0.270); the combined model achieves 0.896 — statistically indistinguishable from gene alone. A permutation test (200 within-gene-group shuffles of the ligand_family label) confirms that ligand family provides no significant additional predictive power once gene identity is known. Normalized mutual information is consistent with this ranking: MI(ligand_family, mechanism) = 0.198, MI(gene, mechanism) = 0.177, MI(clade, mechanism) = 0.025. We therefore state the organizing-unit claim in its tested form: ligand family is not a statistically independent predictor of mechanism beyond receptor gene identity, rather than the stronger, untested claim that ligand family "does not predict" mechanism at all.

2. The organizing unit is the gene/lineage, not the ligand label — a re-test on true orthogroups

The classification in Section 1 assigns each sequence to a reference receptor by a functional-label heuristic (`best_ref`), whose evolutionary purity is low (0.001–0.125). Because a conclusion about the "organizing unit" of mechanism should not rest on a functionally-defined label, we re-derived the grouping from first principles by reconstructing true orthogroups.

Using a DISCO-style decomposition of the gene tree (64.0% of internal nodes assigned to duplication, 36.0% to speciation) we recovered 13,539 orthogroups, of which 91.9% are singletons. Re-computing the mechanism-diversity ranking on these evolutionarily coherent units reproduces the key qualitative conclusions of Section 1: the many-to-many ligand↔mechanism relationship holds, and the high-diversity end of the family ranking

(peptide/chemokine families most diverse) is stable across three independent groupings (best_ref, pooled, and within-group), with the pooled-version ranking agreement rising to ρ = 1.0 as the sample-size threshold increases (Figure 3).

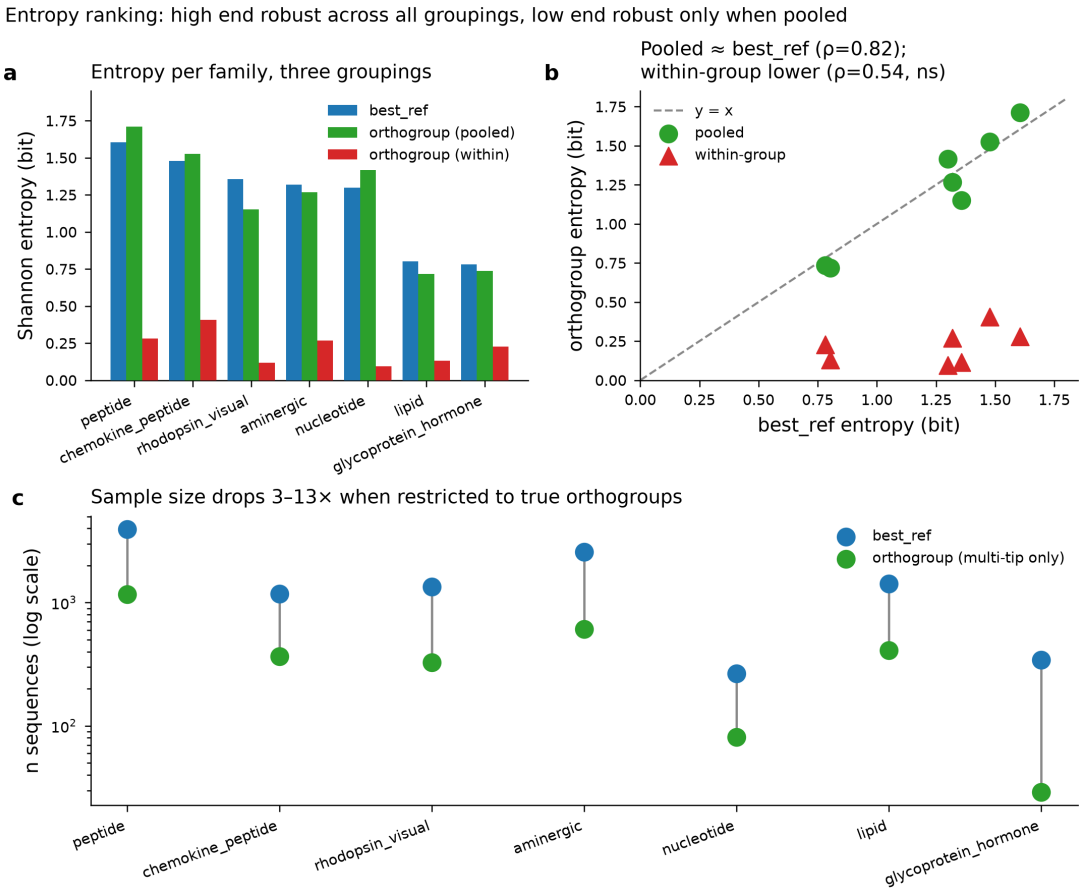

Figure 3. Orthogroup ranking. Entropy ranking across best_ref / pooled / within-group groupings; high-diversity end robust, low-diversity end an honest negative.

We report one negative result plainly: the *low*-diversity end of the ranking is not robust to the strictest (within-group) grouping (ρ = 0.536, p = 0.215, not significant). The claim that certain families are distinctively *low* in mechanistic diversity therefore does not survive the most conservative re-analysis, and we do not make it. The high-diversity end and the many-to-many structure — the load-bearing claims — do survive. This section doubles as a methodological point: the organizing unit of activation mechanism is recoverable only after the functional label is discarded in favor of reconstructed evolutionary units.

3. The switch components coevolve as a coupled network

If activation-switch architecture is not organized by ligand, is it organized at all? We tested whether the three residue systems most consistently implicated in the switch — the 6.30 lock, Tyr5.58, and the sodium-pocket residues — coevolve, after correcting for the phylogenetic non-independence of the sequences.

Three independent pairwise tests (Pagel's method, ARD model, on 11,609–12,484 cross-species sequences) each reject independent evolution decisively: lock–Tyr5.58 (LR = 212.10, $p = 9.37 \times 10^{-45}$), lock–sodium-pocket (LR = 86.88, $p = 6.07 \times 10^{-18}$), and Tyr5.58–sodium-pocket (LR = 264.87, $p = 4.06 \times 10^{-56}$). These three sites, previously linked only at the level of individual structures and molecular-dynamics simulations, are thus coupled at the level of cross-species coevolution, not merely by phylogenetic inertia.

To move beyond three separate pairwise edges to the joint architecture, we fitted four eight-state (lock × Tyr × sodium-pocket) Markov models to the full data set with corHMM and compared them by AICc (Figure 4). The fully-coupled 24-parameter model (M1) is decisively preferred (AICc = 15,554.06). A fully-independent model (M0) is rejected by ΔAICc = 439.0. Two "hub" models — one treating the sodium pocket as the pivot (M2) and one treating Tyr5.58 as the pivot (M3) — are also rejected, by ΔAICc = 102.5 and 53.6 respectively. The data therefore support neither independence nor a single privileged pivot, but the full coupled network. We note, and state in the text, that the ΔAICc of 53.6 separating the Tyr-hub model from the full model, while decisive by conventional thresholds, is not enormous; the data exclude complete independence and a sodium-pocket-hub simplification, but cannot rule out every reduced coupling structure.

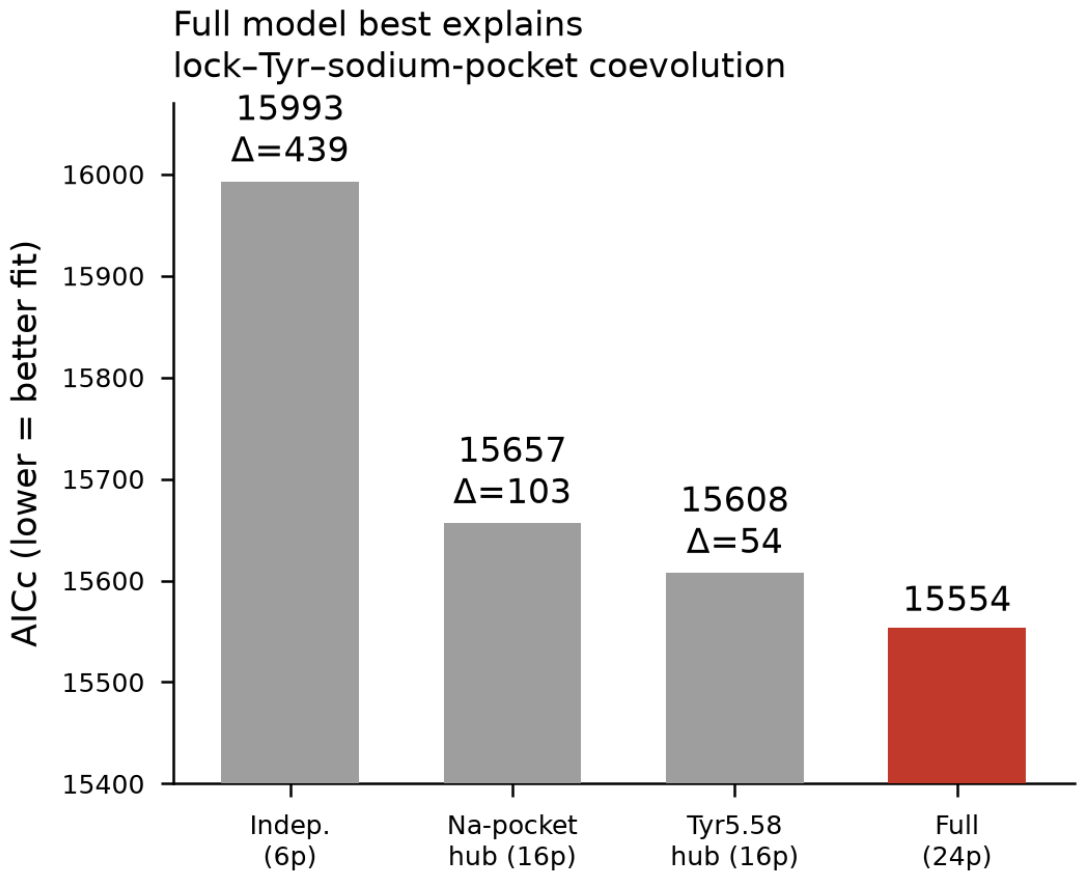

Figure 4. corHMM model comparison. AICc of the four eight-state models M0/M1/M2/M3.

Two robustness checks support the joint result. A parametric bootstrap at a reduced (2,000-tip) scale — necessitated by the ~2.5–3 h cost of a full-scale model pair — gives an observed ΔAICc(M0−M1) of 133.31; across eight null simulations under the fitted independent model, none approaches this value (null mean −16.73, SD 8.32, range [−24.56, −3.44]), placing the observation roughly 18 standard deviations above the null expectation. Second, a leave-one-clade-out jackknife on the same subsample (Figure 5) shows that removing any of the three largest taxonomic groups leaves ΔAICc essentially unchanged (baseline 133.31; −AGTR1 group 112.72; −opsin group 113.47; −5-HT1B group 111.38), all far above the ΔAICc = 10 decisive-support threshold. The coevolution signal is distributed across the phylogeny and is not an artifact of any single over-represented clade.

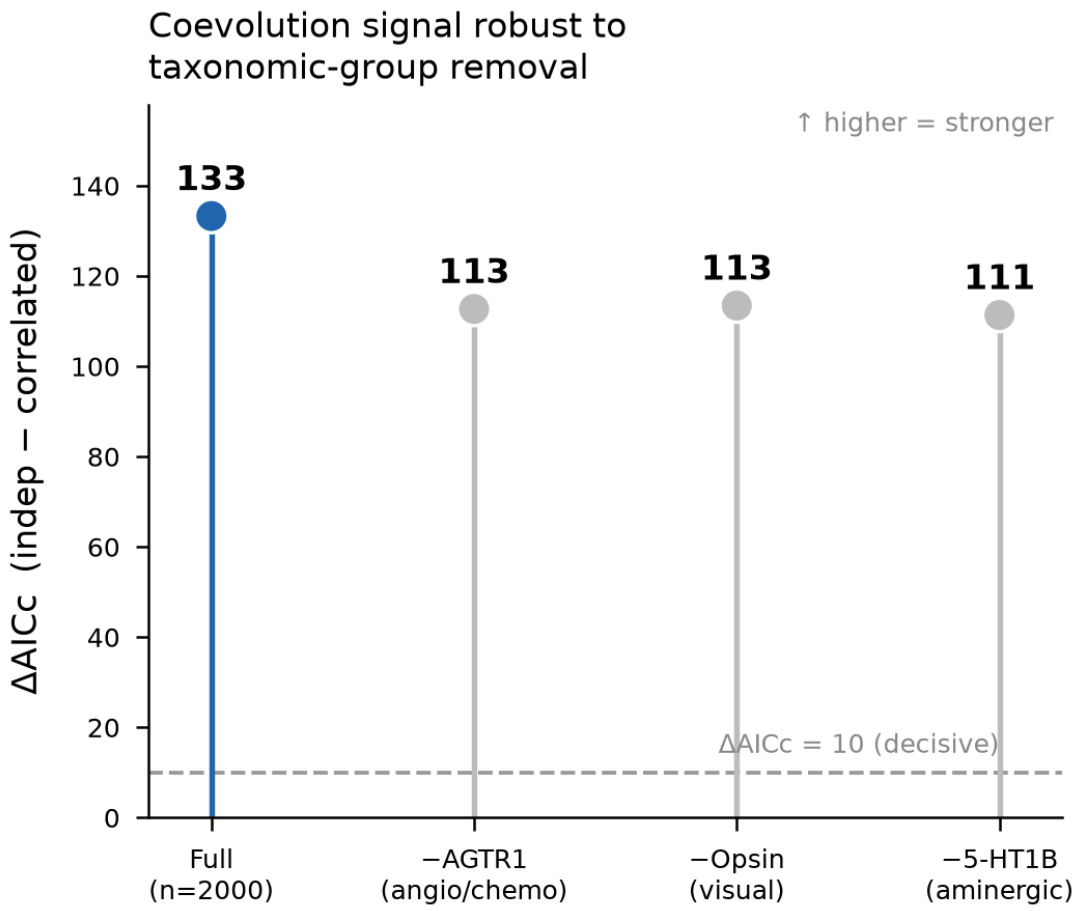

Figure 5. Jackknife robustness. ΔAICc(M0–M1) across leave-one-clade-out removals, with the ΔAICc = 10 decisive-support threshold marked.

Paralog-mixing sensitivity analysis (new in this revision). Because the gene tree used for the tests above mixes paralogs within species (§2), we repeated all three pairwise tests and the joint model comparison restricted to the 1,098 true multi-species orthogroups identified in §2 (4,179 tips, 25.1% of the full library). All three pairwise tests remain highly significant despite the reduced sample: lock–Tyr5.58 (n=3,257, LR=88.95, $p=2.2\times10^{-18}$), lock–sodium-pocket (n=3,304, LR=39.37, $p=5.85\times10^{-8}$), and Tyr5.58–sodium-pocket (n=3,629, LR=147.29, $p=7.76\times10^{-31}$) — effect sizes (LR) reduced by roughly 40-60% relative to the full gene tree, in the direction expected if some of the original signal reflected paralog non-independence, but the qualitative conclusion (rejection of the independence null) is unchanged. The joint corHMM comparison on the same orthogroup-restricted set (n=3,118) shows the same model ranking: M1 (full coupling, 24 parameters) is best-supported (AICc=4,294.27); M0 (independence) is decisively rejected (ΔAICc=220.06, vs. 439.0 on the full tree); M3 (Tyr-hub) is second-best (ΔAICc=12.86, vs. 53.6); M2 (sodium-pocket-hub) is rejected (ΔAICc=56.75, vs. 102.5). The coupled-network conclusion is therefore robust to a substantial (69-75%) sample-size reduction targeting paralog non-independence specifically, though effect sizes should be read as attenuated relative to the full-tree estimates above (an attempt to additionally test three pairwise-coupled models, M4a/b/c, on the full-scale data did not complete in this revision — see Methods and Limitations).

Extended jackknife and bootstrap sensitivity (new in this revision). The leave-one-clade-out jackknife was extended from 3 to 5 of the largest reference clades on the same n=500-2,000 subsample design: removing agtr1_human (ΔAICc=19.17), opsd_bovin (24.43), oprm_mouse (29.51), 5ht1b_human (26.80), or adrb2_human (19.83) all leave

ΔAICc well above the decisive-support threshold of 10, consistent with the original 3-clade result. We also probed how the parametric bootstrap's conclusion depends on subsample size: at the reduced n=300 scale needed to make more than a handful of replicates computationally tractable, the observed ΔAICc(M0-M1) itself falls to 12.84 (vs. 133.31 at n=2,000), and 19 null replicates (root state "111") give a much noisier null distribution (mean -11.83, SD 36.71) in which 2 of 19 replicates exceed the observed value (empirical p≈0.15, not significant at this scale). This is a real sample-size effect, not a contradiction of the n=2,000 result: it demonstrates that the strength of this particular test is sensitive to subsample size, and we report it rather than suppress it. The n=2,000 result (Figure 5) remains the basis for our claim of a decisive coupling signal; the n=300 result shows that claim would not survive naive scaling-down to a much smaller sample, and should be read as a caution on generalizing bootstrap significance across sample sizes rather than as evidence against the coupled-network conclusion, which rests primarily on the full-scale AICc comparison and the pairwise Pagel tests above.

4. A candidate alternative configuration: an acidic residue within TM6 in proximity to Arg6.30

Section 1 identified a "same-charge 6.30 with intra-helical rescue" category — receptors in which the 6.30 position cannot form the classic cross-helical lock, yet an acidic residue elsewhere in TM6 may substitute. We tested whether this arrangement holds at the level of solved structures, replacing the side-chain-centroid distance analysis used in an earlier version of this work with **minimum heavy-atom distances** between the candidate acidic partner (Asp OD1/OD2 or Glu OE1/OE2) and the Arg6.30 guanidinium group (NE/NH1/NH2), computed directly from deposited coordinates (generic numbering from GPCRdb; fusion-protein insertions — T4 lysozyme, rubredoxin, cytochrome b562; PDB residue numbers ≥ 1000 — excluded).

Of five candidate receptors (Table 1), only two — PAR2 (Glu279 = 6.27x28, 3.21 Å, PDB 5NDD) and AGTR2 (Asp252 = 6.31x31, 3.23 Å, 5UNF) — show a heavy-atom distance within classic salt-bridge range (<3.5 Å). CCR2 (Glu235, Arg238, 4.16 Å) is borderline, in an active-state complex rather than a resting state. EDNRA shows no salt-bridge-range contact in either resolved conformational state (inactive 8XVJ: 9.16 Å; active 8HCQ: 9.59 Å to the Cβ, since the Glu296 carboxylate itself is unresolved in this cryo-EM map). CCR5's candidate partner (Glu227) is likewise unresolved beyond Cβ in the deposited structure (5UIW), so no heavy-atom distance can be computed for this case; the distance previously reported for CCR5 rested on an incomplete side-chain model and should not be read as a validated measurement. Both receptors with genuine salt-bridge-range contacts (PAR2, AGTR2) are, in addition, engineered constructs (thermostabilized and BRIL-fusion chimeras respectively) rather than unmodified wild-type receptors, so a native-state distance cannot be independently confirmed from these coordinates alone.

We therefore describe this arrangement as a **candidate intra-TM6 proximity configuration** rather than a demonstrated stabilizing or rescue interaction. Even where a genuine salt-bridge-range contact is present, we have not established — and do not claim — that stabilizing Arg6.30 in this local configuration is functionally equivalent to protecting the DRY-motif Arg3.50 from destabilization; the two residues sit on different helices (TM6 vs. TM3) and no mutagenesis or molecular-dynamics evidence in this study links the two events

directly. The analogy to characterized cases in other receptors (e.g., the histamine H4 receptor's Asp5.69-Arg6.31 substitute salt bridge; Schneider et al. 2010) is suggestive but untested here.

Two receptors originally counted in this category were removed on a separate structural criterion. In both P2RY1 and CCRL2, the candidate acidic partner lies in ICL3, not TM6: GPCRdb places P2RY1's Asp248/Asp250 and CCRL2's Glu231 in the intracellular loop, upstream of where TM6 begins at 6.30. (The AlphaFold monomer does reproduce P2RY1's close Asp250 contact, ~4.5 Å, so the original geometry was not an error; the residue is simply a loop residue, not an intra-helical one.)

We deliberately do not attach a precise genome-wide count to this mechanism. The library-level generic-number assignment is alignment-sensitive at the TM5/TM6 boundary (positions 6.26/6.27), where it can err in both directions — false positives (P2RY1, CCRL2) and false negatives (PAR2, whose genuine TM6 Glu279 the library alignment missed). Of the 890 library sequences carrying a putative intra-TM6 partner, 626 place it at unambiguous deep-TM6 positions (6.28-6.34) while 264 sit at the alignment-sensitive boundary. We therefore present the phenomenon qualitatively and the five structure-tested cases quantitatively, and treat no single library count, nor the label "stabilizes," as established beyond the two confirmed salt-bridge-range cases.

Discussion

Prior work established *how common* the ionic lock is; our results address *what organizes the alternatives* when it is absent. The answer is that the organizing unit is not the ligand family. Within class A, ligand identity is a covariate of activation-switch architecture, not its determinant: a cross-validated multivariate test shows ligand family adds no predictive power for mechanism beyond receptor gene identity, and the unit at which mechanism is actually conserved descends to the gene and lineage. What *is* organized is the coevolutionary coupling of the switch residues themselves — the 6.30 lock, Tyr5.58, and the sodium pocket move together across the tree as a network, independently of ligand, and this coupling survives a paralog-mixing correction restricted to true multi-species orthologous groups (ΔAICc = 220, vs. 439 on the full gene tree; all three pairwise tests remain significant at $p < 10^{-7}$).

Local compensation appears graded rather than binary, but with an unresolved link in the mechanistic chain. Among the five structure-tested intra-TM6 cases, only PAR2 and AGTR2 show a genuine salt-bridge-range heavy-atom contact, and both are engineered rather than wild-type constructs. We have not established — and do not claim — that stabilizing Arg6.30 in this local configuration is functionally equivalent to protecting the DRY-motif Arg3.50; the two residues sit on different helices and no mutagenesis or molecular-dynamics evidence in this study links the two events directly. The analogy to characterized substitute salt bridges in other receptors (e.g., the histamine H4 receptor's Asp5.69-Arg6.31 pairing; Schneider et al. 2010) is suggestive but untested here. Given the small number of structure-tested cases, we offer graded local compensation as a hypothesis, not a conclusion.

The organizing level may itself shift across GPCR classes. Our evidence chain stops at the ligand-family-versus-gene level within class A. It is worth noting that the literature has established a *reversal* at the higher class level (A/B/C/F): Hauser et al. (2021) describe the

class C TM6 mechanism as dominated by dimeric rearrangement rather than monomeric intra-helical rearrangement, and Ellaithy et al. (2020) review the structural evidence that mGluR2 and GABA_B undergo a "TM4/TM5 interface (resting) → TM6 interface (active)" dimeric rearrangement confirmed by multiple cryo-EM structures. Juxtaposing our within-class-A finding with this established cross-class picture suggests that the true level at which constraint mechanisms are organized may be the class itself, with ligand family a local variation beneath it. We flag this explicitly as a literature-supported extrapolation — we do not test class-level causality here, and we introduce no cross-class data of our own. We likewise make no claim about the *origin* (convergent versus ancestral) of any mechanism.

Limitations. The orthogroup re-analysis restricts the coevolution tests to the 8.1% of orthogroups spanning ≥2 species (n=3,118-3,629 sequences, 25-31% of the original samples); the coupling signal survives this restriction at reduced effect size (see above), but the low-diversity end of the family ranking in Section 2 does not survive the strictest (within-group) grouping and we do not make that claim. The triad coupling is supported by pairwise tests and a joint eight-state AICc model across four candidate coupling structures (M0 fully independent, M1 fully coupled, M2/M3 single-trait hubs); M1 remains the best-supported model throughout. An attempt to additionally test three finer-grained pairwise-coupled models (M4a/b/c, designed to probe intermediate coupling structures between M0 and M2/M3) did not complete in this revision due to remote-compute resource constraints (Methods); we therefore did not exhaustively enumerate every possible reduced-coupling parameterization, and this remains an open direction. The parametric bootstrap and extended leave-one-clade-out jackknife were run at reduced sample scale for computational tractability; we report the achieved replicate count and sample size explicitly in Methods and Results rather than implying a larger scale than was run. Notably, the bootstrap's significance is itself scale-sensitive: at n=2,000 the observed ΔAICc(M0-M1)=133.31 is far outside an 8-replicate null distribution, but at n=300 (needed to obtain more replicates in practice) the observed effect shrinks to ΔAICc=12.84 and 19 null replicates give only a marginal empirical p≈0.15 — a caution that this particular robustness check does not generalize cleanly across subsample sizes, though the coupled-network conclusion itself rests primarily on the full-scale AICc comparison and the pairwise Pagel tests, not on the bootstrap. The extended jackknife (5 of 6 planned clades completed; the sixth failed on a data-file naming mismatch) confirms the original 3-clade result across two additional large clades. The intra-TM6 configuration is the weakest evidence in the paper: heavy-atom distance measurement shows only 2 of 5 candidate receptors (both engineered constructs) fall within genuine salt-bridge range, 2 cases have unresolved side chains precluding assessment, and none of the candidate partner residues has published site-directed mutagenesis directly testing its interaction with Arg6.30 or its functional consequence for Arg3.50 protection. We present it as a structural observation whose mechanistic role is not yet experimentally established.

Methods

Sequence data acquisition and filtering

Cross-species class A GPCR sequences were drawn from UniProt (SwissProt + TrEMBL) across 29 species spanning non-bilaterian animals (sponges, ctenophores, placozoans) through protostomes, invertebrate deuterostomes, and vertebrates. The combined library

(`all_species_combined.fasta`) contains 23,036 candidate sequences. Each candidate was pairwise-aligned (Biopython `PairwiseAligner`, BLOSUM62, global mode, open_gap=-11, extend_gap=-1) against 19 GPCRdb reference receptors spanning seven ligand-class families (aminergic, peptide, chemokine-peptide, lipid, nucleotide, glycoprotein-hormone, and rhodopsin-like), assigning each sequence its best-matching reference (`best_ref`) and percent identity. Sequences below 20% identity were excluded; a three-tier confidence label (low 20-25%, moderate 25-35%, high >35%) is retained in the final table (`confidence_tier` column) for sensitivity analysis. Sequences with all six key generic-number positions resolved (1.50x50, 3.50x50, 5.58x58, 6.30x30, 7.49x49, 7.50x50) total 11,609 and form the base sample for most quantitative analyses in Results §1/§3; a three-position subset (6.30/5.58 plus earlier analyses) totals 12,484. Generic (Ballesteros-Weinstein / GPCRdb structure-corrected) numbering was transferred column-by-column from reference-receptor alignments.

Phylogenetic tree construction and preprocessing

Gene tree provenance and known limitation (clarified in this revision). The `species_tree.nwk` file used in this study is a FastTree **gene tree** built from 16,620 candidate sequences (the 11,609/12,484-sequence subsets after `keep.tip` pruning), not a species tree — its default (midpoint) rooting produces a biologically implausible 11,525:84 basal split, indicating the rooting is arbitrary rather than reflecting a true ancestral divergence. On this tree, multiple paralogous genes from the same species co-occur as independent tips, and the `best_ref` functional label is highly polyphyletic (MRCA purity 0.001-0.125) — meaning that phylogenetic comparative tests run directly on this tree treat distinct paralogs as independent evolutionary lineages, risking under- or over-estimated effect sizes if paralog non-independence is not modeled.

Preprocessing pipeline (consistent throughout). (1) `ape::keep.tip` pruning of the 16,620-tip gene tree to the target trait dataset's tip set; (2) `ape::multi2di` random resolution of residual polytomies; (3) flooring of degenerate branch lengths below 1×10^{-5} to 1×10^{-6} to avoid numerical singularities in downstream likelihood calculations.

Methodological correction for paralog mixing (new in this revision). To test directly whether paralog mixing affects the core coevolution conclusions, we reused the project's existing DISCO-style gene-tree/species-tree reconciliation pipeline to decompose the 16,620-tip gene tree node-by-node into duplication/speciation events (any two child clades with overlapping species sets are called a duplication node and the tree is cut there), yielding 13,539 orthogroups — of which 12,441 (91.9%) are singletons unable to support cross-species comparison, and only 1,098 (8.1%) span ≥2 species, totaling 4,179 tips (25.1% of all candidate tips). **This revision adds a repeat of the three pairwise coevolution tests and the joint corHMM model comparison restricted to these 1,098 true orthogroups** (with all three traits resolved) as a paralog-mixing sensitivity analysis (Results §3, new paragraph).

Two tree objects must not be conflated. A separate 28-tip **species-level** skeleton tree (`species_tree_final_rooted_support.nwk`), rooted to minimize duplication-loss cost (DL cost = 23,178), was built for the Direction-1 trait-origin-ordering analysis (which trait arose earlier). Each tip on this tree corresponds to one species, not one sequence; it was **not**, and cannot be, used to carry the sequence-level fitPagel/corHMM tests reported in Results §3 — an ambiguity in the original manuscript's Methods text ("species tree pruned to the sequence set") is clarified here explicitly.

Phylogenetic comparative tests

Trait definitions. Ionic lock (`lock`): 6.30x30 is Asp or Glu (1) vs. other (0). Tyr5.58 (`tyr558`): 5.58x58 is Tyr (1) vs. other (0). Sodium-pocket integrity (`na_pocket`): simultaneous 1.50x50=Asn, 7.49x49=Asn, 7.50x50=Pro (1) vs. other (0).

Pairwise coevolution (Pagel's method). `phytools::fitPagel` (ARD model) tested each of the three pairs — (lock, tyr558), (lock, na_pocket), (tyr558, na_pocket) — for phylogenetically corrected coevolution against an independence null.

Joint eight-state corHMM model comparison. The three binary traits were composited into a single eight-state discrete character ($2^3 = 8$ combinations); four candidate transition-rate matrices were fit with corHMM 2.8 (`rate.cat=1, node.states="none"`) and compared by AICc: M0 (fully independent, 6 parameters), M1 (full ARD/fully coupled, 24 parameters, under a one-step-per-transition/Hamming-distance-1 constraint), M2 (sodium-pocket hub, 16 parameters), M3 (Tyr hub, 16 parameters), and, designed as an addition in this revision, M4a/M4b/M4c (three pairwise-coupled models, each with one trait pair coupled and the third independent, 10 parameters each) intended to probe the intermediate space between M0 and M2/M3 and address the concern that the original candidate set did not exhaust reduced coupling structures. The M4a/b/c rate matrices were constructed and the full-scale fitting run was submitted to remote compute, but did not complete within this revision due to sustained resource contention on the compute host; the AICc comparison reported in Results and Discussion is therefore limited to M0-M3, and M4a/b/c remain designed but unevaluated (see Limitations). All rate-matrix parameter indices were generated programmatically in Python (not hand-drawn) to avoid transcription error.

Parametric bootstrap. Because a single corHMM fit on the full-scale data (11,609-16,620 tips) takes on the order of 2-2.5 hours, bootstrap replicates were run on reduced subsamples (n=300-2,000, depending on remote-compute availability; independent seeds). M0 was fit to the observed data; its converged rate matrix was used to simulate null trait histories (`phytools::sim.history`, fixed root state, default all-negative "000", with a root-state sensitivity check) and M0/M1 were refit to each simulated dataset, yielding a null ΔAICc(M0-M1) distribution compared to the observed value via the plus-one Monte Carlo method (Davison & Hinkley 1997): p = (1 + count of null ≥ observed) / (1 + total replicates).

Leave-one-clade-out jackknife. On the same subsample, the largest `best_ref` reference clades were removed one at a time (extended in this revision from 3 clades to 6 planned, 5 completed — the sixth failed on a data-file naming mismatch), M0/M1 refit, and ΔAICc recorded to test whether the coevolution signal depends on any single over-represented lineage.

Computing environment. All phylogenetic comparative analyses ran on a remote host (`ssh:run`), conda environment `gpcr_phylo` (R 4.3.3, corHMM 2.8, phytools 2.5.2, ape).

Multivariate test of ligand-mechanism predictive power (new in this revision)

To test directly whether ligand family predicts Arg3.50-protection mechanism (rather than inferring this from aggregate entropy statistics alone), we built multivariate classification models on n=11,059 sequences (excluding 550 with a non-Arg 3.50). Using mechanism_top (5 classes) as the outcome, ligand_family (7 classes), gene/best_ref (18 reference-matched classes), and clade (15 lineages) — individually and combined — were fit as predictors with

scikit-learn `LogisticRegression` (one-vs-rest, L2-regularized), evaluated by 5-fold stratified cross-validated log-loss. The key test — whether adding ligand_family provides statistically significant additional predictive power once gene identity is known — used a permutation test (200 within-gene-group shuffles of the ligand_family label, preserving the gene-mechanism association while destroying any independent ligand signal) to build a null distribution against the observed log-loss reduction. Normalized mutual information between ligand_family, gene, clade, and mechanism was also computed as a supplementary descriptive statistic.

Arg3.50-protection mechanism classification

For all 11,609 library sequences, generic-number columns spanning ±4 residues around 3.50 (window 3.46x46-3.54x54) and 6.30 (window 6.26x26-6.34x34) were used to classify each sequence into one of six categories: (1) classic cross-helical lock (6.30 Asp/Glu forming a cross-TM3-TM6 salt bridge with Arg3.50); (2) same-charge intra-helix rescue (6.30 Arg/Lys, same charge as Arg3.50, but an Asp/Glu elsewhere in the window provides local charge neutralization); (3) neutral 6.30 with local acidic residue present; (4) same-charge with no local rescue found; (5) neutral 6.30 with no local acidic residue; (6) non-Arg 3.50 (discussed separately, excluded from the mechanism-diversity statistics of categories 1-5).

Orthogroup reconstruction

A DISCO-style gene-tree decomposition algorithm classified each node of the 16,620-tip gene tree as duplication or speciation (any two child-clade species sets overlapping = duplication, cutting the tree there), yielding non-overlapping orthogroups. The classification was cross-validated two ways: (a) the overall 64.0% duplication / 36.0% speciation ratio matched an independent bitmask-method baseline computed in an earlier project phase; (b) all 13,539 resulting orthogroups were individually verified to contain zero single-copy violations. For each orthogroup, three versions of ligand-family-grouped Shannon entropy of mechanism diversity were computed: (a) best_ref version (original functional-label grouping); (b) orthogroup-pooled version (all sequences within multi-sequence orthogroups pooled by majority ligand_family label before computing entropy); (c) orthogroup within-group version (entropy computed per orthogroup, then averaged by majority label).

Structural analysis

Structure sources. Experimentally determined structures (X-ray/cryo-EM) were retrieved from RCSB PDB (mmCIF format); AlphaFold monomer predictions from the EBI AlphaFold DB API.

Distance measurement (revised in this manuscript: precise heavy-atom distances replace side-chain-centroid distances). For each candidate intra-TM6 acidic-partner case, mmCIF coordinates were parsed with gemmi and the **minimum heavy-atom distance** was computed between the candidate acidic residue (Asp OD1/OD2, or Glu OE1/OE2) and the Arg partner (NE/NH1/NH2), reporting the specific atom pair. For well-resolved cases, a supplementary salt-bridge geometry angle (acidic-residue CG/CD-nearest O atom-Arg nearest N atom) was also computed. Where the candidate residue's side chain was not fully resolved in the experimental density (resolved only to Cβ), this is explicitly flagged as "distance not assessable" rather than substituting a truncated-side-chain proxy distance as if it were a validated measurement.

Construct engineering annotation. For each PDB structure included in Table 1, `_struct_ref_seq_dif` (SEQADV) mmCIF records were used to count engineered mutations, expression tags, and linkers, combined with structure titles/literature to determine whether each structure is a thermostabilized construct, a fusion-protein chimera (e.g., BRIL/T4 lysozyme insertion), or a specific conformational-state complex (active vs. inactive), distinguishing "direct wild-type resting-state evidence" from "engineered-construct inference."

Exclusion criteria. Fusion-protein insertion residues (typically numbered ≥1000, e.g., T4 lysozyme, rubredoxin, cytochrome b562 insertions), hetero residues, and backbone atoms were excluded from the acidic-partner search. Search radius: 12 Å for structural analysis, 15 Å for the library-wide survey.

Data and code availability

See the Data availability statement at the end of this manuscript for deposit details and DOIs.

Table

Table 1. Candidate intra-TM6 acidic residues in proximity to Arg6.30. Five receptors in which an acidic residue within TM6 is a candidate alternative to the classic cross-helical lock. Distances are minimum heavy-atom separations between the acidic partner and the Arg6.30 guanidinium group, computed from deposited coordinates (Methods); fusion-protein insertions excluded. Only PAR2 and AGTR2 fall within classic salt-bridge range (<3.5 Å); both are engineered (non-wild-type) constructs. EDNRA is resolved in two conformational states, neither within salt-bridge range; CCR5 and EDNRA (active) have unresolved acidic side chains and cannot be assessed at heavy-atom resolution.

Receptor	Acidic partner	Generic position	Min. heavy-atom dist. (Å)	Salt-bridge range?	Resolution	Method	Engineered mutations	PDB
PAR2	Glu279	6.27x28	3.21 (OE2-NH1)	Yes (<3.5 Å)	2.80 Å	X-ray	10 (thermostabilized)	5NDD
CCR5	Glu227	6.27x27	not assessable (carboxylate unresolved)	Cannot assess	2.20 Å	X-ray	4	5UIW
CCR2	Glu235	6.28 region	4.16 (OE1-NH2, partner Arg238)	Borderline (<4.5 Å)	2.90 Å	cryo-EM (active complex)	0	7XA3
AGTR2	Asp252	6.31x31	3.23 (OD1-NE)	Yes (<3.5 Å)	2.80 Å	X-ray	6 (BRIL fusion)	5UNF
EDNRA (inactive)	Glu303	6.32x32	9.16 (NE-OE2)	No	3.26 Å	cryo-EM	0	8XVJ
EDNRA (active)	Glu296	6.25x25	not assessable (carboxylate unresolved)	Cannot assess	3.01 Å	cryo-EM	0	8HCQ

Competing interests

The author declares no competing interests.

Funding

This work received no external funding.

Author contributions

H.S. conceived the study, performed all analyses, and wrote the manuscript.

Computational tools and AI-assisted analysis

Data assembly, statistical analysis, and figure generation were implemented in Python with substantial code-development and computational assistance from Claude Science (Anthropic), an AI-based scientific computing agent, used interactively throughout data assembly, analysis, and figure preparation. All analysis code, computed values, and statistical results were reviewed, checked for internal consistency, and independently spot-verified by the author against source data; the author takes full responsibility for the accuracy of all reported results and conclusions. Consistent with ICMJE and COPE guidance, this AI tool is not credited as an author, as it cannot take accountability for the work.

Data availability

All raw classification tables, model-comparison outputs, parametric bootstrap/jackknife results, structural resolution notes, figure source files, the sequence-level gene tree (`species_tree.nwk`), and the per-sequence trait table used for all phylogenetic comparative tests are deposited at Zenodo under an open (CC-BY 4.0) license, under the concept DOI https://doi.org/10.5281/zenodo.21351360 (always resolves to the current version). Version 1 (10.5281/zenodo.21351361, 2026-07-14) is the original submission; version 2 (10.5281/zenodo.21627351, 2026-07-27) adds the data and code from this revision (orthogroup-restricted reruns, the ligand-predictive-power test, precise structural distances, and expanded bootstrap/jackknife results); version 3 (10.5281/zenodo.21899024, 2026-08-12) adds `species_tree.nwk`, the per-sequence trait table (`species_generic_number_transfer_v2_usable.csv`), and the orthogroup tagging table (`orthogroups_disco_tagged.csv`) needed to independently rerun the phylogenetic comparative analyses – inadvertently omitted from versions 1-2.

References

All entries below were independently verified against CrossRef during this revision (author list, journal, volume, and pagination checked against the resolved record for each DOI); two entries in the previous unverified list contained metadata errors, corrected here and noted individually.

- Kling RC, Clark T, Gmeiner P (2016). Comparative MD simulations indicate a dual role for Arg132$^{3.50}$ in dopamine-dependent D2R activation. *PLoS ONE* 11(1):e0146612. doi:10.1371/journal.pone.0146612. [verified; title corrected — the paper is specifically about the D2 dopamine receptor, not "class A GPCR activation" generically]
- Hauser AS, Kooistra AJ, Munk C, Heydenreich FM, Veprintsev DB, Bouvier M, Babu MM, Gloriam DE (2021). GPCR activation mechanisms across classes and macro/microscales. *Nat Struct Mol Biol* 28:879-888. doi:10.1038/s41594-021-00674-7. [verified]
- Ellaithy A, Gonzalez-Maeso J, Logothetis DE, Levitz J (2020). Structural and biophysical mechanisms of class C GPCR function. *Trends Biochem Sci* 45(12):1049-1064. doi:10.1016/j.tibs.2020.07.008. [verified]
- Springael JY, de Poorter C, Deupi X, Van Durme J, Pardo L, Parmentier M (2007). The activation mechanism of chemokine receptor CCR5 involves common structural changes but a different network of interhelical interactions relative to rhodopsin. *Cell Signal* 19(7):1446-1456. doi:10.1016/j.cellsig.2007.01.022. [PMID:17320349] [verified]
- Schneider EH, Schnell D, Strasser A, Dove S, Seifert R (2010). Impact of the DRY motif and the missing "ionic lock" on constitutive activity and G-protein coupling of the human histamine H4 receptor. *J Pharmacol Exp Ther* 333(2):382-392. doi:10.1124/jpet.109.163220. [verified]
- Sotudeh N, Morales P, Hurst DP, Lynch DL, Reggio PH (2019). Towards A Molecular Understanding of The Cannabinoid Related Orphan Receptor GPR18: A Focus on Its Constitutive Activity. *Int J Mol Sci* 20(9):2300. doi:10.3390/ijms20092300. [PMC6539512] [verified; author list resolved and corrected — full author list was not previously located]
- Pagel M (1994). Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc R Soc Lond B* 255:37-45. doi:10.1098/rspb.1994.0006. [verified]
- Revell LJ (2012). phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol* 3:217-223. doi:10.1111/j.2041-210X.2011.00169.x. [verified]
- Beaulieu JM, O'Meara BC, Donoghue MJ (2013). Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habit in campanulid angiosperms. *Syst Biol* 62(5):725-737. doi:10.1093/sysbio/syt034. [verified; title corrected — the original entry dropped the paper's subtitle]
- Willson J, Roddur MS, Liu B, Zaharias P, Warnow T (2022). DISCO: species tree inference using multicopy gene family tree decomposition. *Syst Biol* 71(3):610-629. doi:10.1093/sysbio/syab070. [verified; citation located and corrected — the manuscript's placeholder attribution to Steenwyk & Rokas was wrong]
- Munk C, Isberg V, Mordalski S, Harpsøe K, Rataj K, Hauser AS, Kolb P, Bojarski AJ, Vriend G, Gloriam DE (2016). GPCRdb: the G protein-coupled receptor database — an introduction. *Br J Pharmacol* 173(14):2195-2207. doi:10.1111/bph.13509. [verified]

- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000). The Protein Data Bank. *Nucleic Acids Res* 28(1):235-242. doi:10.1093/nar/28.1.235. [verified]
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596:583-589. doi:10.1038/s41586-021-03819-2. [verified; third author's name corrected from "Pritzker" to "Pritzel"]
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, et al. (2011). Scikit-learn: Machine Learning in Python. *J Mach Learn Res* 12:2825-2830. [verified against the JMLR record]