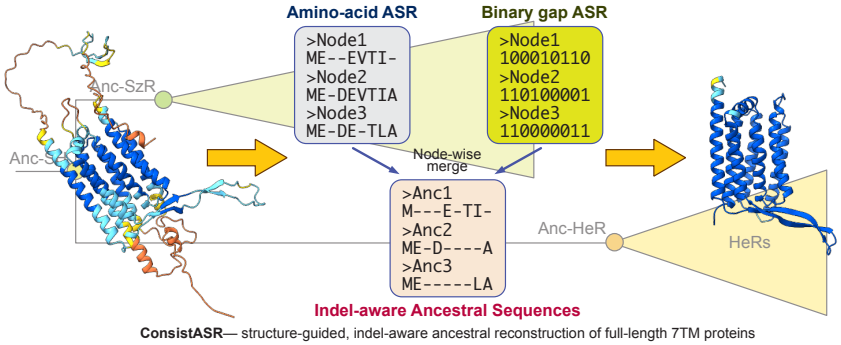

Graphical Abstract

Resurrecting Full-length Ancestral Schizorhodopsins and Heliorhodopsins with Structure-guided, Indel-aware Sequence Reconstruction

Haruto Ishikawa, Yasuhisa Mizutani

Highlights

Resurrecting Full-length Ancestral Schizorhodopsins and Heliorhodopsins with Structure-guided, Indel-aware Sequence Reconstruction

Haruto Ishikawa, Yasuhisa Mizutani

- Ancestral schizorhodopsin and heliorhodopsin are reconstructed.
- Indel-aware refining yields compact ancestors with lengths close to extant proteins.
- AlphaFold models retain lineage-specific extra-membrane secondary structures.
- Reliability is evaluated by mapping PPs together with AlphaFold confidence.
- Anc-SzR and Anc-HeR express as stable colored holoproteins in *E. coli*.

Resurrecting Full-length Ancestral Schizorhodopsins and Heliorhodopsins with Structure-guided, Indel-aware Sequence Reconstruction

Haruto Ishikawa, Yasuhisa Mizutani

[a] *Department of Chemistry, Graduate School of Science, The University of Osaka, 1-1 Machikaneyama, Toyonaka, 560-0043, Osaka, Japan*

Abstract

Microbial rhodopsins exhibit diverse functions ranging from ion pumps and ion channels to light sensors, despite sharing a common seven-transmembrane (7TM) domain structure. Their evolutionary process is of interest in understanding how they have acquired such varied functions. A direct route to inferring and experimentally testing plausible ancestral rhodopsins is provided by ancestral sequence reconstruction (ASR). However, ASR of 7TM proteins is often limited in practice by alignment ambiguity and insertion–deletion (indel) uncertainty, especially in extra-membrane (EM) loops and termini. As a result, many studies focus on trimmed transmembrane (TM) cores and treat EM regions by manual curation, leaving the evolutionary history of full-length architecture difficult to test experimentally. Here we reconstruct and resurrect full-length ancestral schizorhodopsins (Anc-SzR) and heliorhodopsins (Anc-HeR), two microbial rhodopsin families that share a retinal-binding 7TM core but differ in membrane topology and EM secondary-structure elements. Starting from untrimmed alignments, we combine structure-consistent multiple sequence alignments and profile-based evolutionary models with an explicit indel-aware refinement that merges amino-acid ancestral states with binary ancestral gap inference on a fixed topology. Indel-aware refinement prevents artificially overextended ancestors and

⋆Correspondence
Email address: haruto@chem.sci.osaka-u.ac.jp (Haruto Ishikawa)

yields compact full-length sequences. The resulting models form high-confidence 7TM folds, including characteristic β-strands and short helices. Finally, both Anc-SzR and Anc-HeR express in *Escherichia coli* are recovered as stable, colored holoproteins that bind retinal. Together, these results show that full-length, indel-aware ASR can produce experimentally tractable ancestral microbial rhodopsins and enables direct tests of how EM architecture evolves alongside the 7TM core.
Keywords: ancestral sequence reconstruction, microbial rhodopsins, seven-transmembrane proteins, indels, AlphaFold

---

Introduction

Microbial rhodopsins are a broadly distributed family of retinal-binding photoreceptors found across bacteria, archaea, and microbial eukaryotes. Despite wide sequence divergence, they share a conserved seven-transmembrane (7TM) architecture that houses the retinal-binding pocket and the Schiff-base lysine, enabling light-driven ion transport, sensory signaling, and other photobiological functions [1, 2]. Over the past decade, large-scale genome and metagenome sequencing has greatly expanded the catalog of microbial rhodopsins, revealing extensive functional diversity that includes proton, sodium, chloride, and other ion pumps, light-gated channels, and sensory rhodopsins, often with lineage-specific adaptations in both the transmembrane (TM) core and extra-membrane (EM) segments. This explosion of sequence diversity raises a fundamental evolutionary question: how did such a wide range of photochemical and physiological functions emerge from a shared 7TM scaffold? Addressing this question requires approaches that can move beyond comparisons among extant proteins and explicitly reconstruct historical intermediates. Ancestral sequence reconstruction (ASR) provides a direct route to inferring and experimentally testing plausible ancestral rhodopsins, thereby linking phylogenetic history to changes in structure, spectral tuning, oligomerization, and functional mechanisms over evolutionary time.

ASR infers the sequences of extinct proteins at internal nodes of a phylogeny

and enables experimental tests of how protein structure, stability, and function evolved over deep time [3, 4]. In many systems, resurrected ancestors provide tractable reagents for connecting evolutionary hypotheses to biochemical mechanisms. At the same time, ASR outcomes can be strongly affected by practical choices—most notably multiple sequence alignment [5, 6], evolutionary model specification [7, 8], and the treatment of insertions and deletions (indels) [9, 10]. Such choices can alter both tree topology and posterior probabilities (PPs) at internal nodes, yet their consequences are not always apparent from standard ASR diagnostics, motivating explicit evaluations of robustness when ancestors are taken forward for experimental resurrection.

For soluble proteins, the effects of alignment uncertainty, model choice, and indel treatment on ASR have been examined in detail [4, 6]. In contrast, 7TM membrane proteins present a distinct set of practical challenges that remain less systematically explored. 7TM proteins combine a conserved hydrophobic helical core with EM regions that are highly variable in both length and secondary structure, including N- and C-terminal tails, peri-membrane helix termini, amphipathic elements, β-strands, and flexible coils [11, 12]. Because these EM segments are often gap-rich and alignment-sensitive, many studies focus on trimmed TM cores and handle EM regions and termini through manual curation, thereby treating indels only implicitly [13, 14, 15]. Consequently, low overall sequence identity, compositional bias, and topological constraints complicate both alignment and model fitting in many 7TM families [16, 17]. Together, these factors have limited efforts to reconstruct and experimentally test full-length 7TM ancestors and have left EM architecture underused as an explicit source of evolutionary information.

As demonstrated in previous studies, alignment errors can propagate into ASR errors and distort downstream structural and functional inferences, sometimes more strongly than phylogenetic uncertainty itself [18]. In particular, Aadland and Kolaczkowski identified insertion–deletion (indel) errors as a major contributor to ASR inaccuracy across realistic simulation conditions and emphasized that many practical workflows still handle indels only implicitly through

the choice of a single alignment rather than by explicit modeling. This concern is amplified in 7TM proteins, where EM regions are often long and gap-rich: such datasets provide a stringent setting for evaluating how alignment and indel uncertainty affect full-length ancestral reconstructions that are intended for experimental testing.

Schizorhodopsins (SzRs) and heliorhodopsins (HeRs) share a conserved 7TM architecture and the retinal-binding lysine in the seventh transmembrane helix (TM7), yet differ sharply in membrane topology and oligomeric state: SzRs adopt an N-out/C-in orientation and form trimers, whereas HeRs display an inverted N-in/C-out topology and form dimers [19, 20, 21, 22, 23, 24]. Both families also contain characteristic EM secondary-structure elements—including short helices and short or long β-strands—that contribute to intersubunit packing and function [20, 22, 23, 24]. These contrasts provide a controlled setting to (i) compare alignment and model choices, (ii) quantify the impact of indels on full-length reconstructions, and (iii) test whether ancestral sequences can recover not only the 7TM core but also lineage-specific EM architectural motifs with high confidence. (Here and below, TM1–TM7 denote the seven canonical transmembrane helices numbered from the N terminus.)

Recent advances now make it possible to revisit these questions for SzRs and HeRs more systematically. First, profile-based substitution models such as Q.pfam+R better reflect the compositional and site-specific constraints typical of membrane proteins than classical empirical matrices [16], and are now readily usable in routine likelihood-based analyses. Second, explicit binary gap models implemented in RAxML and related tools provide a principled route to reconstructing ancestral indel patterns on a fixed topology [25, 26]. Third, structure prediction with residue-wise confidence measures (pLDDT), such as those provided by AlphaFold [27], offers an independent way to evaluate whether inferred ancestral sequences yield coherent 7TM folds and plausible extra-membrane architectures. Together, these developments enable full-length ancestral reconstructions to be assessed not only by evolutionary fit but also by structural plausibility—an essential step when ancestors are designated for experimental

resurrection.

Accordingly, we reconstruct and experimentally resurrect full-length ancestral schizorhodopsins (Anc-SzR) and heliorhodopsins (Anc-HeR) and evaluate the reliability of both TM and EM regions in this study. Starting from untrimmed, structure-consistent alignments, we infer amino-acid ancestral states under profile-based models, refine ancestral sequences with explicit indel-aware masking based on binary gap reconstructions on a fixed topology, and assess structural plausibility using AlphaFold predictions for monomeric and oligomeric assemblies. To summarize node-level robustness in a transparent way, we report joint sequence–structure confidence by mapping posterior probabilities (PPs) together with pLDDT and interpret these alongside bootstrap support for the corresponding splits.

We heterologously express and purify indel-corrected Anc-SzR and Anc-HeR in *Escherichia coli* (*E. coli*), demonstrating that full-length ancestors inferred from untrimmed alignments can fold as retinal-binding 7TM proteins with characteristic visible absorption spectra. Using these resurrected ancestors as a readout, we evaluate how alignment choice, substitution models, and explicit indel-aware refinement influence node stability, sequence length, and confidence in EM regions, and we compare evolutionary support with structure-based confidence across ancestral depths. Together, our results establish SzR and HeR as a tractable system for full-length 7TM resurrection and show that EM architecture can be reconstructed and evaluated as an explicit component of membrane-protein evolution rather than being routinely trimmed away.

Results

Alignment and model choice shape the stability of Anc-HeR but not Anc-SzR

Multiple sequence alignment had a decisive impact on the stability of HeR–SzR phylogenies, and this effect was concentrated in the HeR lineage. We therefore compared two structure-aware or structure-compatible strategies: PSI/TM-Coffee [28], which incorporates transmembrane-specific constraints and homology

information, and MAFFT L-INS-i [29], a widely used high-accuracy method for protein families with difficulty in sequence alignment. PSI/TM-Coffee consistently preserved TM topology and EM registers across the dataset, whereas L-INS-i produced broadly coherent alignments but introduced localized ambiguity around the HeR TM1–TM2 region. As demonstrated in the following section, these localized differences translate into marked differences in topological stability around Anc-HeR, while Anc-SzR remains robust across alignment choices.

For each alignment, we used ModelFinder in IQ-TREE [30, 31] to select the best-fit amino-acid substitution model under the Bayesian information criterion (BIC). Across the full HeR+SzR+outgroup dataset and the two reduced subsets (HeR+outgroup and SzR+outgroup), profile models from the Q.pfam family with FreeRate heterogeneity (Q.pfam+R5–R7) were consistently favored (Table 1). In the full dataset, Q.pfam+R7 was selected for both PSI/TM-Coffee and MAFFT L-INS-i alignments and improved BIC by roughly 200–250 units relative to the best LG-based FreeRate alternatives, indicating strong support for a membrane-aware profile model over conventional empirical matrices.

However, a single exception was observed in the MAFFT L-INS-i alignment of the HeR+outgroup dataset. In this case, an LG+F+R model was favored, and the HeR ancestral branch collapsed into a near-polytomy in the UFBoot2 consensus tree, indicating locally unstable support around this split [32]. In our analyses, this behavior underscores that alignment quality can dominate model choice in difficult 7TM datasets: when the alignment introduces local inconsistencies—most notably around the HeR TM1–TM2 registers—both the preferred substitution model and the resolution of nearby internal nodes can shift. It is important to note that the overall preference for Q.pfam+R models was unchanged when moderately gapped terminal regions were removed with trimAl `gappyout`, suggesting that the signal favoring Q.pfam+R is not driven solely by gap-rich columns but is supported across both membrane-embedded and extra-membrane regions (Table S1).

For completeness, we also tested a more parameter-rich empirical profile-mixture model (LG+C60+F+R8), which improved overall fit relative to Q.pfam+R7

(Table S1). However, improved fit did not translate into improved stability of the alignment-sensitive HeR region: under LG+C60+F+R8, the Anc-HeR branch remained weak or unstable in the UFBoot2 consensus (absent for PSI/TM-Coffee and 47% for MAFFT L-INS-i), whereas Anc-SzR and Anc-SH, the most recent common ancestor (MRCA) of schizorhodopsin and heliorhodopsin, remained strongly supported. Because LG+C60+F+R8 is substantially more computationally demanding and did not change the qualitative topology or node-level reliability patterns reported below, we used Q.pfam+R7 as the primary model and treated LG+C60+F+R8 as a robustness check (Tables S2 and S3). It is important to note that the consistent preference for Q.pfam+R models also dictated the methodological choice of performing ASR in IQ-TREE, as profile models of this class are not available in classical pipelines based on RAxML or PAML. A direct comparison with a conventional RAxML+PAML workflow under LG-based models is provided in Tables S2 and S3.

These patterns are summarized in Table 1 and visualized in Fig. 1, highlighting the strong dependence of Anc-HeR on alignment quality and the robust behavior of the SzR lineage across all conditions. Accordingly, unless stated otherwise, we used the PSI/TM-Coffee alignment with the Q.pfam+R7 model as the baseline configuration for downstream ASR, indel-aware refinement, and structure-based evaluations.

Indel-aware refinement restores compact, high-confidence ancestral rhodopsins

Indel treatment emerged as a major source of uncertainty in full-length ASR. When ancestral states were inferred in IQ-TREE without explicit indel modeling, the resulting ancestral sequences expanded far beyond extant rhodopsins, with the inflation concentrated in extra-membrane regions. Under the PSI/TM-Coffee alignment, Anc-SzR, Anc-HeR, and Anc-SH reached 401, 434, and 408 residues, respectively, compared with 202 residues for SzR4 and 256 residues for HeR 48C12. This over-extension depressed mean site-wise posterior probabilities (PP; ∼38–66%) and, upon AlphaFold evaluation, produced models that were globally 7TM-like but dominated by elongated, low-confidence tails and EM segments

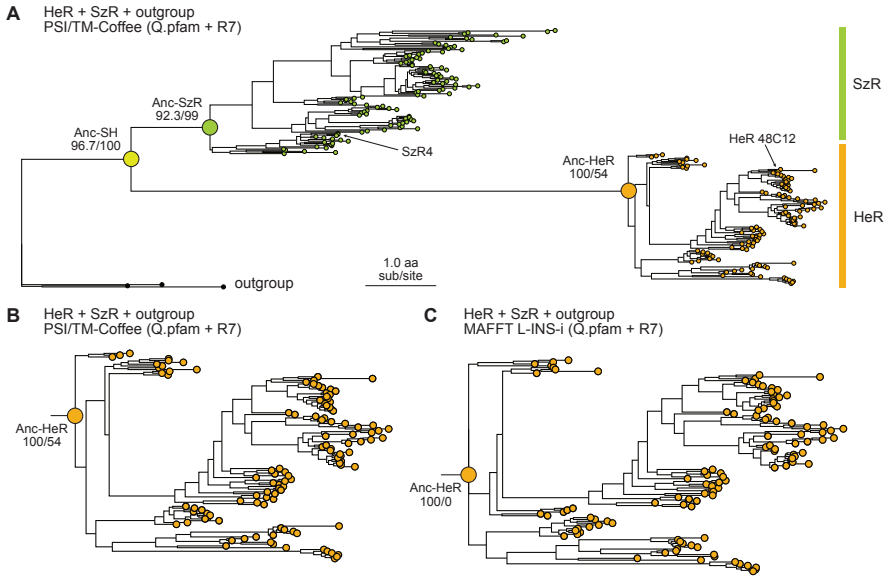

Figure 1: **Phylogenetic context of SzR and HeR ancestors and alignment sensitivity of the HeR lineage.** (A) Maximum-likelihood (ML) tree for the HeR+SzR+outgroup dataset inferred from the PSI/TM-Coffee alignment under Q.pfam+R7. HeR (orange), SzR (green), and outgroup clades are indicated, and the three focal ancestral nodes (Anc-SH, Anc-SzR, and Anc-HeR) are labeled with SH-aLRT/UFBoot2 support. Extant structural representatives (SzR4 and HeR 48C12) are marked. (B) Enlargement of the HeR clade from (A), highlighting internal branching patterns and support values around Anc-HeR under PSI/TM-Coffee. (C) Corresponding enlargement of the HeR clade inferred from the MAFFT L-INS-i alignment under the same model (Q.pfam+R7), showing reduced resolution around Anc-HeR and collapse of the corresponding split in the UFBoot2 consensus. Support values are shown as SH-aLRT/UFBoot2. When a split is absent from the UFBoot2 consensus tree (i.e., collapsed into a local polytomy), its UFBoot2 support is reported as 0 (e.g., Anc-HeR in panel (C)).

(mean pLDDT typically ∼60–70; computed from per-residue Cα pLDDT; see Materials and Methods). The same qualitative behavior was observed with the MAFFT L-INS-i alignment, indicating that—within these reconstructions—gap-unaware ASR tends to yield inflated ancestral length and structurally unreliable termini (Fig. 2A and Table 3).

To address this issue, a two-stage indel-aware refinement pipeline based on binary gap coding was implemented (Fig. 2B) [26]. Briefly, amino-acid ancestral states were inferred in IQ-TREE under profile-based models, while ancestral gap states were inferred from the same alignment recoded as a binary gap/non-gap matrix using a two-state model in RAxML on the fixed IQ-TREE topology. The inferred ancestral gap patterns were then mapped onto the IQ-TREE `.state` output, and sites inferred as gaps for each node were masked to generate node-specific, indel-corrected ancestral sequences (Fig. 2B). This node-specific masking contracts each ancestor according to its inferred indel history, rather than applying a single global gap filter to all nodes.

Indel-aware correction had a pronounced impact on both statistical and structural metrics (Fig. 2C, D and Table 3). For the PSI/TM-Coffee alignment, mean PP increased from 51–66% to 82–91% for Anc-SzR and Anc-HeR, and mean pLDDT rose from ∼62–69 to >93. In parallel, sequence lengths contracted to values closely matching extant structures: 206 residues for Anc-SzR and 259 residues for Anc-HeR, compared with 202 residues for SzR4 and 256 residues for HeR 48C12. Comparable improvements were obtained with the MAFFT L-INS-i alignment, with Anc-SzR and Anc-HeR contracting to 203 and 260 residues, respectively, indicating that the qualitative effect of indel-aware refinement is reproducible across alignment strategies.

AlphaFold models of the indel-corrected ancestors formed compact 7TM folds with uniformly high pLDDT and retained lineage-specific extra-membrane motifs, including the short β-strands between TM2 and TM3 in SzR4 and the combination of a short α-helix between TM2 and TM3 with long β-strands between TM1 and TM2 in HeR 48C12 [23, 22]. In contrast, models predicted from uncorrected ancestors exhibited extended, low-confidence extra-membrane

tails. Thus, node-specific masking based on the inferred ancestral indel state is a key step for obtaining full-length 7TM ancestors that are both statistically well supported and structurally coherent. All three ancestors retained the retinal-binding lysine in TM7; the FSE motif in TM3 was conserved in Anc-SzR but shifted to DSE in Anc-SH.

Beyond global statistics, extra-membrane (EM) secondary-structure patterns between TM1–TM2 and TM2–TM3 provide a concrete example of lineage-specific architectural change captured by the indel-aware reconstructions (Fig. 3). In extant structures, SzR4 displays short β-strands between TM2 and TM3 and lacks extended secondary structure between TM1 and TM2, whereas HeR 48C12 features a short α-helix between TM2 and TM3 together with long β-strands between TM1 and TM2 (Fig. 3A, B). The indel-corrected ancestors recapitulate these motifs in a clade-specific manner: Anc-SzR retains SzR4-like short β-strands between TM2 and TM3 and a loop-like TM1–TM2 region, whereas Anc-HeR retains the HeR-like combination of a TM2–TM3 short helix and TM1–TM2 long β-strands. By contrast, Anc-SH preserves only the SzR-like TM2–TM3 β-strands and lacks the HeR-specific helix/long-β combination, consistent with acquisition of these EM elements along the HeR lineage after divergence from Anc-SH (Fig. 3C). Consequently, indel-aware, node-specific masking preserves informative EM architecture and enables gain/loss of secondary-structure motifs to be traced directly across ancestral nodes.

We next quantified the similarity of indel-corrected ancestors across alignment strategies and relative to extant structural templates. PSI/TM-Coffee– and MAFFT L-INS-i–based reconstructions were highly consistent, with pairwise identities of 88.7% (Anc-SzR), 88.8% (Anc-HeR), and 82.8% (Anc-SH), suggesting convergence on similar solutions despite alignment differences. The indel-corrected ancestors also retained moderate similarity to extant rhodopsins: Anc-SzR (PSI/TM-Coffee) shared 54.0% identity with SzR4, and Anc-HeR shared 51.0% (PSI/TM-Coffee) and 47.1% (MAFFT L-INS-i) identity with HeR 48C12. These identity levels provide a quantitative backdrop for the structural comparisons presented below.

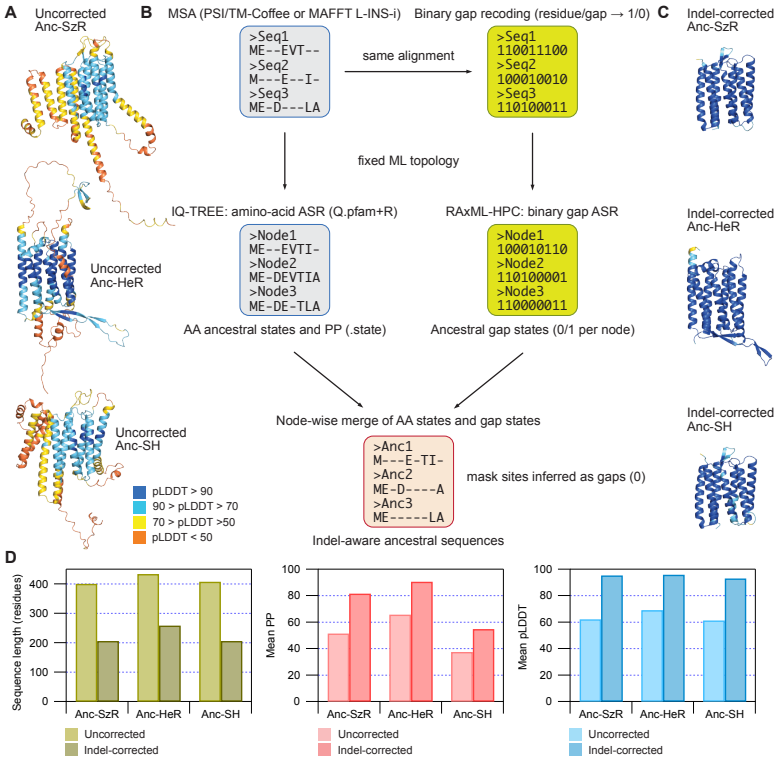

Figure 2: **Indel-aware refinement restores realistic sequence lengths and improves statistical and structural confidence.** (A) AlphaFold monomer models for uncorrected ancestral sequences inferred directly from the full-length PSI/TM-Coffee alignment under Q.pfam+R7, without explicit indel treatment. Structures are colored by per-residue pLDDT using the standard AlphaFold scale. All three ancestors form overall 7TM-like folds but exhibit overextended extra-membrane segments and low-confidence termini. (B) Schematic overview of the indel-aware ASR pipeline. The same full-length amino-acid alignment is used for amino-acid ASR in IQ-TREE and for binary gap coding followed by ancestral gap-state inference in RAxML on the fixed IQ-TREE topology. Node-specific gap reconstructions are then merged with the IQ-TREE `.state` files to mask sites inferred as gaps at each internal node, yielding indel-corrected ancestral sequences. (C) AlphaFold monomer models for the indel-corrected ancestors obtained after node-specific gap masking, colored as in panel (A). Indel-aware correction contracts overextended extra-membrane regions, removes low-confidence tails, and yields compact 7TM folds with uniformly high pLDDT. (D) Effects of indel-aware refinement on sequence length and confidence metrics for Anc-SzR, Anc-HeR and Anc-SH. Bar plots show sequence length, mean posterior probability (PP), and mean pLDDT for uncorrected and indel-corrected ancestors inferred from the PSI/TM-Coffee alignment under Q.pfam+R7. Indel-aware refinement increases both mean PP and mean pLDDT while restoring realistic lengths (see Table 3).

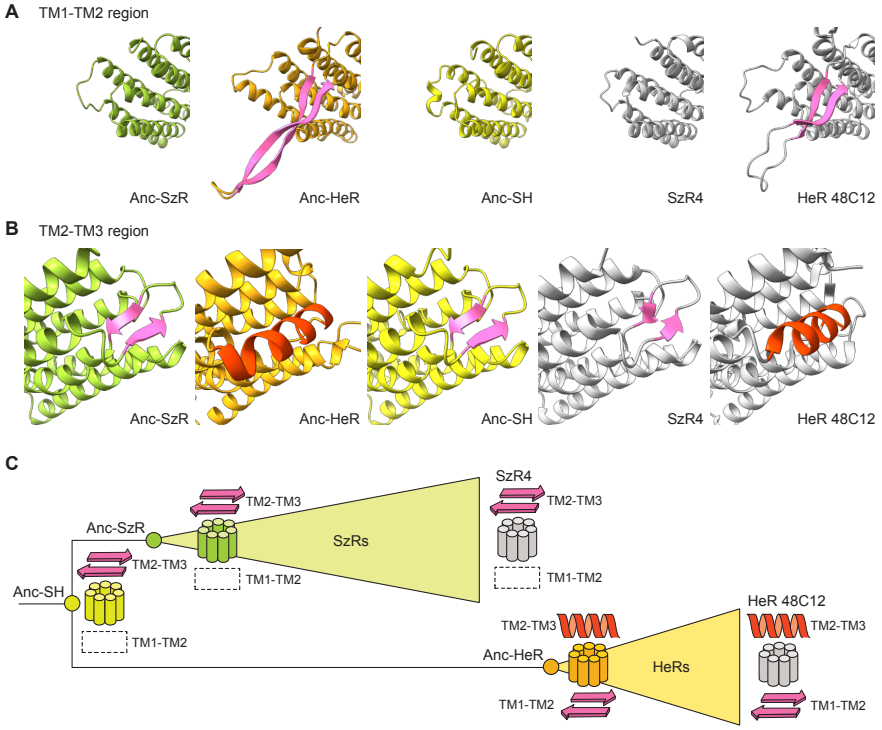

Figure 3: **Evolutionary gain and loss of extra-membrane secondary-structure motifs between TM1–TM2 and TM2–TM3.** (A) Close-up views of the extra-membrane segment between TM1 and TM2 for indel-corrected Anc-SzR (green), Anc-HeR (orange), Anc-SH (light green), and the extant structures SzR4 (PDB 7e4g) and HeR 48C12 (PDB 6su3; both light gray), aligned to their respective structural templates. β-strands in this region are highlighted in pink. Anc-HeR and HeR 48C12 share long β-strands between TM1 and TM2, whereas Anc-SzR, Anc-SH, and SzR4 lack defined secondary structure at this position and form a loop. (B) Corresponding close-up views of the extra-membrane segment between TM2 and TM3. β-strands are highlighted in pink and short α-helices in red. In Anc-SH, Anc-SzR, and SzR4, this region forms short β-strands, whereas in Anc-HeR and HeR 48C12 it forms a short α-helix, indicating a lineage-specific shift in secondary-structure motifs along the HeR branch. (C) Schematic summary of motif gain and loss mapped onto a simplified HeR–SzR phylogeny. Each ancestor or extant protein is represented by a seven-cylinder cartoon indicating the TM helices, with the TM2–TM3 and TM1–TM2 regions annotated. Pink arrows denote β-strands, red ribbons denote short α-helices, and dashed boxes mark loop-like regions lacking defined secondary structure. Together, these panels indicate that Anc-SzR and SzR4 largely retain the Anc-SH pattern (short β-strands between TM2 and TM3 and a loop between TM1 and TM2), whereas the HeR lineage loses the TM2–TM3 β-strands, gains a short α-helix in their place, and acquires long β-strands between TM1 and TM2.

Region-level reliability across membrane-embedded and extra-membrane regions

To quantify region-level reliability, we partitioned each ancestral model into TM and EM regions using membrane boundaries from the Orientations of Proteins in Membranes (OPM) database [33]. For SzR and HeR, we used the OPM entries for SzR4 and HeR 48C12, respectively. TM regions were defined as residues embedded within the OPM membrane slab (i.e., the membrane-embedded cores of the seven helices), whereas EM regions included all residues outside the slab, including helix termini, β-strands, and short coils. According to this convention, helical residues outside the membrane slab are classified as EM.

Using the structure-consistent alignments, OPM-derived TM/EM annotations were transferred from SzR4 and HeR 48C12 to the indel-corrected ancestors, and region-level reliability metrics were computed from PP and AlphaFold confidence (pLDDT) (Fig. 4A, Table 4, and Tables S4–S7). For each region, we summarized mean PP, mean pLDDT, and two combined indices (PP$\times$pLDDT and PP$-$pLDDT). Anc-SH was not included in Table 4 because its hybrid SzR/HeR character precludes a single unambiguous template mapping. However, its global trends are described qualitatively below.

Three consistent patterns emerged. First, TM regions were uniformly high-confidence. Across both alignments, TM segments in Anc-SzR and Anc-HeR exhibited high mean PP (approximately 82–94%) together with near-ceiling mean pLDDT values (>95) (Fig. 4B,C and Table 4). Accordingly, PP$\times$pLDDT was highest in TM regions, indicating strong agreement between sequence-based support and structure-based confidence in the membrane-embedded helix cores.

Second, EM regions also showed consistently high reliability. Although EM values were modestly lower than TM on average, mean PP and pLDDT remained high, with EM segments typically showed a mean PP of $\sim$76–88% and mean pLDDT of $\sim$89–94 (Fig. 4B,C and Table 4). PP$\times$pLDDT likewise indicated that many extra-membrane residues are well supported despite alignment uncertainty. Notably, EM segments corresponding to the characteristic β-strand motifs in SzR4 and HeR 48C12 retained high PP and pLDDT in Anc-SzR and Anc-HeR,

indicating that these lineage-specific EM architectures are recovered as structured elements rather than being erased by indel-aware refinement.

Third, localized reductions in support were concentrated in TM1 and in peri-membrane EM segments flanking TM5. Across ancestors and alignments, PP was consistently lower in these regions, mirrored by reduced PP×pLDDT in the corresponding helix/loop segments (Tables S4 and S6). These sites likely represent evolutionarily labile positions where fold-level constraints are maintained but residue identities vary more freely.

The PP−pLDDT difference provided complementary local insight. While region-averaged PP−pLDDT values were similar between TM and EM segments (Fig. 4B,C and Table 4), site-wise profiles separated two common scenarios: positions with low PP but high pLDDT, where the fold is well defined despite ambiguity in residue identity, and positions with high PP but lower pLDDT, which tend to occur in intrinsically flexible EM segments. In this way, PP−pLDDT helps distinguish sequence-level uncertainty from structural uncertainty at the residue scale.

These region-level analyses showed that OPM-guided TM/EM partitioning provides a more informative view than a simple TM-versus-loop dichotomy, and that the indel-corrected full-length ancestors preserve substantial, lineage-specific extra-membrane architecture.

Node-level reliability: divergence between evolutionary and structural support

We next assessed node-level reliability for three focal ancestors: Anc-SzR, Anc-HeR, and their MRCA Anc-SH. For each node, we summarized (i) branch support from the UFBoot2 consensus tree for the corresponding split, (ii) mean posterior probability (PP) across sites in the indel-corrected ancestor, and (iii) mean pLDDT from the AlphaFold model (Table 2). When a split was not recovered in the UFBoot2 consensus (i.e., collapsed into a polytomy), we set its support to 0 to reflect unresolved topology. UFBoot2 and SH-aLRT values on the ML tree are shown in Fig. 1, and additional support summaries are provided in Table 2.

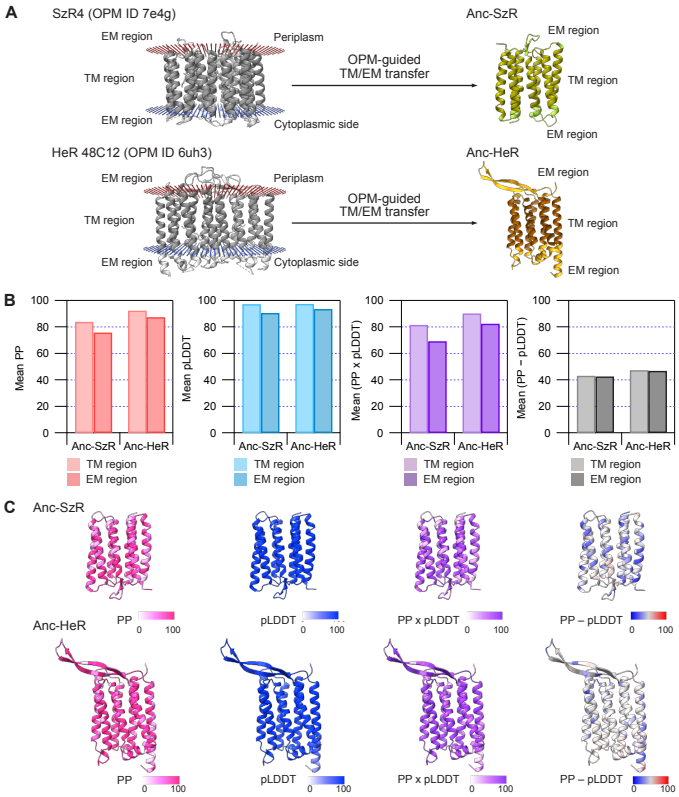

Figure 4: **Region-level reliability of TM and EM segments in Anc-SzR and Anc-HeR.** (A) OPM-guided transfer of transmembrane (TM) and extra-membrane (EM) annotations from extant structures to indel-corrected ancestors. OPM-derived TM/EM boundaries for SzR4 and HeR 48C12 are projected onto Anc-SzR and Anc-HeR via the PSI/TM-Coffee alignment under Q.pfam+R7. In the ancestral models, TM residues are colored dark green (Anc-SzR) or brown (Anc-HeR), and EM residues light green (Anc-SzR) or orange (Anc-HeR). (B) Region-level summary statistics for Anc-SzR and Anc-HeR. For each ancestor, bar plots report mean PP, mean pLDDT, mean (PP×pLDDT), and mean (PP−pLDDT) separately for TM and EM regions defined in panel (a). All metrics are shown on a 0–100 scale. TM segments display uniformly high confidence, whereas EM regions show only slightly reduced but still high values, indicating that many extra-membrane segments are reliably reconstructed. (C) Per-residue reliability mapped onto Anc-SzR and Anc-HeR structures. For each ancestor, four views are shown with Cα atoms colored by PP (white to red), pLDDT (white to blue), PP×pLDDT (white to purple), or PP−pLDDT (blue through gray to red, indicating negative to positive), using the same 0–100 scale. These maps highlight that TM cores are consistently high-confidence, whereas local pockets of reduced PP with high pLDDT in EM segments mark positions where the fold is well defined but residue identity remains more ambiguous. Per-residue reliability mapped onto Anc-SH is shown in Fig. S3.

Anc-SzR showed consistently strong support across all metrics. In the UFBoot2 consensus trees, the Anc-SzR split was recovered with high frequency (99% under PSI/TM-Coffee; 96% under L-INS-i), and the indel-corrected Anc-SzR sequences had high mean PP (80–82%) and mean pLDDT (94–96). TM segments in particular reached near-ceiling pLDDT values together with high PP, consistent with a well-resolved split and a structurally coherent 7TM ancestor. Collectively, these metrics identify Anc-SzR as a robust target for experimental resurrection.

Anc-HeR showed a contrasting pattern: strong sequence/structure support but alignment-sensitive topological support. Under PSI/TM-Coffee, the Anc-HeR split appeared in the UFBoot2 consensus with only moderate support (54%), despite very high mean PP (∼91%) and mean pLDDT (∼96) for the indel-corrected Anc-HeR sequence. Under L-INS-i, the same split collapsed in the UFBoot2 consensus (treated as 0% support), again despite similarly high PP and pLDDT. This discrepancy indicates that bootstrap replicates frequently sample alternative, near-equivalent arrangements within the HeR subtree. Thus, Anc-HeR is well supported conditional on a chosen local topology, but the local HeR branching order remains sensitive to alignment details.

Anc-SH showed the opposite trade-off. The deeper split separating HeRs and SzRs from the outgroup had maximal UFBoot2 support (100%) across all conditions, indicating a robust topology. In contrast, the indel-corrected Anc-SH sequences had substantially lower mean PP (∼55–57%) than the terminal ancestors, even though mean pLDDT remained high (∼93–94) and the overall 7TM fold was well formed. Thus, the deep split is stable, but residue identities at Anc-SH are more diffuse, with many sites supported by multiple plausible amino-acid states.

Together, these contrasts demonstrate that no single metric—branch support, PP, or pLDDT—adequately captures ASR reliability. Stable topology can coexist with diffuse residue support, and high PP/pLDDT can occur despite alignment-sensitive branching. Therefore, we introduce a composite index that integrates all three dimensions into a node-level reliability score.

Composite reliability metric integrating PP, pLDDT, and BS

To integrate evolutionary, statistical, and structural evidence, we defined a composite node-level reliability metric. We first quantified sequence–structure consistency as the mean per-site product $\overline{\mathrm{PP} \times \mathrm{pLDDT}}$, which is high only when both the posterior probability of the inferred residue and the predicted structural confidence at that site are high. We then modulated this value by branch support for the corresponding split, measured as the UFBoot2 frequency in the bootstrap consensus tree, to obtain

$$\mathrm{CompositeScore}_{0\text{–}100} = 100 \times b \times \overline{\mathrm{PP} \times \mathrm{pLDDT}}, \tag{1}$$

where $b = \mathrm{BS}_{\mathrm{UFBoot2}}/100$ is the UFBoot2 support expressed as a fraction (0–1), and PP and pLDDT are each scaled to 0–1 before taking their product and averaging across sites. When a split was absent from the UFBoot2 consensus (collapsed into a polytomy), we set $b = 0$. Composite scores were reported on a 0–100 scale for interpretability (Table 2).

Applying this metric highlights a clear contrast: Anc-SzR is consistently robust, Anc-SH is intermediate, and Anc-HeR is strongly dependent on local topology. For PSI/TM-Coffee, mean PP×pLDDT values were ∼78, 87, and 52 for Anc-SzR, Anc-HeR, and Anc-SH, respectively, while the corresponding UFBoot2 supports were 99%, 54%, and 100% (Fig. 5A; Table 2), yielding composite scores of ∼77, 47, and 52. Under MAFFT L-INS-i, mean PP×pLDDT values were similar (∼76, 86, and 54), but UFBoot2 supports were 96%, 0%, and 100%, giving composite scores of ∼73 for Anc-SzR, 0 for Anc-HeR (collapsed in the consensus), and ∼54 for Anc-SH (Fig. 5B; Table 2). Thus, Anc-SzR remains high across alignments, Anc-SH remains moderate, whereas Anc-HeR collapses when the HeR split is not reproducibly recovered.

This decomposition clarifies why the same dataset can yield "confident" ancestors for different reasons. Anc-HeR combines very high PP×pLDDT with topology-sensitive $\mathrm{BS}_{\mathrm{UFBoot2}}$; accordingly, its composite score is high under PSI/TM-Coffee but strongly suppressed under L-INS-i, matching the local instability described above. By contrast, Anc-SzR pairs high PP×pLDDT with

consistently strong UFBoot2 support under both alignments, making it the most conservative target for experimental resurrection. Anc-SH shows the converse: consistently maximal UFBoot2 support but lower PP×pLDDT, underscoring that deep nodes can be topologically stable yet residue-wise diffuse (Fig. 5C).

We emphasize that this composite score is a decision-support summary rather than a substitute for inspecting BS, PP, and pLDDT individually. Nodes that are collapsed or extensively rearranged in the UFBoot2 consensus—or whose support is clearly alignment-dependent—should be flagged explicitly, and their composite scores interpreted as conditional. In this sense, Anc-HeR represents a "conditionally reliable" ancestor: its site-wise and structural confidence are strong, but its local evolutionary context remains sensitive to alignment choice.

Oligomeric-state prediction supports dimeric HeR and trimeric SzR ancestors

To test whether the resurrected ancestors encode biologically plausible quaternary structure, we predicted homooligomeric assemblies and asked whether the preferred stoichiometries match those observed for extant rhodopsins. Using indel-corrected sequences inferred from the PSI/TM-Coffee alignment, we modeled Anc-SzR, Anc-HeR, and Anc-SH as dimers, trimers, and pentamers with an AlphaFold multimer workflow (Fig. 6). We focused on stoichiometries of 2, 3, and 5 because microbial rhodopsins are most commonly reported as dimers, trimers, or pentamers, whereas tetrameric assemblies are not well established for this family [34, 20]. For each target and stoichiometry, we recorded the predicted complex TM-score (pTM) and interface TM-score (ipTM) as summaries of overall fold confidence and interface quality, respectively [35]. Results for both alignment strategies are reported in Table S8.

Anc-SzR was strongly predicted to form a trimer (Fig. 6A, C). Under PSI/TM-Coffee, the trimer achieved high confidence (ipTM = 0.88; pTM = 0.90), whereas dimeric and pentameric assemblies scored substantially lower (Table S8). MAFFT L-INS-i produced the same ranking, again favoring the trimer over the dimer and pentamer (Table S8). This preference mirrors the trimeric assemblies reported for extant SzRs and suggests that both the 7TM core and the reconstructed EM

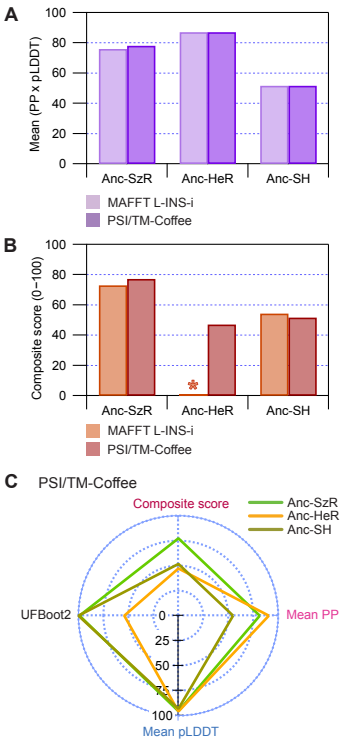

Figure 5: **Node-level sequence–structure consistency and composite reliability across alignments.** (A) Mean PP×pLDDT for Anc-SzR, Anc-HeR and Anc-SH under the PSI/TM-Coffee and MAFFT L-INS-i alignments. Values are averaged over Cα sites and shown on a 0–100 scale. For all three nodes, mean PP×pLDDT is high and broadly similar between alignments, indicating strong overall sequence–structure consistency in the reconstructed ancestors. (B) Composite node-level reliability scores (0–100) combining UFBoot2 branch support and mean PP×pLDDT for the same nodes and alignments. The composite score is defined as $\mathrm{BS}_{\mathrm{UFBoot2}} \times \overline{\mathrm{PP} \times \mathrm{pLDDT}}$ (see Methods). For the MAFFT L-INS-i alignment, the Anc-HeR split is not recovered in the UFBoot2 consensus (collapsed branch; $\mathrm{BS}_{\mathrm{UFBoot2}} = 0$); its composite score is therefore 0 and is marked with an asterisk. Across both alignments, Anc-SzR attains consistently high composite scores, Anc-SH reaches intermediate values, and Anc-HeR is conditionally reliable, with a high composite score only when the HeR split is recovered in the consensus. (C) Radar plot summarizing node-level metrics for the PSI/TM-Coffee alignment. Anc-SzR (green), Anc-HeR (orange) and Anc-SH (dark yellow) are shown with axes corresponding to mean PP, mean pLDDT, mean PP×pLDDT and the composite score (all on a 0–100 scale). The plot highlights that Anc-SzR is uniformly strong across all metrics, Anc-HeR combines very high sequence–structure consistency with alignment-sensitive branch support, and Anc-SH exhibits robust topology but lower residue-level certainty.

elements support a plausible trimeric ancestral interface.

In contrast, Anc-HeR was strongly predicted to form a dimer (Fig. 6B, C). Under PSI/TM-Coffee, the dimer achieved high confidence (ipTM = 0.81; pTM = 0.86), while higher-order assemblies were markedly weaker (Table S8). MAFFT L-INS-i yielded the same ranking, supporting that the preference for a dimeric Anc-HeR is robust to alignment strategy (Table S8). Consistent with this, the dimeric models suggest an interface involving the EM β-strand region between TM1 and TM2, in line with the proposed role of this motif in extant HeRs.

The deeper Anc-SH showed intermediate behavior: trimeric models (ipTM = 0.67; pTM = 0.73) were preferred over dimers and pentamers, but with lower interface confidence than the terminal ancestors (Fig. 6C; Table S8), consistent with Anc-SH's deeper phylogenetic position, lower mean PP, and mixed HeR/SzR character. Overall, these multimeric predictions support the conclusion that the indel-corrected ancestors encode oligomerization interfaces consistent with the trimeric and dimeric states observed for SzRs and HeRs, respectively.

Expression and spectral characterization of ancestral rhodopsins

To evaluate the ability of reconstructed ancestral rhodopsins to be expressed and bind to retinal, we focused on Anc-SzR and Anc-HeR and expressed them heterologously in *E. coli* under multiple induction conditions. Importantly, the constructs were designed to use the ASR-derived sequences essentially "as reconstructed". For each protein, the inferred ancestral sequence (PSI/TM-Coffee reconstruction) was cloned as a single open reading frame into a standard pET vector, with only a vector-encoded His_6 affinity tag. Anc-SzR was inserted into pET-21a(+) with a C-terminal His_6 tag, whereas Anc-HeR was inserted into pET-15b with an N-terminal His_6 tag and a thrombin recognition site upstream of the ancestral open reading frame. No additional fusion partners, solubility tags, truncations, codon optimization, or stabilizing mutations were introduced. Thus, any rhodopsin-like pigmentation and spectral features observed in *E. coli* directly reflect the intrinsic properties of the full-length ancestral sequences produced by our pipeline.

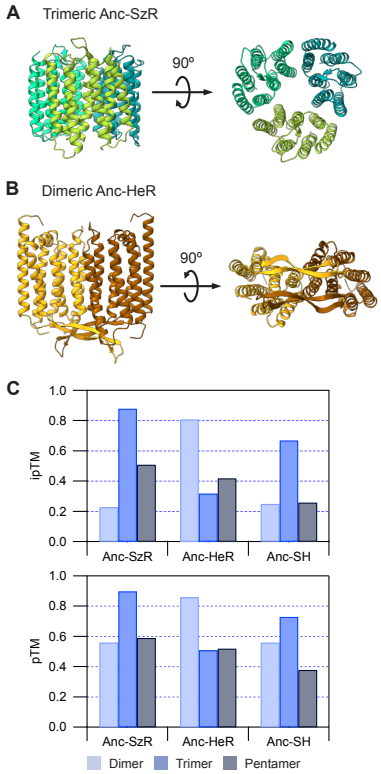

Figure 6: **Oligomeric-state predictions support trimeric Anc-SzR and dimeric Anc-HeR.** (A) AlphaFold-Multimer prediction of trimeric Anc-SzR viewed parallel to the membrane (left) and from the cytoplasmic side (right). Each protomer is shown as a cartoon and colored by chain. The predicted assembly forms a compact trimeric arrangement of seven-helix bundles. (B) AlphaFold-Multimer prediction of dimeric Anc-HeR viewed parallel to the membrane (left) and from the extracellular side (right), with protomers colored by chain. The predicted interface includes extra-membrane elements between TM1 and TM2, consistent with the location of proposed dimerization motifs in extant HeRs. (C) Predicted interface TM-scores (ipTM) and overall complex TM-scores (pTM) for dimeric, trimeric and pentameric assemblies of Anc-SzR, Anc-HeR and Anc-SH, based on indel-corrected sequences inferred from the PSI/TM-Coffee alignment. Bars show the top-ranked AlphaFold-Multimer model for each stoichiometry. Anc-SzR attains the highest ipTM/pTM as a trimer, Anc-HeR as a dimer, and Anc-SH shows intermediate ipTM values with a modest preference for trimeric assemblies. Tetrameric assemblies were not evaluated, as microbial rhodopsins are predominantly reported as dimers, trimers or pentamers rather than tetramers. The same qualitative ranking of preferred stoichiometries was observed for L-INS-i–based ancestors (Table S8).

For both Anc-SzR and Anc-HeR, we compared three induction regimes after addition of IPTG and all-*trans*-retinal: 37 °C for 4 h, 37 °C for 20 h, and 25 °C for 20 h. Under all conditions tested, harvested cell pellets displayed a visible reddish/purple coloration, consistent with expression of retinal-binding rhodopsins. Pigmentation was strongest after 20 h induction (both at 37 °C and 25 °C) and weaker after 4 h at 37 °C. We therefore adopted 25 °C for 20 h as the standard condition for subsequent preparations of both ancestors. Under this condition, both constructs reproducibly yielded strongly pigmented pellets (Fig. 7A), indicating efficient folding and chromophore incorporation in vivo without engineering beyond a simple His_6 tag.

Detergent-solubilized Anc-SzR and Anc-HeR were further assessed by UV–visible absorption spectroscopy (Fig. 7B). Purified Anc-SzR exhibited a dominant absorption maximum at 549 nm, whereas purified Anc-HeR showed a λ_{max} at 543 nm. These maxima are within the ranges reported for extant schizorhodopsins and heliorhodopsins, respectively, supporting that each reconstructed ancestor forms a functional retinal-binding holoprotein with clade-typical spectral properties. The overall spectral line shapes were also characteristic of a protein-bound retinal chromophore in an ordered environment, consistent with the compact, high-confidence 7TM folds predicted by AlphaFold. Photographs of the purified samples (Fig. 7B, inset) further illustrate the characteristic visible coloration of the ancestral rhodopsins.

Discussion

This study provides a full-length case study of ASR in 7TM microbial rhodopsins, focusing on SzRs and HeRs. By combining structure-consistent alignments, profile-based evolutionary models, explicit indel-aware refinement, and structure-based evaluation, we reconstructed compact full-length ancestors (Anc-SzR, Anc-HeR, and Anc-SH) and demonstrated experimental tractability by expressing Anc-SzR and Anc-HeR in *E. coli* as stable, colored holoproteins with rhodopsin-like absorption maxima. These results demonstrate that the

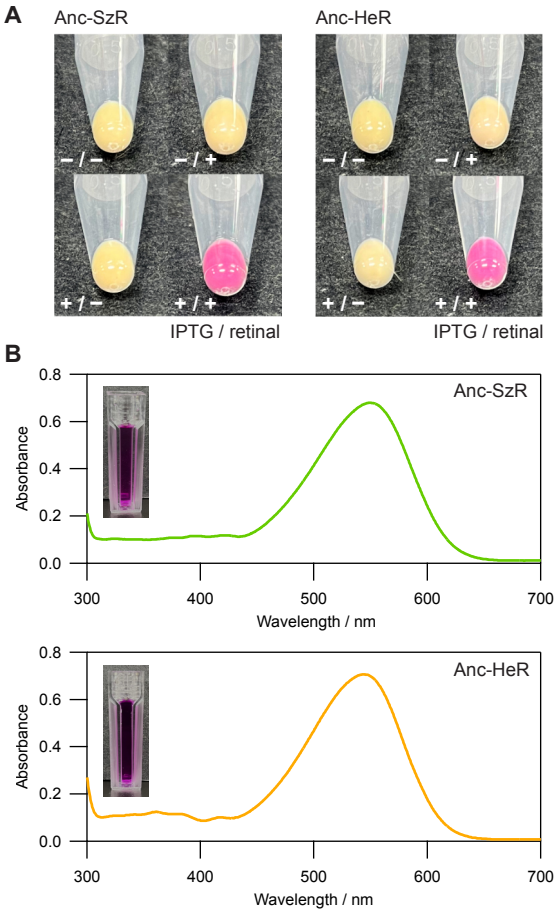

Figure 7: **Expression and spectral properties of Anc-SzR and Anc-HeR.** (A) Whole-cell pigmentation of *E. coli* expressing Anc-SzR or Anc-HeR under four combinations of IPTG induction and all-*trans* retinal supplementation (−/−, −/+, +/−, +/+). Strong visible coloration is observed in induced cells supplied with retinal (+/+), whereas non-induced and/or retinal-free controls remain weakly colored, indicating retinal-dependent formation of holoproteins. (B) UV–visible absorption spectra of purified, detergent-solubilized Anc-SzR and Anc-HeR recorded at room temperature after affinity purification. Insets show photographs of the purified samples in cuvettes, illustrating the characteristic visible colors of the ancestral rhodopsins.

resurrection of full-length 7TM ancestors is feasible without stabilizing engineering and that EM architecture can be reconstructed and evaluated rather than routinely trimmed away.

A central practical finding from the HeR–SzR system is that alignment quality can dominate local phylogenetic stability, especially within the HeR clade. Across datasets and conditions, Q.pfam+R models were consistently preferred over conventional single-matrix models (Table 1), supporting the use of membrane-aware profile models for these sequences. However, the strongest qualitative differences we observed were not between Q.pfam+R and alternative amino-acid models, but between alignment strategies: PSI/TM-Coffee preserved TM registers and produced a HeR subtree that was more consistently recovered in UFBoot2 consensus trees. In contrast, MAFFT L-INS-i yielded a locally unstable resolution around the HeR ancestor (Fig. 1). In other words, for this case study, model choice improved global fit, but alignment quality governed whether key internal HeR splits were repeatedly recovered under resampling.

A second key finding is that explicit indel-aware refinement is essential for realistic full-length ancestors in this system. Gap-unaware ASR on full-length alignments produced markedly overextended ancestors with low-confidence tails in EM regions (Fig. 2A), a failure mode that is easy to miss if one focuses only on the TM core. Indel-aware refinement, implemented here by combining amino-acid ASR from IQ-TREE with node-specific binary gap inference in RAxML on a fixed topology, restored lengths that closely match extant structures and simultaneously increased both PPs and pLDDT in the retained sites (Fig. 2C, D). This "before/after" contrast is especially informative for 7TM proteins because termini and EM motifs can contribute to folding, oligomerization, and function; in the SzR–HeR case, indel-aware correction converts diffuse, elongated ancestors into compact sequences whose 7TM folds and EM motifs are predicted with uniformly high confidence.

Within this full-length framework, the EM regions of SzRs and HeRs proved more interpretable than commonly assumed. Using OPM-defined membrane boundaries to partition membrane-embedded TM cores from all residues outside

the bilayer, we found that EM segments retained high mean PP and high mean pLDDT, only modestly reduced relative to TM regions (Fig. 4 and Table 4). Importantly, EM secondary-structure motifs were recovered in a lineage-specific manner rather than being erased by indel-aware correction. In HeR, the long β-strands between TM1 and TM2 and the short helix between TM2 and TM3 seen in HeR 48C12 were recapitulated in Anc-HeR, whereas SzR-like ancestors retained short β-strands between TM2 and TM3 and a loop-like TM1–TM2 region (Fig. 3A, B). In this case study, these motifs behave as reconstructable evolutionary characters: Anc-SH retains the SzR-like pattern, while the HeR lineage shows a gain/loss and relocation of EM elements along the branch to Anc-HeR (Fig. 3C). This provides a concrete example of how full-length ASR enables evolutionary statements about EM architecture rather than restricting inference to the TM core alone.

By comparing node-level metrics, we also highlight that topological support, residue-level certainty, and structural confidence can diverge in ways that matter for experimental prioritization. Anc-SzR is robust across alignments, with high UFBoot2 support, high mean PP, and high mean pLDDT, making it a conservative resurrection target (Fig. 1 and Table 2). Anc-HeR, by contrast, shows very high PP and pLDDT once a particular local HeR topology is chosen, but its surrounding subtree is alignment-sensitive, leading to moderate or absent support for the corresponding split in bootstrap consensus trees depending on alignment (Fig. 1 and Table 2). Anc-SH illustrates the complementary pattern: the deep split is topologically stable, yet residue identities are more diffuse (lower mean PP) despite a confident overall fold (high pLDDT). For this system, therefore, structural plausibility does not imply a uniquely resolved evolutionary context. Conversely, robust topology does not guarantee high residue-level certainty at deep nodes. The composite BS–PP–pLDDT index used here is best viewed as a compact decision-support summary of these three dimensions, while still requiring inspection of each component for alignment-sensitive nodes such as Anc-HeR (Fig. 5).

Our multimeric predictions further support the biological plausibility of the

reconstructed sequences. AlphaFold-Multimer favored a trimeric assembly for Anc-SzR and a dimeric assembly for Anc-HeR, mirroring the known oligomeric tendencies of extant SzRs and HeRs [19, 20, 21, 22, 23, 24] (Fig. 6 and Table S8). Notably, the preferred oligomeric states were supported by high interface scores (ipTM) only for the corresponding stoichiometries, while alternative assemblies scored substantially lower. In this case study, these predictions strengthen the interpretation that indel-aware full-length ancestors preserve not only monomeric 7TM folds but also EM elements that are compatible with oligomerization interfaces.

A primary contribution of this work is that the computational reconstructions translate into experimentally tractable proteins. Anc-SzR and Anc-HeR were expressed essentially "as reconstructed", using only minimal His_6 tags from standard pET vectors, without solubility fusions, truncations, or stabilizing point mutations. Both constructs yielded visibly pigmented *E. coli* pellets under induction with all-*trans* retinal (Fig. 7A), and purified proteins showed rhodopsin-like absorption maxima consistent with their clades (Fig. 7B). This closes a practical loop that is rarely demonstrated for full-length 7TM ASR: in this SzR–HeR case study, sequence-level support (PP), structure-level confidence (pLDDT), and biochemical tractability (retinal binding and visible spectra) converge on the same reconstructed ancestors.

Placing this result in context, binary gap–coding strategies similar in spirit have been applied previously to microbial rhodopsins to probe ancient photic niches; however, prior work primarily emphasized TM-core inferences (e.g., spectral tuning) and did not evaluate full-length extra-membrane architecture or experimental tractability via protein resurrection [36]. Our results extend these computational reconstructions by showing that full-length, indel-aware ancestors can be expressed as functional holoproteins, enabling joint investigation of EM architecture and the 7TM core within a single workflow.

Finally, we briefly note that Anc-HeR is sensitive not only to alignment choices but also to pipeline/model combinations, which helps interpret the supporting analyses in Table S3. In our comparisons, IQ-TREE-based reconstructions using

Q.pfam+R7 (and robustness checks using LG+C60+F+R8) produced Anc-HeR sequences that were mutually similar across alignments. In contrast, a classical RAxML–PAML reconstruction under an LG-based model generated an Anc-HeR sequence that was far more divergent (Table S3). In this system, the discrepancy is plausibly driven primarily by model differences (profile-based versus LG-type) as much as by software per se; the key point for the case study is that the HeR ancestor is a "fragile" node whose inferred sequence can shift substantially under reasonable alternative modelling choices, while Anc-SzR is comparatively robust. Practically, this motivates treating Anc-HeR as conditionally reliable and explicitly reporting sensitivity analyses for such nodes, rather than assuming that a single pipeline provides a definitive answer.

Several limitations delimit the scope of this case study. First, AlphaFold confidence metrics are strong indicators of structural plausibility but do not replace experimental validation of fold, oligomeric state, and function. Second, our HeR and SzR sampling differs in provenance and annotation depth, and some apparent EM diversity may reflect sampling biases. Third, the composite reliability index inherits assumptions from bootstrap resampling, ASR model specification, and structure prediction, and should be interpreted as an integrative summary rather than a standalone truth metric. Despite these caveats, the SzR–HeR system demonstrates a clear outcome: full-length, indel-aware ancestral reconstruction can yield compact 7TM sequences with coherent EM architecture, and at least two such ancestors can be expressed and recovered as retinal-binding rhodopsins without stabilizing engineering. These results establish a practical path for extending experimental ASR beyond trimmed TM cores and toward direct tests of how extra-membrane architecture co-evolves with the 7TM fold.

Conclusion

In this study, we combined an indel-aware ASR workflow with AlphaFold-based confidence mapping to reconstruct full-length ancestral 7TM microbial rhodopsins, explicitly including not only the transmembrane helices but also

extra-membrane regions. Crucially, the resulting ancestors were experimentally tractable: Anc-SzR and Anc-HeR were expressed essentially "as reconstructed" in *E. coli* and recovered as stable, retinal-binding holoproteins with rhodopsin-like visible absorption maxima. These results demonstrate that full-length, indel-aware ASR can yield compact 7TM ancestors that are supported by both statistical (PP) and structural (pLDDT) evidence and can be directly tested by protein resurrection.

In microbial rhodopsins, functional diversification is often dominated by local changes around the retinal-binding pocket within the TM core, and extracellular loops may appear secondary at first glance. Nevertheless, our reconstructions suggest that substantial portions of EM segments can be inferred with only modestly reduced confidence relative to TM regions and recovered as structurally coherent motifs. These EM elements provide a consistent structural context for interpreting ancestral photochemistry, oligomerization, and potential protein–protein interactions, arguing against routine aggressive trimming in 7TM ASR.

More broadly, many eukaryotic 7TM receptors, including GPCRs, rely on extracellular and cytoplasmic loops and C-terminal tails as primary determinants of ligand recognition and G-protein/arrestin coupling. While extending the present approach to such systems will require careful attention to alignment depth, topology, and lineage-specific indel regimes, our results motivate the use of gap-aware full-length reconstruction strategies in 7TM families where extra-membrane architecture is integral to function. One practical contribution of this work is to provide an explicitly documented, scriptable pipeline (ConsistASR: Consistency-aware ASR), including full control files for PAML-based reconstructions, so that each step from alignment to indel-aware ASR and structural evaluation can be replicated or adapted in other systems. We anticipate that the ConsistASR workflow introduced here will serve as a reproducible starting point for reconstructing and engineering ancestral 7TM proteins, treating both the transmembrane core and extra-membrane segments in a physically and statistically consistent manner.

Materials and Methods

Sequence datasets and curation

HeR and SzR sequences were compiled by combining previously published datasets with additional database searches. Previously characterized HeR and SzR sequences were taken from Inoue et al. and Bulzu et al. [21, 37] and supplemented with homologs identified by BLAST searches against NCBI protein databases and by profile-based searches of MGnify metagenome assemblies [38]. Because SzR-like sequences are sparsely represented in standard protein databases, most additional SzR candidates were obtained from metagenomic assemblies. Several HeR sequences were directly taken from the metagenome-derived dataset of Bulzu et al. [37]; for these entries, we retained the original contig/gene identifiers when a unique public accession could not be assigned. All amino-acid sequences used in this study are provided as Supplementary FASTA files.

Candidate sequences were filtered in four stages: (i) *Length filtering*: sequences shorter than 200 amino acids for HeRs or 150 amino acids for SzRs were discarded to remove obvious fragments and mispredicted multi-pass membrane proteins. (ii) *Functional-residue filtering*: sequences lacking the conserved Schiff-base lysine in TM7 were excluded. (iii) *Redundancy reduction*: highly similar sequences (>95% pairwise identity) were clustered with CD-HIT [39], retaining one representative per cluster. (iv) *Phylogenetic diversity filtering*: Treemmer [40] was applied to a preliminary tree to downsample near-identical sequences while preserving long internal branches.

After automated filtering, we performed a topology-consistency screen using the PSI/TM-Coffee transmembrane-mode alignment [28]. Because microbial rhodopsins have a conserved 7TM core, we required that the alignment contained continuous, gap-poor registers across TM1–TM7. Sequences were excluded if they showed obvious loss of TM registers (e.g., long internal deletions or extensive gaps within one or more predicted TM helices), or if conserved core positions—including the Schiff-base lysine in TM7—were missing due to trunca-

tion or misprediction. This curation step was applied only to remove sequences incompatible with a canonical 7TM rhodopsin fold; no additional per-sequence TM-helix prediction beyond the alignment-based TM-mode constraints was performed. The final non-redundant dataset contained 228 sequences (115 HeRs, 110 SzRs and 3 outgroup microbial rhodopsins). Two reduced datasets, HeR+outgroup and SzR+outgroup, were derived from this set, sharing the same outgroup panel. Sequence lengths ranged from approximately 195 to 305 residues, consistent with known 7TM microbial rhodopsin architectures.

Multiple sequence alignment

Multiple sequence alignments were generated for the full HeR+SzR+outgroup dataset and for the two subset datasets (HeR+outgroup and SzR+outgroup) using PSI/TM-Coffee and MAFFT [28, 29]. Unless otherwise stated, default parameters of the respective programs or web servers were used.

Structure-aware alignments were obtained with PSI/TM-Coffee in transmembrane mode using UniRef100 as the homologous sequence database. In this mode, predicted transmembrane segments and additional homologous sequences are incorporated in the consistency-based alignment procedure, stabilizing TM/EM boundaries. The resulting alignments were exported in full-length FASTA format without manual editing of internal regions.

As a widely used non–structure-guided baseline, we generated MAFFT alignments using the L-INS-i strategy [29]. For exploratory analyses and for a classical workflow comparison, additional alignments were generated with the faster FFT-NS-1 strategy in MAFFT; because FFT-NS-1 showed visible mismatches at TM/EM boundaries and yielded unstable HeR reconstructions in some conditions, it was used only in sensitivity analyses and in the classical pipeline comparison (see below).

To examine the influence of mild gap-based trimming, we also produced trimmed versions of each alignment using trimAl with the `gappyout` option [41]. These trimmed alignments were used solely to assess robustness of model selection and topology. All main ASR, indel-aware refinement, and AlphaFold-

based structural evaluations were performed on untrimmed, full-length alignments (i.e. no aggressive manual trimming of loops or termini was applied).

Phylogenetic inference and model selection

Maximum-likelihood (ML) phylogenies were inferred with IQ-TREE v3.0.1 [31] for the HeR+SzR+outgroup dataset and for the HeR+outgroup and SzR+outgroup subsets. Alignments were analyzed as single-partition protein datasets.

For each alignment, ModelFinder [30] (`-m MFP`) was used to compare empirical matrices (e.g. LG, WAG, JTT) with various rate-heterogeneity schemes and profile families including Q.pfam and related models. Model fit was assessed using the Bayesian information criterion (BIC), and the model with the lowest BIC was selected. In the main analyses we adopted PSI/TM-Coffee with the best-fit Q.pfam+R model (typically Q.pfam+R7) as the baseline evolutionary framework (see Results and Table 1), and used MAFFT L-INS-i as an alignment sensitivity check.

ML tree searches were run in IQ-TREE with ultrafast bootstrap approximation (1,000 replicates) and SH-aLRT support (1,000 replicates) [32]. Three outgroup rhodopsins were specified to root the tree. The best-scoring ML tree (`.treefile`) was used for visualization and as the fixed topology in downstream ASR and binary indel reconstructions. The UFBoot2 consensus tree (`.contree`) was used to determine whether focal branches were resolved or collapsed and to obtain the branch support values used in node-level summaries.

To benchmark the impact of increased model complexity, additional trees were inferred under LG+C60+F+R8 using otherwise identical settings. These analyses were treated as robustness checks rather than the primary framework for ASR.

Ancestral sequence reconstruction in IQ-TREE

ASR was performed in IQ-TREE v3.0.1 on the ML topologies inferred under the baseline Q.pfam+R model. For each dataset (HeR+SzR+outgroup, HeR+outgroup, SzR+outgroup) and each alignment (PSI/TM-Coffee, MAFFT

L-INS-i), we carried out marginal ASR at all internal nodes using `--ancestral`, treating the topology as fixed.

IQ-TREE reports, for each internal node and alignment column, posterior probabilities (PPs) over the 20 amino-acid states and the maximum-posterior state. These values were extracted from the `.state` files. We focused on three internal nodes in the full tree: the SzR ancestor (Anc-SzR), the HeR ancestor (Anc-HeR), and their MRCA (Anc-SH). Sensitivity analyses under LG+C60+F+R8 were performed on the same alignments and topologies to assess model-dependent variation in ancestral states and PP distributions (see Results).

Unless otherwise noted, PP-based summaries and all structural evaluations reported in this study refer to the indel-corrected ancestral sequences obtained after the indel-aware refinement described below.

Indel-aware reconstruction using binary gap models

To make indel treatment explicit and node specific, we implemented a two-stage refinement pipeline inspired by binary gap coding approaches [26]. In this workflow, IQ-TREE is used for amino-acid ASR, and RAxML-HPC v8.2.12 [25] is used for binary gap inference on the same fixed topology.

Binary gap inference. For each amino-acid alignment, we generated a binary gap matrix by recoding each column as 1 (residue present) or 0 (gap). RAxML-HPC was then used to infer ancestral gap states under a two-state (gap/non-gap) model on the fixed IQ-TREE ML topology. This yields, for each internal node and alignment column, a binary ancestral call for presence/absence.

Node mapping and node-specific masking. Because IQ-TREE and RAxML assign different internal node labels, we mapped nodes by bipartitions: for each internal node, the set of descendant tips defines a split, and we matched IQ-TREE nodes to RAxML nodes by exact agreement of these leaf-set bipartitions (with outgroup-aware handling of complementary splits). For each ancestor and alignment column, if the binary reconstruction indicated a gap state (0), the corresponding IQ-TREE amino-acid state and PP were masked and removed; if

a residue state (1) was inferred, the IQ-TREE maximum-posterior residue and its PP were retained. Indel-corrected ancestral sequences were then assembled by concatenating unmasked positions in their original alignment order.

All downstream analyses (sequence identity comparisons, regional PP/pLDDT summaries, composite scores, and structure prediction) used these indel-corrected sequences. Gap recoding, RAxML parsing, bipartition-based node mapping, and masking were implemented in custom Python scripts (see Software and reproducibility).

AlphaFold structure prediction and confidence metrics

Predicted structures for ancestral and extant rhodopsins were obtained using an AlphaFold3 web server [27]. For single-chain modeling, each sequence was submitted as a single input chain; for oligomeric modeling, identical chains were provided as multiple sequences in the same submission to obtain homomeric complexes (stoichiometries of 1, 2, 3, and 5). In this web workflow, multimeric prediction is therefore specified by multi-sequence input rather than by an explicit "Multimer" option.

For each submission, five models were generated, and the top-ranked model (server ranking) was used for downstream analyses. Per-residue pLDDT values were extracted from the B-factor field of the output coordinate files; analyses were restricted to C_α atoms, yielding one pLDDT value per residue. These pLDDT values were paired with site-wise PPs from IQ-TREE using alignment coordinates, restricting the mapping to residues retained after indel-aware masking.

To assess oligomeric preferences, Anc-SzR, Anc-HeR, and Anc-SH were modeled as homomeric assemblies with stoichiometries of 1 (monomer control), 2 (dimer), 3 (trimer), and 5 (pentamer). We focused on 2/3/5 because microbial rhodopsins are experimentally observed as dimers, trimers, or pentamers, whereas tetramers have not been reported. For each stoichiometry and ancestor, we recorded pTM and ipTM for the top-ranked model, which quantify predicted fold accuracy and interface quality. Structural inspection and figure preparation were performed in UCSF ChimeraX and PyMOL.

Definition of TM and EM regions using OPM

To quantify region-specific reliability, we partitioned each ancestral model into transmembrane (TM) and extra-membrane (EM) regions using membrane embeddings from the Orientations of Proteins in Membranes (OPM) database [33]. For SzR, we used the OPM entry for SzR4 (PDB 7e4g). For HeR, we used the OPM entry corresponding to the HeR 48C12 structure (PDB 6uh3) used for TM/EM boundary definition. Residues lying within the hydrophobic core of the OPM membrane slab were classified as TM, and all remaining residues were classified as EM. In this convention, EM includes N- and C-terminal tails, peri-membrane helix termini, β-strands, short coils, and amphipathic helices outside the membrane core.

TM/EM labels were transferred to reconstructed ancestors via the same structure-consistent alignments used in ASR. Labels were propagated column-wise from the extant template sequence to the ancestral sequence; sites aligned to gaps in the ancestor or removed by indel-aware masking were excluded. Using these labels, we computed region-level means for PP, pLDDT, PP$\times$pLDDT, and PP$-$pLDDT for TM and EM residues and, where indicated, for individual TM helices and adjacent EM segments.

Composite reliability metric

To integrate evolutionary, statistical, and structural information at the node level, we defined a composite reliability metric for Anc-SzR, Anc-HeR, and Anc-SH. For each ancestor, site-wise marginal PPs were extracted from IQ-TREE `.state` files and mapped onto the indel-corrected sequence; pLDDT values were taken from the corresponding AlphaFold 3 model and normalized to 0–1.

Let PP_i and pLDDT_i denote the posterior probability and normalized pLDDT (0–1) for residue i, and let N be the number of C_α sites considered. We defined

$$\overline{\mathrm{PP}} = \frac{1}{N}\sum_{i=1}^{N}\mathrm{PP}_i, \quad \overline{\mathrm{pLDDT}} = \frac{1}{N}\sum_{i=1}^{N}\mathrm{pLDDT}_i,$$

$$\overline{\mathrm{PP}\times\mathrm{pLDDT}} = \frac{1}{N}\sum_{i=1}^{N}\big(\mathrm{PP}_i\times\mathrm{pLDDT}_i\big), \quad \overline{\mathrm{PP}-\mathrm{pLDDT}} = \frac{1}{N}\sum_{i=1}^{N}\big(\mathrm{PP}_i-\mathrm{pLDDT}_i\big).$$

In the Results, these quantities are reported on a 0–100 scale by multiplying by 100.

Bootstrap support (BS) for each ancestor was taken from the UFBoot2 consensus tree (`.contree`) under the corresponding alignment and model. BS was defined as the fraction of bootstrap trees containing the relevant split, expressed as $b = \mathrm{BS}/100$ on a 0–1 scale. When a branch was collapsed in the consensus (i.e. the split was absent), b was set to 0.

The composite node-level score was defined as

$$\mathrm{CompositeScore}_{0\text{–}100} = 100 \times b \times \overline{\mathrm{PP} \times \mathrm{pLDDT}}.$$

Composite scores were computed separately for each alignment/model combination; baseline values correspond to PSI/TM-Coffee with Q.pfam+R7.

Expression, purification and spectroscopy of ancestral rhodopsins

Genes encoding Anc-SzR and Anc-HeR were synthesized as contiguous open reading frames corresponding exactly to the inferred amino-acid sequences (GenScript), without additional fusion partners or stabilizing mutations. For Anc-SzR, the gene was cloned into the NdeI and XhoI sites of pET-21a(+), placing the ASR-derived open reading frame in frame with the vector-encoded C-terminal His_6 tag. For Anc-HeR, the gene was cloned into the NdeI and BamHI sites of pET-15b, such that the vector-encoded N-terminal His_6 tag and thrombin recognition sequence precede the ASR-derived open reading frame and the C terminus corresponds exactly to the inferred ancestral sequence. For both constructs, the initiator methionine is encoded by the ATG within the NdeI site (CATATG) and is counted as the first residue of the ASR-derived sequence.

Recombinant Anc-SzR and Anc-HeR were expressed in *E. coli* C43(DE3). Cultures were grown in LB medium supplemented with ampicillin at 37 °C and induced at $\mathrm{OD}_{600} \approx 1.0$ with 1 mM IPTG together with 10 μM all-*trans* retinal (ethanol stock). After induction, cultures were incubated for 20 h at 25 °C with shaking, which yielded strongly pigmented cell pellets for both ancestors (see

Results). For qualitative comparison of expression conditions, additional cultures were induced at 37 °C for 4 or 20 h.

For purification, cells were harvested by centrifugation, resuspended in buffer, and disrupted by sonication. Membrane fractions were collected by ultracentrifugation, resuspended in detergent-containing buffer, and solubilized with n-dodecyl-β-D-maltoside (DDM). Solubilized pigments were purified by Co^{2+}-affinity chromatography and eluted in imidazole-containing buffer with DDM. Eluted fractions were exchanged into a Tris/NaCl buffer containing DDM for spectroscopic measurements. Both Anc-SzR and Anc-HeR showed anomalously low apparent molecular weights by SDS–PAGE compared with calculated masses, despite migrating as single, well-resolved bands (Fig. S4). Detailed buffer compositions and chromatography conditions are provided in the Supplementary Methods.

UV–visible absorption spectra were recorded at room temperature using a 10 mm path length quartz cuvette. Spectra were collected over 300–700 nm with the corresponding buffer used for baseline correction.

Classical RAxML+PAML workflow for comparison

As a classical reference point, we also performed ASR using a RAxML+PAML workflow. ML trees were inferred with RAxML-NG v1.2.2 [42] under an empirical amino-acid model with among-site rate heterogeneity (typically LG+F+G4). These topologies were then supplied to codeml (PAML v4.10.7) [43] for marginal ASR under a fixed LG model with gamma-distributed rate variation (four categories). Full control files and parameter settings are provided as Supplementary Data. Because the classical workflow uses different substitution models from the baseline IQ-TREE analyses (LG-based versus Q.pfam-based profile models), differences in reconstructed ancestors under this comparison reflect a combination of software and model effects rather than a pure implementation effect.

Software and reproducibility

Sequence curation used CD-HIT v4.8.1 [39] and Treemmer v0.3 [40]. Multiple sequence alignments were generated with MAFFT (L-INS-i and FFT-NS-1)

v7.525 [29] and PSI/TM-Coffee v11.0 in transmembrane mode [28]. Alignment trimming used trimAl v1.5 [41]. ML phylogenies, model selection, and ASR were performed with IQ-TREE v3.0.1 [31]; binary gap reconstructions used RAxML-HPC v8.2.12 [25]; classical comparisons used RAxML-NG v1.2.2 [42] and PAML v4.10.7 [43]. Structural predictions were obtained using an AlphaFold3 web server [27], and visualization/figure preparation used UCSF ChimeraX v1.10.1 [44] and PyMOL v3.1.0 (Schrödinger, LLC).

Custom Python 3 scripts were written to (i) convert alignments to binary gap matrices, (ii) map RAxML ancestral gap states to IQ-TREE node labels by bipartitions, (iii) generate indel-corrected ancestral sequences, and (iv) compute PP/pLDDT-derived summaries (PP, pLDDT, PP$\times$pLDDT, PP$-$pLDDT, composite scores). These scripts, configuration files, and example command lines are distributed as the ConsistASR pipeline and as an archival snapshot (see Data and Code Availability).

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

The authors used a generative AI assistant (ChatGPT, GPT-5.1 Thinking; OpenAI) to assist with scriptwriting and to edit parts of the manuscript. (e.g. language polishing and restructuring of paragraphs). All scientific ideas, analyses, and conclusions are the authors' own, and all AI-assisted text was carefully checked and, where necessary, revised by the authors.

Acknowledgments

The authors would like to thank Taito Urui and Yu Iritani for helpful discussions and technical assistance. This research was supported by a Grants-in-Aid for Scientific Research to Y.M. (No. JP23H00285) .

Data availability

Supplementary data contains the sequence, alignment, phylogenetic, ancestral sequence, and structural data underlying the analyses in this manuscript, which are available at Zenodo (DOI: doi.org/10.5281/zenodo.18765174). ConsistASR v1.0.1 is available on GitHub (github.com/harutois/ConsistASR) and archived at Zenodo (DOI: doi.org/10.5281/zenodo.18080327).

Tables

Table 1: Summary of datasets, alignments, and best-fit substitution models used for phylogenetic inference and ancestral sequence reconstruction.

Dataset	Alignment strategy	N sequences	Alignment length (residues)	Best-fit model (BIC)	Notes
HeR+SzR+outgroup	PSI/TM-Coffee	228	449	Q.pfam+R7	Baseline dataset used for joint HeR–SzR phylogeny and ASR. Outgroup sequences were constrained as the external clade.
HeR+SzR+outgroup	MAFFT L-INS-i	228	421	Q.pfam+R7	Alternative alignment used to assess alignment sensitivity under the same model family.
HeR+outgroup	PSI/TM-Coffee	118	533	Q.pfam+R6	HeR-only subtree used for node-level analyses of Anc-HeR.
HeR+outgroup	MAFFT L-INS-i	118	401	LG+F+R6	Alignment/model combination that yielded a near-polytomy around the HeR ancestor in UFBoot2 consensus trees.
SzR+outgroup	PSI/TM-Coffee	113	320	Q.pfam+R5	SzR-only subtree; Anc-SzR showed high and stable bootstrap support.
SzR+outgroup	MAFFT L-INS-i	113	287	Q.pfam+R6	Alignment used to confirm robustness of Anc-SzR to alignment strategy.

Best-fit models are those selected by ModelFinder under the Bayesian information criterion (BIC).

Table 2: Composite node-level reliability metrics integrating posterior probabilities (PP), AlphaFold confidence (pLDDT), and ultrafast bootstrap support (UFBoot) for key ancestral nodes.

Alignment	Ancestor	ML UFBoot (%)	Consensus UFBoot2 (%)	Mean PP (%)	Mean pLDDT	Mean (PP×pLDDT)	Composite score (0–100)
PSI/TM-Coffee	Anc-SzR	99	99	81.6	95.5	78.1	77.3
	Anc-HeR	100	54	90.6	95.9	87.2	47.1
	Anc-SH	100	100	54.9	93.2	51.6	51.6
MAFFT L-INS-i	Anc-SzR	96	96	80.4	94.3	76.1	73.0
	Anc-HeR	100	0[a]	91.4	93.6	86.0	0.0[a]
	Anc-SH	100	100	57.5	93.6	54.3	54.3

Mean PP, mean pLDDT, and mean PP×pLDDT are reported on a 0–100 scale. Composite scores were computed as $100 \times b \times \overline{\mathrm{PP} \times \mathrm{pLDDT}}$, where b is the consensus UFBoot2 support expressed as a fraction ($b = \mathrm{UFBoot2}/100$) and $\overline{\mathrm{PP} \times \mathrm{pLDDT}}$ is the mean product on a 0–1 scale.

[a] For MAFFT L-INS-i, the HeR ancestral branch collapses in the UFBoot2 consensus tree (polytomy), so UFBoot2 support is effectively 0 and the composite score is 0 by definition for this alignment.

Table 3: Global impact of indel-aware reconstruction on ancestral node statistics. Mean posterior probabilities (PP), AlphaFold confidence (pLDDT), and sequence length before and after binary gap-based masking.

Alignment	Ancestor	Before indel-aware masking			After indel-aware masking			Extant length[a]
		Mean PP (%)	Mean pLDDT	Length (residues)	Mean PP (%)	Mean pLDDT	Length (residues)	(aa)
PSI/TM-Coffee	Anc-SzR	51.5	62.3	401	81.6	95.5	206	202
	Anc-HeR	65.8	69.2	434	90.6	95.9	259	256
	Anc-SH	37.6	61.3	408	54.9	93.2	206	–
MAFFT L-INS-i	Anc-SzR	52.4	63.6	386	80.4	94.3	203	202
	Anc-HeR	70.4	75.5	414	91.4	93.6	260	256
	Anc-SH	39.8	63.1	396	57.5	93.6	203	–

Mean PP and mean pLDDT are reported on a 0–100 scale and computed over all Cα positions retained (or predicted) in the corresponding full-length ancestral models, before and after indel-aware masking. [a] Extant lengths correspond to SzR4 (202 residues) for Anc-SzR and HeR 48C12 (256 residues) for Anc-HeR. No direct extant structural reference is available for Anc-SH.

Table 4: Region-level reliability of ancestral 7TM rhodopsins. Mean posterior probabilities (PP), AlphaFold confidence (pLDDT), and derived indices for transmembrane (TM) and extra-membrane (EM) segments.

Alignment	Ancestor	Region	Mean PP	Mean pLDDT	Mean (PP×pLDDT)	Mean (PP−pLDDT)
PSI/TM-Coffee	Anc-SzR	TM	83.9	97.4	81.7	43.2
		EM	75.9	90.7	69.3	42.6
	Anc-HeR	TM	92.5	97.6	90.3	47.4
		EM	87.5	93.7	82.5	46.9
MAFFT L-INS-i	Anc-SzR	TM	82.1	96.2	79.2	42.9
		EM	76.2	89.3	68.0	43.5
	Anc-HeR	TM	93.6	95.5	89.5	49.1
		EM	88.0	91.3	81.0	48.3

TM and EM regions were defined by mapping OPM-derived membrane boundaries for SzR4 and HeR 48C12 onto Anc-SzR and Anc-HeR (see Methods). Values are means over Cα positions retained after indel-aware masking.

References

[1] O. P. Ernst, D. T. Lodowski, M. Elstner, P. Hegemann, L. S. Brown, H. Kandori, Microbial and animal rhodopsins: structures, functions, and molecular mechanisms, Chem. Rev. 114 (2014) 126–163. doi:https://doi.org/10.1021/cr4003769.

[2] A. Rozenberg, K. Inoue, H. Kandori, O. Béjà, Microbial rhodopsins: the last two decades, Ann. Rev. Microbiol. 75 (2021) 427–447. doi:https://doi.org/10.1146/annurev-micro-031721-020452.

[3] J. W. Thornton, Resurrecting ancient genes: experimental analysis of extinct molecules, Nat. Rev. Genet. 5 (2004) 366–375. doi:https://doi.org/10.1038/nrg1324.

[4] G. K. Hochberg, J. W. Thornton, Reconstructing ancient proteins to understand the causes of structure and function, Annu. Rev. Biophys. 46 (2017) 247–269. doi:https://doi.org/10.1146/annurev-biophys-070816-033631.

[5] F. F. Nascimento, M. d. Reis, Z. Yang, A biologist's guide to bayesian phylogenetic analysis, Nat. Ecol. Evol. 1 (2017) 1446–1454. doi:https://doi.org/10.1038/s41559-017-0280-x.

[6] H. Ashkenazy, I. Sela, E. Levy Karin, G. Landan, T. Pupko, Multiple sequence alignment averaging improves phylogeny reconstruction, Syst. Biol. 68 (2019) 117–130. doi:https://doi.org/10.1093/sysbio/syy036.

[7] P. D. Williams, D. D. Pollock, B. P. Blackburne, R. A. Goldstein, Assessing the accuracy of ancestral protein reconstruction methods, PLoS Comput. Biol. 2 (2006) e69. doi:https://doi.org/10.1371/journal.pcbi.0020069.

[8] R. Del Amparo, M. Arenas, Consequences of substitution model selection on protein ancestral sequence reconstruction, Mol. Biol. Evol. 39 (2022) msac144. doi:https://doi.org/10.1093/molbev/msac144.

[9] A. Loytynoja, N. Goldman, Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis, Science 320 (2008) 1632–1635. doi:https://doi.org/10.1126/science.1158395.

[10] C. Iglhaut, J. Pečerska, M. Gil, M. Anisimova, Please mind the gap: indel-aware parsimony for fast and accurate ancestral sequence reconstruction and multiple sequence alignment including long indels, Mol. Biol. Evol. 41 (2024) msae109. doi:https://doi.org/10.1093/molbev/msae109.

[11] K. L. Pierce, R. T. Premont, R. J. Lefkowitz, Seven-transmembrane receptors, Nat. Rev. Mol. Cell Biol. 3 (2002) 639–650. doi:https://doi.org/10.1038/nrm908.

[12] S. Lu, X. He, Z. Yang, Z. Chai, S. Zhou, J. Wang, A. U. Rehman, D. Ni, J. Pu, J. Sun, et al., Activation pathway of a G protein-coupled receptor uncovers conformational intermediates as targets for allosteric drug design, Nature Commun. 12 (2021) 4721. doi:https://doi.org/10.1038/s41467-021-25020-9.

[13] B. S. Chang, K. Jönsson, M. A. Kazmi, M. J. Donoghue, T. P. Sakmar, Recreating a functional ancestral archosaur visual pigment, Mol. Biol. Evol. 19 (2002) 1483–1489. doi:https://doi.org/10.1093/oxfordjournals.molbev.a004211.

[14] D. Kuang, Y. Yao, D. MacLean, M. Wang, D. R. Hampson, B. S. Chang, Ancestral reconstruction of the ligand-binding pocket of Family CG protein-coupled receptors, Proc. Natl. Acad. Sci. 103 (2006) 14050–14055. doi:https://doi.org/10.1073/pnas.0604717103.

[15] G. Blanchet, D. Alili, A. Protte, G. Upert, N. Gilles, L. Tepshi, E. A. Stura, G. Mourier, D. Servent, Ancestral protein resurrection and engineering opportunities of the mamba aminergic toxins, Sci. Rep. 7 (2017) 2701. doi:https://doi.org/10.1038/s41598-017-02953-0.

[16] M. Punta, L. R. Forrest, H. Bigelow, A. Kernytsky, J. Liu, B. Rost, Membrane protein prediction methods, Methods 41 (2007) 460–474. doi:https://doi.org/10.1016/j.ymeth.2006.07.026.

[17] A. J. Situ, T. S. Ulmer, Universal principles of membrane protein assembly, composition and evolution, PLoS One 14 (2019) e0221372. doi:https://doi.org/10.1371/journal.pone.0221372.

[18] K. Aadland, B. Kolaczkowski, Alignment-integrated reconstruction of ancestral sequences improves accuracy, Genome Biol. Evol. 12 (2020) 1549–1565. doi:https://doi.org/10.1093/gbe/evaa164.

[19] A. Pushkarev, K. Inoue, S. Larom, J. Flores-Uribe, M. Singh, M. Konno, S. Tomida, S. Ito, R. Nakamura, S. P. Tsunoda, et al., A distinct abundant group of microbial rhodopsins discovered using functional metagenomics, Nature 558 (2018) 595–599. doi:https://doi.org/10.1038/s41586-018-0225-9.

[20] W. Shihoya, K. Inoue, M. Singh, M. Konno, S. Hososhima, K. Yamashita, K. Ikeda, A. Higuchi, T. Izume, S. Okazaki, et al., Crystal structure of heliorhodopsin, Nature 574 (2019) 132–136. doi:https://doi.org/10.1038/s41586-019-1604-6.

[21] K. Inoue, S. P. Tsunoda, M. Singh, S. Tomida, S. Hososhima, M. Konno, R. Nakamura, H. Watanabe, P.-A. Bulzu, H. L. Banciu, et al., Schizorhodopsins: A family of rhodopsins from Asgard archaea that function as light-driven inward H^+ pumps, Sci. Adv. 6 (2020) eaaz2441. doi:https://doi.org/10.1126/sciadv.aaz2441.

[22] K. Kovalev, D. Volkov, R. Astashkin, A. Alekseev, I. Gushchin, J. M. Haro-Moreno, I. Chizhov, S. Siletsky, M. Mamedov, A. Rogachev, et al., High-resolution structural insights into the heliorhodopsin family, Proc. Natl. Acad. Sci. 117 (2020) 4131–4141. doi:https://doi.org/10.1073/pnas.1915888117.

[23] A. Higuchi, W. Shihoya, M. Konno, T. Ikuta, H. Kandori, K. Inoue, O. Nureki, Crystal structure of schizorhodopsin reveals mechanism of inward proton pumping, Proc. Natl. Acad. Sci. 118 (2021) e2016328118. doi:https://doi.org/10.1073/pnas.2016328118.

[24] J. E. Besaw, S. Peng, A. Kuo, J. Reichenwallner, R. D. Miller, L. S. Brown, O. P. Ernst, Structural insights into an inward proton pumping rhodopsin, Biophys. J. (2025). doi:https://doi.org/10.1016/j.bpj.2025.11.011.

[25] A. Stamatakis, RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models, Bioinformatics 22 (2006) 2688–2690. doi:https://doi.org/10.1093/bioinformatics/btl446.

[26] K. Aadland, C. Pugh, B. Kolaczkowski, High-throughput reconstruction of ancestral protein sequence, structure, and molecular function, in: Computational Methods in Protein Evolution, Springer, 2018, pp. 135–170. doi:https://doi.org/10.1007/978-1-4939-8736-8_8.

[27] J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, et al., Accurate structure prediction of biomolecular interactions with AlphaFold 3, Nature 630 (2024) 493–500. doi:https://doi.org/10.1038/s41586-024-07487-w.

[28] P. Di Tommaso, S. Moretti, I. Xenarios, M. Orobitg, A. Montanyola, J.-M. Chang, J.-F. Taly, C. Notredame, T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension, Nucleic Acids Res. 39 (2011) W13–W17. doi:https://doi.org/10.1093/nar/gkr245.

[29] K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability, Mol. Biol. Evol. 30 (2013) 772–780. doi:https://doi.org/10.1093/molbev/mst010.

[30] S. Kalyaanamoorthy, B. Q. Minh, T. K. Wong, A. Von Haeseler, L. S. Jer-

miin, ModelFinder: fast model selection for accurate phylogenetic estimates, Nat. Methods 14 (2017) 587–589. doi:https://doi.org/10.1038/nmeth.4285.

[31] T. K. Wong, N. Ly-Trong, H. Ren, H. Baños, A. J. Roger, E. Susko, C. Bielow, N. De Maio, N. Goldman, M. W. Hahn, et al., IQ-TREE 3: Phylogenomic inference software using complex evolutionary models, Submitted (2025). doi:https://doi.org/10.32942/X2P62N.

[32] D. T. Hoang, O. Chernomor, A. Von Haeseler, B. Q. Minh, L. S. Vinh, UFBoot2: improving the ultrafast bootstrap approximation, Mol. Biol. Evol. 35 (2018) 518–522. doi:https://doi.org/10.1093/molbev/msx281.

[33] M. A. Lomize, I. D. Pogozheva, H. Joo, H. I. Mosberg, A. L. Lomize, OPM database and PPM web server: resources for positioning of proteins in membranes, Nucleic Acids Res. 40 (2012) D370–D376. doi:https://doi.org/10.1093/nar/gkr703.

[34] M. Shibata, K. Inoue, K. Ikeda, M. Konno, M. Singh, C. Kataoka, R. Abe-Yoshizumi, H. Kandori, T. Uchihashi, Oligomeric states of microbial rhodopsins determined by high-speed atomic force microscopy and circular dichroic spectroscopy, Sci. Rep. 8 (2018) 8262. doi:https://doi.org/10.1038/s41598-018-26606-y.

[35] L. J. Galindo, S. Takaramoto, T. Nagata, A. Rozenberg, H. Takahashi, O. Béjà, K. Inoue, Apusomonad rhodopsins: A new family of ultraviolet to blue light–absorbing rhodopsin channels, Proc. Natl. Acad. Sci. 122 (2025) e2510619122. doi:https://doi.org/10.1073/pnas.2510619122.

[36] C. D. Sephus, E. Fer, A. K. Garcia, Z. R. Adam, E. W. Schwieterman, B. Kacar, Earliest photic zone niches probed by ancestral microbial rhodopsins, Mol. Biol. Evol. 39 (2022) msac100. doi:https://doi.org/10.1093/molbev/msac100.

[37] P.-A. Bulzu, V. S. Kavagutti, A.-S. Andrei, R. Ghai, The evolu-

tionary kaleidoscope of rhodopsins, mSystems 7 (2022) e00405–22. doi:https://doi.org/10.1128/msystems.00405-22.

[38] L. Richardson, B. Allen, G. Baldi, M. Beracochea, M. L. Bileschi, T. Burdett, J. Burgin, J. Caballero-Pérez, G. Cochrane, L. J. Colwell, et al., MGnify: the microbiome sequence data analysis resource in 2023, Nucleic Acids Res. 51 (2023) D753–D759. doi:https://doi.org/10.1093/nar/gkac1080.

[39] W. Li, A. Godzik, Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences, Bioinformatics 22 (2006) 1658–1659. doi:https://doi.org/10.1093/bioinformatics/btl158.

[40] F. Menardo, C. Loiseau, D. Brites, M. Coscolla, S. M. Gygli, L. K. Rutaihwa, A. Trauner, C. Beisel, S. Borrell, S. Gagneux, Treemmer: a tool to reduce large phylogenetic datasets with minimal loss of diversity, BMC Bioinform. 19 (2018) 164. doi:https://doi.org/10.1186/s12859-018-2164-8.

[41] S. Capella-Gutiérrez, J. M. Silla-Martínez, T. Gabaldón, trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses, Bioinformatics 25 (2009) 1972–1973. doi:https://doi.org/10.1093/bioinformatics/btp348.

[42] A. M. Kozlov, D. Darriba, T. Flouri, B. Morel, A. Stamatakis, RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference, Bioinformatics 35 (2019) 4453–4455. doi:https://doi.org/10.1093/bioinformatics/btz305.

[43] Z. Yang, PAML 4: phylogenetic analysis by maximum likelihood, Mol. Biol. Evol. 24 (2007) 1586–1591. doi:https://doi.org/10.1093/molbev/msm088.

[44] E. C. Meng, T. D. Goddard, E. F. Pettersen, G. S. Couch, Z. J. Pearson, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Tools for structure building and analysis, Protein Sci. 32 (2023) e4792. doi:https://doi.org/10.1002/pro.4792.