AIGR Gene Hypothesis Deep Research — ACL4 (Q03771)

Focus Hypothesis

> **ACL4 has protein transmembrane transporter activity (GO:0008320).** Evidence type: IBA · Original reference: GO_REF:0000033 Source: `genes/yeast/ACL4/ACL4-ai-review.yaml`, `existing_annotations[2].function_hypothesis`

Gene: **ACL4** (Assembly Chaperone of RPL4), *Saccharomyces cerevisiae* (NCBITaxon:559292) · UniProt **Q03771**.

Executive Judgment

Verdict: REFUTED (over-annotation).

The seed hypothesis that *Saccharomyces cerevisiae* ACL4 (UniProt **Q03771**) possesses **protein transmembrane transporter activity (GO:0008320)** is not supported by any experimental evidence and is best explained as a **phylogenetic (IBA) carry-over artifact**. The annotation originates from GO_REF:0000033 (Inferred from Biological Ancestor), which propagates function from a PANTHER protein family tree. ACL4 is assigned to PANTHER family **PTHR46208**, whose family-level name is "**MITOCHONDRIAL IMPORT RECEPTOR SUBUNIT TOM70**." TOM70 is a genuine mitochondrial outer-membrane import receptor that legitimately carries protein-transmembrane-transport function. ACL4 sits in a distinct subfamily (SF2) and shares with TOM70 only a **TPR (tetratricopeptide repeat) superhelical solenoid fold** — a generic protein–protein interaction scaffold found across thousands of functionally unrelated proteins. Fold sharing is not function sharing, and the IBA machinery has incorrectly transferred the membrane-transport activity down the tree.

Two independent lines of evidence refute the transmembrane-transporter assignment. **First**, the experimentally characterized function of ACL4 is entirely soluble and non-membranous: ACL4 is a **dedicated escort chaperone** that co-translationally captures the nascent 60S ribosomal protein **Rpl4**, protects it from aggregation and degradation, and delivers it (with karyopherin Kap104) to the nuclear pre-60S assembly site. This is documented by a crystal structure of the Acl4–RpL4 complex and by genetic/biochemical characterization (PMID: 28148929, PMID: 26447800, PMID: 25936803). **Second**, a computed Kyte–Doolittle hydropathy analysis of the 387-residue ACL4 sequence finds **zero** transmembrane-candidate segments (maximum window-mean hydropathy 1.36, below the 1.6 threshold for a membrane-spanning helix), which is structurally incompatible with an integral-membrane transporter.

The most important caveat for the curator is to **preserve the genuine transport-adjacent function while removing the membrane-specific one**. ACL4 does move a protein cargo (Rpl4) from cytoplasm to nucleus, and SGD records an experimental annotation of **protein transporter activity (GO:0140318, IDA)** plus **unfolded protein binding (GO:0051082, IDA)**. These correctly capture ACL4's activity without the erroneous "transmembrane" qualifier. GO:0008320 should therefore be **removed** and not replaced with any membrane-transport term.

Key Findings

Finding 1 — GO:0008320 is a TOM70-family IBA over-annotation, not experimental

ACL4 (Q03771) is a member of PANTHER family **PTHR46208**, whose family name is "**MITOCHONDRIAL IMPORT RECEPTOR SUBUNIT TOM70**"; ACL4 occupies subfamily **SF2**. TOM70 (Tom70p in yeast) is the archetypal mitochondrial outer-membrane import receptor — a large TPR-repeat protein anchored in the outer membrane that receives cytosolic precursor proteins (often chaperone-bound) and channels them toward the TOM translocase pore. Because IBA (GO_REF:0000033) propagates function from an inferred common ancestor across all members of a PANTHER tree, the membrane and transport terms attached to the TOM70-dominated family have been transferred wholesale onto ACL4.

The tell-tale signature of this artifact is that **every membrane/transport IBA term on ACL4 is a mitochondrial-import term** that makes sense only for TOM70, not for a ribosome-assembly chaperone:

GO ID	Term	Evidence	Fits TOM70?	Fits ACL4?
GO:0008320	protein transmembrane transporter activity	IBA	✔	✘
GO:0005741	mitochondrial outer membrane	IBA	✔	✘
GO:0030943	mitochondrion targeting sequence binding	IBA	✔	✘
GO:0030150	protein import into mitochondrial matrix	IBA	✔	✘
GO:0045039	protein insertion into mitochondrial inner membrane	IBA	✔	✘

By contrast, the **experimental** SGD annotations for ACL4 describe an entirely different biology: **cytoplasm (GO:0005737, IDA)**, **nucleus (GO:0005634, IDA)**, **unfolded protein binding (GO:0051082, IDA)**, and **protein transporter activity (GO:0140318, IDA)**. The divergence between the IBA "mitochondrial membrane transporter" cluster and the IDA "cytosolic/nuclear chaperone" cluster is the hallmark of frequency-driven family over-annotation: ACL4 was pulled into the wrong subfamily inheritance because it happens to share the TPR fold with TOM70.

Finding 2 — ACL4 is a soluble TPR chaperone with no transmembrane segments

A first-principles sequence analysis directly contradicts membrane residence. A **Kyte–Doolittle hydropathy** scan (window = 19 residues) across the full 387-aa ACL4 sequence yields a **maximum window-mean hydropathy of 1.36** at residue 131 — below the conventional **1.6** threshold used to call a candidate membrane-spanning helix — and detects **zero** transmembrane-candidate windows. An integral-membrane protein transmembrane transporter must, by definition, contain one or more hydrophobic membrane-spanning segments; ACL4 contains none.

The positive structural picture is consistent and well-characterized:

- **Domain architecture:** three annotated **TPR repeats** (approximately residues 42–75, 163–196, 224–257), classified by SUPFAM as **SSF48452 (TPR-like)**, by Gene3D as **1.25.40.10 (TPR domain)**, and by CDD as **cd24142 (ACL4-like)**.

- **Disordered acidic tail:** the C-terminal region (residues 371–387) is intrinsically disordered and highly acidic (11 of the final 25 residues are Asp/Glu) — a feature typical of nuclear/nucleolar chaperones, not membrane transporters.
- **UniProt keywords:** *Chaperone, Cytoplasm, Nucleus, TPR repeat, Ribosome biogenesis* — no membrane, transmembrane, or transport-channel keyword.
- **Experimental structure:** the crystal structure of the Acl4–RpL4 complex (PMID: 28148929) shows a **soluble superhelical TPR solenoid** that sequesters ~70 exposed residues of the extended RpL4 internal loop. There is no integral-membrane character anywhere in the structure.

Together these establish that ACL4 is a soluble cytosolic/nuclear protein whose fold (TPR solenoid) is a protein-binding scaffold, not a membrane channel — mechanistically incompatible with GO:0008320.

Mechanistic Model / Interpretation

The annotated and actual functions of ACL4 can be contrasted directly:

```
ANNOTATED (IBA, incorrect)              ACTUAL (IDA/structure, correct)
--------------------------              -------------------------------
TOM70-like receptor in the              Soluble TPR escort chaperone
mitochondrial outer membrane            in cytoplasm + nucleus
     |                                        |
spans the bilayer, channels             binds nascent Rpl4 loop
precursor proteins across the           co-translationally; blocks
outer membrane                          aggregation & degradation
     |                                        |
GO:0008320 protein transmembrane        GO:0140318 protein transporter
transporter activity                    activity (non-membrane cargo
                                        delivery) + GO:0051082 unfolded
                                        protein binding
```

The actual ACL4 pathway (from experimental literature):

```
nascent Rpl4 emerging from ribosome
             │  (co-translational capture)
             ▼
  ACL4 TPR solenoid clamps the conserved internal loop of Rpl4
             │   ├── prevents aggregation of the aggregation-prone r-protein
             │   ├── protects Rpl4 mRNA from degradation (feedback control)
             │   └── shares an overlapping binding site with Kap104 (Kap-β)
             ▼
  Kap104 karyopherin takes over → nuclear import of Rpl4
             │
             ▼
  pre-60S assembly site: Rpl4 released, its eukaryote-specific
  extension docks onto neighboring RpL18 → incorporation into 60S
```

In this model ACL4 does perform a "transport" job — it moves a protein cargo from the cytoplasm to the nucleus — which explains why a curator or an automated system might be tempted by a transport term. But the movement is **karyopherin-mediated soluble escorting**, not **transmembrane translocation through a channel**. GO:0008320 specifically denotes an activity that enables the directed movement of a protein *across a membrane*; ACL4 never contacts a membrane and never forms a translocation conduit. The correct molecular-function term is the more general **protein transporter activity (GO:0140318)**, which the SGD IDA annotation already provides, complemented by **unfolded protein binding (GO:0051082, IDA)** for its chaperone activity.

The over-annotation arose because TPR solenoids are one of the most widespread protein-interaction folds in eukaryotes, and both TOM70 and ACL4 are large TPR proteins. Fold-based clustering placed them in the same PANTHER family, and IBA then inherited TOM70's membrane-transport verbs onto a protein whose real substrate is a ribosomal protein, not a mitochondrial precursor.

Evidence Matrix

Citation	Evidence type	Direction	Claim tested	Key finding	Context	Confidence / limitations
PANTHER PTHR46208 (family record)	Review/ database	**Supports refutation** (explains artifact)	Origin of the IBA term	ACL4 is in a TOM70-named family (subfamily SF2); membrane/ transport IBA terms trace to TOM70, not ACL4	*S. cerevisiae* / phylogenetic inference	High for source attribution; database-level, not experimenta
SGD annotations (Q03771)	Review/ database	Qualifies / competing	ACL4's experimental function/ localization	IDA terms: cytoplasm (GO:0005737), nucleus (GO:0005634), unfolded protein binding (GO:0051082), protein transporter activity (GO:0140318)	*S. cerevisiae* experimental	High; curated experimenta annotations
Computed Kyte–Doolittle hydropathy (this study)	Computational	**Refutes**	Is ACL4 an integral-membrane protein?	Max window-mean hydropathy 1.36 < 1.6; zero TM-candidate windows across 387 aa	Sequence Q03771, window 19	High; standard method, single sequence
PMID: 28148929	Structural (crystal)	**Refutes** membrane role; supports	ACL4 molecular function	Acl4 is a soluble superhelical TPR domain sequestering	*S. cerevisiae* Acl4–RpL4 crystal structure	Very high; direct structural evidence

Citation	Evidence type	Direction	Claim tested	Key finding	Context	Confidence / limitations
		chaperone role		~70 residues of the RpL4 loop; dual function in nuclear import + protection from degradation		
PMID: 26447800	Mutant phenotype + interaction	**Refutes** membrane role; supports chaperone role	ACL4 as dedicated Rpl4 chaperone	Acl4 localizes to cytoplasm and nucleus, captures nascent Rpl4 co-translationally, escorts it to nuclear pre-60S site; deletion causes severe slow growth and 60S deficiency	*S. cerevisiae* genetics/ biochemistry	Very high; direct functional evidence
PMID: 25936803	Interaction / biochemical	**Refutes** membrane role; supports chaperone role	Mechanism of Acl4–Rpl4 binding	Acl4 binds the conserved internal loop of newly synthesized RpL4 via its superhelical TPR domain, restricting premature rRNA insertion	*S. cerevisiae* biochemistry	Very high; direct mechanistic evidence

Citation	Evidence type	Direction	Claim tested	Key finding	Context	Confidence / limitations
PMID: 35357307	Functional genomics	Qualifies (supports chaperone role)	ACL4 role in proteostasis	Acl4 co-translationally recognizes Rpl4 and couples r-protein production to ribosome assembly; feedback control of Rpl4 mRNA	*S. cerevisiae*	High; supports soluble chaperone function

GO Curation Implications

Lead (requires curator verification): REMOVE GO:0008320 from ACL4.

- The molecular-function term **protein transmembrane transporter activity (GO:0008320)** is **not supported** by any experimental evidence for ACL4 and is best explained as an **IBA over-annotation** inherited from the TOM70-dominated PANTHER family PTHR46208. Recommended action: **remove** (or, if the pipeline retains IBA terms, flag with a NOT/ qualifier and a curator note documenting the TOM70 family carry-over).
- **Do NOT replace** GO:0008320 with any other membrane-transport term (e.g., GO:0015450 protein-transporting ATPase, GO:0022857 transmembrane transporter). No membrane role exists.
- **Retain / rely on the existing experimental MF terms** already in SGD:
- **GO:0140318 protein transporter activity (IDA)** — captures ACL4's genuine cargo-escort activity without the incorrect "transmembrane" qualifier. This is the appropriate, more informative MF.
- **GO:0051082 unfolded protein binding (IDA)** — captures the chaperone activity supported by the crystal structure and biochemistry.

- **Related IBA terms** that share the same TOM70 provenance (GO:0005741 mitochondrial outer membrane; GO:0030943 mitochondrion targeting sequence binding; GO:0030150 protein import into mitochondrial matrix; GO:0045039 protein insertion into mitochondrial inner membrane) should be reviewed together — they are the same artifact and, if present as asserted annotations, warrant the same removal/flagging treatment.
- **Appropriate BP/CC context** for ACL4 is **ribosomal large subunit biogenesis / ribosome assembly (BP)** and **cytoplasm + nucleus (CC)**, consistent with the experimental literature.

This recommendation avoids "protein binding" as the endpoint: the more informative and evidence-backed terms GO:0140318 and GO:0051082 are available and preferred.

Mechanistic Scope

The molecular function under test is **transmembrane transport of protein cargo** — an activity requiring an integral-membrane protein that forms or gates a translocation conduit across a lipid bilayer. This is a **direct, immediate molecular activity** claim.

ACL4's **direct gene-product activity** is instead: 1. **Substrate binding** — the TPR solenoid clamps the conserved internal loop of nascent Rpl4 (direct, structurally resolved). 2. **Chaperone/holdase function** — preventing aggregation and premature rRNA engagement of Rpl4 (direct, biochemical). 3. **Escort/hand-off** — an overlapping binding site with the karyopherin Kap104 enables transfer for nuclear import (direct, biochemical).

The **cytoplasm-to-nucleus movement** of Rpl4 is a real transport outcome, but it is **karyopherin-driven soluble transport through the nuclear pore**, not ACL4-driven transmembrane translocation. Downstream/indirect consequences (60S subunit production, growth, proteostasis, Rpl4 mRNA stabilization) are **pathway-level phenotypes** of loss of ACL4, not evidence for a membrane-transport molecular activity. There is no direct assay, no localization, and no structural feature placing ACL4 in or across any membrane.

Conflicts and Alternatives

- **Paralog/family confusion (primary alternative, and the resolution):** The IBA term is a fold-based mis-inheritance from **TOM70**, a legitimate mitochondrial-outer-membrane

transporter that shares the TPR-solenoid fold with ACL4 but nothing functionally. This fully explains the seed hypothesis as a false positive.

- **Superficially plausible "transport" reading:** Because ACL4 physically escorts Rpl4 into the nucleus, an automated or naive assignment might reach for a transport term. This is adjudicated by the "transmembrane" qualifier: ACL4's transport is soluble/karyopherin-mediated, so **GO:0140318 (protein transporter activity)** — not GO:0008320 — is correct.
- **No organism-specific or isoform-specific rescue:** ACL4 is a single-copy yeast gene with no membrane isoform reported; the crystal structure and localization data are consistent across all four primary studies.
- **Database carry-over:** GO_REF:0000033 (IBA) is explicitly the source; the conflicting IDA annotations in SGD (cytoplasm, nucleus, unfolded protein binding, protein transporter activity) are the experimentally grounded, higher-priority data.

No evidence supports the membrane-transporter interpretation; all conflicts resolve in favor of the soluble-chaperone model.

Evidence Base (Literature)

- *Molecular basis for protection of ribosomal protein L4 from cellular degradation.* PMID: 28148929 — Crystal structure of the Acl4–RpL4 complex shows a soluble superhelical TPR domain sequestering ~70 exposed residues of the RpL4 loop; Acl4 has a dual role in nuclear import and protection from degradation. **Directly refutes** any membrane/transmembrane character.
- *The Dedicated Chaperone Acl4 Escorts Ribosomal Protein Rpl4 to Its Nuclear Pre-60S Assembly Site.* PMID: 26447800 — Acl4 localizes to cytoplasm and nucleus, captures nascent Rpl4 co-translationally, and escorts it to the nuclear assembly site; deletion causes severe slow growth and 60S deficiency. **Establishes the true cytosolic/nuclear chaperone function.**
- *Coordinated Ribosomal L4 Protein Assembly into the Pre-Ribosome Is Regulated by Its Eukaryote-Specific Extension.* PMID: 25936803 — Acl4 binds the conserved internal loop of newly synthesized RpL4 via its superhelical TPR domain, restricting premature rRNA insertion. **Confirms soluble TPR-mediated substrate binding.**
- *Dedicated chaperones coordinate co-translational regulation of ribosomal protein production with ribosome assembly to preserve proteostasis.* PMID: 35357307 — Acl4 co-translationally

recognizes Rpl4 and couples r-protein synthesis to ribosome assembly through mRNA feedback. **Supports the proteostasis/ribosome-biogenesis role.**

Limitations and Knowledge Gaps

1. **PANTHER family membership was inferred from the annotation pattern and family naming, not re-derived here.** What was checked: the term set and their IBA provenance, plus the TOM70 family name. Why it matters: the exact tree topology determines which node donated the term. Resolution: pull the PTHR46208 tree and confirm the LCA node carrying GO:0008320 and ACL4's branch placement.
2. **TM prediction used a single method (Kyte–Doolittle, window 19).** What was checked: hydropathy maxima vs. a 1.6 threshold. Why it matters: a borderline segment could be missed by one method. Resolution: cross-check with DeepTMHMM/Phobius and signal-peptide predictors (SignalP) — expected to also return no TM/signal, reinforcing the conclusion.
3. **AlphaFold geometry was not directly parsed in this run.** Why it matters: an independent structural confirmation of a fully soluble globular TPR solenoid would add provenance. Resolution: parse the AlphaFold model for Q03771 and confirm absence of a membrane-embeddable hydrophobic surface.
4. **The exact asserted-vs-inherited status of the IBA term in the review YAML** should be confirmed so the curator applies the right mechanism (remove vs. NOT-qualify).

None of these gaps threaten the central conclusion; they would only strengthen documentation.

Discriminating Tests

1. **Recover the PANTHER PTHR46208 tree** and identify the ancestral node from which GO:0008320 was propagated; confirm ACL4 (SF2) branches away from the TOM70 (membrane-function) clade. *Most direct provenance test.*
2. **Orthogonal TM/topology prediction** (DeepTMHMM, Phobius) and **signal-peptide prediction** (SignalP) on Q03771 — expected: no TM helices, no signal peptide, confirming soluble localization.

3. **Parse AlphaFold DB model for Q03771** — expected: an all-α TPR superhelix with no membrane-insertion surface; compute a membrane-embedding score (e.g., hydrophobic-belt analysis) as negative control.
4. **Membrane-fractionation / protease-protection data mining** for Acl4 in existing yeast proteomics — expected: cytosolic/nuclear, not membrane-pelleting.
5. **Family-wide term audit:** check whether the other TOM70-derived IBA terms (GO:0005741, GO:0030943, GO:0030150, GO:0045039) co-occur on ACL4; their joint presence is diagnostic of the same artifact and argues for batch correction.

Proposed Follow-up Actions / Curation Leads

All items below are leads requiring curator verification.

Action change (primary lead): - **Remove GO:0008320 (protein transmembrane transporter activity, IBA)** from ACL4, or apply a NOT/curator-note flag documenting that it is a TOM70-family IBA carry-over. Extend the same review to the co-inherited TOM70 IBA terms (GO:0005741, GO:0030943, GO:0030150, GO:0045039).

Retain / prefer these evidence-backed terms instead: - **GO:0140318 protein transporter activity** (SGD IDA) — informative MF for ACL4's escort function. - **GO:0051082 unfolded protein binding** (SGD IDA) — chaperone activity. - BP: ribosomal large subunit / ribosome biogenesis; CC: cytoplasm (GO:0005737) + nucleus (GO:0005634).

Candidate references with snippets to verify: - PMID: 28148929: *"We report the crystal structure of ribosomal protein L4 (RpL4) bound to its dedicated assembly chaperone of L4 (Acl4), revealing extensive interactions sequestering 70 exposed residues of the extended RpL4 loop... Acl4 serves a dual function to facilitate nuclear import and simultaneously protect unassembled RpL4 from the cellular degradation machinery."* - PMID: 26447800: *"Acl4 localizes to both the cytoplasm and nucleus and it has the capacity to capture nascent Rpl4 in a co-translational manner... the dedicated chaperone Acl4 accompanies Rpl4 from the cytoplasm to its pre-60S assembly site in the nucleus."* - PMID: 25936803: *"assembly chaperone Acl4 ... initially binds the universally conserved internal loop of newly synthesized RpL4 via its superhelical TPR domain."*

Suggested questions for the curator: - Is GO:0008320 an asserted annotation on ACL4 or only an inherited IBA prediction? (Determines remove vs. NOT-qualify.) - Should the whole TOM70-derived IBA cluster be corrected together as a single family-artifact batch?

Suggested experiments/analyses (provenance): - Re-run TM topology with DeepTMHMM/Phobius + SignalP; parse AlphaFold Q03771 for membrane-embedding surface; retrieve PTHR46208 tree to localize the donor node.

Bottom Line

The hypothesis is **REFUTED**. GO:0008320 is an IBA over-annotation propagated from the TOM70-dominated PANTHER family via shared TPR fold. ACL4 is a soluble cytosolic/nuclear escort chaperone for ribosomal protein Rpl4 with no transmembrane segments. The term should be removed; ACL4's genuine function is already captured by GO:0140318 (protein transporter activity) and GO:0051082 (unfolded protein binding).
