Deep Research Report: Does *C. elegans* ubl-5 (P91302) have protein tag activity (GO:0031386)?

Focus: function-assignment — does the gene product directly have GO:0031386 "protein tag activity"? **Source annotation:** `existing_annotations[4].function_hypothesis`, evidence IBA, GO_REF:0000033.

Executive Judgment

Verdict: REFUTED (over-annotated).

The seed hypothesis — that ubl-5 has **protein tag activity (GO:0031386)** — is **not supported** and should be treated as an over-annotation. Three independent, mutually reinforcing lines of evidence converge:

1. **Definitional mismatch.** GO:0031386 is defined as an activity of a protein "that is **covalently attached (AKA tagged or conjugated) to another protein**" (QuickGO), and its curator comment restricts the term to *conjugated* tags. Covalent conjugation is a hard requirement of the term.
2. **Structural signature (computed here).** UBL-5 (P91302, 73 aa) ends in **…FNFEL-Y-Y-Q**, terminating in Gln73 with **no C-terminal di-glycine (GG)**. The di-glycine is mechanistically obligatory for isopeptide-bond conjugation and is present in every canonical covalent tag (ubiquitin, SUMO, NEDD8, ISG15). Its absence in UBL-5 (and in the identical-length human ortholog UBL5) is a diagnostic marker that this subfamily is **not conjugation-competent**.
3. **Direct experimental evidence.** Primary literature across yeast, human, and *C. elegans* states that Hub1/UBL-5 **does not form covalent conjugates** and instead **binds partners non-covalently** through the HIND motif of spliceosomal proteins (Snu66/SART-1, Prp38/PRP-38).

The IBA annotation is a phylogenetic carry-over from the broad ubiquitin superfamily that fails to account for the well-characterized, conjugation-independent divergence of the Hub1/UBL5 clade. The genuinely supported function is **non-covalent binding to HIND-motif spliceosomal factors**, contributing to **mRNA splicing (GO:0000398)** — already annotated (IBA) — not protein tag activity.

Most important caveat: UniProt's human UBL5 entry still carries the legacy keyword "Ubl conjugation pathway," and older yeast reports (early 2000s) initially suggested Hub1 conjugation; these were superseded. This legacy is the likely origin of the over-annotation and should not be weighted against the direct non-covalent evidence.

Evidence Matrix

Citation	Evidence type	Stance	Claim tested	Key finding	Context	Confidence / limitations
This report (UniProt P91302 FASTA)	Structural/ evolutionary (computed)	**Refutes**	Does UBL-5 have a conjugation-competent C-terminus?	73 aa, C-terminus … ELYYQ; **no C-terminal GG**; contrasts ubiquitin/ SUMO/NEDD8/ ISG15 which all end in GG	Sequence-level, worm + human ortholog	High; sequence-based inference of mechanism
Ammon et al. 2014, PMID **24872507**	Direct assay / cell biology	**Refutes**	Is Hub1/ UBL5 a covalent tag?	"Hub1 does not form covalent conjugates with substrates but binds proteins non-covalently"; modulator of alternative splicing	Human + yeast	High; explicit statement of binding mode
Mishra et al. 2011, PMID **21614000**	Structural + biochemical	**Refutes / re-assigns**	Binding mode & function	Hub1 "binds non-covalently to a conserved element termed HIND" (Snu66/Prp38); "non-covalent modification of the spliceosome by an unconventional ubiquitin-like modifier"	*S. cerevisiae*, structural	High; foundational mechanistic paper
Kolathur et al. 2023,		**Refutes (organism-**	Worm UBL-5	"Hub1/UBL-5 associates with	***C. elegans***	High; the directly

Citation	Evidence type	Stance	Claim tested	Key finding	Context	Confidence / limitations
PMID **36480405**	Mutant phenotype + interaction	**specific) / re-assigns**	binding mode & role	proteins **non-covalently**"; binds HIND factors Snu66/ SART-1 and PRP-38; *ubl-5* mutants die at L3 with splicing defects		relevant organism
Oka et al. 2014, PMID **25092792**	Depletion / functional assay	Qualifies (re-assigns)	Metazoan UBL5 function	UBL5 "primarily associates with spliceosomal proteins"; depletion → intron retention, loss of sister-chromatid cohesion via Sororin	Human cells	High; function is splicing, not tagging
Karaduman et al. 2017, PMID **28712727**	Biochemical	Qualifies (re-assigns)	Molecular mechanism	Hub1 binds DEAD-box helicase Prp5 and stimulates its ATPase → error-prone splicing	*S. cerevisiae*	High; non-covalent partner regulation
Haynes et al. 2007, PMID **17925224** (UniProt FUNCTION)	Mutant phenotype	Competing (downstream BP)	Worm-specific role	ubl-5 required for mitochondrial UPR (mtUPR); interacts with dve-1 under stress	*C. elegans*	Downstream biological process, not an MF

Citation	Evidence type	Stance	Claim tested	Key finding	Context	Confidence / limitations
QuickGO GO:0031386 (definition)	Database / ontology	Sets the bar	What the term requires	"covalently attached (… conjugated) to another protein"; comment: "annotate conjugated tags"	Ontology	Definitional; decisive for the call

GO Curation Implications

Lead (requires curator verification): REMOVE / do-not-annotate GO:0031386 (protein tag activity) for ubl-5.

- **MF:** GO:0031386 is **not supported**; the covalent-conjugation requirement of the term is contradicted by direct evidence and by the missing C-terminal di-glycine. Recommend removing the IBA annotation (or entering it as a NOT annotation for the family propagation). If a molecular-function term is desired, the defensible activity is **non-covalent protein binding to spliceosomal HIND-motif partners** — a more informative candidate than bare "protein binding" would be a splicing-factor/spliceosome binding term (e.g., a "spliceosomal complex binding"–type MF) supported by

  P 36480405 /

  21614000, though direct MF terms for UBL5 remain sparse.
- **BP (retain):** GO:0000398 "mRNA splicing, via spliceosome" (IBA) is well supported and now has *C. elegans*-specific backing

  (P 36480405).

  GO:0034514 "mitochondrial unfolded protein response" (IMP) is a valid worm-specific downstream process but is not a molecular tag activity.
- **CC (retain):** cytoplasm/cytosol/nucleus and transcription-regulator-complex localizations are experimentally supported (IDA/IPI, WormBase).

Net: the protein-tag MF should be **removed/generalized away**; the splicing BP and localization CC annotations carry the real function.

Mechanistic Scope

- **Immediate molecular function tested:** whether UBL-5 acts as a covalent post-translational tag (like ubiquitin/SUMO) that is conjugated to substrate lysines to mark them for a cellular fate.
- **What the evidence actually shows:** UBL-5 is a ubiquitin-*fold* protein that engages partners **non-covalently** via the HIND motif (Snu66/SART-1, PRP-38) and modulates spliceosome performance and alternative/weak-intron splicing. It is a **binding modulator**, not a conjugatable marker.
- **Downstream / not the direct MF:** mtUPR regulation (with dve-1/atfs-1), lipid-metabolism modulation upstream of nhr-80, L3 lethality, sister-chromatid cohesion (via Sororin), and stress-response phenotypes are **downstream consequences** of impaired non-covalent splicing/complex function — not evidence of tag activity.

Conflicts and Alternatives

- **Database carry-over / legacy artifact:** Human UBL5 (Q9BZL1) retains the UniProt keyword "Ubl conjugation pathway," and early-2000s yeast work initially proposed Hub1 conjugation. These are the most likely source of the IBA propagation and are contradicted by later definitive studies
  (P 24872507).
  P 21614000
- **Family/paralog over-annotation:** GO:0031386 is a superfamily-level attribute inherited by IBA across ubiquitin-like proteins; the Hub1/UBL5 clade (InterPro IPR039732 Hub1/Ubl5) is the recognized exception. The seed hypothesis reflects clade-blind propagation.
- **No conflicting primary evidence found** that UBL-5 is ever covalently conjugated to a substrate.

Knowledge Gaps

1. **Formal MF term for a non-covalent Ubl modulator.** Checked GO/QuickGO: GO:0031386 requires covalency, so it cannot apply. A precise MF term capturing "non-covalent spliceosome-modulating binding" is not clearly established in GO. *Resolves by:* curator selection of an appropriate binding MF or a new/related term; supported by

   P 36480405

   P 21614000
2. **Direct in vivo conjugation assay in worm.** No mass-spec/isopeptide evidence exists showing UBL-5–substrate conjugates in *C. elegans*. *Matters because* it would be the only way to overturn the non-covalent consensus. *Resolves by:* anti-UBL-5 IP + LC-MS/MS for GG-remnant/branched peptides (expected negative).
3. **C-terminal maturation.** UBL-5 ends in Gln, not a processed Gly; there is no evidence of a protease exposing a conjugatable Gly. *Resolves by:* checking for any C-terminal hydrolase activity (none reported).

Discriminating Tests

- **Isopeptide/GG-remnant proteomics:** IP UBL-5 from worm lysate (± mitochondrial/splicing stress) and search for branched di-glycine remnant peptides — a canonical tag would yield conjugates; UBL-5 should not.
- **C-terminal mutagenesis:** appending or deleting residues at the C-terminus should not affect conjugation (there is none) but should affect fold/partner binding — distinguishing tag vs. binding-module models.
- **Comparative motif/structure analysis:** confirm HIND-binding surface conservation in UBL-5 and absence of a conjugation-competent Gly across the Hub1/UBL5 ortholog set (worm/human/yeast/plant) vs. ubiquitin/SUMO/NEDD8 — already partially done here.
- **AlphaFold geometry:** verify the C-terminal tail is a helix/loop terminating in YYQ (not an extended exposed Gly-Gly), consistent with non-conjugation.

Curation Leads (require curator verification)

- **Action change:** Remove GO:0031386 (protein tag activity, IBA) from ubl-5, or record it as NOT-supported at the family-propagation step. Do **not** replace with bare "protein binding"; prefer a splicing/spliceosome-binding MF if one is warranted.
- **Retain:** GO:0000398 (mRNA splicing, via spliceosome, IBA) — now corroborated in *C. elegans*; GO:0034514 (mtUPR, IMP) as BP; CC terms.
- **Candidate references + exact snippets to verify:**
- PMID **24872507** — "*Different from canonical ubiquitin-like proteins, Hub1 does not form covalent conjugates with substrates but binds proteins non-covalently.*"
- PMID **21614000** — "*Hub1 binds non-covalently to a conserved element termed HIND, which is present in the spliceosomal protein Snu66 in yeast and mammals, and Prp38 in plants.*"
- PMID **36480405** (worm-specific) — "*The ubiquitin-like protein Hub1/UBL-5 associates with proteins non-covalently.*"
- PMID **25092792** — UBL5 "*primarily associates with spliceosomal proteins*"; depletion → intron retention.
- **Suggested curator question:** Is there any primary *C. elegans* evidence of a covalent UBL-5 conjugate? (None found; expected none.)
- **Suggested experiment:** GG-remnant proteomics on UBL-5 pulldowns to formally close the conjugation question.

Provenance (computed)

- UniProt P91302 FASTA retrieved; length 73; C-terminus `…IHEGFNFELYYQ`; no C-terminal Gly-Gly.
- C-terminal comparison panel (mature/precursor last residues): UBL5_worm `FNFELYYQ` (no GG), UBL5_human `MNLELYYQ` (no GG) vs. ubiquitin `…LRGG`, SUMO1 `…QTGG`, NEDD8 `…LRGG`, ISG15 `…LRGG` (all GG).
- GO:0031386 definition (QuickGO): requires covalent attachment/conjugation; comment restricts to conjugated tags.

- Domain assignment: InterPro IPR039732 (Hub1/Ubl5), IPR000626 (Ubiquitin-like domain); PANTHER PTHR13042 (UBIQUITIN-LIKE PROTEIN 5) — the recognized non-conjugated Ubl clade.
