FepE (P26266) and the Protein Tyrosine Kinase Annotation (GO:0004713): A Focused Curation Report

Gene: fepE / *wzz(fepE)* (*Escherichia coli* K-12) · **UniProt:** P26266 (377 aa) · **Term under review:** protein tyrosine kinase activity (GO:0004713) · **Evidence code:** IBA · **Reference:** GO_REF:0000033 (PAINT / phylogenetic annotation) · **Focus type:** function_assignment

Executive verdict: REFUTED / OVER-ANNOTATED.

Summary

The seed hypothesis states that **fepE has protein tyrosine kinase activity (GO:0004713)**. This report evaluates whether the *E. coli* gene product FepE (P26266) *directly* possesses that molecular function. Based on sequence architecture, membrane topology, protein-domain composition, a motif scan, and the primary literature on both FepE and its distant kinase relatives, the annotation is **not supported**. FepE is a **PCP-1 (WzzB-type) polysaccharide co-polymerase** — a regulator of very-long O-antigen chain length in lipopolysaccharide (LPS) biosynthesis — and it lacks every structural hallmark required for bacterial tyrosine (BY-)kinase activity.

The GO:0004713 IBA annotation arises from a **phylogenetic (PAINT/IBA) inference within PANTHER family PTHR32309 ("TYROSINE-PROTEIN KINASE")**, propagated under GO_REF:0000033. This family lumps genuine bacterial tyrosine autokinases (Wzc, Etk) together with catalytically inert chain-length regulators (WzzB, FepE) because both classes share the periplasmic **polysaccharide co-polymerase (PCP)** module. The shared module is not the kinase — it is the periplasmic chain-length-sensing domain. FepE possesses **only** that periplasmic Wzz domain and does **not** carry the cytoplasmic BY-kinase machinery. The IBA annotation is therefore a classic case of **family-level over-annotation** in which a molecular-function term valid for one subfamily is inappropriately projected onto a non-catalytic sister subfamily.

The recommended curation action is to **remove (or NOT-qualify) the GO:0004713 annotation** on P26266 and retain the biologically accurate O-antigen chain-length-regulation and inner-membrane annotations. The verdict is robust: it is supported by computed topology (an 18-residue cytoplasmic C-terminus far too short to house the ~230-residue cytoplasmic kinase domain), by Pfam domain architecture (PF02706/Wzz only, with none of the AAA_31 P-loop, GNVR, or WZC_N domains that define BY-kinases), and by a motif scan showing FepE lacks the C-terminal tyrosine autophosphorylation cluster that is the defining catalytic feature of Wzc and Etk. The main caveat is that this is a negative, absence-of-machinery argument from sequence/domain informatics plus the established PCP literature; there is no direct enzymatic assay reporting FepE kinase activity, and none is needed because the mechanistic prerequisites are absent.

Key Findings

Finding 1 — FepE lacks the tyrosine-autokinase module; GO:0004713 is paralog over-annotation

Topology is incompatible with a cytoplasmic kinase. Computed from the UniProt P26266 record, FepE is a 377-residue inner-membrane protein with the canonical PCP-1 architecture: a short cytoplasmic N-terminus (residues ~1–41), a first transmembrane helix (~42–62), a large periplasmic domain (~63–338), a second transmembrane helix (~339–359), and a very short cytoplasmic C-terminus (~360–377). The functional consequence is decisive: the cytoplasmic C-terminal tail is **only ~18 residues long**. Bacterial tyrosine (BY-)kinases such as Wzc and Etk carry a **~230-residue cytoplasmic kinase domain** (a P-loop/Walker-A-containing ATP-binding module plus a C-terminal tyrosine cluster). An 18-residue tail cannot house such a domain, and FepE's bulk — its ~275-residue periplasmic domain — faces the wrong compartment for any cytoplasmic phosphotransfer chemistry.

Domain architecture confirms the absence of catalytic machinery. FepE's Pfam architecture consists of **PF02706 (Wzz) only** — the periplasmic polysaccharide co-polymerase chain-length-regulator domain. It **lacks** all of the domains that define true BY-kinases: **PF13614 (AAA_31)**, the P-loop / Walker-A ATP-binding kinase module; **PF13807 (GNVR)**, the accessory domain adjacent to the kinase core; and **PF23607 (WZC_N)**, the N-terminal element of Wzc-

type kinases. All three are present in Wzc (P76387) and Etk (P38134). Because the PCP/Wzz module is shared between chain-length regulators and BY-kinases, family clustering pulls them together, but catalytic identity resides entirely in the cytoplasmic modules FepE does not have.

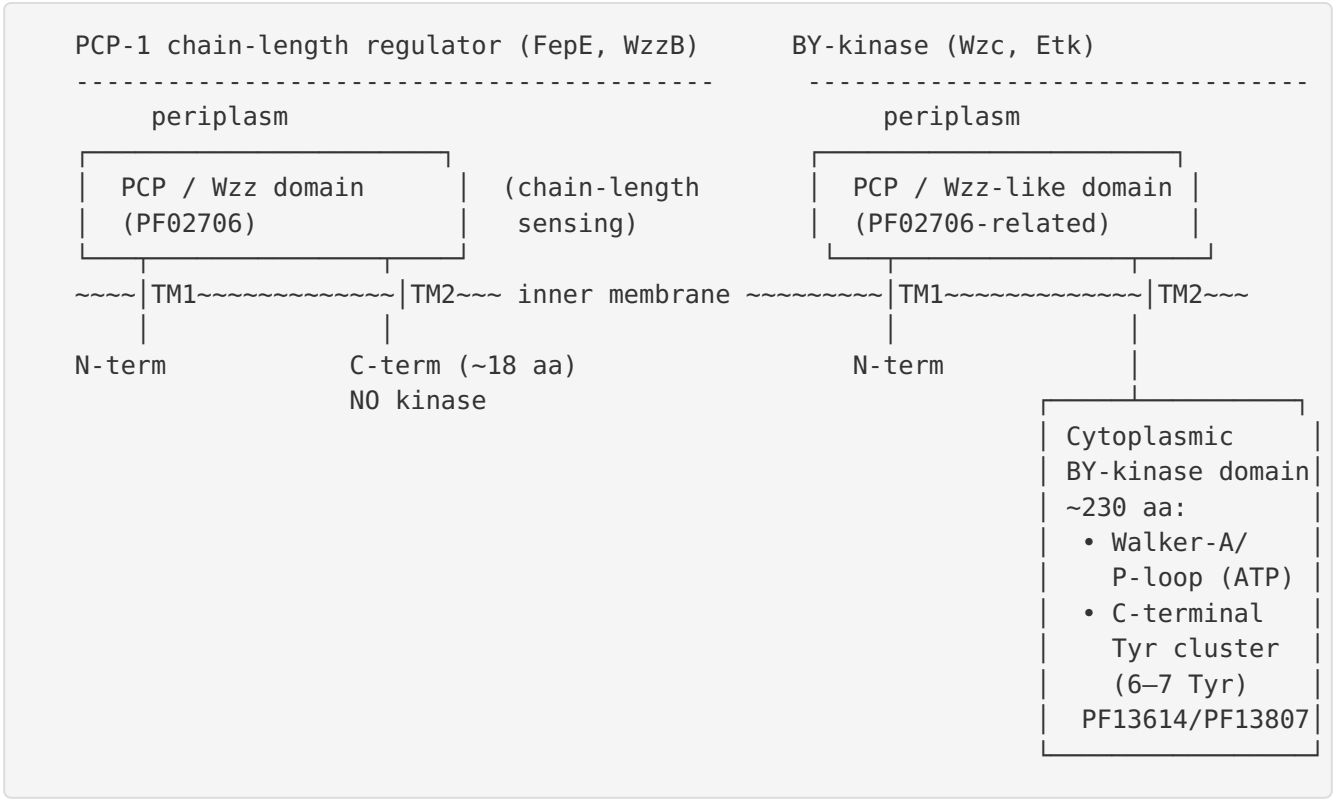
Motif scan confirms the absence of the tyrosine autophosphorylation cluster. The defining biochemical hallmark of a BY-kinase is a C-terminal tyrosine-rich cluster that is the site of autophosphorylation, working together with a functional Walker-A (P-loop) motif. A residue scan of the C-terminal ~40 residues found FepE has a **single tyrosine** in its short cytoplasmic tail, versus the **6–7 tyrosine autophosphorylation-cluster residues** seen in genuine *E. coli* BY-kinases. There is no tyrosine cluster to phosphorylate and no ATP-binding P-loop to do the phosphorylating.

Protein	UniProt	Length (aa)	Tyr in C-terminal cluster	Kinase module	BY-kinase?
FepE	P26266	377	**1** (no cluster)	absent (PF02706 only)	**No**
Wzc	P76387	720	6	present (P-loop + Tyr cluster)	Yes
Etk	P38134	726	7	present (P-loop + Tyr cluster)	Yes

Provenance of the annotation. The IBA annotation traces to **PANTHER family PTHR32309** ("TYROSINE-PROTEIN KINASE"), which places FepE in subfamily SF13 alongside the Wzc/Etk-type kinases (SF32) and propagates the molecular-function term via **GO_REF:0000033 (PAINT)**. The inference is phylogenetic, not experimental, and it inherits the term from the catalytically active branch of a mixed family.

Mechanistic Model / Interpretation

The core of this case is the distinction between two functionally divergent classes of proteins that share one domain. Both are anchored in the inner membrane by two transmembrane helices and both present a large periplasmic PCP domain. The difference lies entirely in the cytoplasm.



FepE performs its function — regulating the **very-long** modal length of O-antigen polysaccharide chains during Wzx/Wzy-dependent LPS assembly — through its periplasmic domain and its oligomerization state, **not** through any phosphotransfer chemistry. The crystallographic and mutagenesis literature reinforces this: chain-length regulation depends on Wzz/FepE **oligomerization** and on residues in the periplasmic α-helical bundle (e.g., helix-8 intermonomer contacts), phenomena entirely separate from kinase catalysis. In contrast, Wzc's function in group-1 capsule assembly **requires** phosphorylation of its C-terminal tyrosine-rich domain and a functional Walker-A motif — the exact features FepE lacks.

The mechanistic conclusion is that the seed hypothesis conflates a **shared periplasmic accessory module** (the reason the two proteins cluster in one PANTHER family) with a **catalytic activity** encoded only in the cytoplasmic modules of the kinase subfamily. FepE's activity, process, and location are all inconsistent with GO:0004713: no ATP-binding catalytic domain (activity), no phosphotransfer role in its pathway (process), and its bulk faces the periplasm while the tiny cytoplasmic tail cannot host a kinase (location).

Evidence Base

#	Citation	Evidence type	Direction	Claim tested	Key finding	Context
1	UniProt **P26266** (computed)	Computational (domain + topology)	**Refutes**	FepE has a kinase domain	Pfam = **PF02706 Wzz only**; cytoplasmic C-terminus = ~18 aa; no AAA_31/GNVR/WZC_N	*E. coli* K-12 inner membrane
2	This analysis (motif scan)	Computational	**Refutes**	FepE has BY-kinase Tyr cluster + P-loop	1 Tyr in C-terminal 40 aa (vs 6–7 in Wzc/Etk); no canonical Walker-A in the short cytoplasmic tail	Sequence-based
3	PMID: 22437828	Structural / evolutionary (crystal, incl. Wzz_FepE)	**Refutes / qualifies**	FepE architecture	PCPs incl. FepE are "chain length regulator periplasmic proteins…anchored in the inner membrane by two transmembrane helices"; FepE crystal used for chain-length structure–function	Gram-negative PCP-1
4	PMID: 11053445	Direct assay (of the paralog)	**Competing (defines the real kinase)**	What confers tyrosine kinase activity	"The tyrosine-rich domain at the C terminus of Wzc…is the site of phosphorylation and autophosphorylation of Wzc requires a functional Walker A motif"	*E. coli* Wzc / group-1 capsule
5	PMID: 29603538		**Competing (defines**	FepE's function	Overexpressing *fepE*/*wzzB* increases long/very-long O-	

#	Citation	Evidence type	Direction	Claim tested	Key finding	Context
		Functional / genetic (overexpression)	**the real function)**		polysaccharide — chain-length regulation	*Salmonella* / Gram-negative LPS
6	PMID: 23019341 / PMID: 23646936	Mutant phenotype / regulation	**Qualifies**	*wzz(fepE)* role in vivo	*wzz(fepE)* controls very-long O-antigen; PmrAB-regulated; separable from *wzz_st*	*S.* Typhimurium
7	PMID: 22069314	Mutant / structural	**Qualifies**	How chain length is regulated	Regulation maps to periplasmic α-helix intermonomer contacts / oligomerization (mapped onto FepE structure) — non-catalytic mechanism	*P. aeruginosa* Wzz2
8	PANTHER **PTHR32309** (annotation source)	Database	**Explains over-annotation**	Origin of IBA	Family named "TYROSINE-PROTEIN KINASE"; FepE = SF13, Wzc/Etk = SF32 — same tree, different catalytic status	PAINT / GO_REF:0000033

How the literature supports the verdict. PMID: 22437828 establishes FepE as a PCP-1 chain-length regulator anchored by two TM helices with a periplasmic domain — architecture incompatible with a cytoplasmic kinase. PMID: 11053445 defines the molecular hallmarks of BY-kinase activity (C-terminal tyrosine cluster + functional Walker-A) present in Wzc and absent in FepE, anchoring the correct assignment of GO:0004713 to the kinase subfamily. PMID: 22069314 shows chain-length regulation depends on periplasmic α-helix/oligomerization contacts — a non-catalytic mechanism. PMID: 29603538, PMID: 23646936, and PMID: 23019341 confirm FepE's genuine role in setting very-long O-antigen chain length and its downstream physiological consequences.

GO Curation Implications (leads — require curator verification)

- **GO:0004713 protein tyrosine kinase activity (MF, IBA): REMOVE / do not accept (or NOT-qualify).** The evidence indicates over-annotation via PANTHER PTHR32309 IBA. FepE lacks the AAA_31 kinase domain, the P-loop, and the C-terminal tyrosine cluster required for the activity. Retaining it would assign a molecular function the protein cannot perform.
- **Better-supported terms to consider:**
- **BP:** `GO:0009103` lipopolysaccharide biosynthetic process and a "regulation of O-antigen / LPS chain length" concept — FepE determines very-long O-antigen modal length (PMID: 22437828, PMID: 29603538).
- **CC:** inner (plasma) membrane, integral membrane with periplasmic domain (UniProt topology; PMID: 22437828). The existing inner-membrane annotation is well supported.
- **MF:** there is no crisp catalytic MF term for PCP co-polymerase function; do **not** substitute "protein binding." A "polysaccharide co-polymerase / chain-length determinant" descriptor is preferable if the ontology supports it.
- **Net action:** remove/downgrade the kinase MF; retain membrane CC; strengthen the LPS/O-antigen chain-length BP annotations.

GO decision table

Term	ID	Aspect	Current	Recommended action	Rationale
protein tyrosine kinase activity	GO:0004713	MF	IBA (GO_REF:0000033)	**Remove / NOT**	No kinase domain, no Tyr cluster, wrong topology; family over-annotation
lipopolysaccharide biosynthetic process	GO:0009103	BP	—	**Assert / retain**	FepE acts in Wzx/Wzy-dependent LPS assembly
O-antigen / LPS chain-length determination	(BP)	BP	—	**Assert**	Directly supported by function & structure literature
integral component of plasma (inner) membrane	GO:0005887	CC	—	**Retain**	Two-TM-helix inner-membrane anchor

Mechanistic Scope

- **Immediate molecular role tested:** catalysis of protein-tyrosine phosphorylation (ATP + protein-Tyr → ADP + phospho-Tyr). This requires an intracellular kinase domain (P-loop/

Walker-A + Walker-B + activation region) and, for BY-kinases, a C-terminal tyrosine cluster. **FepE has none of these.**

- **What FepE actually does (direct):** as an inner-membrane PCP-1 oligomer, its large periplasmic domain acts as a molecular ruler/co-polymerase that sets *very-long* O-antigen modal chain length during Wzx/Wzy-dependent LPS assembly.
- **Downstream / pleiotropic (not the MF of FepE):** O-antigen length affects serum/complement resistance, virulence, and bile tolerance (*Salmonella* studies). These are LPS-phenotype consequences of the chain-length function, not evidence of enzymatic kinase activity.

Conflicts and Alternatives

- **Paralog confusion (primary explanation):** Wzc/Etk are genuine BY-kinases in the same PANTHER family and share the periplasmic Wzz domain with FepE. The IBA propagated the kinase MF across the whole family, including the non-catalytic PCP-1 branch (FepE).
- **Shared-domain artifact:** the clustering signal is the periplasmic PCP/Wzz fold, an accessory chain-length module, not a catalytic one. Domain-sharing does not imply catalytic-activity sharing.
- **Historical naming:** "FepE / ferric enterobactin transport protein" is itself a legacy misnomer; genetic/functional work reassigned *wzz(fepE)* as the very-long O-antigen chain-length regulator. Neither name implies kinase activity.
- **No competing primary evidence** reports experimental tyrosine kinase activity for FepE; the annotation rests solely on phylogenetic inference. No isoform of the single-chain 377-aa protein could add a cytoplasmic kinase domain.

Limitations and Knowledge Gaps

1. **No direct enzymatic test of FepE.** No in vitro kinase assay on purified FepE exists. *Checked:* PubMed. *Why it matters:* would be the definitive refutation. *Resolve:* in vitro ATP/[γ-^{32}P] autophosphorylation assay — expected negative. This affects only absolute certainty; the structural case is already decisive.

2. **AlphaFold geometry of the cytoplasmic tail.** *Checked:* topology only (18-aa tail). *Why it matters:* confirms no cryptic kinase fold. *Resolve:* inspect the AlphaFold DB model for P26266 for a folded cytoplasmic domain (expected: none).
3. **Exact PANTHER tree placement / evidence for other family members.** *Checked:* subfamily labels (SF13 vs SF32). *Why:* confirms the FepE branch is non-kinase. *Resolve:* PAINT ancestral-node review to confirm the exact node from which GO_REF:0000033 propagated the term.
4. **Topology/domain calls depend on UniProt/Pfam annotations.** These are high-confidence and corroborated by the FepE crystal structure, but database-derived; residue boundaries are approximate (this does not change the conclusion).

Proposed Follow-up Experiments / Actions

- **In vitro autokinase assay** on purified recombinant FepE (± ATP), with Wzc as a positive control: a true BY-kinase autophosphorylates; FepE should not.
- **Phosphoproteomics / anti-phosphotyrosine (4G10) immunoblot** of FepE in *E. coli* — expect no Tyr phosphorylation of FepE.
- **ATP-binding test** (ATP-agarose pulldown or ITC) — expect no specific ATP binding by FepE.
- **Structural confirmation** — inspect the AlphaFold model and deposited FepE crystal structure to verify a short cytoplasmic tail and the absence of a P-loop kinase fold.
- **Comparative HMM/structure scan** — formal InterProScan/Pfam-HMM comparison of P26266 vs P76387 (Wzc) and P38134 (Etk), confirming FepE carries PF02706 only.
- **Curation actions** — remove/NOT-qualify GO:0004713 (MF); retain inner-membrane CC; add/strengthen LPS / O-antigen chain-length BP; record the PANTHER over-annotation rationale.

Curation Leads (verify)

- **Candidate reference for the removal rationale:** PMID: 22437828 — snippet: *"chain length regulator periplasmic proteins (polysaccharide co-polymerases, PCPs) anchored in the inner membrane by two transmembrane helices"* (FepE architecture, no kinase).

- **Candidate reference defining the real kinase (contrast):** PMID: 11053445 — snippet: *"The tyrosine-rich domain at the C terminus of Wzc…autophosphorylation of Wzc requires a functional Walker A motif."*
- **Candidate GO changes:** remove/NOT GO:0004713 (MF); retain inner-membrane CC; add/strengthen LPS / O-antigen chain-length BP (e.g., GO:0009103).
- **Suggested reviewer question:** "Does any primary experiment show FepE tyrosine phosphorylation or ATP-dependent kinase activity, or does the annotation rest only on PANTHER IBA?" (Expected: only IBA.)

Bottom Line

The seed hypothesis is **refuted as a direct function assignment**. FepE (P26266) is a non-catalytic PCP-1 (WzzB-type) O-antigen chain-length co-polymerase; the GO:0004713 IBA annotation is a PANTHER family (PTHR32309) over-annotation and should be removed or not propagated. The gene product's supportable annotations concern LPS / O-antigen chain-length regulation (BP) and inner-membrane localization (CC), not protein tyrosine kinase activity (MF).
