Functional Annotation Report: *zwf* (PP_5351, UniProt Q88C32) — Glucose-6-Phosphate 1-Dehydrogenase (G6PDH-C) of *Pseudomonas putida* KT2440

Summary

The gene **zwf** (ordered locus **PP_5351**; UniProt **Q88C32**) of *Pseudomonas putida* strain KT2440 (ATCC 47054 / DSM 6125 / NCIMB 11950) encodes **glucose-6-phosphate 1-dehydrogenase (G6PD/G6PDH; EC 1.1.1.49)**, a cytoplasmic $NADP^+$-dependent oxidoreductase that catalyzes the first committing step of the oxidative branch of the pentose-phosphate pathway. The enzyme oxidizes **D-glucose 6-phosphate to 6-phospho-D-glucono-1,5-lactone**, transferring a hydride to $NADP^+$ to yield NADPH + H^+ (Rhea:15841). This assignment is unambiguous and is supported by convergent lines of evidence from database annotation (UniProt, KEGG, InterPro), purified-enzyme biochemistry, ^{13}C metabolic-flux physiology, and structural/sequence bioinformatics. The gene symbol *zwf* correctly matches the protein description, the organism is confirmed as *P. putida* KT2440, and the protein family/domain content (G6PD family; Rossmann NAD(P)-binding + C-terminal β+α architecture) is fully consistent with the literature. **No mis-identification issue was encountered.**

Critically, PP_5351 is not the organism's only G6PDH. *P. putida* KT2440 carries **three co-existing G6PD isozymes** encoded by *zwfA* (PP_1022), *zwfB* (PP_4042), and *zwfC* (PP_5351). PP_5351 is specifically **G6PDH-C (zwfC)**. Direct enzymology on the purified protein (Volke et al. 2021, *mSystems*) establishes that G6PDH-C is the **most strongly $NADP^+$-specific** of the three isozymes (cofactor specificity constant $\varphi \approx 2082$, roughly five-fold more $NADP^+$-selective than the *E. coli* enzyme) yet is a **catalytically slow, near-dispensable minor isozyme**: its turnover number is below 1 s^{-1} for both cofactors (kcat ≈ 0.77 s^{-1} with NAD^+, 0.54 s^{-1} with $NADP^+$), and deletion of *zwfC* produces no measurable growth phenotype under any tested condition. The bulk of physiological G6PDH flux is carried by the dual-cofactor ZwfA and the NAD^+-preferring ZwfB.

Functionally, G6PDH sits at a pivotal node of *P. putida*'s central carbon metabolism. This bacterium lacks a functional Embden-Meyerhof-Parnas (EMP) glycolysis and catabolizes glucose almost exclusively through the **Entner-Doudoroff (ED) pathway**, routing ~90% of consumed sugar through gluconate/6-phosphogluconate; only ~20% of glucose is phosphorylated to glucose-6-phosphate, the substrate on which G6PDH acts. The enzyme operates within the cyclic **EDEMP** architecture (Entner-Doudoroff + EMP-gluconeogenic + pentose-phosphate reactions) that biases the cell toward net **NADPH formation**, supplying reducing power for biosynthesis and oxidative-stress defense. The *zwfC* gene is monocistronic and lies next to a divergently transcribed RpiR/HexR-family regulator, genomically distinct from the HexR-controlled glucokinase operon that houses the principal isozyme *zwfA*.

Gene/Protein Identity Verification

Before presenting the functional analysis, the mandatory identity checks were completed and all passed:

Check	Result	Status
Gene symbol *zwf* matches protein description	UniProt Q88C32: Name=*zwf*; product = Glucose-6-phosphate 1-dehydrogenase (G6PD), EC 1.1.1.49	Match
Organism correct	*Pseudomonas putida* KT2440 (ATCC 47054 / DSM 6125 / NCIMB 11950); locus PP_5351 (EMBL AAN70916.1)	Confirmed
Protein family/domains align with literature	G6PD family; IPR001282, IPR019796, IPR022674/5, IPR036291; canonical motifs found in sequence	Aligned
Ambiguity / wrong-gene literature	No conflicting gene with same symbol; all literature is on bona fide G6PDH in *Pseudomonas*	No conflict

The one nuance worth flagging is **paralogy, not mis-identification**: because *P. putida* encodes three G6PDH genes (*zwfA/B/C*), care was taken throughout to attribute isozyme-specific properties (kinetics, dispensability, genomic context) specifically to **G6PDH-C** = **zwfC** = **PP_5351**, as distinguished in the primary source (Volke et al. 2021, PMID: 33727391), whose full text states: *"the three isoforms of this enzyme are encoded by zwfA (PP_1022), zwfB (PP_4042), and zwfC (PP_5351)."*

Key Findings

Finding 1 — PP_5351 is G6PDH-C, the third G6PD isozyme of *P. putida* KT2440

PP_5351 (*zwf*/Q88C32) is definitively **glucose-6-phosphate 1-dehydrogenase**, catalyzing the reaction:

> **D-glucose 6-phosphate** + **$NADP^+$** → **6-phospho-D-glucono-1,5-lactone** + **NADPH** + **H^+** (EC 1.1.1.49; Rhea:15841)

This assignment is corroborated across independent databases. UniProt Q88C32 records Name=*zwf*, OrderedLocusNames=PP_5351, EC 1.1.1.49, with the reaction as above. KEGG ppu:PP_5351 assigns orthology **K00036** and maps the gene to the **Pentose phosphate pathway (ppu00030)**, the **Entner-Doudoroff module (M00008)**, and the **oxidative PPP module (M00006)**. The primary literature explicitly names PP_5351 as *zwfC* / G6PDH-C.

The enzymatic role is captured by the source review's opening statement:

> *"Glucose-6-phosphate dehydrogenase (G6PDH) is widely distributed in nature and catalyzes the first committing step in the oxidative branch of the pentose phosphate (PP) pathway, feeding either the reductive PP or the Entner-Doudoroff pathway."* — PMID: 33727391

This defines both the reaction (first committing, oxidative step) and the two downstream metabolic destinations (reductive PP for pentose/NADPH; ED for glycolytic carbon), of which PP_5351/*zwfC* is one contributing isozyme.

Finding 2 — G6PDH-C is strongly $NADP^+$-specific but catalytically slow and physiologically near-dispensable

Volke et al. 2021 purified all three isozymes and characterized their steady-state kinetics. G6PDH-C (zwfC/PP_5351) is a study in contrasts: it has the **strongest cofactor discrimination** but the **weakest catalytic power** of the three.

Parameter	G6PDH-C (zwfC/PP_5351)	Interpretation
Km for $NADP^+$	~3.2 µM (lowest of the three)	Very tight $NADP^+$ binding
Km for NAD^+	~9.5×10^3 µM (highest of the three)	Very poor NAD^+ binding
Cofactor specificity constant φ	≈ 2082	~5× more $NADP^+$-selective than *E. coli* G6PDH (φ ≈ 410)
kcat (NAD^+)	0.77 s^{-1}	Very low turnover
kcat ($NADP^+$)	0.54 s^{-1}	Very low turnover
Activity on ribose	near-zero	Not induced/active under ribose growth
Δ*zwfC* growth phenotype	negligible under all conditions	Physiologically dispensable

The strong $NADP^+$ preference (φ ≈ 2082) means that when G6PDH-C is active it channels reducing equivalents specifically toward the **NADPH pool** rather than NADH. However, its sub-unity kcat marks it as a **minor, low-flux isozyme**; the physiologically dominant G6PDH activity — and all measurable growth phenotypes — are carried by **ZwfA (dual-cofactor)** and **ZwfB (NAD^+-preferring)**. This positions PP_5351 as a specialist "fine-tuning" isozyme rather than a workhorse. The source paper that generated these measurements is the same review that defines the family and pathway role (PMID: 33727391).

Finding 3 — Spatial and metabolic context: a cytoplasmic enzyme feeding the ED/EDEMP cycle and NADPH supply

G6PDH-C carries out its reaction in the **cytosol** (GO:0005829). Its metabolic setting is unusual and important for interpreting its role. *P. putida* KT2440 **lacks a functional EMP glycolysis** and catabolizes glucose almost exclusively via the **Entner-Doudoroff pathway**. ^{13}C metabolic-flux analysis (Nikel et al. 2015, *J Biol Chem*) showed:

> *"Metabolic flux analysis demonstrated that 90% of the consumed sugar was converted into gluconate, entering central carbon metabolism as 6-phosphogluconate and further channeled into the ED pathway."* — PMID: 26350459

Because only ~20% of glucose is phosphorylated to glucose-6-phosphate (the G6PDH substrate) while ~90% is oxidized to gluconate, G6PDH operates on a **minority carbon stream**. It sits at the connecting node of the ED, EMP-gluconeogenic, and PP reactions that together form the **EDEMP cycle**, a recycling architecture biased toward NADPH production:

> *"Cells growing on glucose thus run a biochemical cycle that favors NADPH formation."* — PMID: 26350459

This makes the **$NADP^+$-specificity of G6PDH-C mechanistically coherent** with the cell's need to generate NADPH for anabolism and redox defense. The glucokinase/G6PDH genes that yield 6-phosphogluconate are under transcriptional control of the HexR repressor:

> *"HexR controls the genes that encode glucokinase/glucose 6-phosphate dehydrogenase that yield 6-phosphogluconate"* — PMID: 18245293

GO annotation summarizes the enzyme: function = glucose-6-phosphate dehydrogenase activity (GO:0004345), NADP binding (GO:0050661); process = oxidative branch of the pentose-phosphate shunt (GO:0009051); location = cytosol (GO:0005829).

Finding 4 — Sequence/structure evidence: all canonical G6PD signatures are present

The 485-residue Q88C32 sequence carries every diagnostic hallmark of the G6PD family, providing structure-based confirmation of function independent of the biochemical assay:

1. **Rossmann dinucleotide-binding fingerprint** — the N-terminal glycine-rich motif **GxxGDLA (GGTGDLA, residues 16–22)** within the Pfam G6PD_N domain (PF00479); this is the NAD(P)-binding element.
2. **Substrate-binding motif** — the conserved glucose-6-phosphate-binding **EKPIG (residues 148–152)**.
3. **PROSITE PS00069 "G6P_DEHYDROGENASE" active-site signature** — spanning **DHYLGKE (residues 178–184)**, containing the invariant **catalytic histidine His179** that serves as the general base/proton acceptor.

InterPro domain content is fully canonical: IPR001282 (G6P_DH), IPR019796 (G6P_DH active site), IPR022674 (NAD-binding), IPR022675 (C-terminal domain; PF02781 G6PD_C), and IPR036291 (NAD(P)-binding Rossmann-like superfamily). The ~485-aa length and two-domain architecture (N-terminal Rossmann + C-terminal β+α domain) are textbook G6PD. UniProt appends a CAUTION that the sequence *"lacks conserved residue(s) required for the propagation of feature annotation"* — a note that dovetails with the enzyme's measured divergence (low turnover, atypical isozyme role). The family/functional assignment implied by these motifs matches the enzymatic definition from the source review (PMID: 33727391).

Finding 5 — *zwfC* is monocistronic and genomically distinct from operon-embedded *zwfA*

The three *zwf* paralogues differ in genomic organization, which underlies their distinct regulation. From KEGG genomic coordinates:

```
PP_5350  (complement 6098222..6099088)  RpiR/HexR-family transcriptional regulator  <- divergent
PP_5351  (6099177..6100634)             zwfC / G6PDH-C  (this gene)
PP_5352  (6100849..6103035)             uvrD-like DNA helicase II (repair, unrelated)
```

Thus **zwfC is a standalone (monocistronic) gene** flanked on one side by a divergent RpiR/HexR-family regulator and on the other by an unrelated DNA-repair helicase. This contrasts sharply with the principal isozyme **zwfA (PP_1022)**, which lies within the **HexR-regulated glucokinase operon** and is glucose-induced. The genomic separation of *zwfC* is consistent with it being differentially (and likely more weakly/conditionally) regulated than the main glycolytic isozyme, supporting its status as a specialist rather than the constitutive workhorse. HexR regulation of the glucokinase/G6PD (zwfA) operon is documented directly:

> *"HexR controls the genes that encode glucokinase/glucose 6-phosphate dehydrogenase that yield 6-phosphogluconate"* — PMID: 18245293

Finding 6 — High-confidence AlphaFold model supports the predicted two-domain G6PD fold

No experimental (X-ray or cryo-EM) structure of *P. putida* G6PDH-C has been deposited in the PDB. However, the **AlphaFold DB entry AF-Q88C32-F1 (model v6)** covers all 485 residues with a **global pLDDT of 93.2** — very high confidence. Per-residue analysis: 84% of Cα atoms have pLDDT ≥ 90 (very high), 98% ≥ 70 (confident), and only ~1% below 50. This strongly supports

the predicted canonical two-domain G6PD fold (N-terminal Rossmann NAD(P)-binding domain + C-terminal β+α domain). KEGG orthology K00036 assigns both EC 1.1.1.49 and EC 1.1.1.363 (the $NAD(P)^+$ dual-specificity variant), consistent with G6PD family members capable of using either cofactor.

Finding 7 — Concluding synthesis: annotation resolved across four independent evidence types

The identity and function of Q88C32/PP_5351 are corroborated by four convergent evidence classes:

Evidence type	Key result
Database annotation	UniProt (EC 1.1.1.49, Rhea:15841, cytosol); KEGG K00036 (ppu00030, ED module M00008); InterPro G6PD domains
Purified-enzyme enzymology	Volke et al. 2021: strongly $NADP^+$-preferring ($\varphi \approx 2082$) but slow (kcat < 1 s^{-1}) and dispensable (no Δ*zwfC* growth phenotype)
Physiological / flux context	Nikel et al. 2015 (^{13}C-MFA): cytoplasmic G6PDH acts on the ~20% phosphorylated glucose fraction, feeding the ED pathway and NADPH-generating EDEMP cycle; HexR regulation (del Castillo et al. 2008)
Structure / evolution	All canonical motifs present (GGTGDLA, EKPIG, PS00069/His179); high-confidence AlphaFold model (pLDDT 93.2); one of three paralogues arising from gene duplication with relaxed/specialized cofactor specificity

Mechanistic Model / Interpretation

The reaction and its place in central metabolism

G6PDH-C catalyzes the oxidation of the C1 hydroxyl of glucose-6-phosphate to a lactone, with hydride transfer to $NADP^+$:

```
                G6PDH-C (zwfC / PP_5351)
                NADP+ -> NADPH + H+
                       |
D-glucose 6-phosphate -+--> 6-phospho-D-glucono-1,5-lactone
 (catalytic base: His179; substrate anchored by EKPIG motif;
  cofactor bound by Rossmann GGTGDLA fingerprint)
```

The product 6-phosphogluconolactone is hydrolyzed (by 6-phosphogluconolactonase) to 6-phosphogluconate, the pivotal intermediate that either continues through the oxidative PP pathway or is dehydrated/cleaved by the **Entner-Doudoroff** enzymes (Edd/Eda) to pyruvate + glyceraldehyde-3-phosphate.

Why a slow, hyper-specific third isozyme?

The *P. putida* glucose metabolic map explains the seemingly paradoxical properties of G6PDH-C:

```
 glucose --(90%)--> gluconate --> 2-ketogluconate / 6-phosphogluconate --> ED (Edd/Eda)
    |                                                 ^
 (~20% phosphorylated)                                |
    v                                                 |
glucose-6-P --[G6PDH: ZwfA/ZwfB/ZwfC]--> 6-P-gluconolactone --> 6-P-gluconate
                  NADP+/NAD+ -> NADPH/NADH
                                                      |
                        EDEMP cycle (ED + EMP-gluconeogenic + PP)  -->  net NADPH
```

- Because most glucose is oxidized peripherally to gluconate (bypassing G6PDH), the phosphorylative G6PDH branch handles a **minority flux**. A high-Vmax isozyme is not strictly required here, which is consistent with the **low kcat** and **dispensability** of G6PDH-C.
- The **EDEMP cycle biases the cell toward NADPH**. G6PDH-C's extreme $NADP^+$ selectivity ($\varphi \approx 2082$) makes it a "clean" NADPH-generating catalyst that does not squander flux producing NADH. It plausibly acts as a **fine-tuning / redox-balancing specialist** — engaged under specific conditions (e.g., oxidative stress, particular carbon regimes) to nudge the NADPH/NADH ratio — while ZwfA (dual-cofactor, glucose-induced, operon-embedded) and ZwfB (NAD^+-preferring) carry the constitutive load.
- The **divergent genomic context** (monocistronic, next to an RpiR/HexR-family regulator, separate from the glucokinase operon) reinforces the interpretation that *zwfC* is under distinct regulatory control, i.e., a conditionally deployed accessory isozyme.

Isozyme comparison

Isozyme	Locus	Cofactor preference	Catalytic role	Genomic context
ZwfA (G6PDH-A)	PP_1022	Dual (NAD^+/ $NADP^+$)	Predominant workhorse; carries growth phenotype	HexR-regulated glucokinase operon; glucose-induced
ZwfB (G6PDH-B)	PP_4042	NAD^+-preferring	Major contributor to flux	—
ZwfC (G6PDH-C)	**PP_5351**	**Strongly $NADP^+$ ($\varphi \approx 2082$)**	**Minor, slow (kcat<1 s^{-1}), dispensable specialist**	**Monocistronic; divergent RpiR/HexR-family regulator**

Evidence Base

PMID	Title (abbrev.)	How it supports the findings
33727391	*Cofactor Specificity of G6PD Isozymes in P. putida Reveals a General Principle Underlying Glycolytic Strategies in Bacteria* (Volke et al. 2021, mSystems)	**Primary source.** Defines the reaction and pathway role; systematically characterized all three isozymes; provides the kinetics (φ, Km, kcat) and dispensability data specific to G6PDH-C/zwfC/PP_5351.
26350459	*Pseudomonas putida KT2440 Metabolizes Glucose through a Cycle Formed by ED, EMP, and PP Pathways* (Nikel et al. 2015, J Biol Chem)	Establishes the EDEMP cycle and that ~90% of glucose is routed through gluconate/ED; shows the cycle favors NADPH formation — the metabolic context for G6PDH's $NADP^+$ specificity.
18245293	*A set of activators and repressors control peripheral glucose pathways in P. putida* (del Castillo et al. 2008)	Documents HexR transcriptional regulation of the glucokinase/G6PD locus (zwfA operon), contrasting with the distinct genomic context of zwfC.
36354357	*G6PD ZwfA, a Dual Cofactor-Specific Isozyme… in P. bharatica CSV86*	Comparative support: confirms Zwf/G6PDH function (G6P → 6-phosphogluconolactone; NAD(P)H generation; ED + oxidative PPP; oxidative-stress role) across *Pseudomonas*.
24951791	*The functional structure of central carbon metabolism in P. putida KT2440*	Context: central carbon metabolism regulation is largely post-transcriptional/metabolic, relevant to isozyme deployment.
22434849	*Regulatory tasks of the PTS of P. putida in central carbon metabolism*	Context: carbon flux split between ED, PP, and EMP pathways and its regulation.

Additional literature reviewed for context on *Pseudomonas* central metabolism and multi-omics flux:

P 30936206,

P 30831266,

P 41259139.

Limitations and Knowledge Gaps

1. **No isozyme-specific in vivo condition identified.** Deletion of *zwfC* has no detectable growth phenotype under all tested conditions, so the precise physiological situation in which G6PDH-C is deployed (if any) remains undefined. Its extreme $NADP^+$ specificity suggests a redox/oxidative-stress niche, but this has not been directly demonstrated for *zwfC*.
2. **No experimental structure.** The structural model rests on a high-confidence AlphaFold prediction (pLDDT 93.2); no X-ray or cryo-EM structure of *P. putida* G6PDH-C exists. Catalytic-residue assignments (e.g., His179) are by homology/motif inference, not by crystallographic or mutagenic proof for this specific protein.
3. **Regulation of *zwfC* is inferred, not measured.** The divergent RpiR/HexR-family regulator (PP_5350) adjacent to *zwfC* is a plausible dedicated regulator, but no experimental data confirm that it controls *zwfC* expression or identify inducing conditions.
4. **Oligomeric state and allostery unknown.** Many G6PDHs function as dimers/tetramers with cofactor-dependent assembly; the quaternary structure and any allosteric regulation of G6PDH-C are uncharacterized.
5. **UniProt CAUTION.** The annotation notes the sequence lacks some conserved residues used for feature propagation, consistent with its divergence — a caveat for automated functional transfer.

Proposed Follow-up Experiments / Actions

1. **Conditional phenotyping of Δ*zwfC*.** Screen the deletion strain under oxidative stress (H_2O_2, paraquat, diamide), nitrogen/carbon limitation, and diverse carbon sources to reveal a niche where $NADP^+$-specific G6PDH-C matters, using growth, redox ($NADPH/NADP^+$ ratio), and survival readouts.
2. **Regulator characterization.** Delete/overexpress PP_5350 (RpiR/HexR-family) and measure *zwfC* transcript/protein (qRT-PCR, reporter fusions) to test whether it is the dedicated regulator and identify effector metabolites.

3. **Structural determination.** Solve the crystal or cryo-EM structure of purified G6PDH-C (ideally with G6P and $NADP^+$) to confirm the AlphaFold fold, the His179 catalytic role, and the structural basis of extreme $NADP^+$ selectivity.
4. **Site-directed mutagenesis.** Mutate His179 and the cofactor-discriminating residues to validate catalysis and to test whether relaxing $NADP^+$ specificity alters in vivo redox balance.
5. **Flux re-partitioning.** Use ^{13}C-MFA in triple/double *zwf* mutant backgrounds to quantify the specific flux contribution of G6PDH-C and its impact on the NADPH budget under engineered conditions.
6. **Enzyme-engineering context.** Given *P. putida*'s role as a metabolic-engineering chassis, evaluate whether overexpressing the hyper-$NADP^+$-specific G6PDH-C can boost NADPH supply for reductive biotransformations.

Conclusion

P. putida KT2440 **zwf/PP_5351 (Q88C32)** encodes **glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)**, specifically the **G6PDH-C / ZwfC** isozyme — a cytoplasmic enzyme that oxidizes D-glucose-6-phosphate to 6-phospho-D-glucono-1,5-lactone with reduction of $NADP^+$ to NADPH, initiating the oxidative pentose-phosphate branch and feeding carbon into the dominant Entner-Doudoroff pathway within the NADPH-favoring EDEMP cycle. Among the organism's three G6PDH isozymes it is the most strongly $NADP^+$-selective ($\varphi \approx 2082$) but the slowest ($k_{cat} < 1\ s^{-1}$) and physiologically near-dispensable, functioning as a specialist NADPH-generating/redox-tuning accessory while ZwfA and ZwfB carry the constitutive flux. The annotation is fully resolved by convergent database, enzymatic, physiological-flux, and structural/sequence evidence.
