# Functional Annotation Report: *ilvC* / Ketol-Acid Reductoisomerase (Q88DZ0) in *Pseudomonas putida* KT2440

## Summary

The gene ***ilvC*** (ordered locus **PP_4678**; UniProt **Q88DZ0**) of *Pseudomonas putida* strain KT2440 (ATCC 47054 / DSM 6125 / NCIMB 11950) encodes **ketol-acid reductoisomerase (KARI)**, also called **acetohydroxy-acid isomeroreductase (AHIR/AHAIR)**, **EC 1.1.1.86**. This is the **second enzyme of the branched-chain amino acid (BCAA) biosynthesis pathway**, which produces L-valine, L-isoleucine, and L-leucine. The enzyme is a soluble, cytoplasmic, 338-residue **class I ("type 1")** KARI that uses **two $Mg^{2+}$ ions and NADPH** to carry out, within a single active site, two chemically distinct reactions: an **$Mg^{2+}$-dependent alkyl (methyl or ethyl) migration/ isomerization** followed by an **NADPH-dependent reduction**. Gene identity was rigorously verified: the gene symbol, organism, EC number, protein family (ketol-acid reductoisomerase family), and the diagnostic InterPro domains (KARI_N IPR013116, KARI_C IPR000506, KARI_prok IPR014359) are all internally consistent, and the recovered literature describes exactly this enzyme class. There is no ambiguity — *ilvC* unambiguously denotes KARI.

Mechanistically, KARI converts **(S)-2-acetolactate → (2R)-2,3-dihydroxy-3-isovalerate** (the valine/leucine branch) and **(S)-2-aceto-2-hydroxybutyrate → (2R,3R)-2,3-dihydroxy-3-methylvalerate** (the isoleucine branch). It sits immediately downstream of acetohydroxyacid synthase (AHAS, encoded by *ilvBN*, which supplies its substrates) and immediately upstream of dihydroxyacid dehydratase (DHAD, *ilvD*). In *P. putida* KT2440, PP_4678 (*ilvC*) is genomically clustered with AHAS genes in a canonical ***ilvBN-ilvC*-type arrangement**, placing the enzyme physically and transcriptionally adjacent to its substrate-producing partner. The product ultimately feeds not only the three BCAAs but also 2-ketoisovalerate-derived **pantothenate and coenzyme A biosynthesis**.

The enzyme carries out its function in the **cytoplasm**, where the entire soluble BCAA biosynthetic pathway resides. Because this pathway is present in plants, fungi, and bacteria but **absent in animals**, KARI is a validated target for antimicrobials and herbicides, and — because of its NADPH dependence — a heavily engineered node in microbial **isobutanol**

**biofuel** production. The findings below are supported by four converging lines of evidence: (1) authoritative enzymological and structural literature on KARI mechanism, (2) the curated UniProt/HAMAP annotation specific to Q88DZ0, (3) genomic-context analysis of the PP_4678 locus in KT2440, and (4) direct bioinformatic analysis of the Q88DZ0 sequence identifying its cofactor-binding and metal-coordinating residues.

---

## Key Findings

### F001 — *ilvC* encodes ketol-acid reductoisomerase, the second step of BCAA biosynthesis

The core functional assignment is definitive. Q88DZ0 is annotated as **EC 1.1.1.86** under HAMAP-Rule MF_00435. KARI is a bifunctional enzyme that catalyzes two chemically distinct steps within a single active site: **(1)** an **$Mg^{2+}$-dependent alkyl migration (isomerization)**, in which a methyl or ethyl group migrates to convert (S)-2-acetolactate to 3-hydroxy-3-methyl-2-oxobutanoate, and **(2)** an **NADPH-dependent reduction** of the resulting keto group. The net conversions are:

- **Valine/leucine branch:** (S)-2-acetolactate → **(2R)-2,3-dihydroxy-3-isovalerate**
- **Isoleucine branch:** 2-aceto-2-hydroxybutyrate → **(2R,3R)-2,3-dihydroxy-3-methylvalerate**

The substrate specificity is strict for the acetohydroxy-acid intermediates of the *ilv* pathway. This is directly established by primary enzymology: KARI "catalyzes the conversion of 2-acetolactate into (2R)-2,3-dihydroxy-3-isovalerate or the conversion of 2-aceto-2-hydroxybutyrate into (2R,3R)-2,3-dihydroxy-3-methylvalerate. KARI catalyzes two reactions—alkyl migration and reduction—and requires Mg(2+) and NADPH for activity" (PMID: 19362563). Its ordinal position is equally clear: "Ketol-acid reductoisomerase (KARI) is the second enzyme in the branched-chain amino acid biosynthesis pathway" (PMID: 23036858).

### F002 — KARI is an $Mg^{2+}$/NADPH-dependent, two-domain enzyme using an induced-fit mechanism

Structurally, KARI is built from two domains: an **N-terminal Rossmann-like NAD(P)H-binding domain** (KARI_N, IPR013116) and a **C-terminal knotted α-helical domain** (KARI_C, IPR000506) that houses the two catalytic $Mg^{2+}$ ions. The active site lies **at the interface between the two domains**. Crystallographic studies demonstrate that the binding of $Mg^{2+}$ and

NADPH triggers a **domain-closure/reorganization (induced fit)** that assembles the catalytically competent active site. As stated for the plant enzyme, "the binding of Mg(2+) and NADPH opens the interface between the N- and C-domains, thereby allowing access for the substrates to bind" (PMID: 23036858). For the class I enzymes to which the prokaryotic *ilvC* belongs, "the class I KARI structures indicate that the active sites close upon binding NAD(P)H" (PMID: 25849365), directly confirming the induced-fit closure mechanism. Isothermal titration calorimetry on the *E. coli* enzyme further shows $Mg^{2+}$ binding increases conformational disorder while NADPH binding increases rigidity, consistent with the ordered assembly of the active site.

## F003 — Q88DZ0 is a class I / "type 1" prokaryotic KARI

UniProt annotates Q88DZ0 explicitly as "Ketol-acid reductoisomerase type 1 / type I" and assigns the **prokaryotic KARI signature (KARI_prok, IPR014359)**. KARIs partition into two classes: **class I** (short, single ~340-residue subunit, predominant in bacteria and fungi) and **class II** (arising from an internal duplication of the C-domain, ~490+ residues, found in plants and some bacteria such as *E. coli*). Q88DZ0's 338-residue length and prokaryotic signature place it firmly in class I. The literature notes that "most sequenced members of the industrially important ketol-acid reductoisomerase (KARI) family are class I enzymes, [while] structural studies to date have focused primarily on the class II KARIs, which arose through domain duplication" (PMID: 25849365). The class distinction is functionally relevant because it correlates with the length of the β2αB cofactor-specificity loop that governs NADPH-versus-NADH preference.

## F004 — KARI operates in the cytoplasm; it is a validated antimicrobial target absent from animals and a key biofuel-engineering node

The BCAA biosynthesis pathway — and KARI with it — is "found in plants, fungi and bacteria but not in animals. This difference in metabolism between animals and microorganisms makes KARI an attractive target for the development of antimicrobial agents" (PMID: 23036858). The functional importance of KARI in vivo is demonstrated genetically across taxa: in *Mycobacterium tuberculosis*, "Mtb shows survival deficit in macrophages and in mice after ketol-acid reductoisomerase down-regulation" (PMID: 36354071); and in the fungal pathogen *Fusarium graminearum*, deletion of the KARI homologue FgIlv5 abolishes growth without exogenous isoleucine and valine and reduces virulence (PMID: 24493249). As a soluble cytosolic enzyme, KARI is also a central node in engineered isobutanol production. Because glycolysis yields NADH while wild-type KARI uses NADPH, cofactor-swapped variants were

engineered: "an NADH-dependent pathway enables anaerobic isobutanol production at 100% theoretical yield" (PMID: 21515217), underscoring KARI's NADPH specificity as the natural default and its role as an engineerable bottleneck.

## F005 — Protein-specific annotation: 338-residue two-domain enzyme, 2 $Mg^{2+}$/subunit, step 2/4 of both Val and Ile routes

The curated UniProt Q88DZ0 record (HAMAP-Rule MF_00435) provides protein-specific detail. The sequence is **338 amino acids** long, consistent with a compact single-subunit class I KARI. Its domain architecture comprises an **N-terminal KARI Rossmann domain (~residues 1–181)** containing the Rossmann NAD(P)-binding fingerprint (...GYGSQGHA... around residues 22–29) and a **C-terminal KARI knotted domain (~182–327)**. An active-site residue is annotated at position **107**, with multiple metal/substrate binding sites (residues 24–27, 47, 50, 52, 82–85, 133, 190, 194, 226, 230, 251). The enzyme binds **two $Mg^{2+}$ per subunit** and uses **NADPH**. The curated Rhea reactions are:

- **RHEA:22068** — (2R)-2,3-dihydroxy-3-methylbutanoate + $NADP^+$ = (2S)-2-acetolactate + NADPH + $H^+$
- **RHEA:13493** — (2R,3R)-2,3-dihydroxy-3-methylpentanoate + $NADP^+$ = (S)-2-ethyl-2-hydroxy-3-oxobutanoate + NADPH + $H^+$

Pathway assignments are **L-valine biosynthesis from pyruvate, step 2/4** and **L-isoleucine biosynthesis from 2-oxobutanoate, step 2/4**.

## F006 — PP_4678 is genomically clustered with acetohydroxyacid synthase genes and feeds Val/Ile and pantothenate/CoA pathways

Genomic-context analysis of *P. putida* KT2440 (KEGG/GenBank) places **PP_4678 (*ilvC) at complement(5316083..5317099), immediately preceded on the same minus strand by PP_4679 (~163 aa, acetohydroxyacid synthase small/regulatory subunit, ilvN/ilvH-like; 57-bp gap to ilvC) and PP_4680 (~574 aa, acetohydroxyacid synthase large/catalytic subunit, ilvB/ilvI-like; ~2-bp gap to PP_4679). The transcription order on the minus strand is therefore AHAS-large → AHAS-small → ilvC***, a canonical ***ilvB(I)-ilvN(H)-ilvC* cluster**. This physically co-locates KARI with the AHAS enzyme that produces its substrates. KEGG annotations assign KO **K00053** (KARI, EC 1.1.1.86); modules **M00019** (Val/Ile biosynthesis) and **M00570** (Ile biosynthesis from threonine); and pathways **ppu00290** (Val/Leu/Ile biosynthesis),

**ppu00770** (pantothenate & CoA biosynthesis), **ppu01210** (2-oxocarboxylic acid metabolism), and **ppu01230** (amino-acid biosynthesis). The pantothenate/CoA link reflects that the KARI-DHAD product 2-ketoisovalerate is a branch-point precursor for pantothenate.

### F007 — Sequence analysis predicts NADPH specificity and the two-$Mg^{2+}$ acidic ligand cluster

Direct bioinformatic analysis of the Q88DZ0 sequence corroborates the annotation. **(1)** A canonical Rossmann dinucleotide-binding fingerprint **GXGXXG** is present at residues 23–28 (G23-Y-G25-S-Q-G28) in the N-domain. **(2)** The cofactor-specificity region immediately downstream (β2αB loop) contains basic residues **Arg47 and Lys48** (…GLRKGSAT…, residues 45–52) plus Ser50/Thr52; positively charged residues at these positions are the structural hallmark that coordinates the **2′-phosphate of NADPH**, indicating NADPH (not NADH) preference — consistent with the enzyme's formal name "Ketol-acid reductoisomerase (NADP(+))." **(3)** The C-terminal domain carries a conserved cluster of acidic residues — **Asp190, Glu194, Glu226, Glu230** (all UniProt-annotated binding sites) — forming two Asp/Glu pairs that coordinate the two catalytic $Mg^{2+}$ ions. **(4)** An annotated active-site residue **His107** lies at the N/C-domain interface. This is grounded in the finding that specificity-loop residues determine cofactor preference in class I KARIs: "insertions in the specificity loops that confounded previous attempts to classify them according to loop length" (PMID: 25849365).

---

## Mechanistic Model / Interpretation

### Position in the branched-chain amino acid pathway

KARI (*ilvC*) is the second of four common enzymatic steps that build all three branched-chain amino acids from pyruvate (and, for isoleucine, 2-oxobutanoate). The pathway architecture in *P. putida* KT2440 is:

```
                    pyruvate + pyruvate  (Val/Leu branch)
                    pyruvate + 2-oxobutanoate  (Ile branch)
                              |
                              |  [ Step 1 ]  AHAS  (ilvBN / PP_4680-PP_4679)
                              v
     (S)-2-acetolactate        OR   (S)-2-aceto-2-hydroxybutyrate
                              |
                              |  [ Step 2 ]  KARI  (ilvC / PP_4678)  <-- THIS GENE
                              |     Mg2+-dependent alkyl migration
                              |     + NADPH-dependent reduction
                              v
(2R)-2,3-dihydroxy-isovalerate  OR  (2R,3R)-2,3-dihydroxy-3-methylvalerate
                              |
                              |  [ Step 3 ]  DHAD  (ilvD)  -- dehydration
                              v
     2-ketoisovalerate       OR       2-keto-3-methylvalerate
                  |                                |
                  |  [ Step 4 ] transaminase       | [ Step 4 ] transaminase
                  v                                v
               L-VALINE                      L-ISOLEUCINE
                  |
                  +--> (via ilvD product 2-ketoisovalerate)
                       --> LEUCINE biosynthesis (leuABCD)
                       --> PANTOTHENATE / CoA biosynthesis (panBCD)
```

KARI performs the pivotal chemistry that converts the acetohydroxy-acid intermediates into dihydroxy-acids. Its bifunctionality is remarkable: a single active site accommodates two mechanistically unrelated reactions. The **isomerization** step is an alkyl migration (methyl for the valine branch, ethyl for the isoleucine branch), positioning the substrate for the second reaction; this step requires the **two active-site $Mg^{2+}$ ions** coordinated by the Asp190/Glu194 and Glu226/Glu230 pairs. The **reduction** step then uses hydride from **NADPH** bound in the N-terminal Rossmann domain.

## Structure–function logic

| Feature | Structural element (Q88DZ0) | Functional consequence |
|---|---|---|
| NAD(P)H binding | N-terminal Rossmann domain; GXGXXG at res 23–28 | Binds nicotinamide cofactor for the reduction step |
| NADPH vs NADH selection | β2αB loop basic residues Arg47/ Lys48 | Coordinate 2′-phosphate of NADPH → NADPH preference |
| Two $Mg^{2+}$ ions | Acidic cluster Asp190/Glu194/ Glu226/Glu230 (C-domain) | Catalyze the $Mg^{2+}$-dependent alkyl migration; position substrate |
| Active site | His107 at N/C-domain interface | Catalytic residue at the inter-domain cleft |
| Class I identity | 338 aa, single subunit, KARI_prok signature | Compact bacterial enzyme (no C-domain duplication) |
| Induced fit | N/C-domain interface | Closes on $Mg^{2+}$ + NADPH binding to assemble competent site |

## Localization and pathway integration

KARI is a **soluble cytoplasmic enzyme** — the entire BCAA biosynthetic pathway operates in the cytosol, with no signal peptide or membrane-targeting features. Its genomic clustering with AHAS (*ilvBN*) in KT2440 supports co-regulation and metabolic channeling of substrates directly from AHAS to KARI. Downstream, the shared intermediate 2-ketoisovalerate is a branch point not only for valine/leucine but also for **pantothenate and coenzyme A biosynthesis**, explaining KARI's assignment to pathway ppu00770 in addition to the BCAA pathways.

## Note on catabolism versus biosynthesis

Much of the *P. putida*-specific historical literature (PMIDs 5030618, 4150713, 4150714, 4352175, 10217783, 19910413) concerns branched-chain amino acid **catabolism** (the *bkd* operon, branched-chain keto acid dehydrogenase, transaminases). These pathways **degrade** BCAAs and are distinct from the **biosynthetic** role of *ilvC*. KARI is unambiguously a biosynthetic enzyme; the catabolic literature is contextually relevant to *P. putida* BCAA metabolism but does not describe *ilvC* function directly.

## Evidence Base

| PMID | Title (abbrev.) | How it supports the findings |
|---|---|---|
| 19362563 | *Conformational changes in a plant KARI upon $Mg^{2+}$ and NADPH binding* | Defines exact substrates/products and the two-reaction (alkyl migration + reduction) chemistry requiring $Mg^{2+}$ and NADPH (F001) |
| 23036858 | *Bacterial and plant KARIs have different mechanisms of induced fit* | Establishes KARI as the second enzyme of BCAA biosynthesis; two-domain architecture with interface active site; taxonomic distribution absent in animals (F001, F002, F004) |
| 25849365 | *Cofactor specificity motifs and induced fit in class I KARIs* | Defines class I vs class II; induced-fit active-site closure on NAD(P)H binding; specificity-loop residues determine cofactor preference (F002, F003, F007) |
| 21515217 | *Engineered KARI and ADH enable anaerobic isobutanol at theoretical yield* | Shows NADPH specificity as wild-type default and KARI as engineerable biofuel node (F004) |
| 36354071 | *M. tuberculosis KARI down-regulation affects persistence* | Genetic evidence that KARI is important for bacterial fitness/persistence in vivo (F004) |
| 24493249 | *FgIlv5 required for BCAA biosynthesis and virulence in F. graminearum* | Deletion abolishes growth without Ile+Val; confirms biosynthetic role and phenotype of KARI loss (F004) |
| 25081555 | *2-ketoisovalerate pathway enzymes in R. eutropha* | Confirms the AHAS(ilvBH)→AHAIR(ilvC)→DHAD(ilvD) enzyme order and feedback regulation; contextual for F006 |
| 27600064 | *Poly(2-hydroxyisovalerate-co-lactate) biosynthesis in E. coli* | Uses *ilvCD* alongside *ilvBN* for 2-hydroxyisovalerate production; contextual application of KARI |

Supporting *P. putida* BCAA-catabolism literature (contextual, not describing *ilvC* directly): 5030618, 4150713, 4150714, 4352175, 10217783, 19910413.

### Convergence of evidence

The functional assignment rests on **four independent, converging evidence lines**: 1. **Enzymology/structure** — primary literature defines KARI's reaction, cofactors, and mechanism. 2. **Curated annotation** — UniProt/HAMAP-Rule MF_00435 assigns EC 1.1.1.86, class I identity, sequence length, and catalytic residues specifically to Q88DZ0. 3. **Genomic context** — PP_4678 sits in a canonical *ilvBN-ilvC* cluster in KT2440. 4. **Sequence analysis** — the Rossmann fingerprint, NADPH-specificity loop residues, and two-$Mg^{2+}$ acidic ligand cluster are all identifiable in the Q88DZ0 primary sequence.

---

## Limitations and Knowledge Gaps

1. **No direct experimental characterization of the *P. putida* KT2440 enzyme.** All mechanistic and structural detail is inferred from orthologues (plant, *E. coli*, *M. tuberculosis*, *R. eutropha*, *F. graminearum*) and from bioinformatic/curated annotation. There is no published crystal structure, kinetic study, or knockout phenotype specifically for Q88DZ0/PP_4678.
2. **Cofactor specificity is predicted, not measured.** NADPH preference for Q88DZ0 is inferred from the β2αB loop basic residues (Arg47/Lys48) and the enzyme name. Direct kinetic determination of NADPH vs NADH preference for this specific enzyme has not been done.
3. **Substrate-branch kinetics unknown.** The relative activity toward the valine-branch (2-acetolactate) versus isoleucine-branch (2-aceto-2-hydroxybutyrate) substrate has not been measured for this enzyme; branch partitioning is assumed to follow the general KARI paradigm.
4. **Residue numbering caveats.** Active-site (His107) and metal-ligand residue positions are drawn from UniProt annotation and sequence analysis; they have not been validated by mutagenesis in this organism.
5. **Regulation not experimentally defined.** The genomic clustering with *ilvBN* implies co-regulation, but promoter mapping, transcriptional regulation, and feedback inhibition of KARI in KT2440 specifically remain uncharacterized (though feedback inhibition of KARI/AHAIR by BCAAs is documented in related bacteria such as *R. eutropha* and *C. glutamicum*).

---

## Proposed Follow-up Experiments / Actions

1. **Recombinant expression and steady-state kinetics.** Express Q88DZ0 (His-tagged) in *E. coli*, purify, and determine k_cat/K_M for both branch substrates with NADPH and NADH; confirm the predicted strong NADPH preference and quantify the NADH/NADPH selectivity ratio.
2. **Metal and cofactor dependence.** Measure $Mg^{2+}$ (and $Mn^{2+}$ substitution) dependence and stoichiometry by ITC; verify the two-metal requirement and the roles of Asp190/Glu194/Glu226/Glu230 via alanine mutagenesis.
3. **Structural determination.** Solve the crystal structure (or generate an AlphaFold model followed by validation) of Q88DZ0 with bound $Mg^{2+}$/NADPH to confirm the class I two-domain fold, the interface active site, and induced-fit closure.
4. **Genetic knockout in KT2440.** Delete PP_4678 and test for BCAA auxotrophy (growth rescue by valine + isoleucine), confirming the essential biosynthetic role in situ.
5. **Operon/transcription mapping.** Define the PP_4680–PP_4679–PP_4678 transcript(s) by RT-PCR/RNA-seq and identify the promoter and any BCAA-responsive regulation, testing co-regulation with AHAS.
6. **Metabolic-engineering assessment.** Evaluate Q88DZ0 (wild-type and NADH-swapped variants) as a KARI module for isobutanol / 2-ketoisovalerate-derived product biosynthesis in *P. putida*, exploiting the organism's robust solvent tolerance.

---

## Conclusion

*ilvC* (Q88DZ0, PP_4678) of *Pseudomonas putida* KT2440 encodes **ketol-acid reductoisomerase (KARI/AHAIR, EC 1.1.1.86)**, a soluble cytoplasmic, 338-residue class I enzyme that catalyzes the **second step of branched-chain amino acid biosynthesis**. Using two $Mg^{2+}$ ions and NADPH in a single active site, it performs an alkyl-migration isomerization followed by reduction, converting (S)-2-acetolactate to (2R)-2,3-dihydroxyisovalerate (valine/leucine branch) and (S)-2-aceto-2-hydroxybutyrate to (2R,3R)-2,3-dihydroxy-3-methylvalerate (isoleucine branch). It acts downstream of acetohydroxyacid synthase (*ilvBN*, encoded immediately upstream) and upstream of dihydroxyacid dehydratase (*ilvD*), also feeding

pantothenate/CoA biosynthesis. The assignment is supported by convergent enzymological, curated-annotation, genomic-context, and sequence-analysis evidence, though no experiment has yet been performed on the KT2440 enzyme itself.

---