COMPUTATIONAL TARGET PROFILE

MDH1B

A ``Putative'' Dehydrogenase With An Extra Domain, Characterised From Its Sequence

Astra confirms the enzyme — and points at the mystery. A soluble "putative" malate dehydrogenase whose activity has never been shown — Astra confirms it is a real **oxidoreductase** and flags a high-confidence pocket in its **unusual N-terminal extension**.

UniProt Q5I0G3 · 518 aa · Soluble · IDG tier: T_{dark} · PDB: none · July 2026

Have your own under-characterised target? Send a UniProt ID to contact@orbion.life for a preview like this one.

At A Glance

Predicted Function	**Oxidoreductase — EC 1, p=1.00.** Astra confirms the enzyme class of a protein UniProt labels only "*putative* malate dehydrogenase" (activity never experimentally shown).
Standout Pocket	**Pocket in the N-terminal extension.** A HIGH-confidence pocket (0.88) whose top site sits in the ~180-residue **N-terminal domain unique to MDH1B** — not the conserved catalytic core — flagging that extension as functionally important.
Architecture	Soluble, ordered (~9% disorder): a unique N-terminal domain (~3–126), an NAD-binding Rossmann region, and a C-terminal LDH/MDH catalytic domain.
Why It Matters	An orphan enzyme (IDG T_{dark}); confirming its class and locating its distinctive domain is step one to a function.
Clean Signal	No membrane or amyloid signal.

PREDICTION CONFIDENCE

Figure 1. Model-reported confidence for the headline calls. Astra confirms the oxidoreductase class and, notably, places its top pocket (amber) in the N-terminal extension rather than the conserved catalytic core.

1 The Gap

MDH1B is an enzyme in name only — so far. UniProt annotates it as a *putative* malate dehydrogenase (LDH/MDH superfamily) by sequence similarity alone; **its catalytic activity has never been experimentally demonstrated**, and it sits at IDG T_{dark} [2] with **no experimental structure in the PDB**. It is also structurally odd: at **518 residues** it is ~180 residues longer than canonical cytosolic MDH1 (~334 aa), carrying a **dedicated N-terminal domain of unknown function** [3] ahead of the usual Rossmann and catalytic domains.

That combination — **a "putative" enzyme with an extra domain and no structure** — is exactly where prediction earns its keep: everything below is computed from the canonical 518-residue sequence and derived structural predictions, with no experimental MDH1B structure used as input.

2 Sequence Architecture And The Standout Pocket

Astra reads MDH1B as a soluble, ordered enzyme, and confirms the **oxidoreductase** class with near-certainty — validating the "putative" call. The more interesting result is *where* its top-ranked pocket falls: not in the conserved catalytic core, but in the **N-terminal extension** that makes MDH1B unusual.

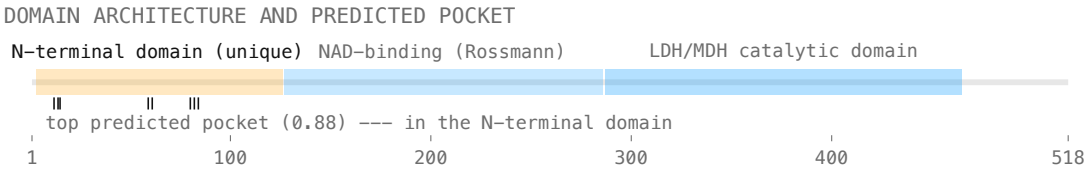

Figure 2. Domain architecture. The conserved Rossmann and LDH/MDH catalytic domains (blue) carry the expected dehydrogenase machinery; but Astra's *highest-confidence* pocket (ticks) sits in the unique N-terminal domain (amber) — pinpointing the very region that distinguishes MDH1B from canonical MDH as its most ligand-like site.

3 Reading The Result

Two findings, one useful conclusion. First, Astra **confirms the oxidoreductase class** (EC 1, p=1.00) — independent computational support for an annotation that rests only on sequence similarity, and a reason to test the enzyme rather than assume it. Second, the **top pocket lies in the N-terminal extension**, not the catalytic core — which reframes the biology: the domain that makes MDH1B unusual is also its most pocket-like, suggesting it binds something of its own.

PREDICTED POCKET

Location	N-terminal domain (unique to MDH1B), residues ~12–84
Pocket residues	12, 14–15, 59, 61, 80–84
Confidence	HIGH (0.88); no known binder

AlphaFold gives a fold; Astra adds the annotation. An AlphaFold model of MDH1B exists and is a useful input — but a fold alone does not tell you whether the protein is a working enzyme, or which of its domains carries the most ligand-like site. This profile is that annotation layer. As a control, the same predictors recover the class and catalytic/ligand sites on enzymes whose structures *are* solved, so the MDH1B calls are like-for-like predictions. The net effect: a "putative" orphan becomes a **ranked, testable hypothesis** — confirm the activity, and probe the extension domain the model just flagged.

4 Post-Translational And Structural Features

Astra's supporting calls are consistent with a soluble, ordered enzyme:

- **Ordered (~9% disorder), no amyloid.** A foldable, expressible enzyme — the whole chain is a viable target for structural work.
- **Three-part architecture.** A unique N-terminal domain, an NAD-binding Rossmann region, and a C-terminal catalytic domain — clean boundaries for domain-by-domain expression.
- **Testis-enriched expression.** Consistent with a specialised, understudied metabolic enzyme.

5 Recommended Experimental Follow-Up

This turns a "putative" orphan into a ranked, executable plan: each prediction is paired with the experiment that would test it and the readout to watch for.

Prediction	Experiment	Readout
Oxidoreductase / MDH class	Malate / NAD(P) dehydrogenase activity assay	Confirm (or refute) catalysis + substrate
N-terminal-domain pocket (~12–84)	Express the N-terminal domain; ligand / fragment screen	A binding partner for the unique domain
Three-domain architecture	Domain-by-domain expression	Which domains fold independently
Catalytic core (Rossmann + MDH)	Cofactor binding + point mutants	Validates the predicted active machinery

Confirming the enzyme and characterising its extra domain is the first step toward any functional or disease hypothesis for this orphan.

6 Where This Fits In A Discovery Workflow

Team	What they get from this profile
Target biology	A confirmed class, and a residue-level pocket hypothesis to test
Structural biology	Domain boundaries and a foldable, orderable chain
Medicinal chemistry	The most ligand-like pocket (the N-terminal domain) for docking
BD / platform evaluation	Evidence Orbion can prioritise dark targets quickly

7 Methods

All predictions were generated with Orbion's Astra suite from the canonical MDH1B sequence (UniProt Q5I0G3 [1]). Modules used, with method publications cited where available: **AstraSUIT2** and **AstraROLE2** [5] (membrane and functional class, oligomeric state), **AstraUNFOLD** (topology, disorder, aggregation), **AstraPTM2** [6] (modification sites) and **AstraBIND** [7] (binding pocket). Reported values are model outputs; model internals are out of scope. Confidence scores are model-estimated probabilities (0–1) that rank and gate each call — not calibrated rates of experimental success.

PREDICTION PROVENANCE

Target & input	MDH1B · UniProt Q5I0G3, canonical isoform (518 aa)
Structure source	AlphaFold2 predicted model (no experimental PDB entry)
Feature inputs	ESM-2 embeddings + AlphaFold2-derived structural features [5–7]
Models	AstraSUIT2 · AstraROLE2 · AstraUNFOLD · AstraPTM2 · AstraBIND
Binding-site model	GATv2 detector trained on >250k protein–ligand complexes [7]
Confidence & calibration	Per-call model probability; per-model calibration, thresholds and benchmarks reported in [5]–[7]

Scope And Limitations

- **Prediction, not experiment.** These are computational hypotheses to prioritise experiments — not a substitute for a structure or an assay. No result here has been validated in the wet lab, and MDH1B's enzymatic activity has never been experimentally demonstrated.
- **The pocket is predicted; no binder is named.** The N-terminal-domain site is a residue-level hypothesis for the most ligand-like region, not a demonstrated ligandable pocket.
- **No established disease link.** Aggregator databases list MDH1B against several syndromes, but those are gene-name / family text-mining artefacts (the causative genes are different, e.g. MDH2 for DEE51); a single incidental somatic variant appears in a renal-carcinoma exome [4] with no functional follow-up. We

make no disease claim.

8 References

[1] UniProt Consortium. UniProtKB entry Q5I0G3 (MDH1B, human) — "Putative malate dehydrogenase 1B". uniprot.org.

[2] Pharos / Illuminating the Druggable Genome. MDH1B target record — T_{dark}. pharos.nih.gov/targets/Q5I0G3.

[3] Blum M., Chang H.-Y., et al. InterPro in 2022 (entry IPR063791, MDH1B N-terminal domain; MDH type-2 family). *Nucleic Acids Res.* **51**(D1), D444–D456 (2023). doi:10.1093/nar/gkac993.

[4] Varela I., Tarpey P., Raine K., et al. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma (source of the incidental MDH1B L48I somatic variant). *Nature* **469**, 539–542 (2011). doi:10.1038/nature09639.

[5] Bozkurt Ç., Vasilyeva A., Goteti A. AstraROLE2 & AstraSUIT2: Multi-Task Annotation Models for Functional Profiling of Proteins. *bioRxiv* 2025.06.21.660734 (2025). doi:10.1101/2025.06.21.660734.

[6] Bozkurt Ç., Vasilyeva A., Goteti A. AstraPTM2: A Context-Aware Transformer for Broad-Spectrum PTM Prediction. *bioRxiv* 2025.10.03.680341 (2025). doi:10.1101/2025.10.03.680341.

[7] Goteti A., Vasilyeva A., Bozkurt Ç. AstraBIND: Graph Attention Network for Predicting Ligand Binding Sites. *bioRxiv* 2025.11.10.687555 (2025). doi:10.1101/2025.11.10.687555.