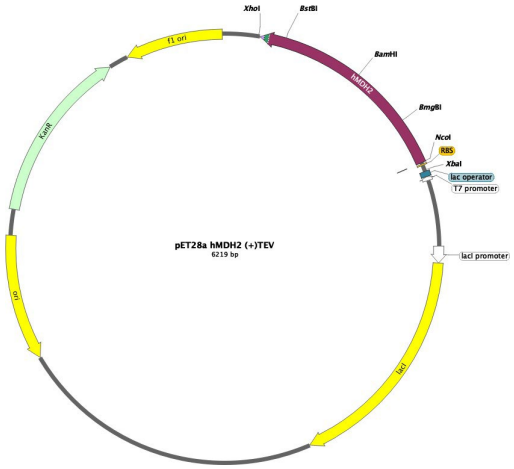
Human Mitochondrial MDH Protein/Clone Information Sheet hMDH2 splice variant 1

Protein Name: Mitochondrial Human Malate Dehydrogenase 2 (hMDH2) transcript variant 1
Organism: Homo sapiens (human) MDH2 **Plasmid Name**: pET28a hMDH (+)TEV aka hMDH2 and hMDH2V1

Clone/Plasmid History: Human Malate dehydrogenase 2 gene was synthesized after codon-optimization of human MDH2 transcript variant 1 without the first 24 amino acids (mito targeting sequence) and adding a Met before endogenous aa 25 for expression in BL21 (DE3) and cloned into pET28 vector using a NcoI/XhoI digested pET28a. Created June 13, 2017. The TEV recognition site was also added between the His tag and MDH. Both the TEV and the His tag are C terminus of hMDH1. The N terminus remains unaltered. *Because the gene is synthesized and codon optimize, the nucleotide sequence will not match the accession number*. Please refer to the linked snapgene file or FASTA formatted file linked below for the DNA sequence of the coding region for hMDH2.

NCBI / Gene Accession Number: https://www.ncbi.nlm.nih.gov/gene/4191 (see note above – this is the endogenous human sequence NOT the sequence found in this vector). hMDH2 variant 1 synthesized coding region

SnapGene Plasmid Map: *Downloadable file includes*:
Resistance, Promotor (for bacterial or mammalian), Sequencing primers, RBS and Kozak sequence, History of cloning, Annotated start and stop of protein, Highlighted tags or TEV/Thrombin sites



NCBI Protein Sequence Accession: CAG38785.1 (hMDH2 with the targeting sequence, not included in this construct)

UniProt Protein Page: P40926 (MDHM_HUMAN) (The transcript variant 1 is considered the canonical isoform. This is the mature clone without the transit peptide aa 1-24.) Variant 2 is missing 41 amino acids (144-185) as a result of alternative splicing. COUNTING STARTS AFTER TRANSIT WITH MET.

RCSB PDB Page: 2DFD (hMDH2 mature): includes structures without and with several ligands

Interesting Publications:

Shen Q, Teng L, Wang Y, Guo L, Xu F, Huang H, Xie W, Zhou Q, Chen Y, Wang J, Mao Y, Chen J, Jiang H. Integrated genomic, transcriptomic and metabolomic analysis reveals MDH2 mutation-induced metabolic disorder in recurrent focal segmental glomerulosclerosis. Front Immunol. 2022 Sep 8;13:962986. doi: 10.3389/fimmu.2022.962986. PMID: 36159820; PMCID: PMC9495259.

Ticci C, Nesti C, Rubegni A, Doccini S, Baldacci J, Dal Canto F, Ragni L, Cordelli DM, Donati MA, Santorelli FM. Bi-allelic variants in MDH2: Expanding the clinical phenotype. Clin Genet. 2022 Feb;101(2):260-264. doi: 10.1111/cge.14088. Epub 2021 Nov 22. PMID: 34766628.

Jungtrakoon Thamtarana P, Marucci A, Pannone L, Bonnefond A, Pezzilli S, Biagini T, Buranasupkajorn P, Hastings T, Mendonca C, Marselli L, Di Paola R, Abubakar Z, Mercuri L, Alberico F, Flex E, Ceròn J, Porta-de-la-Riva M, Ludovico O, Carella M, Martinelli S, Marchetti P, Mazza T, Froguel P, Trischitta V, Doria A, Prudente S. Gain of Function of Malate Dehydrogenase 2 and Familial Hyperglycemia. J Clin Endocrinol Metab. 2022 Feb 17;107(3):668-684. doi: 10.1210/clinem/dgab790. PMID: 34718610; PMCID: PMC8852227.

Available Mutations: None at this time, will become available upon publication.

Protein Notes: This is variant 1 of the hMDH 2(mitochondrial) isoform. The version included in this construct does NOT include the mitochondrial targeting sequence (mlsalvrpvsaalrrsfstsaqnn). A Met has been added to the mature hMDH2V1 found in the mitochondria without the targeting sequence. Isoform 2 does not include 41 internal amino acids not present on MD21 isoform

1. hMDH2V1 is a 361 amino acid (with the additional glycine in the N term to maintain reading frame and both a TEV protease recognition site and a 6X His tag placed on the C terminus of hMDH2). Human MDH2V1 has a predicted mw = 34.8 kDa. hMDH2 is reported to be a dimer .

Human Mitochondrial MDH construct:

pI = 8.33/ ε_{280} , = 0.257 (assuming reduced disulphides) or 0.271 (assuming disulphides intact) $mL.mg^{-1}.cm^{-1}$ extinction coefficient (280 nm: calculated using ProtParam.)

Key amino acids / functions studied include

Residue	Flexible Loop	Aspartate	Arginine	Arginine	Aspartate	Arginine	Histidine
Watermelon-g MDH 1sev/1smk Equivalent	117-140	D77	R124	R130	D193	R196	H220
hCytosolic 7rm9.pdb or 7rm9repaired.pdb	85-108	D42	R92	R98	D159	R162	H187
hCytoConstruct.pdb		D61	R111	R117	D178	R1181	H226
hMitochondrial 2DFD.pdb	79-102	D39	R86	R92	D155	R158	H182
hMitoConstruct.pdb		D34	R81	R87	D150	R153	H177
Plasmodium falciparum: 5NFR.pdb	74-97	D32	R81	R87	D147	R150	H174
Ignicoccus Islandicus 6qss,pdb	77-100	D37	R86	R92	D151	R154	H178
Function	Closes over active site on substrate binding	Governs specificity for NAD(H)	Malate/Oxaloacetate/Citrate Binding	Malate/Oxaloacetate/Citrate Binding	Alters Basicity of Catalytic Histidine	Malate/Oxaloacetate/Citrate Binding	Catalytic Base

Clone FAQ and Important Points: Strong protein expression at 37°C 1mm IPTG for 3-4 hour induction. Stronger expression at 20°C (room temp) for 14-24 hrs. ~0.5 mg or more per ml of culture. Stronger expression at 20°C (room temp) for 14-24 hrs. pET28a (Novagen) is a low copy plasmid (~40) and will not give high yields of DNA preps. Kan Resistant. Do not freeze thaw purified protein. Purification easily performed in column or batch format. Concentrations approaching 1-1.25 mg/ml will precipitate over a short time. Dilute immediately after purification and before dyalysis to 1 mg/ml or less. See MDH Stability Datasheet for more information. Stable at 4° for 6-8 weeks dialyzed against (10 mM K phosphate, 0.1 mM EDTA, 20% glycerol, pH 8.0). Long term storage -20 to -80°C (10% Glycerol, 50 mM NaCl, 1 mM β-ME in 10 mM K phosphate, pH 8.0). See MDH Stability Datasheet for more information.

See Snap Gene File for details.

Construct Amino Acid Coding

Sequence:

MAKVAVLGASGGIGQPLSLLLKNSPLVSRLTLYDIAHTPGVAADLSHIETKAAVKGYLGPEQLPDCLKGCDVVVIPAGVP
RKPGMTRDDLFNTNATIVATLTAACAQHCPEAMICVIANPVNSTIPITAEVFKKHGVYNPNKIFGVTTLDIVRANTFVAE
LKGLDPARVNVPVIGGHAGKTIIPLISQCTPKVDFPQDQLTALTGRIQEAGTEVVKAKAGAGSATLSMAYAGARFVFSLV
DAMNGKEGVVECSFVKSQETECTYFSTPLLLGKKGIEKNLGIGKVSSFEEKMISDAIPELKASIKKGEDFVKTLKENLYF QGHHHHHH

```
2DFD.pdb Amino Acid Sequence
MHHHHHHSSGVDLGTENLYFQSMSAQNNAKVAVLGASGGIGQPLSLLLKNSPLVSRLTLYDIAHTPGVAA
DLSHIETKAAVKGYLGPEQLPDCLKGCDVVVIPAGVPRKPGMTRDDLFNTNATIVATLTAACAQHCPEAM
ICVIANPVNSTIPITAEVFKKHGVYNPNKIFGVTTLDIVRANTFVAELKGLDPARVNVPVIGGHAGKTII
PLISQCTPKVDFPQDQLTALTGRIQEAGTEVVKAKAGAGSATLSMAYAGARFVFSLVDAMNGKEGVVECS
FVKSQETECTYFSTPLLLGKKGIEKNLGIGKVSSFEEKMISDAIPELKASIKKGEDFVKTLK
```

```
hMitoconstructAmino     ----------------------------MAKVAVLGASGGIGQPLSLLLKNSPLVSRLTLY   33
2DFD.pdb                MHHHHHHSSGVDLGTENLYFQSMSAQNNAKVAVLGASGGIGQPLSLLLKNSPLVSRLTLY   60
                                                     ********************************

hMitoconstructAmino     DIAHTPGVAADLSHIETKAAVKGYLGPEQLPDCLKGCDVVVIPAGVPRKPGMTRDDLFNT   93
2DFD.pdb                DIAHTPGVAADLSHIETKAAVKGYLGPEQLPDCLKGCDVVVIPAGVPRKPGMTRDDLFNT   120
                        ************************************************************

hMitoconstructAmino     NATIVATLTAACAQHCPEAMICVIANPVNSTIPITAEVFKKHGVYNPNKIFGVTTLDIVR   153
2DFD.pdb                NATIVATLTAACAQHCPEAMICVIANPVNSTIPITAEVFKKHGVYNPNKIFGVTTLDIVR   180
                        ************************************************************

hMitoconstructAmino     ANTFVAELKGLDPARVNVPVIGGHAGKTIIPLISQCTPKVDFPQDQLTALTGRIQEAGTE   213
2DFD.pdb                ANTFVAELKGLDPARVNVPVIGGHAGKTIIPLISQCTPKVDFPQDQLTALTGRIQEAGTE   240
                        ************************************************************

hMitoconstructAmino     VVKAKAGAGSATLSMAYAGARFVFSLVDAMNGKEGVVECSFVKSQETECTYFSTPLLLGK   273
2DFD.pdb                VVKAKAGAGSATLSMAYAGARFVFSLVDAMNGKEGVVECSFVKSQETECTYFSTPLLLGK   300
                        ************************************************************

hMitoconstructAmino     KGIEKNLGIGKVSSFEEKMISDAIPELKASIKKGEDFVKTLKENLYFQGHHHHHH        328
2DFD.pdb                KGIEKNLGIGKVSSFEEKMISDAIPELKASIKKGEDFVKTLK-------------        342
                        ****************************************
```

Coding Plasmid Sequence:

TTAGTGATGGTGGTGATGATGACCCTGGAAGTACAGGTTCTCTTTCAGGGTCTTCACGAAATCTTCACCCTTCTTGATGC
TCGCTTTCAGTTCCGGAATTGCGTCGGAGATCATCTTCTCTTCGAAAGAAGAAACTTTACCGATGCCCAGGTTCTTCTCG
ATGCCTTTCTTGCCCAGCAGCAGTGGAGTGCTGAAGTAGGTGCATTCGGTTTCTTGAGACTTCACGAAGGAACATTCCAC
TACACCTTCCTTGCCGTTCATTGCGTCAACCAGAGAGAATACGAAACGAGCACCAGCGTAGGCCATGCTCAGAGTCGCAG
AGCCAGCACCTGCTTTCGCTTTCACCACTTCAGTGCCTGCCTCCTGGATACGACCAGTCAGAGCGGTCAGCTGGTCCTGT
GGGAAGTCTACTTTCGGAGTGCACTGGGAAATCAGCGGAATGATGGTCTTGCCAGCGTGACCACCGATTACCGGAACGTT
TACGCGAGCTGGATCCAGACCTTTCAGTTCAGCCACAAAGGTGTTTGCACGCACGATATCCAGGGTGGTAACACCGAAGA
TCTTGTTCGGGTTGTAAACACCGTGTTTCTTGAATACCTCGGCAGTAATCGGGATGGTGCTGTTAACCGGATTAGCAATC
ACACAGATCATAGCTTCCGGACAGTGCTGAGCGCAAGCAGCGGTCAGGGTCGCAACGATAGTCGCATTAGTATTGAACAG
GTCATCACGGGTCATACCTGGTTTACGTGGAACACCTGCTGGGATAACTACCACGTCACAACCTTTCAGACAATCTGGCA
GCTGTTCTGGACCCAGGTAGCCTTTCACCGCAGCTTTAGTCTCAATGTGGGACAGGTCAGCAGCAACGCCTGGAGTATGT
GCGATATCGTACAGGGTCAGACGGCTAACCAGTGGAGAGTTCTTCAGCAGCAGGGACAGCGGTTGACCAATACCACCAGA
TGCACCCAGAACAGCTACTTTAGCCAT