Structures of *Medicago truncatula* L-Histidinol Dehydrogenase Show Rearrangements Required for NAD\(^+\) Binding and the Cofactor Positioned to Accept a Hydride

Milosz Ruszkowski & Zbigniew Dauter

Plants, lower eukaryotes, bacteria, and archaeabacteria synthesise L-histidine (His) in a similar, multistep pathway that is absent in mammals. This makes the His biosynthetic route a promising target for herbicides, antifungal agents, and antibiotics. The last enzyme of the pathway, bifunctional L-histidinol dehydrogenase (HDH, EC 1.1.1.23), catalyses two oxidation reactions: from L-histidinol (HOL) to L-histidinaldehyde and from L-histidinaldehyde to His. Over the course of the reaction, HDH utilises two molecules of NAD\(^+\) as the hydride acceptor. The object of this study was the HDH enzyme from the model legume plant, *Medicago truncatula* (MtHDH). Three crystal structures complexed with imidazole, HOL, and His with NAD\(^+\) provided in-depth insights into the enzyme architecture, its active site, and the cofactor binding mode. The overall structure of MtHDH is similar to the two bacterial orthologues whose three-dimensional structures have been determined. The three snapshots, with the MtHDH enzyme captured in different states, visualise structural rearrangements that allow for NAD\(^+\) binding for the first time. Furthermore, the MtHDH complex with His and NAD\(^+\) displays the cofactor molecule situated in a way that would allow for a hydride transfer.

L-Histidine (His) biosynthesis is a part of the primary anabolism of archaeabacteria, bacteria, lower eukaryotes, and plants. In a multistep pathway, His is synthesised from 5-phosphoribosyl-1-pyrophosphate and ATP, which links amino acid and nucleotide metabolism\(^1,2\). The absence of the His-biosynthetic pathway in mammals makes it a good target for antibiotics, herbicides, and antifungal agents\(^3-5\). In multiple examples, the His-auxotrophic mutants, lacking one or more of the His-synthesising enzymes, are unable to tolerate His starvation. For instance, in *Mycobacterium tuberculosis* and *Brucella suis*, His-synthesising enzymes are essential for survival of those bacterial species, responsible for very severe infections, tuberculosis and brucellosis, respectively\(^6,7\). Among all enzymes from the His-biosynthetic pathway, the ultimate catalyst, L-histidinol dehydrogenase (HDH), deserves special attention\(^8\). Blocking the last step of a biosynthetic pathway not only hinders the final product but also results in the accumulation of intermediate metabolites, which may increase the toxic effect resulting from the lack of the pathway end-product alone. In fact, HDH has been prioritised among the top 50 targets against totally drug-resistant strains of *Mycobacterium tuberculosis*\(^9\).

HDH (EC 1.1.1.23) is a bifunctional enzyme that oxidises L-histidinol (HOL) via L-histidinaldehyde (HAL) into His in two sequential reactions, each coupled with a reduction of one nicotinamide adenine dinucleotide (NAD\(^+\)) molecule to NADH (Fig. 1)\(^10\). The enzyme operates via the Bi-Uni-Uni-Bi Ping Pong mechanism, and, rather than leaving the catalytic site, the intermediate HAL is immediately converted to His\(^10-13\). It is also known that NAD\(^+\) binding is effective only to HOL- or HAL-complexed HDHs\(^11,12\), which is uncommon for NAD-dependent dehydrogenases. In most cases, the cofactor binds first. HDHs are metalloenzymes, binding a single Zn\(^{2+}\) cation per subunit\(^14,15\). However, Zn\(^{2+}\) is not directly involved in catalysis. Instead, it plays a structural role by facilitating and stabilising the binding of HOL, HAL, and His\(^16\).

Synchrotron Radiation Research Section of MCL, National Cancer Institute, Argonne, IL, USA. Correspondence and requests for materials should be addressed to M.R. (email: milosz.ruszkowski@nih.gov)
Structural information is crucial for a rationale-based drug design. To date, there have been only two publications presenting crystal structures of HDH enzymes, both from bacterial sources. The enzyme from E. coli (EcHisD) (in this work, we use the HisD acronym for bacterial HDHs, which relates to his operon organisation in many prokaryotic species) was studied by Barbosa et al. The authors presented four structures: (i) unliganded EcHisD (Protein Data Bank, PDB ID: 1k75) and complexes with (ii) Zn\(^{2+}\), HOL, and NAD\(^+\) (PDB ID: 1kae), and (iii) Zn\(^{2+}\) and L-histidine (1kar) and (iv) Zn\(^{2+}\) and His (1kah). This set of structures revealed, for the first time, both the overall fold of an HDH enzyme and the location and composition of the active site. Unfortunately, as the authors acknowledged, the enzyme in the structure with HOL and NAD\(^+\) had been trapped in a state that would not allow the hydride transfer to occur, as the HOL carbon that is to be oxidised and the C4 atom of NAD\(^+\) are too distant (approximately 4.9 Å from each other). It is worth noting that the distance between the hydride donor and the acceptor should be approximately 3 Å and has been calculated to be 2.7 Å for the transition state. Because the distance in the EcHisD complex was too long, the substrate (HOL) was found in the active site instead of the product (His). The authors attributed this situation to the crystal packing, which prevented small rearrangements that would otherwise allow NAD\(^+\) to approach and oxidise HOL. In other words, soaking EcHisD crystals with HOL and NAD\(^+\) did not drive the necessary conformational changes, as the protein chain had been locked by the structure of the crystal lattice.

Two other structures of HDH, from B. suis (BsHisD), show the enzyme unliganded (PDB ID: 4g07) and with a very potent inhibitor, (3S)-3-amino-4-(1H-imidazol-5-yl)-1-[4-(phenylmethoxy)phenyl]-butan-2-one (4g09). Two other reports that deal with HDH structures are based on computational modelling. In both cases, as the template for molecular modelling, the EcHisD complex with Zn\(^{2+}\), HOL, and NAD\(^+\) [PDB ID: 1kae, ref. 16] was used in the locked state. Interestingly, the compounds that exhibited the highest inhibition of HDH were predicted to simultaneously occupy the substrate binding site as well as a part of the NAD\(^+\) binding cleft.

Since bacterial HDHs are perceived as very promising targets for antibiotics, a parallel approach seems reasonable for seeking novel herbicides. Studies of plant His-synthesizing pathways not only expand the knowledge about this fragment of plant metabolism but are also needed to deduce the differences and similarities between the prokaryotic and eukaryotic routes.

Overall, the plant pathway leading to His biosynthesis is very similar to the bacterial pathway. Most of the enzymes that take part in plant His biosynthesis are encoded by single genes, unlike for many other plant metabolic pathways. All enzymes required for His biosynthesis in plants are located in chloroplasts and have been identified: HisN1, HisN2, HisN3, HisN4, HisN5, HisN6, HisN7, and HisN8. However, structural knowledge of plant His-synthesising enzymes is very scarce, with only the structures of Arabidopsis thaliana HisN5, HisN7, and Medicago truncatula HisN8 reported to date.

This article is dedicated to structural studies of HisN8, a HDH from the model legume plant M. truncatula, which from now on will be referred to as MhHDH. This is the first report of an HDH structure from a plant source. We show three crystal structures of Zn\(^{2+}\)-bound MhHDH complexed with: (i) imidazole (IMD); (ii) HOL; and (iii) His with NAD\(^+\). These structures, representing snapshots over the course of the reaction, expand the knowledge of HDH enzymes by showing the movement of the loops triggered by HOL binding as well as revealing the location of NAD\(^+\) with C4, the hydride-accepting atom within a reacting range to the HOL carbon that is oxidised. Additionally, because HDH enzymes from plants and bacteria are similar, many of the results presented herein should concern prokaryotic HDHs in addition to the orthologs of plant origin. It is important to note that our structures represent the active state of the enzyme, which may serve as an updated model for designing HDH-targeted herbicides and antibiotics.

Results and Discussion

General properties of the MhHDH structure. UniProt search revealed two HDH enzymes in the M. truncatula proteome [B7FNC7 and G7IKX3]. The only difference between the two sequences is an Asp79Asn mutation that lies outside the PCR primer-banding region of the MhHDH open reading frame (ORF). DNA sequencing verified that the amplified ORF used in this study corresponds to the UniProt G7IKX3 entry without the predicted chloroplast-targeting signal peptide.

The sequences of plant HDHs are highly homologous. Even between the plant and bacteria kingdoms, the similarities are significant (Fig. 2). For instance, identities/similarities between the Medicago enzyme and orthologs from A. thaliana, B. oleracea, and E. coli are 79%/88%, 78%/87%, and 46%/59%, respectively.
Figure 2. Sequence alignment of HDH enzymes from various sources: MtHDH, Medicago truncatula [UniProt accession number: G7IKX3]; AtHDH, Arabidopsis thaliana [Q9C5U8-1]; BoHDH, Brassica oleracea [P24226]; EcHisD, Escherichia coli (strain K12) [P06988]; MtHisD, Mycobacterium tuberculosis (strain H37Rv) [P9WNW9]; StHisD, Salmonella typhimurium (strain LT2) [P10370]; BsHisD, Brucella suis biovar 1 (strain 1330) [Q8G2R2]. For clarity of the alignment, the N-terminal signal peptides of At and Bo enzymes were truncated at positions 31 and 32, respectively, which correspond to Pro42 in Mt structure, which is the first residue visible in the electron density. Domains of MtHDH are coloured: I, orange; II, magenta; III, turquoise; IV, chartreuse. Elements of the secondary structure are: α-helices, yellow; 310 helices, red; β-strands, cyan. Helices are numbered consecutively, regardless of the type.
MtHDH crystallises in the \( P2_1 \) space group. The three structures reported herein are isomorphous, with three dimers (AB, CD, and EF) forming the asymmetric unit. The dimeric quaternary structure agrees with the results of size-exclusion chromatography (not shown) as well as the oligomeric state reported for other HDH enzymes. The MtHDH homodimeric assembly (Fig. 3) measures approximately 90 × 70 × 50 Å. Based on PISA calculations, the inter-subunit interface is above 5,200 Å\(^2\) (per monomer), which is approximately 25% of the total solvent-accessible area. Along with numerous hydrophobic contacts, 25 salt bridges and at least 90 hydrogen bonds stabilise a MtHDH dimer. For clarity, the following structural analyses are based on the A subunits (of the three structures), unless noted otherwise.

The MtHDH monomer is composed of four domains. Domain I (residues 66–138, 160–275, and 422–427) contains a core with a Rossmann-fold-like super-secondary structure. The six-stranded β-sheet (ordered 2–6–5–4–7–8), which lies in the centre of domain I, is mostly parallel except for the antiparallel strand 2 at the edge of the β-sheet. The β-sheet is sandwiched between helices \( \alpha 7 \) and \( \alpha 8 \) from one side and \( \alpha 9 \) and \( \alpha 10 \) from the other. The Rossmann-fold core of domain I is surrounded by helix \( \alpha 6 \), which is interacting with \( \alpha 7 \) and \( \alpha 8 \), and by a V-shaped pair, \( \alpha 3 \) and \( \alpha 4 \), that shields \( \alpha 9 \) and \( \alpha 10 \). Additionally, a short 310 helix \( \eta 18 \) is formed at the interface between domains I and III.

Domain II (residues 36–65 and 276–421), also adopts the Rossmann-like fold in its core, despite a very different sequence. The six-stranded, parallel β-sheet (ordered 1–11–10–9–12–13) is surrounded by helices \( \alpha 11, \alpha 12, \alpha 13, \) and \( \alpha 16 \) on one side and \( \eta 1, \alpha 2, \alpha 14, \eta 15, \) and \( \eta 17 \) on the other. Domains I and II together form an extended globular structure with a cleft at the interface.

**Figure 3.** Structure of the MtHDH homodimer. Chain A, cyan, in pipes-and-planks representation and chain B, yellow, with the protein surface are shown. NAD\(^+\) and L-histidinol (HOL) from superposed structures show locations of the active sites, one per monomer. White-on-black denotation of domains corresponds to the subunit B.
In the \( \textit{M}. \text{HDH} \) dimer, the \( \beta \)-sheet of domain II is extended from the edge of the \( \beta_{13} \) strand by two strands belonging to domain III (residues 139–159 and 428–437): \( \beta_{14} \) (parallel) and \( \beta_{3} \) (anti-parallel) of the second protein subunit. Domains III and IV (residues 438–478) are almost perpendicular to each other. They form an L-shaped structure that is mutually swapped between the subunits and occupies a cleft between domains I and II of the dimer-mate subunit. The C-terminal domain IV (helices \( \alpha_{19} \) and \( \alpha_{20} \)) is V-shaped and interacts mostly with domain II of the other subunit, completing the active site of its dimer-mate (see below).

Complex with imidazole illustrates a state before the reaction. Each monomer of \( \textit{M}. \text{HDH} \) binds a single \( \text{Zn}^{2+} \) cation in a deep pocket between domains I, II, and IV* (asterisk indicates an element of the other protein subunit of the dimer), which is the active site of HDH enzymes\(^{15} \) (Fig. 4). In the structure with IMD, \( \text{Zn}^{2+} \) is octahedrally coordinated by two water molecules (Wat1 and Wat2), \( \text{N} \) of His302, \( \text{O} \delta \) of Asp401, \( \text{N} \varepsilon \) of His460*, and \( \text{N} \) of IMD.

IMD has been known as a weak, competitive inhibitor of HisD from \( \textit{S. typhimurium} \), with \( K_i = 1.5 \text{mM} \)\(^{14} \). In the \( \textit{M}. \text{HDH} \) complex, IMD creates two polar interactions, one with \( \text{Zn}^{2+} \) and one with Glu455*, which bind indirectly to domain II and directly to domain IV* (Fig. 4). IMD faces domain I but does not interact with it. It is possible that, for stability of the dimer, there must be a link between \( \text{Zn}^{2+} \) and Glu455*, as we failed to obtain a structure without IMD (or an IMD-derivative) added to the crystallisation solution.

Binding of L-Histidinol drives structural rearrangements. HOL coordinates \( \text{Zn}^{2+} \) with \( \text{N} \) and \( \text{N} \delta \) atoms. More precisely, the N atom occupies the same position as Wat1 in the \( \textit{M}. \text{HDH}/\text{IMD} \) complex, whereas the

Figure 4. The active site of \( \textit{M}. \text{HDH} \). Chain A is coloured cyan whereas chain B is yellow. \( \text{Zn}^{2+} \) coordination and imidazole (IMD) binding is shown in the A panel. The panel B depicts binding of the substrate, histidinol (HOL, ball-and-stick, green mesh is the omit \( F_o - F_c \) electron density map, contoured at 7\( \sigma \) level). Gln299 from superposed \( \textit{M}. \text{HDH}/\text{IMD} \) on \( \textit{M}. \text{HDH}/\text{HOL} \) complex is shown as thin black sticks in the panel B for comparison.
imidazole moiety of HOL is bound in a manner similar to IMD (Fig. 4B). Like IMD, HOL also forms an H-bond with Glu455*. The O atom of HOL interacts with carbonyl O of His408 and Nε of His368.

It has been noted that the crystals were grown in acidic conditions (pH 5.2) in which the imidazole ring should be predominantly double-protonated. Based on the crystallisation screening, acidic conditions are necessary for MtHDH to pack into a crystal lattice. This would greatly disfavour HOL binding. However, we observed that the best crystals were grown when MtHDH was incubated with the ligands at pH 7.5 for 24–48 hours before the crystallisation was set up, which permits the single-protonated HOL to bind to the Zn2+ occupied site before crystallisation.

Upon HOL binding, a significant change takes place in the coordination sphere of Zn2+. The Oε of Gln299 substitutes Wat2, meaning that, in the HOL complex, the metal is coordinated only by the protein and the substrate atoms (Figs 4 and 5). It has been known that NAD+ does not effectively bind to HDH enzymes in the absence of HOL, but the structural reason for that remained elusive. Based on the MtHDH structures reported herein, we now have a better understanding of that feature. A close examination of the aligned structures allowed us to detect several differences (Fig. 5). One important change is a different conformation of Gln299. In the MtHDH/IMD complex, Gln299 Nε interacted with the carboxylic O of Glu397, whereas the Oε was H-bonded to the Oγ of Ser277 and the Oγ of Ser306. In the complex with HOL, where Gln299 no longer binds Ser277, the latter has an altered conformation. It appears that Gln299 and Ser277 are the key players, because Ser277 takes part in NAD+ binding (see below). Although Gln299 belongs to domain II, this domain shows few conformational changes, whereas the majority of the rearrangements concern domain I. The most prominent movement involves a loop region from Pro273 to Ser277 (Fig. 5). A corresponding fragment was briefly mentioned by Barbosa et al.16. Other conformational changes involve loop Pro168-Val173 and the loop with the N-terminal side of helix α10, Gly248-Tyr253. Rearrangement of the latter region flips the ψ angle of Pro249 by almost 155°, from −53.8° in the MtHDH/IMD complex to 151.3° in HOL complex. This feature has not been observed in EcHisD16. The three mostly loop regions are very conserved among HDH enzymes (Fig. 2) and, as presented below, contribute to NAD+ binding.

The structure of MtHDH shows NAD+ binding that allows for a hydride transfer. To visualise the cofactor binding site, MtHDH was cocrystallised after incubation with the product (His) to mimic the substrate (HOL) and NAD+ at pH 7.5 to allow for binding. If HOL was added instead of His, the increase in absorbance at 340 nm related to the production of NADH (not shown) confirmed that the enzymatic reaction had taken place. In the MtHDH/His/NAD+ complex, His binds in the same manner as HOL, with the HOL and His O atoms lying at corresponding positions and creating similar contacts with His368 and carbonyl O of His408. Gln299 coordinates Zn2+ via Oε as it does in the complex with HOL.
In all six subunits in the asymmetric unit, a clear electron density—which agrees with the structure of \( \text{NAD}^+ \)—was found in clefts within domain I that lead to the His binding site (Fig. 6). The \( \text{NAD}^+ \) adenine moiety is stacked between Phe96 (inter-ring distance \( \approx 3.5 \) Å) and Tyr253 (\( \approx 3.4 \) Å). Going towards the His-binding site, there are 15 polar interactions between \( \text{MtHDH} \) and \( \text{NAD}^+ \). The adenosine ribose \( O_2' \) interacts with \( \text{N_e} \) of Gln288 and \( O_3' \) with \( \text{O_6} \) of Asp97, whereas \( O_4' \) and \( O_5' \) both interact with \( N_8 \) of Asn225. The \( \text{NAD}^+ \) pyrophosphate binds to \( O_1' \) of Tyr166, \( O_7' \) of Thr171, and \( N_6 \) of Asn225, in addition to the backbone amides of Gly169, Gly170, Thr171, and Asn225. The nicotinamide-adjacent ribose, by its \( O_2' \), binds to \( O_7' \) and the amide of Ser277 (mentioned previously). Nicotinamide moiety binds to \( O_5' \) of Glu367 and Leu410, which belong to domain II. More specifically, the nicotinamide \( \text{NH}_2 \) group interacts with carbonyl \( O \) and nicotinamide \( O \) with the backbone amide of Thr277 (mentioned previously). It is important to note that most of the residues that bind \( \text{NAD}^+ \) (Gly169, Gly170, Thr171, Asn225, Ryr253, and Ser277) are within the loops pinpointed in the previous section as the fragments undergoing the most significant rearrangements upon \( \text{HAL} \) binding that “prepare” the enzyme for interaction with \( \text{NAD}^+ \). In other words, thanks to the \( \text{MtHDH} \) complexes, we know which conformational changes triggered by \( \text{HAL} \) binding enable the enzyme to bind \( \text{NAD}^+ \).

The distance between the C4 atom of nicotinamide and the C of His is approximately 3.1 Å. Such proximity would allow for a hydride transfer from the C atom of \( \text{HAL} \) to the C4 of \( \text{NAD}^+ \) and, subsequently, from \( \text{HAL} \) to another \( \text{NAD}^+ \) molecule. As mentioned in the Introduction, the corresponding distance in the only other \( \text{NAD}^+ \)-containing HDH structure, \( \text{EChisD} \) [PDB ID: 1kae, ref. 16], was 4.9 Å, which resulted in the presence of unreacted \( \text{HAL} \) in the active site. In that structure, there are also far fewer binding interactions between the enzyme and \( \text{NAD}^+ \). Only six direct and two water-mediated H-bonds anchor the cofactor to \( \text{EChisD} \). Among them, conserved between \( \text{EChisD} \) and \( \text{MtHDH} \) (corresponding residue parenthesised), are the following interactions: \( O_2' \) of adenosine ribose and \( \text{Gln}188 \) (\( \text{Gln}228 \)); \( O_4' \) and \( O_5' \) of the same ribose with \( N_6 \) of Asn211 (Asn225); pyrophosphate with backbone \( N \), \( N_6 \) of Asn211 (Asn225), and \( O_7 \) of Tyr130 (Tyr166). Neither nicotinamide nor nicotinamide ribose interacted with the protein atoms in \( \text{EChisD} \). This means that our structures of \( \text{MtHDH} \) complement complexes of \( \text{EChisD} \) by showing the location and conformation of \( \text{NAD}^+ \) that is very likely to reflect the reactive state.

Because \( \text{NAD}^+ \) was located far from the reactive carbon of \( \text{HAL} \) in the \( \text{EChisD} \) complex and the nicotinamide ring was oriented nearly perpendicular to the C–Cα bond of \( \text{HAL} \), it was difficult to determine on which face the hydride(s) might be accepted. Based on the \( \text{MtHDH} \) structures, it is possible to unambiguously determine that the hydride is abstracted by the \( \text{re} \) face of nicotinamide C4 (Fig. 6).

The root-mean-square-deviation (rmsd) between the \( \text{MtHDH}/\text{HAL} \) and \( \text{MtHDH}/\text{His}/\text{NAD}^+ \) complexes is 1.1 Å (419 Cα pairs within 3 Å distance), which would indicate more profound differences than between the \( \text{MtHDH}/\text{IMD} \) and \( \text{MtHDH}/\text{HAL} \) complexes (0.64 Å). However, as shown in the previous section, \( \text{HAL} \) binding involved crucial conformational changes, such as mainchain flips, whereas the subsequent binding of \( \text{NAD}^+ \) only causes shifts of region, particularly the V-shaped pair of helices α3 and α4 (Fig. 7). The lock-and-key analogy postulated by Emil Fischer in 1894 and commonly used for enzymes can be modified for \( \text{MtHDH} \) to lock-key-and-door. \( \text{HAL} \) binding unlocks the door and turns the knob (detailed but, nomen est amen, key changes) and allows \( \text{NAD}^+ \) to bind, which “opens the door.” However, in the case of \( \text{MtHDH} \), “opening” does not mean that there are hinges between domains because the relative positions of the four domains remain unchanged throughout the course of the reaction.

**Plant versus bacterial HDHs.** Biochemical studies of cabbage enzyme (BoHDH) that is nearly identical to the \( \text{MtHDH} \) showed that, functionally, plant HDHs behave similarly to the bacterial orthologs 30, 31, 38. With the knowledge of the 3-D structures of \( \text{MtHDH} \), it is possible to adopt and update the mechanism proposed...
previously\(^1,12,16,39\) (Fig. 1). Namely, a proton is withdrawn from the HOL O atom by N\(_\varepsilon\) of His368 (Base 1) that becomes double-protonated (it must also be protonated at N\(_\delta\), as it forms an H-bond with the carbonyl O of Glu278), and one hydride is abstracted by the first NAD\(^+\) molecule. The hybridisation of reactive carbon changes from sp\(^3\) to sp\(^2\), and HAL is formed. The “used” NADH dissociates and is replaced by the second NAD\(^+\) molecule. A water molecule (Wat3 in Fig. 4) is activated by Glu367 (Base 2) and performs a nucleophilic attack on the reactive carbon, forming a new C–O bond. Simultaneously, the HAL oxygen withdraws the proton back from N\(_\varepsilon\) of His368 (H:Base 3, now acting as an H-donating acid), resulting in the formation of a gem-diol HAL hydrate with sp\(^3\)-hybridised carbon. In the next step, His368 (Base 4) abstracts a proton from one of the hydroxyl groups of HAL hydrate, whereas the second NAD\(^+\) removes hydride from the reactive carbon, changing its hybridisation to sp\(^2\) and producing His.

Activation of HDH enzymes by Mn\(^{2+}\), reported for bacterial orthologues\(^15\), is very unlikely to be physiologically relevant in plants. Only a 10% increase of the \textit{Salmonella typhimurium} enzyme activity was observed in 50 \(\mu\)M Mn\(^{2+}\), and a 50% increase was observed in 500 \(\mu\)M Mn\(^{2+}\). Such a high concentration corresponds to Mn-toxic acid soils, and while chloroplasts are second to vacuoles among organelles that accumulate Mn\(^{2+}\), a 10 \(\mu\)M concentration is already considered high\(^40\).

Structurally, \textit{Mt} HDH is similar to the two bacterial HDHs, \textit{Ec}HisD and \textit{Bs}HisD. There are, however, a few differences, mainly within domains I and III, as revealed by the superpositions of the \textit{Mt}HDH/IMD complex with \textit{Ec}HisD (PDB ID: 1k75) and \textit{Bs}HisD (4g07). They are most prominent between residues 98–117 in \textit{Mt}HDH and the corresponding residues 58–81 in \textit{Ec}HisD and \textit{Bs}HisD (Fig. 8). The two bacterial HDHs lack the 32 strand (in \textit{Mt}HDH topology) and the short \(\gamma5\) helix that \textit{Mt}HDH has. Furthermore, the subsequent residues 111–115 form a loop region in \textit{Mt}HDH, whereas the corresponding fragment of prokaryotic HDHs is longer by two amino acids (Fig. 2) and forms a helix. Another significant difference is the presence of a two-stranded \(\beta\)-sheet within the domain III in \textit{Mt}HDH, as opposed to the three-stranded \(\beta\)-sheet in \textit{Ec}HisD and \textit{Bs}HisD. In \textit{Mt}HDH, only Val145 interacts as in a \(\beta\)-sheet. However, in all three enzymes, domain III extends the \(\beta\)-sheet of domain II (see above).

It is also very interesting to note that the correlation of the sequence alignment (Fig. 2) with a structural comparison shows that HDHs often share a more significant similarity within the loop regions than within fragments of a more compact secondary structure.

Most plant chloroplast dehydrogenases prefer NADP(H) over NAD(H) as the dinucleotide cofactor. HDHs are unusual from that perspective. Cabbage \textit{Bo}HDH, for instance, displays a 870-fold preference for NAD\(^+\) over NADP\(^+\)\(^41\). The \textit{Mt}HDH/His/NAD\(^+\) structure explains this preference. A formation of the \textit{Mt}HDH/NADP\(^+\) complex is highly disfavoured because of severe steric hindrances between the O2’ phosphate and Gln228 (Fig. 6), in addition to the negative–negative charge repulsion.

HDH inhibitors have the potential to be antibacterial agents\(^6–9\). It will be exciting to see if similar compounds act against plant enzymes, too. Undoubtedly, further studies are required to examine the possibilities.
light, this work, which updates and rectifies previous observations, particularly those related to the cofactor binding, is important because it may serve as a new scaffold for structure-based drug design. Now, thanks to the MtHDH structures, the design may also include transition-state analogues, which are generally considered to be very powerful enzyme inhibitors42, 43.

Methods

Cloning, overexpression, and purification of MtHDH. The total RNA was isolated from M. truncatula roots using the RNaseasy Plant Mini Kit (Qiagen), and the transcriptome was transcribed into the coding DNA (cDNA) with SuperScript II reverse transcriptase (Life Technologies) using oligo dT (15 and 18) primers. The cDNA suited as a template for amplification of the sequence coding for MtHDH ORF without the predicted N-terminal signal peptide (35 amino acids) by polymerase chain reaction (PCR). The signal peptide was recognised using the TargetP 1.1 server44, 45. The primers used (Forward: TACTTCCAATCCAATGCCTCCATTTCCATGGCAAATCCAATCAAAAC, Reverse: TTATCCACTTCCAATGTTATCATCTTTGAAACCTGTCTGGCTTCTATG) allowed us to incorporate the insert into the pMCSG68 vector (Midwest Center for Structural Genomics) using a ligase-independent cloning method46. The pMCSG68 vector introduces a His6-tag, followed by the Tobacco Etch Virus (TEV) protease cleavage site and the Ser-Asn-Ala linker, which precedes the N-terminus of the expressed protein. The correctness of the insert was confirmed by DNA sequencing.
Overexpression was carried out in LB media supplemented with 150 μg/mL ampicillin in BL21 Gold E. coli cells (Agilent Technologies). The bacteria were cultured with shaking at 210 rpm at 37 °C until the OD 600 reached 1.0. Then, the cultures were cooled down to 18 °C, and MtHDH overexpression was induced by the addition of isopropyl-D-thiogalactopyranoside at a final concentration of 0.5 mM, which continued for 18 h. The cell pellet from the 2 L culture was centrifuged at 3,500 x g for 20 min at 4 °C and resuspended in 35 mL of binding buffer [50 mM Hepes-NaOH pH 7.5; 500 mM NaCl; 20 mM IMD; 1 mM tris(2-carboxyethyl) phosphine (TCEP)] and stored at −80 °C. The samples were thawed and the cells were disrupted by sonication (4 min of probe working time), using bursts of 4 s and 26 s intervals for cooling in an ice/water bath. The cell debris was pelleted by centrifugation at 25,000 x g for 40 min at 4 °C. The supernatant was applied on a 50 ml column packed with 4 mL of HisTrap HP resin (GE Healthcare) and plugged into VacMan (Promega), with a vacuum pump setup to accelerate the process. The resin-bound MtHDH was washed five times with 40 mL of the binding buffer. The His 6-tagged protein was eluted with 20 mL of elution buffer (50 mM Hepes-NaOH pH 7.5; 500 mM NaCl; 400 mM IMD; 1 mM TCEP). The His 6-tag was cleaved with TEV protease (final concentration 0.1 mg/mL), and the IMD concentration was lowered to 20 mM by simultaneous dialysis overnight at 4 °C. The solution was applied again to HisTrap HP resin to remove the cleaved His 6-tag and the His 6-tagged TEV protease. The flow-through was collected, concentrated to 2.4 mL, applied on a HiLoad Superdex 200 16/60 column (GE Healthcare), and equilibrated with a buffer composed of 25 mM Hepes-NaOH pH 7.5, 100 mM KCl, 50 mM NaCl, and 1 mM TCEP.

### Crystallisation and diffraction data collection.

A homogenous, dimeric fraction of MtHDH was concentrated using Amicon concentrators (Millipore) to 8 mg/mL (based on A 280 with the extinction coefficient of 37,400). Screening for crystallisation conditions was performed using a robotic sitting drop vapor diffusion (Mosquito). The most promising hits were optimized manually in hanging drops. For complexes with HOL and His/NAD+, the ligands (20 mM final concentration) were added as buffered solutions (in 100 mM Hepes pH 7.5) to the protein sample and incubated overnight. The solution was applied again to HisTrap HP resin to remove the cleaved His-tag and the His-tagged TEV protease. The flow-through was collected, concentrated to 2.4 mL, applied on a HiLoad Superdex 200 16/60 column (GE Healthcare), and equilibrated with a buffer composed of 25 mM Hepes-NaOH pH 7.5, 100 mM KCl, 50 mM NaCl, and 1 mM TCEP.

| MOHDH | Zn²⁺, IMD | Zn²⁺, HOL | Zn²⁺, His, NAD⁺ |
|-------|----------|----------|-----------------|
| **Data collection** | | | |
| Wavelength (Å) | 1.0000 | 0.9763 | 1.0000 |
| Space group | P2₁ | P2₁ | P2₁ |

### Table 1. Data collection and refinement statistics. Values in parentheses correspond to the highest resolution shell. *R*<sub>meas</sub> = redundancy independent R-factor<sup>80</sup>.

|                  | MtHDH/IMD | MtHDH/HOL | MtHDH/His/NAD⁺ |
|------------------|-----------|-----------|-----------------|
| **Wavelength (Å)** | 1.0000 | 0.9763 | 1.0000 |
| **Space group** | P2₁ | P2₁ | P2₁ |
| a, b, c (Å) | 105.8, 142.8, 105.4 | 102.9, 139.2, 102.7 | 103.7, 139.1, 103.6 |
| α, β, γ (°) | 90, 120.2, 90 | 90, 119.2, 90 | 90, 119.5, 90 |
| Resolution (Å) | 40–2.25 (2.39–2.25) | 50–1.97 (2.09–1.97) | 50–2.59 (2.75–2.59) |
| Reflections collected/unique | 622093/120737 | 1080052/175983 | 219001/76791 |
| Completeness (%) | 94.7 (73.8) | 99.4 (97.4) | 96.4 (90.9) |
| Multiplicity | 5.2 (3.5) | 6.1 (5.7) | 2.9 (2.7) |
| R<sub>meas</sub> (%) | 8.6 (71.5) | 8.5 (99.6) | 10.7 (64.8) |
| <Iσ(I)> | 15.7 (2.1) | 13.7 (1.9) | 10.3 (2.0) |
| **Refinement** | | | |
| R<sub>free</sub> reflections | 1208 | 1232 | 1152 |
| No. of atoms (non-H) | | | |
| protein | 19734 | 19694 | 19668 |
| ligands | 78 | 66 | 425 |
| solvent | 656 | 625 | 149 |
| R<sub>free</sub>/R<sub>work</sub> (%) | 18.1/23.4 | 17.9/22.9 | 22.1/26.4 |
| Average B-factor (Å²) | 42.0 | 47.0 | 57.0 |
| RMSD from ideal geometry | | | |
| bond lengths (Å) | 0.014 | 0.015 | 0.013 |
| bond angles (°) | 1.65 | 1.69 | 1.74 |
| Ramachandran statistics (%) | | | |
| favored | 96.8 | 96.9 | 97.3 |
| allowed | 3.2 | 3.1 | 2.7 |
| outliers | 0.0 | 0.0 | 0.0 |
| PDB code | 5vlb | 5vlc | 5vld |
Determination and refinement of the crystal structures. The crystal structure of MtHDH was solved by molecular replacement with Balbex, using 2.7 Å resolution data (not reported here). Phenix AutoBuild was used to build the initial model, which afterwards was placed inside the unit cell with the ACHESYM server. COOT was used for manual fitting in the electron density maps between rounds of model refinement in Refmac. A single TLS group was added to each chain in every structure, whereas non-crystallographic symmetry restraints were added only for the lowest-resolution MtHDH/His/NAD complex. The possibility of crystal twinning was taken into account because the a- and c-unit cell dimensions are similar, and the β angle is very close to 120° (Table 1). Twinning was, however, excluded during the structure refinement in Refmac. Models were validated in MolProbity. The refinement statistics are listed in Table 1.

Other software used. Molecular illustrations were created with UCSF Chimera, which also served for calculations of rmsds. Sequence alignment was performed in ClustalW, whereas identities/similarities were calculated in BLAST.

References
1. Winkler, M. E. In Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology (ed F. C. Neidhardt) (Am. Soc. Microbiol., 1996).
2. Fink, G. R. Gene-Enzyme Relations in Histidine Biosynthesis in Yeast. Science 146, 525–527 (1964).
3. Lunardi, J. et al. Targeting the histidine pathway in Mycobacterium tuberculosis. Carr Top Med Chem 13, 2866–2884 (2013).
4. Jastrzebowska, K. & Gabriel, I. Inhibitors of amino acids biosynthesis as antifungal agents. Amino Acids 47, 227–249 (2015).
5. Gohda, K., Ohta, D., Iwasaki, G., Ertl, P. & Jacob, O. Computational modeling of a binding conformation of the intermediate L-histidinol to histidinol dehydrogenase. J Chem Inf Comput Sci 43, 196–201 (2003).
6. Kohler, S. et al. The analysis of the intramolecular virulome of Brucella suis deciphers the environment encountered by the pathogen inside the macrophage host cell. Proc Natl Acad Sci USA 99, 15711–15716 (2002).
7. Parish, J. Travarvation survival response of Mycobacterium tuberculosis. J Bacteriol 185, 6702–6706 (2003).
8. Monti, S. M., De Simone, G. & D'Ambrosio, K. L-Histidinol Dehydrogenase as a New Target for Old Diseases. Carr Top Med Chem 16, 2369–2378 (2016).
9. Aguero, F. et al. Genomic-scale prioritization of drug targets: the TDR Targets database. Nat Rev Drug Discov 7, 900–907 (2008).
10. Adams, E. L-Histidinol, a biosynthetic precursor of histidine. J Biol Chem 217, 325–344 (1955).
11. Grubmeyer, C. T., Chu, K. W. & Insinga, S. Kinetic mechanism of histidinol dehydrogenase: histidinol binding and exchange reactions. Biochemistry 26, 3369–3373 (1987).
12. Grubmeyer, C. & Teng, H. Mechanism of Salmonella typhimurium histidinol dehydrogenase: kinetic isotope effects and pH profiles. Biochemistry 38, 7355–7362 (1999).
13. Gorisch, H. Steady-state investigations of the mechanism of histidinol dehydrogenase. Biochem J 181, 153–157 (1979).
14. Grubmeyer, C. T., Insinga, S., Bhatia, M. & Mozami, N. Salmonella typhimurium histidinol dehydrogenase: complete reaction stereochemistry and active site mapping. Biochemistry 28, 8174–8180 (1989).
15. Grubmeyer, C., Skiadopoulos, M. & Senior, A. E. L-histidinol dehydrogenase, a Zn2+–metalloenzyme. Arch Biochem Biophys 272, 311–317 (1989).
16. Barbosa, J. A. et al. Mechanism of action and NAD+–binding mode revealed by the crystal structure of L-histidinol dehydrogenase. Proc Natl Acad Sci USA 99, 1859–1864 (2002).
17. Hammes-Schiffer, S. Comparison of hydride, hydrogen atom, and proton–coupled electron transfer reactions. Chemphyschem 3, 33–42 (2002).
18. D'Ambrosio, K. et al. Structural basis for the rational design of new anti-Brucella agents: the crystal structure of the C366S mutant of L-histidinol dehydrogenase from Brucella suis. Biochemistry 97, 114–120 (2014).
19. Abdo, M. R. et al. Anti-virulence strategy against Brucella suis: synthesis, biological evaluation and molecular modeling of selective histidinol dehydrogenase inhibitors. Org Biomol Chem 9, 3681–3690 (2011).
20. Nunes, I. E. et al. Molecular, kinetic, thermodynamic, and structural analyses of Mycobacterium tuberculosis hisD-encoded metal-dependent dimeric histidinol dehydrogenase (EC 1.1.1.123). Arch Biochem Biophys 512, 143–153 (2011).
21. Stepansky, A. & Leustek, T. Histidine biosynthesis in plants. Amino Acids 30, 127–142 (2006).
22. Egle, R. A. Histidine biosynthesis. Arabidopsis Book 9, e0141 (2011).
23. Wiater, A., Krajewska-Grynikiewicz, K. & Klopotsowski, T. Histidine biosynthesis and its regulation in higher plants. Acta Biochim Pol 48, 299–307 (1991).
24. Fujimori, K. & Ohta, D. Isolation and characterization of a histidine biosynthetic gene in Arabidopsis encoding a polypeptide with two separate domains for phosphoribosyl-ATP pyrophosphohydrolase and phosphoribosyl-AMP cyclohydrolase. Plant Physiol 118, 275–283 (1998).
25. Fujimori, K., Tada, S., Kanai, S. & Ohta, D. Molecular cloning and characterization of the gene encoding N'-([5'-phosphoribosyl]formimino)-5-aminooimidazole-4-carboxamido ribonucleotide (BBM II) isomerase from Arabidopsis thaliana. Mol Gen Genet 259, 216–223 (1998).
26. Fujimori, K. & Ohta, D. An Arabidopsis cDNA encoding a bifunctional glutamine amidotransferase/cyclase suppresses the histidine auxotrophy of a Saccharomyces cerevisiae his7 mutant. FEBS Lett 428, 229–234 (1998).
27. El Malki, F., Frankard, V. & Jacobs, M. Molecular cloning and expression of a cDNA sequence encoding histidinol phosphatase aminotransferase from Nicotiana tabacum. Plant Mol Biol 37, 1013–1022 (1998).
28. Petersen, L. N. et al. The missing link in plant histidine biosynthesis: Arabidopsis myoinositol monophosphatase-like2 encodes a functional histidinol-phosphate phosphatase. *Plant Physiol* **152**, 1186–1196 (2010).
29. Wong, Y. S. & Mazelis, M. Detection and properties of L-histidinol dehydrogenase in wheat-germ. *Phytochemistry* **20**, 1831–1834 (1981).
30. Nagai, A. et al. Structural and functional conservation of histidinol dehydrogenase between plants and microbes. *Proc Natl Acad Sci USA* **88**, 4133–4137 (1991).
31. Nagai, A. & Scheidegger, A. Purification and characterization of histidinol dehydrogenase from cabbage. *Arch Biochem Biophys* **284**, 127–132 (1991).
32. Bisson, C. et al. Crystal Structures Reveal that the Reaction Mechanism of Imidazolglycerol-Phosphate Dehydratase Is Controlled by Switching Mn(II) Coordination. *Structure* **23**, 1236–1245 (2015).
33. Glynn, S. E. et al. Structure and mechanism of imidazolglycerol-phosphate dehydratase. *Structure* **13**, 1809–1817 (2005).
34. Ruszowski, M. & Dauter, Z. Structural Studies of Medicago truncatula Histidinol Phosphate Phosphatase from Inositol Monophosphatase Superfamily Reveal Details of Penultimate Step of Histidine Biosynthesis in Plants. *J Biol Chem* **291**, 9960–9973 (2016).
35. UniProt, C. UniProt: a hub for protein information. *Nucleic Acids Res* **43**, D204–212 (2015).
36. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797 (2007).
37. Rao, S. T. & Rossmann, M. G. Comparison of super-secondary structures in proteins. *J Mol Biol* **76**, 241–256 (1973).
38. Khetromoom, A. et al. Steady-state kinetics of cabbage histidinol dehydrogenase. *Arch Biochem Biophys* **312**, 493–500 (1994).
39. Teng, H. & Grubmeyer, C. Maturation of histidinol dehydrogenase reveals roles for conserved histidine residues. *Biochemistry* **38**, 7363–7371 (1999).
40. Pittman, J. K. Managing the manganese: molecular mechanisms of manganese transport and homeostasis. *New Phytol* **167**, 733–742 (2005).
41. Nagai, A. et al. In Research in photosynthesis (ed N Murata) 95–98 (1993).
42. Davis, J. P., Cain, G. A., Pitts, W. J., Magolda, R. L. & Copeland, R. A. The immunosuppressive metabolite of leflunomide is a potent inhibitor of human dihydroorotate dehydrogenase. *Biochemistry* **35**, 1270–1273 (1996).
43. Schramm, V. L. Enzymatic transition states, transition-state analogs, dynamics, thermodynamics, and lifetimes. *Annu Rev Biochem* **80**, 703–732 (2011).
44. Emanuelsen, O., Nielsen, H., Brunak, S. & von Heijne, G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol* **300**, 1005–1016 (2000).
45. Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* **10**, 1–6 (1997).
46. Kim, Y. et al. High-throughput protein purification and quality assessment for crystallization. *Methods* **55**, 12–28 (2011).
47. Gorrec, F. The MORPHEUS protein crystalization screen. *J Appl Crystallogr* **42**, 1035–1042 (2009).
48. Kabsch, W. *X-ray Acta Cryst. D* **66**, 125–132 (2010).
49. Long, E., Vagin, A. A., Young, P. & Murshudov, G. N. BALBES: a molecular-replacement pipeline. *Acta Cryst. D* **64**, 125–132 (2008).
50. Terwilliger, T. C. et al. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Cryst. D* **64**, 61–69 (2008).
51. Kowiel, M., Jaskolski, M. & Dauter, Z. ACHESYM: an algorithm and server for standardized placement of macromolecular models in the unit cell. *Acta Cryst. D* **70**, 3290–3298 (2014).
52. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Cryst. D* **66**, 486–501 (2010).
53. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Cryst. D* **67**, 355–367 (2011).
54. Winn, M. D., Murshudov, G. N. & Papiz, M. Z. Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol* **374**, 300–321 (2003).
55. Winn, M. D., Isupov, M. N. & Murshudov, G. N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Cryst. D* **57**, 122–133 (2001).
56. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Cryst. D* **66**, 12–21 (2010).
57. Pettersen, E. F. et al. UCSF Chimera-a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612 (2004).
58. MA, L. et al. ClustalW and ClustalX version 2. *Bioinformatics* **23**, 2947–2948 (2007).
59. Autschbl, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J Mol Biol* **215**, 403–410 (1990).
60. Diederichs, K. & Karplus, P. A. Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol* **4**, 269–275 (1997).

Acknowledgements

This project was supported in part by the Intramural Research Program of the NCI Center for Cancer Research. Diffraction data were collected at the SER-CAT beamline 22-ID at the Advanced Photon Source, Argonne National Laboratory, supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract W-31–109-Eng-38, and at the Structural Biology Center at the Advanced Photon Source, operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. We are thankful to Dr. Bartosz Sekula for a careful reading of the manuscript.

Author Contributions

M.R. designed and performed the studies, analysed the results, and wrote the manuscript. Z.D. analysed the results and supervised the work.

Additional Information

Competing Interests: The authors declare that they have no competing interests.

Accession Numbers: Coordinates and structure factors of the related structures were deposited in the Protein Data Bank (PDB): MtHDH/IMD complex, 5vib; MtHDH/HOL, 5vlc; MtHDH/His/NAD++, 5vld.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
