



OPEN

SUBJECT AREAS:
ENZYMES
PHYLOGENETICSReceived
8 October 2014Accepted
23 February 2015Published
16 March 2015Correspondence and
requests for materials
should be addressed to
G.Z. (gpz1996@
yahoo.com)

Novel Type II and Monomeric NAD⁺ Specific Isocitrate Dehydrogenases: Phylogenetic Affinity, Enzymatic Characterization, and Evolutionary Implication

Peng Wang, Changqi Lv & Guoping Zhu

Institute of Molecular Biology and Biotechnology, College of Life Sciences, Anhui Normal University, No.1 Beijing East Road, Wuhu 241000, Anhui, China.

NAD⁺ use is an ancestral trait of isocitrate dehydrogenase (IDH), and the NADP⁺ phenotype arose through evolution as an ancient adaptation event. However, no NAD⁺-specific IDHs have been found among type II IDHs and monomeric IDHs. In this study, novel type II homodimeric NAD-IDHs from *Ostreococcus lucimarinus* CCE9901 IDH (OIIDH) and *Micromonas* sp. RCC299 (MiIDH), and novel monomeric NAD-IDHs from *Campylobacter* sp. FOBRC14 IDH (CaIDH) and *Campylobacter curvus* (CcIDH) were reported for the first time. The homodimeric OIIDH and monomeric CaIDH were determined by size exclusion chromatography and MALDI-TOF/TOF mass spectrometry. All the four IDHs were demonstrated to be NAD⁺-specific, since OIIDH, MiIDH, CaIDH and CcIDH displayed 99-fold, 224-fold, 61-fold and 37-fold preferences for NAD⁺ over NADP⁺, respectively. The putative coenzyme discriminating amino acids (Asp326/Met327 in OIIDH, Leu584/Asp595 in CaIDH) were evaluated, and the coenzyme specificities of the two mutants, OIIDH R³²⁶H³²⁷ and CaIDH H⁵⁸⁴R⁵⁹⁵, were completely reversed from NAD⁺ to NADP⁺. The detailed biochemical properties, including optimal reaction pH and temperature, thermostability, and metal ion effects, of OIIDH and CaIDH were further investigated. The evolutionary connections among OIIDH, CaIDH, and all the other forms of IDHs were described and discussed thoroughly.

The progressive sequencing of complete biological genomes has dramatically increased the size of protein databases. The expansion of protein information, most of which are functionally annotated by computational techniques, will consequently increase the diversity of each protein family, thus providing us with an opportunity to extend and refine the classification of protein families. Protein phylogenetic analysis has always been important, because it can provide insight into protein evolution and further implicate protein function. In the present study, we applied this principle to explore novel isocitrate dehydrogenases (IDHs).

IDH is a key enzyme in the tricarboxylic acid (TCA) cycle. It catalyzes the oxidative decarboxylation of isocitrate to α -ketoglutarate (α -KG) and CO₂, which is accompanied by the reduction of NAD(P⁺) to NAD(P)H. The IDH reaction provides organisms with not only energy but also biosynthetic precursors, such as α -KG, for metabolism. Thus, these metabolic pathways are among the first to have evolved^{1,2}. Consequently, IDHs are ubiquitously distributed throughout the three domains of life: Archaea, Bacteria, and Eukarya. Based on coenzyme specificity, the IDH family can be divided into NAD⁺-dependent IDHs (EC 1.1.1.41, NAD-IDHs) and NADP⁺-dependent IDHs (EC 1.1.1.41, NADP-IDHs). IDHs with different coenzyme dependencies play varying roles *in vivo*. NAD-IDH catalysis generates NADH, which participates in energy metabolism. NADP-IDH catalysis generates NADPH, which is an important source of reducing power. NADPH also plays a role in the cellular defense against oxidative damage and the detoxification of reactive oxygen species³⁻⁵.

Most studies have focused on the general understanding of the biochemistry, structure, and evolution of IDH. Meanwhile, the potential applications of IDH in biotechnology and drug design against pathogens have also been recently investigated. An NAD-IDH with poor performance in decarboxylating from *Zymomonas mobilis* serves



as a potential genetic modification target towards optimized *Z. mobilis* strains to produce ethanol⁶. The characterization of NADP-IDH from *Microcystis aeruginosa* may provide new ideas for controlling blue-green algae through biological techniques⁷. IDHs from pathogenic bacteria, such as *Plasmodium falciparum*, *Mycobacterium tuberculosis*, and *Leptospira interrogans*, have been reported as drug targets, because they stand at a branch point of the TCA cycle and glyoxylate shunt^{8–10}. IDHs are also ideal immuno-diagnostic candidates, due to their highly conserved housekeeping function. For example, *M. tuberculosis* IDHs elicit strong B-cell responses in tuberculosis (TB)-infected populations and can differentiate between healthy vaccinated and TB populations¹¹. In addition, *Helicobacter pylori* IDH can be an immunogen that interacts with the host immune system to subsequently lead to possible autolytic release and significantly elicit humoral responses in individuals with invasive *H. pylori* infection¹².

Besides pathogenic bacterial IDH, human cytosolic NADP-IDH (IDH1) and mitochondrial NADP-IDH (IDH2) have been considered as drug targets. Mutations in IDH1 and IDH2 are frequently identified in various cancers, such as glioblastoma multiforme and acute myeloid leukemia^{13,14}. Heterozygous IDH mutations are remarkably specific to a single codon in the conserved and functionally important arginine 132 residue (R132) of IDH1 and 172 residue (R172) of IDH2. Mutations result in the simultaneous loss of normal IDH catalytic activity. However, the production of α -KG and NADPH grants mutated IDHs with the neomorphic activity of reducing α -KG to 2-hydroxyglutarate (2-HG), which is accompanied by the oxidation of NADPH to NADP⁺^{15,16}. The accumulation of 2-HG competitively inhibits α -KG-dependent enzymes, thus causing cellular alterations in epigenetics, collagen maturation, and hypoxia signaling^{17–19}.

As an ancient enzyme, IDH acquired various primary structures and different oligomeric states through evolution. Four kinds of IDHs have been reported: monomer, homo-dimer, homo-tetramer, and hetero-oligomer. Monomeric IDHs have been characterized from various eubacteria, and all of them are highly specific to NADP⁺^{20–22}. Because the amino acid sequence identities are <10% between monomeric IDHs and other types of IDHs, this group has been recognized as a separate clade that evolved independently^{20,23}. Dimeric and multimeric IDHs have been divided into three phylogenetic subfamilies^{23–25}. Subfamily I is a prokaryotic group, in which NAD⁺ and NADP⁺ usage is widespread within archaeal and eubacterial homo-dimeric enzymes. Subfamily II is mainly composed of eukaryotic homo-dimeric NADP-IDHs, with a small number of eubacterial homo-dimeric NADP-IDHs. Subfamily III is comprised of mitochondrial hetero-oligomeric NAD-IDHs and eubacterial homo-tetrameric enzymes with either NAD⁺ or NADP⁺ as the cofactor²⁶.

Because subfamily III IDHs share more than 30% sequence identities with subfamily I members but less than 15% sequence identities with subfamily II counterparts, subfamilies III and I can be combined accordingly. Therefore, we simply divided the IDH protein family into type I IDHs (subfamily I and III) and type II IDHs (subfamily II) in our previous study⁴. Three total types of IDH can be distinguished: type I IDH, type II IDH, and monomeric IDH. Interestingly, both NAD⁺ and NADP⁺ are broadly utilized among type I IDHs. However, homo-dimeric type II IDHs and monomeric IDHs are all NADP⁺ specific⁴. We previously demonstrated that NAD⁺ use is an ancestral trait of IDH, and the NADP⁺ phenotype of eubacterial dimeric NADP-IDH arose on or about the time that eukaryotic mitochondria first appeared (about 3.5 billion years ago) to synthesize NADPH for the adaptation of bacterial growth on acetate⁴. Consequently, NADP⁺-specific type II IDHs and monomeric IDHs may also have their own NAD⁺-specific ancestors. However, no NAD-IDHs that belong to these two subfamilies have been explored.

The rapid growth of the IDH protein “pool”, which has been contributed by various genome projects, should allow us to search

for putative NAD⁺-specific type II IDHs and monomeric IDHs. In the present study, four novel NAD-IDHs that belong to the families of type II IDHs and monomeric IDHs were reported for the first time. Enzymatic properties of these NAD-IDHs, including the oligomeric state in solution, optimum pH and temperature for catalysis, thermostability, and metal ion dependency, were characterized thoroughly. Kinetic parameters of the two IDHs towards coenzymes were determined in detail, and their coenzyme-binding sites were evaluated by site-directed mutagenesis. The discovery of these novel NAD-IDHs will refine the chemistry and phylogeny of IDH and provide new insights into the evolution of this ancient protein family.

Results

Phylogenetic Analysis and Sequence Alignment. In our previous study, we divided the IDH protein family into type I IDHs and type II IDHs based on sequence homology⁴. Herein, we expanded the classification of this ancient family by incorporating new members (Fig. 1). Monomeric IDHs were included, although their overall sequence homologies with other forms of IDH are relatively low ($\approx 10\%$)^{22,27}. Consequently, they constituted a monophyletic clade on the phylogenetic tree (Fig. 1). Therefore, three main IDH subfamilies (type I IDH, type II IDH and monomeric IDH) were identified to comprise the whole IDH protein family. Members in IDH subfamily were further classified into eleven subgroups with regard to different coenzyme specificity, different oligomeric states and diverse resources (Fig. 1). Different tree building algorithms were also employed, such as UPGMA method, Maximum Likelihood method and Minimum Evolution method. All trees contain three well-supported monophyletic groups: type I IDHs, type II IDHs, and monomeric IDHs (see Supplementary Fig. S3, Fig. S4 and Fig. S5 online). When an outgroup of malate dehydrogenase (MDH) was added into the phylogenetic study, the overall topology of the evolutionary tree remained very similar (see Supplementary Fig. S6 online).

In the type I subgroup, the most common ones are the eubacterial homodimeric NADP- and NAD-IDHs, which are represented by *E. coli* IDH and *Z. mobilis* IDH, respectively^{6,23}. The pairwise amino acid sequence identity among this subgroup was more than 50%. Mitochondrial hetero-oligomeric NAD-IDHs, together with a batch of eubacterial homotetrameric NAD-IDHs, were grouped into a single branch in the type I subfamily. More than 45% identity exists between these eukaryotic and prokaryotic NAD-IDHs²⁶. Although this branch was designated as subfamily III in other IDH phylogenetic studies^{20,24,25,28}, it is reasonable to encompass it into the type I subfamily, because proteins in this branch exhibit more than 30% identity with type I eubacterial homodimeric IDHs. A small group of IDHs, which were represented by *Rickettsia* IDH and *Thermus thermophilus* IDH, was branched before the mitochondrial group. These IDHs were structurally distinguished from other dimeric IDHs, as they were longer in the C-terminal region. Although they were assigned as subfamily IV IDHs in two recent studies^{28,29}, they were included as type I IDHs in our study, as they shared considerable sequence identities (>40%) with mitochondrial NAD-IDHs.

The type II subfamily was comprised of homodimeric NADP-IDHs from eukaryotes and eubacteria (Fig. 1). A small group of NADP-IDHs, which are represented by *Thermotoga maritima* IDH (TmIDH)³⁰, branched clearly before the well-characterized homodimeric NADP-IDH clade. Although the TmIDH-like proteins were highly identical (>50%) to type II homodimeric NADP-IDHs, these two kinds of IDHs branched separately in the type II subtree. This separate branching may be due to the fact that TmIDH-like IDHs can form homotetramers in solution under some conditions^{25,30}. Because all reported type II IDHs have been NADP⁺ specific, no NAD-IDHs have been found in this subfamily. In this study, we identified the type II NAD-IDH group for the first time. This small group of distinctive IDHs was represented by IDH from *O.*

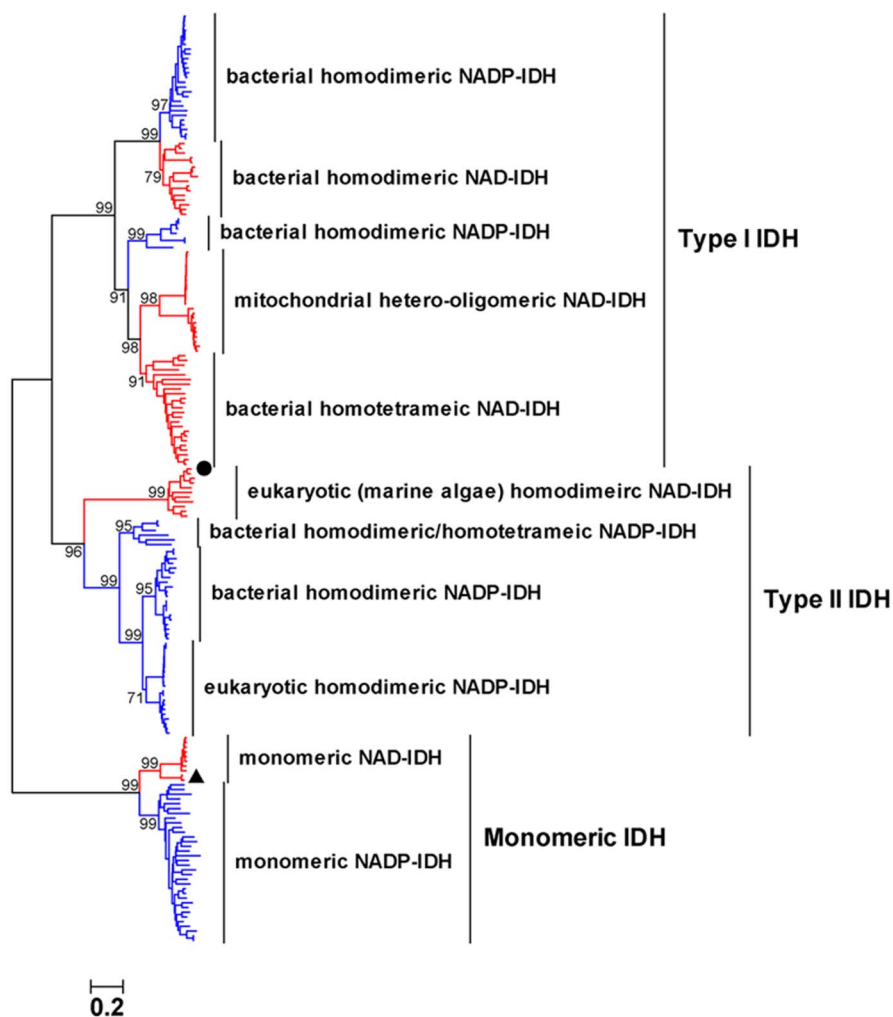


Figure 1 | Evolutionary relationships of 197 IDHs from diverse background. The evolutionary history was inferred using the Neighbor-Joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. Phylogenetic analyses were conducted in MEGA6. The IDH sequences used were listed in Table S1. OIIDH and CaIDH were marked by “●” and “▲” respectively.

lucimarinus CCE9901 (OIIDH), along with several counterparts that are derived from marine algae, such as *Micromonas* sp. RCC299 (MiIDH, GenBank accession number: XP_002502450). OIIDH demonstrated substantial sequence identity (>30%) with typical type II homodimeric NADP-IDHs and low identity with type I IDHs (<15%). Therefore, OIIDH clustered into a unique clade among the type II subfamily.

Although OIIDH was recognized as a member of the type II subfamily, its coenzyme specificities seemed to be different from those of the other type II NADP-IDHs. The crystal structure of human cytosolic homodimeric NADP-IDH, which is the most investigated type II IDH, shows that Arg353 and His354 were involved directly in coenzyme discrimination³¹. The conservation of these two NADP⁺-binding residues was confirmed by the structures of type II NADP-IDHs from *M. tuberculosis*⁹, *Desulfotalea psychrophila*³², *Thermotoga maritima*³⁰, porcine mitochondria³³, and yeast mitochondria³⁴. However, the corresponding amino acid residues in OIIDH were substituted with Asp326 and Met327 (Fig. 2). The replacement of two positively charged amino acids with one negatively charged amino acid (Asp) and one neutral amino acid with a large side chain (Met) suggests that OIIDH should favor its binding to NAD⁺ over NADP⁺, because Asp in the coenzyme binding site will repel the 2'-phosphate of NADP⁺ but can properly contact with NAD⁺^{35,36}. Other suspected type II NAD-IDHs that were grouped with OIIDH

also had Asp and Met (or Leu) in the corresponding sites (see Supplementary Fig. S1 online), suggesting that they share the same coenzyme-binding mechanism.

All monomeric IDHs were separated into a monophyletic group in the phylogenetic tree (Fig. 1). Two subgroups could be distinguished in the clade clearly, one of which was represented by the well-studied monomeric NADP-IDHs from *Azotobacter vinelandii* (AvIDH)²¹ and *Corynebacterium glutamicum* (CgIDH)²⁰. The other, however, was newly discovered and was proposed to be NAD⁺ specific. The representative member of this special subgroup was the *Campylobacter* sp. FOBR14 IDH (CaIDH). CaIDH was 731 amino acids in length, which is typical for monomeric IDH. However, CaIDH shared less than 50% sequence identity with monomeric NADP-IDHs, whereas monomeric NADP-IDHs shared more than 70% sequence identity among themselves. Sequence alignment results show that the key coenzyme-binding residues, His589 and Arg600, in AvIDH²⁰, which are absolutely conserved in all monomeric NADP-IDHs, had been replaced by Leu and Asp at the corresponding sites of CaIDH (Fig. 2). The presence of negatively charged Asp and neutral Leu with a large side chain eliminates the possibility of NADP⁺ use by CaIDH, thus making CaIDH the first putative NAD⁺-specific monomeric IDH. Besides CaIDH, some other homogeneous monomeric NAD-IDHs, which are derived from *Campylobacter* species, such as *Campylobacter curvus* (CcIDH, GenBank accession number: WP_018136314),

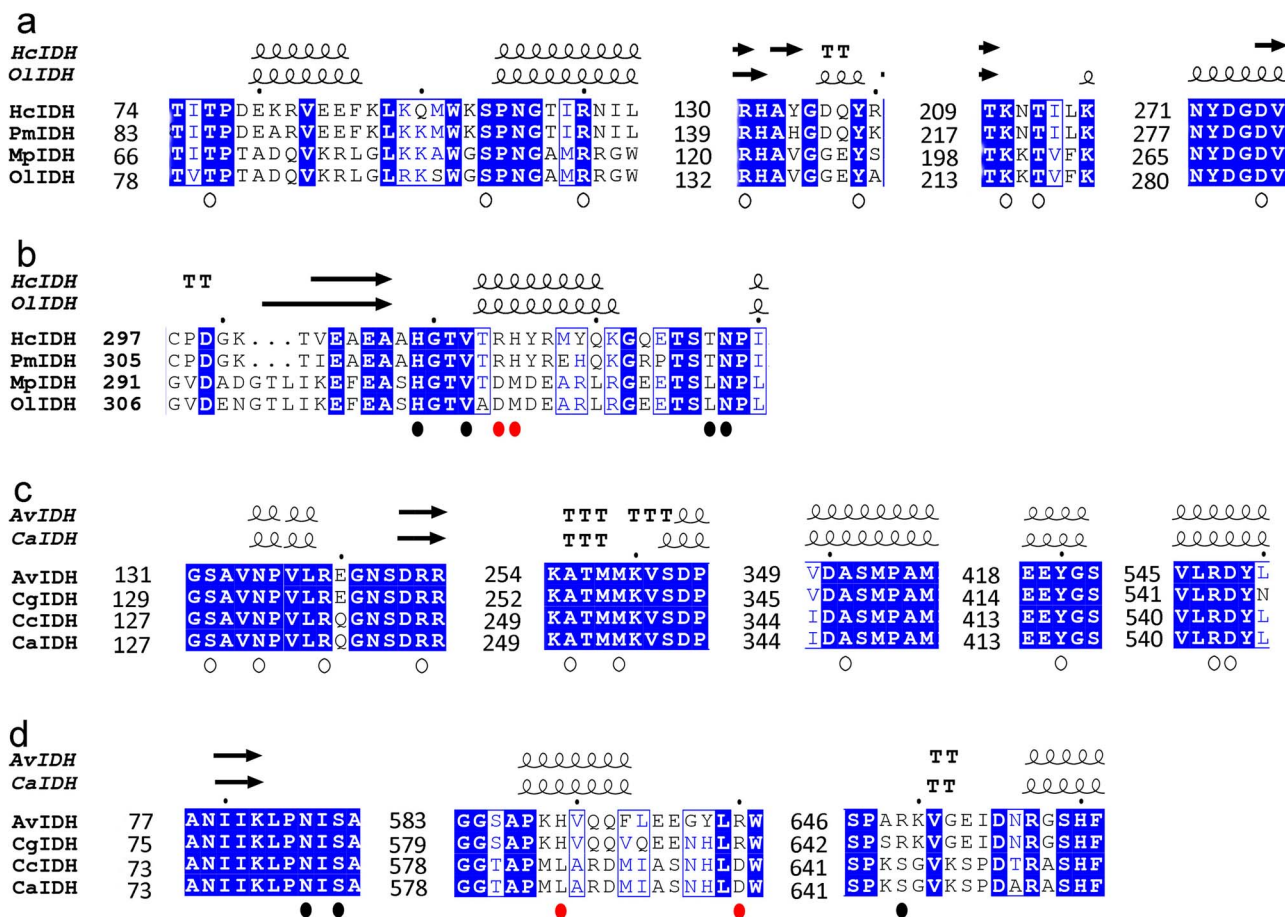


Figure 2 | Secondary structure based protein sequence alignments. (a, b), The comparison of OIIDH with its homologous IDH from *Micromonas pusilla* (MpIDH, GenBank Accession No. XP_003056989.1) and two typical type II homodimeric NADP-IDHs from human cytosol (HcIDH, GenBank Accession No. NP_001269316.1) and porcine mitochondria (PmIDH, GenBank Accession No. P33198.1). (c, d), The comparison of CaIDH with its homologous IDH from *Campylobacter curvus* (CcIDH, GenBank Accession No. WP_018136314.1) and two typical monomeric NADP-IDHs from *Azotobacter vinelandii* (AvIDH, GenBank Accession No. BAA11169.1) and *Corynebacterium glutamicum* (CgIDH, GenBank Accession No. WP_011013800.1). High-resolution structure of HcIDH (PDB ID, 1T09) and AvIDH (PDB ID, 1J1W) were downloaded from the PDB database. The homology models of OIIDH and CaIDH were generated by SWISS-MODEL server (<http://swissmodel.expasy.org/>). Invariant residues are highlighted by shaded blue boxes and conserved residues by open blue boxes. The conserved residues involved in the substrate binding are indicated by empty circle (○). The core cofactor binding sites, which are the targets of the site directed mutagenesis, are indicated by red dots (●) while other cofactor binding pocket residues are indicated by black dots (●). The figure was made with ESPript 2.2.

were also found (see Supplementary Fig. S2 online). These characteristic IDHs shared more than 70% sequence identity with each other, and they all had Leu and Asp (or Leu) at the putative coenzyme-binding sites (see Supplementary Fig. S2 online).

Overexpression, Purification, and Oligomeric State Determination.

The recombinant OIIDH and CaIDH that were tagged with 6×His were successfully produced in *E. coli* and then purified to homogeneity (Fig. 3a). Purified OIIDH and CaIDH gave a single band around 45 kDa and 80 kDa in SDS-PAGE, respectively, which compared well with the theoretical molecular mass of 6×His-tagged OIIDH (46 kDa) and CaIDH (81 kDa). Size exclusion chromatography (SEC) was then performed to estimate the oligomerization status of OIIDH and CaIDH in solution. A single-elution peak was observed for OIIDH, and its native molecular mass was estimated to be 76 kDa (Fig. 3b). This data can be interpreted as the protein being presented as a homodimer in solution, since the calculated mass of a monomer is 46 kDa. As for CaIDH, a dominant peak, corresponding to a molecular weight of 81 kDa, was eluted (Fig. 3c). This proved that the overwhelming majority of CaIDH presented as a monomer in solution. The fraction (<5%) that eluted before the main peak was calculated to be 150 kDa, and this represented a dimeric CaIDH. The

dimeric CaIDH was more likely to be induced by factors, such as protein concentration, salt, temperature, and pH.

However, molecular weight estimation by SEC is limited by the premise that the protein interacts with the column resin in an ideal way (e.g., no electrostatic or hydrophobic interactions)^{37,38}. In our experiment, the disagreement between the SEC estimation (76 kDa) and the deduced molecular weight of the dimeric OIIDH (92 kDa) may be due to non-ideal interactions between the protein and the SEC media. In order to accurately assign the molecular masses of the recombinant OIIDH and CaIDH, we performed the MALDI-TOF/TOF mass spectrometry. The molecular weight of the recombinant OIIDH and CaIDH were precisely determined to be 93 kDa and 82 kDa, respectively (Fig. 4), demonstrating the homodimeric structure of OIIDH and the monomeric structure of CaIDH.

Enzyme Activity and Kinetic Characterization. The specific activity of purified recombinant OIIDH was 72.3 U/mg with NAD⁺ and only 3.8 U/mg with NADP⁺. Recombinant CaIDH showed similarly high specific activity with NAD⁺ (53.2 U/mg) and very low activity with NADP⁺ (8.7 U/mg). This observation primarily confirmed the NAD⁺ preference of OIIDH and CaIDH, as suggested by the bioinformatic analysis. Kinetic characterization results show that the

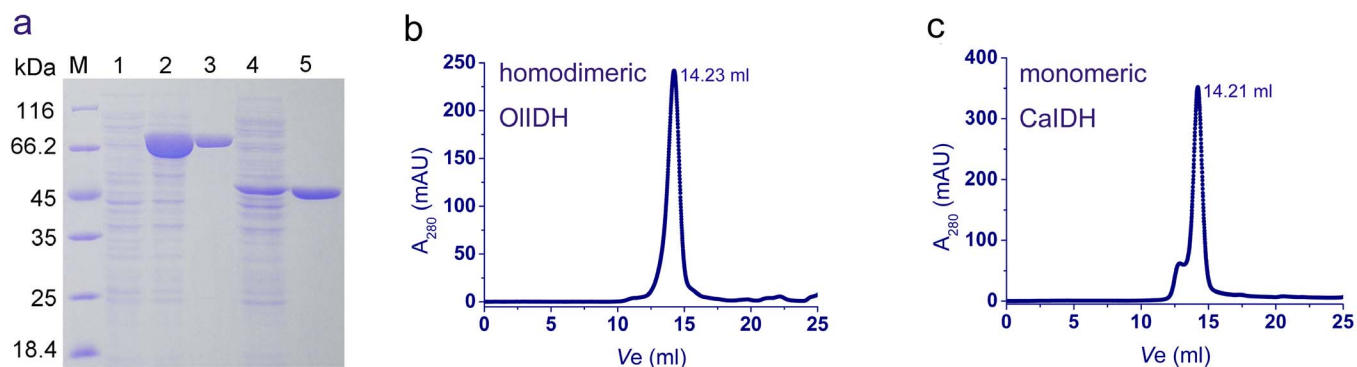


Figure 3 | Overexpression, purification and oligomeric state determination of the recombinant CaIDH and OIIDH. (a), Protein purity was detected by 12% SDS-PAGE. M, protein marker; lane 1, crude extracts of cells harboring pET-28b(+) with IPTG induction; lane 2, crude extracts of cells harboring recombinant plasmid pET-CaIDH with IPTG induction; lane 3, purified recombinant CaIDH; lane 4, crude extracts of cells harboring recombinant plasmid pET-OIIDH with IPTG induction; lane 5, purified recombinant OIIDH. (b) and (c), SEC analysis of the recombinant OIIDH and CaIDH. The flow rate was $0.5 \text{ ml}\cdot\text{min}^{-1}$ and the proteins in the fractions were monitored at 280 nm. V_e of the recombinant OIIDH and CaIDH were 14.23 ml and 14.21 ml, respectively.

K_m of OIIDH for NADP^+ was over 16-fold greater than the K_m for NAD^+ . The coenzyme specificity (k_{cat}/K_m) of OIIDH was 99-fold greater for NAD^+ than NADP^+ (Table 1). Consequently, OIIDH showed a high preference for NAD^+ , thus becoming the first NAD^+ -specific IDH in the type II subfamily. As expected, CaIDH was also characterized as an NAD^+ -specific IDH by kinetic analysis,

and its coenzyme specificity was 61-fold greater towards NAD^+ than NADP^+ (Table 1). Hence, CaIDH represents the first known monomeric NAD^+ -IDH.

Because the newly defined type II NAD^+ -IDHs and monomeric NAD^+ -IDHs stand for important IDH subfamilies, identification of one enzyme for each is not sufficient enough to support their dis-

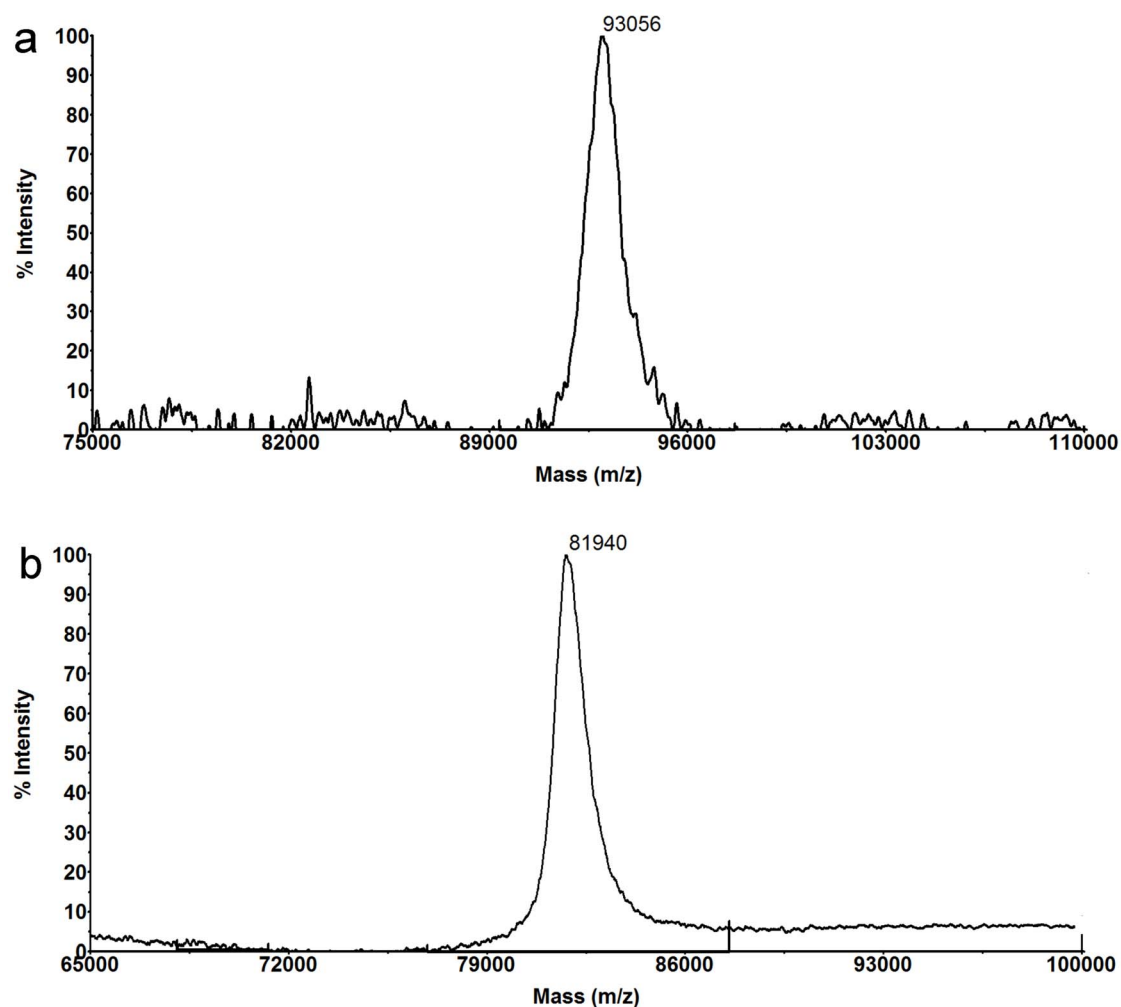


Figure 4 | MALDI-TOF/TOF MS spectra of the recombinant OIIDH (a) and CaIDH (b).



Table 1 | The kinetic parameters of the wild-type and mutant enzymes

| Enzyme | NAD ⁺ | | | NADP ⁺ | | | Specificity | |
|---|-------------------------|--------------------------------------|---|-------------------------|--------------------------------------|---|-------------|-------|
| | K_m (μM) | k_{cat} (s^{-1}) | k_{cat}/K_m (A) ($\mu\text{M}^{-1} \text{s}^{-1}$) | K_m (μM) | k_{cat} (s^{-1}) | k_{cat}/K_m (B) ($\mu\text{M}^{-1} \text{s}^{-1}$) | (A/B) | (B/A) |
| OIIDH | 136.6 ± 13.6 | 60.6 ± 3.8 | 0.444 | 2211 ± 75 | 10.0 ± 1.6 | 0.0045 | 99 | 0.009 |
| MiIDH | 126 ± 5.3 | 22.5 ± 2 | 0.179 | 1827 ± 42.4 | 1.4 ± 0.2 | 0.0008 | 224 | 0.004 |
| CaIDH | 28.9 ± 4.5 | 7.0 ± 0.8 | 0.242 | 513.2 ± 14.9 | 1.9 ± 0.1 | 0.004 | 61 | 0.017 |
| CcIDH | 74.2 ± 8.2 | 10.8 ± 1.4 | 0.146 | 475.9 ± 6.5 | 2 ± 0.2 | 0.004 | 37 | 0.027 |
| OIIDH R ³²⁶ H ³²⁷ | 2941.3 ± 75.8 | 29.1 ± 7.9 | 0.01 | 8.2 ± 1.9 | 1.8 ± 0.2 | 0.22 | 0.045 | 22 |
| CaIDH H ⁵⁸⁴ R ⁵⁹⁵ | 492.8 ± 52.5 | 9.5 ± 1.8 | 0.0193 | 11.4 ± 4.6 | 4.2 ± 1.0 | 0.368 | 0.052 | 19 |

tinctiveness. We therefore characterized another two NAD-IDHs that belong to these two novel subfamilies, respectively. MiIDH, an OIIDH analog from *Micromonas* sp. RCC299, and CcIDH, a CaIDH analog from *Campylobacter curvus*, were also produced in *E. coli* and purified to homogeneity. Kinetic analysis showed that both MiIDH and CcIDH are NAD⁺-specific as expected, because their preference for NAD⁺ was 224-fold and 37-fold over NADP⁺, respectively (Table 1). The conformation of the other two NAD⁺-specific members from type II IDH and monomeric IDH subfamilies further validated the novelty of our finding.

Coenzyme Binding Site. To evaluate the significance of the putative coenzyme-determining sites (Asp326 and Met327 in OIIDH, Leu584 and Asp595 in CaIDH), each mutant enzyme containing two point mutations, R³²⁶H³²⁷ for OIIDH and H⁵⁸⁴R⁵⁹⁵ for CaIDH, was constructed, based on the protein sequence alignment (Fig. 2). The mutated enzymes were successfully produced in *E. coli* and purified to homogeneity. CD spectroscopy was performed to determine the secondary structure of wild-type and mutant enzymes (Fig. 5). The results show that OIIDH R³²⁶H³²⁷ and CaIDH H⁵⁸⁴R⁵⁹⁵ mutants were very similar to the wild-type enzyme, thus indicating that mutations that occur at key activity sites do not cause significant changes in protein secondary structure.

The kinetic characterization results are reported in Table 1. The OIIDH R³²⁶H³²⁷ mutant displayed a 22-fold higher K_m value for NAD⁺ than that of the wild-type enzyme. Meanwhile, the mutant enzyme showed a greatly increased affinity to NADP⁺, as demonstrated by a 270-fold decrease in K_m value. The k_{cat}/K_m of OIIDH R³²⁶H³²⁷ towards NADP⁺ was 49-fold higher than that of the wild-type enzyme, whereas the k_{cat}/K_m for NAD⁺ underwent a 44-fold decrease. Consequently, the overall specificity of the OIIDH R³²⁶H³²⁷ mutant was 22-fold greater for NADP⁺ than that for NAD⁺.

Therefore, the two point mutations in OIIDH completely altered its coenzyme specificity, which demonstrates that Asp326 and Met327 were key specificity determinants for OIIDH.

The importance of Leu584 and Asp595 in the direct binding of NAD⁺ to CaIDH was also confirmed by the mutagenesis study. The CaIDH H⁵⁸⁴R⁵⁹⁵ mutant's affinity to NAD⁺ was loosened, as evidenced by a 17-fold elevation in K_m , as compared to that of the wild-type enzyme. By contrast, the mutant enzyme displayed a 45-fold decrease in K_m for NADP⁺. The k_{cat}/K_m of CaIDH H⁵⁸⁴R⁵⁹⁵ towards NADP⁺ was 92-fold higher than that of the wild-type enzyme, whereas the k_{cat}/K_m for NAD⁺ underwent a 13-fold decrease. Consequently, the overall specificity of the CaIDH H⁵⁸⁴R⁵⁹⁵ mutant was 19-fold greater for NADP⁺ than that for NAD⁺. Thus, the monomeric NAD⁺-specific CaIDH was converted to an NADP⁺-dependent enzyme by two mutations in the coenzyme binding sites.

Biochemical Characterization. The effects of pH on OIIDH and CaIDH activities were determined for the NAD⁺-linked reaction. OIIDH exhibited slightly different pH activity profiles and optimum pH using Mg²⁺ or Mn²⁺ as its cofactor. The optimum pH for OIIDH was pH 9.0 or pH 8.5 in the presence of Mg²⁺ or Mn²⁺, respectively (Fig. 6a). Furthermore, the optimum pH for CaIDH was pH 8.0 or pH 7.5 in the presence of Mg²⁺ or Mn²⁺, respectively (Fig. 6d). OIIDH showed similar pH-activity correlation with the homodimeric NAD⁺-specific *Z. mobilis* IDH (8.5 with Mg²⁺ and 8.0 with Mn²⁺) and *Streptococcus suis* IDH (7.0 with Mg²⁺ and 8.5 with Mn²⁺) from the type I subfamily^{6,39}. CaIDH showed a slightly lower optimum pH when compared to the monomeric NADP-IDH from *Streptomyces lividans* TK54 (9.0 with Mg²⁺ and 8.5 with Mn²⁺)²².

The optimum reaction temperature for OIIDH was around 40°C with either Mg²⁺ or Mn²⁺ as the cofactor (Fig. 6b). Results from heat

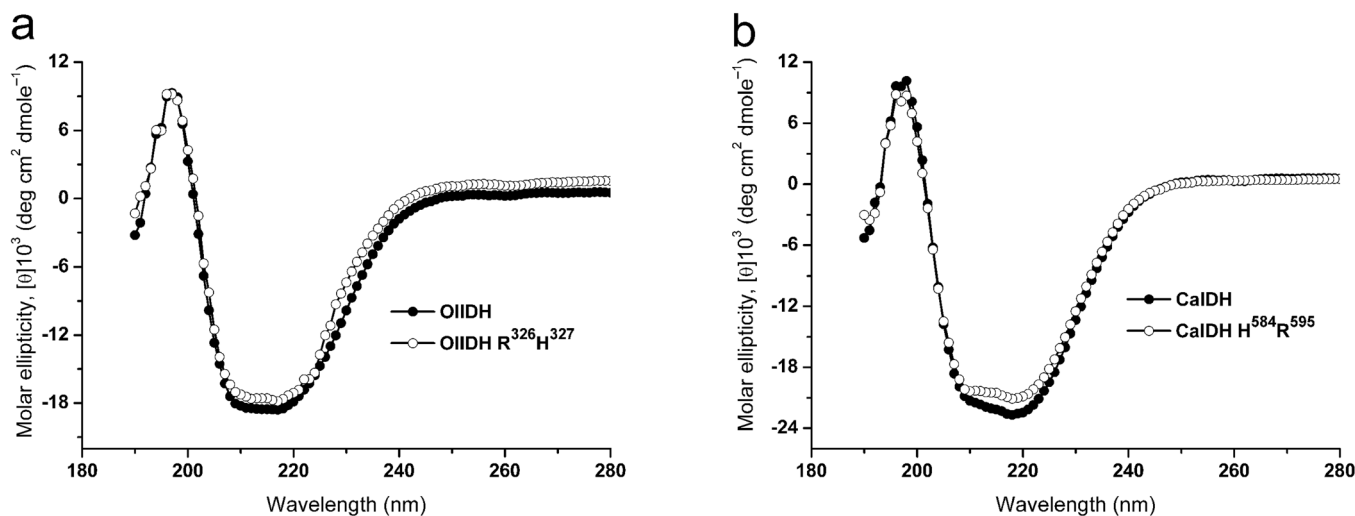


Figure 5 | Circular dichroism (CD) spectra of the wild-type OIIDH, CaIDH and their mutants, OIIDH R³²⁶H³²⁷ and CaIDH H⁵⁸⁴R⁵⁹⁵. The CD was measured and the molar ellipticity was calculated as described in Methods section.

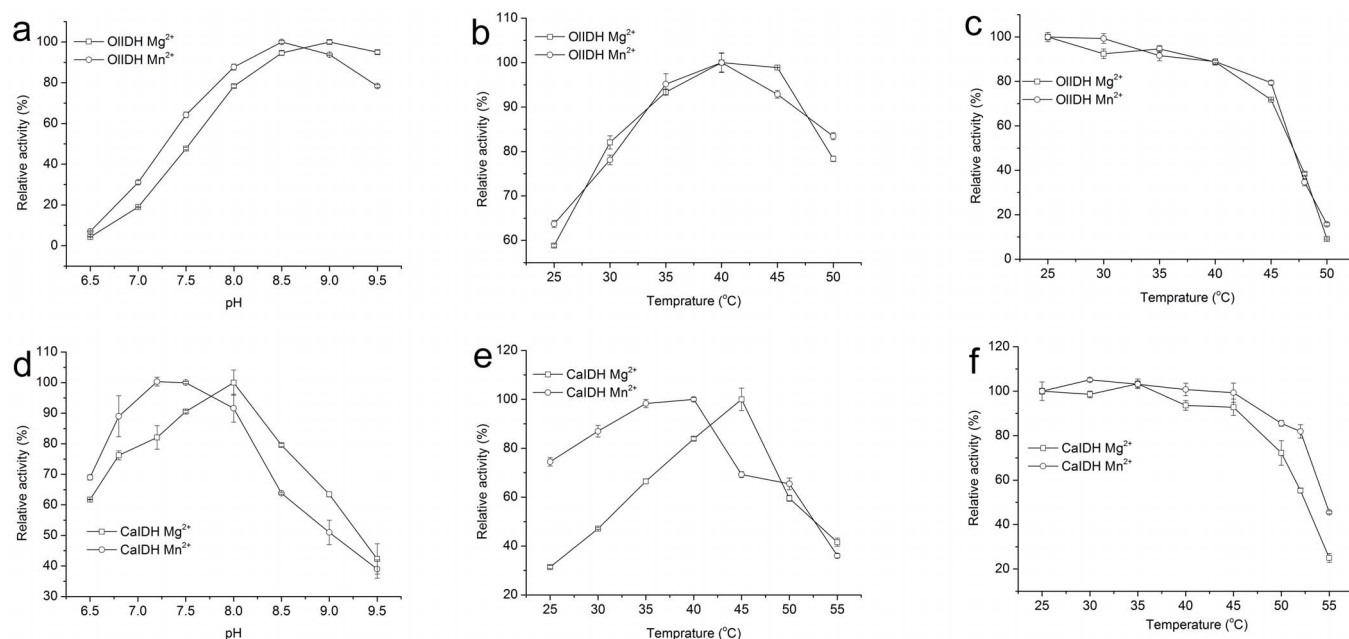


Figure 6 | Effects of pH and temperature on the activity of the recombinant OIIDH and CaIDH. (a) and (d), Effects of pH on the activity of OIIDH and CaIDH from pH 6.5 to 9.5 in the presence of Mg^{2+} (\square) and Mn^{2+} (\circ), respectively. (b) and (e), Effects of temperature on activity of OIIDH from 25°C to 50°C and CaIDH from 25°C to 55°C in the presence of Mg^{2+} (\square) and Mn^{2+} (\circ), respectively. (c) and (f), Heat-inactivation profiles of the recombinant OIIDH from 25°C to 50°C and CaIDH from 25°C to 55°C in the presence of Mg^{2+} (\square) and Mn^{2+} (\circ), respectively.

inactivation studies demonstrate that recombinant OIIDH was stable below 45°C, but its activity rapidly declined as the temperature was raised. Incubation at 45°C for 20 min caused a 28% or 21% loss of activity in the presence of Mg^{2+} or Mn^{2+} (Fig. 6c), respectively, whereas incubation at 50°C caused a 91% or 84% loss of activity in the presence of Mg^{2+} or Mn^{2+} , respectively (Fig. 6c). The optimum temperature for CaIDH activity was around 45°C or 40°C in the presence of Mg^{2+} or Mn^{2+} , respectively (Fig. 6e). Recombinant CaIDH retained the majority of the activity below 45°C. However, its activity dropped rapidly as the temperature was raised. Incubation

at 55°C for 20 min caused a 55% or 75% loss of activity in the presence of Mg^{2+} or Mn^{2+} , respectively (Fig. 6f).

The effects of different metal ions on the activities of OIIDH and CaIDH were examined (Table 2). Both enzymes needed the presence of a divalent cation for catalysis, although fractions of activities were observed for OIIDH (5.8%) and CaIDH (7.9%) when no metal ions were added. Mn^{2+} was the most favorable cation for both OIIDH and CaIDH, and its role could be largely replaced by Mg^{2+} (68.6% for OIIDH and 78.3% for CaIDH). Mn^{2+} has also been found to be the preferred cation for other homodimeric NAD-IDHs from

Table 2 | Effects of metal ions on the activity of the recombinant OIIDH and CaIDH

| Metal ions | Relative activity (%) OIIDH | Relative activity (%) CaIDH |
|---------------------|-----------------------------|-----------------------------|
| None | 5.8 ± 1.4 | 7.9 ± 2.7 |
| Mn^{2+} | 100.0 ± 3.2 | 100.0 ± 3.7 |
| Mg^{2+} | 68.6 ± 0.2 | 78.3 ± 2.8 |
| Co^{2+} | 2.8 ± 0.1 | 35.5 ± 0.4 |
| Ca^{2+} | 13.6 ± 2.1 | 0 |
| Zn^{2+} | 0 | 0 |
| Cu^{2+} | 0 | 0 |
| Ni^{2+} | 4.1 ± 0.2 | 8.2 ± 0.9 |
| K^+ | 8.7 ± 1.4 | 1.5 ± 0.1 |
| Na^+ | 8.7 ± 0.6 | 3.6 ± 0.2 |
| Rb^+ | 11.0 ± 0.9 | 2.1 ± 1.8 |
| Li^+ | 8.7 ± 0.3 | 1.3 ± 0.5 |
| Mn^{2+} | 100 ± 3.2 | 100 ± 3.7 |
| $Mn^{2+} + Mg^{2+}$ | 90.8 ± 0.8 | 83.9 ± 2.1 |
| $Mn^{2+} + Co^{2+}$ | 20.5 ± 0.1 | 50.9 ± 2.4 |
| $Mn^{2+} + Ca^{2+}$ | 89.1 ± 2.9 | 36.4 ± 0.4 |
| $Mn^{2+} + Zn^{2+}$ | 0 | 0 |
| $Mn^{2+} + Cu^{2+}$ | 0 | 22.6 ± 2.6 |
| $Mn^{2+} + Ni^{2+}$ | 61.0 ± 3.5 | 75.2 ± 3.0 |
| $Mn^{2+} + K^+$ | 96.7 ± 4.1 | 105.9 ± 3.7 |
| $Mn^{2+} + Na^+$ | 94.4 ± 0.2 | 101.4 ± 0.9 |
| $Mn^{2+} + Rb^+$ | 88.8 ± 0.5 | 102.3 ± 1.4 |
| $Mn^{2+} + Li^+$ | 91.2 ± 2.6 | 109.4 ± 1.4 |



Streptococcus mutans, *S. suis* and *Z. mobilis*^{6,39,40}. In addition, Mn²⁺ has been determined as the most favored metal ion for monomeric NADP-IDHs^{22,40}. OIIDH was completely inactivated by either 2 mM Cu²⁺ or Zn²⁺, whereas Co²⁺ and Ni²⁺ caused a 79.5% and 39% loss in activity, respectively. Other metal ions, such as Ca²⁺, Na⁺, and K⁺, slightly affected OIIDH activity in the presence of Mn²⁺. CaIDH was eliminated by 2 mM Zn²⁺ and severely down regulated by Co²⁺ (49%), Ca²⁺ (74%), and Cu²⁺ (77%). Monovalent ions (K⁺, Na⁺, Rb⁺, Li⁺) slightly improved CaIDH activity.

Discussion

In the present study, four novel IDHs, belong to two novel IDH subfamilies, were reported for the first time. Two of them were the NAD⁺-specific homodimeric IDHs from marine alga, *O. lucimarinus* (OIIDH) and *Micromonas* sp. RCC299 (MiIDH), and the other two were the NAD⁺-specific monomeric IDH from pathogens *Campylobacter* sp. FOBR14 (CaIDH) and *Campylobacter curvus* (CcIDH). OIIDH and MiIDH were found to be the first NAD-IDHs in the type II subfamily, as all members of this subfamily were previously thought to be NADP⁺ specific. CaIDH and CcIDH, however, were found to be the first NAD⁺ specific monomeric IDHs. These four IDHs, together with their NAD⁺ specific counterparts, constituted two separate branches on the phylogenetic tree. Most previous studies have proposed that IDH favors NADP⁺ over NAD⁺ as a coenzyme^{22,25,27,31,34,40}. However, as more and more NAD-IDHs from diverse backgrounds being reported, NAD⁺ appears to be widely used by IDH through nature^{6,26,39,41}.

By adding two groups of OIIDH-like and CaIDH-like NAD-IDHs, we have expanded and refined the evolutionary classification of the IDH protein family. The phylogenetic tree presented in Fig. 1 contains three well-supported monophyletic groups: type I IDH, type II IDH, and monomeric IDH. Monomeric IDHs were also included in the present phylogenetic analysis for the integrity of IDH family, and the same principle was applied by Delbaere et al.²⁰. The most important advancement was the discovery of type II homodimeric NAD-IDHs and monomeric NAD-IDHs, which completed the classification of the IDH protein family in the view of coenzyme specificity. Neither of these groups of NAD-IDH have been reported in previous studies^{4,23–25}. IDHs with NAD⁺ specificity are ancestral to NADP-IDHs, and this evolutionary hypothesis has been demonstrated by experimental reverse evolution, which was applied to the typical type *E. coli* NADP-IDH^{4,23,35}. The findings of NAD⁺-specific OIIDH and CaIDH, together with their homologous proteins, will help identify the possible ancestors of type II IDHs and monomeric IDHs, respectively.

By searching the genome of *O. lucimarinus* CCE9901 (GenBank Assembly ID: GCA_000092065.1), one copy of *IDH* gene could be found, which suggests that OIIDH is likely the only functional IDH isozyme in this marine algae. The genus *Ostreococcus* is composed of a group of globally distributed, photosynthetic, unicellular green algae, and these cells are the smallest known eukaryotes⁴². Because all known NAD-IDHs from eukaryotes are hetero-oligomeric and consist of at least two different subunits^{3,43}, OIIDH represents the first eukaryotic homodimeric NAD-IDH. Furthermore, IDHs from some other marine algae, such as *Micromonas* sp. RCC299, *Emiliania huxleyi* and *Thalassiosira oceanica* (see Supplementary Table S1 online), showed high homology to OIIDH (>70%) and branched together with OIIDH on the phylogenetic tree. Considering the antiquity of oceanic algae, the ancient trait of NAD⁺ specificity that was possessed by IDHs in these organisms can be fairly explained.

Similar to OIIDH, the monomeric CaIDH is likely the only active IDH isozyme found in *Campylobacter* sp. FOBR14 (GenBank Assembly ID: GCA_000287855.1). The monomeric NAD-IDH group seemed to be very small in composition, because only IDHs from the genus *Campylobacter* were included (see Supplementary Table S1 online). These monomeric IDHs shared very high sequence

identity (>60%), and the putative NAD⁺-binding sites were conservative (Fig. 2). *Campylobacter* is the most common cause of bacterial foodborne illness and has drawn a lot of attention in recent years^{44,45}. As a pathogen, it is surprising that the glyoxylate bypass is absent in *Campylobacter* sp. FOBR14, as no genes encoding isocitrate lyase and malate synthase can be found in its genome. The glyoxylate bypass ensures the bypass of two oxidative steps of the TCA cycle and permits the net incorporation of carbon during the growth of most microorganisms on acetate or fatty acids as the primary carbon source. The end products of the bypass can be used for gluconeogenesis and other biosynthetic processes⁴⁶. Most intracellular human pathogens, such as *Salmonella typhimurium*⁴⁷, *Burkholderia pseudomallei*⁴⁸ and *M. tuberculosis*⁴⁹, need the glyoxylate bypass for their virulence, because fatty acids are the only abundant sources of C2 carbon in mammalian tissues⁵⁰. Interestingly, the existence of the glyoxylate bypass in microorganisms has always been accompanied by at least one NADP-IDH isozyme, which provides the majority of NADPH to support bacterial growth on limited carbon sources⁴. Because the glyoxylate bypass could not be detected in *Campylobacter* sp. FOBR14, it is understandable that an NAD⁺-specific IDH, rather than an NADP⁺-specific IDH, was found in this organism. The finding also suggests that the infection mechanism of *Campylobacter* sp. FOBR14 may be different from that of pathogens with the glyoxylate bypass.

As an eukaryotic NAD-IDH, OIIDH shared very similar kinetic properties with hetero-oligomeric NAD-IDHs from other eukaryotic cells. The K_m value for NAD⁺ of OIIDH (136.6 μM) was in a similar range of those determined for *Yarrowia lipolytica* yeast (136 μM)⁵¹ and rats (148.9 μM)⁵². It was higher than that for humans (70 μM)⁵³ and lower than that for budding yeast (210 μM)⁵⁴. When compared to homodimeric NADP-IDHs of type II subfamily, OIIDH showed much lower affinity to its coenzyme than NADP-IDHs do to NADP⁺, such as in the wild pig (5.6 μM)⁵⁵, rat (11.5 μM)⁵⁶ and budding yeast (20 μM)⁵⁷. Due to the decrease in cofactor affinity, OIIDH has much lower k_{cat}/K_m (0.444 $\mu\text{M}^{-1} \text{s}^{-1}$) than its type II NADP-IDH counterparts (5.96 $\mu\text{M}^{-1} \text{s}^{-1}$ for wild pig and 9.1 $\mu\text{M}^{-1} \text{s}^{-1}$ for rat)^{55,57}. The poor performance of OIIDH in catalysis may be understood as a latent ancient phenotype, thereby providing more evidence for the age of OIIDH among the type II subfamily. The comparison of kinetic parameters shows that the preference of CaIDH for NAD⁺ over NADP⁺ (61-fold) was significantly lower than that of monomeric NADP-IDHs, such as *S. lividans* TK54 (85,000-fold)²² and *C. glutamicum* (50,000-fold)²⁷, thus making CaIDH an old and ineffective enzyme in using NAD⁺. Both OIIDH and CaIDH represent ancient members in the type II and monomeric subfamily, respectively. The modern, sophisticated type II homodimeric NADP-IDH and monomeric NADP-IDH are very possibly refined from old NAD⁺-utilizing ancestors through evolution, as partially evidenced by the fact that just two point mutations in the coenzyme-binding sites of OIIDH and CaIDH were sufficient in converting them to NADP⁺-utilizing enzymes (Table 1).

Conclusion

In the present study, we refined and expanded the phylogenetic classification of the IDH protein family by dividing the type I, type II, and monomeric subfamilies and identifying two new groups: one group of type II NAD-IDHs, represented by OIIDH, and one monomeric NAD-IDHs, represented by CaIDH. Thus, the classification of the IDH protein family in the view of coenzyme specificity is now complete. OIIDH and CaIDH were heterologously produced and enzymatically characterized in detail. Although the NAD⁺ specificity of the two enzymes was confirmed by kinetic analysis, both enzymes were demonstrated to be ineffective NAD⁺-utilizing enzymes. The coenzyme specificity of both enzymes could be completely altered from NAD⁺ to NADP⁺ by merely mutating two coenzyme-binding amino acids, thus suggesting the ancestral positions of OIIDH and



CaIDH in the type II and monomeric subfamilies, respectively. Further studies are clearly needed to understand these two novel groups of IDH in the areas of structure determination and catalytic mechanism investigation.

Methods

Gene Synthesis. IDHs from *Ostreococcus lucimarinus* CCE9901 (OIIDH, GenBank accession number: ABP01147), *Micromonas* sp. RCC299 (MiIDH, GenBank accession number: XP_002502450), *Campylobacter* sp. FOBR14 (CaIDH, GenBank accession number: EJP74315) and *Campylobacter curvus* (CcIDH, GenBank accession number: WP_018136314) were the four targets of this study. Full-length genes encoding these four proteins were synthesized through the gene synthesis service by Generey Biotech Co., Ltd. (Shanghai, China). The coding sequences for four genes were codon optimized by selecting only the most preferential codons according to the *Escherichia coli* bias. The artificial genes were then inserted into the expression vector, pET-28b (+), between the *Nde* I and *Xho* I sites, thus generating four recombinant plasmids, pET-OIIDH, pET-MiIDH, pET-CaIDH and pET-CcIDH. The gene sequences were confirmed by sequencing.

Site-Directed Mutagenesis. Point mutations (Asp326Arg and Met327His) were introduced into OIIDH by overlap extension, PCR-based, site-directed mutagenesis. The oligonucleotides that were used to generate the OIIDH mutant were as follows: forward, 5'-CACGGACAGGCCACCGTCATTATCTGCGGTATCTC-3', and reverse, 5'-GAGATACCGCAGATAATGACCGTGGGCCGTGCCGTG-3'. The underlined codons are mutated sequences. The mutated version of CaIDH harboring the Leu584His and Asp594Arg mutations was constructed by two rounds of site-directed mutagenesis. The first round of mutation introduced the Leu584His change, using the following oligonucleotides: forward, 5'-GCGGCACTGCTCCGATGCAC-GCACGCGATATGATCG-3', and reverse, 5'-CGATCATATCGCGTGCCTGCAT-CGGAGCAGTCCCGC-3'. The second round of mutation introduced the Asp594Arg change, using the following oligonucleotides: forward, 5'-CGAGCA-ACCACCTGCGTTGGGATAGCCTGGGCGAG-3', and reverse, 5'-CTCGC-CCAGGCTATCCCAACGCGAGTGGTTGCTCG-3'. The underlined codons are mutated sequences. Mutated genes were then inserted into the expression vector, pET-28b (+), and confirmed by DNA sequencing.

Overexpression and Purification of Wild-type and Mutated Enzymes. *E. coli* Rosetta (DE3) cells were transformed with pET-OIIDH, pET-MiIDH, pET-CaIDH, pET-CcIDH, or recombinant plasmids carrying the mutated IDH genes and grown at 37°C with vigorous shaking in LB medium containing 30 µg/ml kanamycin and 25 µg/ml chloramphenicol. Then, cells were inoculated in 50 ml fresh LB media with the same antibiotic. When the OD₆₀₀ of the culture reached 0.6, isopropyl-1-thio-β-D-galactopyranoside was added to the culture at a final concentration of 0.1 mM with subsequent cultivation overnight at 20°C. Cells were harvested by centrifugation at 4,000 rpm for 15 min and then resuspended in lysis buffer (50 mM Tris-HCl, pH 7.5, 500 mM NaCl). The insoluble debris was removed by centrifugation at 12,000 g for 20 min at 4°C. Then, enzymes with 6×His-tag were purified by using BD TALON Metal Affinity Resins (Clontech, USA), according to the manufacturer's instructions. The expression abundance and purification homogeneity were verified by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE).

Enzyme Assay. The activities of wild-type and mutant enzymes were assayed by a modification of the method by Cvitkovich et al.⁵⁸. Activity assays were carried out in 25°C 1-ml cuvettes (1-cm light path) containing 35 mM Tris-HCl buffer (pH 7.5), 2 mM MgCl₂ or MnCl₂, 1.5 mM DL-isocitrate, and 1.0 mM NAD⁺. The increase in NADH was monitored at 340 nm with a thermostated Cary 300 UV-Vis spectrophotometer (Varian, USA), using a molar extinction coefficient of 6.22 mM⁻¹cm⁻¹. One unit of enzyme activity represented the reduction of 1 µM of NAD⁺ per minute. Protein concentrations were determined using the Bio-Rad protein assay kit (Bio-Rad, USA) with bovine serum albumin as the standard.

Kinetic Analysis. To measure the Michaelis constant (K_m) values of the wild-type and mutant enzymes for NAD⁺ and NADP⁺, the isocitrate concentration was kept fixed at 1.0 mM with varying cofactor concentrations. Apparent maximum velocity (V_{max}) and K_m values were calculated by nonlinear regression using Prism 5.0 (Prism, USA). All kinetic parameters were obtained from at least three measurements.

Temperature and pH Effects. The effects of temperature and pH on the activity of recombinant OIIDH and CaIDH were carried out using the assay method described above. The activities of purified recombinant OIIDH and CaIDH were assayed in 35 mM Tris-HCl buffer between pH 6.5 and 9.5 in the presence of Mn²⁺ (Mg²⁺). The optimum temperature was determined at temperatures that ranged from 25°C–55°C. The thermostability of recombinant OIIDH and CaIDH through heat inactivation were determined by incubating enzyme aliquots at 25°C–50°C for 20 min. After incubation, the aliquots were immediately cooled on ice, and the residual enzyme activity was measured by using the standard enzyme assay.

Metal Ion Effects. The effects of different metal ions on the activities of recombinant OIIDH and CaIDH were determined using the standard assay method, including

2 mM monovalent ions (K⁺, Li⁺, Na⁺, and Rb⁺) and divalent ions (Ca²⁺, Co²⁺, Cu²⁺, Mg²⁺, Mn²⁺, Ni²⁺, and Zn²⁺).

Size Exclusion Chromatography (SEC). The molecular masses of recombinant OIIDH and CaIDH were detected by SEC on a HiLoad™ 10/300 Superdex 200 column (Amersham Biosciences, Germany), which was equilibrated with 0.05 M potassium phosphate buffer (pH 7.0) containing 0.15 M NaCl and 0.01% sodium azide. Protein standards for calibrating gels were ovalbumin (45 kDa), conalbumin (75 kDa), aldolase (158 kDa), ferritin (440 kDa), and thyroglobulin (669 kDa).

Circular Dichroism Spectroscopy. Circular dichroism (CD) spectroscopy was conducted using a Jasco model J-810 spectropolarimeter (Oklahoma City, OK, USA). The ellipticity measurements, as a function of wavelength, were performed as described previously⁵⁹. Briefly, purified protein samples (0.3 mg/mL) were prepared in 50 mM sodium phosphate and 60 mM NaCl (pH 7.5). The ellipticity (θ) was obtained by averaging three scans of the enzyme solution between 200 and 260 nm at 0.5-nm increments. The mean molar ellipticity [θ] (deg cm² dmole⁻¹) was calculated from $[\theta] = \theta/10nCl$, where θ was the measured ellipticity (millidegrees), C was the molar concentration of protein, l was the cell path length in centimeters (0.1 cm), and n was the number of residues per subunit of enzyme (415 for OIIDH and the mutant, 737 for CaIDH and the mutant).

MALDI-TOF/TOF mass spectrometry. Mass spectrometry analyses were conducted using an AB SCIEX MALDI TOF-TOF 5800 Analyzer (AB SCIEX, USA) equipped with a neodymium: yttrium-aluminum-garnet laser (laser wavelength was 349 nm), in linear high mass positive-ion mode. SA was used as the matrix and TFA was applied as an ionization auxiliary reagent. The TOF/TOF calibration mixtures (AB SCIEX) were used to calibrate the spectrum to a mass tolerance within 10 ppm. The MS spectra were processed using TOF-TOF Series Explorer software (V4.0, AB SCIEX).

Sequence Alignments and Phylogenetic Analysis. The X-ray structures of *M. tuberculosis* NADP⁺-IDH1 (MtIDH, 4HCX) and *Azotobacter vinelandii* (AvIDH, 1J1W) were downloaded from the Protein Data Bank database (<http://www.rcsb.org/pdb/>). The homology models of OIIDH and CaIDH were generated by the SWISS-MODEL server (<http://swissmodel.expasy.org>). The structure-based amino acid sequence alignment was conducted with the Clustal X program (<ftp://ftp.ebi.ac.uk/pub/software/clusterw2>) and ESPript 3.0 web tool (<http://esprict.ibcp.fr/ESPrict/ESPrict/>)^{60,61}.

The typical IDH was used as bait to identify similar IDH sequences in protein database by performing BLAST Link search (<http://www.ncbi.nlm.nih.gov/sutils/blink.cgi?mode=query>). The redundancy of the query results was eliminated by keeping one IDH sequence for each species, while removing all the other identical IDH sequences for the same species. IDH sequences with relative high homology to the bait sequence were taken for phylogenetic analysis, and their coenzyme binding sites were evaluated by sequence alignment in the first place, in order to confirm their identical coenzyme specificity with the query IDH. Other IDH sequences among the search results were discarded either because their coenzyme usages were different from the bait IDH as predicted by coenzyme binding sites alignment, or because their sequence identities with the query IDH were relatively low, which may suggest their different distribution on the phylogenetic tree. By applying this principle, 197 IDH sequences in total, representing the eleven subgroups encompassed by the three IDH subfamilies, were chosen for phylogenetic analysis. IDH sequences from diverse resources were downloaded from GenBank via the National Center for Biotechnology Information web site (<http://www.ncbi.nlm.nih.gov/>). The bootstrapped neighbor-joining tree was constructed with the MEGA 6 software (<http://www.megasoftware.net/>), based on the sequence alignment by Clustal X program (<ftp://ftp.ebi.ac.uk/pub/software/clusterw2>)^{60,62}. In order to improve the accuracy of the phylogenetic analysis, some other tree building algorithms were also employed, such as UPGMA method, Maximum Likelihood method and Minimum Evolution method. One outgroup of malate dehydrogenase (MDH) was tried into the phylogenetic study in order to examine whether the IDH tree will be disturbed by adding different proteins.

- Amend, J. P. & Shock, E. L. Energetics of amino acid synthesis in hydrothermal ecosystems. *Science* **281**, 1659–1662 (1998).
- Melendez-Hevia, E., Waddell, T. G. & Cascante, M. The puzzle of the Krebs citric acid cycle: assembling the pieces of chemically feasible reactions, and opportunism in the design of metabolic pathways during evolution. *J. Mol. Evol.* **43**, 293–303 (1996).
- Taylor, A. B., Hu, G., Hart, P. J. & McAlister-Henn, L. Allosteric motions in structures of yeast NAD⁺-specific isocitrate dehydrogenase. *J. Biol. Chem.* **283**, 10872–10880, doi:10.1074/jbc.M708719200 (2008).
- Zhu, G., Golding, G. B. & Dean, A. M. The selective cause of an ancient adaptation. *Science* **307**, 1279–1282, doi:10.1126/science.1106974 (2005).
- Jo, S. H. et al. Control of mitochondrial redox balance and cellular defense against oxidative damage by mitochondrial NADP⁺-dependent isocitrate dehydrogenase. *J. Biol. Chem.* **276**, 16168–16176, doi:10.1074/jbc.M010120200 (2001).
- Wang, P., Jin, M. & Zhu, G. Biochemical and molecular characterization of NAD⁺-dependent isocitrate dehydrogenase from the ethanologenic bacterium



- Zymomonas mobilis. *FEMS Microbiol. Lett.* **327**, 134–141, doi:10.1111/j.1574-6968.2011.02467.x (2012).
7. Jin, M. M. *et al.* Biochemical characterization of NADP(+)-dependent isocitrate dehydrogenase from *Microcystis aeruginosa* PCC7806. *Mol. Biol. Rep.* **40**, 2995–3002, doi:10.1007/s11033-012-2371-8 (2013).
 8. Chan, M. & Sim, T. S. Recombinant *Plasmodium falciparum* NADP-dependent isocitrate dehydrogenase is active and harbours a unique 26 amino acid tail. *Exp. Parasitol.* **103**, 120–126 (2003).
 9. Quartararo, C. E., Hazra, S., Hadi, T. & Blanchard, J. S. Structural, kinetic and chemical mechanism of isocitrate dehydrogenase-1 from *Mycobacterium tuberculosis*. *Biochemistry* **52**, 1765–1775, doi:10.1021/bi400037w (2013).
 10. Zhao, X., Wang, P., Zhu, G., Wang, B. & Zhu, G. Enzymatic characterization of a type II isocitrate dehydrogenase from pathogenic *Leptospira interrogans* serovar Lai strain 56601. *Appl. Biochem. Biotechnol.* **172**, 487–496, doi:10.1007/s12010-013-0521-7 (2014).
 11. Banerjee, S. *et al.* *Mycobacterium tuberculosis* (Mtb) isocitrate dehydrogenases show strong B cell response and distinguish vaccinated controls from TB patients. *Proc Natl Acad Sci U S A* **101**, 12652–12657, doi:10.1073/pnas.0404347101 (2004).
 12. Hussain, M. A. *et al.* Isocitrate dehydrogenase of *Helicobacter pylori* potentially induces humoral immune response in subjects with peptic ulcer disease and gastritis. *PLoS One* **3**, e1481, doi:10.1371/journal.pone.0001481 (2008).
 13. Parsons, D. W. *et al.* An integrated genomic analysis of human glioblastoma multiforme. *Science* **321**, 1807–1812, doi:10.1126/science.1164382 (2008).
 14. Mardis, E. R. *et al.* Recurring mutations found by sequencing an acute myeloid leukemia genome. *N Engl J Med* **361**, 1058–1066, doi:10.1056/NEJMoa0903840 (2009).
 15. Dang, L. *et al.* Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **462**, 739–744, doi:10.1038/nature08617 (2009).
 16. Ward, P. S. *et al.* The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. *Cancer Cell* **17**, 225–234, doi:10.1016/j.ccr.2010.01.020 (2010).
 17. Lu, C. *et al.* IDH mutation impairs histone demethylation and results in a block to cell differentiation. *Nature* **483**, 474–478, doi:10.1038/nature10860 (2012).
 18. Turcan, S. *et al.* IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature* **483**, 479–483, doi:10.1038/nature10866 (2012).
 19. Reitman, Z. J. *et al.* Cancer-associated Isocitrate Dehydrogenase 1 (IDH1) R132H Mutation and d-2-Hydroxyglutarate Stimulate Glutamine Metabolism under Hypoxia. *J. Biol. Chem.* **289**, 23318–23328, doi:10.1074/jbc.M114.575183 (2014).
 20. Imabayashi, F., Aich, S., Prasad, L. & Delbaere, L. T. Substrate-free structure of a monomeric NADP isocitrate dehydrogenase: an open conformation phylogenetic relationship of isocitrate dehydrogenase. *Proteins* **63**, 100–112, doi:10.1002/prot.20867 (2006).
 21. Yasutake, Y. *et al.* Crystal structure of the monomeric isocitrate dehydrogenase in the presence of NADP+: insight into the cofactor recognition, catalysis, and evolution. *J. Biol. Chem.* **278**, 36897–36904, doi:10.1074/jbc.M304091200 (2003).
 22. Zhang, B. *et al.* Enzymatic characterization of a monomeric isocitrate dehydrogenase from *Streptomyces lividans* TK54. *Biochimie* **91**, 1405–1410, doi:10.1016/j.biochi.2009.07.011 (2009).
 23. Dean, A. M. & Golding, G. B. Protein engineering reveals ancient adaptive replacements in isocitrate dehydrogenase. *Proc Natl Acad Sci U S A* **94**, 3104–3109 (1997).
 24. Steen, I. H., Lien, T. & Birkeland, N. K. Biochemical and phylogenetic characterization of isocitrate dehydrogenase from a hyperthermophilic archaeon, *Archaeoglobus fulgidus*. *Arch. Microbiol.* **168**, 412–420 (1997).
 25. Steen, I. H. *et al.* Comparison of isocitrate dehydrogenase from three hyperthermophiles reveals differences in thermostability, cofactor specificity, oligomeric state, and phylogenetic affiliation. *J. Biol. Chem.* **276**, 43924–43931, doi:10.1074/jbc.M105999200 (2001).
 26. Stokke, R. *et al.* Biochemical characterization of isocitrate dehydrogenase from *Methylococcus capsulatus* reveals a unique NAD+ dependent homotetrameric enzyme. *Arch. Microbiol.* **187**, 361–370, doi:10.1007/s00203-006-0200-y (2007).
 27. Chen, R. & Yang, H. A highly specific monomeric isocitrate dehydrogenase from *Corynebacterium glutamicum*. *Arch Biochem Biophys* **383**, 238–245, doi:10.1006/abbi.2000.2082 (2000).
 28. Vinekar, R., Verma, C. & Ghosh, I. Functional relevance of dynamic properties of Dimeric NADP-dependent Isocitrate Dehydrogenases. *BMC Bioinformatics* **13** Suppl 17, S2, doi:10.1186/1471-2105-13-S17-S2 (2012).
 29. Kumar, S. M. *et al.* Crystal structure studies of NADP+ dependent isocitrate dehydrogenase from *Thermus thermophilus* exhibiting a novel terminal domain. *Biochem. Biophys. Res. Commun.* **449**, 107–113, doi:10.1016/j.bbrc.2014.04.164 (2014).
 30. Karlstrom, M. *et al.* The crystal structure of a hyperthermostable subfamily II isocitrate dehydrogenase from *Thermotoga maritima*. *FEBS J.* **273**, 2851–2868, doi:10.1111/j.1742-4658.2006.05298.x (2006).
 31. Xu, X. *et al.* Structures of human cytosolic NADP-dependent isocitrate dehydrogenase reveal a novel self-regulatory mechanism of activity. *J. Biol. Chem.* **279**, 33946–33957, doi:10.1074/jbc.M404298200 (2004).
 32. Leiros, H. K., Fedoy, A. E., Leiros, I. & Steen, I. H. The complex structures of isocitrate dehydrogenase from *Clostridium thermocellum* and *Desulfotalea psychrophila* suggest a new active site locking mechanism. *FEBS Open Bio* **2**, 159–172, doi:10.1016/j.fob.2012.06.003 (2012).
 33. Ceccarelli, C., Grodsky, N. B., Ariyaratne, N., Colman, R. F. & Bahnson, B. J. Crystal structure of porcine mitochondrial NADP+-dependent isocitrate dehydrogenase complexed with Mn2+ and isocitrate. Insights into the enzyme mechanism. *J. Biol. Chem.* **277**, 43454–43462, doi:10.1074/jbc.M207306200 (2002).
 34. Peng, Y., Zhong, C., Huang, W. & Ding, J. Structural studies of *Saccharomyces cerevisiae* mitochondrial NADP-dependent isocitrate dehydrogenase in different enzymatic states reveal substantial conformational changes during the catalytic reaction. *Protein Sci.* **17**, 1542–1554, doi:10.1110/ps.035675.108 (2008).
 35. Chen, R., Greer, A. & Dean, A. M. A highly active decarboxylating dehydrogenase with rationally inverted coenzyme specificity. *Proc Natl Acad Sci U S A* **92**, 11666–11670 (1995).
 36. Imada, K. *et al.* Structure and quantum chemical analysis of NAD+-dependent isocitrate dehydrogenase: hydride transfer and co-factor specificity. *Proteins* **70**, 63–71, doi:10.1002/prot.21486 (2008).
 37. Folta-Stogniew, E. & Williams, K. R. Determination of molecular masses of proteins in solution: Implementation of an HPLC size exclusion chromatography and laser light scattering service in a core laboratory. *Journal of biomolecular techniques* : *JBT* **10**, 51–63 (1999).
 38. Arakawa, T., Ejima, D., Li, T. & Philo, J. S. The critical role of mobile phase composition in size exclusion chromatography of protein pharmaceuticals. *J. Pharm. Sci.* **99**, 1674–1692, doi:10.1002/jps.21974 (2010).
 39. Wang, P. *et al.* Enzymatic characterization of isocitrate dehydrogenase from an emerging zoonotic pathogen *Streptococcus suis*. *Biochimie* **93**, 1470–1475, doi:10.1016/j.biochi.2011.04.021 (2011).
 40. Wang, A. *et al.* Heteroexpression and characterization of a monomeric isocitrate dehydrogenase from the multicellular prokaryote *Streptomyces avermitilis* MA-4680. *Mol. Biol. Rep.* **38**, 3717–3724, doi:10.1007/s11033-010-0486-3 (2011).
 41. Wang, P., Song, P., Jin, M. & Zhu, G. Isocitrate dehydrogenase from *Streptococcus mutans*: biochemical properties and evaluation of a putative phosphorylation site at Ser102. *PLoS One* **8**, e58918, doi:10.1371/journal.pone.0058918 (2013).
 42. Lanier, W., Moustafa, A., Bhattacharya, D. & Comeron, J. M. EST analysis of *Ostreococcus lucimarinus*, the most compact eukaryotic genome, shows an excess of introns in highly expressed genes. *PLoS One* **3**, e2171, doi:10.1371/journal.pone.0002171 (2008).
 43. Hu, G., Lin, A. P. & McAlister-Henn, L. Physiological consequences of loss of allosteric activation of yeast NAD+-specific isocitrate dehydrogenase. *J. Biol. Chem.* **281**, 16935–16942, doi:10.1074/jbc.M512281200 (2006).
 44. Deshpande, N. P. *et al.* Sequencing and validation of the genome of a *Campylobacter concisus* reveals intra-species diversity. *PLoS One* **6**, e22170, doi:10.1371/journal.pone.0022170 (2011).
 45. Deshpande, N. P., Kaakoush, N. O., Wilkins, M. R. & Mitchell, H. M. Comparative genomics of *Campylobacter concisus* isolates reveals genetic diversity and provides insights into disease association. *BMC Genomics* **14**, 585, doi:10.1186/1471-2164-14-585 (2013).
 46. Dunn, M. F., Ramirez-Trujillo, J. A. & Hernandez-Lucas, I. Major roles of isocitrate lyase and malate synthase in bacterial and fungal pathogenesis. *Microbiology* **155**, 3166–3175, doi:10.1099/mic.0.030858-0 (2009).
 47. Tchawa Yimga, M. *et al.* Role of gluconeogenesis and the tricarboxylic acid cycle in the virulence of *Salmonella enterica* serovar Typhimurium in BALB/c mice. *Infect. Immun.* **74**, 1130–1140, doi:10.1128/IAI.74.2.1130-1140.2006 (2006).
 48. van Schaik, E. J., Tom, M. & Woods, D. E. Burkholderia pseudomallei isocitrate lyase is a persistence factor in pulmonary melioidosis: implications for the development of isocitrate lyase inhibitors as novel antimicrobials. *Infect. Immun.* **77**, 4275–4283, doi:10.1128/IAI.00609-09 (2009).
 49. Munoz-Elias, E. J. & McKinney, J. D. *Mycobacterium tuberculosis* isocitrate lyases 1 and 2 are jointly required for in vivo growth and virulence. *Nat. Med.* **11**, 638–644, doi:10.1038/nm1252 (2005).
 50. McKinney, J. D. *et al.* Persistence of *Mycobacterium tuberculosis* in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. *Nature* **406**, 735–738, doi:10.1038/35021074 (2000).
 51. Morgunov, I. G., Solodovnikova, N. Y., Sharyshev, A. A., Kamzolova, S. V. & Finogenova, T. V. Regulation of NAD(+)-dependent isocitrate dehydrogenase in the citrate producing yeast *Yarrowia lipolytica*. *Biochemistry (Mosc)* **69**, 1391–1398 (2004).
 52. Qi, F., Chen, X. & Beard, D. A. Detailed kinetics and regulation of mammalian NAD-linked isocitrate dehydrogenase. *Biochim. Biophys. Acta* **1784**, 1641–1651, doi:10.1016/j.bbapap.2008.07.001 (2008).
 53. Soundar, S., O'Hagan, M., Fomulu, K. S. & Colman, R. F. Identification of Mn2+-binding aspartates from alpha, beta, and gamma subunits of human NAD-dependent isocitrate dehydrogenase. *J. Biol. Chem.* **281**, 21073–21081, doi:10.1074/jbc.M602956200 (2006).
 54. Dean, A. M. A molecular investigation of genotype by environment interactions. *Genetics* **139**, 19–33 (1995).
 55. Huang, Y. C. & Colman, R. F. Location of the coenzyme binding site in the porcine mitochondrial NADP-dependent isocitrate dehydrogenase. *J. Biol. Chem.* **280**, 30349–30353, doi:10.1074/jbc.M505828200 (2005).
 56. Jennings, G. T., Minard, K. I. & McAlister-Henn, L. Expression and mutagenesis of mammalian cytosolic NADP+-specific isocitrate dehydrogenase. *Biochemistry* **36**, 13743–13747, doi:10.1021/bi970916r (1997).



57. Lu, Q., Minard, K. I. & McAlister-Henn, L. Dual compartmental localization and function of mammalian NADP⁺-specific isocitrate dehydrogenase in yeast. *Arch Biochem Biophys* **472**, 17–25, doi:10.1016/j.abb.2008.01.025 (2008).
58. Cvitkovitch, D. G., Gutierrez, J. A. & Bleiweis, A. S. Role of the citrate pathway in glutamate biosynthesis by *Streptococcus mutans*. *J. Bacteriol.* **179**, 650–655 (1997).
59. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411–2423, doi:10.1002/pro.5560041120 (1995).
60. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948, doi:10.1093/bioinformatics/btm404 (2007).
61. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–324, doi:10.1093/nar/gku316 (2014).
62. Tamura, K., Stecher, G., Peterson, D., Filipowski, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729, doi:10.1093/molbev/mst197 (2013).

Acknowledgments

This research was supported by the National High Technology Research and Development Program (“863” Program: 2012AA02A708), the National Natural Science Foundation of China (31170005; 31400003), the Natural Science Foundation of Anhui Province of China

(1308085QC67), and the Provincial Project of Natural Science Research for Colleges and Universities of Anhui Province of China (KJ2013A128).

Author contributions

P.W. and C. L. performed the experiments. P.W. and G. Z. planned the project. P. W. and G. Z. wrote the manuscript. All authors reviewed the manuscript.

Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Wang, P., Lv, C. & Zhu, G. Novel Type II and Monomeric NAD⁺ Specific Isocitrate Dehydrogenases: Phylogenetic Affinity, Enzymatic Characterization, and Evolutionary Implication. *Sci. Rep.* **5**, 9150; DOI:10.1038/srep09150 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>