



OPEN

Molecular evolutionary insight of structural zinc atom in yeast xylitol dehydrogenases and its application in bioethanol production by lignocellulosic biomass

Kentaroh Yoshiwara^{1,5}, Seiya Watanabe^{1,2,3,5}✉ & Yasunori Watanabe⁴

Xylitol dehydrogenase (XDH) catalyzes the NAD⁺-dependent oxidization of xylitol into D-xylulose, and belongs to a zinc-dependent medium-chain dehydrogenase/reductase family. This protein family consists of enzymes with one or two zinc atoms per subunit, among which catalytic zinc is necessary for the activity. Among many XDHs from yeast and fungi, XDH from *Pichia stipitis* is one of the key enzymes for bioethanol production by lignocellulosic biomass, and possesses only a catalytic zinc atom. Despite its importance in bioindustry, a structural data of XDH has not yet been available, and little insight into the role of a second zinc atom in this protein family is known. We herein report the crystal structure of XDH from *P. stipitis* using a thermostabilized mutant. In the refined structure, a second zinc atom clearly coordinated with four artificially introduced cysteine ligands. Homologous mutations in XDH from *Saccharomyces cerevisiae* also stabilized and enhanced activity. The substitution of each of the four cysteine ligands with an aspartate in XDH from *Schizosaccharomyces pombe* contributed to the significantly better maintenance of activity and thermostability than their substitution with a serine, providing a novel hypothesis for how this zinc atom was eliminated.

The utilization of ethanol produced from plant biomass (so-called “bioethanol”), which is derived from the fixation of atmospheric CO₂, as an industrial carbon source and car fuel is one of the most important research issues for the realization of a sustainable global environment. Bioethanol is mainly produced from agricultural crop biomass, the biological fermentation of which is easy, but it commercially competes as food and animal feed resources. Alternatively, “lignocellulosic biomass”, such as woods and agricultural residues, represents an attractive feedstock that consists of cellulose (60% of the mass), hemicellulose (30%), and lignin (10%). Hemicellulose comprises pentoses, such as D-xylose (and L-arabinose), as well as hexoses, and D-xylose accounts for approximately 25% of the total sugar content in lignocellulosic biomass^{1,2}.

Yeast, particularly *Saccharomyces cerevisiae*, has long been used to produce alcoholic beverages because of its ability to produce high concentrations of ethanol and its high inherent tolerance of ethanol. However, native strains cannot ferment D-xylose as a carbon source. Therefore, many studies have attempted to overcome the limitations associated with the utilization of D-xylose by introducing its metabolic pathway from other microorganisms^{3,4}. Although the biological degradation of D-xylose in microorganisms is classified into phosphorylated and non-phosphorylated pathways^{5,6}, only the former, which is further classified into two different pathways, is used for this purpose. In the “isomerase pathway”, D-xylose is directly converted into D-xylulose by D-xylose isomerase (XI; EC 5.3.1.5) without any cofactors (Fig. 1a). Although this pathway mostly operates in

¹Department of Bioscience, Graduate School of Agriculture, Ehime University, 3-5-7 Tarumi, Matsuyama, Ehime 790-8566, Japan. ²Faculty of Agriculture, Ehime University, 3-5-7 Tarumi, Matsuyama, Ehime 790-8566, Japan. ³Center for Marine Environmental Studies (CMES), Ehime University, 2-5 Bunkyo-Cho, Matsuyama, Ehime 790-8577, Japan. ⁴Faculty of Science, Yamagata University, 1-4-12 Kojirakawa-Machi, Yamagata, Yamagata 990-8560, Japan. ⁵These authors contributed equally: Kentaroh Yoshiwara and Seiya Watanabe. ✉email: irab@agr.ehime-u.ac.jp

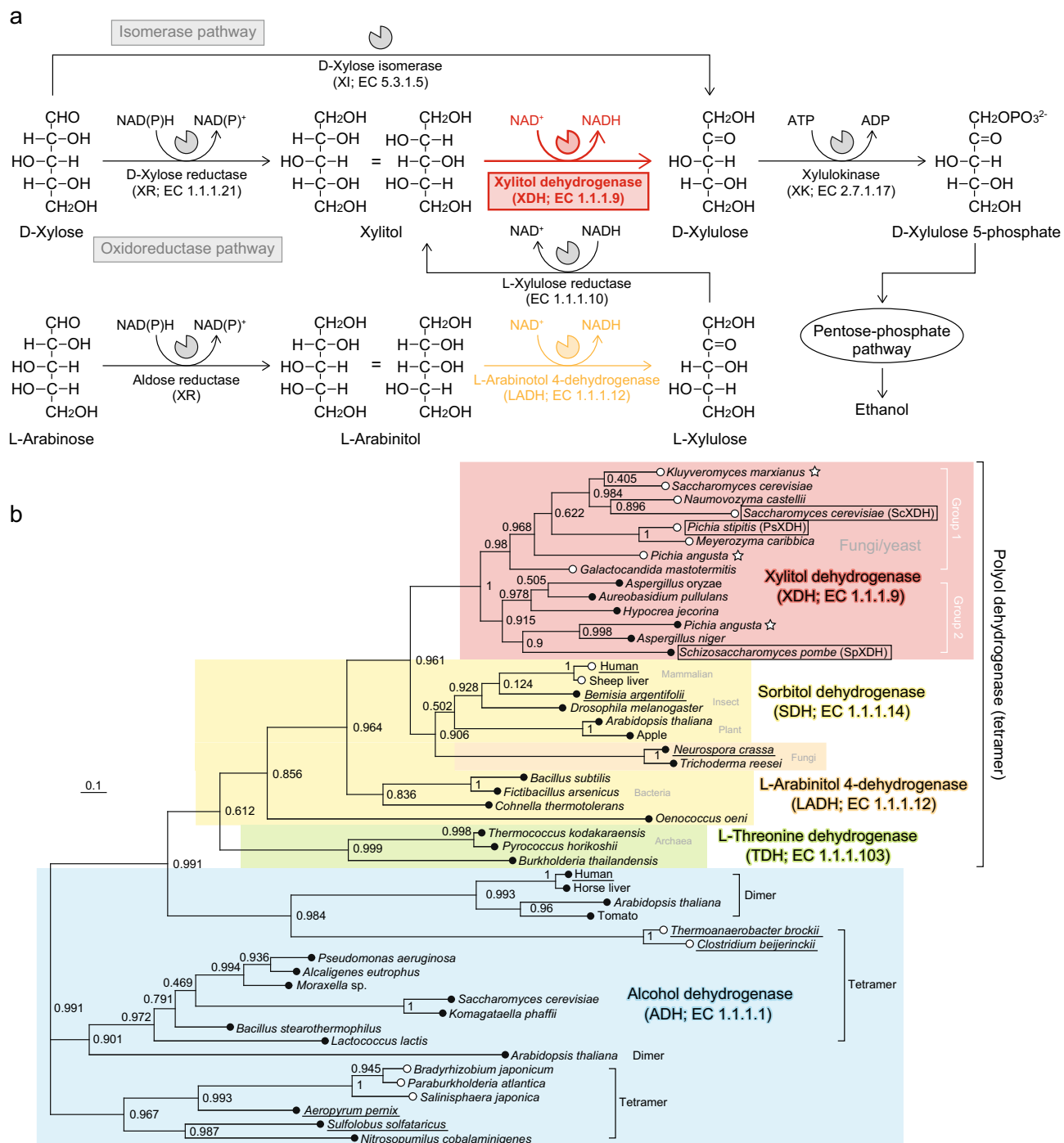


Figure 1. D-Xylose metabolism by yeast and fungi. **(a)** Metabolic network of D-xylose and L-arabinose. There are two routes of D-xylose metabolism, with XDH being involved in the oxidoreductase pathway. This pathway partially overlaps with L-arabinose metabolism, in which LADH belongs to the same protein family as XDH, described in **(b)**. **(b)** A phylogenetic tree of the zinc-dependent MDR superfamily group. The number on each branch indicates the bootstrap value. Subfamilies of XDH, SDH, LADH, TDH, and ADH are colored in red, yellow, orange, yellow-green, and cyan, respectively. Open and closed circles at the end of each branch are enzymes in the absence and presence of structural zinc, respectively. XDHs with asterisks are from thermotolerant yeast. Proteins in the box were purified and characterized in the present study. Underlined proteins were used for discussions in the text.

bacteria, a few fungi possess the bacterial type of XI⁷. Alternatively, in the “oxidoreductase pathway”, D-xylose reductase (XR; EC 1.1.1.21) catalyzes the reduction of the C1 carbonyl group of D-xylose, yielding xylitol as

the product (Fig. 1a). Xylitol is then oxidized by xylitol dehydrogenase (XDH; EC 1.1.1.9) to give D-xylulose. Xylulokinase (XK; EC 2.7.1.17) commonly phosphorylates D-xylulose into D-xylulose 5-phosphate, which is metabolized further via the pentose-phosphate pathway. Although *S. cerevisiae* possesses the endogenous oxidoreductase pathway consisting of YHR104w (GRE3), YLR070c (XYL2), and YGR194c (XKS1; XK) as XR, XDH, and XK, respectively, the rate of D-xylose metabolism by strains overexpressing them has not yet reached industrially competitive levels⁸. Alternatively, XR and XDH genes from the native D-xylose-metabolizing yeast *Pichia stipites* (*Scherratomyces stipites*; PsXR and PsXDH) are mostly used together with the endogenous XK gene.

XDHs from yeast and fungi, including PsXDH, belong to a polyol dehydrogenase (PDH) subfamily in the zinc-dependent group of the medium-chain dehydrogenase/reductase (MDR) superfamily^{9,10}, together with sorbitol dehydrogenase (SDH; EC 1.1.1.14)^{10–12} and L-arabinol 4-dehydrogenase (LADH; EC 1.1.1.12) from several organisms^{13–15} (Fig. 1b). Among them, LADH is involved in L-arabinose metabolism by fungi, which partially overlaps with the oxidoreductase pathway of D-xylose (Fig. 1a). All MDR enzymes utilize NAD⁺(H) or NADP⁺(H) as a cofactor, and one zinc atom with catalytic functions (so-called “catalytic zinc”) is present in the active center of PDH, alcohol dehydrogenase (ADH; EC 1.1.1.1)^{16,17}, and L-threonine dehydrogenase (TDH; EC 1.1.1.103)¹⁸. Many PDHs and ADHs also have a second zinc atom (so-called “structural zinc”), which generally coordinates with four cysteine ligands. Since it is impossible to remove this zinc atom without the loss of stable folding or enzyme activity by site-directed mutagenesis^{19,20}, “why” and “how” MDR enzymes appeared in the evolutionary stage in the absence of structural zinc remain unclear.

Although XR-XDH achieves higher metabolic fluxes than XI, the excretion of xylitol occurs during D-xylose fermentation by *S. cerevisiae*²¹. An intercellular redox imbalance due to the different coenzyme specificities of XR (with NADPH) and XDH (with NAD⁺) has been suggested as one of the main factors^{4,5}. Furthermore, a relationship has been reported between stability and intercellular expression levels. Therefore, protein engineering to change (modify) coenzyme specificity and/or increase thermostability represents an attractive approach. A unique NADP⁺(H)-dependent SDH from insects¹² is used as a reference enzyme, through which the complete reversal of coenzyme specificity towards NADP⁺ is achieved²². On the other hand, the introduction of four cysteine residues provided additional zinc-binding sites and significantly increased thermostability²².

Despite its importance in bioindustry, a crystallographic analysis of PsXDH has not yet been performed, and there is currently no structural evidence for protein engineering, particularly the artificial introduction of structural zinc. We herein report for the first time the crystal structure of PsXDH using a thermostabilized mutant. A second zinc atom coordinated with four artificially introduced cysteine ligands. The substitution of each of the four cysteine ligands with an aspartate in XDH from *Schizosaccharomyces pombe* contributed to the significantly better maintenance of activity and thermostability than their substitution with a serine, providing a novel hypothesis for how enzymes in the absence of structural zinc, such as PsXDH, appeared.

Results

Overall structure of the PsXDH_{C4} mutant. The so-called PsXDH_{C4} mutant was previously constructed by substituting Ser96, Ser99, and Tyr102 with cysteine residues in the wild-type (WT) enzyme (Fig. 2)²². The crystal structure of the apo-form of PsXDH_{C4} was refined at 2.80 Å resolution after molecular replacement using the coordinates of SDH from sheep liver (PDB ID; 3QE3) as a research model¹¹. Data collection and refinement statistics are summarized in Table 1. Each monomer contained a bidomain architecture composed of a large “catalytic domain” (Thr2-Val163 and Arg300-Pro362) and a smaller “coenzyme-binding domain” (Gly164-Phe299), with a large cleft separating them (Fig. 3a). The former consisted of an α/β fold with a similar structure to that in other MDR enzymes, and the latter had the characteristic α/β Rossmann fold. The region corresponding to Ser121–Gly124 within a long loop between α3 and β7 in the catalytic domain was not built owing to invisible density map.

Analytical size-exclusion chromatographic studies demonstrated that PsXDH_{C4} formed a homotetramer in solution. The four subunits (A–D) in the biological tetramer generated from crystallographic 222 symmetry were regarded as a dimer of identical A/B and C/D dimers (Fig. 3b). The contact region between each monomer of the dimer was strand β14, which formed a seven parallel β-sheet (β11–β10–β9–β12–β13–β15–β14), and packed antiparallel with strand β14 of the second monomer. During the formation of the tetramer, the subunits A/B dimer contacted the subunits C/D dimer at two regions. Region 1 corresponded to a loop that protruded from the catalytic domain, at which a zinc atom other than the catalytic zinc atom bound, as described below (Fig. 3c). The side chain of Arg97 in subunit A (D) formed a salt bridge with the side chain of Asp141 in the same subunit A (D), which also interacted with the side chain Lys103 in subunit D (A) of the other dimer; symmetrical interactions are shown in parentheses. Region 2 was located between two loops of the coenzyme-binding domain in subunit A (C), and one helix of the catalytic domain in subunit C (A). Therefore, these two types of contact were achieved by different subunits in the opposing dimers.

As described in “Introduction”, the zinc-dependent MDR superfamily group is phylogenetically classified into three subclasses; PDH, ADH, and TDH (Fig. 1b). As expected, higher (structural-based) sequence homologies were noted in the same PDHs (r.m.s.d. values of 1.3–2.2 Å and sequence identities of 36–42%) than in ADHs and TDHs (r.m.s.d. values of 2.1–3.5 Å and sequence identities of 21–31%) (Fig. 4 and Supplementary Table S1). PDHs and TDHs, and ADHs from bacteria and archaea are homotetramer with the same crystallographic symmetry as PsXDH_{C4} (Fig. 3d). On the other hand, ADHs from eukaryote and plant tend to be homodimer with two subunits (Fig. 1b), which corresponds to the A/B (or C/D) dimer of tetrameric MDR enzymes; therefore, neither regions 1 nor 2 contribute to formation of the tertiary structure (Fig. 3e).

Catalytic zinc binding in the PsXDH_{C4} mutant. An anomalous difference Fourier map using data collected at the K-edge of zinc (wavelength of 1.280 Å) showed two clear peaks within the electron density map of

Enzyme	Source organisms	Four cysteine ligands			
XDH	<i>P. stipites</i>	Ser96	Ser99	Tyr102	Cys110
XDH	<i>K. marxianus</i>	Ser97	Ser100	Thr103	Cys111
XDH	<i>S. cerevisiae</i>	Asp99	Ser102	Met105	Asp113
XDH	<i>M. morgani</i>	Asp93	Ser96	Ser99	Asp107
XDH	<i>S. pombe</i>	Cys97	Cys100	Cys103	Cys111
SDH*	<i>B. argentifolii</i>	Cys96	Cys98	Cys102	Cys110
SDH	Human	Arg99	Asp102	Cys105	Ser113
SDH	Sheep liver	Arg98	Asp101	Cys104	Ser112
LADH*	<i>N. crassa</i>	Cys108	Cys111	Cys114	Cys122
ADH*	Horse liver	Cys97	Cys100	Cys103	Cys111
ADH*	<i>S. solfataricus</i>	Glu98	Cys101	Cys104	Cys112
AraDH*	<i>S. solfataricus</i>	Asp94	Cys97	Cys100	Cys108
ADH*	<i>A. pernix</i>	Asp109	Cys112	Cys115	Cys123
ADH	<i>T. brockii</i>	Asp89	Thr92	Val95	Ser103
ADH	<i>C. beijerinckii</i>	Asp89	Ser92	Val95	Ser103
(ADH)	<i>O. oeni</i>	Arg94	Cys97	Cys100	Cys108
(ADH)	<i>P. atlantica</i>	Ser110	Cys113	Cys116	Ala124

Figure 2. Four cysteine ligands for structural zinc binding in MDR enzymes. Cysteine, serine, acidic, and basic residues are shadowed in gray, yellow-green, red, and blue, respectively. Proteins with asterisks possess structural zinc based on a crystallographic analysis.

PsXDH_{C4} (Fig. 5a,b). Among them, one peak, located at the bottom of the catalytic domain, corresponded to catalytic zinc, based on inference from other related structures (Fig. 4). Catalytic zinc was coordinated by interactions with Cys41 (distance of 2.3 Å), His66 (2.3 Å), Glu67 (2.2 Å), and a water molecule (Wat6; 2.6 Å) (Fig. 5a). The carboxyl group of Glu67 further formed a hydrogen bond with the side chain of Lys356, and nearby Glu159 was linked to the zinc atom through the wat6 molecule. All of these residues (and neighboring Ser43 and Asp44) were highly conserved in PDH members. XDH from the yeast *Galactocandida mastotermitis* (54% sequence homology with PsXDH) contained ~6 Mg²⁺ ions, which were selectively removed by dialysis without a loss of activity²³. No magnesium ions were found in the crystal structure of the apo-form of PsXDH_{C4}.

Second zinc binding in the PsXDH_{C4} mutant. As described above, PsXDH_{C4} contained another peak in the anomalous difference Fourier map at the K-edge of zinc (Fig. 5b). Since binding sites were artificially constructed, another metal may have (concomitantly) been present; cobalt may have been alternatively introduced into the same binding sites of catalytic zinc as the native enzyme of ADH¹⁶. On the other hand, in previous biochemical study, the zinc content of the PsXDH_{C4} mutant was estimated to be ~1.9 mol of zinc/mol of subunit, which was close to 2.0, by atomic absorption spectrophotometry²². Furthermore, since zinc compounds were not used in the purification or crystallization protocols, we concluded that the metal must be zinc and intrinsically contained in the protein. To the best of our knowledge, this is the first structural evidence for artificially introducing structural zinc into MDR superfamily enzymes.

The second zinc atom was bound within a loop that protruded from the catalytic domain of PsXDH_{C4} (Fig. 3a), at which it was ligated by the enzyme residues Cys96, Cys99, Cys102, and Cys110 (distance of 2.3 Å) (Fig. 5b). Superimposition to the crystal structures of other MDR enzymes revealed that this zinc atom was equivalent to (inherently bound) structural zinc (Fig. 4b), and there was no significant difference in their binding loops regardless of zinc (Fig. 5c–g). SDH from humans and ADH from *Thermoanaerobacter brockii* possessed a salt bridge and/or hydrogen bond inside and/or outside of this loop (Fig. 5c,e).

When Phe98 and Glu101, neighboring residues of four cysteine ligands, were both substituted to arginine and phenylalanine residues in the PsXDH_{C4} mutant, respectively, the resultant double mutant (C4/F98R/E101F) further increased thermostability (Fig. 6a)²⁴. Mutated Arg98 in subunit A may have formed a salt bridge with Asp141 in subunit D of the other dimer, and mutated Phe101 hydrophobically interacted with Leu109 within the same loop to significantly enhance dimer-dimer interaction(s) for the formation of tetramers (double-headed dashed arrow) (Fig. 3c).

Introduction of four cysteine ligands for structural zinc in ScXDH. As described in “Introduction”, the YLR070c protein from *S. cerevisiae* (45% sequence homology with PsXDH; Fig. 1b) functions as XDH in the endogenous D-xylose pathway^{8,25}, and possessed none of the four cysteine ligands (Fig. 2); therefore, we designed it the D99C/S102C/M105C/D113C mutant (referred to as ScXDH_{C4}). Each WT and C4 mutant enzyme of ScXDH (and also PsXDH) was successfully expressed in *Escherichia coli* cells and purified using the same procedure as that for PsXDH. The k_{cat}/K_m values of ScXDH_{C4} and PsXDH_{C4} for xylitol were 7.3- and 4.3-fold higher, respectively, than those of ScXDH_{WT} and PsXDH_{WT}, and these differences were attributed to 10- and

XDH C4 apo form (PDB code 7Y9P)	
Data collection	
Space group	P622
<i>a, b, c</i> (Å)	158.27, 158.27, 65.50
Wavelength (Å)	1.28000
Resolution range (Å)	50–2.80 (2.95–2.80)
Rmerge	0.457 (4.19)
Rmeas	0.461 (4.23)
CC1/2	0.997 (0.856)
<i>I</i> / σ (<i>I</i>)	16.6 (2.6)
Completeness (%)	100 (100)
Redundancy	57.24 (60.46)
Refinement	
<i>R</i> / <i>R</i> _{free}	0.223/0.274 (0.343/0.367)
No. of atoms	
Protein	2618
Zinc	2
Glycerol	6
Sulfate ion	5
Polyethylene glycol	7
Water	15
<i>B</i> -factors (Å ²)	
Protein	64.08
Zinc	87.95
Glycerol	63.15
Sulfate ion	78.36
Polyethylene glycol	66.54
Water	50.91
r.m.s. deviations	
Bond lengths (Å)/angles (°)	0.004/0.688
Ramachandran plot	
Favored/allowed/outliers (%)	97.17/2.83/0.00
Rotamer outliers (%)	2
Clashscore	8

Table 1. Data collection and refinement statistics.

22-fold higher k_{cat} values, respectively (Table 2). The inactivation of ScXDH_{C4} or PsXDH_{C4} was not detected after an incubation at 45 °C for 10 min, whereas the activities of each WT enzyme were decreased to 89 and 48%, respectively, by the same treatment (Fig. 6a). Collectively, the introduction of four cysteine ligands increased activity and thermostability in not only PsXDH, but also ScXDH, which appeared to be due to the introduction of a second zinc atom.

Serine mutants of four cysteine ligands for structural zinc in SpXDH. The hypothetical protein SPBC1773.05c from *S. pombe* (40% sequence homology with PsXDH; Fig. 1b) had four cysteine ligands at positions 97, 100, 103, and 111 (Fig. 2); therefore, we selected it as a target for the enzyme in the presence of structural zinc. The k_{cat}/K_m value of xylitol was similar to that of PsXDH (31.3 and 79.9 min⁻¹ mM⁻¹, respectively), suggesting its function as XDH (referred to as SpXDH) (Table 2). To elucidate the physiological role of structural zinc, we initially substituted each cysteine ligand with a serine residue. A gel filtration analysis using the sample purified by Ni-NTA revealed that, compared with the WT, the ratio of active molecular species with tetramer structure was significantly decreased. Therefore, we performed a kinetic analysis using the fraction with the highest activity (Supplementary Fig. S1).

Among the four serine mutants, the k_{cat}/K_m values of the C97S and C103S mutants (29.6 and 14.0 min⁻¹ mM⁻¹) were similar to that of the WT enzyme (68.5 min⁻¹ mM⁻¹), whereas that of the C100S mutant was markedly lower (0.743 min⁻¹ mM⁻¹) and the C111S mutant was completely inactive (Table 2). A heat treatment analysis indicated their significant decrease of thermostability; the half-life time for inactivation at 50 °C was estimated to be within 1 min (Fig. 6b,d), whereas the inactivation of WT was not detected by the same treatment. Similar results were observed at the optimum temperatures for activity. Collectively, these results indicated that all serine mutations markedly affected thermostability, and that the four cysteine ligands had varying impact on the levels of activity.

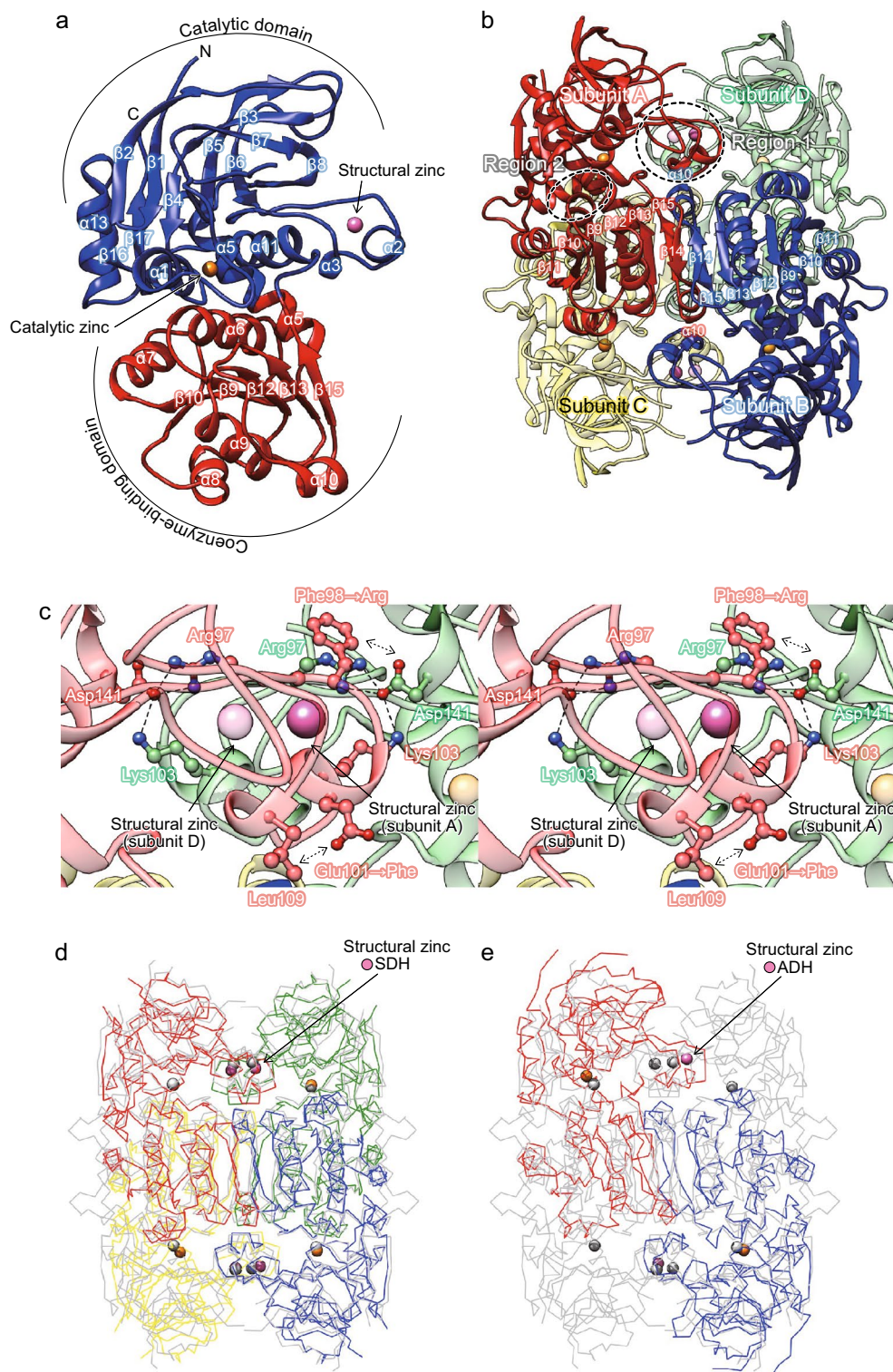


Figure 3. Crystal structure of the PsXDH_{C4} mutant. **(a)** Monomer structure. The catalytic domain housing both catalytic and structural zinc atom (orange and magenta spheres, respectively) is shown in blue, and the cofactor-binding domain is shown in red. The figure was prepared using PyMOL⁴⁹. **(b)** Structure of the biological tetramer generated from crystallographic symmetry. The tetramer is a dimer of dimers (A/B and C/D). The dashed line ellipse indicates two contact regions 1 (c) and 2 between two dimers. In (c), hydrogen bonds are shown as black dashed lines. It is likely that the mutated Arg98 in subunit A can form a salt bridge with Asp141 in subunit D of another dimer, and the mutated Phe101 interacts with Leu109 within the same loop hydrophobically. Superimposed Ca traces of the structures of tetrameric PsXDH_{C4} (gray) on tetrameric SDH from silverleaf whitefly (1E3J) **(d)** and dimeric ADH from *Arabidopsis thaliana* (2CF5) **(e)**.

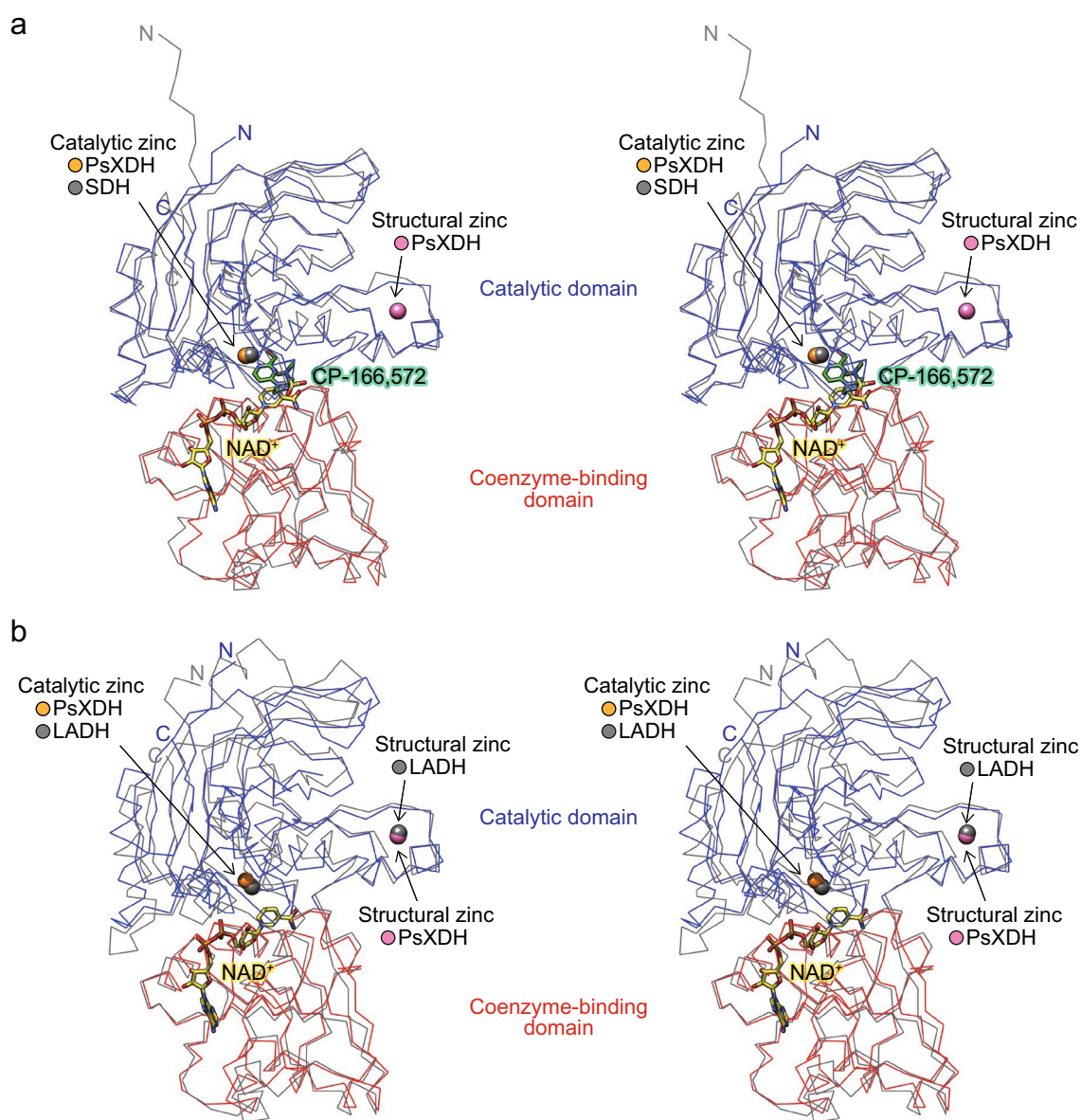


Figure 4. Stereo view of the superimposed Ca traces of the coenzyme binding domains of PsXDH_{C4} and SDH from human in complex with the competitive inhibitor (CP-16657213) and NAD⁺ (1PL6; gray) (a), or LADH from *Neurospora crassa* in complex with NAD⁺ (3M6I; gray) (b).

Other mutants of four cysteine ligands for structural zinc in SpXDH. In some (putative) ADH subfamily enzymes in the MDR superfamily, one of the four cysteine ligands was substituted with an aspartate, glutamate, or arginine residue; D-C-C-C, E-C-C-C, or R-C-C-C, respectively (Fig. 2). In the crystal structures of (hyper)thermophilic archaeal enzymes, aspartate and glutamate residues coordinated with structural zinc (Fig. 5f,g)^{26–29}. Therefore, Cys97 in SpXDH was changed to design the C97D, C97E, and C97R mutants. Among them, the courses of the thermal inactivation of the C97D and C97E mutants were enhanced (Fig. 6c).

An aspartate residue was frequently detected in some substitution patterns of the four cysteine ligands in MDR enzymes; D-S-M-D, D-S-S-D, and R-D-C-S. (Fig. 2). Therefore, Cys100, Cys103, and Cys111 in SpXDH were further substituted with an aspartate. The k_{cat}/K_m values of the C97D, C100D, and C103D mutants increased and were 57%, 23%, and 65% that of WT, respectively (Table 2). Furthermore, the C111D mutant was significantly active, which differed from the C111S mutant. In the heat treatment analysis at 50 °C, losses in activity of 27%, 50%, and 23% were observed in the C97D, C100D and C103D mutants, respectively, which were less than those in each serine mutant (91%, 100%, and 82%), and their half-live times for inactivation at 50 °C were estimated to be longer than 1 h (Fig. 6d). Similar results were obtained for the optimum temperatures for activity. Regarding the C97D and C103D mutants, samples purified by Ni-NTA contained a large amount of the active tetramer, which may have been due to thermostabilization, as described below (Supplementary Fig. S1).

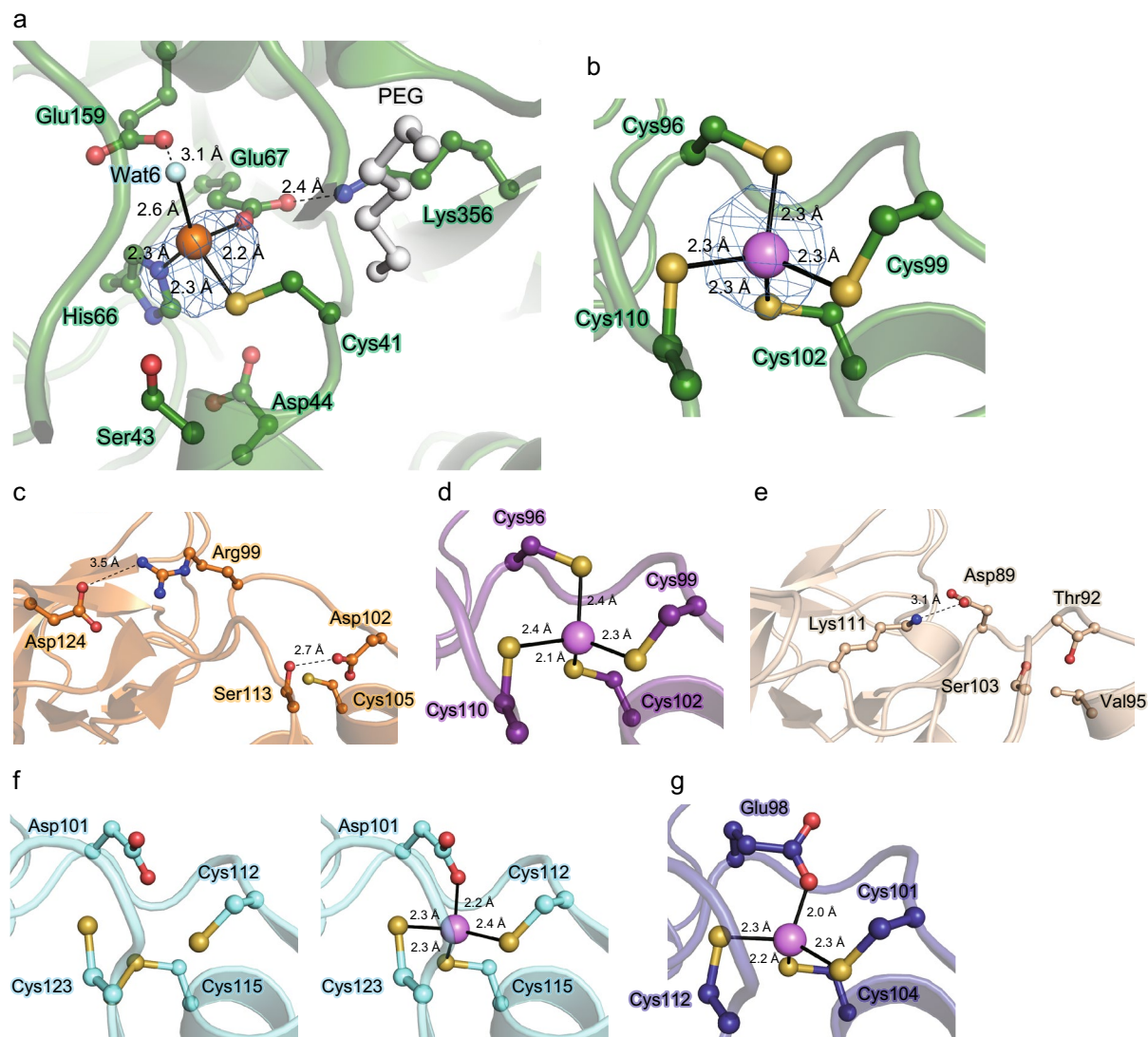


Figure 5. Zinc binding in the PsXDH_{C4} mutant. **(a)** Amino acid residues involved in chelating the catalytic zinc atom in the active site, together with some neighboring residues. PEG is derived from crystallization solution. Anomalous difference Fourier maps, contoured at 3.0 σ and 10 σ , suggest peaks for the catalytic zinc **(a)** and structural zinc atoms **(b)**, respectively, and are shown as a blue mesh. **(b)** Structural zinc binding sites, and comparisons with SDH from human (1PL6) **(c)**, SDH from silverleaf whitefly (1E3J) **(d)**, and ADHs from *T. brockii* (1YKF) **(e)**, *A. pernix* (1H2B) **(f)**, and *S. solfataricus* (1JVB) **(g)**. In **(f)**, there are two structures in the presence of zinc (right panel) and in its absence showing the disulfide bond (left panel).

To investigate the effects of the aspartate ligand in more detail, (the mutated) Cys96 in PsXDH_{C4} was further changed to an aspartate residue. The resultant C4/S96D mutant (equivalent to the S96D/S99C/Y102C mutant) exhibited similar thermotolerance to WT (Fig. 6a), whereas the k_{cat}/K_m value increased by 2.7-fold, which was caused by a marked increase in the k_{cat} value, similar to the C4 mutant (Table 2). On the other words, a change from C4 to C4/S96D in PsXDH had similar effects on WT and the C97D mutant of SpXDH, suggesting no difference of functions between artificial and inherent structural zinc.

Intracellular expression level of XDH. A Western blot analysis using an anti (His)₆-tag antibody showed that the PsXDH_{C4}, ScXDH_{C4}, and SpXDH_{WT} proteins were more highly expressed in *E. coli* cells at 37 °C than the PsXDH_{WT}, ScXDH_{WT}, and SpXDH_{C100S} proteins, respectively (Fig. 7a–c). When any of the four cysteine ligands in SpXDH was substituted with a serine residue, translational levels in *E. coli* cells at 25 °C were higher than those at 37 °C, whereas no significant difference was observed between them (Fig. 7d). Therefore, the peak with a high molecular weight in gel filtration using the sample purified by Ni-NTA appeared to be due to the aggregation of XDH (but not contaminant proteins) (Supplementary Fig. S1). Collectively, these results suggest a relationship between stability in vitro and intercellular expression levels in vivo.

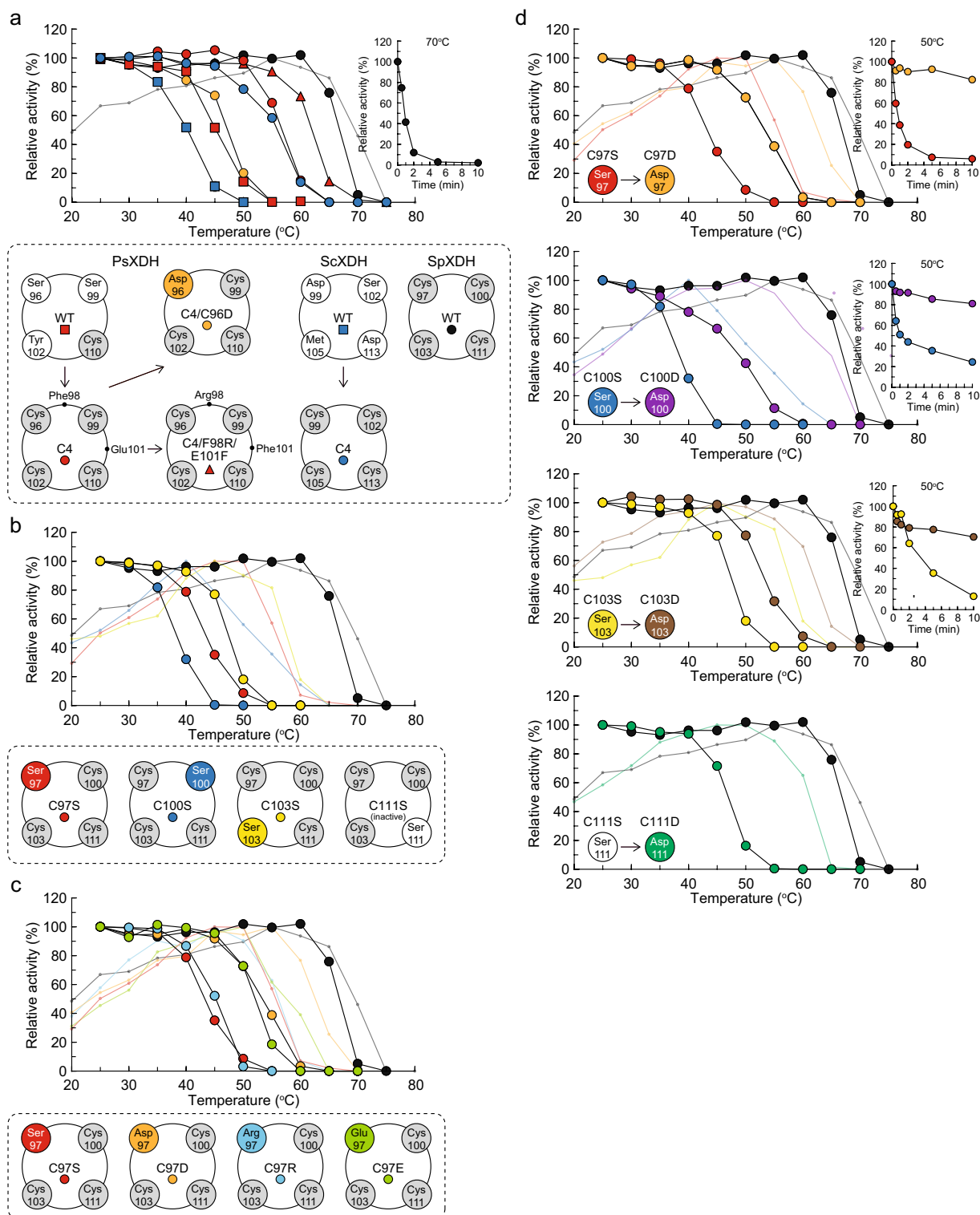


Figure 6. Courses of thermal inactivation. Purified enzymes were dialyzed against 50 mM Tris-HCl (pH 8.0) at 4 °C overnight. Dialyzed enzymes were incubated for 10 min at each temperature or at the indicated temperature for each time (inset). Enzyme activities were shown as relative average values ($n=3$) expressed as a percent of the controls without a heat treatment. The background graph shows the effects of temperature on activity. (a) WT and mutants of PsXDH and ScXDH and WT of SpXDH. (b) Four serine mutants of SpXDH. (c) Four Cys97 mutants. (d) Four aspartate mutants and comparisons with each serine mutant.

Enzyme	Mutation(s)	Pattern ^a	<i>K_m</i> (mM)	<i>k_{cat}</i> (min ⁻¹)	<i>k_{cat}/K_m</i> (min ⁻¹ mM ⁻¹)	Relative value (%) ^c	<i>T_m</i> value (°C) ^f
PsXDH	WT	S-S-Y-C	2.05 ± 0.20 ^b	37.8 ± 3.5	18.4 ± 0.2	100	45.2 (0.0) ^g
	S96C/S99C/Y102C	C-C-C-C	4.94 ± 1.62 ^b	384 ± 111	79.1 ± 4.0	430	56.8 (+ 11.6)
	S96C/S99C/Y102C	C-C-C-C	4.94 ± 1.62 ^b	384 ± 111	79.1 ± 4.0	430	56.8 (+ 11.6)
	S96D/S99C/Y102C	D-C-C-C	11.0 ± 1.2 ^c	550 ± 38	50.2 ± 1.7	270	47.2 (+ 2.0)
ScXDH	WT	D-S-N-D	1.82 ± 0.42 ^b	1.09 ± 0.19	0.605 ± 0.042	100	40.2 (0.0)
	D99C/S102C/M105C/D113C	C-C-C-C	5.44 ± 0.78 ^b	24.1 ± 3.0	4.44 ± 0.10	730	55.9 (+ 15.7)
SpXDH	WT	C-C-C-C	6.10 ± 0.34 ^c	417 ± 18	68.5 ± 1.0	100	66.8 (0.0)
	C97S	S-C-C-C	12.3 ± 2.3 ^c	362 ± 45	29.6 ± 1.6	43	43.3 (- 23.5)
	C100S	C-S-C-C	16.2 ± 5.0 ^c	12.1 ± 4.0	0.743 ± 0.030	1.1	38.2 (- 28.6)
	C103S	C-C-S-C	9.01 ± 0.74	125 ± 6	14.0 ± 0.4	20	47.3 (- 19.5)
	C111S	C-C-C-S	N.D. ^d	N.D.	N.D.	N.D.	N.D.
	C97R	R-C-C-C	8.24 ± 1.42 ^c	50.8 ± 5.7	6.21 ± 0.35	9.1	45.2 (- 21.6)
	C97E	E-C-C-C	13.4 ± 0.8 ^c	3.26 ± 0.17	0.242 ± 0.003	0.4	52.1 (- 14.7)
	C97D	D-C-C-C	10.4 ± 0.1 ^c	407 ± 3	39.0 ± 0.4	57	53.4 (- 13.4)
	C100D	C-D-C-C	8.26 ± 0.89 ^c	130 ± 9	15.8 ± 0.6	23	48.5 (- 18.3)
	C103D	C-C-D-C	10.5 ± 0.7 ^c	469 ± 18	44.6 ± 1.4	65	53.0 (- 13.8)
	C111D	C-C-C-D	9.76 ± 0.15 ^c	216 ± 4	22.1 ± 0.1	32	47.0 (- 19.8)

Table 2. Kinetic parameters of WT and mutant XDH enzymes for xylitol. Activity was assayed spectrophotometrically. Values are means ± SD, *n* = 3. ^aFour amino acid residues at equivalent positions to Ser96, Ser99, Y102, and Cys110 in PsXDH. ^{b,c}Eight different concentrations of xylitol between 0.1 and 1 mM (b) or 1 and 10 mM (c) in the presence of 1.5 mM NAD⁺ were used. ^dNot assessed due to trace activity. ^eRelative value of *k_{cat}/K_m* to each WT. ^fThe apparent half denaturation temperature obtained from Fig. 5 ^gThe difference from each WT enzyme.

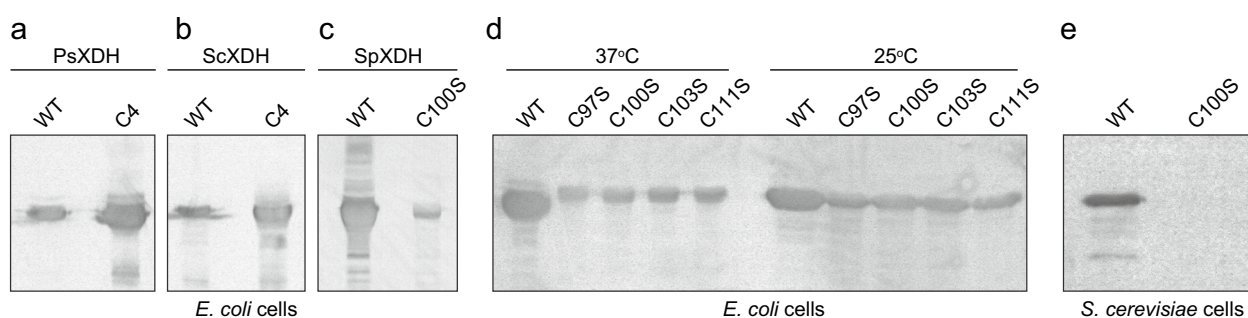


Figure 7. Intercellular expression level by an immunoblot analysis. Fifty micrograms each of cell-free extracts of transformed *E. coli* (a–d) or *S. cerevisiae* cells (e) was applied. A Western blot analysis was performed using the ECL Western blotting system (GE Healthcare) and Anti-Penta-His antibody (Qiagen) according to the manufacturer's instructions.

Discussion

Molecular evolution of structural zinc.

Only two studies previously investigated the artificial removal of structural zinc by site-directed mutagenesis, similar to the present study. Mutations in any of the four cysteine ligands to an alanine residue in $\beta\beta$ and $\chi\chi$ ADHs from humans¹⁹ or phenylacetaldehyde reductase (long-chain ADH) from *Corynebacterium* sp.²⁰ resulted in no expression in *E. coli* cells or a marked decrease in activity to less than 4% that of the WT enzyme, suggesting the impossibility of removing the zinc atom without the loss of stable folding or enzyme activity. Alternatively, the C97S and C103S mutants of SpXDH maintained folding and activity; however, their thermostabilities decreased. In other words, if decreased stability is not a significant issue for enzyme function under physiological conditions, such a mutation may be neutral (but not negative). In spite of the (possible) removal of structural zinc, all SpXDH mutants maintained similar thermostabilities to PsXDH_{WT} and ScXDH_{WT} (Fig. 6a,b).

Any substitution(s) of the four cysteine ligands with an aspartate (and glutamate) residue in SpXDH_{WT} prevented a decrease in thermostability (the C97D, C100D, and C103D mutants), and/or enhanced correct structural folding (the C111D mutant). These acidic residues may have been alternatively coordinated to structural zinc, similar to some MDR enzymes (Fig. 5f,g)^{26–29}. In other words, the four cysteine ligands may have been primarily

modified via an aspartate residue, but not by random mutations in any residue(s). Since the structural lobe surrounding zinc formed one of the major points of contact in the XDH tetramer (Fig. 3b,c), a serine residue with a similar sized side chain to cysteine must have contributed to the maintenance of the integrity of this lobe after the loss of zinc. Therefore, the aspartate and serine residues, which are often found in the substitution patterns of the four cysteine ligands in PDHs, may be traced for the hypothetical evolutionary process; S-S-Y-C, S-S-T-C, D-S-M-D, D-S-S-D, and R-D-C-S (Fig. 2).

The introduction of the four cysteine ligands increased thermostability not only in PsXDH (S-S-Y-C), but also in ScXDH (D-S-M-D) (Fig. 6a). Since C4 mutations reversely mimic the molecular evolution described above, and do not have to be generated by a random mutagenesis method, this strategy may be broadly applicable to other MDR enzymes. XDHs from yeast and fungi are further classified into two groups, which correspond to enzymes in the absence (group 1) or presence of structural zinc (group 2), respectively (Fig. 1b). Among them, group 1 contains some enzymes from “thermotolerant” yeasts with the ability to grow and ferment at higher temperatures (50 °C), including *Kluyveromyces marxianus*³⁰ and *P. angusta*³¹. Their thermostabilities are similar to PsXDH_{C4}, indicating that these properties were acquired later by a strategy other than C4 mutations, such as refining of the structural zinc binding loop²⁴, as described above (Fig. 3c).

Application of bioethanol production by lignocellulose biomass. Although *S. cerevisiae* co-expressing the PsXR and PsXDH genes ferments (metabolizes) D-xylose, additional genetic introductions and/or deletions have been shown to result in an increased ethanol yield, concomitant with a decreased byproduct yield, including xylitol, glycerol, and acetate^{32–34}. Another strategy is to modify the intercellular amount of XDH (and also XR) by plasmid copy numbers and the promoter control^{35,36}. Alternatively, since increases in genetic expression levels may be (partially) compensated for by the intercellular lifetime of the translated protein (Fig. 7a–d), SpXDH may be significantly useful for D-xylose fermentation because of its markedly higher thermostability than not only PsXDH_{WT} but also PsXDH_{C4}; the optimum temperature for activity was 55–65 °C and thermal inactivation was eventually observed at 70 °C (Fig. 6a). In a preliminary experiment, the SpXDH_{WT} gene was successfully expressed in *S. cerevisiae* cells under a constitutive phosphoglycerate kinase (PGK) promoter³⁷, whereas no expression of the thermolabile C100S mutant was noted (Fig. 7e).

L-Arabinose accounts for approximately 28% of the hemicellulose fraction of corn fiber (14%). The efficient fermentation of L-arabinose by *S. cerevisiae* has been achieved by using the bacterial pathway consisting of AraABD³⁸. On the other hand, the co-expression of LADH and L-xylulose reductase genes, involved in the fungal pathway (Fig. 1a), along with PsXR, PsXDH, and ScXK enabled *S. cerevisiae* to ferment L-arabinose; however, ethanol production occurred at a very low rate³⁹. PsXDH exhibited native activity for L-arabinol (data not shown), and showed high sequence homology with XDH from *Meyerozyma caribbica* (70%) (Fig. 1b), which corresponds to “LADH” purified from yeast cells grown on L-arabinose as a sole carbon source⁴⁰. In this regard, PsXDH is suitable for generating a bifunctional dehydrogenase for xylitol and L-arabinol, based on the structural data in this study, which is useful for breeding D-xylose and L-arabinose co-fermenting *S. cerevisiae*.

Materials and methods

Expression and purification of recombinant proteins. The primer sequences used in the present study are shown in Supplementary Table 2. Each (putative) XDH gene of *P. stipitis* (encoded by PICST_86924 gene), *S. cerevisiae* (YLR070c), and *S. pombe* (SPBC1773.05c) was introduced into pQE-81L (Qiagen), a plasmid vector for conferring an N-terminal (His)₆-tag on expressed proteins, to yield pQE-PsXDH_{WT}, pQE-ScXDH_{WT}, and pQE-SpXDH_{WT}, respectively. *E. coli* strain DH5α harboring the pQE-based vector was grown at 37 °C to a turbidity of 0.8 at 600 nm in LB medium containing ampicillin (50 mg/l). After the addition of 1 mM isopropyl-β-D-thiogalactopyranoside, the culture was grown at 37 °C for 6 h or at 20 °C for 18 h to induce the expression of the respective (His)₆-tagged protein. Cells were harvested and resuspended in Buffer A (50 mM sodium phosphate buffer (pH 8.0) containing 300 mM NaCl and 10 mM imidazole). Cells were then disrupted by sonication and the solution was centrifuged. The supernatant was loaded onto a Ni-NTA Superflow column (Qiagen), which was then washed with Buffer B (pH 8.0, Buffer A containing 25 mM imidazole instead of 10 mM imidazole). Enzymes were eluted with Buffer C (pH 8.0, Buffer A containing 250 mM imidazole instead of 10 mM imidazole), and the elutant was loaded onto a HiLoad 16/600 Superdex 200 pg column (GE Healthcare) equilibrated with Buffer D (20 mM Tris-HCl (pH 8.0) containing 150 mM NaCl). The main single-peak fractions were collected and concentrated by ultrafiltration with Amicon Ultra-15 (Millipore).

Site-directed mutagenesis. Several mutants of PsXDH, ScXDH, and SpXDH were constructed by a PCR-based method with the mutated sense and antisense primers (Supplementary Table S2), and pQE/PsXDH_{WT}, pQE/ScXDH_{WT}, or pQE/SpXDH_{WT} as a template, respectively.

Enzyme assay. Dehydrogenase activity for xylitol was measured using a continuous spectrophotometric assay at 340 nm at 30 °C in 50 mM Tris-HCl buffer (pH 8.0) containing 100 mM xylitol and 1 mM NAD⁺.

Crystallization and X-ray crystallography. All crystallization trials were performed at 20 °C using the sitting-drop vapor diffusion method. Drops (0.5 μL) of ~20 mg/mL PsXDH_{C4} protein in Buffer D were mixed with equal amounts of reservoir solution, and equilibrated against 70 μL of the same reservoir solution by vapor diffusion. The initial trial was performed using Index HT and Crystal Screen (Hampton Research). The best crystal of PsXDH_{C4} was obtained within 1 week under the following conditions: 100 mM Hepes-NaOH (pH 7.0), 2 M ammonium sulfate, and 2.5% (w/v) polyethylene glycol 400. The crystals obtained were cryoprotected

with reservoir solution supplemented with 15% (w/v) glycerol, and flash-cooled and kept in a stream of nitrogen gas at 100 K during data collection.

Diffraction data were collected with the PILATUS 6 M detector of BL45XU at SPring-8 (Hyogo, Japan), and the processed ZOO system and XDS^{41–43}. The structure of the apo-form of PsXDH_{C4} was solved by the molecular replacement method using the molecular-replacement pipeline program BALBES⁴⁴ with the structure of SDH from sheep liver (PDB ID 3QE3)¹¹ as the search model. Further model building for all structures was performed manually with COOT⁴⁵ and crystallographic refinement with PHENIX⁴⁶. Detailed data collection and processing statistics are shown in Table 1.

Overexpressing XDH genes in *S. cerevisiae*. Each DNA fragment of (His)₆-PsXDH_{WT}, (His)₆-PsXDH_{C4}, (His)₆-SpXDH_{WT} and (His)₆-SpXDH_{C100S} was amplified by PCR using the pQE-based vector as a template and was then introduced into EcoRI-HindIII sites between the PGK expression cassettes in the plasmid YEpPGK³⁷. *S. cerevisiae* D452-2 strains (*MATa leu2 his3 ura3 can1*) harboring the YEpPGK-based vector were grown in minimal medium supplemented with 2% (w/v) glucose as a sole carbon source at 30 °C. Cells were harvested, resuspended in 50 mM Tris-HCl (pH 8.0), and vortexed together with an equal volume of glass beads (diameter of 0.5 mm). Cell debris and glass beads from the cell extract were separated by centrifugation and the remaining supernatant was used for enzyme assessments.

Sequence comparison. Protein sequences were analyzed using the Protein-BLAST and Clustal W programs distributed by the Kyoto Encyclopedia of Genes and Genomes (KEGG) of Japan (www.kegg.jp/kegg/kegg1.html)^{47,48}.

Data availability

Coordinates and structural factors have been deposited in the Protein Data Bank (www.pdb.org), under the accession code 7Y9P.

Received: 24 August 2022; Accepted: 31 January 2023

Published online: 02 February 2023

References

1. Servigce, R. F. Cellulosic ethanol. Biofuel researchers prepare to reap a new harvest. *Science* **315**, 1488–1491 (2007).
2. Rubin, E. M. Genomics of cellulosic biofuels. *Nature* **454**, 841–845 (2008).
3. Jeffries, T. W. & Jin, Y. S. Metabolic engineering for improved fermentation of pentoses by yeasts. *Appl. Microbiol. Biotechnol.* **63**, 495–509 (2004).
4. Hahn-Hägerdal, B., Karhumaa, K., Fonseca, C., Spencer-Martins, I. & Gorwa-Grauslund, M. F. Towards industrial pentose-fermenting yeast strains. *Appl. Microbiol. Biotechnol.* **74**, 937–953 (2007).
5. Matsushika, A., Inoue, H., Kodaki, T. & Sawayama, S. Ethanol production from xylose in engineered *Saccharomyces cerevisiae* strains: Current state and perspectives. *Appl. Microbiol. Biotechnol.* **84**, 37–53 (2009).
6. Watanabe, S. *et al.* Novel non-phosphorylative pathway of pentose metabolism from bacteria. *Sci. Rep.* **9**, 155 (2019).
7. Kuyper, M. *et al.* High-level functional expression of a fungal xylose isomerase: the key to efficient ethanolic fermentation of xylose by *Saccharomyces cerevisiae*?. *FEMS Yeast Res.* **4**, 69–78 (2003).
8. Toivari, M. H., Salusjärvi, L., Ruohonen, L. & Penttilä, M. Endogenous xylose pathway in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **70**, 3681–3686 (2004).
9. Riveros-Rosas, H., Julián-Sánchez, A., Villalobos-Molina, R., Pardo, J. P. & Piña, E. Diversity, taxonomy and evolution of medium-chain dehydrogenase/reductase superfamily. *Eur. J. Biochem.* **270**, 3309–3334 (2003).
10. Pauly, T. A. *et al.* X-ray crystallographic and kinetic studies of human sorbitol dehydrogenase. *Structure* **11**, 1071–1085 (2003).
11. Yennawar, H., Moller, M., Gillilan, R. & Yennawar, N. X-ray crystal structure and small-angle X-ray scattering of sheep liver sorbitol dehydrogenase. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 440–446 (2011).
12. Banfield, M. J., Salvucci, M. E., Baker, E. N. & Smith, C. A. Crystal structure of the NADP(H)-dependent ketose reductase from *Bemisia argentifolii* at 2.3 Å resolution. *J. Mol. Biol.* **306**, 239–250 (2001).
13. Richard, P., Londesborough, J., Putkonen, M., Kalkkinen, N. A. & Penttilä, M. Cloning and expression of a fungal L-arabinitol 4-dehydrogenase gene. *J. Biol. Chem.* **276**, 40631–40637 (2001).
14. Sullivan, R. & Zhao, H. Cloning, characterization, and mutational analysis of a highly active and stable L-arabinitol 4-dehydrogenase from *Neurospora crassa*. *Appl. Microbiol. Biotechnol.* **77**, 845–852 (2007).
15. Bae, B., Sullivan, R. P., Zhao, H. & Nair, S. K. Structure and engineering of L-arabinitol 4-dehydrogenase from *Neurospora crassa*. *J. Mol. Biol.* **402**, 230–240 (2010).
16. Schneider, G., Eklund, H., Cedergren-Zeppezauer, E. & Zeppezauer, M. Crystal structures of the active site in specifically metal-depleted and cobalt-substituted horse liver alcohol dehydrogenase derivatives. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 5289–5293 (1983).
17. Korkhin, Y. *et al.* NADP-dependent bacterial alcohol dehydrogenases: Crystal structure, cofactor-binding and cofactor specificity of the ADHs of *Clostridium beijerinckii* and *Thermoanaerobacter brockii*. *J. Mol. Biol.* **278**, 967–981 (1998).
18. Ishikawa, K., Higashi, N., Nakamura, T., Matsuura, T. & Nakagawa, A. The first crystal structure of L-threonine dehydrogenase. *J. Mol. Biol.* **366**, 857–867 (2007).
19. Jeloková, J., Karlsson, C., Estonius, M., Jörnval, H. & Höög, J. O. Features of structural zinc in mammalian alcohol dehydrogenase. Site-directed mutagenesis of the zinc ligands. *Eur. J. Biochem.* **225**, 1015–1019 (1994).
20. Wang, J. C., Sakakibara, M., Matsuda, M. & Itoh, N. Site-directed mutagenesis of two zinc-binding centers of the NADH-dependent phenylacetaldehyde reductase from styrene-assimilating *Corynebacterium* sp. strain ST-10. *Biosci. Biotechnol. Biochem.* **63**, 2216–2218 (1999).
21. Karhumaa, K., Garcia Sanchez, R., Hahn-Hägerdal, B. & Gorwa-Grauslund, M. F. Comparison of the xylose reductase-xylylitol dehydrogenase and the xylose isomerase pathways for xylose fermentation by recombinant *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **6**, 5 (2007).
22. Watanabe, S., Kodaki, T. & Makino, K. Complete reversal of coenzyme specificity of xylylitol dehydrogenase and increase of thermostability by the introduction of structural zinc. *J. Biol. Chem.* **280**, 10340–10349 (2005).
23. Lunzer, R., Mamnun, Y., Haltrich, D., Kulbe, K. D. & Nidetzky, B. Structural and functional properties of a yeast xylylitol dehydrogenase, a Zn²⁺-containing metalloenzyme similar to medium-chain sorbitol dehydrogenases. *Biochem. J.* **336**, 91–99 (1998).

24. Annaluru, N. *et al.* Thermostabilization of *Pichia stipitis* xylitol dehydrogenase by mutation of structural zinc-binding loop. *J. Biotechnol.* **129**, 717–722 (2007).
25. Richard, P., Toivari, M. H. & Penttilä, M. Evidence that the gene *YLR070c* of *Saccharomyces cerevisiae* encodes a xylitol dehydrogenase. *FEBS Lett.* **457**, 135–138 (1999).
26. Guy, J. E., Isupov, M. N. & Littlechild, J. A. The structure of an alcohol dehydrogenase from the hyperthermophilic archaeon *Aeropyrum pernix*. *J. Mol. Biol.* **331**, 1041–1051 (2003).
27. Esposito, L. *et al.* Crystal structure of the alcohol dehydrogenase from the hyperthermophilic archaeon *Sulfolobus solfataricus* at 1.85 Å resolution. *J. Mol. Biol.* **318**, 463–477 (2002).
28. Brouns, S. J., Turnbull, A. P., Willemen, H. L., Akerboom, J. & van der Oost, J. Crystal structure and biochemical properties of the D-arabinose dehydrogenase from *Sulfolobus solfataricus*. *J. Mol. Biol.* **371**, 1249–1260 (2007).
29. John, J., Crennell, S. J., Hough, D. W., Danson, M. J. & Taylor, G. L. The crystal structure of glucose dehydrogenase from *Thermoplasma acidophilum*. *Structure* **2**, 385–393 (1994).
30. Lulu, L. *et al.* Identification of a xylitol dehydrogenase gene from *Kluyveromyces marxianus* NBRC1777. *Mol. Biotechnol.* **53**, 159–169 (2013).
31. Biswas, D., Datt, M., Ganesan, K. & Mondal, A. K. Cloning and characterization of thermotolerant xylitol dehydrogenases from yeast *Pichia angusta*. *Appl. Microbiol. Biotechnol.* **88**, 1311–1320 (2010).
32. Walfridsson, M., Hallborn, J., Penttilä, M., Keränen, S. & Hahn-Hägerdal, B. Xylose-metabolizing *Saccharomyces cerevisiae* strains overexpressing the *TKL1* and *TAL1* genes encoding the pentose phosphate pathway enzymes transketolase and transaldolase. *Appl. Environ. Microbiol.* **61**, 4184–4190 (1995).
33. Runquist, D., Fonseca, C., Rådström, P., Spencer-Martins, I. & Hahn-Hägerdal, B. Expression of the *Gxf1* transporter from *Candida intermedia* improves fermentation performance in recombinant xylose-utilizing *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **82**, 123–130 (2009).
34. Roca, C., Nielsen, J. & Olsson, L. Metabolic engineering of ammonium assimilation in xylose-fermenting *Saccharomyces cerevisiae* improves ethanol production. *Appl. Environ. Microbiol.* **69**, 4732–4736 (2003).
35. Ho, N. W., Chen, Z. & Brainard, A. P. Genetically engineered *Saccharomyces* yeast capable of effective cofermentation of glucose and xylose. *Appl. Environ. Microbiol.* **64**, 1852–1859 (1998).
36. Zha, J. *et al.* Balance of *XYL1* and *XYL2* expression in different yeast chassis for improved xylose fermentation. *Front. Microbiol.* **3**, 355 (2012).
37. Kang, Y. S., Kane, J., Kurjan, K., Stadel, J. M. & Tipper, D. J. Effects of expression of mammalian Gα and hybrid mammalian-yeast Gα proteins on the yeast pheromone response signal transduction pathway. *Mol. Cell Biol.* **10**, 2582–2590 (1990).
38. Becker, J. & Boles, E. A modified *Saccharomyces cerevisiae* strain that consumes L-arabinose and produces ethanol. *Appl. Environ. Microbiol.* **69**, 4144–4150 (2003).
39. Richard, P., Verho, R., Putkonen, M., Londesborough, J. & Penttilä, M. Production of ethanol from L-arabinose by *Saccharomyces cerevisiae* containing a fungal L-arabinose pathway. *FEMS Yeast Res.* **3**, 185–189 (2003).
40. Sukpipat, W., Komeda, H., Prasertsan, P. & Asano, Y. Purification and characterization of xylitol dehydrogenase with L-arabitol dehydrogenase activity from the newly isolated pentose-fermenting yeast *Meyerozyma caribbica* 5XY2. *J. Biosci. Bioeng.* **123**, 20–27 (2017).
41. Hirata, K. *et al.* ZOO: An automatic data-collection system for high-throughput structure analysis in protein microcrystallography. *Acta Crystallogr. D Struct. Biol.* **75**, 138–150 (2019).
42. Yamashita, K., Hirata, K. & Yamamoto, M. KAMO: Towards automated data processing for microcrystals. *Acta Crystallogr. D Struct. Biol.* **74**, 441–449 (2018).
43. Kabsch, W. Xds. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
44. Long, F., Vagin, A., Young, P. & Murshudov, G. N. BALBES: A molecular-replacement pipeline. *Acta Crystallogr. D Biol. Crystallogr.* **64**, 125–132 (2008).
45. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
46. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
47. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–D462 (2016).
48. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
49. DeLano, W. PyMOL, DeLano Scientific, San Carlos, CA (2002).

Acknowledgements

We would like to thank the staff at beamline BL45XU of SPring-8, Japan for their support with data collection. The X-ray diffraction experiments were carried out with the approval of Japan Synchrotron Radiation Research Institute (proposal No. 2022B2534, Japan). Our thanks extend to especially Dr. Miyuki Kawano-Kawada and Prof. Takayuki Sekito, Ehime University, for the gift of genomic DNA from *S. pombe*.

Author contributions

S.W. and K.Y. mainly performed the experiments. S.W. conceived and designed the experiments and wrote the manuscript. K.Y. and Y.W. contributed to the X-ray crystallographic analysis.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-29195-7>.

Correspondence and requests for materials should be addressed to S.W.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023