STRUCTURAL INSIGHTS INTO THE ENZYME MECHANISM OF A NEW FAMILY OF D-2-HYDROXYACID DEHYDROGENASES, A CLOSE HOMOLOG OF 2-KETOPANTOATE REDUCTASE

SUKANTA MONDAL suku@nibio.go.jp
KENJI MIZUGUCHI kenji@nibio.go.jp

National Institute of Biomedical Innovation, 7-6-8 Saito-Asagi, Ibaraki-shi, Osaka 567-0085, Japan

A newly identified family of NAD-dependent D-2-hydroxyacid dehydrogenases (D-2-HydDHs) catalyzes the stereo-specific reduction of branched-chain 2-keto acids with bulky hydrophobic side chains to 2-hydroxyacids. They are promising targets for industrial/practical applications, particularly in the stereo-specific synthesis of C3-branched D-hydroxyacids. Comparative modeling and docking studies have been performed to build models of the enzyme-cofactor-substrate complexes and identify key residues for cofactor and substrate recognition. To explore large conformational transitions (domain motions), a normal mode analysis was employed using a simple potential and the protein models. Our analysis suggests that the new D-2-HydDH family members possess the N-terminal NAD(H) binding Rossmann-fold domain and the α-helical C-terminal substrate binding domain. A hinge bending motion between the N- and C-terminal domains was predicted, which would trigger the switch of the conserved essential Lys to form a key hydrogen bond with the C2 ketone of the 2-keto acid substrates. Our findings will be useful for site-directed mutagenesis studies and protein engineering.

Keywords: Rossmann-fold; 2-keto acids; homodimer; hinge bending; normal mode analysis.

1. Introduction

D-2-hydroxyacid dehydrogenases (D-2-HydDHs) catalyze the stereo-specific and reversible reduction of 2-ketoacids to 2-hydroxyacids. These NAD(H)-dependent oxidoreductases perform various physiological roles, with distinct preferences for the substrate side chains. Therefore, these enzymes are of interest in a variety of fields; they are valuable catalysts for the production of the stereo-specific isomers of 2-hydroxyacids, which are used in the production of semi-synthetic antibiotics or pharmaceuticals [1]; and they are involved in the biopreservation properties of lactic acid bacteria, as certain 2-hydroxyacids exhibit antifungal and antilisterial activities [13, 25].

It has been reported recently that the gene LL1323 from *Lactococcus lactis* encodes a D-hydroxyacid dehydrogenase (hdhD) [4] and its homologous gene from *Enterococcus faecalis* encodes a D-mandelate dehydrogenase (D-manDH2) [22]. These enzymes exhibit high catalytic activity toward various 2-ketoacid substrates with bulky hydrophobic side chains, particularly C3-branched substrates such as 2-ketoisocaproate (KIC), 2-ketoisovalerate (KIV) and benzylformate (BEN), and strict coenzyme specificity for NADH and NAD^+. These enzymes have been shown to constitute a new family of D-2-HydDHs [4, 22], which differs from the previously characterized D-2-hydroxyisocaproate dehydrogenase (D-HicDH) family. The putative enzymes of this new

family are widely distributed in various organisms and display a high diversity in their sequences. They showed a marked sequence similarity with *Escherichia coli* 2-ketopantoate reductase (KPR; EC 1.1.1.169), encoded by the *panE* gene, and previously, these putative enzymes were incorrectly annotated as *panE*. KPR catalyzes the NADPH-dependent reduction of ketopantoate to pantoate, an essential step in the biosynthesis of pantothenate (vitamin B5) [5]. However, the differences in enzymatic function between KPR and D-2-HydDH are due to only a few amino acid substitutions. Based on a comparative sequence analysis, Wada *et al.* [22], proposed that Asp30 and Met205 of D-manDH2 are the crucial residues for cofactor (NAD(H)) and substrate specificity. Currently, KPR and its close homologs (non-KPRs including D-2-HydDHs) are annotated to possess the same Pfam [11] domains, namely, ApbA (PF02558) and ApbA_C (PF08546). It has been reported that only minor structural changes, such as, amino acid replacements, cause a drastic change in the enzymatic function in D-2HydDH family, such as, the conversions from a D-lactate dehydrogenase (D-LDH) to a D-HicDH [21], and even from a formate dehydrogenase to a D-2-HydDH [20].

Given the practical importance of these enzymes, a better understanding of the structural basis for coenzyme and substrate specificity and the reaction mechanism would be beneficial for protein engineering applications. In this report, we utilize comparative modeling and bioinformatics tools and propose a novel reaction mechanism for these enzymes.

2. Material and Methods

2.1. *Comparative Modeling and Structural Analysis*

Protein sequences of the members of the family of D-2-HydDHs members were downloaded from UniProt [6]. We have used the PDB entry 2EW2 from *E. faecium* (UniProt Q831Q5) as a template for modeling a representative member of the new family, *L. lactis* hdhD (UniProt Q9CFY8) [4] (see section 3.1 for details). The coordinates of the analyzed protein structures were downloaded from the RCSB Protein Data Bank (PDB) [2]. Correctly aligning the query sequence with the template is a crucial step in comparative modeling. For the present study, reliable alignments of query sequence with template 2EW2 were generated using the sequence-structure homology recognition software FUGUE [17]. Based on this alignment, we constructed models using MODELLER 9v5 [10] and the quality of models was assessed with JOY [15], Verify3D [9] and Prosa-web [24]. MUSTANG [12] was used for comparing the model and template structures. Dimeric interfaces were analyzed using DIMPLOT [23]. Domain movements were analyzed with the DynDom server [14].

The structural similarity between the N- and C-terminal domains of KPR and 2EW2 enabled us to carry out molecular modeling to identify the cofactor and substrates binding site, since the structure of the ternary complex of *E. coli* KPR with $NADP^+$ and pantoate (PDB code 2OFP) has been determined [5]. The N-terminal domain of KPR

(2OFP_A) was fitted to the modeled hdhD structure, and the $NADP^+$ cofactor extracted from KPR and transferred to the modeled structure. The $NADP^+$ molecule was edited to NAD^+ and NADH. Hydrogen atoms were added to the modeled complexes with the SYBYL Biopolymer module (Tripos, 1998) and the cofactor and its neighboring amino acid residues were optimized using the Tripos force field. Similar steps were taken for modeling substrate complexes, where the pantoate molecule was edited to KIC, KIV and BEN. Protein-ligand interactions were studied using LIGPLOT [23] and NCONT from the CCP4 suit [16]. Pymol [7] was used for visualization and image generation.

2.2. *Normal Mode Analysis*

Normal mode analysis (NMA) of proteins is based on the hypothesis that the vibrational normal modes exhibiting the lowest frequencies describe the largest and functionally relevant movements in a protein. NMA has proven successful in representing domain and hinge-bending motions in proteins [18].

The PDB formatted dimeric enzyme structures were submitted to the *ElNemo* [19], a web interface to the *Elastic Network Model*, for computing low frequency normal modes. Input parameters included the request for the first five nontrivial normal modes, a DQMIN to DQMAX range of -500 to 500, and a DQSTEP of 50. The NRBL setting was set to 'auto', and the default cutoff of 8 Å was used to identify the elastic interaction ranges.

3. Results

3.1. *Comparative Modeling*

EF2445 from *E. faecium* (UniProt Q831Q5) is a member of this new family [22], although previously it had been incorrectly annotated as a putative KPR. The crystal structure of EF2445 is available from a structural genomics project (PDB code 2EW2) and is similar to that of *E. coli* KPR. *E. coli* KPR has been shown to adopt open and closed conformations [5]. The structure of EF2445 is an *apo* (unliganded) form and shows an open conformation.

The sequence lengths of EF2445 (313 amino acids) and hdhD (312 amino acids) are nearly equal. Their full-length percentage sequence identity (PID) is 61.3. Thus, we have used the PDB entry 2EW2 as a template and built comparative models of hdhD.

The asymmetric unit of the crystal structure of EF2445 contains two protein molecules, burying a large amount of solvent accessible surface areas. Their dimeric interface consists of a core hydrophobic patch and a ring of polar residues. These residues are conserved in this family (see Figure 1) and solution studies also suggest that hdhD forms a homodimer [4].

3.2. *Hinge Bending*

Although EF2445 is a homodimer, the relative orientations of the N- and C-terminal domains are different in the two subunits. By comparing these two subunits, the DynDom server [14] detected a hinge bending motion, with the inter-domain bending residues of 183-184. Normal mode analysis of the dimer of EF2445 also suggests a collective motions, with collectivity of 0.70 for the lowest three modes and a correlation between predicted and experimental B-factors being 0.70 for 626 C-α atoms. Normal mode analysis of the modeled homodimer of hdhD also reveals collective motions (Figure 2A).

Figure 1. Multiple sequence alignment of *E. coli* KPR (PDB code 2OFP) and representative new family members of *Enterococcus faecalis* (*Streptococcus faecalis*) (UniProt Q831Q5, PDB code 2EW2), *Clostridium perfringens* (UniProt Q0TT01), *Streptococcus pyogenes* (UniProt P65666), *Lactococcus lactis* (hdhD, UniProt Q9CFY8), *Lactobacillus sakei* (UniProt Q38ZN5), *Enterococcus faecium* DO (UniProt Q3Y316, same as D-manDH2), *Lactococcus lactis* subsp. cremoris (UniProt Q02YL7), *Lactobacillus casei* (UniProt Q039A2), *Salmonella typhi* (UniProt Q8Z4L0), *Salmonella typhimurium* (UniProt Q8ZN23), *Staphylococcus aureus* (UniProt Q2FVH3). Starting residue numbers are shown in parenthesis; GXGXXG motif, cofactor and substrates binding residues are shaded in gray. Dimeric interface residues of 2EW2 are marked as #. The alignment is annotated with JOY with the following formatting convention: red, α helices; blue, strands; upper case letters, solvent inaccessible; lower case letters, solvent accessible; bold type, hydrogen bonds to mainchain amides; underlining, hydrogen bonds to mainchain carbonyls; italic, positive mainchain torsion angles φ.

3.3. *Cofactor Recognition*

hdhD and EF2445 have an Asp at position 30 on the β2-α2 loop (Figures 1 & 2B), suggesting that they are NAD-dependent. NAD-specific enzymes commonly possess an acidic amino acid, Asp or Glu, in the β2-α2 loop of the Rossmann fold, which forms a

hydrogen bond with the 2'-hydroxyl group of NAD, but not with the 2'-phosphate of NADP due to electrostatic repulsion [3].

Lys72 of *E. coli* KPR is crucial for cofactor recognition [5], although its side chain forms a hydrogen bond with $NADP^+$ pyrophosphates only in the closed form. In hdhD, the structurally equivalent Lys79 is conserved and would play a similar role (Figure 1). Two other residues, Asn98 and Glu256 of *E. coli* KPR, form hydrogen bonds with the 3' hydroxyl group of nicotinamide-ribose and the equivalent residues [5], Asn105 and Glu273 of hdhD, would be crucial for cofactor recognition (Figures 1 & 2B).

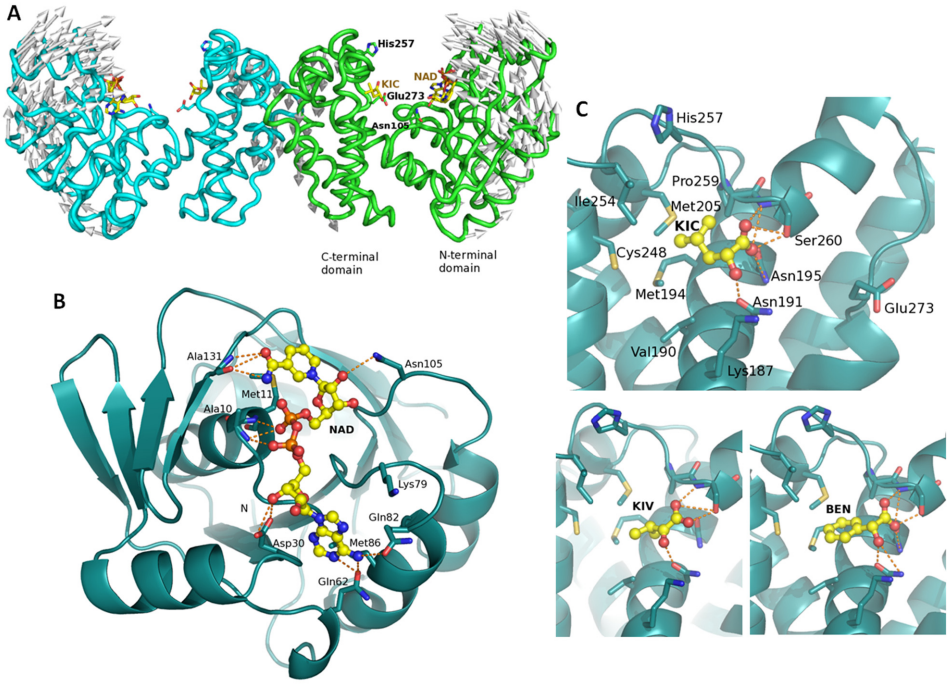

Figure 2. (A) Domain movement of hdhD as seen in the first nontrivial normal mode. Proposed models for interaction of the (B) cofactor and (C) substrates with hdhD.

3.4. *Substrate Recognition*

Residues Asn191, Asn195 and Ser260 of hdhD are conserved and structurally equivalent to Asn180, Asn184 and Ser244 of *E. coli* KPR and these residues form hydrogen bonds with the hydroxyacid moiety of pantoate [5]. Therefore, these residues are predicted to recognize the keto acid moiety of the 2-keto acid substrates of hdhD. Hydrophobic side chains of hdhD substrates can be nicely accommodated in the conserved hydrophobic pocket (Pro259, Met205, Ile254, Cys248, Met194 and Val190 of hdhD) (Figure 2C). Consistent with this observation, a recent report suggests that a Cys residue is involved in

the active center of hdhD [4]. His257 of hdhD in the C-terminal domain is structurally equivalent to Asn241 of the *E. coli* KPR, which interacts with both the cofactor and substrate in the closed form. Although His257 in the modeled structure is >8 Å away from the substrate, this residue may interact with the N-terminal domain in the closed form and stabilize that conformation. A recent study has shown that a His reagent significantly lowers the catalytic activity of hdhD [4] and this observation may be explained by the predicted role of His257. However, this His residue would not be able to interact with the keto acid moiety, even in the closed form, because of its spatial position. No other His residue is found near the active site and therefore, no His residue in hdhD appears to be directly involved in the catalysis, in contrast to other known D-2-HydDHs, which possess the characteristic feature of His-Asp/Glu pair, as observed in D-HicDH from *L. casei* [8].

4. Discussion

hdhD possesses the cofactor (NAD(H)) binding N-terminal Rossmann-fold domain, and the substrate binding α-helical C-terminal domain. Homodimerization *via* the C-terminal domains appears to be important for facilitating the open-close transition, as shown in our normal mode analysis.

Based on our comparative analysis, we have proposed a reaction mechanism for hdhD, which is expected to hold true for the other members of this new family. The binding of NAD(H) and the substrate induces conformational changes to provide an active environment that favors 2-keto acids (KIC, KIV or BEN) binding. In the closed form, Asn105 and Glu273 provide an extended hydrogen bonding network, which orients the ribose-nicotinamide moiety of the cofactor in the conformation required for productive D-substrate binding and optimal hydride transfer. In other known D-2-HydDHs, a His residue plays an essential role in fixing the proper orientation of the substrate [8] but in hdhD, we predict the conserved Lys187 to play an equivalent role by forming a key hydrogen bond with the C2 ketone of 2-keto acid substrates, in analogy with KPR. The substrate is locked during each step of the reaction via hydrogen bonds of its carboxylate group to the conserved Ser260. Asn105 and Ly187 may play a central role in the reaction mechanism, and Asn105, Glu273 and His257 may play a crucial role in opening and closing the enzyme.

5. Conclusions

NAD-dependent hdhD and other D-2-HydDH family members are of interest in a variety of fields including the production of semi-synthetic antibiotics or pharmaceuticals and biopreservation. Despite one unliganded (and unannotated) crystal structure, little is known about the catalytic mechanism of the new family of hdhD. Our structure-based analysis revealed that NAD(H) would bind in the N-terminal Rossmann-fold domain and substrates would bind in the α-helical C-terminal domain. A hinge bending motion between the N- and C-terminal domains was predicted, which would trigger the switch of

the essential Lys187 of hdhD to form a key hydrogen bond with the substrates. We proposed a reaction mechanism for this family based on the analogy to *E. coli* KPR. Our models elucidated the roles of active site residues and this information would be useful for planning mutagenesis studies and protein engineering.

Acknowledgments

This work was in part supported by the Knowledge Cluster Initiative (stage II) by the MEXT (Ministry of Education, Culture, Sports, Science and Technology). We thank Dr. Lokesh P. Tripathi for carefully reading the manuscript.

References

[1] Kallwass, H.K.W., Potential of R-2-hydroxyisocaproate dehydrogenase from *Lactobacillus casei* for stereospecific reductions, *Enzyme Microb. Technol.*, 14:28-35, 1992.

[2] Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. and Bourne, P.E., The Protein Data Bank, *Nucleic Acids. Res.*, 28(1):235-242, 2000.

[3] Carugo, O. and Argos, P., NADP-dependent enzymes. I: Conserved stereochemistry of cofactor binding, *Proteins*, 28(1):10-28, 1997.

[4] Chambellon, E., Rijnen, L., Lorquet, F., Gitton, C., van Hylckama, Vlieg, J.E., Wouters, J.A. and Yvon, M., The D-2-hydroxyacid dehydrogenase incorrectly annotated PanE is the sole reduction system for branched-chain 2-keto acids in *Lactococcus lactis*, *J. Bacteriol.*, 191(3):873-881, 2009.

[5] Ciulli, A., Chirgadze, D.Y., Smith, A.G., Blundell, T.L. and Abell, C., Crystal structure of *Escherichia coli* ketopantoate reductase in a ternary complex with $NADP^+$ and pantoate bound: substrate recognition, conformational change, and cooperativity, *J. Biol. Chem.*, 282(11):8487-8497, 2007.

[6] Consortium, U., The Universal Protein Resource (UniProt) 2009, *Nucleic Acids Res.*, 37(Database issue):D169-174, 2009.

[7] DeLano, W.L., The PyMOL Molecular Graphics System, DeLano Scientific, Palo Alto, CA, USA, 2002.

[8] Dengler, U., Niefind, K., Kiess, M. and Schomburg, D., Crystal structure of a ternary complex of D-2-hydroxyisocaproate dehydrogenase from *Lactobacillus casei*, NAD^+ and 2-oxoisocaproate at 1.9 Å resolution, *J. Mol. Biol.*, 267(3):640-660, 1997.

[9] Eisenberg, D., Luthy, R. and Bowie, J.U., VERIFY3D: assessment of protein models with three-dimensional profiles, *Methods Enzymol.*, 277:396-404, 1997.

[10] Eswar, N., Eramian, D., Webb, B., Shen, M.Y. and Sali, A., Protein structure modeling with MODELLER, *Methods Mol. Biol.*, 426:145-159, 2008.

[11] Finn, R.D., Tate, J., Mistry, J., Coggill, P.C., Sammut, J.S., Hotz, H.R., Ceric, G., Forslund, K., Eddy, S.R., Sonnhammer, E.L. and Bateman, A., The Pfam protein families database, *Nucleic Acids Res.*, 36(Database Issue):D281-D288, 2008.

[12] Konagurthu, A.S., Whisstock, J.C., Stuckey, P.J. and Lesk, A.M., MUSTANG: a multiple structural alignment algorithm, *Proteins*, 64(3):559-574, 2006.
[13] Lavermicocca, P., Valerio, F. and Visconti, A., Antifungal activity of phenyllactic acid against molds isolated from bakery products, *Appl. Environ. Microbiol.*, 69(1):634-40, 2003.
[14] Lee, R.A., Razaz, M. and Hayward, S., The DynDom database of protein domain motions, *Bioinformatics*, 19(10):1290-1291, 2003.
[15] Mizuguchi, K., Deane, C.M., Blundell, T.L., Johnson, M.S. and Overington, J.P., JOY: protein sequence-structure representation and analysis, *Bioinformatics*, 14(7):617-623, 1998.
[16] Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Cowtan, K., Emsley, P., Murshudov, G.N., Cohen, S., Perrakis, A. and Noble, M., Developments in the CCP4 molecular-graphics project, *Acta Crystallogr. D Biol Crystallogr.*, 60:2288-2294, 2004.
[17] Shi, J., Blundell, T.L. and Mizuguchi, K., FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties, *J. Mol. Biol.*, 310(1):243-257, 2001.
[18] Skjaervena, L., Hollupc, S.M. and Reuter, N., Normal mode analysis for proteins, *J. Mol. Struct.: THEOCHEM*, 898(1-3):42-48, 2009.
[19] Suhre, K. and Sanejouand, Y.H., *ElNemo*: a normal mode web-server for protein movement analysis and the generation of templates for molecular replacement, *Nucleic Acids Res.*, 32(Web Server issue):W610-W614, 2004.
[20] Tishkov, V.I. and Popov, V.O., Protein engineering of formate dehydrogenase. *Biomol. Eng.*, 23(2-3):89–110, 2006.
[21] Tokuda, C., Ishikura, Y., Shigematsu, M., Mutoh, H., Tsuzuki, S., Nakahira, Y., Tamura, Y., Shinoda, T. and Taguchi, H., Conversion of *Lactobacillus pentosus* D-lactate dehydrogenase through a single amino acid replacement, *J. Bacteriol.*, 185(16):5023–5026, 2003.
[22] Wada, Y., Iwai, S., Tamura, Y., Ando, T., Shinoda, T., Arai, K. and Taguchi, H., A new family of D-2-hydroxyacid dehydrogenases that comprises D-mandelate dehydrogenases and 2-ketopantoate reductases. *Biosci. Biotechnol. Biochem.*, 72(4):1087-1094, 2008.
[23] Wallace, A.C., Laskowski, R.A. and Thornton, J.M., LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions, *Protein Eng.*, 8(2):127-134, 1995.
[24] Wiederstein, M. and Sippl, M.J., ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, *Nucleic Acids Res.*, 35(Web Server issue):W407-410, 2007.
[25] Wilson, A.R., Sigee, D. and Epton, H.A., Anti-bacterial activity of *Lactobacillus plantarum* strain SK1 against *Listeria monocytogenes* is due to lactic acid production, *J. Appl. Microbiol.*, 99(6):1516-1522, 2005.