Biochimica et Biophysica Acta 1844 (2014) 681–693

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

Aldehyde dehydrogenase enzyme ALDH3H1 from *Arabidopsis thaliana*: Identification of amino acid residues critical for cofactor specificity

Naim Stiti, Karolina Podgórska, Dorothea Bartels *

Institute of Molecular Physiology and Biotechnology of Plants, University of Bonn, Kirschallee 1, 53115 Bonn, Germany

ARTICLE INFO

Article history:
Received 12 June 2013
Received in revised form 9 January 2014
Accepted 13 January 2014
Available online 23 January 2014

Keywords:
Aldehyde
Cofactor
Dehydrogenase activity
NAD^+
$NADP^+$

ABSTRACT

The cofactor-binding site of the NAD^+-dependent *Arabidopsis thaliana* aldehyde dehydrogenase ALDH3H1 was analyzed to understand structural features determining cofactor-specificity. Homology modeling and mutant analysis elucidated important amino acid residues. Glu149 occupies a central position in the cofactor-binding cleft, and its carboxylate group coordinates the 2′- and 3′-hydroxyl groups of the adenosyl ribose ring of NAD^+ and repels the 2′-phosphate moiety of $NADP^+$. If Glu149 is mutated to Gln, Asp, Asn or Thr the binding of NAD^+ is altered and rendered the enzyme capable of using $NADP^+$. This change is attributed to a weaker steric hindrance and elimination of the electrostatic repulsion force of the 2′-phosphate of $NADP^+$. Simultaneous mutations of Glu149 and Ile200, which is situated opposite of the cofactor binding cleft, improved the enzyme efficiency with $NADP^+$. The double mutant $ALDH3H1_{Glu149Thr/Ile200Val}$ showed a good catalysis with $NADP^+$. Subsequently a triple mutation was generated by replacing Val178 by Arg in order to create a "closed" cofactor binding site. The cofactor specificity was shifted even further in favor of $NADP^+$, as the mutant $ALDH3H1_{E149T/V178R/I200V}$ uses $NADP^+$ with almost 7-fold higher catalytic efficiency compared to NAD^+. Our experiments suggest that residues occupying positions equivalent to 149, 178 and 200 constitute a group of amino acids in the ALDH3H1 protein determining cofactor affinity.

1. Introduction

Aldehyde dehydrogenases (ALDHs): (E.C 1.2.1.3) constitute a large evolutionary conserved superfamily of enzymes that catalyze the irreversible oxidation of a wide range of aldehydes to their corresponding carboxylic acids using $NAD(P)^+$ as cofactors. Aldehydes are widespread molecules that take part in different physiological processes. Aldehydes become toxic when they accumulate to excessive, nonphysiological levels. Hence, ALDHs are considered to be detoxifying enzymes which maintain physiological levels of aldehydes [1].

ALDHs are ubiquitous and are found throughout prokaryotic and eukaryotic organisms. With the completion of several genome sequencing projects, the number of newly identified *ALDH* genes from different taxa is steadily increasing, providing new insights into sequence plasticity and evolution [2–9]. Each ALDH monomeric subunit consists of three domains: the catalytic domain, the $NAD(P)^+$-binding domain and the oligomerization domain. Subunits are assembled by hydrogen bonds into enzymatically active homotetramers for members of families 1, 2, 7, and 9 [10–13], whereas members of family3 function as homodimers [14,15].

Plant ALDHs have been classified into 13 distinct protein families, where seven families (ALDH2, ALDH3, ALDH5, ALDH6, ALDH7, ALDH11 and ALDH18) have mammalian orthologs, and six families (ALDH10, ALDH12, ALDH21, ALDH22, ALDH23 and ALDH24) are specific for plants. Analysis of the genome sequence of *Arabidopsis thaliana* revealed 16 *ALDH* genes encoding members of ten ALDH protein families [3].

Plant ALDHs are localized in different subcellular compartments including cytosol, mitochondria, plastids, peroxisomes and microsomes [16–23]. This differential localization indicates functional specialization. Different cell compartments require appropriate ALDH enzymes with distinct physico-chemical characteristics. Developmental-dependent and organ-specific expression patterns of ALDHs have been reported [20,22–25].

The role of plant ALDHs in different physiological processes has become an emerging theme in plant physiology during the last years. The first plant ALDH to which a function was assigned was the spinach ALDH10 (betaine aldehyde dehydrogenase) which is involved in the synthesis of the osmoprotectant glycine betaine [26]. Later, the maize mitochondrial ALDH2B2/RF2A gene was identified as a nuclear restorer of cytoplasmic male sterility [27,28]. The mitochondrial class 2 ALDH of rice was reported to be essential for rapid detoxification of acetaldehyde which is produced upon re-aeration after submergence [29].

Expression of members of plant ALDHs from families 2, 3, 5, 7 and 10 is responsive to oxidative and abiotic stresses [25,26,29–31]. The involvement of these genes in stress protection has been demonstrated

Abbreviations: ALDH, aldehyde dehydrogenase; AMADH, aminoaldehyde dehydrogenase; BADH, betaine aldehyde dehydrogenase; ROS, reactive oxygen species

* Corresponding author. Tel.: +49 228 73 2070; fax: +49 228 73 1697.
E-mail address: dbartels@uni-bonn.de (D. Bartels).

1570-9639/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bbapap.2014.01.008

by analyzing transgenic plants. Overexpression of *Arabidopsis* ALDHs of families 3 and 7 confers tolerance to heavy metal, osmotic and oxidative stresses [21,32]. Likewise, ectopic expression of a soybean *ALDH7* enhanced stress tolerance [33]. Involvement of ALDHs in stress tolerance was further corroborated by the analysis of *Arabidopsis ALDH3I1*, *ALDH7B4* T-DNA and rice *OsALDH7* knockout mutants which exhibited higher stress sensitivity [21,25].

Although physiological data imply an important function of plant ALDHs in growth and environmental stress adaptation, the structural properties of plant ALDHs remain mostly unveiled [25,34–36]. Solved crystal structures of ALDHs [10,11,14,15,37–39] revealed common structural features. Dehydrogenases bind nicotinamide cofactors in a region containing a core Rossmann-type fold domain, which consists of six parallel β-strands interspersed by α-helices that appear on both sides of the six-stranded β-strand. A typical glycine-rich motif (Gly_1-X-Gly_2-X-X-Gly_3) forms a tight turn between the end of the first β-strand (β1) and the beginning of the so-called "dinucleotide binding helix" (αA) [40]. Structural variations have been found for the Rossmann fold in dehydrogenases. The Rossmann fold of ALDHs is different from other dehydrogenases. It is made up of a five stranded open α/β domain. Moreover, the NAD^+-binding fingerprint (Gly_1-X-Gly_2-X-X-Gly_3) which forms a tight turn and is usually present between β1 and αA is missing. Consequently, the loop located between β1 and αA is more extended and bulkier. The classical tight turn is found between β4 and αD in ALDHs and corresponds to the characteristic glycine-rich sequence Gly_1-X-X-X-Gly_2 involved in interaction with the cofactor nicotinamide ring. The $NAD(P)^+$ binding cleft is positioned between the helices αC and αD instead of the central helices as in alcohol dehydrogenases [14].

The relevance of different residues from the five open α/β strands of the Rossmann fold in cofactor recognition and accommodation was studied using site-directed mutagenesis approaches [34,41–43]. The residue located immediately downstream of the β2-strand functions as part of the NAD^+ recognition site, since the 2′-hydroxyl group of the ribose moiety forms a hydrogen bond with its side chain [11,14]. The mutation of this residue changed the binding and the catalytic efficiency of the nicotinamide cofactor. Substitution of the equivalent residue of the rat ALDH3a1 in either Glu140 (Uniprot: P11883) by Asn, Gln or Thr favored $NADP^+$ binding as shown by catalytic efficiency [41]. On the other hand, mutating the equivalent residue, Thr175 in the $NADP^+$-dependent *Vibrio harveyi* ALDH (VIBHA_ALDH Uniprot: Q56694) to Glu or Gln strengthened interaction with NAD^+ [42]. The mutation of the adjacent Lys (Lys137 in ALDH3a1) to Glu affected cofactor-binding abilities negatively [41]. Similar observations were made when the equivalent residue in human liver mitochondrial ALDH2 was mutated (K192Q) [43]. In contrast, mutation of the last residue of the glycine rich sequence, Gly_1_X-X-X-X-Gly_2 in ALDH3a1, (Gly192) did not affect enzyme efficiency [41].

Very little is known about biochemical characteristics of families 2 and 3 plant ALDHs. Most biochemical studies in plants consist of kinetic analyses where substrate specificities were investigated taking into account the aliphatic chain length or their degree of unsaturation [34,35]. *Arabidopsis* family 3 ALDHs (ALDH3H1 and ALDH3I1) were found to be able to oxidize medium- to long-chain aliphatic aldehydes, with a preference for long-chain aldehydes. Affinities were also found to be generally lower with unsaturated than saturated aldehydes [34].

Crystal structures of ALDHs from different organisms were solved in their apo or holo forms, but there are very few reports on structural data of plant ALDHs. Homology-based structural reconstitutions of rice and maize ALDHs have been reported [44,45]. Recently, the first crystal structure of a chloroplastic ALDH10 from spinach SoBADH (Uniprot: 17202, PDB: 4A0M) was solved and amino acid residues critical for high BADH activity and substrate affinity have been identified [46].

Information is scarce about amino acid residues which are involved in modulating plant ALDH activities or which are critical in nicotinamide cofactor recognition and binding. We recently identified residues involved in redox regulation of the family3 stress-responsive *A. thaliana* ALDHs, ALDH3H1 and ALDH3I1 and evidence was provided that Ile200 is a key residue responsible for the NAD^+-dependency in *A. thaliana* ALDH3H1 [34].

In this study, we continued our investigation in cofactor binding modalities in the *A. thaliana* ALDH3H1 enzyme based on knowledge of cofactor specificities in other ALDHs, sequence comparisons, and site-directed mutagenesis. We engineered an active $NADP^+$-specific enzyme, where substitutions were introduced in the cofactor binding pocket. Alteration of nicotinamide cofactor specificity provided new insights into plant family 3 ALDHs, as it identified amino acids which determine nicotinamide cofactor binding sites.

2. Results

2.1. Sequence alignment and prediction of amino acid residues involved in cofactor binding sites

Recently we have demonstrated that the amino acid residue Ile200 is responsible for the Arabidopsis ALDH3H1 NAD^+-dependency [34]. Ile200 is specific for ALDH3H1 and is replaced by Val in enzymes which are able to use NAD^+ or Gly in the $NADP^+$-dependent ALDH of *V. harveyi* VIBHA_ALDH, while it is present in the $NADP^+$-dependent glyceraldehyde-3-phosphate dehydrogenase from *Streptococcus mutans* GAPN_STRMU (Fig. 1). We had indications that besides Ile200 also other amino acids contribute to the cofactor interaction. Therefore sequence alignments and cofactor binding site structure comparisons were performed to identify amino acids which are additionally involved in the ALDH3H1 cofactor binding.

Amino acid sequences encompassing the cofactor binding site were aligned from the following enzymes: the $NADP^+$-preferring ALDHs *Rattus norvegicus* ALDH3a1 (UniProt: P11883), *Homo sapiens* ALDH3A1 (UniProt: P30838), the $NADP^+$-dependent *S. mutans* glyceraldehyde-3-phosphate dehydrogenase (UniProt: Q59931), $NADP^+$-specific fatty aldehyde dehydrogenase from *V. harveyi* (UniProt: Q56694), NAD^+-dependent *A. thaliana* ALDH3H1 (UniProt: Q70DU8), NAD^+-preferring *A. thaliana* ALDH3I1 (UniProt: Q8W033) and the NAD^+-dependent aldehyde dehydrogenase from *Rhizobium meliloti* (UniProt: Q930S8). These alignments revealed the presence of a Thr residue in the $NADP^+$-dependent dehydrogenases in a position where Glu is present in ALDHs which use NAD^+ (Fig. 1) and site-directed mutagenesis experiments showed that the Thr residue contributes to the $NADP^+$-dependency [38,42].

Secondary and tertiary structural models of *Arabidopsis* ALDH3H1 were built using crystal structures of rat ALDH3a1 (PDB: 1AD3) and human ALDH3A1 (PDB: 3SZB) as templates (Fig. 2). The stereochemical quality of the predicted protein structure was evaluated by analyzing residue-by-residue geometry and overall structure geometry using tools as ANOLEA energy [47], GROMOS force field energy [48] provided by SWISS-MODEL server and PROCHECK [49] which attest favorable energy environment showing the absence of misfolding of the whole enzyme and the absence of abnormal packing energy or aberrant geometries for amino acids of the cofactor binding site.

Residues from 10 to 90 form five helices α1 (10–25), α2 (27–26), α3 (32–60), α4 (64–69) and α5 (73–90). The strand β0 (105–112) forms an arm-like stretch which bridges the α5 helix and the cofactor binding domain. The open α/β nicotinamide cofactor binding Rossmann fold consists of the five β-strands β1 (115–119), β2 (142–146), β3 (170–174), β4 (191–195) and β5 (215–218) which are interspersed by four α-helices αA (127–138), αB (152–165), αC (178–186) and αD (198–211) (Fig. 2A). The glycine-rich motif Gly_1-X-X-X-Gly_2, positioned between β4 and αD and characteristic for ALDHs corresponds to the sequence GSSKIG (residues 196–201). The catalytic region is made up of six parallel β-strands β6 (224–227), β7 (258–261), β9 (332–334), β10 (352–358), β11 (376–380), β12 (398–401), the antiparallel strand β8 (315–318) and six helices α10 (233–244), α11