Chemico-Biological Interactions 234 (2015) 45–58

Contents lists available at ScienceDirect

Chemico-Biological Interactions

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

Amino acid residues that affect the basicity of the catalytic glutamate of the hydrolytic aldehyde dehydrogenases

Rosario A. Muñoz-Clares [a,*], Lilian González-Segura [a], Héctor Riveros-Rosas [b], Adriana Julián-Sánchez [b]

[a] Departamento de Bioquímica, Facultad de Química, Universidad Nacional Autónoma de México, México D. F. 04510, Mexico
[b] Departamento de Bioquímica, Facultad de Medicina, Universidad Nacional Autónoma de México, México D. F. 04510, Mexico

ARTICLE INFO

Article history:
Available online 21 January 2015

Keywords:
Crystal structures
Theoretical pK_a
Catalytic mechanism
Phylogenetic analysis
Multiple sequence alignments

ABSTRACT

In the catalytic mechanism of hydrolytic aldehyde dehydrogenases (ALDHs) the role of Glu268 (mature human ALDH2 numbering) as a general base is of major relevance. Since Glu268 basicity depends on its protein environment, here we explore its interactions with other amino acid residues in the three different conformations observed in ALDH crystal-structures: "inside", "intermediate" and "outside". In all of them Glu268 is in a hydrophobic environment. In the "inside" conformation, the theoretical pK_a estimated by PROPKA3 is the result of the effects of hydrogen bonds with the protonated thiol of the catalytic Cys302 and/or the main-chain amide nitrogen of the highly conserved Gly270, and of charge-charge interactions with neighboring side-chains—Lys178, Glu/Asp476, His465 or Glu399 depending on the enzyme. In the "intermediate" conformation Glu268-carboxyl pK_a is influenced by interactions with Glu/Asp476, Arg/Lys475, Lys/Arg178, His465 or Arg459, also depending on the enzyme. In the "outside" conformation, the effects on Glu268-carboxyl pK_a arise from hydrogen bonds with the side chains of the strictly conserved Thr224 and/or of Lys/Arg178, and from charge-charge interactions with Lys/Arg/Asp178, Glu476, or Arg459. The estimated pK_as and interactions of Glu268-carboxyl in the "intermediate" and "outside" conformations are consistent with their previously proposed roles in activating the hydrolytic water and in a proton relay mechanism, respectively. Water channels connecting Glu268 with the bulk water were found in all hydrolytic ALDHs. In the "inside" conformation the theoretical pK_as of the Glu268-carboxyl and Cys302-thiol groups suggest that the carboxyl cannot receive the proton from the thiol. We propose that a protonated Cys302 might perform the nucleophilic attack on the aldehyde, which can be facilitated by Glu268 in the "intermediate" conformation. Finally, the conservation of the residues influencing Glu268 basicity between and within ALDH families suggests that these residues, not previously studied, are important for the catalytic mechanism of many ALDH enzymes.

1. Introduction

In all organisms throughout the phylogenetic scale, the oxidation of aldehydes to their corresponding acids is mainly catalyzed by aldehyde dehydrogenase (ALDH) enzymes, which group into three categories: phosphorylating, hydrolytic, and CoA-acylating. These enzymes use NAD(P)$^+$ as cofactor and their catalytic mechanism involves the formation of two intermediaries covalently bound to the enzyme: a thiohemiacetal resulting from the nucleophilic attack of the catalytic cysteine on the carbonyl carbon of the aldehyde substrate, and a thioester resulting from the oxidation of the thiohemiacetal, which transfer a hydride to NAD(P)$^+$. The release of the acid product of the reaction requires the breakage of the thioester bond in the deacylation step, which is achieved by phosphorolysis in the case of phosphorylating ALDHs, hydrolysis in the case of hydrolytic ALDHs, or thiolysis by Coenzyme A in the case of CoA-acylating ALDHs. Evolutively, these enzymes belong to two different superfamilies [1]: the glyceraldehyde 3-P dehydrogenase superfamily and the canonical ALDH superfamily, which includes the hydrolytic and CoA-acylating enzymes.

In spite of the wealth of information regarding the structure and catalysis of the enzymes of the ALDH superfamily, there are still several aspects that have not been studied in depth. Of major relevance are the structural factors that allow the proposed roles of the catalytic Glu (Glu268, *Hs*ALDH2 numbering, which is the one used throughout the paper) as a general base in the hydrolytic ALDHs activating the catalytic thiol and/or the hydrolytic water molecule by abstracting a proton from either of them (reviewed in [2]). The activation of the hydrolytic water is essential for the

Abbreviations: ALDH, aldehyde dehydrogenase; *Hs*ALDH2, mature human ALDH2; PDB, Protein Data Bank.

* Corresponding author. Tel.: +52 55 56225276.
E-mail address: clares@unam.mx (R.A. Muñoz-Clares).

http://dx.doi.org/10.1016/j.cbi.2015.01.019
0009-2797/

deacylation step in these enzymes, as indicated by the strict conservation of Glu268 in them and its absence in CoA-acylating ALDH enzymes, such as those of the ALDH6 and ALDH20 families. However, the role of Glu268 in activating the catalytic thiol is somehow controversial, or at least not general [3–5]. In any case, for an efficient transfer of a proton from either the catalytic thiol or a water molecule to Glu268, the pK_a of its side-chain carboxyl group must be significantly higher than the typical value of ~4.5 in water, and higher than, or at least similar to, that of the thiol group or the water molecule. This increase in Glu268-carboxyl pK_a could be in part achieved by the hydrophobic environment of the active site of ALDH enzymes, but other effects arising from specific interactions of the carboxyl group with neighboring amino acid residues surely are also important in determining its pK_a value. These interactions have not been studied so far, and therefore how the pK_a perturbation of Glu268 is achieved in the active site of the hydrolytic ALDH enzymes is not fully understood.

It is well accepted that, during the catalytic cycle, the side chain of Glu268 may sample different protein environments that would affect the basicity of the carboxyl group. In the reported crystal structures of hydrolytic ALDHs, Glu268 has been observed in three stable conformations, which we named "inside", "intermediate", and "outside" [6] on the basis of the distance of the carboxyl group from the catalytic Cys302-thiol, being the "inside conformation" the closer one. The importance of the conformational flexibility of the side-chain of Glu268 for the different catalytic roles that this residue can perform has been previously discussed [7]. It has been proposed that Glu268 participates in activation of the catalytic cysteine, i.e., in forming the thiolate anion, when it adopts the "inside" conformation; in the activation of the hydrolytic water molecule, i.e., in forming the hydroxide anion, when it adopts the "intermediate" conformation; and in the release of the proton taken from either nucleophile, when it adopts the "outside" conformation (reviewed in [2] and [7]). These roles depend not only on the position of the Glu268 carboxyl group inside the active site but also on the pK_a of this carboxyl, which in turn is mainly determined by its environment. Undoubtedly, the ratio of protonated/deprotonated carboxyl species will directly impact on the rate of the catalytic steps in which Glu268 in any of its three conformations participates, and the different environments of the carboxyl group in the active site will affect this ratio.

In the present work, we aimed to get a deeper understanding of these important mechanistic aspects of the reaction catalyzed by hydrolytic ALDHs, by: (i) investigating the different interactions of Glu268-carboxyl group in any of the three conformations observed in the available ALDH crystal structures; (ii) analyzing how these interactions may affect the carboxyl group basicity–which we assessed by estimating its theoretical pK_a value—and therefore its probable function; and (iii) performing multiple amino acid-sequence alignments in order to know the degree of conservation of the residues identified in the previous analyses, and therefore to be able to infer how general in the ALDH superfamily the interactions observed in the crystal structures are. Finally, we discuss the consistency of the results of these analyses with the possible roles of Glu268.

2. Methods

2.1. Structural comparisons and theoretical pK_a determinations

The known three-dimensional structures of ALDH enzymes from different ALDH families—74 non-redundant crystal structures from 61 organisms—were compared using PyMOL (http://www.pymol.org) and Coot [8]. Water channels were searched in the crystal structures using CAVER 3.0 [9]. Theoretical calculations of the pK_a values of side-chain groups were carried out using the program PROPKA versions 3.0 and 3.1 [10,11]. The latter includes the effects of protein ligands in the pK_a estimation, while the former only considers the protein residues. None of the two PROPKA versions take into account the effects of ordered water molecules.

2.2. Sequence analyses

ALDH amino acid sequences were retrieved by Blastp searches at the UniProt site [12] (http://www.uniprot.org./blast). To identify protein sequences as members of the different reported ALDH families we carried out phylogenetic analyses using the MEGA6 software [13] (http://www.megasoftware.net). The unrooted phylogenetic tree was inferred from 500 replicates, using the Maximum Likelihood method. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 1.0567). Progressive multiple protein sequence alignments were performed with ClustalX version 2 [14] (http://www.clustal.org). Analysis of amino acid residues at selected positions in the alignment was performed with ProfileGrid [15] (http://www.profilegrid.org). Protein ALDH families are defined according to criteria established by the ALDH Gene Nomenclature committee [16] as described in Black and Vasiliou [17], with the exception of the ALDH14 family described by Peng et al. [18], and ALDH25 ALDH26 and ALDH27 described by Riveros-Rosas et al. [19]. The ALDH14 family includes orthologues to the *phnN* gene product from *Sphingomonas* sp. 14DN61, which has broad substrate specificity towards various aromatic and short-chain aliphatic aldehydes [18]. The ALDH25 family comprises orthologues to enzymes exhibiting betaine aldehyde dehydrogenase (BADH) activity characterized in the Gram positive bacteria *Bacillus subtilis* [20], *Staphylococcus xylosus* [21], and *Staphylococcus aureus* (PDB code 4QTO). The ALDH26 family consists of orthologues to the *ycdW* gene product from *Escherichia coli*, which has aminoaldehyde dehydrogenase activity [22,23]. The ALDH27 family contains orthologues to the *kauB* gene product from *Pseudomonas aeruginosa* PAO1 [24,25] and *puuC* gene product from *E. coli* [26], both exhibiting aminoaldehyde dehydrogenase activity. Because some characterized ALDHs with a resolved crystal structure do not belong to any of these previously reported ALDH families [16–19], the protein family assigned by the Conserved Domain Database (http://www.ncbi.nlm.nih.gov/cdd) was given. This Conserved Domain database was used because the protein classification included in it agrees with the previously reported ALDH families and with the results of the phylogenetic analysis performed here. Sequence logos were constructed using the WebLogo server [27] (http://weblogo.threeplusone.com).

3. Results and discussion

3.1. Interactions of the catalytic glutamate in the "inside" conformation

In this conformation, the carboxyl group of Glu268 is buried and the effects of desolvation estimated by PROPKA3 result in a theoretical pK_a around four pH units above the typical 4.5 value of a carboxyl group exposed to water. However, in many of the structures that have this conformation of Glu268, this pK_a is lowered by hydrogen bonds with the main-chain amide nitrogen of Gly270 and/or the side-chain thiol group of the catalytic Cys302 (Fig. 1 and Table 1). The hydrogen bond between Glu268-carboxyl and the amide nitrogen of Gly270 can be formed only when the 269–270 peptide bond has a rotated ("flipped") conformation, as observed for the first time in the crystal structure of the ALDH4 from *Thermus termophilus* (PDB 1UZB) [28]. The presence of a Gly