PALACKÝ UNIVERSITY OLOMOUC

Faculty of science

Department of Biochemistry

Study on three aldehyde dehydrogenase families from moss (*Physcomitrella patens*)

Author:	**Bc. Jan Vilím**
Study programme:	B1406 Biochemistry
Study branch:	Biochemistry
Form of study:	Full-time
Supervisor:	**Mgr. Martina Kopečná, Ph.D.**
Year:	2015

I hereby declare that I have written this thesis solely by myself and all the sources used in the thesis are properly cited and listed. I agree with publishing this thesis according to the law no. 111/1998 Sb., including the later additions. I acknowledge the rights and obligations resulting from the law no. 121/2000 Sb., including the later additions.

In Olomouc ....

I would like to thank Mgr. Martina Kopečná, Ph.D. and Mgr. David Kopečný for the three years that I have spent under their supervision. It was a great journey that gave me insight to the world of scientific research and countless new information and great experiences. Further I would like to thank Mgr. Radka Končitíková for help with various problems that emerged during the work in our lab, Pierre Briozzo for kind supervision during my stay in Versailles, Mgr. Filip Dyčka, Ph.D. for help with MALDI-TOF analysis and of course to my family and friends.

Bibliografická identifikace

Jméno a příjmení autora	Jan Vilím
Název práce	Studium tří rodin aldehyddehydrogenas z mechu (*Physcomitrella patens*)
Typ práce	Diplomová
Pracoviště	Katedra biochemie
Vedoucí práce	Mgr. Martina Kopečná, Ph.D.
Rok obhajoby práce	2015
Abstrakt	Byla vypracovana literarni reserse shrnující poznatky o mechu *Physcomitrella patens* a nadrodině aldehyddehydrogenas. Byla izolována RNA, která byla následně přepsána do cDNA a bylo provedeno klonování *PpALDH10*, *PpALDH12* a *PpALDH21*. Rekombinantní enzymy byly produkovány v expresních buňkách *Escherichia coli*. Při nedostatečné produkci byla provedena optimalizace volbou vektoru a odstraněním signální sekvence. Tyto změny umožnily dvou krokovou purifikaci enzymů a jejich následnou molekulární a biochemickou charakterizaci. Mezi stanovované vlastností patří oligomerizační status, pH optimum, termostabilita, substrátová specificita a kinetické parametry. Nakonec byly krystalovány dvě zkoumáné PpALDH a byla získána rentgenostrukturní data pro PpALDH21 při rozlišení až 2.8Å.
Klíčová slova	aldehyddehydrogenasa, mech, purifikace, enzymová kinetika, krystalizace
Počet stran	68
Počet příloh	-
Jazyk	anglický

Bibliographical identification

Author's first name and surname	Jan Vilím
Title	Study on three aldehyde dehydrogenase families from moss (*Physcomitrella patens*)
Type of thesis	Diploma
Department	Department of Biochemistry
Supervisor	Mgr. Martina Kopečná, Ph.D.
The year of presentation	2015
Abstract	Theoretical part summarizes known data related to a moss *Physcomitrella patens* and aldehyde dehydrogenase (ALDH) superfamily. Practical part deals with cDNA preparation and cloning of three *ALDH* genes from Physcomitrella, namely *PpALDH10*, *PpALDH12* and *PpALDH21*. Recombinant enzymes were produced in *Escherichia coli* expression cells. To counter very low production of the first two ALDHs, optimization was performed by modifying putative signal sequences and used vector. These changes allowed for two-step purification and further the molecular and biochemical characterization of all three PpALDHs. Oligomerization state, pH optimum, thermostability, substrate specificity and kinetic parameters were determined. Finally, two studied PpALDHs were crystallized and X-ray diffraction data for PpALDH21 were collected up to 2.8 Å resolution.
Keywords	aldehyde dehydrogenase, moss, purification, enzyme kinetics, crystallization
Number of pages	68
Number of supplements	-
Language	English

Table of content

1. Aims of the work

- A review of known data related to *Physcomitrella patens* and aldehyde dehydrogenase superfamily.

- Cloning of *PpALDH10*, *PpALDH12* and *PpALDH21* followed by the expression and purification steps.

- Biochemical characterization of produced enzymes (thermal stability, pH optimum) and determination of substrate specificity and kinetic parameters.

- Crystallization of produced ALDHs.

2. Introduction

Aldehyde dehydrogenase (ALDH) superfamily is composed of wide variety of enzymes catalyzing conversion of various aldehydes in NAD^+/$NADP^+$ dependent manner. In recent years, the main aim of research concerning the plant ALDHs at our Department in Olomouc were enzymes belonging to ALDH10 family – aminoaldehyde dehydrogenases or betaine aldehyde dehydrogenases due to their involvement in glycine betaine (GB) and γ-aminobutyric acid (GABA) formation. This work is focused on three ALDHs, namely ALDH10, ALDH12 and ALDH21 from moss *Physcomitrella patens*, which has been used as a plant evolution model, more specifically, for uncovering the conquest of land by plants. The non-vascular plants including moss contain up to date uncharacterized family ALDH21. There is no information about substrate preferences or physiological role. Other studied ALDHs include those from family 10 and 12 because there are no kinetic data for non-vascular plant ALDH10 and ALDH12 proteins. Obtained data were compared to those from higher plants including ALDH10 members from *P. sativum, Z. mays* and *Solanum lycopersicum* with known crystal structures (Tylichová *et al.*, 2010; Kopečný *et al.*, 2011; Kopečný *et al.*, 2013).

Theoretical background

2.1. Physcomitrella patens

Physcomitrella patens is a small green moss (bryophyte) used widely as model for plant evolutionary studies. Its genome has been published in 2008 (Rensing *et al.*, 2008). *P. patens* also stands at the beginning of reverse genetics era, because of the discovery of highly efficient homologous recombination (Kammerer and Cove, 1996). With the development of reliable transformation methods, the possibilities of genetic engineering and with the mosses being the intermediate between primitive unicellular photosynthetic organisms and terrestrial plants, evolutionary studies became focused on *P. patens*. With many techniques to analyze gene function, these things together gave rise to the new versatile model plant (Strobek *et al.*, 2013).

2.1.1. Life cycle

The *P. patens* life cycle is like of the other mosses different from seed plants. The life cycle composes of two alternating heteromorphic and heterophasic generations (Figure 1A), with the gametophyte representing the predominant generation. Haploid spore germination starts the gametophytic generation (Figure 1B). The spore forms a branched filamentous protonema tissue (Figure 1C). Two different cell types are present in protonema: chloronema – the chloroplast enriched cells with perpendicular cross walls and caulonema – cells containing few chloroplasts and oblique cross-walls. Transition from protonema to gametphore is initiated by formation of meristematic buds with three-faced apical cells. Gametophores develop from these buds (Figure 1D) and consist of shoot-like stem with leaf shaped phylloids and rhizoids (Figure 1E). *P. patens* is a monoecious species, so the sexual reproduction is initiated when both a male (antheridia) and female (archegonia) sex organs at the tip of the gametophore are developed (Figure 1F). Spermatozoids formed in antheridia are able to move in the presence of water and their aim is to fertilize single egg cell inside the archegonium. The diploid sporophyte develops from the zygote formed after the fertilization. It builds up a short seta with spore capsule where haploid spores are produced by meiosis (Figure 1G) (Strobek *et al.*, 2013).

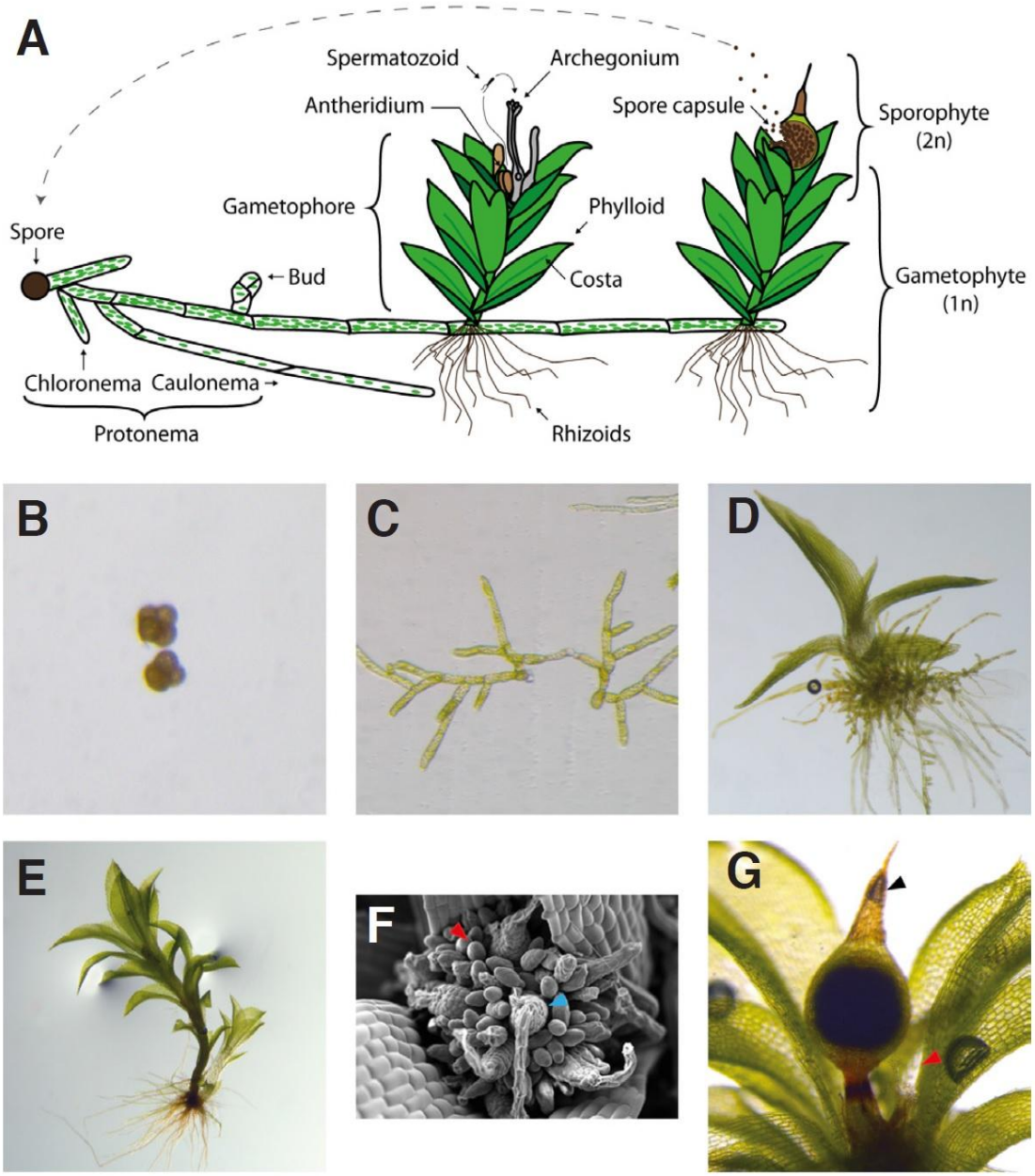

Figure 1. Life cycle of *P. patens* and its phases. (A) Scheme of *P. patens* life cycle. (B) Tetrades of haploid spores originated from meiotic cell division of mother spores released by the spore capsule of sporophyte. (C) Branched protonema cells grown in liquid culture. (D) Gametophore in development. (E) Gametophores at late (left) and early (right) development stages. (F) Top view of gametangia by scanning electron microscopy, antheridia is indicated by red and archegonia by blue arrow. (G) Fully developed sporophyte with calyptra (black arrow) and unfertilized archegonia (red arrow) (Modified from Strobek *et al.*, 2013).

2.1.2. Model for evolutionary studies

Mosses occupy key position on evolutionary track from aquatic plants to terrestrial plants associated with acquiring the special body features separating vascular and primitive single cell photosynthetic organisms. 500 million years ago (MYA), within the Charophyta evolved the first feature of higher plants body: a three-dimensional and multicellular gametophyte consisting of cell wall lined cells interconnected with plasmodesmata (Figure 2). Embryophyta (the first land plants) evolved multicellular diploid sporophyte 470 MYA and later on, 440 MYA, the Tracheophyta gained special water conducting cells giving base to extended longitudinal growth. First seeds evolved approximately 400 MYA, followed by first flowers (Magnoliophyta) 100 MYA later. It is understandable that models representing key evolutionary steps in plant development were chosen for genome sequencing (Floyd and Bowman, 2007).

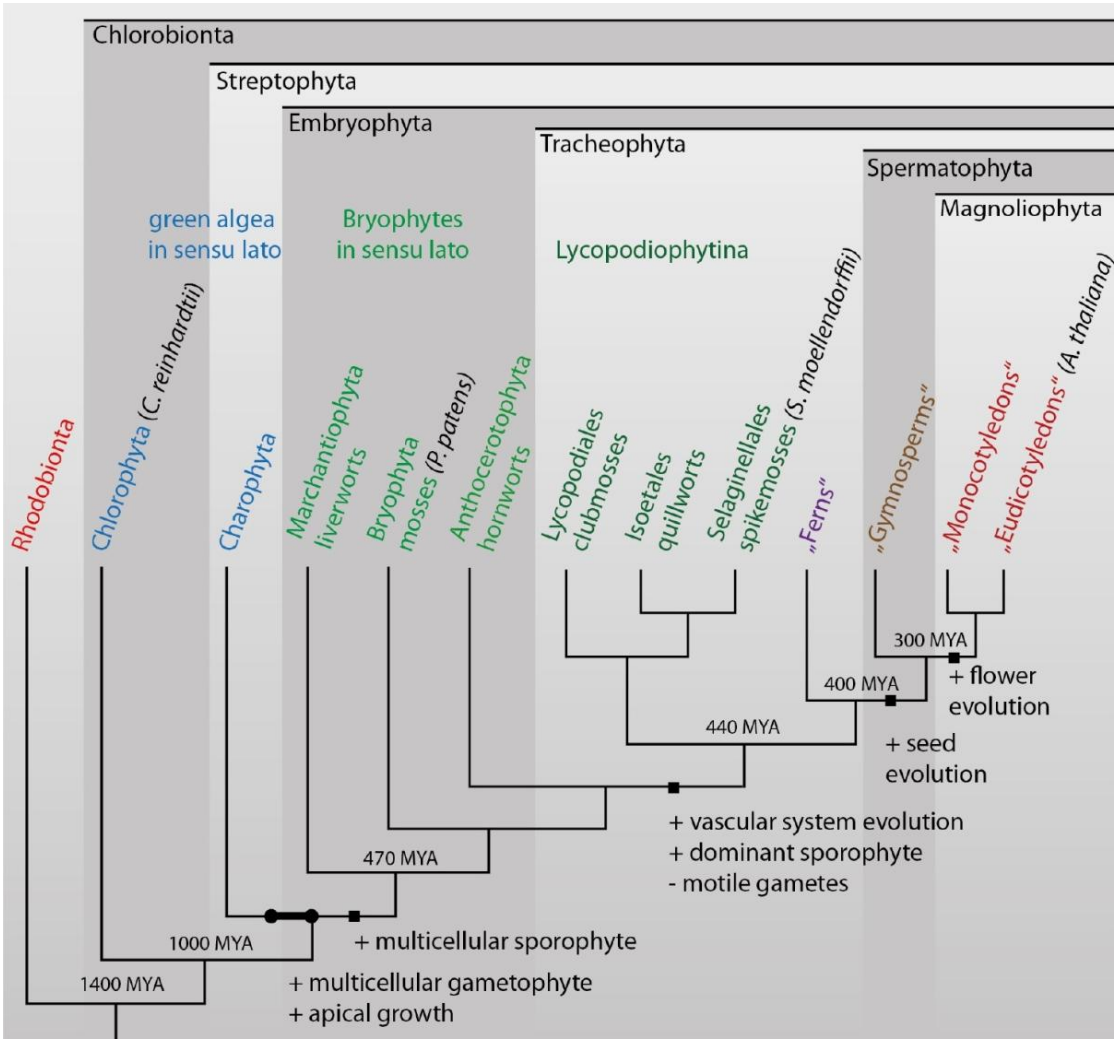

Figure 2. Plant evolution. Phylogenetic tree representing evolutionary relationships from red algae (Rhodobionta) to angiosperms (Magnobiontha). Listed model organisms represent key stages in plant evolution (Strobek *et al.*, 2013).

Concerning the evolutionary studies, *P. patens* with other used models such as aquatic green algae representative *Chlamydomonas reindhartii*, early tracheophyte representative *Selaginella moellendorfii* and high seed plants model *Arabidopsis thaliana* cover plant evolutionary route (Figure 2), thus allowing study of features necessary to adapt to the terrestrial life and of course abiotic stresses associated with limited water supply on land, strong sun light irradiation and varying temperatures (Floyd and Bowman, 2007). Comparative analysis of gene function in species representing different evolutionary steps allows to differentiate gene families with highly conserved function through the evolution and novel gene families that occurred as adaptation process.

2.1.3. Biotechnological potential

P. patens is a very versatile plant. Apart from being the model for evolutionary studies, it can be also utilized as a model in biotechnological research, particularly in therapeutic and diagnostic protein expression (Koprivova *et al.*; 2004, Parsons *et al.*, 2012). The group of recombinant protein produced in plants is increasing in numbers and some plant-made pharmaceuticals have already reached the clinical trials stage (Faye and Gomord, 2010). *P. patens* being multicellular eukaryote has advantage over unicellular expression system such as bacteria and yeast in ability to perform extensive posttranslational modifications including glycosylation, disulfide bridge formation and multimeric assembly (Raskin *et al.*, 2002; Gomord and Faye, 2004). Other positive sides of using moss expression system can be reduction of processing cost and exclusion of possible contamination by human pathogens (Raskin *et al.*, 2002; Fischer *et al.*, 2004). Important disadvantage of recombinant proteins expressed in plants is a possible allergic reaction after application due to the presence of different *N*-glycans (Foetisch *et al.*, 2003). However, *P. patens* knockout strain without the plant specific *N*-glycosyl transferases was generated and expressed proteins lack immunogenicity (Koprivova *et al.*, 2005). Furthermore, improved moss strain with incorporated 1,4-galactosyl transferase instead of plant specific transferases was obtained. The result was a proper human-type protein lacking immunogenic residues (Huether *et al.*, 2005). Possibility of low-cost production of non-allergenic therapeutic proteins makes *P. patens* highly attractive not only to researches but also to pharmaceutical industry.

2.2. Aldehyde dehydrogenases (ALDHs)

Aldehyde dehydrogenase (ALDH) superfamily is a wide group of eukaryotic and prokaryotic enzymes catalyzing conversion of aldehydes to their respective carboxylic acids in NAD^+/$NADP^+$ dependent manner. Up to date, 26 distinct families have been reported and half of them are present in plants: ALDH2, ALDH3, ALDH5, ALDH6, ALDH7, ALDH10, ALDH11, ALDH12, ALDH18, ALDH21, ALDH22, ALDH23 and ALDH24. ALDH enzymes are present in every plant species, the number of genes and the family spectrum indeed differ from species to species. As a fact demonstrating huge diversity in plant ALDH superfamily is that families ALDH21, ALDH23 and ALDH24 appear to be restricted to the primitive terrestrial plants and that the *Volvox carteri* algae contains only 7 *ALDH* genes and families (Brocker *et al.*, 2013).

2.2.1. Nomenclature of ALDHs

The ALDH superfamily nomenclature was established in 1999 by Vassiliou *et al.* It was developed following the Human Gene Nomenclature Comitee recommendations. The nomenclature is used for all species across prokaryotes and eukaryotes. The base for every gene name are letters „ALDH“, followed by family number written in arabic numeral, then, the capital letter stating respective subfamily and finally the number identifying the gene in subfamily. Prefix is formed from the first letters of the source organism name. Families are numbered chronologically according to the identification. Genes are written in cursive and capital letters are used, whereas proteins are written in capital letters only – e.g. ALDH10 gene from corn (*Zea mays*) - *ZmALDH10A9*. Classification is determined according to amino acid sequence similarity. Proteins sharing 40% of their amino acid sequence are placed into same family and proteins sharing 60% of the sequence are placed into same subfamily. Number of identified ALDH genes in whole plant genomes is growing, up to date there were characterized ALDH genes from corn (*Zea mays*) (Jimenez-Lopez *et al.*, 2010), rice (*Oryza sativa*) (Gao and Han, 2009), thale cress (*Arabidopsis thaliana*) (Kirch *et al.*, 2004), soybean (*Glycine max*) (Kotochoni *et al.*, 2012), cotton (*Gossypium raymondii*) (He *et al.*, 2014), foxtail millet (*Setaria italica L.*) (Zhu *et al.*, 2014), grape (*Vitis vinifera*) (Zhang *et al.*, 2012), *Physcomitrella patens*, *Chlamydomonas reinhardtii*, *Ostrococcus tauri* (Wood and Duff, 2009) and some others. Non-respecting of nomenclature and inaccurately identified genes bring confusion. Therefore genomes for some of the listed species were re-examined and renamed accordingly in comprehensive plant ALDH study by Brocker *et al.* (2013).

2.2.2. Role in metabolism

Compounds with the aldehyde moiety are generated in various metabolic pathways acting either as intermediates or emerging due to various stressful conditions. The aldehydes are very reactive and can be a source of great damage to organism; therefore cytoprotective action of ALDHs is very important. Aldehydes produced in oxidation of lipid membranes, otherwise known as lipid peroxidation (LPO), such as 4-hydroxynoneal, 4-oxononenal and malondialdehyde readily form adducts with nucleophiles (proteins, nucleic acids), and thus altering or disabling their normal function. In case of nucleic acids aldehydes cause mutations and disturb the homeostasis. The significant role of ALDHs in coping with stressing conditions has been proven many times, mostly using knock-out lines, overexpressing key proteins or by monitoring upregulation of gene expression during stress.

2.2.3. ALDHs in *P. patens*

21 ALDH genes were identified in genome of *P. patens* (Table 1) (Brocker *et al.*, 2013). That is one more than was identified by Wood (Figure 3) (Wood and Duff, 2009) due to newly added gene belonging to ALDH18 family. PpALDHs are divided into eleven families with four of them consisting of multiple members – ALDH2, ALDH3, ALDH5 and ALDH11. *P. patens* contains two families that were identified only in primitive terrestrial plants. ALDH21 family was identified in *P. patens* (Wood and Duff, 2009), *S. moellendorffii* (Brocker *et al.*, 2013), *Syntrichia caninervis* (Yang *et al.*, 2012) and *Tortura ruralis* (Chen *et al.*, 2002). ALDH23 gene was identified only in *P. patens* in one copy and in *S. moellendorffii* in two copies (Brocker *et al.*, 2013).

Table 1. ALDH genes in *P. patens*. List of ALDH genes in *P. patens* and putative function of encoded proteins. ALDH – aldehyde dehydrogenase; SSDH – succinic semialdehyde dehydrogenase; MMSDH – methylmalonyl semialdehyde dehydrogenase; AMADH – aminoaldehyde dehydrogenase; GANDH – non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase; P5CDH - Δ^1-pyrroline-5-carboxylate dehydrogenase; P5CS - Δ^1-pyrroline-5-carboxylate synthase (modified from Brocker *et al.*, 2013).

ALDH family	Gene name	NCBI Gene ID	NCBI Protein ID	Putative function	Number of AA
Family 2	*PpALDH2B1*	5930606	XP_001767457	ALDH [NAD(P)+]	553
	PpALDH2B2	5948865	XP_001785650	ALDH [NAD(P)+]	535
Family 3	*PpALDH3H1*	5928120	XP_001764841	ALDH [NAD(P)+]	492
	PpALDH3H2	5943334	XP_001780129	ALDH [NAD(P)+]	583
	PpALDH3K1	5920386	XP_001757163	ALDH [NAD(P)+]	497
	PpALDH3K2	5930362	XP_001767194	ALDH [NAD(P)+]	485
	PpALDH3K3	5933556	XP_001770374	ALDH [NAD(P)+]	467
Family 5	*PpALDH5F1*	5916037	XP_001752876	SSDH	492
	PpALDH5F2	5943671	XP_001780466	SSDH	498
Family 6	*PpALDH6B1*	5920495	XP_001757403	MMSDH	574
Family 7	*PpALDH7B4*	5941578	XP_001778351	Antiquitin	511
Family 10	*PpALDH10A1*	5919866	XP_001756623	AMADH	559
Family 11	*PpALDH11A5*	5935487	XP_001772261	GANDH	504
	PpALDH11A1	5928282	XP_001765101	GANDH	503
	PpALDH11A2	5918399	XP_001755163	GANDH	496
	PpALDH11A3	5916920	XP_001753784	GANDH	496
	PpALDH11A4	5923860	XP_001760657	GANDH	496
Family 12	*PpALDH12A1*	5923366	XP_001760169	P5CDH	571
Family 18	*PpALDH18B1*	5919578	XP_001756289	P5CS	757
Family 21	*PpALDH21A1*	5932362	XP_001769187	GANDH	497
Family 23	*PpALDH23A1*	5918738	XP_001755511	ALDH [NAD(P)+]	494

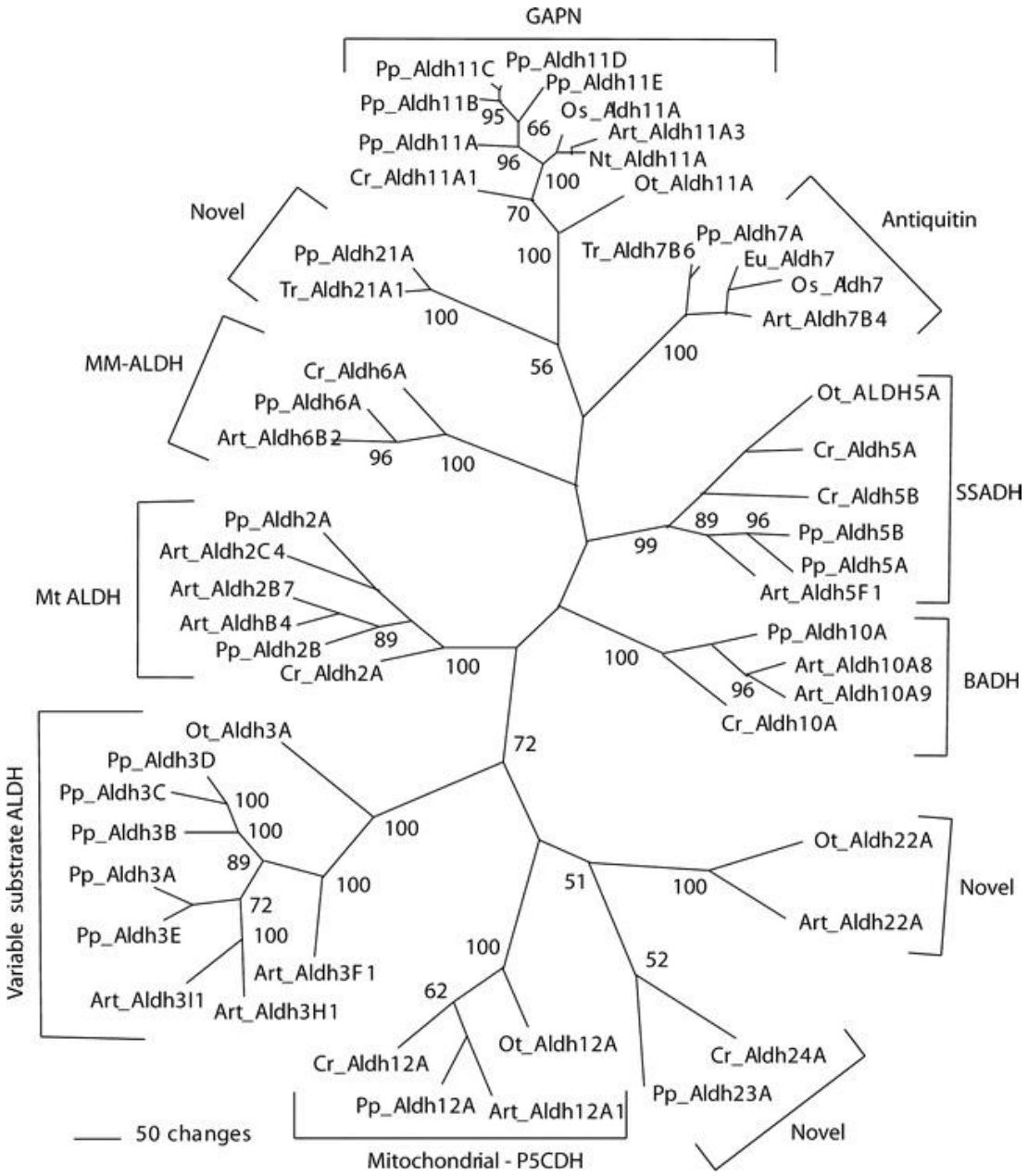

Figure 3. Phylogenetic relations in the ALDH superfamily of *P. patens* (Wood *et al.*, 2009). Numbers at nodes are bootstrap support percentages from 500 bootstrap replicates using maximum parsimony (upper) and neighbor-joining methods (lower). Dashes represent support of less than 50%. GAPN – glyceraldehyde-3-phosphate dehydrogenase, SSADH – succinic semialdehyde dehydrogenase, BADH – betaine aldehyde dehydrogenase., MM-ALDH – methylmalonyl dehydrogenase, Mt ALDH – mitochondrial aldehyde dehydrogenase.

2.2.4. ALDH10

Aminoaldehyde dehydrogenase (EC 1.2.1.19), ALDH10 family members, were in past years also known as 4-aminobutyryl dehydrogenase (ABALDH, EC 1.2.1.19), 4-guanidinebutyraldehyde dehydrogenase (GBALDH, EC 1.2.1.54), 4-trimethylaminobutyraldehyde dehydrogenase (TMABALDH, EC 1.2.1.47) and betaine aldehyde dehydrogenase (BADH, EC 1.2.1.8). All of AMADHs metabolize various aminoaldehydes emerging from several key metabolic pathways (Figure 4).

The ALDH10 family can be divided into two groups regarding the substrate specificity. First group preferentially oxidizes reactive ω-aminoaldehydes originating from polyamine catabolism – AMADHs from pea, maize or tomato (Tylichová *et al.*, 2010; Kopečný *et al.*, 2011 and 2013) – degradation of polyamines by polyamine oxidases (PAO) (Šebela *et al.*, 2000). The second group exhibits apart from ω-aminoaldehyde activity also high activity with betaine aldehyde (BAL), a precursor of the versatile osmo-protectant glycine betaine – AMADHs from mangrove, spinach or amaranth (Hibino *et al.*, 2001; Incharoensakdi *et al.*, 2000; Valenzuela-Soto a Munoz-Clares, 1993), and are therefore referred to as betaine aldehyde dehydrogenases (BADHs). Although ALDH10 proteins are plant specific, AMADHs are not. Their function in mammals belongs to the ALDH9 family, which preferentially oxidizes TMABAL. This determines the role of non-plant AMADHs, which mainly participate in the carnitine metabolism (Vaz *et al.*, 2000).

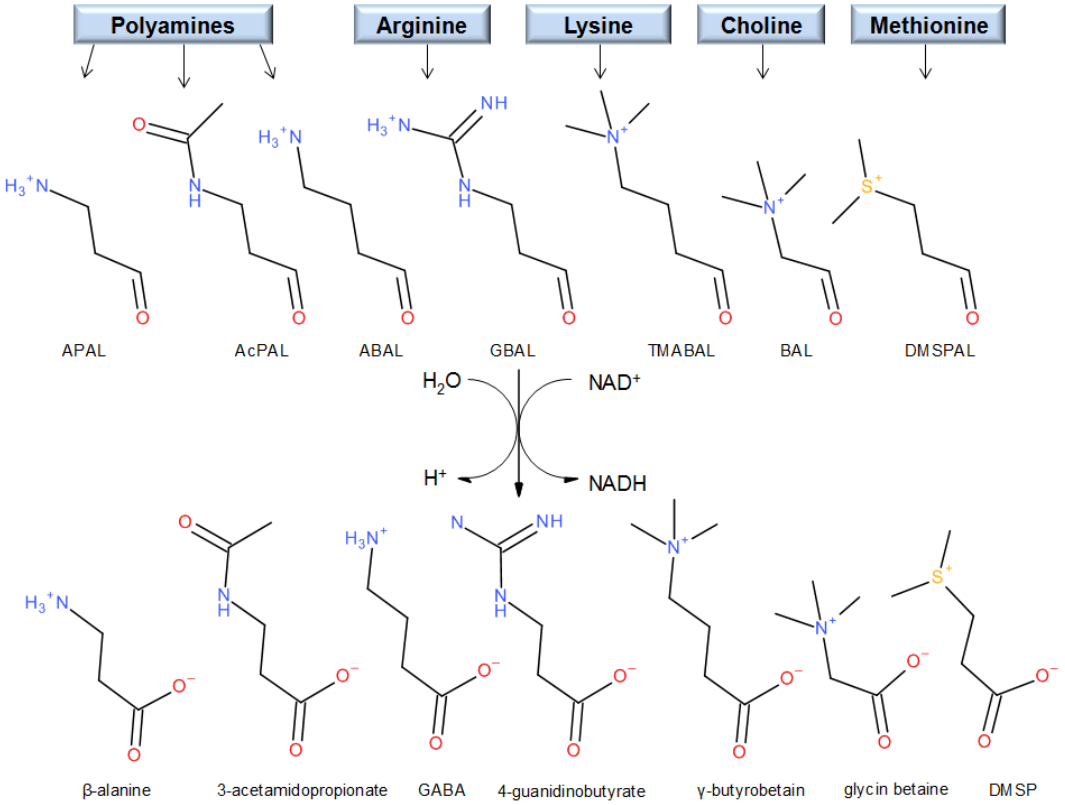

Figure 4. AMADH native substrates reaction and their metabolic origin. APAL – 3-aminopropionaldehyde, AcAPAL – *N*-acetyl-3-aminopropionaldehyde, ABAL – 4-aminobutyraldehyde, GBAL – guanidinobutyraldehyde, TMABAL – *N*,*N*,*N*-trimethyl-3-aminobutyraldehyde, BAL – betaine aldehyde, DMSPAL – S,S-dimethylsulfopropionaldehyde, GABA – γ-aminobutyric acid, DMSP – S,S-dimethylsulfopropionic acid (modified from Kopečný *et al.*, 2013).

Substrate specificity and active site residues

The interest in GB oxidation gave rise to the need to unravel the molecular basis of the substrate specificity. Up to date, several plant AMADH structures have been solved using X-ray crystallography – two AMADHs from pea (PsAMDH1, PDB number - 3IWJ and PsAMADH2, 3IWK, both Tylichová *et al.*, 2010), from spinach (SoBADH, 4A0M, Díaz-Sánchez *et al.*, 2012) and from maize and tomato (ZmAMADH1a, 4I8P, SlAMADH1, 4I9B, both Kopečný *et al.*, 2013). These crystal structures allowed for the identification of active site residues responsible for substrate binding.

Almost all so far characterized members of ALDH10 family have the highest activity with APAL, except for AMADH1 from tomato (SlAMADH1), which preferentially oxidizes 4-aminobutyryl aldehyde (ABAL). That is due to wider substrate channel caused by mutation of conserved cysteine residue in active site (C453→T, PsAMDHs numbering), which reacts with oligomerization domain of neighboring subunit thus leading to wider and less polar channel interior (Figure 5) (Kopečný *et al.*, 2013). The high activity rate with aminoaldehydes is caused by carboxylate amino acid residues at entrance to substrate channel and binding the protonated nitrogen. Mechanism was proven by site-directed mutagenesis, mutation of all three residues to alanine (E106→A, D110→A and D113→A) completely eradicated high activity and low K_m values for ω-aldehydes. Aromatic residues placed in substrate channel of AMADH active site anchor the carbon chain of the substrate *via* van der Waal's interactions - residues Y163 and Y170 (Kopečný *et al.*, 2011). Affinity to BAL lies apparently in amino acid residue on position 444 (in PsAMADHs). Residue on this position in AMADHs with low BAL affinity is isoleucine. In those AMADHs, which are able to oxidize BAL, I444 residue is mutated to alanine or cysteine enabling widening the substrate channel for bulky trimethylammonium moiety of BAL. Such widening of substrate channel is possible due to shift of T459 residue by 1.6 Å away from T163. The shift is not possible with I444 due to the steric hindrance (Kopečný et al, 2013). However, this mutation has undergone long evolutionary route. Originally, the I444 was in place and ability to utilize BAL was acquired by plants after *ALDH10* gene duplication occurred (Munoz-Clares *et al.*, 2014).

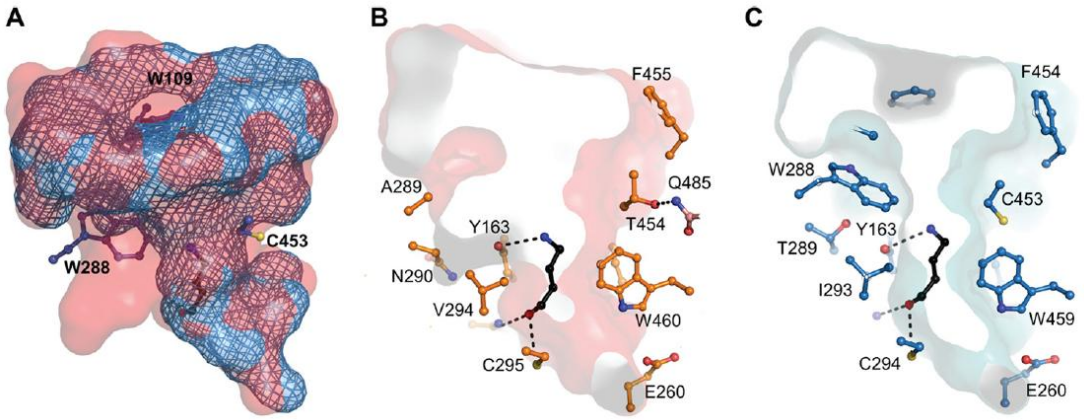

Figure 5. Substrate channels of SlAMADH1 and PsAMADH2. (A) Comparison of SlAMADH1 (red) and PsAMADH2 (blue) substrate channel cavities. (B) SlAMADH substrate channel with docked ABAL. (C) Substrate channel of PsAMADH2 with docked ABAL (Kopečný *et al.*, 2013).

Apart from main natural substrates, AMADH are able to oxidize wide spectrum of other aldehydes. SlAMADH1 shows broad substrate specificity thanks to previously mentioned wider substrate channel and oxidizes 3- and 4-pyridine carboxaldehydes at high rate (Kopečný *et al.*, 2013). Aldehydes derived from pyridine, purine, 7-deazapurine and pyrimidine are oxidized by AMADHs from pea and tomato. Their K_m values suggest that they probably are not *in vivo* substrates, but could be used when detoxifying exogenous aldehydes (Frömmel *et al.*, 2011). Moreover, AMADHs are able to oxidize APAL and ABAL derivatives including AcAPAL being *in vivo* substrate (Kopečný *et al.*, 2013, Frömmel *et al.*, 2015).

Coenzyme binding and preference

AMADH preferentially functions with NAD^+ than with $NADP^+$ (Tylichová *et al.*, 2010; Kopečný *et al.*, 2011 and 2013). In general, the coenzyme is bound in coenzyme cavity located on the opposite side to substrate channel (Figure 6A). Preferential binding of NAD^+ by various ALDH superfamily members is determined by the presence of conserved glutamate residue (E188 in PsAMADHs) sterically and electrostatically repelling the 2‘-phosphate group of ribose in $NADP^+$ (Perozich *et al.*, 2000). However, there are also other residues involved because ALDH members with glutamate can also use the $NADP^+$ (Perozich *et al.*, 2000, Velasco-García *et al.*, 2000). Different residues at respective position may shift the preference towards $NADP^+$ (Figure 6B) (González-Segura *et al.*, 2015). Also NAD^+ analogs can serve as coenzymes and for example the hypoxanthine derivative of NAD^+ (deaminoNAD^+) shows 1.5 fold higher reaction rate compared to NAD^+ with both PsAMADHs (Tylichová *et al.*, 2010).

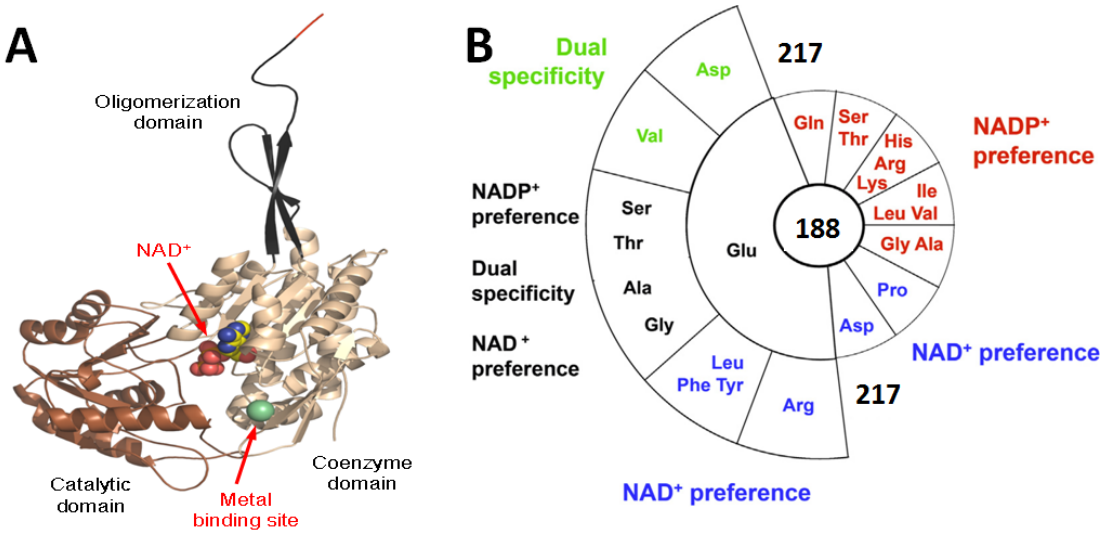

Figure 6. Coenzyme binding. (A) NAD^+ bound to subunit of PsAMADH. Oligomerization unit is depicted in black ribbon, catalytic domain in brown and coenzyme domain in beige, respectively. Sodium ion (green) and NAD^+ (atoms in respective element's color) are shown in balls (Tylichová *et al.*, 2010) (B) Schematic representation of possible coenzyme preference based on AA residues on positions 188 and 217 (PsAMADH numbering) (modified from González-Segura *et al.*, 2015).

AMADH in biotechnology

AMADHs have been in recent years extensively studied for their role in elimination of reactive ω-aminoaldehydes. The source compounds, polyamines, belong to the key molecules taking part in cell-division, embryogenesis, flowering, fruit-ripening and others (Buchanan *et al.*, 2000). On the contrary, ω-aminoaldehydes are cytotoxic (Agnostelli *et al.*, 2004; Wood *et al.*, 2007), which makes the ability of AMADHs to detoxify them almost essential. This is well documented on down-regulation of *OsAMADH1* gene in rice. These transgenic lines exhibit significantly reduced tolerance to drought, salinity and osmotic stress and also a lower crop yield (Tang *et al.*, 2014). Slower oxidation of substrates such as APAL and ABAL leads to the reduced formation of osmoprotectants (Trossat *et al.*, 1997) and it is likely the reason for the reduced stress tolerance.

Osmoprotectant glycine betaine (GB) is synthesized from BAL, product of choline oxidation. Choline can by synthesized in two ways, either from phosphoethanol amine methylated to phosphocholine and subsequently hydrolyzed to choline (spinach, sugar beet) or from membrane lipid phosphatidylcholine (barley) (Buchannan *et al.*, 2000). BAL is formed from choline by choline monoxygenase, which is a monomeric, ferredoxin dependent enzyme (Rathinasabapathi *et al.*, 1997).

GB can be utilized by all plants, but is not produced by all. Therefore, plants can be divided into GB accumulating and GB non-accumulating. Examples of accumulating plant species are spinach (*Spinacea oleracea*) or amaranth (*Amaranthus hypochondricanus*) and to

the non-accumulating group belong maize (*Zea mays*), but some maize cultivars accumulate the GB. GB non-accumulating plants include also rice (*Oryza sativa*) or tomato (*Solanum lycopersicum*) (Rhodes and Hanson, 1993). It is evident that the introduction of the GB synthetic pathway to the non-accumulators can be effective way to improve stress tolerance. This has been successfully performed with many different plant species, tobacco (Holmstrom *et al.*, 2000, Shen *et al.*, 2002, maize (Quan *et al.*, 2004) rice (Sakamoto *et al.*, 1998; Su *et al.*, 2006) and potato (Ahmad *et al.*, 2008). Also the introduction of AMADH with ability to oxidize BAL can significantly improve stress tolerance. Sweet potato expressing *BADH* from spinach was able to withstand harsher conditions than non-transgenic line (Fan *et al.*, 2012).

Another interesting feature of AMADHs is their involvement in fragrance in rice. The source of characteristic aroma of basmati or jasmin rice is called 2-acetyl-1-pyrroline (2-AP). The main cause of 2-AP accumulation is loss of AMADH activity due to a mutation in *OsBADH2* gene (Bradbury *et al.*, 2005, Bradbury *et al.*, 2008). The connection between accumulation of 2-AP and loss of AMADH activity was confirmed in soybean (*Glycine max*) (Arikit *et al.*, 2010). The inactivity of *OsBADH2* gene is widespread in fragrant rice varieties and probably causes their smaller tolerance to stresses when compared to non-fragrant varieties (Fitzgerald *et al.*, 2010). Also the transgenic soybean showed lower seed production when exposed to the salt stress having *AMADH2* gene inactivated (Arikit *et al.*, 2010).

2.2.5. ALDH12 (P5CDH)

ALDH12 gene codes for Δ^1-pyrroline-5-carboxylate dehydrogenase (P5CDH, EC 1.5.1.12), which catalyzes NAD^+-dependent oxidation of Δ^1-pyrroline-5-carboxylate (P5C) to glutamate (Figure 7). However, P5C is not oxidized in its cyclic form, in solution it exists in equilibrium with the glutamate semialdehyde (GSAL), which is the substrate for ALDH12. Mammals have also enzymes catalyzing the oxidation of P5C but due to low similarity with plant enzymes they are classified as ALDH4 family (Yoshida et al, 1998). P5CDH is a part of proline degradation pathway localized in mitochondria. The localization has been confirmed for P5CDH from rat liver (Hasslet *et al.*, 2004) and *A. thaliana* (Deuschle *et al.*, 2001) and also newly annotated genes are predicted code for mitochondrial ALDHs.

Proline metabolism in plants is located in cytosol and in mitochondria (Figure 8). In cytosol or chloroplast pathway synthesizing proline is located and proline degradation is localized in mitochondria. Proline synthesis is divided into two steps. First of all, GSAL (P5C) is synthesized from glutamate by Δ^1-pyrroline-5-carboxylate synthase (P5CS, EC 1.2.1.88). Then the proline is formed *via* Δ^1-pyrroline-5-carboxylate reductase (P5CR).

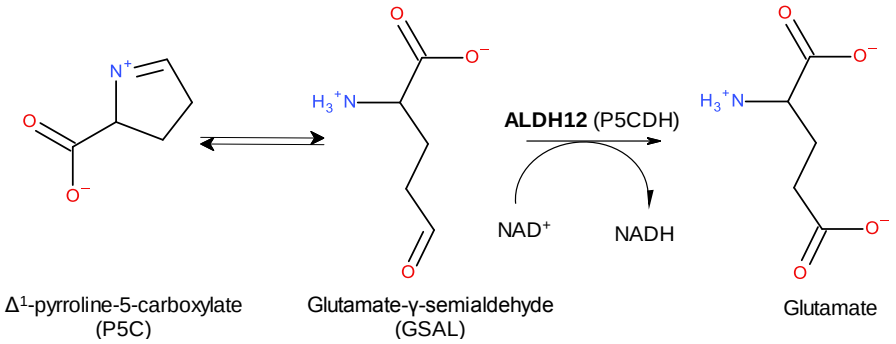

Figure 7. Reaction catalyzed by ALDH12. Reaction scheme depicting transformation of P5C to glutamate catalyzed by ALDH12 family members.

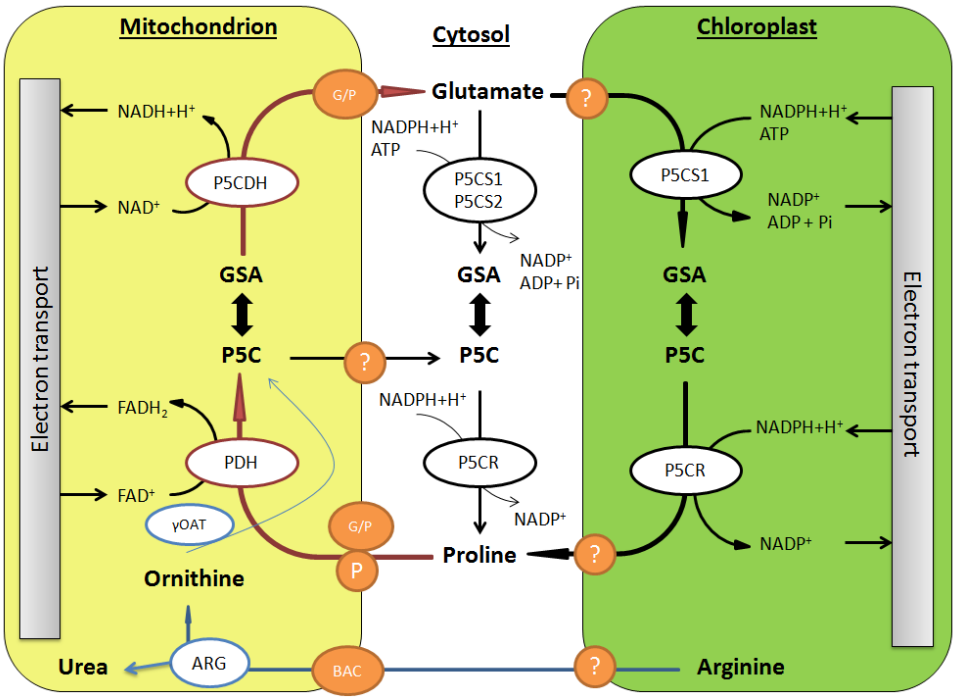

Figure 8. Model of proline metabolism in higher plants. Biosynthetic pathway is represented by black lines, catabolic pathway in red lines and ornithine pathway in blue lines. Enzymes are depicted as ellipses with white inside and transporter proteins are in orange. PDH – Proline dehydrogenase, P5CDH - Δ^1-pyrroline-5-carboxylate dehydrogenase, P5CR - Δ^1-pyrroline-5-carboxylate reductase, P5CS(1,2) - Δ^1-pyrroline-5-carboxylate synthase, ARG – arginase, γOAT – ornithine aminotransferase, BAC – Basic amino acid transporter involved in arginine and ornithine exchange, P – mitochondrial proline transporter, G/P – glutamate-proline antiporter, ? – predicted transporter (modified from Szabados and Savouré, 2010).

Degradation part of proline metabolism is mediated by proline dehydrogenase (PDH, EC 1.5.5.2) and emerging P5C (GSAL) is then oxidized to glutamate by P5CDH (Stein *et al.*, 2011). ALDH12 genes show high degree of sequence homology and analyzed plants usually have only one copy of *ALDH12* gene. Such a high degree of conservation between evolutionary distant species suggests existence of strong selective pressure to maintain gene function (Brocker *et al.*, 2013).

Proline role in plants

Proline is a proteinogenic amino acid accumulating in plants during various environmental stresses (Figure 9). Accumulation during drought (Choudhary *et al.*, 2005), high salinity (Yoshiba *et al.*, 1995), high light and UV irradiation (Saradhi *et al.*, 1995), heavy metals presence (Schat *et al.*, 1997), in response to oxidative stress (Yang *et al.*, 2009) and biotic stresses (Fabro *et al.*, 2004). However, the correlation between proline accumulation and the stress tolerance is not absolute. High levels of proline can be detected in cold hypersensitive *A. thaliana* mutants (Chen *et al.*, 2007) or drought tolerant rice (Choudhary *et al.*, 2005) but do not occur in salt stressed barley (Widodo *et al.*, 2009). Nevertheless, the proline metabolism has a complex effect on stress responses and development, and proline accumulation helps to deal with certain stresses inducing environmental conditions (Mattioli *et al.*, 2008; Székely *et al.*, 2008). Although the proline accumulation is one of the stress coping mechanisms, high concentration of proline in non-stressed plants is toxic (Hare *et al.*, 2002). An exogenous proline application to *p5cdh* mutants leads to slow seed maturation and cell death symptoms underlining the importance of P5CDH (Deuschle *et al.*, 2004).

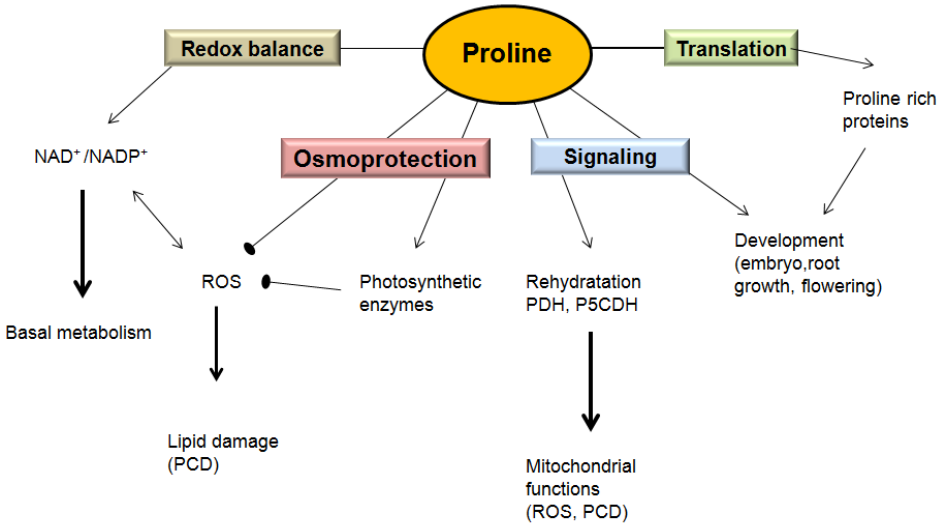

Figure 9. Scheme of multiple proline role in plants. PDH – Proline dehydrogenase, P5CDH - Δ^1-pyrroline-5-carboxylate dehydrogenase, ROS – reactive oxygen species, PCD – programmed cell death (Modified from Szabados and Savouré, 2010).

In plant development, proline and hydroxyproline rich proteins are important part of the cell wall. Inhibited proline synthesis results in the malformation of a leaf parenchyma and vascular tissues accompanied by decreased proline and hydroxyproline content in the cell wall (Nanjo *et al.*, 1999). On the other hand, enhanced synthesis leads to delayed senescence and affects a leaf size and tissue formation (Mattioli *et al.*, 2008). Proline also

takes part in flower development (Kavi Kishor *et al.*, 1995) after observing that tobacco plants overexpressing *P5CS* from moth bean developed more flowers. Overexpression of *AtP5CS1* leads to early flowering phenotype (Mattioli *et al.*, 2008).

P5CDH and substrate preferences

Up to date there are no kinetic and substrate specificity studies devoted to plant P5CDHs as well as determined protein structures. Only human P5CDH (hALDH4) was deeply characterized. The reaction mechanism is sequential and not different from other ALDHs (Inagaki *et al.*, 2006). The enzyme is non-competitively inhibited by the reaction product (Forte-McRobbie and Pietruszko, 1989). P5CDHs were reported as dimeric in human, *Thermus thermophillus*, *Bacillus halodurans* and *Mus musculus* (Forte-McRobbie and Pietruszko, 1989, Inagaki *et al.*, 2006, Luo *et al.*, 2013 and Pemberton and Tanner, 2013). However Luo *et al.* (2013) also reported that TtP5CDH and DrP5CDH (*Deinococcus radiodurans*) form hexamers consisting of three dimers

hALDH4 oxidizes also glutaric semialdehyde, adipic semialdehyde and succinic semialdehyde apart from GSAL (Forte-McRobbie and Pietruszko, 1986). Rat liver P5CDH was reported to oxidize the same substrates with kinetic constants similar to those of hALDH4 (Small and Jones, 1990). Inhibition-kinetic study performed on P5CDH from *Mus musculus* (MmP5CDH) revealed an inhibition by a short dicarboxylate compound malonate and a single carboxyl-compound glyoxalate (Pemberton and Tanner, 2013). Structural basis of the coenzyme preference was elucidated by Inagaki *et al.* (2007). Authors determined again the importance of glutamate in P5CDH from *Thermus thermophilus*. Although the K_m values for NAD^+ and $NADP^+$ were very similar to each other, the turnover rate was 4-fold lower in case of $NADP^+$.

2.2.6. ALDH21

The first ALDH21 family member was found in *T. ruralis* (TrALDH21), a moss studied to understand molecular and biochemical mechanisms of desiccation tolerance (Oliver *et al.*, 2000). The phylogenetic analysis suggests a close relationship of TrALDH21 to ALDH11 family and further to ALDH5, ALDH7 (antiquitin, aminoadipic semialdehyde dehydrogenase) and ALDH10 (aminoaldehyde dehydrogenase) families (Chen *et al.*, 2002). So far, ALDH21 family members were identified in four organisms, namely in *P. patens*, *T. ruralis*, *S. moelendorffii* and *S. caninervis*. The role of ALDH21 has not yet been well understood but there are some information available. The expression of *TrALDH21* (Chen *et al.*, 2002) and *ScALDH21* (Yang *et al.*, 2012) is up-regulated during water deficit stress. Yang *et al.* also reported up-regulation of *ScALDH21* expression by the abscisic acid (ABA), which is used by plants as a mediator of water-stress inducible genes e.g. in *P. patens* or

Funaria hygrometica (Ingram and Bartels, 1996). To understand possible roles and properties of ALDH21, it is necessary to look at two ALDH families closest to ALDH21.

ALDH11 gene family consists of non-phosphorylating glyceraldehyde-3-phosphate (GAP) dehydrogenases (GAPNs, EC 1.2.1.9). These cytosolic ALDHs catalyze irreversible oxidation of GAP to 3-phosphoglycerate in $NADP^+$-dependent manner (Figure 10A) (Valverde *et al.*, 1999). There are also phosphorylating glyceraldehyde-3-phosphate dehydrogenases (GAPDH, EC 1.2.1.12), catalyzing the conversion to 1,3-bisphosphoglycerate. GAPNs can bypass the reaction catalyzed by GADPHs (Rius *et al.*, 2006). The bypass is unique to photosynthetic eukaryotes like plants and microalgae (Plaxton *et al.*, 1996; Valverde *et al.*, 1999). The NADPH formed in the reaction is utilized in mannitol biosynthesis by many plant species via the action of mannose-6-phosphate reductase converting mannose-6-phosphate to mannitol-1-phosphate. Mannitol is after dephosphorylated and represents the major carbohydrate produced by photosynthesis as observed for example in celery leaves (Gao and Loescher, 2000). Interesting is a high number of *ALDH11* genes in moss genomes. *P. patens* (Wood and Duff, 2009) and *S. moelendorffii* (Brocker *et al.*, 2013) carry five and six *ALDH11* genes, respectively.

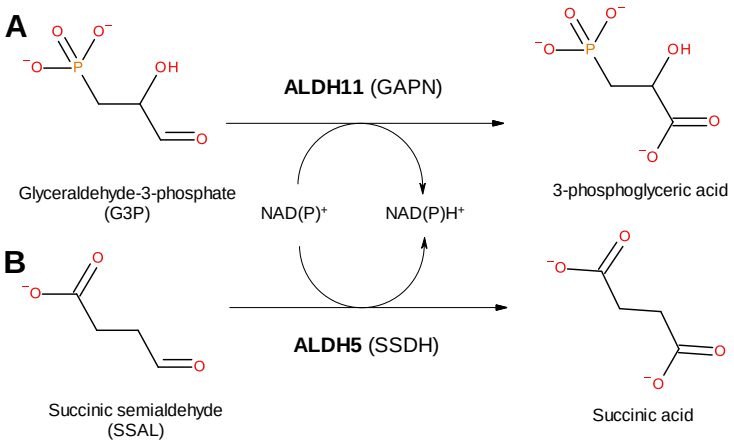

Figure 10. Reactions catalyzed by ALDH11 and ALDH5. (A) ALDH11 – glycerladehyde-3-phosphate is oxidized to 3-phopshogylecric acid. (B) ALDH5 – succinic semialdehyde is oxidized to succinate.

ALDH5 was identified as the second most closely related family to ALDH21 (Wood and Duff, 2009). *ALDH5* orthologues were found in various species, except for *Volvox carteri* genome (Brocker *et al.*, 2013). ALDH5 stands for a succinic semialdehyde dehydrogenase (SSALDH or SSDH, EC 1.2.1.24) catalyzing the conversion of succinic semialdehyde (SSAL) to succinic acid (Figure 10B). SSADH belongs to enzymes participating in GABA shunt (Figure 11). The first enzyme glutamate decarboxylase (GAD, EC 4.1.1.15) catalyzes the α-decarboxylation of glutamate to GABA followed by GABA transaminase (GABA-T, EC 2.6.1.19) that converts GABA to SSAL using either pyruvate or 2-oxoglutarate as amino acceptors. The final step is mediated by SSALDH (Shelp *et al.*,

1999). GABA shunt appears to be interlinked with amino acids metabolism, major and minor carbohydrate metabolism, glycolysis, oxidative pentose phosphate pathway and lipid metabolism (Fait *et al.*, 2008). The knock-out of *SSALDH* gene in *Arabidopsis thaliana* leads to necrotic lesions due to higher accumulation of γ-hydroxybutyric acid emerging from SSAL and reactive oxygen species, respectively (Fait *et al.*, 2005).

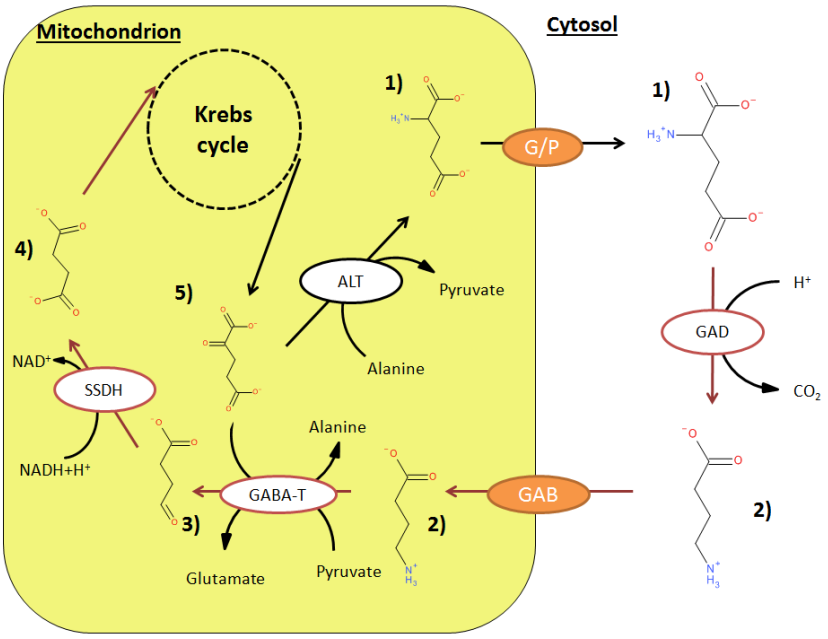

Figure 11. **Scheme of GABA shunt and its connection to Krebs cycle and glutamate synthesis.** GABA shunt is depicted in red arrows and glutamate synthesis from Krebs cycle intermediate in black arrows. Enzymes are marked as ellipses with white inside and transport channels as orange ellipses. 1) glutamate 2) GABA 3) SSAL 4) Succinic acid 5) 2-oxoglutarate. GAD – glutamate decarboxylase, GABA-T – GABA transaminase, SSDH – succinic semialdehyde dehydrogenase, ALT – alanine transaminase, G/P – glutamate/proline antiporter, GAB – putative mitochondrial GABA transporter (Modified from Shelp *et al.*, 1999 and Fait *et al.*, 2008).

The GABA, in mammals mainly posing as neurotransmitter (Varju *et al.*, 2001), plays important role in the plant development and stress toleration. GABA gradient is required for guiding pollen tubes to ovaries (Palanivelu *et al.*, 2003), concentration dependent regulation of pollen tube growth (Ling at al., 2013), cryoprotection (Heber *et al.*, 1971) including alleviation of cold injuries symptoms when added to bananas (Wang *et al.*, 2014), or restriction of reactive oxygen species (Bouché *et al.*, 2003). SSALDHs were found localized in mitochondria of *A. thaliana* (Busch and Fromm, 1999; Bouché *et al.*, 2003). Also either in rice (Gao and Han, 2009) or maize (Jimenez-Lopez *et al.*, 2010) there are ALDHs with predicted mitochondrial localization.

6. Conclusion

- The cDNA of *P. patens* was synthesized and *PpALDH10*, *PpALDH12* and *PpALDH21* were cloned.

- The production of recombinant enzymes was optimized to obtain amounts sufficient for kinetic and structural studies.

- Substrate specificity, pH optima, thermostability and kinetic parameters were determined.

- PpALDH21 was successfully crystallized and X-ray data were collected up to 2.8 Å resolution.

7. References

Agostinelli E., Arancia G., Dalla Vedova L., Belli F., Marra M., Salvi M., Toninello A. (2004) The biological functions of polyamine oxidation products by amine oxidases: Perspectives of clinical applications. *Amino Acids* **27**, 347–358.

Ahmad R., Kim M.D., Back K.H., Kim H.S., Lee H.S., Kwon S.Y., Murata N., Chung W.I., Kwak, S.S. (2008) Stress-induced expression of choline oxidase in potato plant chloroplasts confers enhanced tolerance to oxidative, salt, and drought stresses. *Plant Cell Reports* **27**, 687–698.

Arikit S., Yoshihashi T., Wanchana S., Uyen T.T., Huong N.T.T., Wongpornchai S., Vanavichit A. (2011) Deficiency in the aminoaldehyde dehydrogenase encoded by *GmAMADH2*, the homologue of rice *Os2AP*, enhances 2-acetyl-1-pyrroline biosynthesis in soybeans (*Glycine max* L.). *Plant Biotechnology Journal* **9**, 75-87.

Bradbury L.M.T., Fitzgerald T.L., Henry R.J., Jin Q.S., Waters D.L.E. (2005) The gene for fragrance in rice. *Plant Biotechnology Journal* **3**, 363–370.

Bradbury L.M.T., Gillies S.A., Brushett D.J., Waters D.L.E., Henry R.J. (2008) Inactivation of an aminoaldehyde dehydrogenase is responsible for fragrance in rice. *Plant Molecular Biology* **68**, 439-449.

Briesemeister S., Rahnenführer J., Kohlbacher O. (2010) Going from where to why - interpretable prediction of protein subcellular localization. *Bioinformatics* **26**, 1232-1238.

Brocker C., Vasiliou M., Carpenter S., Carpenter C., Zhang Y., Wang X., Kotchoni S.O., Wood A.J., Kirch H.H., Kopečný D. Nebert D.W. Vasiliou V. (2013) Aldehyde dehydrogenase (ALDH) superfamily in plants: gene nomenclature and comparative genomics. *Planta* **237**, 189-210.

Bouché N., Fait A., Bouchez D., Møller S. G., Fromm H. (2003) Mitochondrial succinicsemialdehyde dehydrogenace of the γ-aminobutyrate shunt is required to restrict levels of reactive oxygen intermediates in plants. *Plant Biology* **100**, 6843-6848.

Buchanan B.B., Gruissem W., Russel J.L. (2000) Biochemistry and molecular biology of plants, pp. 911-914 and 1165-1168, American Society of Plant Physiology, Rockville, Maryland, USA.

Bullheller B.M., Hirst J.D. (2009) DichroCalc – circular and linear dichroism online. *Bioinformatics* **25**, 539-540.

Busch K.B., Fromm H. (1999) Plant succinic semialdehyde dehydrogenase. Cloning, purification, localization in mitochondria, and regulation by adenine nucleotides. *Plant Physiology* **121**, 589-597.

Chen X., Zeng Q., Wood A.J. (2002) The stress-responsive *Tortura ruralis* gene *ALDH21A1* describes a novel eukaryotic aldehyde dehydrogenase protein family. *Journal of Plant Physiology* **159**, 677-684.

Chen Z., Cuin T.A., Zhou M., Twomey A., Naidu B.P., Shabala S. (2007) Compatible solute accumulation and stressmitigating effects in barley genotypes contrasting in their salt tolerance. *Journal of Experimental Botany* **58**, 4245–4255.

Choudhary N.L., Sairam R.K., Tyagi A. (2005) Expression of delta1-pyrroline-5-carboxylate synthetase gene during drought in rice (*Oryza sativa* L.). *Indian Journal of Biochemistry and Biophysics* **42**, 366–370.

Claros M.G., Vincens P. (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. *European Journal of Biochemistry* **241**, 779-786 .

Díaz-Sánchez A.G., Gonzalez-Segura L., Mujica-Jimenez C., Rudino-Pinera E., Montiel C., Martinez-Catilla L.P., Munoz-Clares R.A. (2012) Amino acid residues critical for the specificity for betaine aldehyde of the plant ALDH10 isoenzyme involved in the synthesis of glycine betaine. *Plant Physiol* **158**, 1570-1582.

Deuschle K., Funck D., Hellmann H., Däscher K., Binder S., Frommer W.B. (2001) A nuclear gene encoding mitochondrial Δ^1-pyrroline-5-carboxylate dehydrogenase and its potentional role in protection from proline toxicity. *The Plant Journal* **27**, 345-455.

Deuschle K., Funck D., Forlani G., Stransky H., Biehl A., Leister D., van der Graaf E., Kunze R., Frommer W.B. (2004) The role of Δ^1-pyrroline-5-carboxylate dehydrogenase in proline degradation. *The Plant Cell* **16**, 3413-3425.

Edelhoch H. (1967) Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry* **6**, 1948-1954.

Edgar R.C. (2004), MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792-97.

Emanuelsson O., Brunal S., von Heijne G., Nielsen H. (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nature Protocols* **2**, 953-971.

Fabro G., Kovács I., Pavet V., Szabados L., Alvarez M.E. (2004) Proline accumulation and AtP5CS2 gene activation are induced by plant-pathogen incompatible interactions in Arabidopsis. *Molecular Plant–Microbe Interactions* **17**, 343–350.

Fait A., Yellin A., Fromm H. (2005) GABA shuntdeficiencies and accumulation of reactive oxygen intermediates: insight fro *Arabidopsis* mutants. *FEBS Letters* **579**, 415-420.

Fait A., Fromm H., Walter D., Galili G., Fernie A.R. (2008) Highway or byway: the metabolic role of the GABA shunt in plants. *Cell Press* **13**, 14-19.

Fan W., Zhang M., Zhang H., Zhang P. (2012) Improved tolerance to various abiotic stresses in transgenic sweet potato (*Ipmoea batatas)* expressing spinach betaine aldehyde dehydrogenase. *Plos ONE* **7**, e73744. doi:10.1371/journal.pone.0037344

Faye L., Gomord V. (2010) Succes stories inmolecular farming – brief overview. *Plant Biotechnology Journal* **8**, 525-528.

Fischer R., Stoger E., Schillberg S., Christou P., Twyman R.M. (2004) Plant-based production of biopharmaceuticals. *Current Oppinion in Plant Biology* **7**, 152-158.

Fitzgerald T.L., Waters D.L.E., Brooks L.O., Henry R.J. (2010) Fragrance in rice (*Oryza sativa*) is associated with reduced yield under salt treatment. *Environmental and Experimental Botany* **68**, 292-300.

Floyd S.K., Bowman J.L. (2007) The ancestral developmental tool kit of land plants. *International Journal of Plant Sciences* **168**, 1-35.

Foetisch K., Westphal S., Lauer I., Retzek M., Altmann F., Kolarich D., Scheurer S., Vieths S. (2003) Biological activity of IgE specific for cross-reactive carbohydrate determinants. *Journal of Allergy and Clinical Immunology* **111**, 889-896.

Forte-McRobbie C.M.F., Pietruszko R. (1986) Purification and characterization of human liver „high K_m" aldehyde dehydrogenase and its identification as glutamic γ-semialdehyde dehydrogenase. *Journal of Biological Chemistry* **261**, 2154-2163.

Forte-McRobbie C.M.F., Pietruszko R. (1989) Human glutamic γ-semialdehyde dehydrogenase. *Biochemistry Journal* **261**, 935-943.

Franc V., Řehulka P., Raus M., Stulík J., Novak J., Renfrow M.B., Šebela M. (2013) Elucidating heterogeneity of IgA1 hinge-region O-glycosylation by use of MALDI-TOF/TOF mass spectrometry: role of cysteine alkylation during sample processing. *Journal of Proteomics* **92**, 299–312.

Franc V., Šebela M., Řehulka P., Končitíková R., Lenobel R., Madzak C., Kopečný D. (2012) Analysis of N-glycosylation in maize cytokinin oxidase/dehydrogenase 1 using a manual microgradient chromatographic separation coupled offline to MALDI-TOF/TOF mass spectrometry. *Journal of Proteomics* **75**, 4027–4037.

Frömmel J., Soural M., Tylichová M., Kopečný D., Demo G., Wimmerová M., Šebela M. (2011) Plant aminoaldehyde dehydrogenases oxidize a wide range of nitrogenous heterocyclic aldehydes. *Amino Acids* **43**, 1189-1202.

Frömmel J., Šebela M., Demo G., Lenobel R., Pospíšil T., Soural M., Kopečný D. (2015) *N*-acyl-ω-aminoaldehydes are efficient substrates of plant aminoaldehyde dehydrogenase. *Amino Acids* **47**, 175-187.

Gao Z., Loescher W.H. (2000) NADPH supply and mannitol biosynthesis. Characterization, cloning, and regulation of the nonreversible glyceraldehyde-3-phosphate dehydrogenase in celery leaves. *Plant Physiol*ogy **124**, 321–330.

Gao C., Han B. (2009) Evolutionary and expression study of the aldehyde dehydrogenase (ALDH) gene family in rice (*Oryza sativa*). *Gene* **431**, 86-94.

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. In: The Proteomics Protocols Handbook (Walker J. M. ed), Humana Press, USA, 571-607

Gomord V., Faye L (2004) Posttranslational modification of theraputic proteins in plants. *Current Oppinions in Plant Biology* **7**, 171-181.

Gómez D. F. C., Sesma J. I., Iglesias A. A. (2000) Structural and kinetic characterization characterization of NADP-dependent non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase from celery leaves. *Plant Science* **154**, 107-115.

González-Segura L., Riveros-Rosas H., Julián-Sánchez A., Munoz-Clares R. (2015) Residues that influence coenzyme preference in the aldehyde dehydrogenases. *Chemico-Biological Interactions* **234**, 59-74.

Guindon S., Dufayard J.F., Lefort V., Anisimova M., Hordijk W., Gascuel O. (2010) New algorithms and methods to estimate maximum-likehood phylogenesis: assesing the performance of PhyML 3.0. *Systematic Biology* **59**, 307-21.

Hall T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41,** 95-98.

Hare P.D., Cress W.A., van Staden J. (2002) Disruptive effects of exogenous proline on chloroplast and mitochondrial ultrastructure in Arabidopsis leaves. *South African Journal of Botany* **68,**393–396.

Haslett M., Pink D., Walters B., Brosnan M.E. (2004) Assay and subcellular localization of pyrroline-5-carboxylate dehydrogenase in rat liver. *Biochimica et Biophysica Acta* **1675**, 81-86.

Hasse K., Schuhre K. (1962) Derivative des gamma-aminobutyraldehyds und delta-aminovaleraldehyds. *Biochemische Zeitschrift* **336**, 20-34.

He D., Lei Z., Xing H., Tang B. (2014) Genome-wide identification and analysis of the aldehyde dehydrogenase (ALDH) gene superfamily of *Gossypium raimondii*. *Gene* **549**, 123-133.

Heber U., Tyankova L. and Santarius K.K. (1971) Stabilization and inactivation of biological membranes during freezing in the presence of amino acids. *Biochimica Biophysica Acta* **241**, 578–582.

Hibino T., Meng Y.L., Kawamitsu Y. Uehara N., Matsuda N., Tanaka Y., Ishikawa H., Baba S., Takaba T., Wada K., Ishii T. (2001) Molecular cloning and functional characterization of two kinds of betaine-aldehyde dehydrogenase in betaine-acumulating mangrove *Avicennia marina*. *Plant Molecular Biology* **45**, 353-363.

Holmstrom K.O., Somersalo S., Mandal A., Palva T.E., Welin B. (2000) Improved tolerance to salinity and low temperature intransgenic tobacco producing glycine betaine. *Journal of Experimental Botany* **51**, 177–185.

Huether C.M., Lienhart O., Baura S., Stemmer C., Gorr G., Reski R., Decker E.L. (2005) Glyco-engineering of moss lacking plant-specific sugar residues. *Plant Biology* **7**, 292-299.

Inagaki E., Ohshima N., Takahashi H., Kuroishi C., Yokoyama S., Tahirov T.H. (2006) Crystal structure of *Thermus thermophilus* Δ^1-pyrroline-5-carboxylate dehydrogenase. *Journal of Molecular Biology* **362**, 490-501.

Inagaki E., Ohshima N., Sakamoto K., Babayeva N.D., Kato H., Yokoyama S., Tahirov T.H. (2007) New insights into the binding mode of coenzymes: structure of *Thermus thermophilus* Δ^1-pyrroline-5-carboxylate dehydrogenase cemplexed with $NADP^+$. *Acta Crystallographica Section F Structural Biology and Crystallization Communications* **63**, 462-465.

Incharoensakdi A., Matsuda N., Hibino T., Meng Y.L., Ishikawa H., Hara A., Funaguma T., Takabe T. (2000) Overproduction of spinach betaine aldehyde dehydrogenase in *Escherichia coli*. Structural and functional properties of wild-type, mutants and E. coli enzymes. *Eur J Biochem/FEBS* **267**, 7015–7023.

Ingram J., Bartels D. (1996) The molecular basis of dehydratation tolerance in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**, 377-403

Jimenez-Lopez J.C., Gachomo E.W., Seufferheld M.J., Kotochoni S.O. (2010) The maize ALDH protein superfamily: linking structural features to functional specificities. *BMC Struct Biol* **10**, 43.

Kammerer W., Cove D.J. (1996). Genetic analysis of the effects of re-transformation of transgenic lines of the moss *Physcomitrella patens*. *Molecular and General Genetics* **250**, 380-382.

Kavi Kishor P.B., Hong Z., Miao G.H., Hu C.A.A., Verma D.P.S. (1995) Overexpression of D1-pyrroline-5-carboxylate synthetase increases proline production and confers osmotolerance in transgenic plants. *Plant Physiology* **108**, 1387–1394.

Kirch H.H., Bartels D., Wei Y., Schnable P.S., Wood A.J., (2004) The ALDH gene superfamily of arabidopsis. *Trends in Plant Science* **9**, 371-377.

Kotochoni S.O., Jimenez-Lopez J.C., Kayodé A.P.P., Gachomo E.W., Baba-Moussa L. (2012) The soybean aldehyde dehydrogenase (ALDH) protein subfamily. *Gene* **495**, 128-133.

Kopečný D., Tylichová M., Snegaroff J., Popelková H., Šebela M. (2011) Carboxylate and aromatic active-site residues are determinants of high affinity bonding of ω-aminoaldehydes to plant aminoaldehyde dehydrogenases. *Febs J* **278**, 3130-3139.

Kopečný D., Končitíková R., Tylichová M., Vigouroux A., Moskalíková H., Soural M.,Šebela M., Moréra S. (2013) Identyfying critical residues for substrace specifity and trapping a thiohemicatal intermediate. *J Biol Chem* **288**, 9491-9507.

Koprivova A., Stemmer C., Altmann F., Hoffmann A., Kopriva S., Gorr G., Reski R., Decker E.L. (2004) Targeted knockouts of *Physcomittrella* lacking plant-specific immunogenic *N*-glycans. *Plant Biotechnology Journal* **2**, 517-523.

Ling Y., Chen T., Jing Y., Fan L., Wan Y., Lin J. (2013) γ-aminobutyric acid (GABA) homeostasis regulates pollen germination and polarized growth in *Pice wilsonii*. *Planta* **238**, 831-843.

Louis-Jeune C., Andrade-Navarro M.A., Perez-Iratxeta C. (2012) Predictionof protein secondary structure using theoretically derived specrta. *Proteins* **80**, 374-381.

Luo M., Singh R.K., Tanner J.J. (2013) Structural determinants of oligomerization of Δ^1-pyrroline-5-carboxylate dehydrogenase: identification of a hexamerization hot spot. *Journal of Molecular Biology* **425**, 3106-3120.

Mattioli R., Marchese D., D'Angeli S., Altamura M.M., Costantino P., Trovato M. (2008) Modulation of intracellular proline levels affects flowering time and inflorescence architecture in Arabidopsis. *Plant Molecular Biology* **66**, 277–288.

Michels S., Scagliarini S., Della Seta F., Carles C., Riva M., Trost P. (1994) Arguments against a close relationship between non-phosphorylating and phos- phorylating glyceraldehyde-3-phosphate dehydrogenases. *FEBS Letters* **339**, 97–100.

Moravcová D., Kahle V., Řehulková H., Chmelík J., Řehulka P. (2009) Short monolithic columns for purification and fractionation of peptide samples for matrix-assisted laser desorption/ionization time-of-flight/time-of-flight mass spectrometry analysis in proteomics. *Journal of Chromatography A* **1216**, 3629–3636.

Munoz-Clares R.A., Riveros-Rosas H., Garza-Ramos G., González-Segura L., Mújica-Jiménez C., Julián-Sánchez A. (2014) Exploring the evolutionary route of the acquisition of betaine aldehyde dehydrogenase activity by plant ALDH10 enzymes: implications for the synthesis of osmoprotectant glycine betaine. *BMC Plant Biology* **14**, 1471-2229.

Nanjo T., Kobayashi M., Yoshiba Y., Sanada Y., Wada K., Tsukaya H., Kakubari Y., Yamaguchi-Shinozaki K., Shinosaki K. (1999) Biological functions of proline in morphogenesis and osmotolerance revealed in antisense transgenic Arabidopsis thaliana. *Plant Journal* **18**, 185–193.

Oliver M.J., Velten J., Wood A.J. (2000) Bryophytes as experimental models for the study of environmental stress tolerance: desiccation-tolerance in mosses. *Plant Ecology* **151**, 73–84.

Pace C.N., Vajdos F., Fee L., Grimsley G., Gray T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci*. **11**, 2411-2423.

Palanivelu R., Brass L., Edlund A.F., Preuss D. (2003) Pollen tube growth and guidance is regulated by POP2, an Arabidopsis gene that controls GABA levels. *Cell* **114**, 47-59.

Parsons J., Altmann F., Arrenberg C.K., Koprivova A., Beike A.K., Stemmer C., Gorr G., Reski R., Decker E.L. (2012) Moss-based production of asialo-erythropoetin devoid ofLewis A and other plant-typical carbohydrate determinants. *Plant Biotechnology Journal* **10**, 851-861.

Pemberton T.A. and Tanner J.J. (2013) Structural basis of Δ^1-pyrroline-5-carboxylate dehydrogenase (ALDH4A1): Semialdehyde chain length. *Archives of Biochemistry and Biophysics* **538**, 34-40.

Perozich J., Kuo I., Wang B.C., Boesch J.S., Lindahl R. , Hempel J. (2000) Shifting the NAD/NADP preference in class 3 aldehyde dehydrogenase. *European Jornal of Biochemistry* **267**, 6197–6203.

Quan R., Shang M., Zhang J. (2004) Improved chilling tolerance by transformation with *betA* gene for the enhancement of glycinebetaine synthesis in maize. *Plant Sci*ence **166**, 141–149.

Raskin I., Ribnicky D.M., Komarnytsky S., Ilic N., Poulev A., Borisjuk N., Brinker A., Moreno D.A., Ripoll C., Yakoby N., O'Neal J.M., Cornwell T., Pastor I., Fridlender B. (2002) Plants and human health in twenty-first century. *Trends in Biotechnology* **20**, 522-531.

Rathinasabapathi B, Burnet M, Russell BL, Gage DA, Liao PC, Nye GJ, Scott P, Golbeck JH, Hanson AD (1997) Choline monooxygenase, an unusual iron–sulfur enzyme catalyzing the first step of glycine betaine synthesis in plants:prosthetic group characterization and cDNA cloning, *Plant Biology* **94**, 3454–3458.

Rensing S.A., Lang D., Zimmer A.D., Terry A., Salamov A., Shapiro H., Nishiyama T. *et al.* (2008) The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* **319**, 64-69.

Rius S.P., Casati P., Iglesias A.A., Gomez-Casati D.F.(2006) Characterization of an *Arabidopsis thaliana* mutant lacking a cytosolic non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase. *Plant Molecular Biology* **61**, 945–957.

Rhodes D., Hanson A.D. (1993) Quarternary ammonium and tertiary sulfonium compounds in higher plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **44**, 357-384.

Rossmann M.G., Moras D., Olsen K.W. (1974) Chemical and biological evolution of a nucleotide-binding protein. *Nature* **250,** 194–199.

Sakamoto A., Alia, Murata N. (1998) Metabolic engineering of rice leading to biosynthesis of glycinebetaine and tolerance to salt and cold. *Plant Molecular Biol*ogy **38**, 1011–1019.

Saradhi P.P., Alie, Arora S., Prasad K.V. (1995) Proline accumulates in plants exposed to UV radiation and protects them against UV induced peroxidation. *Biochemical and Biophysical Research Communications* **209**, 1–5.

Schat H., Sharma S.S., Vooijs R. (1997) Heavy metal-induced accumulation of free proline in a metal-tolerant and a nontolerant ecotype of Silene vulgaris. *Physiologia Plantarum* **101**, 477–482.

Shelp B.J., Bown A.W., McLean M.D. (1999) Metabolism and functions of gamma-aminobutyric acid. *Trends in Plant Science* **4**, 446-452.

Shen Y.G., Du B.X., Zhang W.K., Zhang J.S., Chen S.Y. (2002) AhCMO, regulated by stresses in *Atriplex hortensis*, can improve drought tolerance in transgenic tobacco. *Theoretical and Applied Genet*ics **105**, 815–821.

Small W.C., Jones M.E. (1990) Pyrroline 5-carboxylate dehydrogenase of the mitochondrial matrix of rat liver. *Journal of Biological Chemistry* **260**, 18668-18672.

Strobek C., Krinninger S., Frank W. (2013) The moss *Physcomitrella patens*: methods and tools from cultivation to targeted analysis of gene function. *International Journal of Developmental Biology* **57**, 533-564.

Su J., Hirji R., Zhang L., He C., Selvaraj G., Wu R. (2006) Evaluation of the stress-inducible production of choline oxidase in transgenic rice as a strategy for producing the stress-protectant glycine betaine. *Journal of Experimental Bot*any **57**, 1129–1135.

Suckau D., Resemann A., Schuerenberg M., Hufnagel P., Franzen J., Holle A. (2003) A novel MALDI LIFT-TOF/TOF mass spectrometer for proteomics. *Annals of Bioanalytical Chemistry* **376**, 952–965.

Székely G., Abrahám E., Cséplo A., Rigó G., Zsigmond L., Csiszár J., Ayaydin F., Strizhov N., Jásik J., Schmelzer E., Koncz C., Szabados L. (2008) Duplicated P5CS genes of Arabidopsis play distinct roles in stress regulation and developmental control of proline biosynthesis. *Plant Journal* **53**, 11–28.

Šebela M., Brauner F., Radova A., Jacobsen S., Havlis J., Galuszka P., Pec P. (2000) Characterisation of a homogeneous plant aminoaldehyde dehydrogenase. *Biochimica et Biophysica Acta* **1480**, 329–341

Talavera G. and Castresana J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology* **56**, 564-577.

Tamura K, Stecher G, Peterson D, Filipski A, and Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* **30**, 2725-2729.

Trossat C., Rathinasabapathi B., Hanson A.D. (1997) Transgenically expressed betaine aldehyde dehydrogenase efficiently catalyzes oxidation of dimethylsulfoniopropionaldehyde and _-aminoaldehydes. *Plant Physiology* **113**, 1457–1461.

Tylichová M, Kopečný D, Moréra S, Briozzo P, Lenobel R, Snégaroff J,Sebela M (2010) Structural and functional characterization of plantaminoaldehyde dehydrogenase from Pisum sativum with a broad specificity for natural and synthetic aminoaldehydes. *Journal of Molecular Biology* **396**, 870–882.

Tang W., Sun J., Liu F., Yan J., Gou X., Lu B., Liu Y. (2014) RNAi-directed downregulation of *betaine aldehyde dehydrogenase 1* (*OsBADH1*) results in decreased stress tolerance and increased oxidative markers without affecting glycine betaine biosynthesis in rice (*Oryza staiva)*. *Plant molecular biotechnology* **86**, 443-454.

Valenzuela-Soto E.M, Munoz-Clares R.A. (1993) Betaine-aldehyde dehydrogenase from leaves of *Amaranthus hypocondriacus L.* exhibits an iso ordered bi bi steady-state mechanism. *J Biol Chem* **268**, 23818-23823.

Valverde F., Losada M., Serrano A. (1999) Engineering a central metabolic pathway: glycolysis with no net phosphorylation in an Escherichia coli gap mutant complemented with a plant GapN gene. *FEBS Lett*ers **449**, 153–158.

Varju P., Katarova Z., Madarasz E., Szabo G. (2001) GABA signalling during development:new data and old questions. *Cell Tissue Research* **305**, 239–246.

Vasiliou V., Bairoch A., Tipton K.F., Nebert D.W. (1999) Eucaryotic aldehyde dehydrogenase (ALDH) genes:human polymorphism, and recomended nomenclature based on divergent evolution and chromosomal mapping. *Pharmacogenetics* **9**, 421-434.

Vaz F.M., Fouchier S.W., Ofman R., Sommer M., Wanders R.J. (2000) Molecular and biochemical characterization of rat_-trimethylaminobutyraldehyde dehydrogenase and evidence for the involvement of human aldehyde dehydrogenase 9 in carnitine biosynthesis. *Journal of Biological Chemistry* **275**, 7390–7394.

Velasco-Garcia R., Gonzalez-Segura L., Munoz-Clares R.A. (2000) Steady-state kinetic mechanism of the $NADP^+$- and NAD^+-dependent reactions catalysed by betaine aldehyde dehydrogenase from Pseudomonas aeruginosa. *Biochemistry Journal* **352**, 675–683.

Wang Y., Luo Z., Huang X., Yang K., Gao S., Du R. (2014) Effects of exogenous γ-aminobutyric acid (GABA) treatment on chilling injury and antioxidant capacity in banana peel. *Scientia Horticulturae* **168**, 132-137.

Wei A., He C., Li B., Li N. Zhang J. (2011) The pyramid of transgenes *TsVP* and *BetA* effectively enhances the drought tolerance of maize plants. *Plant Biotechnology Journal* **9**, 216-299.

Widodo, Patterson J.H., Newbigin E., Tester M., Bacic A., Roessner U. (2009) Metabolic responses to salt stress of barley (Hordeum vulgare L.) cultivars, Sahara and Clipper, which differ in salinity tolerance. Journal of Experimental Botany **60**, 4089–4103.

Wood A.J., Duff R.J. (2009) The aldehyde dehydrogenase (ALDH) gene superfamily ofthe moss *Physcomitrella patens* and the algae *Chlamydomonas reindhardtii* and *Ostreococcus tauri*. *The Bryologist* **112**, 1-11.

Wood P.L., Moskal J. R. Khan A. M. (2007) The concept of „aldehyde load" in neurodegenerative mechanisms: Cytotoxicity of the polyamine degradation products hydrogen peroxide, acrolein, 3-aminopropanal, 3-acetamidopropanal and 4-amidobutanal in a retinal ganglion cell line. *Brain Research* **1145**, 150-156.

Yang H., Zhang D., Wang J. (2012) Molecular cloning of a stress-responsive aldehyde dehydrogenase gene *ScALDH21* from the dessication-tolerant moss *Syntrichia caninervis* and its responses to different stresses. *Molecular Biology Reports* **39**, 2645-2652.

Yang S.L., Lan S.S., Gong M. (2009) Hydrogen peroxide-induced proline and metabolic pathway of its accumulation in maize seedlings. *Journal of Plant Physiology* **166**, 1694–1699.

Yoshiba Y., Kiyosue T., Katagari T., Ueda H., Mizoguchi T., Yamaguchi-Shinozaki K., Wada K., Harada Y., Shinozaki K. (1995) Correlation between the induction of a gene for Δ^1-pyrroline-5-carboxylate synthetase and the accumulation of proline in *Arabidopsis thaliana* under osmotic stress. *The Plant Journal* **7**, 751-760.

Zhang Y., Mao L., Wang H., Brocker C., Yin X., Vasiliou V., Fei Z., Wang X. (2012) Genome-wide identification and analysis of grape aldehyde dehydrogenase (ALDH) gene superfamily. *PLoS ONE* **7**, e32153. doi:10.1371/journal.pone.0032153

Zhu C., Ming C., Zhao-shi X., Lian-cheng L., Xue-ping C., You-zhi M. (2014) Characteristics and Expression Patterns of the Aldehyde Dehydrogenase (ALDH) Gene Superfamily of Foxtail Millet (*Setaria italica L.*). *PLoS ONE* **9**, e101136. doi:10.1371/journal.pone.0101136

8. Abbreviations

2-AP	2-acetyl-1-pyrroline
AA	Aminoacid
AASAL	α-aminoadipic semialdehyde
AcAPAL	*N*-acetylaminopropionaldehyde
ABAL	4-aminobutyraldehyde
ALDH	Aldehyde dehydrogenase
AMADH	Aminoaldehyde dehydrogenase
APAL	3-aminopropionaldehyde
BAL	Betaine aldehyde
Benz-Ald	Benzaldehyde
bp	Basepair
C3	Propionaldehyde
C4	Butyraldehyde
C5	Valeraldehyde
C6	Hexanal
C7	Heptanal
C8	Octanal
C9	Nonanal
CHCA	α-cyanohydroxycinnamic acid
Da, kDa	Dalton, kilodalton
DMSP	*S*,*S*-dimethylsulfopropionate
DMSPAL	*S*,*S*-dimethylsulfopropionaldehyde
G3P	Glyceraldehyde-3-phosphate
GABA	γ-amonibutyric acid
GAPN	Non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase
GB	Glycine betaine
GBAL	Guanidinobutyraldehyde
GOI	Gene of interest
GSAL	Glutamate-γ-semialdehyde
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
LB	Lauren-Broth medium
LPO	Lipid peroxidation
MASAL	Adipic semialdehyde methyl ester
MDA	Malondialdehyde
MME	Monomethylether

MYA	Millions years ago
P5C	Δ^1-pyrroline-5-carboxylate
P5CDH	Δ^1-pyrroline-5-carboxylate dehydrogenase
PEG	Polyethyleneglycol
PAO	Polyamine oxidase
SSAL	Succinic semialdehyde
SSDH	Succinic semialdehyde dehydrogenase
TFA	Trifluoric acid
TMABAL	*N*,*N*,*N*-trimethylaminobutyraldehyde