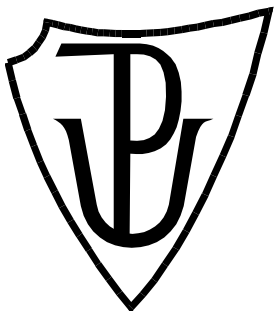
PALACKÝ UNIVERSITY OLOMOUC

Faculty of Science

Department of Biochemistry



Cloning and expression of ALDH selected from families 2 and 3 from moss

MASTER THESIS

Author: **Bc. Adéla Hýlová**

Programme: N1406 Biochemistry

Study branch: Biochemistry

Form of study: Full-time

Supervisor: **Mgr. Martina Kopečná, Ph.D.**

Academic year: 2015

I hereby declare that I wrote this thesis solely by myself and that all the sources cited in this work are listed in the reference part. I agree with publishing this thesis according to the Act no. 111/1998 Coll., as subsequently amended. I am aware of the fact that my work is subject to the rights and obligations arising from the Act no. 121/2000 Coll., Copyright Act, as subsequently amended.

In Olomouc, 20th July 2015

Acknowledgement

I would like to thank my supervisors Mgr. Martina Kopečná, Ph.D. and Mgr. David Kopečný, Ph.D. for their professional guidance, patience and help with processing my thesis. Another thank you belongs to Mgr. Radka Končitíková for kind help and a lot of useful advice. Thank you also goes to Mgr. Filip Dyčka, Ph.D. for help and assistance during MALDI experiments. At least but not last, thank you Dr. Klaus von Schwartzenberg and Dr. Ann-Cathrin Lindner from University Hamburg for guidance through microscopic part of my thesis and teaching me new methods.

Great thanks belong to my parents and friends for moral support in difficult times.

Bibliografická identifikace

Jméno a příjmení autora	Bc. Adéla Hýlová
Název práce	Klonování a exprese vybraných ALDH z rodin 2 a 3 z mechu
Typ práce	Diplomová
Pracoviště	CRH
Vedoucí práce	Mgr. Martina Kopečná, Ph.D.
Rok obhajoby práce	2015
Abstrakt	Geny kódující tři vybrané aldehyddehydrogenasy (ALDH) z mechu *Physcomitrella patens* byly klonovány a exprimovány. Rekombinantní proteiny PpALDH2A, PpALDH2B a PpALDH3D byly produkovány v *E. coli.* Produkce těchto proteinů byla optimalizována modifikací signální sekvence. Rekombinantní proteiny byly purifikovány afinitní chromatografií. Byly studovány vlastnosti hlavně pro PpALDH2A včetně subcelulární lokalizace, molekulové hmotnosti, pH optima a substrátové specificity vůči různým alifatickým a aromatickým aldehydům. Byly změřeny saturační křivky s možnými fyziologickými substráty a určeny hodnoty K_m and V_{lim}.
Klíčová slova	aldehyddehydrogenasa 2, aldehyddehydrogenasa 3, *Physcomitrella patens*, metabolismus aldehydů, lokalizace, enzymová aktivita
Počet stran	74
Počet příloh	0
Jazyk	Anglický

Bibliographical identification

Author's first name and surname	Bc. Adéla Hýlová
Title	Cloning and expression of ALDH selected from families 2 and 3 from moss
Type of thesis	Master
Department	CRH
Supervisor	Mgr. Martina Kopečná, Ph.D.
The year of presentation	2015
Abstract	Three genes coding for aldehyde dehydrogenases (ALDH) from moss *Physcomitrella patens* were cloned and expressed. Recombinant proteins PpALDH2A, PpALDH2B and PpALDH3D were produced in *E. coli*. Their production was optimized by modifying a putative signal sequence. Recombinant ALDHs were purified by affinity chromatography. Properties of especially ΔPpALDH2A were studied. These included localization, molecular weight, pH optimum and substrate specificity towards various aliphatic and aromatic aldehydes. Saturation curves for several physiological aldehydes were measured and kinetic properties K_m and V_{lim} were determined.
Keywords	aldehyde dehydrogenase 2, aldehyde dehydrogenase 3, *Physcomitrella patens*, aldehyde metabolism, localization, enzyme activity
Number of pages	74
Number of appendices	0
Language	English

Contents

1 The aims of the thesis

Theoretical part

- Description of plant ALDH2 and ALDH3 families
- Description of reaction mechanism of aldehyde dehydrogenases (ALDH)
- Description of model organism *Physcomitrella patens*

Experimental part

- Cloning of several *PpALDH* genes
- Optimizing production of PpALDHs in *Escherichia coli*
- Analysis of substrate specificity of purified PpALDHs
- Subcellular localization of fusion PpALDHs proteins in moss protoplasts

2 Introduction

Aldehydes are generated from various endogenous sources, including for example metabolism of biogenic amines, amino acids, sugars or lipids also known as lipid peroxidation, resulting in 4-hydroxy-2-nonenal and malondialdehyde (Esterbauer *et al.,* 1991). Other possibility represent exogenous precursors, such as aldehydes formed in xenobiotic metabolism. Aldehydes are highly reactive molecules due to the electrophilic nature of their carbonyl group. Although some of them are essential in physiological processes, such as vision or neurotransmission, many have also cytotoxic, carcinogenic, mutagenic or genotoxic properties (Nadkarni and Sayre, 1995; Brooks and Theruvathu, 2005) and therefore their levels must be tightly regulated. One of the aldehyde metabolic pathways is oxidation by aldehyde dehydrogenases (ALDHs, EC 1.2.1.3), which are $NAD(P)^+$-dependent enzymes that irreversibly oxidize various types of aliphatic and aromatic aldehydes to carboxylic acids (reviewed in Lindahl, 1992). ALDHs play an important role in detoxification processes.

19 human ALDHs belonging to 11 distinct families have been described so far (Marchitti *et al.*, 2008). 13 families form ALDH superfamily in plants (Brocker *et al.*, 2013). There is significant variability of the total number of *ALDH* genes present in different plant species. A water-to-land transition of plants is considered to be connected with an increase of *ALDH* genes number (Brocker *et al.*, 2013). This work focuses on plant ALDHs, particularly ALDH2 and ALDH3 from the moss *Physcomitrella patens*, which is regarded as one of the first land colonizers and is phylogenetically situated between algae and seed plants, making it an important model organism in terms of evolutionary comparisons between species (Rensing *et al.,* 2008).

The current state of the research

2.1 *Physcomitrella patens*

Bryophyte *Physcomitrella patens* (Hedw.) B.S.G. is monoecious and self-fertile plant. It belongs to class *Bryopsida*, order *Funariales* and family *Funariaceae*. *P. patens* grows close to water areas and is widely distributed around the world. It is an ephermal plant that starts to develop from spores in early summer. The sporophytes are formed in the late summer.

2.1.1 Life cycle

P. patens has simple and short life cycle. It occurs mostly in haploid phase during the cycle. The plant develops from haploid spores (Fig. 1.1), which germinate under favourable environment conditions and generate filamentous protonema (Fig. 1.2, 1.3). There are two types of protonema. The first type is the chloronema (Fig. 1.4), which contains high number of chloroplasts and cell plates that form between the two cells are transverse to the axis of the cell. The second cell type is caulonema (Fig. 1.5), containing fewer less-developed plastids and their cell plates are at an angle to the longer axis of the cell. Tissue differentiation of apical cells from chloronema to caulonema is auxin-dependent (Ashton et al., 1979). The buds (Fig. 1.6) are formed from protonemal tissue and leafy shoots called gametophores (Fig. 1.7) develop later, which is regulated by cytokinins (Ashton et al., 1979). Gametophores make colonies (Fig. 1.8) that produce reproductive organs – female gametes (archegonia, Fig. 1.9) and male gametes (antheridia, Fig. 1.10). The antheridia have flagella, therefore are motile in water and swim towards the archegonia, which leads to fertilization and formation of the zygote. Diploid sporophyte (Fig. 1.11) develops directly on the gametophores and produces several thousand spores by meiosis, which is the final step of the life cycle of the moss. The whole cycle can be reached in the laboratory conditions approximately within 12 weeks (reviewed in Cove, 2005).

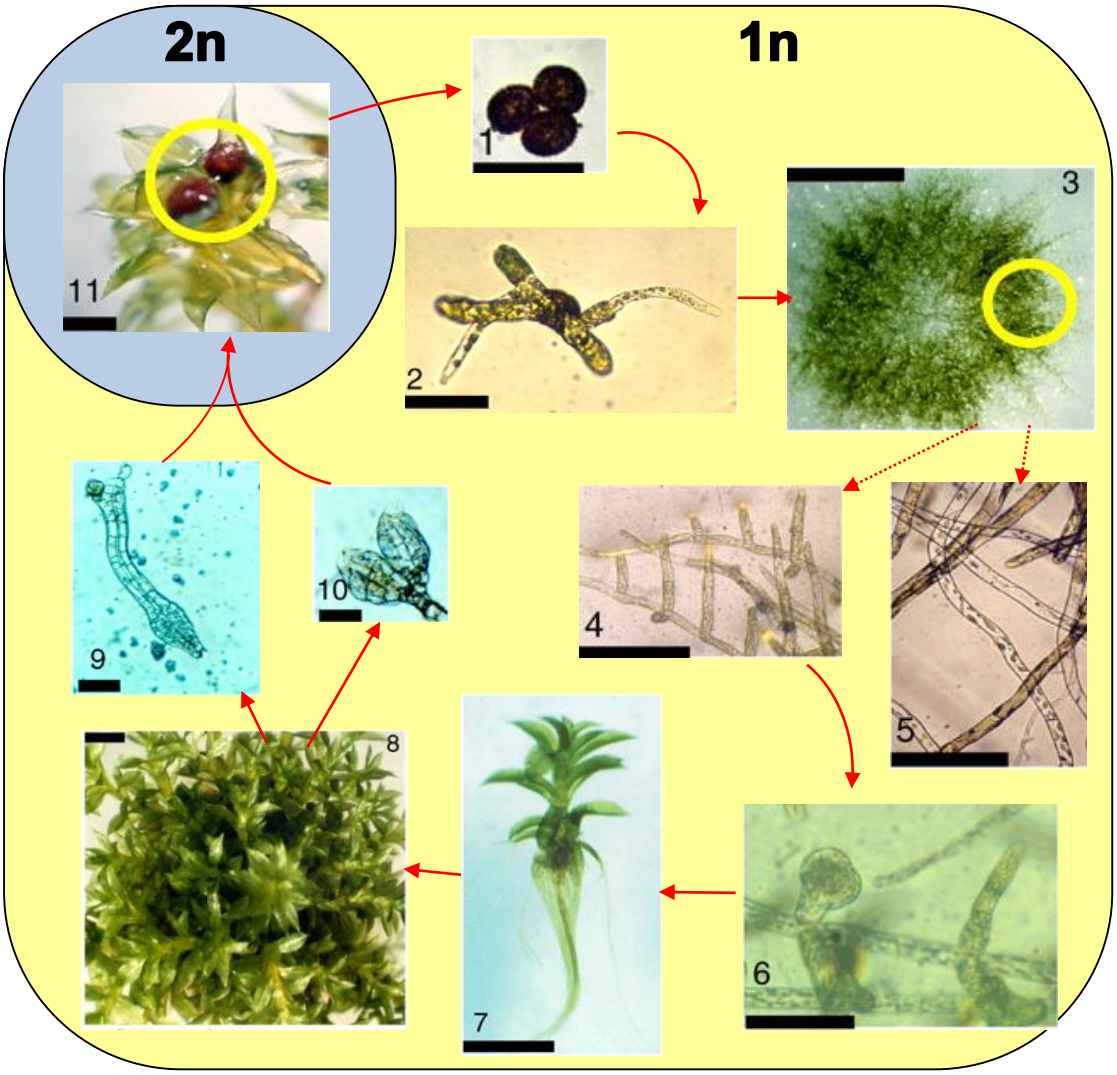


Fig. 1. The life cycle of *Physcomitrella patens*. (1) Haploid (1n) spores (50 µm), (2) Primary filamentous chloronema (50 µm), (3) Protonema (1 cm), that can divide into (4) chloronema or (5) caulonema (0.5 mm), (6) Young bud (0.5 mm), (7) Single gametophore with rhizoids (0.5 mm), (8) Gametophore colony (1 mm) produces (9) archegonia and (10) antheridia (50 µm), (11) Diploid (2n) sporophyte (modified according to Schaefer and Zrÿd, 2001).

2.1.2 Evolution

Terrestrialization dates back to approximately 460 million years ago (MYA) (Kenrick and Crane, 1997). The land plants (embryophytes) diversified from charophycean algae (Qiu, 2008; Finet *et al.,* 2010). The major bryophyte lineages – liverworts, mosses and hornworts – separated and are considered as the first terrestrial plants. The haploid gametophyte remained dominant till 440 MYA, when diploid sporophytic generation predominated in the life cycle

and gave rise to the vascular plants (tracheophytes). The first seed plants (gymnosperm) evolved approximately 400 MYA, comprising orders *Cycads*, *Pinus* and *Gnetum*. Angiosperms, the flowering plants, occurred 300 MYA following the separation of monocots and eudicots later on (Fig. 2) (Strotbek *et al.,* 2013).

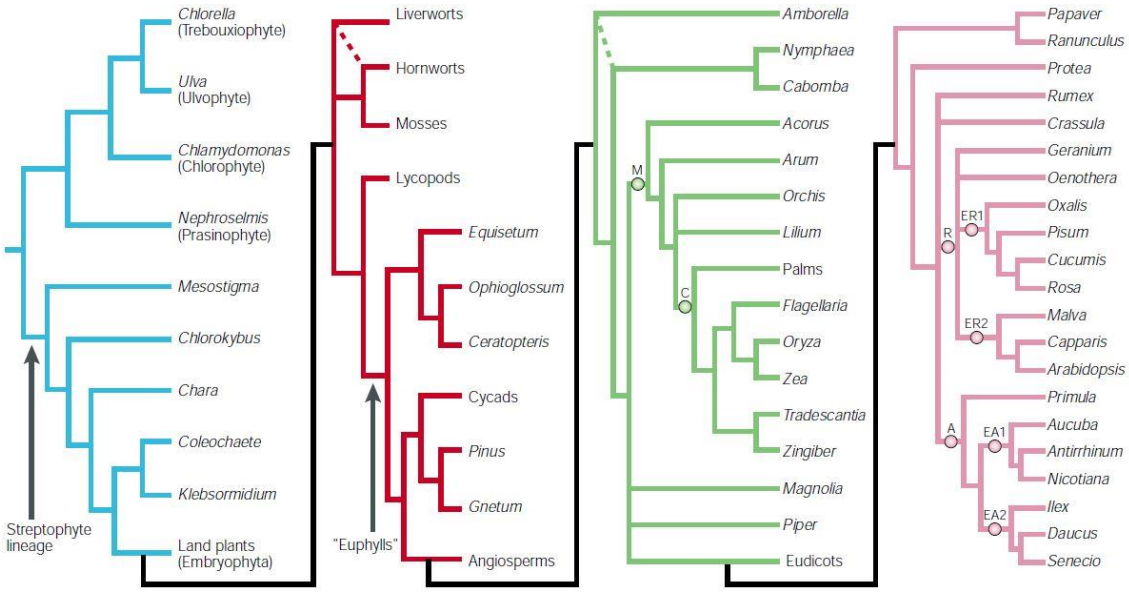


Fig. 2. The phylogeny tree of green plants. *Chara*, belonging to algae (blue), are the last common ancestors of land plants (red). Liverworts and hornworts were proved to be basal embryophyta. *Equisetum* and gymnosperms *Cycads*, *Pinus* and *Gnetum* are sister to seed plants. Angiosperms (green and pink) contain the basal *Amborella*, *Nymphaea* and *Cabomba*. Monocots (M) and eudicots include various genera (Cronk, 2001).

P. patens position in the phylogeny tree makes it an ideal model organism for studying specific adaptations connected with water-to-land transition, such as various modifications induced by abiotic stresses – water insufficiency or temperature fluctuations.

2.1.3 *Physcomitrella* as a model organism

The moss *Physcomitrella patens* is becoming a widely known model plant in the last two decades due to its specific attributes and recently sequenced genome (Rensing *et al.*, 2008). It is a great model system for the studies of gene function and physiology. *P. patens* is an attractive plant for genetic studies, because of its simple life cycle and due to the fact that it occurs

dominantly in the haploid phase of the gametophyte stage. Gene targeting is a method used to study gene function, which provides the generation of loss-of-function mutants or point mutations in the gene of study. Multicellular eukaryotes are not accessible systems for targeted mutagenesis, because foreign DNA integrates at random locations in the genome. *P. patens* is the only plant capable of integration of foreign DNA at targeted locations with high frequencies, as it is able to make homologous recombination (Schaefer and Zrÿd, 1997). It is also first moss that was successfully transformed (Schaefer *et al.,* 1991). Another potential utility of *P. patens* is studying development and metabolism, because of the ease of cultivation in liquid or solid mineral medium without hormones under defined and controlled conditions.

2.1.4 Utilization of *Physcomitrella* in biotechnology

Molecular farming is becoming popular recently and *P. patens* has been developed into good expression system for recombinant proteins. Humanized recombinant proteins can be produced by knockout strains deficit in xylosylation and fucosylation activity (Kopřivová *et al.,* 2004), but with galactosylation that was acquired by adding human galactosyltransferase to the xylosyl and fucosyltransferase locus (Huether *et al.,* 2005). Antibodies IgG1, IgG4, human vascular endothelial growth factor (VEGF) or erythropoietin (Decker and Reski, 2008) can be produced and secreted into culture medium (Schaaf *et al.,* 2005). Biopharmaceuticals production by knockout/knockin *P. patens* strain is safe due to humanized pattern of posttranslational modifications thus non-immunogenic response in human body. Another biotechnological usage of the moss was suggested by Anterola *et al.* (2009). *P. patens* was stably transformed with taxadiene synthase gene isolated from *Taxus brevifolia* to produce taxa-4(5),11(12)-diene, the precursor of paclitaxel, which is used as anticancer drug. The mechanism of action lies in inhibition of cell proliferation by microtubules stabilization (Schiff *et al.,* 1979).

2.2 Aldehyde dehydrogenases (ALDHs)

ALDHs oxidize a broad range of aldehydes to corresponding carboxylic acids in NAD(P)$^+$-dependent manner. NAD$^+$ is cofactor in plant and mammal ALDHs, whereas NADP$^+$ is typical for bacteria. General mechanism of the reaction catalysed by ALDHs involves cofactor binding, which results in a conformational change of catalytic cysteine prior to substrate binding. Catalytic cysteine acts as a nucleophile and attacks the carbonyl group of the substrate. Oxyanion thiohemiacetal intermediate is formed and stabilized by two NH-groups of catalytic Asn residue. Hydrogen atom is transferred to the cofactor and thioacylenzyme intermediate is formed. Active-site Glu residue acts as a general base, therefore catalyses hydrolysis of the thioacylenzyme intermediate, which leads to release of carboxylic acid product. The cofactor probably dissociates from the enzyme afterwards (Fig. 3) (Marchitti *et al.,* 2008).

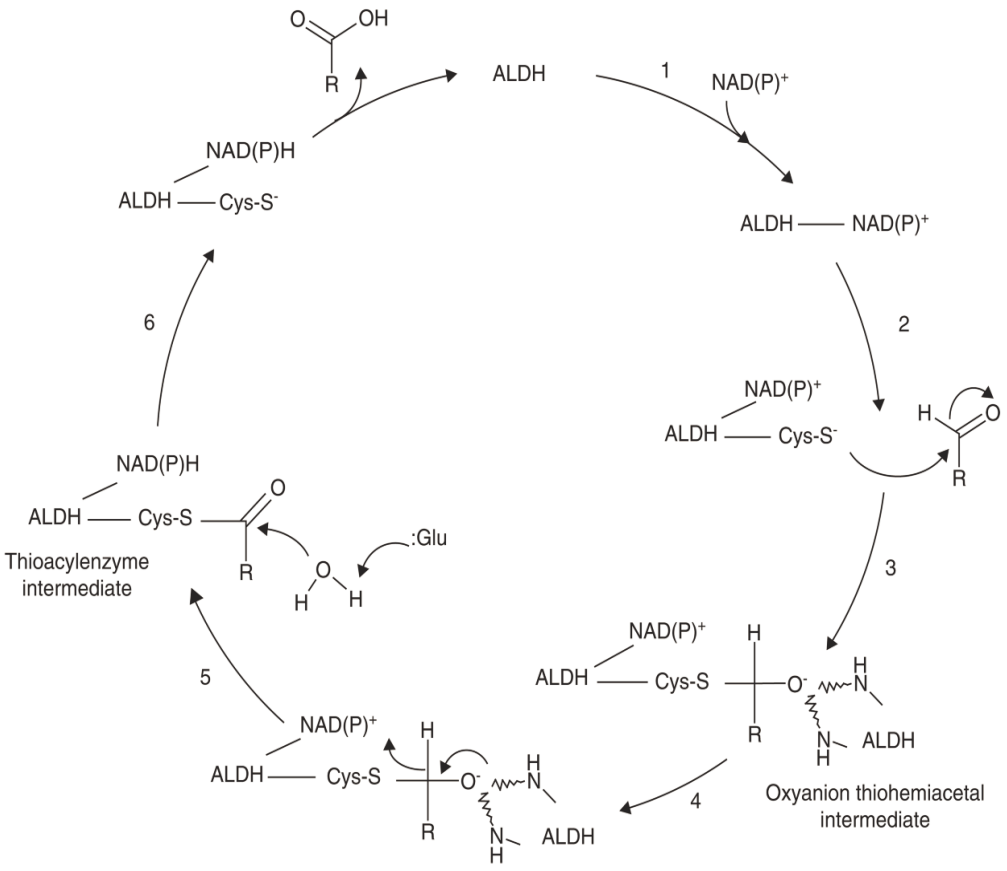


Fig. 3. Scheme of non-CoA dependent ALDH catalysis mechanism. (1) Cofactor binds to enzyme, which activates catalytic thiol (Cys-S$^-$); (2) An aldehyde is attacked; (3) Oxyanion is stabilized; (4) Hydride transfer causes cofactor reduction; (5) Thioacylenzyme intermediate is hydrolysed by the catalysis of glutamate; (6) Carboxylic acid is released followed by cofactor (Marchitti *et al.,* 2008).

2.3 Nomenclature of ALDH superfamily

The nomenclature system was set up according to Human Gene Nomenclature Guidelines (http://www.genenames.org/about/guidelines) (Vasiliou *et al.,* 1999). ALDH amino acid sequences that share more than 40 % identity are classified as the same family. Subfamilies consist of ALDHs sharing 60 % sequence identity or higher. As mentioned above, 13 ALDH families were identified so far: ALDH2, ALDH3, ALDH5, ALDH6, ALDH7, ALDH10, ALDH11, ALDH12, ALDH18, ALDH21, ALDH22, ALDH23 and ALDH24. The ALDH10, ALDH12, ALDH21, ALDH22, ALDH23 and ALDH24 are plant-specific, but the ALDH2, ALDH3, ALDH5, ALDH6, ALDH7 and ALDH18 orthologues occur in mammals (Brocker *et al.,* 2013). ALDH superfamilies were described previously in several plant species. The *Arabidopsis thaliana* ALDHs were described in Kirch *et al.* (2004) and maize (*Zea mays*) in Jimenez-Lopez *et al.* (2010). Description of the rice (*Oryza sativa*) ALDH superfamily was finished in 2010 (Kotchoni *et al.,* 2010). *P. patens* and the algaes *Chlamydomonas reinhardtii* and *Osteococcus tauri* ALDH families were characterized by Wood and Duff, 2009. ALDH family of common grape vine (*V. vinifera*) was described later (Zhang *et al.,* 2012). The number of ALDHs within families 2 and 3 in various species is shown in Tab. 1. The number of ALDH2 family members differs significantly between species because of widening during the evolution of land plants.

Tab. 1. Number of ALDHs from families 2 and 3 and total ALDH numbers identified in various species. ALDH familes 2 and 3 in moss are coloured in red and orange for clarity. (modified according to Brocker *et al.,* 2013).

Species	ALDH2	ALDH3	All ALDHs
A. thaliana	3	3	16
C. reinhardtii	1	-	9
O. sativa	5	5	20
P. patens	**2**	**5**	**21**
V. vinifera	5	4	25
S. moellindorffii	6	2	24
Z. mays	6	5	22
H. sapiens	7	4	19

2.4 ALDHs found in *Physcomitrella patens*

P. patens genome comprises 21 genes included in 11 families (Tab. 2). *ALDH21* and *ALDH23* genes were identified recently and very little is known about their function (Brocker *et al.,* 2013). Only two *ALDH2* family members were identified in *P. patens*, whereas the lycophyte *Selaginella moellindorffii* contains six *ALDH2* genes (Tab. 1). Lycophytes are regarded as fundamental vascular plants. Expansion of *ALDH2* family is possibly the reason of evolution of basal terrestrial plants into basal vascular plants. ALDH2B2 (PpALDH2B) from *P. patens* is phylogenetically related to mitochondrial ALDH2 from *Z. mays*. ALDH2B1 (PpALDH2A) was recently found clustering with ALDH2C subfamily members from other plants, therefore is believed to be cytosolic (Fig. 4) (Končitíková *et al.,* 2015).

Tab. 2. Gene and protein information about ALDHs found in *Physcomitrella patens* (Brocker *et al.,* 2013). ALDH families 2 and 3 are coloured in red and orange for clarity.

Family	Gene name	NCBI gene ID	Other names	NCBI protein ID	Phytozome ID	Amino acids #
2	*ALDH2B1*	5930606	Pp_Aldh2A	XP_001767457	Pp1s93_81V6	553
	ALDH2B2	5948865	Pp_Aldh2B	XP_001785650	Pp1s496_1V6	535
3	*ALDH3H1*	5928120	Pp_Aldh3A	XP_001764841	Pp1s71_128V6	492
	ALDH3H2	5943334	Pp_Aldh3E	XP_001780129	Pp1s272_3V6	583
	ALDH3K1	5920386	Pp_Aldh3B	XP_001757163	Pp1s26_253V6	479
	ALDH3K2	5930362	Pp_Aldh3D	XP_001767194	Pp1s90_226V6	485
	ALDH3K3	5933556	Pp_Aldh3C	XP_001770374	Pp1s124_90V6	467
5	*ALDH5F1*	5916037	Pp_Aldh5B	XP_001752876	Pp1s6_180V6	492
	ALDH5F2	5943671	Pp_Aldh5A	XP_001780466	Pp1s279_72V6	498
6	*ALDH6B1*	5920495	Pp_Aldh6A	XP_001757403	Pp1s27_198V6	574
7	*ALDH7B4*	5941578	Pp_Aldh7A	XP_001778351	Pp1s237_29V6	511
10	*ALDH10A1*	5919866	Pp_Aldh10A	XP_001756623	Pp1s23_38V6	559
11	*ALDH11A5*	5935487	Pp_Aldh11E	XP_001772261	No entry	504
	ALDH11A1	5928282	Pp_Aldh11A	XP_001765101	Pp1s73_32V6	503
	ALDH11A2	5918399	Pp_Aldh11B	XP_001755163	Pp1s16_156V6	496
	ALDH11A3	5916920	Pp_Aldh11D	XP_001753784	Pp1s10_228V6	496
	ALDH11A4	5923860	Pp_Aldh11C	XP_001760657	Pp1s44_64V6	496
12	*ALDH12A1*	5923366	Pp_Aldh12A	XP_001760169	Pp1s41_177V6	571
18	*ALDH18B1*	5919578	-	XP_001756289	Pp1s22_40V6	757
21	*ALDH21A1*	5932362	Pp_Aldh21A	XP_001769187	Pp1s111_161V6	497
23	*ALDH23A1*	5918738	Pp_Aldh23A	XP_001755511	Pp1s18_148V6	494

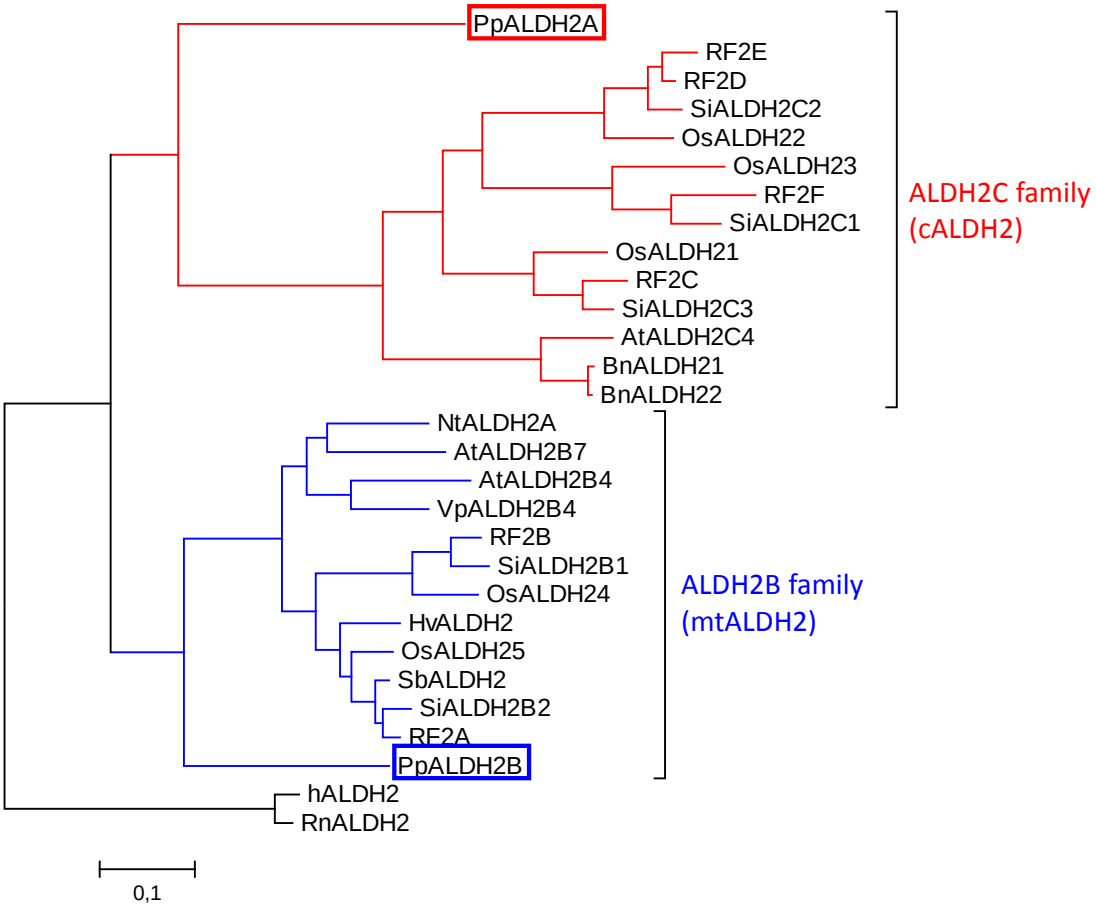


Fig. 4. Phylogenetic tree of the ALDHs found in *P. patens* and other well-characterized plant ALDHs. Families ALDH2C and ALDH2B are highlighted with red and blue, respectively. ALDH sequences from *Antirrhinum majus (*FJ151199), *Arabidopsis thaliana* (At3g48000, At1g23800 and At3g24503), *Brassica napus* (FN995990 and FN995991), *Homo sapiens* (AY621070), *Hordeum vulgare* (BAB62757), *Nicotiana tabacum* (CAA71003), *Oryza sativa* (Os06g15990, Os02g49720, Os01g40860, Os01g40870 and Os06g39230), *Physcomitrella patens* (XP_001767457 and XP_001785650), *Rattus norvegicus* (P11884), *Sorghum bicolor* (BAB92019), *Vitis pseudoreticulata* (DQ150256), *Setaria italica* (XM_004953741, XM_004965148, XM_004965940, XM_004968994 and XM_004968990) and *Zea mays* (NP_001105891, AHM26657, AHM26658, AIV00510, AIV00511, AHM26659) were used to construct the tree in MEGA 6.0 (modified according to Končitíková *et al.,* 2015).

2.5 ALDH2 family

Plant ALDH2 family members (EC 1.2.1.-) have been identified as mitochondrial (mt) and cytosolic (c) and have been sorted into ALDH2B and ALDH2C subfamilies, respectively. Their amino acid sequence is identical approximately from 54 to 63 % to human ALDH2 (hALDH2), which is included in

ALDH2A subfamily (Končitíková *et al.,* 2015). hALDH2 is involved in ethanol metabolism, because it catalyses conversion of acetaldehyde to acetate (Klyosov, 1996). ALDH2 enzymes are homotetrameric and broad substrate specificity is their characteristic feature.

The *ALDH2* genes were studied in many plant species. The first *ALDH2* found were described as fertility-restorers in maize. A mutation in one of the two restorer genes, *rf2,* lead to cytoplasmic male sterility linked to a production of non-viable pollen (Cui *et al.,* 1996). RF2 (later renamed RF2A) was characterized as mtALDH (Liu *et al.,* 2001). Two *RF2* genes coding for mtALDHs were identified in *Zea mays*, *RF2A* and *RF2B* (Liu and Schnable, 2002), and were named as *ALDH2B2* and *ALDH2B5,* respectively, according to Human Gene Nomenclature Guidelines (Brocker *et al.,* 2013). Two mtALDH2 from *O. sativa* have been found abundant in panicles (Nakazono *et al.,* 2000). Two mtALDH2 from tobacco (*Nicotiana tabacum*) are involved in pollen growth, because of their accumulation in male and female reproductive organs (op den Camp and Kuhlemeier, 1997). *A. thaliana* also contains two mtALDH2 members (Skibbe *at al.,* 2002). Another mtALDH2 was characterized in snapdragon (*Antirrhinum majus*) exhibiting BALDH activity (Long *et al.,* 2009). Five mtALDH2 enzymes are present in soybean (*Glycine max*), where they metabolize aldehydes in various cellular compartments (Kotchoni *et al.,* 2012). Major function of all plant mtALDH2 enzymes is catalysing the conversion of acetaldehyde generated in ethanolic fermentation to acetate. Acetate is subsequently used for CoA synthesis bypassing pyruvate dehydrogenase pathway (Fig. 5) (op den Camp and Kuhlemeier, 1997; Wei *et al.,* 2009). RF2A is capable of oxidizing aromatic aldehydes, such as anisaldehyde involved in plant-insect interactions (Teulon *et al.,* 1993; Kubo and Kinst-Hori, 1998) and redox cycling of H_2O_2 (Guillen and Evans, 1994), or cinnamaldehyde being involved in lignin biosynthesis (Kajita *et al.,* 1996).

Cytosolic ALDH2C4 from *A. thaliana* and oilseed rape (*Brassica napus*) was identified and characterized as reduced epidermal fluorescence 1 (REF1) and is involved in biosynthesis of ferulic and sinapic acids. Nair *et al.* (2004) proved that *ref1* mutant plants have reduced levels of cell wall-linked ferulate

esters to 50% than wild-type plants and also exhibit decreased sinapaldehyde dehydrogenase specific activity to 20%. The *fah1* mutant of *A. thaliana* lacks ferulate-5-hydroxylase and did not accumulate sinapic acid-derived metabolites (Meyer *et al.,* 1996) but trace amounts of feruoylmalate were present (Hemm *et al.,* 2003). Leaf extracts of double mutant *ref1 fah1* were analysed by HPLC and no feruoylmalate was detected (Nair *et al.,* 2004). These experiments provide evidence that REF1 is essential for the biosynthesis of ferulate and sinapate esters that contribute to the formation of a cell wall (Grabber et al., 2000). These compounds are generated from coniferaldehyde and sinapaldehyde. Sinapaldehyde and coniferaldehyde originates in phenylpropanoid pathway as well as cinnamaldehyde. cALDH2s from *Z. mays* comprise RF2C, RF2D, RF2E and RF2F, named ALDH2C1, ALDH2C2, ALDH2C4 and ALDH2C5, respectively (Skibbe *et al.,* 2002; Končitíková *et al.,* 2015). They oxidize broad range of aromatic and also aliphatic aldehydes (Fig. 5).

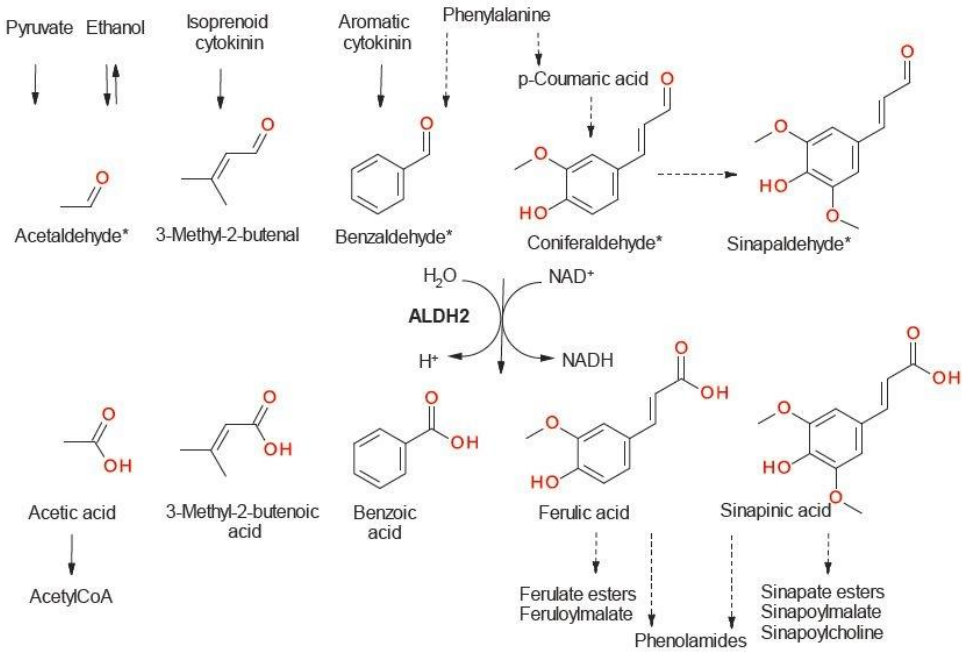


Fig. 5. Conversions of acetaldehyde, 3-methyl-2-butenal, benzaldehyde, coniferylaldehyde and sinapaldehyde catalysed by plant ALDH2. (Končitíková *et al.,* 2015). The substrates marked with an asterisk were confirmed *in vivo* (Wei et al., 2009; Long et al., 2009; Nair et al., 2004; Mittasch et al., 2013).

Maize ALDH2C isoforms could be involved in the benzenoid and phenylpropanoid pathways, as they are capable of oxidizing benzaldehyde and

phenylacetaldehyde, volatile compounds responsible for the floral scent in *Petunia hybrida* (Končitíková *et al.,* 2015; Boatright *et al.,* 2004). Benzaldehyde and phenylacetaldehyde arise from phenylalanine generated in shikimate or arogenate pathways (Tzin and Galili, 2010; Maeda and Dudareva, 2012). Benzaldehyde is also generated by the oxidation of aromatic cytokinins. The phenylpropanoid pathway gives rise to many secondary metabolites, e.g. flavonoids, coumarins, phenolamides or monolignol, the last named being part of the lignin biosynthesis. Short-chained aliphatic aldehydes, such as butyraldehyde or valeraldehyde, are derived from branched-chain amino acids by decarboxylation followed by deamination or directly by transamination followed by carboxylation (Dudareva *et al.,* 2013). Aliphatic 3-methyl-2-butenal (isopentenal) originates from the oxidation of isoprenoid cytokinins (Kopečný *et al.,* 2005). Other aliphatic aldehydes have been found as substrates to maize ALDH2C isoforms, such as hexanal, nonanal and their unsaturated forms *trans*-2-hexenal and *trans*-2-nonenal, which originates from C_{18} unsaturated linolenic or linoleic acids *via* the lipoxygenase pathway (Dudareva *et al.,* 2013) (Fig. 6). These compounds are volatile giving fruits and vegetables their characteristic aroma.

Human liver mtALDH2 crystal structure was solved by Ni *et al.* (1999) and shares 95% amino acid sequence identity to bovine mtALDH2 (Steinmetz et al., 1997). Active site of these enzymes possesses residues T244, E268, E476 and cluster of three neighbouring cysteine residues (C302, C303 and C304).

Crystal structures of two cALDH2 members RF2C and RF2F were solved only recently (Končitíková *et al.,* 2015). Active sites of RF2C and RF2F differ significantly (Fig. 7A). The substrate cavity in RF2F is wider than in RF2C due to the nonpolar residues V192 and M477, which corresponds to the presence of aromatic residues F178 and F460, respectively, in RF2C. A316 and E495 (I302 and M478 in RF2C) also cause widening of the substrate channel. This fact probably leads to weaker nonpolar interactions with diverse substrates demonstrated by higher K_m values. Active site residues in RF2D and RF2E are not different, which corresponds with similar kinetic properties. A model of RF2E

was created by SWISS-MODEL using RF2C as a template (Fig. 7B). The cavity is slightly wider due to L177 (F178 in RF2C).

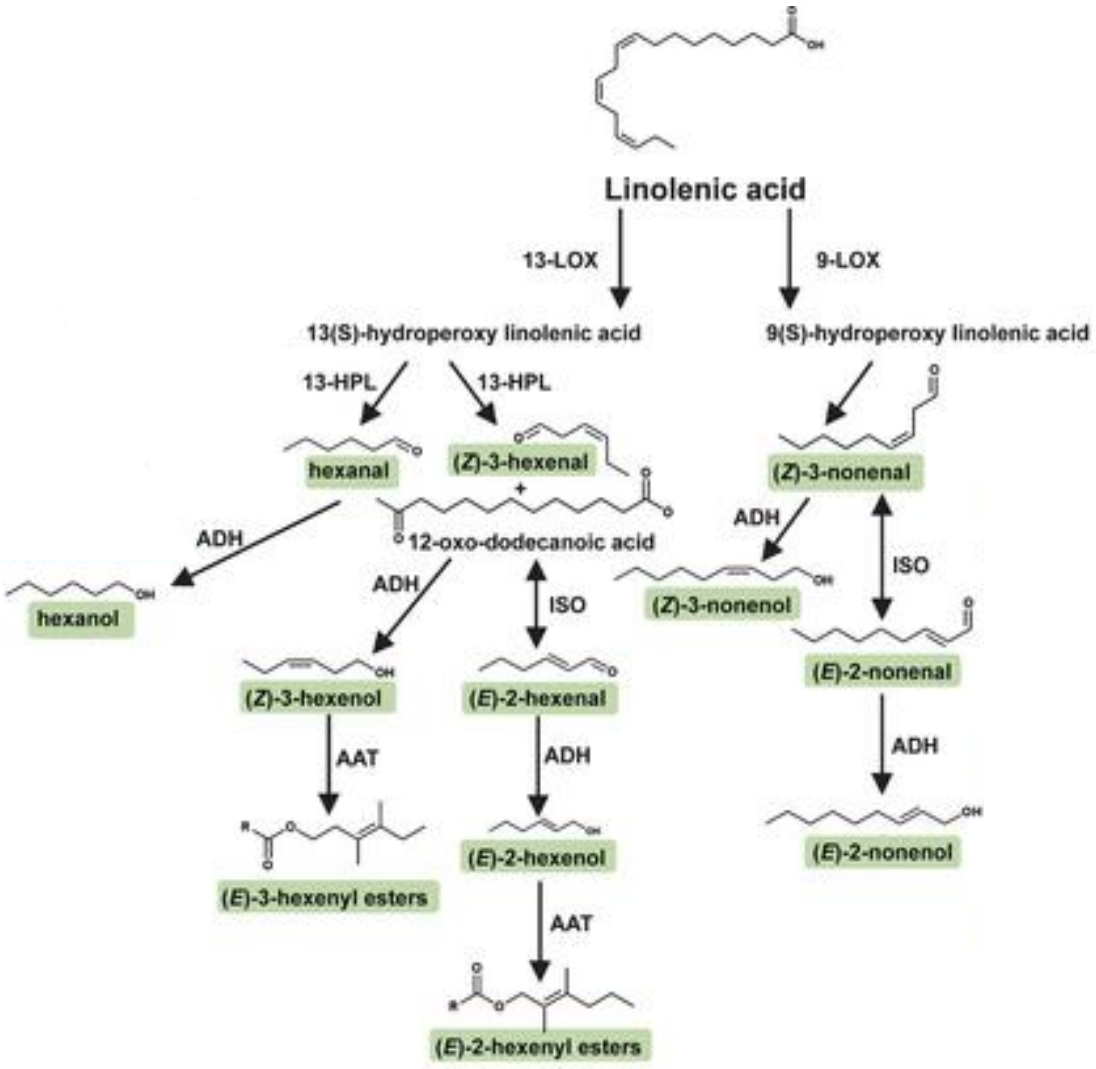


Fig. 6. Biosynthesis and metabolic pathways of several aliphatic aldehydes. Hexanal, *trans*-2-hexenal and *trans*-2-nonenal are substrates to maize ALDH2C isoforms. Abbreviations: 13-LOX – 13-lipoxygenase, 13-HPL – 13-hydroperoxylyase, 9-LOX – 9-lipoxygenase, AAT – alcohol acyltransferase, ADH – alcohol dehydrogenase, ISO – isomerase (Dudareva *et al.,* 2013; Končitíková *et al.,* 2015).

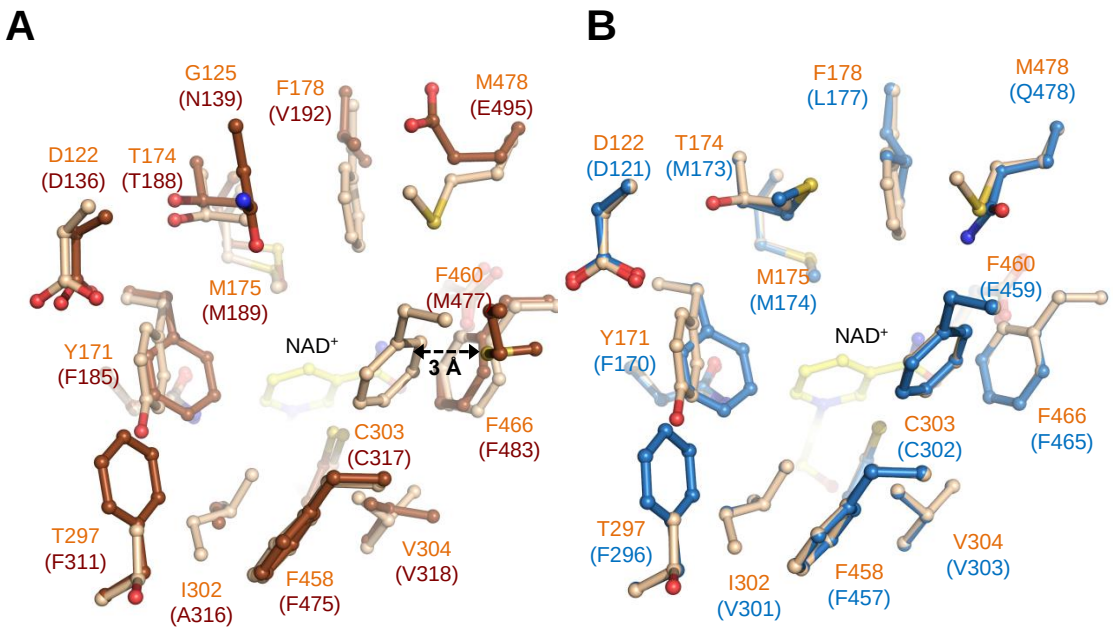


Figure 7. Position of critical residues within the active sites of two isoenzymes from maize. (A) A top view on substrate channel of RF2C (light brown with residues labelled in orange) and RF2F (dark brown, residues shown in brackets). (B) Model of the substrate cavity of RF2E (blue) made using SWISS-MODEL (http://swissmodel.expasy.org/) with RF2C as a template (light brown) (Končitíková *et al.,* 2015).

2.6 ALDH3 family

Mammal ALDH3 family members are cytosolic and mitochondrial enzymes (EC. 1.2.1.5) (Marchitti *et al.,* 2008). Various subcellular localization was predicted for plant ALDH3 homologues. They can be located in cytosol, chloroplasts and mitochondria (Jimenez-Lopez *et al.,* 2010; Kirch *et al.,* 2004; Stiti *et al.,* 2011). It is believed that many *ALDH3* genes are expressed under regulation of the abscisic acid as a response to stress due to the analysis of *CpALDH*, which was isolated from the resurrection plant *Craterostigma plantagineum.* Kinetic measurements showed that *CpALDH* might be part of detoxification system of reactive aldehydes generated by lipid peroxidation, which is a response to oxidative stress (Kirch *et al.,* 2001).

The plant *ALDH3* family is divided into six subfamilies – *ALDH3E*, *3F*, *3H*, *3I*, *3J*, and *3K*. Therefore it is a very expanded family. Five genes coding for *ALDH3* homologues were found in *Z. mays*, *O. sativa* and *P. patens*. On the

contrary, *A. thaliana* contains three *ALDH3* genes (Tab. 1). Presence of *ALDH3* members in algaes *C. reinhardtii* and *V. carteri* was not proved and hence it was suggested, that diversification of the genes within this family set in with water-to-land transition of plants (Kirch *et al.,* 2005). It refers to high probability of evolution of these genes as a result of functional specialization within specific tissues and subcellular compartments.

2.6.1 ALDH3 from *Arabidopsis thaliana*

A. thaliana ALDH3 family consists of three members (Brocker *et al.,* 2013). Analysis of transcript accumulation revealed that *ALDH3I1* from *A. thaliana* respond to stress conditions induced by ABA treatment, heavy metals Cu^{2+} and Cd^{2+}, salt stress, dehydration, hydrogen peroxide or chemicals inducing oxidative stress. *ALDH3H1* responded to these types of stress weakly, whereas *ALDH3F1* is expressed constitutively and therefore it seems not to be related to stress (Kirch *et al.,* 2001; 2005; Sunkar *et al.,* 2003). *ALDH3I1* T-DNA knockout mutant lines exhibited higher sensitivity to dehydration and salt than wild-type plants. Transgenic *A. thaliana* overexpressing *ALDH3I1* showed decreased levels of malondialdehyde (MDA), which is the end product of lipid peroxidation, and H_2O_2 causing oxidative stress. Knockout *ALDH3I1* lines were more sensitive to MDA and H_2O_2 than wild-type plants (Kotchoni *et al.,* 2006). This fact indicates that ALDH3I1 plays role as reactive oxygen species scavenger and also inhibits lipid peroxidation.

Substrate specificity studies of AtALDH3H1 and AtALDH3I1 were carried out by Stiti *et al.* (2011b). Both enzymes were found to be able to oxidize medium- to long-chain aliphatic aldehydes with a preference for long-chain aldehydes. Short-chain aldehydes were not good substrates due to high K_m values, e.g. propionaldehyde with K_m of 510 ± 59 μM. Both enzymes preferentially use NAD^+. The results suggested preference of both enzymes for saturated aldehydes, such as dodecanal being the best substrate (K_m = 5 ± 1 μM and 1.3 ± 0.2 μM for ALDH3H1 and ALDH3I1, respectively). K_m values were also found to be generally higher with unsaturated than saturated

aldehydes, for example 180 ± 24 µM for ALDH3H1 with *trans*-2-hexenal but 71 ± 12 µM with hexanal.

ALDH3F1, ALDH3H1 and ALDH3I1-GFP fusion proteins were localized in different compartments of the plant cells (Fig. 8) (reviewed in Stiti *et al.,* 2011). That refers to functional specialization of individual ALDH3 members and to the fact that different cellular compartments (chloroplasts and cytosol) possess ALDH enzymes with specific biochemical properties.

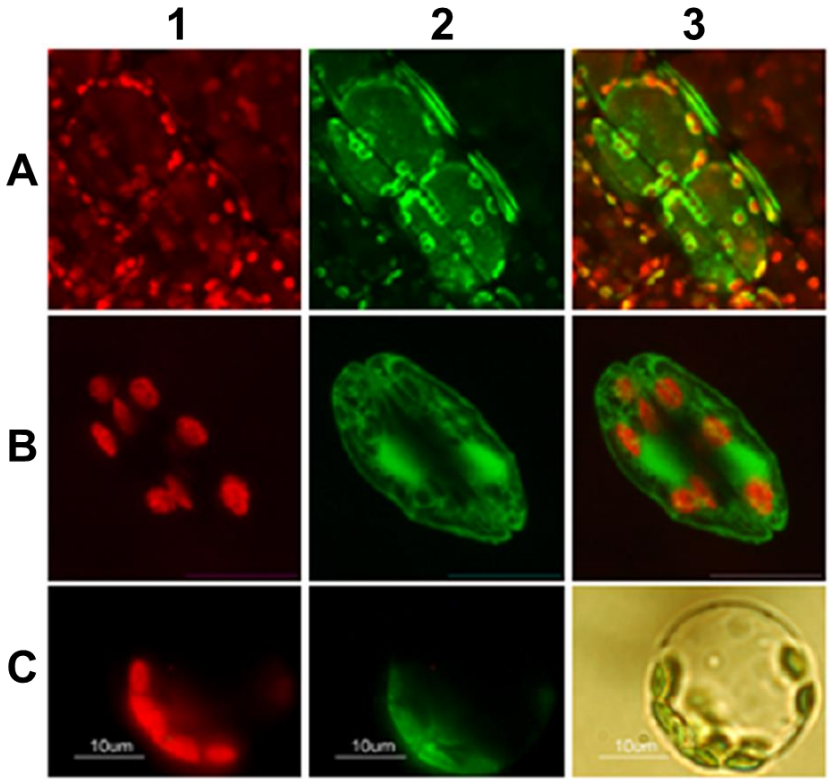


Fig. 8. Subcellular localization of AtALDH3-GFP fusion proteins. 1, 2 and 3 show chlorophyll autofluorescence, GFP-signal and merged signals, respectively. (C3) shows the transformed protoplast in bright field to make chloroplasts visible. (A) ALDH3F1 and (B) ALDH3H1 are located in cytosol, (C) ALDH3I1 is located in chloroplasts (modified according to Stiti *et al.,* 2011).

2.7 Utilization of ALDH enzymes

Many studies suggested that ALDHs could play an important role in abiotic stress-response pathways in plants. Stress-tolerant crops with increased

resistance to unfavourable and stressful conditions, e.g. dehydration or high soil salinity, could be developed based on this knowledge.

Ferulic acid derivatives were proved to be the part of cross-linking of polysaccharides to lignin (Grabber *et al.,* 2000). REF1 enzyme was found to be required for ferulic acid but not lignin biosynthesis (Nair *et al.,* 2004). This points at the alternative way how the content of phenylpropanoid compounds can be affected in plant cell wall. Therefore digestibility of forages without compromising agronomic performance could be improved. The quality of vegetables could be modified and enhanced due to the fact that ferulic acid derivatives present in cell wall contribute to its texture (Waldron *et al.,* 1997).

Silencing of *ref1* in *B. napus* lead to lower accumulation of sinapoylmalate and cell wall-associated phenolic compounds in leafs. This fact brings the potential to increase the sensitivity of plants lacking *BnREF1* against pathogen stress and UV-B (Mittasch *et al.,* 2013). Content of sinapine, which is abundant antinutritive compound in the seeds of *B. napus*, can be reduced by mutation in both *BnREF1* and *BnSGT* genes resulting in improved quality of oilseed rape as an animal feed or in food industry (Emrani *et al.,* 2015).

3 Summary

Theoretical part of this work comprises characterization of model organism *Physcomitrella patens* and description of family 2 and 3 plant ALDHs studied so far as well as their reaction mechanism and various properties.

Cloning and expression of *PpALDH3D* and two isoforms of *ALDH2* genes, *PpALDH2A* and *PpALDH2B* was carried out in experimental part. After optimizing the expression, recombinant proteins were produced and purified by affinity chromatography. Various properties of recombinant proteins were estimated such as molecular weight, folding and subcellular localization in moss protoplasts. Substrate specificity was determined for ΔPpALDH2A and ΔPpALDH2B.

4 References

Anterola, A., Shanle, E., Perroud, P. F., Quatrano, R. (2009) Production of taxa-4(5),11(12)-diene by transgenic *Physcomitrella patens. Transgenic Res*. **18**, 655-660.

Ashton, N. W., Cove, D. J. (1977) Isolation and preliminary characterization of auxothophic and analog resistant mutants of moss, *Physcomitrella patens. Molec. gen. Genet.* **154**, 87-95.

Ashton, N. W., Grimsley, N. H., Cove, D. J. (1979) Analysis of gametophytic development in the moss, *Physcomitrella patens*, using auxin and cytokinin resistant mutants. *Planta* **144**, 427-435.

Awal H. M. A., Kinoshita T., Yoshida I., Doe M., Hirasawa E. (1997) Aminoaldehyde dehydrogenase of pea epicotyls. *Phytochemistry* **44**, 997-1000.

Brocker, C., Vasiliou, M., Carpenter, S., Carpenter, C., Zhang, Y. C., Wang, X. P., 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.

Brooks, P. J., Theruvathu, J. A. (2005) DNA adducts from acetaldehyde: implications for alcohol-related carcinogenesis. *Alcohol* **35**, 187-193.

Cove, D. J. (2005) The moss *Physcomitrella patens*. *Annu. Rev. Genet.* **39**, 339-358.

Cronk, Q. C. B. (2001) Plant evolution and development in a post-genomic context. *Nat. Rev. Genet.* **2**, 607-619.

Cui, X. Q., Wise, R. P., Schnable, P. S. (1996) The *rf2* nuclear restorer gene of male-sterile T-cytoplasm maize. *Science* **272**, 1334-1336.

Decker, E. L., Reski, R. (2008) Current achievements in the production of complex biopharmaceuticals with moss bioreactors. *Bioprocess. Biosyst. Eng.* **31**, 3-9.

Dudareva, N., Klempien, A., Muhlemann, J. K., Kaplan, I. (2013) Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytol.* **198**, 16-32.

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

Emrani, N., Harloff, H. J., Gudi, O, Kopisch-Obuch, F., Jung, C. (2015) Reduction in sinapine content in rapeseed (*Brassica napus* L.) by induced mutations in sinapine biosynthesis genes. *Mol. Breed.* **35**, 37.

Esterbauer, H., Schaur, R. J., Zollner, H. (1991) Chemistry and biochemistry of 4-hydroxynonenal, malondialdehyde and related aldehydes. *Free Radic. Biol. Med.* **11**, 81-128.

Finet, C., Timme, R. E., Delwiche, C. F., Marleta, F. (2010) Multigene phylogeny of the green lineage reveals the origin and diversification of land plants. *Curr*. *Biol*. **20**, 2217-2222.

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. *J. 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. *J. Proteomics* **75**, 4027–4037.

Grabber, J. H., Ralph, J., Hatfield, R. D. (2000) Cross-linking of maize walls by ferulate dimerization and incorporation into lignin. *J. Agric. Food Chem*. **48**, 6106-6113.

Guillen, F., Evans, C. S. (1994) Anisaldehyde and veratraldehyde acting as redox cycling agens for H_2O_2 production by *Pleurotus eryngii. Appl. Environ. Microb.* **60**, 2811-2817.

Hemm, M. R., Ruegger, M. O., Chapple, C. (2003) The *Arabidopsis ref2* mutant is defective in the gene encoding CYP83A1 and shows both phenylpropanoid and glucosinolate phenotypes. *Plant Cell* **15**, 179-194.

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

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

Kajita, S., Katayama, Y., Omori, S. (1996) Alterations in the biosynthesis of lignin in transgenic plants with chimeric genes for 4-coumarate:Coenzyme A ligase. *Plant Cell Physiol.* **37**, 957-965.

Kenrick, P., Crane, P. R. (1997) The origin and early evolution of plants on land. *Nature* **389**, 33-39.

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

Kirch, H. H., Nair, A., Bartels, D. (2001) Novel ABA- and dehydration-inducible aldehyde dehydrogenase genes isolated from the resurrection plant *Craterostigma plantagineum* and *Arabidopsis thaliana. Plant J.* **28**, 555-567.

Kirch, H. H., Schlingensiepen, S., Kotchoni, S. O., Sunkar, R., Bartels, D. (2005) Detailed expression analysis of selected genes of the aldehyde dehydrogenase (ALDH) gene superfamily in *Arabidopsis thaliana. Plant Mol. Biol*. **57**, 315-332.

Kitson, T. M., Hill, J. P., Midwinter, G. G. (1991) Identification of a catalytically essential nucleophilic residue in sheep liver cytoplasmic aldehyde dehydrogenase. *Biochem. J.* **275**, 207-210.

Klyosov, A. A. (1996) Kinetics and specificity of human liver aldehyde dehydrogenases toward aliphatic, aromatic, and fused polycyclic aldehydes. *Biochemistry* **35**, 4457-4467.

Končitíková, R., Vigoroux, A., Kopečná, M., Andree, T., Bartoš, J., Šebela, M., Moréra, S., Kopečný, D. (2015) Role and structural characterization of plant aldehyde dehydrogenases from family 2 and family 7. *Biochem. J.* **468**, 109-123.

Kopečný, D., Pethe, C., Šebela, M., Houba-Hérin, N., Madzak, C., Majira, A., Laloue, M. (2005) High-level expression and characterization of *Zea mays* cytokinin oxidase/dehydrogenase in *Yarrowia lipolytica. Biochimie* **87**, 1011-1022.

Kopřivová, A., Stemmer, C., Altmann, F., Hoffmann, A., Kopřiva, S., Gorr, G., Reski, R., Decker, E. L. (2004) Targeted knockouts of *Physcomitrella* lacking plant specific immunogenic N-glycans. *Plant Biotechnol. J.* **2**, 517-523.

Kotchoni, S. O., Kuhns, C., Ditzer, D., Kirch, H. H., Bartels, D. (2006) Over-expression of different aldehyde dehydrogenase genes in *Arabidopsis thaliana* confers tolerance to abiotic stress and protects plants against lipid peroxidation and oxidative stress. *Plant Cell Environ.* **29**, 1033-1048.

Kotchoni, S. O., Jimenez-Lopez, J. C., Gao, D. Y., Edwards, V., Gachomo, E. W., Margam, V. M., Seufferheld, M. J. (2010) Modeling-Dependent protein characterization of the rice aldehyde dehydrogenase (ALDH) superfamily reveals distinct functional and structural features. *PLoS One* **5**, e11516.

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

Kubo, I., Kinst-Hori, I. (1998) Tyrosinase inhibitors from anise oil. *J. Agr. Food Chem.* **46**, 1268-1271.

Lindahl, R. (1992) Aldehyde dehydrogenases and their role in carcinogenesis. *Crit. Rev. Biochem. Mol. Biol.* **27**, 283-335.

Liu, F., Cui, X. Q., Horner, H. T., Weiner, H., Schnable, P. S. (2001) Mitochondrial aldehyde dehydrogenase activity is required for male fertility in maize. *Plant Cell* **13**, 1063-1078.

Liu, F., Schnable, P. S. (2002) Functional specialization of maize mitochondrial aldehyde dehydrogenases. *Plant Physiol.* **130**, 1657-1674.

Long, M. C., Nagegowda, D. A., Kaminaga, Y., Ho, K. K., Kish, C. M., Schnepp, J., Sherman, D., Weiner H., Rhodes, D., Dudareva, N. (2009) Involvement of snapdragon benzaldehyde dehydrogenase in benzoic acid biosynthesis. *Plant J.* **59**, 256-265.

Maeda, H., Dudareva, N. (2012) The shikimate pathway and aromatic amino acid biosynthesis in plants. *Annu. Rev. Plant Biol.* **63**, 73-105.

Marchitti, S. A., Brocker, C., Stagos, D., Vasiliou, V. (2008) Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase superfamily. *Expert Opin. Drug Metab. Toxicol.* **4**, 697-720.

Meyer, K., Cusumano, J. C., Somerville, C., Chapple, C. C. S. (1996) Ferulate-5-hydroxylase from *Arabidopsis thaliana* defines a new family of cytochrome P450-dependent monoxygenases*. Proc. Natl. Acad. Sci. U. S. A.* **93**, 6869-6874.

Mittasch, J., Böttcher, C., Frolov, A., Strack, D., Milkowski, C. (2013) Reprogramming the phenylpropanoid metabolism in seeds of oilseed rape by supressing the orthologs of reduced epidermal fluorescence1. *Plant Physiol.* **161**, 1656-1669.

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. *J. Chromatogr. A* **1216**, 3629–3636.

Nadkrani, D. V., Sayre, L. M. (1995) Structural definition of early lysine and histidine adduction chemistry of 4-hydroxynonenal. *Chem. Res. Toxicol.* **8**, 284-291.

Nair, R. B., Bastress, K. L., Ruegger, M. O., Denault, J. W., Chapple, C. (2004) The *Arabidopsis thaliana REDUCED EPIDERMAL FLUORESCENCE1* gene encodes an aldehyde dehydrogenase involved in ferulic and sinapic acid biosynthesis. *Plant cell* **16**, 544-554.

Ni, L., Zhou, J. Z., Hurley, T. D., Weiner, H. (1999) Human liver mitochondrial aldehyde dehydrogenase: Three-dimensional structure and the

restoration of solubility and activity of chimeric forms. *Protein Sci.* **8**, 2784-2790.

op den Camp, R. G. L., Kuhlemeier, C. (1997) Aldehyde dehydrogenase in tobacco pollen. *Plant Mol. Biol.* **35**, 355-365.

Peitsch, M. C., Schwede, T., Guex, N. (2000) Automated protein modelling: the proteome in 3D. *Pharmacogenomics* **1**, 257-266.

Qiu, Y. L. (2008) Phylogeny and evolution of charophytic algae and land plants. *J. Syst. Evol.* **46**, 287-306.

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

Reski, R., Abel, W. O. (1985) Induction of budding on chloronemata and caulonemata of the moss, *Physcomitrella patens*, using isopentenyladenine. *Planta* **165**, 354-358.

Schaaf, A., Tintelnot, S., Baur, A., Reski, R., Gorr, G., Decker, E. L. (2005) Use of endogenous signal sequances for transient production and efficient secretion by moss (*Physcomitrella patens*) cells. *BMC Biotechnol.* **5**(30).

Schaefer, D. G., Zrÿd, J. P. (1997) Efficient gene targeting in the moss *Physcomitrella patens. Plant J*. **11**, 1195-1206.

Schaefer, D. G., Zrÿd, J. P. (2001) The moss *Physcomitrella patens*, now and then. *Plant Physiol.* **127**, 1430-1438.

Schaefer, D. G., Zrÿd, J. P., Knight, C. D., Cove, D. J. (1991) Stable transformation of the moss *Physcomitrella patens. Mol. Gen. Genet.* **226**, 418-424.

Schiff, P. B., Fant, J., Horwitz, S. B. (1979) Promotion of microtubule assembly *in vitro* by taxol. *Nature* **277**, 665-667.

Schween, G., Hohe, A., Kopřivová, A., Reski, R. (2003) Effects of nutrients, cell density and culture techniques on protoplast regeneration and early protonema development in a moss, *Physcomitrella patens*. *J. Plant Physiol.* **160**, 209-212.

Steinmetz, C. G., Xie, P. G., Weiner, H., Hurley, T. D. (1997) Structure of mitochondrial aldehyde dehydrogenase: The genetic component of ethanol aversion. *Structure* **5**, 701-711.

Stiti N., Missihoun, T. D., Kotchoni, S. O., Kirch, H. H., Bartels, D. (2011) Aldehyde dehydrogenases in *Arabidopsis thaliana*: biochemical requirements, metabolic pathways, and functional analysis. *Frontiers Plant Sci.* **2**(65).

Stiti, N., Adewale, I. O., Petersen, J., Bartels, D., Kirch, H. H. (2011b) Engineering the nucleotide coenzyme specificity and sulfhydryl redox sensitivity of two stress-responsive aldehyde dehydrogenase isoenzymes of *Arabidopsis thaliana. Biochem. J.* **434**, 459-471.

Strotbek, C., Krinniger, S., Frank, W. (2013) The moss *Physcomitrella patens*: methods and tool from cultivation to targeted analysis of gene function. *Int. J. Dev. Biol.* **57**, 553-564.

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

Sunkar, R., Bartels, D., Kirch, H. H. (2003) Overexpression of a stress-induced aldehyde dehydrogenase gene from *Arabidopsis thaliana* in transgenic plants improves stress tolerance. *Plant J*. **35**, 452-464.

Teulon, D. A. J., Penman, D. R., Ramakers, P. M. J. (1993) Volatile chemicals for thrips (*Thysanoptera*, *Thripidae*) host-finding and applications for thrips pest-management. *J. Econ. Entomol.* **86**, 1405-1415.

Tzin, V., Galili, G. (2010) New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants. *Mol. Plant.* **3**, 956-972.

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

Waldron, K. W., Smith, A. C., Parr, A. J., Ng, A., Parker, M. L. (1997) New approaches to understanding and controlling cell separation in relation to fruit and vegetable texture. *Trends Food Sci. Technol.* **8**, 213-221.

Wei, Y. L., Lin, M., Oliver, D. J., Schnable, P. S. (2009) The roles of aldehyde dehydrogenases (ALDHs) in the PDH bypass of *Arabidopsis*. *BMC Biochem*. **10**(7).

Weng, J. K., Chapple, C. (2010) The origin and evolution of lignin biosynthesis. *New Phytol.* **187**, 273-285.

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

Zhang, Y. C., Mao, L. Y., Wang, H., Brocker, C., Yin, X. J., Vasiliou, V., Fei, Z. J., Wang, X. P. (2012) Genome-wide identification and analysis of grape aldehyde dehydrogenase (ALDH) gene superfamily. *PLoS One* **7**, e32153.

5 List of abbreviations

ABA	abscisic acid
ALDH	aldehyde dehydrogenase
BALDH	benzaldehyde dehydrogenase
BSA	bovine serum albumin
cALDH	cytosolic aldehyde dehydrogenase
CD	circular dichroism
CLSM	confocal laser scanning microscopy
CoA	coenzyme A
GFP	green fluorescent protein
GPC	gel permeation chromatography
hALDH	human aldehyde dehydrogenase
HRP	horseradish peroxidase
mtALDH	mitochondrial aldehyde dehydrogenase
MDA	malondialdehyde
NAD	nicotinamide adenine dinucleotide
PEG	polyethylene glycol
PVDF	polyvinylidene difluoride
SDS	sodium dodecyl sulfate
TBS	Tris Buffered Saline
TFA	trifluoroacetic acid