

Supplemental Figures

Suppl. Fig. 1: Phylogenetic analysis of *Arabidopsis* Acd proteins.

The protein sequences of all *Arabidopsis* sHsps and Acd proteins including AtHsp15.7 and AtAcd31.2 (Scharf et al., 2001) were retrieved from Genbank and the E value of the Hsp20 domain determined (HMMER, Suppl. Table II). In case of improved gene structure prediction, former acronyms (Scharf et al., 2001) were corrected for the new molecular mass of the full-length protein (see Suppl. Table II). Proteins with an E value higher than 0.01 of the Hsp20 domain (see Suppl. Table II) were not considered significant and excluded from phylogenetic analyses. The *Arabidopsis* proteins were supplemented by predicted protein sequences containing an Acd domain from the red algae *Cyanidioschyzon merolae* (merolae.biol.s.u-tokyo.ac.jp) and *Galdieria sulphuraria* (genomics.msu.edu/galdieria/, assembly of ESTs), yeast (*S. cerevisiae*, e.g. ScHsp26 and ScHsp42, Haslbeck et al., 2004; *Ashbya gossypii*, *Candida glabrata*, *Kluyveromyces lactis*, *Yarrowia lipolytica*), fungi (*Neurospora crassa*, *Aspergillus fumigatus*, and *Magnaporthe grisea*), apicomplexa (*Theileria annulata*; *Plasmodium falciparum*, *Trypanosoma brucei*; *Toxoplasma gondii*), the protists *Dictyostelium discoideum* and *Schistosoma japonicum*, *Homo sapiens*, and two members each for the subdivisions of alpha (*Caulobacter crescentus* and *Rickettsia prowazekii*), beta (*Nitrosomonas europaea* and *Ralstonia eutropha*), gamma (*Escherichia coli* and *Pseudomonas syringae*), and delta proteobacteria (*Desulfovibrio vulgaris* and *Myxococcus xanthus*), and cyanobacteria (*Nostoc* sp. and *Synechocystis* sp.). Notably, Hsp homologs with a putative PTS1 are also detected in a few other eukaryotes other than *Arabidopsis* (Dd_Acd20.0, SKL>; Dd_Acd48.1, AHF>, both in *Dictyostelium discoideum*; Kl_Acd22.2, AKL> in *Kluyveromyces lactis*), possibly indicating that these homologs are also targeted to peroxisomes. The sequences comprising the Hsp20 domain of these proteins (CDART) were aligned using ClustalX, analyzed by the neighbor joining method, and an unrooted tree was drawn by treeview. Abbreviations and accession numbers are as follows: Af_Acd20.5 (XP_754172), .Af_Acd37.7 (XP_753951), Af_Acd23.0 (XP_753614), .Af_Acd21.6 (XP_750539), *Aspergillus fumigatus* Af293; Ag_Acd18.7 (AAS54004), Ag_Acd22.8 (AAS53808), Ag_Acd19.1 (AAS53138), Ag_Acd18.2 (AAS53062), Ag_Acd20.2 (AAS50238), *Ashbya gossypii* ATCC 10895; Cc_Hsp17.8 (NP_422386), Cc_Hsp16.9 (NP_421061), *Caulobacter crescentus* CB15; Cg_Acd42.9 (CAG58642), *Candida glabrata* CBS138; Cm_Acd27.1 (CMJ101C), Cm_Acd20.6 (CMJ100C), *Cyanidioschyzon merolae*; Dd_Acd14.8 (XP_647206), Dd_Acd31.9 (XP_644663), Dd_Acd21.1 (XP_643793), Dd_Acd23.5 (XP_642604), Dd_Acd20.0 (XP_643557, SKL>), Dd_Acd12.5 (XP_642989), Dd_Acd18.9 (XP_641631), Dd_Acd48.1 (XP_641311, AHF>), Dd_Acd137.0 (XP_637736), Dd_Acd25.8 (XP_636488), *Dictyostelium discoideum*; Dv_Hsp13.0 (YP_011654), Dv_Hsp15.4 (YP_011653), Dv_Hsp13.3 (YP_010690), Dv_Hsp16.7 (YP_009880), Dv_Hsp15.4 (YP_009879), *Desulfovibrio vulgaris subsp. vulgaris* str. Hildenborough; Ec_lbpA15.8 (BAB38050), Ec_lbpB16.3 (BAB38049), *Escherichia coli* O157:H7; Gs_Acd17.4 (Gs_11670, A4_12B06), Gs_Acd18.3 (GS02700; A4_10A12), *Galdieria sulphuraria*; Hs_HspB1 (NP_001531), Hs_HspB2 (NP_001532), Hs_HspB3 (NP_006299), Hs_HspB4 (NP_000385), Hs_HspB5 (NP_001876), Hs_HspB6 (NP_653218), Hs_HspB7 (NP_055239), Hs_HspB8 (NP_055180), Hs_HspB9 (NP_149971), Hs_ODF1 (NP_077721), *Homo sapiens*; Kl_Acd22.2 (CAG98636, AKL>), Kl_Acd44.6 (CAH00928), *Kluyveromyces lactis* NRRL Y-1140; Mg_Acd42.9 (XP_361913), Mg_Acd27.1 (XP_360345), *Magnaporthe grisea* 70-15; Mx_Hsp16.6 (AAX51976), *Myxococcus xanthus*; Nc_Acd25.3 (P19752), Nc_Acd25.6 (XP_327518), *Neurospora crassa*; Ne_Hsp15.9 (NP_842084), *Nitrosomonas europaea* ATCC 19718; Nsp_Hsp17.8

(NP_485849), Nsp_Hsp17.1 (NP_484330), *Nostoc* sp. PCC 7120; Pf_Acd25.2 (CAD52179), Pf_Acd19.4 (CAD51208), *Plasmodium falciparum* 3D7; Ps_Hsp16.6 (AAY37023), *Pseudomonas syringae* pv. *syringae* B728a; Re_Hsp15.6 (ZP_00170836), Re_Hsp14.8 (ZP_00170055), Re_Hsp8.3 (ZP_00170413), *Ralstonia eutropha* JMP134; Rp_Hsp18.7 (CAA14735), *Rickettsia prowazekii*; Sc_Hsp26 (P15992), Sc_Hsp42 (NP_010456), *Saccharomyces cerevisiae*; Sj_Acd36.6 (AAW26535), Sj_Acd39.2 (AAW25890), Sj_Acd38.7 (AAW25825), Sj_Acd18.6 (AAW25494), Sj_Acd40.9 (AAW25328), Sj_Acd52.1 (AAW24545), *Schistosoma japonicum*; Ssp_Hsp16.6 (NP_440316), *Synechocystis* sp. PCC 6803; Ta_Acd23.1 (AI76223), *Theileria annulata*; Tb_Acd15.8 (AAZ10367), *Trypanosoma brucei*; Tg_Bag1_Acd25.0 (CAA57695), Tg_Acd21.3 (AAV28541), Tg_Acd29.7 (AAU20421), Tg_Acd28.6 (AAT67148), Tg_Acd20.3 (AAT66039), *Toxoplasma gondii*; YI_Acd16.9 (XP_504108), YI_Acd17.9 (XP_501399), YI_Acd17.2 (XP_501398), *Yarrowia lipolytica*.

Suppl. Fig. 2: Conservation of putative PTS in plant homologs of AtHsp15.7 and AtAcd31.2.

Plant EST sequences that shared high sequence similarity with AtHsp15.7 (A) and AtAcd31.2 (B) and are presumably orthologous were retrieved from organism-specific plant EST databases by TblastN and reversely translated from the putative start methionine that aligned with the start methionine of AtHsp15.7 and AtAcd31.2 to the stop codon that aligned with the C-terminal end of the query proteins (www.expasy.ch). Full-length protein sequences were assembled by alignment of overlapping ESTs and supplemented by the genomic full-length ORF from *Oryza sativa* (Os_Hsp16.0). The molecular mass of the proteins was calculated (www.expasy.ch) and the proteins abbreviated following the nomenclature introduced by Scharf et al. (2001) (e.g. Bv_Acd14.5). For AtAcd31.2 full-length homologs (e.g. Am_Acd30.7) assembled by overlapping ESTs were supplemented by partial N-terminal (e.g. Ah_N-EST) or C-terminal ESTs (e.g. Br_C-EST). The alignment presented comprises the N- and C-terminal 30 residues of AtHsp15.7 and AtAcd31.2. Abbreviations and accession numbers are as follows: (A) AtHsp15.7-CI (At5g37670, *Arabidopsis thaliana*), Bg_Hsp16.0, BP943597, *Bruguiera gymnorhiza*; Bp_Hsp15.6, CD278503, *Betula pendula*; Cs_Hsp14.6, CK932957, *Citrus sinensis*; Gm_Hsp16.2, BF066236, *Glycine max*; In_Hsp16.1, BJ568951, *Ipomoea nil*; Le_Hsp16.1, AW929113, *Lycopersicon esculentum*; Lj_Hsp15.9, BI418471, *Lotus japonicus*; Ls_Hsp15.6, BQ989674, *Lactuca sativa*; Mt_Hsp16.0, AJ500975, *Medicago truncatula*; Os_Hsp16.0, BAD46159, *Oryza sativa*; Pv_Hsp16.2, CV532726, *Phaseolus vulgaris*; Qr_Hsp16.1, DN950655, *Quercus robur*; Ta_Hsp15.1, CD868971, *Triticum aestivum*; Vv_Hsp15.8, CD719440, *Vitis vinifera*; (B) Amaj_Acd30.7, assembled from AJ792671 and AJ790535, *Antirrhinum majus*; Ah_N-EST, CD038232, *Arachis hypogaea*; Amar_N-EST, BM173239, *Avicennia marina*; Bm_N-EST, BP941734, *Burma mangrove*; Br_C-EST, CO749426, *Brassica rapa*; Bv_N-EST, BQ585004, *Beta vulgaris*; Ds_N-EST, BG321516, *Descurainia sophia*; Gr_Acd28.3, CO094312 and CO074697, *Gossypium raimondii*; Gm_Acd28.7, BI427588 and BM178755, *Glycine max*; Hp_N-EST, CF086378, *Helianthus paradoxus*; Ls_N-EST, BQ993055, *Lactuca sativa*; Le_Acd26.5, AW037627, AI780912, and AI782263, *Lycopersicon esculentum*; Mc_Acd30.7, BF479323 and CA834824, *Mesembryanthemum crystallinum*; Mp_C-EST, CN446454, *Malva pusilla*; Mt_Acd28.6, CX518484, and CF068210, *Medicago truncatula*; Vv_Acd28.6, CA808634 and CB003599, *Vitis vinifera*.

Suppl. Fig. 3: Relative tissue-specific and developmental expression levels of AtHsp15.7 and AtAcd31.2.

The relative expression of AtHsp15.7 and AtAcd31.2 was analyzed in different plant tissues and at various stages of plant (A) and leaf development (B) based on publicly available microarray expression data as of

January 2006 (2204 arrays) using GENEVESTIGATOR (www.geneinvestigator.ethz.ch, Zimmermann et al., 2004).

Suppl. Fig. 4: Comparative analysis of heat-inducible expression of *Arabidopsis* sHsp and Acd genes.

The relative expression of *Arabidopsis* sHsp and Acd genes was analyzed under normal conditions in representative organs like seeds and rosette leaves (A) and under heat stress conditions (B) based on publicly available microarray expression data as of January 2006 (2204 arrays) using GENEVESTIGATOR (www.geneinvestigator.ethz.ch, Zimmermann et al., 2004). sHsps and Acd proteins that are not listed were not found on microarrays. According to improved gene prediction the former acronyms of some sHsps and Acd proteins (Scharf et al., 2001) were corrected for some proteins: AtAcd22.1 (At3g22530), AtAcd22.3 (formerly AtAcd21.4, At1g54850), AtAcd25.1 (At2g27140), AtAcd27.7 (formerly AtAcd26.0, At2g03020), AtAcd31.2-Px (At1g06460), AtAcd39.4/30.4 (formerly AtAcd39.0, At1g54840), AtAcd41.3 (At5g04890), AtAcd44.3 (At1g20910), AtAcd48.0 (formerly AtAcd55.2, At1g76510), AtAcd51.9 (At1g20870), AtAcd55.3 (At3g12570), AtAcd55.8/39.8 (formerly AtAcd55.8, At2g37570), AtAcd56.6 (formerly AtAcd55.5, At5g02480), and AtAcd86.6 (AtAcd32.4, see also Suppl. Table II).

Suppl. Table I: List of primers

Suppl. Table II: Conservation of the Hsp20/ α -crystallin domain in *Arabidopsis* sHsps and Acd proteins